UNIVERSITY OF LJUBLJANA
FACULTY OF PUBLIC ADMINISTRATION

UNIVERSITY OF RIJEKA
FACULTY OF ECONOMICS AND BUSINESS

Doctoral dissertation

A COMPREHENSIVE MEASUREMENT FRAMEWORK FOR PERFORMANCE ASSESSMENT OF THE NATIONAL INNOVATION SYSTEMS IN EU COUNTRIES

Jasmina Popovska

Ljubljana, May 2026

UNIVERSITY OF LJUBLJANA
FACULTY OF PUBLIC ADMINISTRATION

UNIVERSITY OF RIJEKA
FACULTY OF ECONOMICS AND BUSINESS

Doctoral dissertation

A COMPREHENSIVE MEASUREMENT FRAMEWORK FOR PERFORMANCE ASSESSMENT OF THE NATIONAL INNOVATION SYSTEMS IN EU COUNTRIES

Candidate:	Jasmina Popovska
Enrolment number:	04190474
Study programme:	Joint doctoral study programme Governance and Economics in the Public Sector
Mentor:	Assoc. Prof. Dr. Lan Umek
Co-Mentor:	Prof. Dr. Marko Jaklič

Ljubljana, May 2026

DECLARATION OF AUTHORSHIP OF THE DOCTORAL DISSERTATION

I, the undersigned Jasmina Popovska, doctoral candidate of the joint doctoral study programme Governance and Economics in the Public Sector, with enrolment number 04190474, I am the author of the doctoral dissertation titled “A comprehensive measurement framework for performance assessment of the national innovation systems in EU countries”.

By signing, I hereby guarantee that:

- the work submitted is solely the result of my own research,
- I have ensured that the works and opinions of other authors used in the submitted work are cited or quoted in accordance with the Faculty's guidelines,
- I have ensured that all works cited and opinions expressed by other authors are included in the list of references and sources, which is an integral part of the submitted work and is written in accordance with the Faculty's instructions,
- I have obtained all permissions to use the copyright works that are fully transferred to the submitted work, and I have made this clear in the submitted work,
- I am aware that plagiarism - the presentation of another's work, either verbatim or in the form of almost verbatim paraphrasing or in graphic form, whereby another's thoughts or ideas are presented as my own - is punishable by law (Copyright and Related Rights Act, Official Gazette of the Republic of Slovenia, No. 21/95) and that the offence is also punishable by action under the rules of the University of Ljubljana and the Faculty of Administration,
- I am aware of the consequences that proven plagiarism may have on the work submitted and on my status at the Faculty of Administration,
- I agree to the publication of the work in the University of Ljubljana Repository.

The doctoral dissertation was proofread by Assoc. Prof. Dr. Marijana Klemenchich, English Language Teaching Department, International Balkan University, Skopje.

Ljubljana, June 2026

Signature of the author:

Jasmina Popovska

ABSTRACT

This dissertation develops a conceptually coherent and empirically robust framework for assessing the performance of National Innovation Systems (NIS) in the European Union (EU). It addresses a central constraint of existing benchmarks, such as the European Innovation Scoreboard, regarding the conflation of innovation outcomes with inputs, and framework conditions, which limits diagnostic value and policy relevance. Building on innovation systems theory, the study distinguishes between NIS's performance, efficiency and capacity.

Methodologically, the dissertation introduces Multi-Cluster Feature Selection for transparent, data-driven indicator selection; applies a two-stage Data Envelopment Analysis model distinguishing knowledge production from commercialisation; and proposes the Efficiency-Adjusted Result-based Performance Index, which integrates performance and efficiency through a performance-first aggregation rule. The framework was empirically implemented for the EU-27 over 2017-2024.

Findings reveal that result-based performance and efficiency are empirically independent dimensions, commercialisation constitutes the binding structural constraint across most EU countries, confirming the European Paradox at the efficiency level, and innovation systems cluster into four distinct performance-efficiency configurations rather than converging around the EU average. Robustness checks confirm ranking stability and construct validity.

The framework advances innovation measurement by providing policymakers with stage-specific diagnostics for evidence-based policy design and structurally grounded peer benchmarking. Limitations include the sample-dependent nature of DEA efficiency, normative weighting choices, and a fixed two-year transformation lag. Future research could extend the framework through meta-frontier approaches, dynamic efficiency analysis, sector-differentiated lag structures, and application to regional or sector-specific innovation systems.

Keywords: Innovation policy, innovation performance, innovation efficiency, innovation capacity, national innovation system, innovation index, Data Envelopment Analysis.

POVZETEK

CELOVIT OKVIR MERJENJA ZA OCENJEVANJE USPEŠNOSTI NACIONALNIH INOVACIJSKIH SISTEMOV V DRŽAVAH EU

Ta disertacija razvija konceptualno koherenten in empirično robusten okvir za ocenjevanje uspešnosti nacionalnih inovacijskih sistemov (NIS) v Evropski uniji (EU). Obravnava osrednjo omejitev obstoječih meril, kot je Evropska tabela inovacijskih kazalnikov, glede združevanja inovacijskih rezultatov z vložki in okvirnih pogojev, kar omejuje diagnostično vrednost in ustreznost za politiko. Študija, ki temelji na teoriji inovacijskih sistemov, razlikuje med uspešnostjo, učinkovitostjo in zmogljivostjo NIS.

Metodološko disertacija uvaja izbor večklasterskih značilnosti za pregleden izbor kazalnikov, ki temelji na podatkih; uporablja dvostopenjski model analize ovoja podatkov, ki ločuje proizvodnjo znanja od komercializacije; in predlaga indeks uspešnosti, prilagojen rezultatom, ki združuje uspešnost in učinkovitost s pravilom agregacije, ki daje prednost uspešnosti. Okvir je bil empirično implementiran za EU-27 v obdobju 2017-2024.

Ugotovitve kažejo, da sta uspešnost in učinkovitost, ki temeljita na rezultatih, empirično neodvisni dimenziji, da komercializacija predstavlja zavezujočo strukturno omejitev v večini držav EU, kar potrjuje evropski paradoks na ravni učinkovitosti, inovacijski sistemi pa se združujejo v štiri različne konfiguracije uspešnosti in učinkovitosti, namesto da bi se zbližali okoli povprečja EU. Preverjanja robustnosti potrjujejo stabilnost uvrstitve in veljavnost konstrukcije.

Okvir izboljšuje merjenje inovacij tako, da oblikovalcem politik zagotavlja diagnostiko, specifično za posamezne faze, za oblikovanje politik, ki temeljijo na dokazih, in strukturno utemeljeno primerjalno analizo med vrstniki. Omejitve vključujejo naravo učinkovitosti DEA, odvisno od vzorca, možnosti normativne uteži in fiksni dvoletni zamik pri preobrazbi. Prihodnje raziskave bi lahko okvir razširile z meta-mejami, analizo dinamične učinkovitosti, sektorsko diferenciranimi strukturami zamika in uporabo v regionalnih ali sektorsko specifičnih inovacijskih sistemih.

Ključne besede: Inovacijska politika, inovacijska uspešnost, inovacijska učinkovitost, inovacijska zmogljivost, nacionalni inovacijski sistem, inovacijski indeks, Analiza podatkovne učinkovitosti.

SAŽETAK

SVEOBUHVATNI OKVIR MJERENJA ZA PROCJENU USPJEŠNOSTI NACIONALNIH INOVACIJSKIH SUSTAVA U ZEMLJAMA EU

Ova disertacija razvija konceptualno koherentan i empirijski robustan okvir za procjenu uspješnosti nacionalnih inovacijskih sustava (NIS) u Europskoj uniji (EU). Obrađuje središnje ograničenje postojećih mjerila, poput Europske ploče inovacijskih rezultata, u vezi s povezivanjem inovacijskih ishoda s ulazima i okvirnim uvjetima, što ograničava dijagnostičku vrijednost i relevantnost za politiku. Nadovezujući se na teoriju inovacijskih sustava, studija razlikuje uspješnost, učinkovitost i kapacitet NIS-a.

Metodološki, disertacija uvodi odabir više klasterskih značajki za transparentan odabir pokazatelja temeljen na podacima; primjenjuje dvostupanjski model analize obuhvatnosti podataka koji razlikuje proizvodnju znanja od komercijalizacije; te predlaže indeks uspješnosti temeljen na rezultatima prilagođen učinkovitosti, koji integrira uspješnost i učinkovitost putem pravila agregacije koje prvo stavlja uspješnost na prvo mjesto. Okvir je empirijski implementiran za EU-27 u razdoblju 2017-2024.

Nalazi otkrivaju da su rezultati i učinkovitost empirijski neovisne dimenzije, komercijalizacija predstavlja obvezujuće strukturno ograničenje u većini zemalja EU-a, potvrđujući europski paradoks na razini učinkovitosti, a inovacijski sustavi grupiraju se u četiri različite konfiguracije učinka i učinkovitosti umjesto da konvergiraju oko prosjeka EU-a. Provjere robusnosti potvrđuju stabilnost rangiranja i valjanost konstrukcije.

Okvir unapređuje mjerenje inovacija pružajući kreatorima politika dijagnostiku specifičnu za fazu za dizajn politika temeljen na dokazima i strukturno utemeljeno usporedno uspoređivanje. Ograničenja uključuju ovisnu o uzorku prirodu učinkovitosti DEA-e, normativne izbore ponderiranja i fiksni dvogodišnji transformacijski zaostatak. Buduća istraživanja mogla bi proširiti okvir kroz meta-granične pristupe, analizu dinamičke učinkovitosti, sektorski diferencirane strukture zaostajanja i primjenu na regionalne ili sektorski specifične inovacijske sustave.

Ključne riječi: Politika inovacija, učinak inovacija, učinkovitost inovacija, kapacitet inovacija, nacionalni inovacijski sustav, indeks inovacija, Analiza omota podataka.

CONTENTS

LIST OF ILLUSTRATIONS

LIST OF FIGURES

LIST OF TABLES

LIST OF ARTICLES PUBLISHED FROM THE DOCTORAL DISSERTATION

Popovska, J., and Umek, L. (2026). Advancing DEA assessment of innovation efficiency through feature selection. Central European Public Administration Review (CEPAR) *(publishing in progress).*

LIST OF ABBREVIATIONS

BERD - Business Expenditure on Research and Development
CEE - Central and Eastern Europe
CIS - Community Innovation Survey
CLRM - Classical Linear Regression Model
COVID-19 - Coronavirus Disease 2019
CRS - Constant Returns to Scale
DEA - Data Envelopment Analysis
DMU - Decision Making Unit
DNDEA - Dynamic Network Data Envelopment Analysis
EC - European Commission
ECM - Efficiency Contribution Measure
EIB - European Investment Bank
EII - Efficiency Innovation Index
EIS - European Innovation Scoreboard
EIT - European Institute of Innovation and Technology
EPO - European Patent Office
EU - European Union
FDI - Foreign Direct Investment
GCR - Global Competitiveness Report
GDP - Gross Domestic Product
GE - Government Effectiveness
GEE - Generalized Estimating Equations
GERD - Gross Expenditure on Research and Development
GII - Global Innovation Index
GNI - Gross National Income
GVC - Global Value Chain
H1 - Hypothesis 1
H2 - Hypothesis 2
HDI - Human Development Index
HHI - Herfindahl-Hirschman Index
ICT - Information and Communication Technology
IEI - Innovation Efficiency Index
IE_IRPI - Efficiency-adjusted Innovation Result-based Performance Index
IKC - Innovation Knowledge/technological Commercialisation subindex
IKP - Innovation Knowledge/technological Production subindex
IMD - International Institute for Management Development
IOI - Innovation Output Indicator
IP - Intellectual Property
IPR - Intellectual Property Rights

IRPI - Innovation Result-based Performance Index
IUS - Innovation Union Scoreboard
KCEI - Knowledge/technological Commercialisation Efficiency subindex
KPEI - Knowledge/technological Production Efficiency subindex
LASSO - Least Absolute Shrinkage and Selection Operator
MC - Monte Carlo
MCFS - Multi-Cluster Feature Selection
NIS - National Innovation System
NSE - National System of Entrepreneurship
OECD - Organisation for Economic Co-operation and Development
OLS - Ordinary Least Squares
PCA - Principal Component Analysis
PCT - Patent Cooperation Treaty
PII - Productivity Innovation Index
PPS - Purchasing Power Standard
R&D - Research and Development
RIS - Regional Innovation Scheme
RQ - Research Question
RQa - Research Question a (innovation results patterns)
RQb - Research Question b (efficiency bottlenecks)
RQc - Research Question c (joint configurations)
RTDI - Research, Technology, Development and Innovation
S3 - Smart Specialisation Strategy
SE - Standard Error
SII - Summary Innovation Index
SME - Small and Medium-sized Enterprise
SPSS - Statistical Package for the Social Sciences
STEM - Science, Technology, Engineering and Mathematics
STI - Science, Technology and Innovation
SURE - Support to mitigate Unemployment Risks in an Emergency
TFP - Total Factor Productivity
TRL - Technology Readiness Level
UN - United Nations
UNDP - United Nations Development Programme
UNECE - United Nations Economic Commission for Europe
VC - Venture Capital
VIF - Variance Inflation Factor
VRS - Variable Returns to Scale
WIPO - World Intellectual Property Organisation

1 INTRODUCTION

Innovation is widely recognised as a central mechanism of long-run economic growth from neoclassical perspectives (Solow, 1956; Romer, 1986) and from evolutionary perspectives that emphasise structural change (Schumpeter, 1934) because it shifts the production frontier through new products, processes, and organisational forms. In this perspective, innovation is not merely an outcome of market activity but a driver of structural change, as it shapes productivity dynamics, competitiveness, and the capacity of economies to renew themselves over time. Beyond growth, innovation is increasingly positioned as a means of addressing major social and environmental challenges, including those articulated in the United Nations (UN) Sustainable Development Goals (Schot and Steinmueller, 2018; Lundvall, 2024). This expanded role provides a basis for rethinking government intervention in innovation at national and supranational levels.

The scientific motivation of this dissertation is grounded in theoretical traditions that conceptualise innovation as dynamic, cumulative, and system-dependent rather than as a linear input-output relation. Schumpeter (1934) conceptualised innovation as the engine of structural economic change, while endogenous growth theory formalised the role of knowledge accumulation and technological progress as mechanisms of sustained development (Romer, 1986). National Innovation Systems (NIS) research extends this reasoning by arguing that countries differ not only in how much they invest in knowledge creation, but in how their institutional structures and actor interactions enable the conversion of knowledge into economic and social value (Freeman, 1987; Lundvall, 1992, 2007; Nelson, 1993). Empirical work repeatedly reports substantial cross-country variation in innovation outcomes even under broadly comparable resource conditions (Fagerberg, 1994; Edquist, 2011). In NIS terms, this pattern implies that performance cannot be inferred from resources alone and that transformation processes matter analytically and politically.

Similar to Jaklič (2017), the NIS approach provides the appropriate conceptual framework for this task because it defines NIS performance as an emergent property of interactions among actors, institutions, and learning dynamics. Therefore, the NIS approach is understood as an analytical lens rather than a single unified theory. It frames innovation as institutionally embedded and systemically produced rather than reducible to firm-level activity or research and development (R&D) spending alone. Freeman (1987) characterises NIS as a network of institutions whose activities and interactions initiate, import, and diffuse new technologies; Lundvall (1992) emphasises interactive learning among users, producers, and research organisations in the production, diffusion, and use of economically useful knowledge; and Nelson (1993) highlights that the innovative performance of national firms depends on institutional interaction patterns.

A central foundation is evolutionary economics, particularly the work of Nelson and Winter (1982), which conceptualises innovation as the outcome of cumulative learning, routines, and bounded rationality rather than instantaneous optimisation. From this perspective, innovation unfolds through path-dependent trajectories, incremental experimentation, and the gradual accumulation of firm- and system-level capabilities. This evolutionary logic provides the theoretical basis for analysing innovation at the system level, as outcomes cannot be understood by examining individual actors or isolated inputs alone.

Building on this foundation, institutional economics highlights the role of formal and informal institutions in shaping innovation dynamics. Institutions such as regulatory frameworks, education systems, funding arrangements, and governance structures influence incentives, knowledge flows, and coordination among actors. As Edquist and Johnson (1997) argue, systems of innovation are defined in institutional terms, with institutions and organisations jointly conditioning how knowledge is generated, diffused, and exploited. They identify three core functions of institutions: reducing uncertainty, managing conflict and cooperation, and providing incentives. Institutions can both support and retard innovation depending on the match between institutional set-ups and prevailing technologies, making the alignment of institutions, organisations, and technological trajectories a key condition for NIS performance over time.

Theories of interactive learning further extend this view by stressing that innovation is fundamentally a collective and relational process. Lundvall (1992, 2007) characterises modern economies as “learning economies” in which competitive advantage depends on the capacity to generate, absorb, and apply knowledge through interaction. Innovation arises from user-producer relationships, inter-firm collaboration, networks of research organisations, labour mobility, and human capital formation. These learning processes are cumulative and socially embedded, explaining why innovation systems differ persistently across countries even under similar economic conditions.

From an innovation policy perspective, governments’ interventions are justified not only by market failures, but by NIS-level failures that affect learning, coordination, and knowledge diffusion. As Edquist and Johnson (1997) argue innovation policy can reduce uncertainty through information, foster cooperation, and provide incentives that respond to capacity, network, institutional failures, that would otherwise keep investment below the social optimum.

Within the EU, innovation policy has been elevated to a strategic priority since at least the Lisbon Agenda for Jobs and Growth (European Commission, 2005) which explicitly positioned innovation as a shared European objective. European innovation policies have been operationalised through a dense landscape of innovation programmes and strategies (e.g. Innovation Union, Horizon 2020, Horizon Europe, European Institute of Innovation and Technology - EIT RIS). In July 2025, the European Commission (EC) proposed nearly doubling the budget for research and innovation to €175 billion for 2028-2034 under the next Framework Programme, complemented by a €409 billion European Competitiveness Fund (European Commission, 2025). As public commitments expand, the logic of innovation policy shifts from whether to intervene toward how to intervene effectively. The core policy problem is no longer only underinvestment, but the possibility that resources are not translated into knowledge and economic value with sufficient effectiveness. In this context, the pressure for evaluation is structural: when budgets grow, accountability requires credible evidence about what innovation systems actually deliver and how well they transform public and private resources into innovation results.

As public budgets for innovation policy increase, governments face growing pressure to evaluate and measure the results and outcomes of such policies and programs. Since the 1990s, Research, Technology, Development and Innovation (RTDI) policies and programs have been subject to evaluations due to the need to better allocate what are increasingly scarce public resources, the need for accountability and transparency in spending taxpayers' money and the intensive public debate about justification, direction and advantages of public investments in research and innovation (Gutiérrez et al., 2017). In addition to policy evaluations and given the objective of innovation policy to influence the development and diffusion of innovations, authors conclude that the relative innovation output should be known to policy analysts, policymakers, and politicians. This will motivate them to improve innovation intensities by means of policy (Edquist et al., 2018). Patel and Pavitt (1994) note the importance of constructing and using frameworks to measure national innovative performance in order to design evidence-based innovation policy. The expansion of RTDI spending and the Europeanisation of innovation policy have been accompanied by governance mechanisms designed for coordination and mutual learning, including benchmarking and performance measurement tools (Wintjes and Nauwelaers, 2008).

Latest research confirmed that the purpose of performance measurement is therefore not ranking per se, but diagnostic capacity: identifying where system functioning breaks down and which interventions are likely to improve innovation results. Long-term priorities such as digital transformation, technological sovereignty, and the green transition depend on an understanding of which parts of the system support innovation and which impede it. At the same time, shocks and structural pressures such as economic disruptions, demographic change, and geopolitical instability, raise the value of measurement tools that can distinguish between limited outcomes caused by constrained system conditions and limited outcomes caused by weak transformation efficiency. However, the literature remains divided regarding the effectiveness of the existing "soft" coordination instruments in the European Union (EU), particularly the

European Innovation Scoreboard (EIS), which has become a dominant reference for cross-country NIS performance assessment.

Research Gap

The theoretical traditions jointly imply that NIS performance must be assessed as the outcome of sequential transformation processes embedded in institutional and learning structures. The core research problem is therefore not the absence of benchmarking tools or empirical studies, but the absence of an integrated, multidimensional measurement framework that reflects the internal logic of innovation systems. Recent research tends to focus on single dimensions in isolation, most prominently innovation efficiency, while treating innovation outcomes and system capacity implicitly or as background conditions. In practice, composite indices emphasise aggregated performance outcomes, whereas efficiency studies focus narrowly on input-output transformation. These approaches have evolved largely in parallel, without a framework capable of jointly analysing what innovation systems achieve, how efficiently they achieve it, and under which systemic conditions. As a result, the literature does not provide policymakers with tools to simultaneously diagnose weak outcomes, inefficient transformation processes, and constrained system capacities. This lack of multidimensional integration limits both analytical coherence and policy relevance.

Innovation research continues to exhibit a persistent gap between how innovation is conceptualised in theory and how it is operationalised in empirical measurement. While innovation systems theory frames innovation as a systemic, sequential, and institutionally embedded process, prevailing measurement frameworks rely on aggregated representations that collapse structure, process, and outcomes. This misalignment, explicitly noted in the literature (Edquist et al., 2018), limits the ability of existing benchmarking instruments to support targeted policy intervention and structurally meaningful peer learning. Despite extensive empirical work and widespread use of innovation composite indicators, these limitations point to a deeper structural gap in how innovation systems are assessed.

This misalignment between innovation systems theory and developed and applied measurement frameworks implies that existing tools are insufficient for diagnosing where and why NIS underperform and for identifying structurally comparable peers for policy learning. Addressing this limitation requires a multidimensional measurement framework that preserves analytical separation while enabling synthesis across dimensions. Specifically, NIS performance assessment must distinguish result-based innovation performance from innovation efficiency, explicitly model the two-stage structure of the innovation process, from knowledge production to commercialisation and diffusion, and reorganise existing indicators into conceptually coherent capacity, performance, and efficiency domains.

This limitation is most evident in existing EU measurement frameworks. The EIS is built on 32 indicators organised into 12 innovation dimensions across four categories, such as framework conditions, investments, innovation activities, and impacts. While this structure acknowledges the multidimensional nature of innovation, the Summary Innovation Index (SII) aggregates these dimensions into a single composite score. In doing so, it combines framework conditions and investments with realised innovation impacts, effectively treating NIS performance as a unidimensional construct (Edquist et al., 2018).

This aggregation creates conceptual ambiguity with direct policy implications. It becomes unclear whether countries perform well because they achieve strong innovation results or because they possess favourable framework conditions and high investment levels. High SII scores may therefore reflect resource endowments or supportive environments rather than effective transformation of knowledge into outcomes, while lower scores may conflate limited capacity with inefficient system functioning. Moreover, despite distinguishing between framework conditions, activities, and impacts at the indicator level, the SII does not model the sequential nature of innovation processes. Applied benchmarking thus implicitly adopts a single-stage logic, despite the long-standing emphasis in NIS theory on innovation as a multi-stage and interactive process encompassing knowledge production, diffusion, and commercialisation (Kline and Rosenberg, 1986; Lundvall, 1992). This obscures stage-specific bottlenecks and reduces the diagnostic value of rankings for targeted policy design.

This dissertation is grounded in a conceptual framework derived from NIS theory, which conceptualises NIS performance as an emergent, systemic outcome of interactions among actors, institutions, and learning processes, unfolding over time and across sequential stages. Building on evolutionary and institutional perspectives, the framework treats innovation not as a single input-output relation but as a multi-stage transformation process, in which knowledge creation, diffusion, and commercialisation are analytically distinct yet interdependent. Within this perspective, NIS performance is understood to comprise three analytically separable dimensions: result-based innovation performance, referring to observable innovation results; innovation efficiency, capturing how effectively systems transform resources into these results across stages; and innovation capacity, defined as a latent system property reflecting institutional quality, governance coherence, learning capabilities, and systemic complementarities. These properties of the capacity are only partially observable and cannot be consistently aggregated without conflating inputs, processes, and outcomes. For this reason, capacity is retained as a conceptual and interpretive dimension. While prior research has examined these dimensions individually, NIS theory implies that meaningful assessment requires their joint consideration within a coherent analytical structure.

The purpose of this dissertation is to develop and apply a conceptually coherent and theoretically grounded framework for assessing the NIS performance in a way that is consistent with innovation systems theory. Specifically, the study aims to operationalise NIS performance as a multidimensional construct by analytically distinguishing between the observable results produced by innovation systems, the efficiency with which resources are transformed into those results across sequential stages of the innovation process, and the underlying system capacity that conditions both performance and efficiency. By explicitly separating outcomes from transformation processes, the framework is designed to overcome the conceptual limitations of existing measurement approaches that conflate inputs, framework conditions, and results. The overarching objective is to provide a more diagnostically informative basis for cross-country comparison that enables identification of stage-specific bottlenecks in knowledge production and commercialisation, supports structurally meaningful peer learning, and improves the analytical foundations of innovation policy evaluation. While the empirical application focuses on EU member countries, the framework is developed at a level of abstraction that allows transferability to other national contexts. A central premise of the dissertation is that NIS performance is inherently multidimensional and cannot be meaningfully assessed through a single analytical lens. A comprehensive evaluation of an innovation system must simultaneously consider:

- the results it produces (scientific outputs, technological developments, innovation outcomes);
- the efficiency with which it transforms inputs into outputs and outcomes, recognising the sequential nature of innovation processes; and
- the systemic capacity that enables these transformations, rooted in institutions, skills, networks, governance, and coordination mechanisms.

A well-functioning NIS must therefore not only generate strong innovation outcomes but do so efficiently and consistently across the stages of the innovation process. Approaches that isolate one dimension, whether performance or efficiency, obscure the mechanisms through which innovation systems succeed or fail.

Against this background, the dissertation responds to three interrelated gaps in the innovation measurement literature and in applied policy frameworks:

1. *Unidimensional focus*: Existing efficiency studies examine either performance or efficiency in isolation, but do not integrate these dimensions within a single analytical framework. This limits the ability to distinguish high-performing systems from highly efficient ones.

2. *Lack of stage differentiation*: Despite the recognition of innovation as a sequential process, most measurement frameworks adopt a single-stage logic,

masking whether weaknesses arise in knowledge production, commercialisation, or diffusion.

3. *Limited diagnostic value for policy learning*: Current rankings of EU member countries by overall scores do not differentiate between performance levels and efficiency patterns, restricting the identification of structurally appropriate peers and targeted interventions.

Addressing these gaps requires a conceptual and methodological framework capable of differentiating what innovation systems achieve (performance), how they achieve it (efficiency), and why systems differ in their ability to transform inputs into outcomes (capacity).

This dissertation proposes a multidimensional measurement approach that integrates three interconnected dimensions of NIS functioning:

- Innovation performance (result-based) that is capturing the observable results of NIS (scientific outputs, technological advances, economic gains).
- Innovation efficiency for assessing how productively NIS transform inputs into outputs and outcomes across the two sequential stages of knowledge production and commercialisation.
- Innovation capacity is defined as the latent systemic ability of a NIS to mobilise and combine resources, such as knowledge, skills, institutions, networks, to support high performance and efficient transformation processes.

Together, these dimensions provide a more theoretically grounded explanation of cross-country differences in innovation outcomes. Throughout this dissertation, the term innovation capacity refers to an underlying system property used for conceptual framing and interpretation, while the empirical measurement is restricted to performance outcomes and efficiency in transforming resources into results.

Building on the premise that existing innovation measurement frameworks conflate what NIS achieve with the processes through which they achieve it, thereby limiting diagnostic value for policy, this dissertation aims to examine the structure of result-based innovation performance and efficiency across EU member countries and to determine whether their joint consideration, interpreted through an innovation capacity lens, provides diagnostic insights unavailable from one-stage and single-dimension assessment. The conceptual framework distinguishes the three analytically separable dimensions, however only performance and efficiency are directly measured, while capacity provides the interpretive lens through which performance-efficiency configurations are understood.

The research design follows a twin approach of assessing both result-based innovation performance and stage-specific transformation efficiency as dimensions of NIS performance. The performance-efficiency framework is critical in examining whether countries achieving similar outcomes operate through fundamentally different transformation processes, a question with direct implications for peer selection and policy learning. The capacity dimension, though not directly measured, informs the interpretation of why countries with similar resources achieve different results and why some systems sustain high performance over time while others do not.

Hence, the research approach consists of the following stages:

1. *Identification of performance patterns across NIS* through examining how countries differ in their balance between knowledge production results and commercialisation results, mapping the structure of result-based innovation performance.
2. *Assessment of stage-specific efficiency patterns* through examining how NIS differ in their transformation productivity at the knowledge production stage versus the commercialisation stage, identifying where bottlenecks occur.
3. *Examination of the performance-efficiency relationship* through analysing how these dimensions relate empirically and what configurations emerge when countries are classified by both dimensions jointly.
4. *Identification of peer group structures* through comparing groupings based on joint performance-efficiency assessment with those based on aggregate rankings alone.
5. *Interpretation through the capacity lens* to understand persistent performance-efficiency configurations by reference to underlying systemic conditions, such as accumulated capabilities, institutional coherence, and structural connectivity.

The focus on dimensional separation aspires to provide a diagnostic framework that moves beyond ranking toward structural understanding. Addressing the conflation problem through an explicit analytical distinction is expected to yield benefits for policy evaluation by revealing heterogeneity that conventional indices obscure.

In this context, the research questions aim to discover facts about the structure and variation of NIS functioning across EU member countries:

Main Research Question (RQ): *How do NIS differ in their innovation results and in the productivity with which they transform resources into those results, and what does this variation reveal about cross-country heterogeneity?* This overarching question is explored through three specific research questions:

1. RQa: What patterns of variation exist in how NIS achieve innovation results, and what configurations characterise the balance between generating *new knowledge and translating it into economic value*?

2. RQb: What efficiency patterns exist across NIS *in the sequential transformation of resources into innovation results,* and where do bottlenecks concentrate?
3. RQc: What structural characteristics distinguish countries when *both their innovation results and transformation productivity are considered together?*

The first research question (RQa) aims to map the empirical landscape of innovation results across EU countries. It explores whether NIS differ in how they balance generating new knowledge, reflected in scientific publications, patents, and technological developments, with translating that knowledge into economic value through commercial applications and market returns. This question does not presuppose a particular pattern but seeks to discover what configurations actually exist: whether countries that excel at producing scientific and technological knowledge are equally successful at bringing it to market, or whether systematic imbalances characterise different system types.

The second research question (RQb) explores transformation productivity across the innovation process. It examines where bottlenecks occur in the sequential transformation of resources into innovation results, whether NIS struggle primarily with generating new knowledge from available resources, or with bringing that knowledge to market. This question is motivated by the "European Paradox", which suggests that EU countries may excel at knowledge generation but encounter constraints in commercialisation. RQb seeks to discover whether this pattern holds systematically across EU member countries and which countries deviate from it.

The third research question (RQc) explores the policy implications of considering both innovation results and transformation productivity together. It examines what structural characteristics emerge when countries are grouped by both dimensions simultaneously: strong results achieved efficiently, strong results achieved inefficiently, modest results with high productivity, and so forth, and whether these groupings reveal different peer structures than conventional rankings based on aggregate performance alone. This question discovers whether considering both dimensions yields substantively different and potentially more meaningful country groupings for policy learning.

In this research, innovation capacity is treated as a latent systemic property that conditions both observed results and transformation productivity. It encompasses three interrelated components identified in NIS theory: accumulated capabilities (stocks of knowledge, skills, and routines), institutional coherence (alignment and stability of formal and informal rules), and structural connectivity (density and quality of actor networks). Although capacity is not measured as a separate empirical index, it provides the interpretive vocabulary for understanding why the patterns discovered through RQa-RQc occur. The capacity lens is applied in Chapter 6 and Chapter 7, where persistent configurations of results and productivity are explained by reference to underlying systemic conditions.

In the same direction, hypothesis development relies on previous research in formulating testable relationships among variables. Prior studies on innovation measurement stress the importance of distinguishing performance from efficiency and suggest that these dimensions may capture distinct aspects of system functioning (Edquist, 2011; Edquist et al., 2018; Barbero et al., 2021; Zofio et al., 2023). The impact of dimensional conflation on diagnostic value has been documented across various measurement contexts (Grupp and Schubert, 2010). Furthermore, the innovation efficiency literature demonstrates that two-stage decomposition can reveal patterns invisible in aggregate analysis (Guan and Chen, 2012; Carayannis et al., 2016). Studies also find that countries with similar aggregate scores often differ markedly in their underlying system functioning, suggesting structural heterogeneity beneath surface-level similarity (Fagerberg and Srholec, 2008; Bartels et al., 2012).

Based on this evidence, the hypotheses predict specific relationships that complement the research questions. While the research questions discover what patterns exist, the hypotheses test whether theoretically expected relationships hold:

(Hypothesis 1) “A multidimensional framework that integrates result-based innovation performance and innovation efficiency measures, interpreted through an innovation capacity perspective, provides a more comprehensive and diagnostically informative assessment of NIS than existing frameworks”.

H1 predicts that the dimensional separation proposed in the conceptual framework will prove empirically meaningful, that performance and efficiency capture distinct rather than redundant aspects of NIS functioning, and that their joint assessment reveals structural patterns obscured in one-stage single-dimension frameworks. H1 is supported if:

- The correlation between result-based innovation performance and innovation efficiency is weak, indicating empirical distinctiveness;
- Systematic differences exist between knowledge production efficiency and commercialisation efficiency, revealing stage-specific patterns; and
- The joint performance-efficiency distribution reveals meaningful country configurations aligned with capacity-related characteristics.

(Hypothesis 2) “The innovation indices based on the multidimensional theoretical framework for performance measurement of NIS enable more accurate identification of peer countries for benchmarking and policy learning”.

H2 predicts that peer groups identified through joint performance-efficiency assessment will be more structurally coherent than those identified through aggregate rankings alone. H2 is supported if:

- The efficiency-performance composite indicator produces stable country rankings under reasonable parameter variation;
- Countries in the same performance-efficiency quadrant share structural characteristics (e.g., similar bottleneck patterns, similar capacity-related determinants); and
- Peer groups based on joint assessment differ substantively from those based on aggregate performance rankings (e.g., SII).

Equally important, hypothesis testing must specify conditions under which the framework would be disconfirmed. Beyond supporting evidence, the empirical analysis evaluates explicit falsification criteria on performance-efficiency independence, stage asymmetry, ranking stability and peer group differentiation in Chapter 5.

Overall, the research design adheres to a clear analytical sequence: theoretical problem → research questions → methods → results → interpretation, ensuring that methodological choices remain subordinated to substantive analytical objectives. The framework maintains a clear distinction between what is measured and what is interpreted as presented in Table 1 below.

Table 1: NIS performance measurement framework

Dimension	Role	Empirical Status
Performance	What NIS achieve	Directly measured (IRPI)
Efficiency	How productively NIS transform resources	Directly measured (IEI)
Capacity	Why NIS differ in their ability to perform and transform	Interpretive lens; inferred from patterns

Source: Own

By integrating performance, efficiency, and capacity perspectives, the dissertation advances a conceptually coherent framework for assessing NIS that aligns innovation systems theory with the practical requirements of policy evaluation and comparative benchmarking.

Methodologically, answering the research questions requires indicators that are non-redundant, empirically justified, and aligned with the sequential structure of innovation processes, requirements that expert-driven composite indices do not satisfy. The dissertation therefore employs a data-driven indicator selection procedure and integrates result-based performance measurement with two-stage efficiency analysis, not as ends in themselves, but as instruments enabling the diagnostic granularity necessary to distinguish bottlenecks in knowledge production from those in commercialisation. Empirically, the dissertation applies this framework to EU member countries, generating evidence on the relationship between result-based innovation performance and efficiency that is not available from existing benchmarking instruments. The expected contribution is a more diagnostically informative basis for cross-country comparison, one that reveals structural heterogeneity obscured by conventional indices and supports the identification of policy-relevant peer groups based on system functioning rather than aggregate rank proximity. Together, these contributions are intended to provide policymakers and researchers with an analytically grounded framework that moves beyond ranking toward structural diagnosis, supporting evidence-based intervention design rather than generic benchmarking.

The methodological framework of the dissertation is explicitly designed to operationalise this conceptual logic without conflating dimensions. Rather than proposing a single measurement technique, the research design will require a dimension-specific methodological orientation, in which different analytical tools are employed to address distinct analytical questions. Result-based innovation performance and transformation efficiency are treated as empirically observable and therefore subject to quantitative assessment, while innovation capacity functions as an interpretative construct that informs comparative analysis rather than a directly measured outcome. This approach aligns with earlier calls in the innovation systems literature for more analytically disciplined measurement architectures (e.g. Grupp and Schubert, 2010; Edquist, 2011; Edquist et al., 2018). In this dissertation, the term *framework* refers to the conceptual structure guiding assessment, while *methods* and *measurement instruments* denote the empirical techniques used to operationalise this structure, which are introduced in subsequent chapters.

The methods employed in this dissertation: composite index construction for measuring innovation results, data-driven indicator selection using Multi-Cluster Feature Selection (MCFS), and two-stage Data Envelopment Analysis (DEA) for assessing transformation efficiency, are adopted strictly as analytical instruments to address the research questions, not as ends in themselves. Each method is selected because it corresponds to a specific analytical requirement derived from the research questions: MCFS addresses the need for a transparent, non-arbitrary procedure for identifying relevant indicators from a larger pool; composite indexing provides the basis for constructing the Innovation Result-based Performance Index (IRPI) as a coherent result-based performance measure; and construction of the Innovation Efficiency Index (IEI) through separate assessment of efficiency in knowledge production and commercialisation based on two-stage DEA, reflecting the sequential logic of innovation processes established in NIS theory. These two indices are subsequently integrated into the Efficiency-adjusted Innovation Result-based Performance Index (IE_IRPI), which synthesises performance and efficiency into a single diagnostic measure while preserving their analytical separation. These techniques do not define the research contribution, but they operationalise the conceptual distinctions that constitute it. The scientific value of the dissertation lies in the theoretical framework that separates result-based innovation performance from innovation efficiency, in the empirical patterns this separation reveals across EU member countries, and in the enhanced diagnostic capacity it provides for policy learning.

This dissertation makes three interrelated contributions to the literature on NIS assessment. Conceptually, it advances NIS performance measurement by enforcing strict analytical separation between result-based innovation performance (observable results), innovation efficiency (transformation processes), and innovation capacity (latent system conditions). These dimensions are frequently conflated in existing measurement frameworks. By restoring this separation, the framework realigns measurement practice with the foundational premises of innovation systems theory, which conceptualises innovation as a sequential, cumulative, and institutionally embedded process. This conceptual discipline addresses a persistent gap in the literature: the absence of measurement architectures that distinguish what innovation systems achieve from how they achieve it and under which systemic conditions.

The following chapters develop this framework in a systematic manner. Chapters 2 to 4 establish the conceptual, theoretical, and empirical foundations underpinning the research questions. Chapter 5 presents the methodological instruments used to operationalise the analytical framework, while Chapter 6 reports the empirical results. Chapters 7 and 8 interpret these findings, discuss their theoretical and policy implications, and synthesise the main contributions of the dissertation.

The structure of the dissertation follows a cumulative analytical logic in which each chapter addresses a necessary precondition for the next. The sequence reflects the theoretical premise that meaningful measurement must be grounded in conceptual clarity before methodological operationalisation.

Chapter 1 establishes the research problem and articulates why existing measurement frameworks fail to capture the multidimensional nature of innovation system performance. It introduces the three-dimensional analytical structure: performance, efficiency, and capacity, which organises the subsequent chapters.

Chapters 2 and 3 together establish what is known and what remains unresolved. Chapter 2 critically reviews the innovation measurement literature, demonstrating that composite indices, efficiency studies, and result-based indicators each capture only partial aspects of system functioning and cannot be combined without a prior conceptual framework. Chapter 3 provides bibliometric evidence that these strands have developed in parallel, confirming the fragmentation documented in Chapter 2.

Together, these chapters justify why incremental refinement of existing approaches is insufficient.

Chapter 4 responds to this gap by developing the conceptual framework that structures the empirical analysis. It formally distinguishes result-based innovation performance (observable results), innovation efficiency (transformation processes), and innovation capacity (latent system conditions), and specifies how these dimensions relate theoretically. This chapter is essential because it provides the analytical architecture that disciplines indicator selection, stage classification, and index construction in Chapter 5.

Chapter 5 operationalises the framework through three measurement instruments: IRPI for result-based performance, IEI for stage-specific efficiency, and IE_IRPI for their integration. Each instrument is explicitly linked to the conceptual distinctions established in Chapter 4 and to the research questions formulated in Chapter 1.

Chapter 6 reports the empirical results for EU member countries, presenting performance and efficiency scores, stage-specific patterns, and the relationship between IRPI and IEI. These findings provide the evidence base for testing the hypotheses and answering the research questions.

Chapter 7 interprets these findings through the lens of innovation systems theory, discussing what the observed patterns reveal about NIS functioning, where the proposed framework succeeds and where it encounters limitations, and what implications follow for policy and future research.

Chapter 8 synthesises the dissertation's contributions and reflects on its broader significance for innovation measurement and policy evaluation.

This chapter has established the research problem, introduced the three-dimensional analytical structure (performance, efficiency, capacity), formulated the research questions and hypotheses, and outlined the dissertation's cumulative logic. The following chapters develop this framework systematically. Chapter 2 begins this development by reviewing the theoretical foundations of innovation measurement and critically examining existing approaches. Its purpose, as stated above, is to demonstrate that composite indices, efficiency studies, and result-based indicators each capture only partial aspects of NIS functioning and cannot be combined without a prior conceptual framework. This critical review provides the foundation upon which the dissertation's measurement architecture is built.

2 THEORETICAL FOUNDATIONS AND CHALLENGES IN INNOVATION SYSTEM ASSESSMENT

Chapter 1 established the research problem: existing innovation benchmarks conflate enabling conditions, transformation processes, and realised results, limiting their diagnostic value for policy. This chapter provides the theoretical foundations necessary for constructing an alternative measurement architecture.

The chapter proceeds in three steps. Section 2.1 reviews the theoretical premises that matter for measurement, establishing innovation as systemic, cumulative, and sequential. Section 2.2 specifies what the NIS perspective implies for assessment, distinguishing system structure, transformation processes, and outcomes. Section 2.3 critically examines the dominant measurement paradigms, such as composite indices such as the EIS, result-based indices, and efficiency approaches, documenting their structural limitations. Section 2.4 synthesises these limitations into a single research gap: the absence of a measurement framework that maintains analytical separation between performance, efficiency, and capacity while enabling their joint interpretation.

This chapter, together with Chapter 3's bibliometric analysis, establishes what is known and what remains unresolved in innovation measurement, thereby justifying the conceptual framework developed in Chapter 4.

2.1 INNOVATION: THEORETICAL FOUNDATIONS

The innovation theory provides a rich understanding of innovation as a systemic, cumulative, and institutionally embedded process. It underlines that innovation cannot be treated as a static input-output relationship, and system-level assessment must reflect the multi-stage, process-oriented nature of innovation and distinguish realised outcomes from the conditions and mechanisms that enable them. In Schumpeter's seminal contribution in "*The theory of economic development"* (Schumpeter, 1934), innovation is defined as the introduction of "new combinations" of production factors. These include new products, new methods of production, new markets, new sources of supply, and new forms of industrial organisation. In Schumpeter's framework, innovation constitutes the central mechanism of economic development by generating discontinuous change (Schumpeter, 1934) through processes of creative destruction (Schumpeter, 1942). A key conceptual distinction introduced by Schumpeter (1934) is that between invention and innovation: invention refers to the emergence of a new idea or principle, while innovation denotes its economic application and commercialisation. This distinction establishes innovation as an inherently multi-stage phenomenon, progressing from knowledge creation to market realisation. Subsequent contributions formalised this understanding. For example, Fagerberg (2003) characterises innovation as the first commercialisation of an idea, reinforcing the focus on realised outcomes rather than inventive activity alone. Schumpeter's perspective (1942) provides a foundational link between innovation, entrepreneurship, competition, and long-run growth, and underpins much of the macro-oriented innovation literature.

Rogers (1995) and Dosi (1988) researched the processes through which innovations spread and generate impact. Rogers' "*Diffusion of Innovations"* (1995) synthesised extensive empirical evidence to explain how new ideas and technologies are adopted within social systems over time, emphasising communication channels, social networks, and adopter behaviour. This work extended the analysis of innovation beyond creation to include diffusion and use, highlighting that economic and social effects depend on adoption dynamics. In parallel, economic analyses of technological change advanced the understanding of innovation trajectories. Dosi (1988) elaborates the concepts of technological paradigms and trajectories, arguing that innovation follows cumulative problem-solving paths shaped by existing knowledge bases and socio-economic conditions. These contributions reinforced the view of innovation as a historically conditioned and directional process rather than a sequence of isolated events. Evolutionary economics further consolidated this process-oriented view. Nelson and Winter (1982) conceptualised firms as organisations characterised by routines, learning, and bounded rationality. Innovation emerges through processes of search, selection, and retention, rather than through optimisation under perfect information. This perspective emphasises cumulative learning, heterogeneity, and persistence in innovation behaviour across firms and technologies. For the present dissertation, evolutionary theory provides the conceptual foundation for treating innovation as a dynamic transformation process, unfolding over time and shaped by institutional and organisational contexts.

Building on these insights, the *system-of-innovation* approach conceptualises innovation as the outcome of interactions among firms, universities, research organisations, financial institutions, and public authorities within a national context (Freeman, 1987; Lundvall, 1992; Nelson, 1993; Groenewegen and van der Steen, 2006; Maloney, 2017). NIS theory highlights that countries differ not only in the scale of their innovation activities but also in the organisation, coordination, and effectiveness of their systems. Lundvall's (1992) emphasis on interactive learning and user-producer relationships underscores the importance of systemic interactions in shaping innovation outcomes, while Nelson's comparative analyses demonstrate how institutional configurations influence national innovation trajectories. These contributions constitute the primary theoretical foundation of this dissertation, motivating both the national level of analysis and the focus on system functioning rather than isolated actors.

Given the central role of innovation in economic growth and policy, a substantial body of research has focused on measurement and performance assessment (Freeman, 1987, Freeman and Soete, 1997, Freeman and Louçã, 2001; Edquist et al., 2018;

Vértésy and Tarantola, 2014). In parallel, international organisations contributed to the standardisation of innovation concepts and data collection practices. The Oslo Manual, developed by the OECD and Eurostat and most recently updated in 2018, provides a harmonised definition of innovation as the implementation of new or significantly improved products, processes, organisational methods, or marketing methods. By focusing on implementation, the Oslo Manual reinforces the emphasis on observable outcomes and underpins empirical instruments such as the Community Innovation Surveys (CIS), which form an important empirical basis for system-level analysis. International indicator frameworks such as the EIS and global indices seek to capture innovation activities and outcomes at the country level using composite measures. Furthermore, econometric and efficiency-based approaches analyse how effectively inputs are transformed into outputs and impacts. These empirical efforts draw directly on the systems and evolutionary traditions by attempting to operationalise abstract concepts such as knowledge, capabilities, and interaction patterns. They reflect a long-standing effort within the literature to translate the theoretical understanding of innovation as a systemic and process-driven phenomenon into tools suitable for comparative analysis.

In Table 2 a synthesis of schools of thought is presented with their measurement implications.

Table 2: Theoretical traditions and their measurement implications

Theoretical tradition	Core assumption	Implication
Schumpeterian economics (Schumpeter, 1934, 1942)	Innovation is discontinuous creative destruction.	Performance should capture transformative outcomes, not incremental improvements.
Evolutionary economics (Nelson and Winter, 1982; Metcalfe, 1995, 1998)	Innovation emerges through cumulative learning and routines.	Capacity reflects accumulated capabilities, it cannot be inferred from current inputs alone.
NIS theory (Freeman and Soete, 1997; Lundvall, 1992; Nelson, 1993)	Innovation is systemically produced through actor interactions.	Aggregate indices obscure system functioning, and stage-specific assessment is required.
Interactive learning (Lundvall,1992)	Innovation depends on knowledge flows and user-producer relationships.	Efficiency depends on coordination quality, not just resource levels.
Functional approach (Edquist, 2011)	Systems should be evaluated by their ability to perform innovation functions.	Diagnosis requires identifying which functions fail, not just overall performance.

Source: Own

Across these traditions, three shared premises emerge that carry direct implications for measurement design. *First*, innovation is conceptualised as a transformation process, not a static endowment, implying that measurement must capture how resources are converted into outcomes, not merely what resources exist or what outcomes result. *Second*, this transformation is sequential and cumulative, unfolding through stages of knowledge creation, diffusion, and commercialisation, implying that single-stage assessment obscures where system functioning breaks down. *Third*, system-level outcomes are emergent, arising from coordination, learning, and institutional quality rather than from aggregated firm-level activity, implying that capacity is a latent property that conditions performance but cannot be directly observed without conflation. These premises are widely accepted in innovation theory but remain inadequately reflected in prevailing measurement frameworks, which either collapse process into outcome (composite indices) or abstract from outcomes altogether (efficiency studies). The design challenge is therefore not to select among existing approaches but to construct a measurement architecture that honours what innovation theory has established about how innovation systems function.

2.2 NATIONAL INNOVATION SYSTEM: STRUCTURE, PROCESSES AND OUTCOMES

According to Jaklič (2017), the NIS approach is best understood as a conceptual framework or analytical lens, rather than a single unified or closed-form theory. Its central contribution lies in highlighting that NIS performance depends on the structure, quality, and coherence of interactions among key actors and institutions involved in the creation, diffusion, and application of knowledge. From this perspective, innovation outcomes reflect not only resource endowments, but also the effectiveness of coordination, learning, and institutional alignment within the system.

NIS theory conceptualises innovation as a systemic and institutionally embedded phenomenon. Innovation outcomes are understood to emerge from interactions among heterogeneous actors, embedded in specific institutional, organisational, and policy contexts. Rather than treating innovation as the result of isolated inputs or individual firm behaviour, the NIS perspective emphasises coordination, learning, and cumulative processes at the level of the national system (Freeman, 1987; Lundvall, 1992; Nelson, 1993). This systemic understanding provides the central theoretical foundation for analysing performance of innovation systems at the country level.

Within NIS theory, innovation systems comprise a broad constellation of actors, institutions, infrastructures, and relational mechanisms. These elements can be analytically grouped into four interrelated components: innovation actors, networks and relationships, institutions, and infrastructure. Table 3 summarises these core components and their respective roles within the system.

Table 3: Key components of NIS

Innovation actors	Networks and relationships
– Firms which are producers, adopters, and modifiers of innovation; – Universities and Research and technology organisations (RTOs) that are sources of scientific knowledge; – Government and agencies such as regulatory bodies, ministries, funding agencies; – Financial institutions including venture capital, banks, public funding mechanisms; – Intermediaries, ex. technology transfer offices, incubators, innovation agencies; and – Users and consumers as co-creators of innovation through feedback and demand.	Innovation is generated through: – collaboration, – knowledge sharing, – joint ventures, – public-private partnerships, and/or – user-producer interactions.
Institutions	**Infrastructure**
Institutional elements include: – education and training systems, – intellectual property rights regimes, – industrial policies, – research funding frameworks, and – cultural norms shaping cooperation and trust.	– Digital infrastructure, – R&D facilities, – communication systems, and – transportation and logistics networks.

Source: Own

Together, these components describe the structural configuration of a NIS. Firms, universities, research organisations, financial actors, public authorities, intermediaries, and users interact through formal and informal networks shaped by institutional frameworks and are supported by physical and knowledge infrastructure. Innovation, in this view, is generated through collaboration, knowledge exchange, and user-producer interaction rather than through linear transmission mechanisms.

NIS theory has exerted a profound influence on innovation policy design. It underpins policy approaches that emphasise horizontal coordination, public-private partnerships, long-term capacity building, and systemic interventions targeting structural bottlenecks. The institutionalisation of the NIS concept by the OECD (2018) facilitated its application in international comparative analysis and reinforced its relevance for policy benchmarking and learning across countries. While NIS theory offers a rich conceptual account of how innovation systems function, empirical assessment of system performance requires analytical frameworks capable of capturing system structure, transformation processes, and observable outcomes in a coherent and separable manner.

At the same time, the NIS framework does not imply that innovation systems are fully contained within national borders. Innovation activities are increasingly embedded in cross-border knowledge flows, multinational firms' R&D networks, global value chains, and international research collaboration. Consequently, NIS performance must be interpreted in light of the degree of internationalisation of an innovation system, including outward and inward linkages, cross-border R&D activity, mobility of talent, and openness to external knowledge sources. This consideration is particularly important for small and open economies, whose innovation outcomes depend critically on their ability to access, absorb, and integrate knowledge generated beyond national boundaries. More broadly, in a context where technological capabilities are becoming central to economic resilience and geopolitical positioning, the interaction between domestic system structures and international knowledge networks represents an integral dimension of national innovation competitiveness.

A central insight shared across foundational innovation theories is that NIS performance ultimately concerns what systems achieve. Schumpeter (1934) defined innovation as the introduction of new combinations into economic life, while Nelson (1993) emphasised that national systems differ most clearly in the results they produce. From this perspective, NIS performance reflects the outputs of knowledge production and the outcomes of commercialisation and diffusion. Empirical studies have therefore highlighted the importance of measuring observable innovation outputs and outcomes to assess whether new knowledge has been successfully transformed into economic and technological results. However, the literature exhibits substantial heterogeneity in how performance-related concepts are defined and operationalised, as discussed below.

Despite broad agreement on the systemic nature of innovation, the literature exhibits substantial heterogeneity in how performance-related concepts are defined and operationalised. Scholars differ in their use of terminology and in their conceptual distinctions between innovation performance, innovation efficiency, and innovation capacity (Freeman, 1987; Lundvall, 1988, 1992; Edquist et al., 2018; Huang et al., 2011; Vértésy and Tarantola, 2014). As a result, empirical studies often draw on overlapping indicator sets and apply different labels to analytically distinct phenomena. While the literature provides extensive insights into system functioning, it offers less explicit guidance on how to maintain conceptual consistency when assessing innovation systems quantitatively.

These theoretical traditions converge on a common understanding of innovation as a cumulative, interactive, and multi-stage process. At the same time, they imply, often implicitly rather than explicitly, that different analytical dimensions must be distinguished when assessing innovation systems.

The literature thus implies that three analytically distinct dimensions must be considered when assessing innovation systems: the results systems achieve (performance), how productively resources are transformed (efficiency), and the underlying conditions that enable innovation activity (capacity). However, existing measurement frameworks have not adequately maintained separation between these dimensions, as the following sections demonstrate.

Several scholars have conceptualised innovation capacity as a latent system property. Furman et al. (2002) define national innovation capacity as a country's long-term ability to generate and commercialise new technologies, shaped by its innovation infrastructure, industrial environment, and linkages between actors. However, the literature has not resolved the fundamental measurement challenge: any observable indicator that might proxy for capacity (such as R&D personnel, university quality, or institutional trust) simultaneously functions either as an input to the innovation process or as a reflection of prior innovation success. This identification problem explains why existing measurement frameworks either conflate capacity with inputs or omit it altogether. Chapter 4 addresses this challenge by treating capacity as an interpretive construct rather than a directly measured dimension.

The efficiency literature has developed a distinct analytical approach. Drawing on Farrell's (1957) concept of efficiency as distance from the best-practice frontier and Griliches' (1990) interpretation in innovation contexts, scholars assess how effectively available resources are converted into innovation results. Contemporary research recognises that this transformation involves at least two sequential stages: knowledge production and commercialisation (Guan and Chen, 2012). However, as Section 2.3.3 will demonstrate, efficiency studies have developed largely in isolation from performance measurement, without integration into a coherent multidimensional framework.

The importance of distinguishing these dimensions becomes evident when considering cross-country differences in innovation outcomes. Freeman (1987), Lundvall (1992), and Nelson (1993) demonstrate that countries with similar resource endowments often exhibit markedly different innovation results due to differences in institutional arrangements, learning processes, and interaction patterns. Edquist (2011) further argues that innovation systems should be evaluated by their ability to perform essential innovation functions, as functional weaknesses frequently explain underperformance in input-rich systems. Performance measurement thus serves a diagnostic role, revealing whether challenges originate in knowledge production, diffusion, or commercialisation.

The NIS perspective therefore implies that measurement must distinguish between system structure (institutions and capabilities), transformation processes (knowledge production and diffusion), and observable outcomes. However, as the following section demonstrates, existing measurement approaches have not maintained this separation, creating a gap that motivates the conceptual framework developed in Chapter 4.

2.3 MEASURING PERFORMANCE OF NATIONAL INNOVATION SYSTEMS

Composite indicator approaches represent the dominant paradigm in performance measurement of NIS, particularly in policy-oriented contexts (e.g. Desai et al., 2002; Archibugi and Coco, 2004; Saltelli, 2007; Nardo et al., 2008; Hollanders et al., 2024; WIPO, Cornell University, & INSEAD, 2019; Vértésy and Tarantola, 2014). Their appeal lies in their capacity to condense heterogeneous indicators into a single comparative metric, facilitating benchmarking and communication (Nardo et al., 2008). However, from an innovation systems perspective, this strength is also their central limitation. By aggregating indicators that capture framework conditions, resources, activities, and outcomes into a unitary score, composite indices implicitly assume that these dimensions are analytically interchangeable. This assumption conflicts with the core premise of NIS theory, which emphasises that innovation outcomes emerge through structured transformation processes rather than from static endowments (Zabala-Iturriagagoitia et al., 2007, 2021; Edquist et al., 2018).

An additional critical reason for advancing NIS performance measurement in the EU lies in the persistent and well-documented NIS performance gap between EU member countries. The EC has repeatedly highlighted, in documents such as the EIS (Hollanders et al., 2024) and the New European Innovation Agenda (European Commission, 2022), that innovation leaders and moderate or emerging innovators continue to diverge rather than converge, threatening the EU's internal cohesion and its ability to compete globally. Research by Archibugi et al. (2009) similarly shows that technological divergence within the EU reduces the Union's collective capacity to respond to global competitive pressures. In this context, benchmarking instruments are used extensively, but their policy value depends on whether they diagnose system functioning rather than merely ranking countries.

Composite indicators have proven effective for comparison and communication, yet their aggregation logic conflicts with the process-oriented foundations of innovation systems theory. This tension highlights the need for a framework that retains the empirical richness of composite indicators while avoiding conceptual conflation between resources, activities, and outcomes.

2.3.1 What the European Innovation Scoreboard enables and what it cannot explain

The EIS constitutes the most influential application of the composite indicator paradigm in the EU policy context. The SII is organised into a hierarchical framework consisting of 12 innovation dimensions and 32 indicators. The dimensions include human resources, research system quality, innovation-friendly conditions, financial support, firm-level investments, use of information technologies, innovation activities among SMEs, collaborative linkages, intellectual property generation, employment impacts, sales impacts, and indicators related to environmental performance. The indicators' definitions and data sources are provided in EIS 2024 Methodology Report published by the Directorate General for Research and Innovation, EC on 8th of July 2024. Table 4 presents the indicators.

Table 4: EIS Indicators list

Code	Indicator
1.1.1	New doctorate graduates in science, technology, engineering, and mathematics (STEM) per 1000 population aged 25-34
1.1.2	Percentage of population aged 25-34 having completed tertiary education
1.1.3	Percentage population aged 25-64 participating in lifelong learning
1.2.1	International scientific co-publications per million population
1.2.2	Scientific publications among the top 10% most cited publications worldwide as percentage of total scientific publications of the country
1.2.3	Foreign doctorate students as a percentage of all doctorate students
1.3.1	Broadband penetration
1.3.2	Individuals who have above basic overall digital skills (% share)
2.1.1	R&D expenditure in the public sector (percentage of GDP)
2.1.2	Venture capital expenditures (percentage of GDP)
2.1.3	Direct government funding and government tax support for business R&D (percentage of GDP)
2.2.1	R&D expenditure in the business sector (percentage of GDP)
2.2.2	Non-R&D innovation expenditures (percentage of turnover)
2.2.3	Innovation expenditures per person employed
2.3.1	Enterprises providing training to develop or upgrade ICT skills of their personnel
2.3.2	ICT specialists (as a percentage of total employment)
3.1.1	SMEs introducing product innovations (percentage of SMEs)
3.1.2	SMEs introducing business process innovations (percentage of SMEs)
3.2.1	Innovative SMEs collaborating with others (percentage of SMEs)
3.2.2	Public-private co-publications per million population
3.2.3	Job-to-job mobility of Human Resources in Science & Technology
3.3.1	PCT patent applications per billion GDP (in PPS)
3.3.2	Trademark applications per billion GDP (in PPS)
3.3.3	Design applications per billion GDP (in PPS)
4.1.1	Employment in knowledge-intensive activities (percentage of total employment)
4.1.2	Employment in innovative enterprises
4.2.1	Exports of medium and high technology products as a share of total product exports
4.2.2	Knowledge-intensive services exports as percentage of total services exports

Code	Indicator
4.2.3	Sales of new-to-market and new-to-enterprise innovations as percentage of turnover
4.3.1	Resource productivity
4.3.2	Air emissions by fine particulate matter (PM2.5) in Industry
4.3.3	Development of environment-related technologies, percentage of all technologies

Source: Own and Hollanders et al., 2024

EU countries are assigned to four performance groups: Innovation Leaders, Strong Innovators, Moderate Innovators, and Emerging Innovators, based on their relative position to the EU average. These classifications serve descriptive purposes and support comparative analysis.

Although EIS underlying structure recognises multiple dimensions of innovation, distinguishing between framework conditions, investments, innovation activities, and impacts, this multidimensionality is not grounded in any theoretical framework or preserved at the level of performance assessment. The SII collapses these dimensions into a single composite score, thereby reintroducing the very conceptual conflation that the dimensional structure initially seeks to avoid.

The limitation of the EIS does not lie in its indicator base, which is comprehensive and methodologically robust, but in its aggregation logic and lack of theoretical explanation. By combining enabling conditions with realised outcomes, the SII obscures whether differences in country performance reflect effective transformation processes or simply favourable structural endowments. As a result, the EIS provides limited insight into innovation system functioning, despite its widespread policy relevance.

Although the SII provides a structured and widely used composite measure of NIS performance within the EIS, numerous methodological and conceptual limitations affect its interpretability, analytical precision, and suitability for assessing the functioning of NIS. These limitations have been well documented in the literature on composite indicators, innovation system evaluation, and science and innovation policy analysis.

The SII operationalises selected components of NIS by measuring human resources, research excellence, firm activities, linkages, intellectual assets, and innovation outputs. These correspond to important institutional and structural features highlighted in NIS theory (Freeman and Soete, 1997; Lundvall, 1992; Nelson, 1993). However, as emphasised by Edquist et al. (2018), SII fails to embody the systemic nature of innovation, because it aggregates inputs, framework conditions, and outputs without distinguishing between innovation capacity and innovation results. The index therefore provides useful benchmarking information but does not capture the efficiency, interactions, or the dynamic multi-stage processes through which national systems generate and commercialise knowledge.

Furthermore, the SII places greater emphasis on innovation inputs and framework conditions, while offering less systematic coverage of innovation outputs, impacts, and innovation system efficiency. Although output indicators such as sales of innovative products or employment in knowledge-intensive activities are included, they represent only a narrow subset of outputs and do not capture broader societal or economic impacts. Scholars argue that innovation measures must assess not only resource availability but also the effectiveness of converting inputs into measurable innovation outcomes (Hagedoorn and Cloodt, 2003). The SII therefore provides limited insight into the efficiency, productivity, or transformative capacity of NIS.

Some of the flaws and limitations are methodological such as the lack of a variable (indicator) selection method, and consequential equal weighting applied to all indicators within each dimension and to all dimensions within the composite score. Although this choice enhances transparency, it assumes without theoretical justification that all indicators contribute equally to innovation performance. This approach can distort underlying constructs by inflating the impact of redundant, overlapping, or weakly correlated indicators while downplaying variables empirically shown to be more central to innovation processes (Grupp and Schubert, 2010; Saisana and Tarantola, 2002).

The index continues to face data availability and imputation challenges, especially for smaller or newer EU member countries. The EIS applies imputation methods such as rolling averages and proportional adjustments to address missing data, yet these techniques can introduce systematic bias, estimation error, or artificial smoothing of national trajectories (Nardo et al., 2008). Consequently, countries with weaker statistical infrastructures may receive SII scores that obscure volatility, structural weaknesses, or discontinuities in their innovation systems.

Like most composite indicators, the SII is prone to over-interpretation and policy misuse. The danger is that policymakers may focus on improving individual indicators to achieve better rankings, rather than addressing deeper systemic bottlenecks or institutional deficiencies (Roper and Hewitt-Dundas, 2006; Edquist et al., 2018).

These limitations are not about the indicators themselves but about the EIS aggregation logic: the conflation of inputs, framework conditions, intermediate outputs, and final outcomes into a unitary index obscures the transformation processes central to innovation system analysis. Consequently, the limitation of the EIS lies not in its data foundation but in its aggregation logic, which motivates the development of an alternative, stage-differentiated measurement framework. Composite indices, as implemented in the EIS/SII, cannot diagnose whether weak outcomes reflect limited capacity or inefficient transformation.

2.3.2 Result-based innovation performance indicators

Alongside composite indices and efficiency-oriented approaches, a distinct strand of the literature argues that NIS performance should be assessed through the results and consequences of innovation processes rather than through inputs or framework conditions (e.g. Edquist et al., 2018; Janger et al., 2017). In this view, NIS performance refers to what innovation systems ultimately achieve in terms of scientific outputs, technological advances, and economic or societal outcomes.

The most prominent policy-oriented implementation of this logic is the Innovation Output Indicator (IOI), developed by the EC in 2013 at the request of the European Council (European Commission, 2013). The IOI was designed to complement R&D intensity measures by focusing explicitly on observable innovation outputs, with the objective of benchmarking national innovation policies and monitoring the EU's performance relative to major trading partners (Vértésy and Tarantola, 2014). In this sense, the IOI represents an explicit attempt to operationalise NIS performance in practice through the construction of a composite index centered on outcomes rather than inputs.

However, similar to the SII, the IOI is primarily a pragmatic policy instrument rather than the outcome of a well-developed theoretical literature on NIS performance measurement. While its focus on outputs improves conceptual clarity relative to input-based indices, the construction of the IOI is not grounded in a coherent theoretical framework that explains why particular outputs are selected, how they relate to different stages of the innovation process, or how they collectively represent innovation system performance. As a result, the index reflects expert judgement and data availability more than explicit theoretical reasoning.

This limitation is compounded by the absence of a transparent and systematic method for output and outcome indicator selection. The IOI does not employ a formal procedure to assess the representativeness, redundancy, or explanatory relevance of its indicators. Consequently, although the IOI isolates outcomes more clearly than broader composite indices, it remains vulnerable to conceptual arbitrariness and implicit assumptions about what constitutes innovation success.

Result-based indicators therefore play an important role in clarifying what innovation systems achieve, but they cannot explain how these outcomes are generated or whether they reflect efficient transformation processes. Nor do they resolve the deeper measurement challenge identified in the literature: the lack of theoretically grounded and methodologically disciplined frameworks that integrate performance outcomes with transformation efficiency while preserving analytical separation between dimensions.

These limitations motivate efficiency-oriented literature, which shifts analytical attention from what innovation systems achieve to how effectively they transform resources into results.

2.3.3 Innovation efficiency: analytical relevance and structural limitations

In contrast to composite indicator approaches, there are NIS efficiency studies that explicitly focus on the transformation of resources into innovation outputs and outcomes. This focus aligns more closely with the core insight of innovation systems theory that NIS performance cannot be inferred from inputs or outcomes alone, but from the processes that link them.

Within this literature, innovation efficiency is commonly conceptualised as the extent to which observed innovation outputs approach a feasible maximum given available inputs, or conversely, the minimum inputs required to generate a given level of outputs (Guan and Chen, 2012). From this perspective, weak NIS performance may reflect not only insufficient resources, but ineffective transformation processes. This distinction is analytically and politically significant, as it directly informs whether policy responses should prioritise capacity expansion or improvements in system functioning.

Over the past two decades, a substantial body of research has applied this efficiency logic to NIS, typically by separating indicators into inputs and outputs and estimating relative efficiency using non-parametric methods such as DEA (e.g. Guan and Chen, 2012; Aristovnik, 2014; Edquist et al., 2018; Ansaloni et al., 2019; Anderson and Stejskal, 2019; Alnafrah, 2021). In policy contexts, this logic has also been reflected in efficiency ratios embedded within composite frameworks, most notably the Innovation Efficiency Ratio in the Global Innovation Index. Taken together, these studies have strengthened the case for analysing NIS performance as a transformation problem rather than as a static endowment.

However, despite their conceptual proximity to innovation systems theory, efficiency studies introduce a different form of analytical reduction. By construction, efficiency models prioritise the input-output relationship, often abstracting from the substantive nature of innovation outcomes and from the institutional and structural conditions under which transformation occurs. As a result, efficiency scores capture how intensively resources are transformed, but not what is being achieved, nor why certain transformation patterns emerge in specific national contexts.

A second limitation is that the efficiency literature remains predominantly method-driven rather than framework-driven. While increasingly sophisticated DEA variants have been developed, including two-stage, network, dynamic, and multi-objective models, their application is rarely embedded within an explicit conceptual architecture that distinguishes result-based innovation performance, innovation efficiency, and innovation capacity as analytically separate dimensions. Consequently, efficiency scores are frequently interpreted as proxies for overall innovation performance, despite capturing only one aspect of system functioning. This conflation mirrors, in a different form, the aggregation problem observed in composite indicator approaches.

A further and largely unresolved limitation of innovation efficiency studies concerns the selection of input and output indicators. While most efficiency analyses rely on the same broad pools of indicators drawn from sources such as the EIS, or the GII, the criteria guiding indicator choice are rarely made explicit. In practice, variables are typically selected based on availability, precedent in earlier studies, or authors' theoretical judgement regarding representativeness, rather than through a systematic or data-driven selection procedure. As a result, efficiency estimates are highly sensitive to discretionary modelling choices that are often insufficiently justified or scrutinised.

This lack of transparency has important analytical consequences. Different studies operationalise innovation inputs and outputs in substantially different ways, even when analysing similar country samples, leading to efficiency scores that are not directly comparable across studies. Moreover, overlapping or weakly informative indicators are frequently included without assessing redundancy or relevance to the specific transformation stage under analysis. Consequently, efficiency results may reflect artefacts of indicator choice rather than genuine differences in innovation system functioning. Despite extensive methodological refinement of DEA models, the literature has paid comparatively little attention to disciplining the indicator space itself, leaving a critical component of efficiency analysis conceptually underdeveloped.

Recent methodological advances have attempted to better reflect the sequential nature of innovation processes by distinguishing between knowledge production and commercialisation stages (e.g. Guan and Chen, 2012; Carayannis et al., 2016; Tziogkidis et al., 2020; Anouze et al., 2024). Table 5 summarises methodological choices in comparable DEA studies on NIS efficiency. This comparison demonstrates that methodological heterogeneity produces incomparable results, reinforcing the need for a conceptually prior framework rather than a particular DEA variant.

Table 5: Selected scientific articles for comparison of DEA methodology

Scientific articles	**Edquist et al. (2018)**	**Carayannis et al. (2016)**	**Anouze et al. (2024)**	**Guan and Chen (2012)**	**Hudec and Procházková (2013)**	**Liou (2009)**
Indicators Selection method	No	No	No	No	No	No
Data sources	IUS	IUS, WIPO	GII	OECD, IMD	OECD, Eurostat	OECD
Period	2014-2015	2007-2011	2016-2021	1999-2003	2004-2010	2000-2005
Two-stage DEA	No, one	Yes	Yes	Yes	Yes	Yes
Applied Method	Input oriented, CRS DEA	MOLP-DEA, VRS, input oriented - stage 1, output oriented - stage 2	DNDEA-VRS, non-oriented	PLSR, VRS and CRS	CRS and VRS, input-oriented	Malmquist DEA
Analysed NISs	EU-28 countries	EU-23 countries	33 oil countries	OECD-22 countries	EU-19 countries	15 OECD countries

Source: Own

These two-stage and network DEA models represent an important step toward aligning efficiency analysis with the process logic emphasised in innovation systems theory. At the same time, they amplify sensitivity to indicator selection, classification of intermediate outputs, and assumptions regarding comparability across heterogeneous national systems. As a result, efficiency estimates become increasingly conditional on modelling choices, reinforcing the need for careful interpretation.

The literature also highlights that inefficiency identified by DEA should not be equated with managerial failure. In the context of NIS, inefficiency often reflects structural and institutional constraints such as weak coordination, rigid governance arrangements, or path-dependent specialisation, rather than suboptimal resource allocation in a narrow sense. Several authors therefore caution that efficiency analysis is most informative when used as a diagnostic tool for identifying systemic bottlenecks, not as a basis for ranking countries by overall performance.

The reviewed efficiency studies provide indispensable insights into transformation processes but remain analytically incomplete when applied in isolation. By focusing on transformation intensity, they tend to understate outcome quality and broader system conditions, and by privileging method over framework, they blur the distinction between efficiency and performance. As a result, the efficiency literature has developed largely in parallel with composite performance measurement, without a coherent structure that integrates both perspectives.

This fragmentation constitutes a central research gap. Existing approaches cannot simultaneously answer three questions implied by innovation systems theory: what innovation systems achieve, how efficiently they transform resources into outcomes, and how underlying system conditions shape these processes. Addressing this gap requires a multidimensional measurement architecture that preserves analytical separation between performance and efficiency while enabling their joint interpretation.

In response, this dissertation adopts innovation efficiency not as a substitute for performance measurement, but as a complementary analytical dimension embedded within a broader framework. By explicitly separating result-based innovation performance from transformation efficiency, and by interpreting both through an innovation systems lens, the framework developed in this study seeks to overcome the limitations of existing efficiency-only and performance-only approaches without discarding their empirical contributions.

2.3.4 Synthesis: the need for a new measurement architecture

The limitations documented in Sections 2.3.1-2.3.3 are not isolated methodological shortcomings but reflect a structural misalignment between how innovation is conceptualised in theory and how it is operationalised in measurement. Composite indices conflate inputs, framework conditions, and outcomes, efficiency studies treat efficiency as a proxy for performance, and result-based indicators ignore transformation processes. None of these approaches maintains the analytical separation that NIS theory implies is necessary.

Addressing this misalignment requires a measurement framework that: maintains strict separation between performance, efficiency, and capacity; reflects the sequential nature of innovation through stage-specific assessment; employs transparent indicator selection; and treats capacity as a latent construct to avoid conflation. Chapter 4 develops such a framework, and Chapter 5 operationalises it through specific methodological choices.

2.4 SUMMARY OF THEORETICAL FOUNDATIONS AND RESEARCH LIMITATIONS

This dissertation most directly extends the functional approach to NIS assessment articulated by Edquist et al. (2018) and the two-stage efficiency strand developed by Guan and Chen (2012) and subsequent scholars. From the functional approach, it inherits the premise that innovation systems should be evaluated by their ability to perform essential functions, and that diagnostic value lies in identifying which functions fail, not merely in ranking overall performance. From the two-stage efficiency literature, it adopts the recognition that transformation processes must be assessed separately for knowledge production and commercialisation. However, the dissertation departs from both strands by integrating them within a unified measurement architecture that preserves the separation between performance and efficiency that efficiency-only studies collapse, employs systematic indicator selection that neither strand has adopted, and treats capacity explicitly as a latent construct rather than as a measured input or background condition.

The critical review of innovation measurement approaches confirms and deepens the research gap introduced in Chapter 1. The literature is characterised less by contradiction than by fragmentation: composite indices prioritise breadth and comparability but sacrifice the ability to distinguish outcomes from enabling conditions; efficiency studies prioritise transformation analysis but abstract from outcome quality and system context; and result-based indicators capture what systems achieve but not how or why. Each strand offers partial insight, yet none provides the integrated diagnostic capacity that innovation systems theory implies is necessary.

The evidence reviewed in Sections 2.3.1-2.3.3 explains why this fragmentation persists. Composite indicators such as the SII aggregate framework conditions, inputs, activities, and impacts into a single score, obscuring whether high performance reflects effective transformation or favourable endowments (Edquist et al., 2018). Efficiency studies, while closer to the process-oriented logic of NIS theory, typically treat efficiency scores as proxies for overall performance and rely on DEA without embedding it in a conceptual architecture that distinguishes performance from efficiency from capacity (Hollanders and Celikel Esser, 2007; Guan and Chen, 2012; Aparicio et al., 2020). Result-based indicators such as the IOI isolate innovation outcomes more clearly but lack transparent selection procedures and do not explain how outcomes relate to transformation processes across stages (Janger et al., 2017; Vértésy and Tarantola, 2014).

The recurring weakness across all three strands is the implicit treatment of enabling conditions as performance outcomes and the absence of architectures that maintain analytical separation while enabling synthesis. This pattern is not incidental: it reflects the difficulty of operationalising the multi-stage, institutionally embedded conception of innovation that NIS theory articulates. The literature has refined individual measurement techniques extensively but has not resolved the prior conceptual challenge of specifying how performance, efficiency, and capacity relate and how each should be assessed without contaminating the others.

In the EU policy context, where innovation measurement informs cohesion funding, cross-country coordination, and mutual learning, these limitations carry practical consequences. Current tools rank countries but do not reveal whether weak outcomes reflect resource constraints, transformation inefficiencies, or bottlenecks at specific stages of the innovation process. Policymakers seeking to design targeted interventions lack the diagnostic granularity that structurally meaningful benchmarking requires (Freeman, 1987; Lundvall, 1992; Edquist et al., 2018).

This chapter has established that the central challenge in innovation measurement is architectural rather than empirical. Existing approaches, such as composite indices, efficiency studies, and result-based indicators, each capture partial aspects of NIS functioning but lack the structural integration that innovation systems theory requires. The critical review documented three specific limitations: (1) composite indices conflate inputs with outcomes, obscuring whether performance reflects effective transformation or favourable endowments; (2) efficiency studies treat transformation productivity as a proxy for overall performance without distinguishing outcome levels; and (3) these research strands have developed largely in parallel without conceptual integration.

Chapter 3 tests whether this fragmentation reflects a selective reading of the literature or a structural characteristic of the research field itself. Through bibliometric analysis, examining publication patterns, citation networks, and thematic clustering it provides empirical evidence on how research on NIS performance measurement is organised. If the outcome-oriented and efficiency-oriented streams identified in this chapter's critical review have indeed developed with limited cross-fertilisation, this pattern confirms that the measurement gap is structural and justifies the framework development undertaken in Chapter 4.

3 BIBLIOMETRIC ANALYSIS OF PERFORMANCE ASSESSMENT OF NATIONAL INNOVATION SYSTEMS

Chapter 2 identified structural limitations in existing innovation measurement approaches through critical literature review. This chapter provides complementary empirical evidence for these limitations by mapping the intellectual structure of the research field itself. Bibliometric analysis serves a specific diagnostic function within this dissertation. By examining how research on NIS performance measurement is organised, which thematic clusters dominate, which influential works shape the field, and where integration between research streams is weak, the analysis tests whether the fragmentation documented in Chapter 2 reflects the broader structure of the literature. If outcome-oriented and efficiency-oriented research streams have developed in parallel with limited cross-fertilisation, this pattern confirms that the measurement gap identified in Chapter 2 is not merely a selective interpretation but a structural characteristic of the field.

The chapter is structured as follows: Section 3.1 describes the data sources, search strategy, and bibliometric methods. Section 3.2 reports descriptive patterns including publication dynamics and citation distributions. Section 3.3 analyses the intellectual structure through citation and keyword mapping. Section 3.4 examines influential works that have shaped the field. Section 3.5 synthesises the findings and draws out implications for the conceptual framework developed in Chapter 4.

3.1 INTRODUCTION, DATA COLLECTION AND BIBLIOMETRIC METHODS

The bibliometric analysis in the field of performance measurement of NIS complements the findings in previous literature review. Bibliometric analysis is a quantitative method for evaluating and mapping scientific knowledge through the statistical examination of publications, citations, and research networks. Bibliometric analysis uses quantitative indicators (publication counts, citations, H-index) and network mapping techniques (keyword co-occurrence, co-authorship networks) to reveal the intellectual structure of research fields. This approach enables systematic identification of thematic clusters, influential works, and gaps in the literature that qualitative reviews alone cannot capture. As López-Rubio et al. (2022) note, the complexity and interdisciplinarity of NIS research, spanning economics, management, policy, and technology studies, make bibliometric methods particularly valuable for mapping how the field has evolved and where it fragments.

Within the domain of innovation studies, bibliometric analysis has mapped how the NIS concept evolved from its theoretical foundations (Freeman, 1987, Freeman and Soete, 1997; Lundvall, 1992; Nelson, 1993) into a multi-dimensional empirical research field. López-Rubio et al. (2022), covering over 1,100 publications, identified four contemporary research streams through bibliographic coupling: factors influencing innovation systems, national capabilities, the dynamics of innovation, and innovation systems. Building on this picture, Jansen (2025) identifies eight contemporary clusters, with NIS, Institutional Cooperation Networks, and Frameworks and Process Modelling emerging as the central, most interconnected hubs of activity. These bibliometric insights reveal how innovation measurement, once focused on R&D intensity and patent outputs, has shifted toward integrated system performance and policy coherence, yet, as Chapter 2 documented, without achieving the analytical integration this dissertation pursues.

The bibliometric analysis in this chapter was conducted using Scopus as the bibliographic database due to its broad coverage of social sciences, business, and economics. Records were retrieved on 24th August 2025. The search strategy targeted publications containing the terms innovation, measurement, framework, index, indicator, output, scoreboard, and performance or efficiency in titles, abstracts, and keywords. The final query[1] provided 4,492 documents. To focus on socio-economic and policy dimensions rather than technical or engineering aspects, disciplines such as engineering, medicine, computer science, and natural sciences were excluded, resulting in a final sample of 2,569 documents spanning 1972 to 2025.

The analysis employs annual output and impact metrics (H- and G-indices, citation distributions), collaboration patterns, and thematic structures via keyword co-occurrence and country co-authorship networks. Science mapping uses association-

---

[1] TITLE-ABS-KEY(nation* OR countr*) AND TITLE-ABS-KEY(innovati*) AND TITLE-ABS-KEY measur* OR framework OR index OR indicator OR output OR scoreboard) AND TITLE-ABS-KEY performance OR efficiency OR success) AND (EXCLUDE (SUBJAREA,"ENGI") OR EXCLUDE (SUBJAREA, "ENVI") OR EXCLUDE(SUBJAREA,"MEDI") OR EXCLUDE (SUBJAREA,"COMP") OR EXCLUDE (SUBJAREA,"ENER") OR EXCLUDE (SUBJAREA,"AGRI") OR EXCLUDE (SUBJAREA,"PSYC") OR EXCLUDE (SUBJAREA,"EART") OR EXCLUDE (SUBJAREA,"MULT") OR EXCLUDE (SUBJAREA,"NURS") OR EXCLUDE (SUBJAREA,"PHYS") OR EXCLUDE (SUBJAREA,"MATH") OR EXCLUDE (SUBJAREA, "BIOC") OR EXCLUDE (SUBJAREA,"CENG") OR EXCLUDE (SUBJAREA,"PHAR") OR EXCLUDE (SUBJAREA,"CHEM") OR EXCLUDE (SUBJAREA, "MATE") OR EXCLUDE (SUBJAREA,"HEAL") OR EXCLUDE (SUBJAREA, "IMMU") OR EXCLUDE (SUBJAREA,"DENT") OR EXCLUDE (SUBJAREA, "VETE")) AND (LIMIT-TO (LANGUAGE, "English"))

strength normalisation and community detection to identify research fronts. The bibliometric analysis was conducted in Biblium, a Python library for scientometrics (Umek and Ravšelj, 2025). This toolkit was chosen both for the objectives of this dissertation and for its capacity to map fragmentation across composite-index, result-based, and efficiency-oriented strands, the three approaches whose limitations Chapter 2 identified. By integrating bibliometric indicators with theoretical insights, this chapter situates NIS performance assessment within the broader landscape of innovation theory and policy, providing a solid empirical foundation for subsequent analytical discussion.

3.2 DESCRIPTIVE BIBLIOMETRIC RESULTS

The query used for the bibliometric analysis identified literature that examines how innovation at the national or country level is assessed using different forms of measurement tools such as frameworks, indices, indicators, outputs, or scoreboards and how these tools are linked to evaluating performance, efficiency, or success. By excluding technical subject areas like engineering, medicine, computer science, energy, and natural sciences, the focus narrows to studies rooted in social sciences, economics, business, and policy analysis. This refined scope captures work on innovation measurement in relation to competitiveness, governance, institutional frameworks, and economic outcomes.

Table 6 below provides a comprehensive overview of bibliometric indicators drawn from a dataset of 2,569 documents published between 1972 and 2024. In terms of temporal distribution, the average year of publication is around 2018, with the median in 2019 and the mode in 2024, reflecting a strong recent growth in research output. Missing values are evident in source titles (770 cases) and especially in unrestricted access data, where nearly three quarters (73.73%) of documents lack clear information. This highlights both the long historical span of the collection and challenges in metadata completeness, particularly regarding access and publication details.

Table 6: Summary of bibliometric indicators

Cited by	Mean	19.44
	Median	3
	Mode	0
	Max	1122
	Percentiles - 10%	0
	Percentiles - 25%	0
	Percentiles - 50%	3
	Percentiles - 75%	12
	Percentiles - 90%	38
	C1	1848 (71.93%)

	C5	1083 (42.16%)
	C10	759 (29.54%)
	C20	469 (18.26%)
	C50	198 (7.71%)
	C100	94 (3.66%)
	C1000	1 (0.04%)
Performance indicator	Number of documents	2569
	Total citations	49942
	H-index	97
	Average year	2017.70
	Collaboration index	3.18
	Documents per active year	47.57
	Citations per active year	924.85
Time series analysis	Timespan	1972-2024
	Most Productive Year	2024 (313 documents)
	Average Growth (Last 3 Years)	27.22%
	Average Growth (Last 5 Years)	17.80%
	Average Growth (Last 10 Years)	11.65%
	Most Influential Year	1997 (389.67 citations/doc)
	Most Influential (Last 3 Years)	2022 (10.16 citations/doc)
	Most Influential (Last 5 Years)	2020 (14.85 citations/doc)
	Most Influential (Last 10 Years)	2016 (34.57 citations/doc)

Source: Own

Citation analysis indicates a skewed distribution, with an average of about 19.4 citations per document but a median of only 3, showing that while some works are highly cited, many receive little attention. The maximum citation count reaches 1,122, and percentile values reveal that half of the documents have three or fewer citations, while only 10% exceed 38 citations. The concentration of impact is further emphasised by cumulative citation measures: the share of documents that received at least 1 citation is 71.93% (C1), the documents with at least 5 citations are 42.16% (C5), and less than 4% had more than a hundred citations (C100). This confirms a highly unequal distribution of scholarly influence, consistent with bibliometric laws such as Lotka's.

The collaboration index of about 3.18 reflects a moderate tendency toward co-authorship. The documents are spread across 47.6 active publication years, averaging about 925 citations annually. Time-series analysis highlights sturdy growth in recent decades, with average annual growth rates of 27.22% (last three years), 17.8% (last five), and 11.65% (last ten), culminating in 2024 as the most productive year with 313 publications. However, influence peaks earlier, with 1997 standing out as the most influential year (389.67 citations per document), while 2016 remains the most influential in the last decade. More recent years like 2020 and 2022 show lower but still relevant citation averages, reflecting the typical citation lag for newer publications.

The performance indicators in Table 7 provide further context: the dataset accounts for nearly 50,000 citations in total, with an H-index of 97, meaning 97 documents received at least 97 citations each. The dataset exhibits a highly skewed citation distribution, where a small number of papers account for most citations. This is reflected in a Gini coefficient of 0.83 and an HHI of 54.3, consistent with Lotka’s bibliometric law.

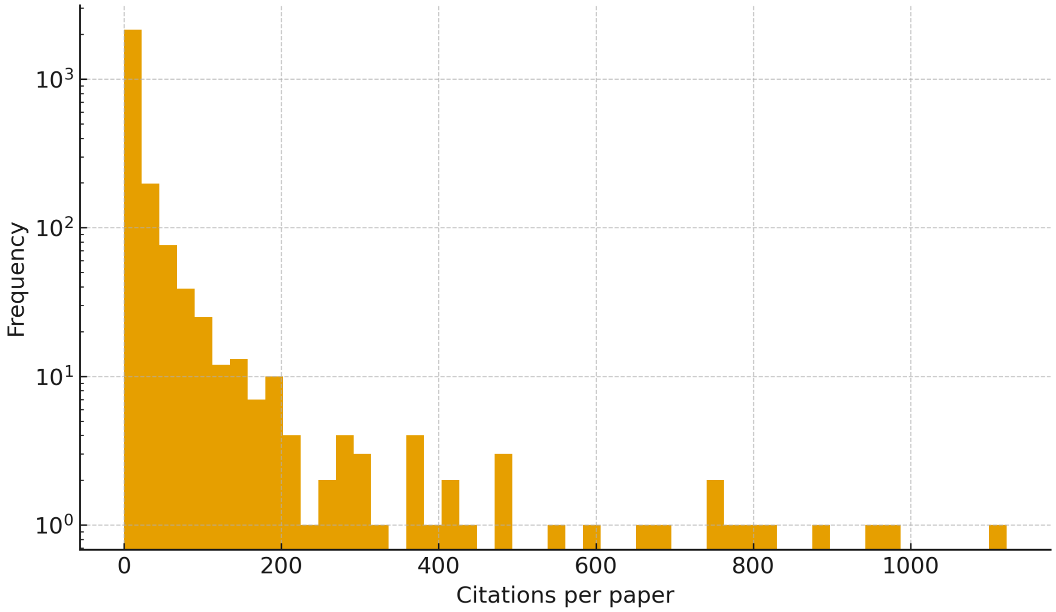
Table 7: Core Indicators

Indicator	Value	Interpretation
H-index	97	Number of papers with ≥ h citations
G-index	182	Top g papers collectively received ≥ g² citations
C100	30.52%	Share of total citations held by top 1% of papers
C5	59.53%	Share of total citations held by top 5% of papers
C1	73.20%	Share of total citations held by top 10% of papers
Gini coefficient	0.83	Measure of citation inequality (0=equal, 1=unequal)
Herfindahl-Hirschman Index (HHI)	54.33	Concentration of citations among few papers

Source: Own

The Citation distribution (log scale) in Figure 1 shows a steep right-skewed pattern, with most papers receiving few or zero citations and a small number of papers driving the overall impact.

Figure 1: Citation distribution (log scale)

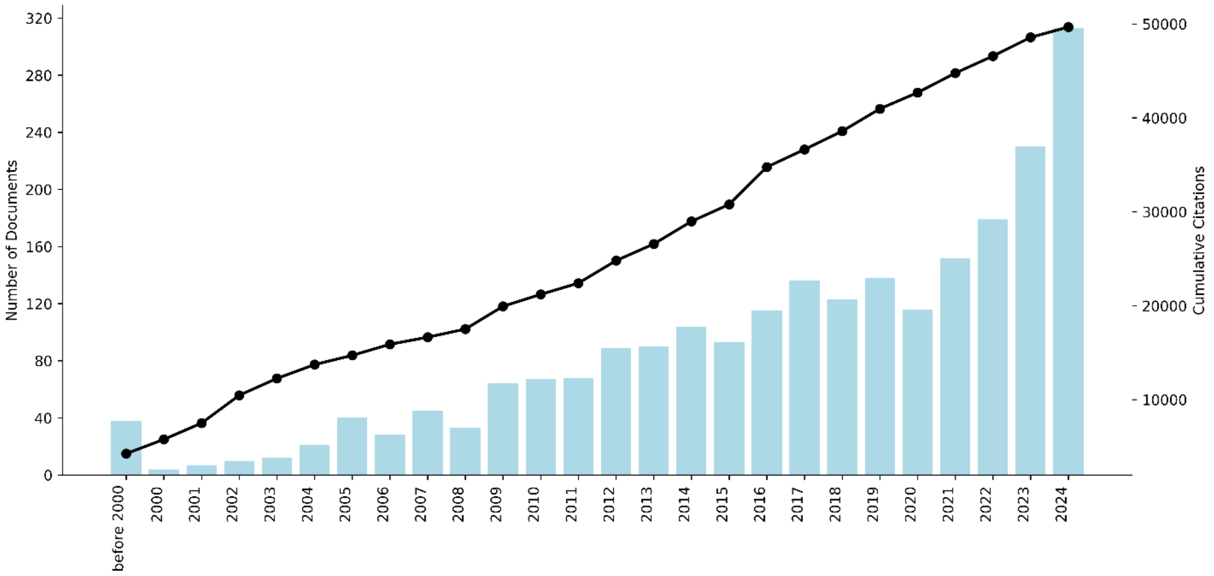


Source: Own

The temporal analysis indicates that research on NIS performance assessment remained sparse until the early 2000s, followed by a rapid expansion in the 2010s and 2020s. The number of publications peaked in 2024 (313 documents), while cumulative citations exceeded 49,000 by 2025. The field exhibits sustained growth and diversification, reflecting its increasing relevance for innovation policy and national competitiveness studies. The analysis of scientific production traces the annual distribution of documents and their cumulative citations from 1972 to 2025. In the early years, publication activity was sparse, with only isolated outputs before the 1990s, and cumulative citations grew slowly during this period. A gradual rise becomes visible in the 1990s, but it was only after 2000 that publication volumes started to grow consistently, reflecting the expansion of research activity in this field. For visualisation purposes, all documents published before 2000 were aggregated together in the plot to provide a clearer picture of trends in the modern period.

From 2000 onwards, the dataset shows a strong upward trajectory in both document production and cumulative citations. While the early 2000s recorded fewer than 50 documents annually, by 2009 the annual output exceeded 60, and in the 2010s the numbers grew steadily, surpassing 100 documents per year after 2014. The trend accelerated further in the late 2010s and early 2020s, culminating in a peak of 313 documents in 2024. Cumulative citations mirrored this growth, rising from just over 5,700 in 2000 to nearly 50,000 by 2025, with marked jumps in years of higher publication output. Figure 2 below shows that the cumulative citations have consistent compounding growth. The year 2025 was excluded from the plotted visualisation, despite being included in calculations, to avoid skewing trends with incomplete citation data for the current year. While 2025 already shows 254 documents and contributes to raising the cumulative citation total to 49,942, its omission in the plot ensures that growth patterns are based on complete and stable data. The result is a clear visualisation of a field that remained small and scattered before 2000 but has since entered a phase of sustained expansion and citation impact, particularly in the last decade.

Figure 2: Cumulative citations over time



Source: Own

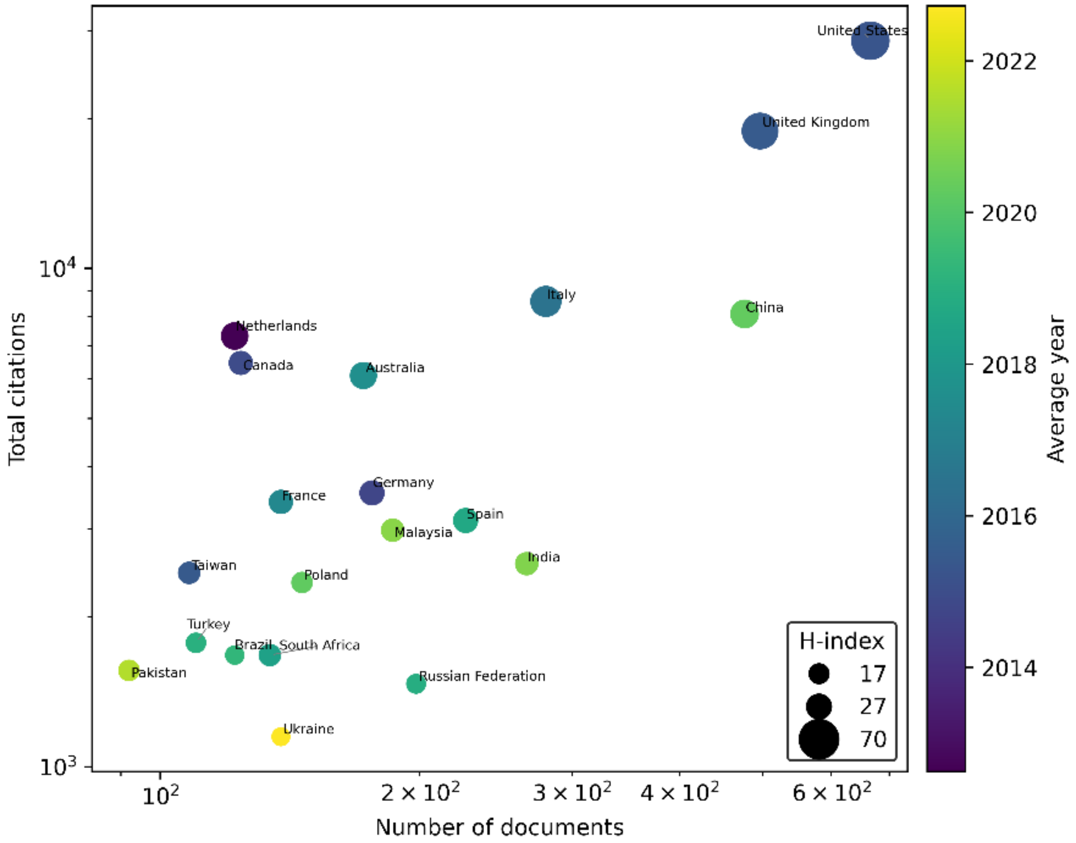
The descriptive patterns reveal a rapidly growing but thematically fragmented field. The high publication growth rate (17.8% annually over the past five years) indicates sustained scholarly interest in innovation measurement, while the skewed citation distribution (median of 3 citations versus mean of 19.4) suggests that influence is concentrated in a small number of influential works. For the present dissertation, this pattern implies that the field's intellectual foundations are established by a few highly cited works and warrant careful examination to identify what they achieve and where gaps remain. The analysis of top-cited documents in Section 3.4 therefore serves not merely descriptive purposes but diagnostic ones: identifying which measurement challenges the influential literature has addressed and which remain unresolved.

3.3 SCIENCE MAPPING AND NETWORK STRUCTURES

The global distribution of publications and citations reveals strong geographical patterns, with the United States and the United Kingdom acting as dominant contributors and emerging research activity increasingly originating from Asia and Eastern Europe. These trends point to a growing internationalisation of scholarship on innovation measurement and performance. To move beyond descriptive statistics and understand how this research landscape is organised, science mapping techniques are applied to examine co-authorship networks, collaboration structures, and thematic clusters. This approach clarifies how knowledge flows between countries and how research communities form around shared topics.

Figure 3 below illustrates the distribution of publications and citations by country, based on the affiliations of all co-authors. The United States is by far the most prominent contributor, with 666 documents and over 28,000 citations, resulting in the highest H-index (70). The United Kingdom follows with 496 documents and nearly 19,000 citations, also showing a high H-index (64). Italy, China, and India appear among the most productive nations in terms of publication numbers (266-476 documents each), though their citation counts and H-indices are more modest compared to the U.S. and U.K. Countries such as the Netherlands and Canada, while producing fewer documents (around 120 each), stand out for their high citation impact, suggesting fewer but more influential contributions.

Figure 3: Publications and citations per country



Source: Own

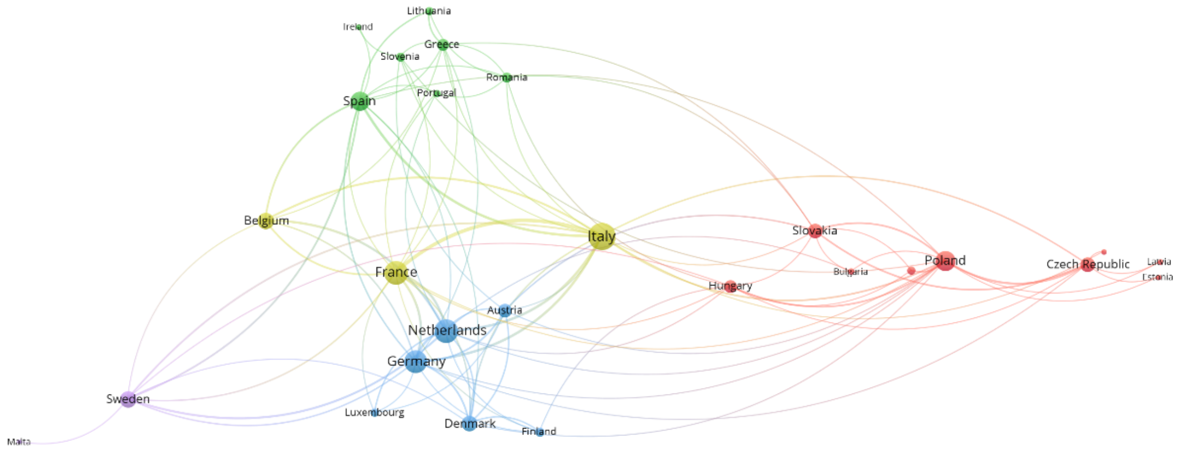
The temporal distribution, reflected in the “average publication year” values, indicates that much of this research activity has accelerated in the last decade. Countries like Malaysia, India, Pakistan, Ukraine, and Poland show very recent average years (2020-2022), highlighting their emergence as newer players in this field. By contrast, the Netherlands, Germany, and Canada have earlier average years (2012-2015), pointing to an established but slightly older body of contributions. The figure shows a strong dominance of Western countries in terms of citations and impact, but also reveals an increasing diversification of research contributions, with rising outputs from Asia, Eastern Europe, and developing economies in the last few years. This pattern reflects the globalisation of research on innovation measurement and performance.

Focusing the analysis specifically on EU member countries is warranted both by the objectives of this thesis and by the distinctive policy architecture that characterises the European innovation landscape. The European Union (EU) has established a coordinated and institutionally embedded framework for innovation governance reflected in instruments such as the EIS, Horizon Europe funding programmes, and the Smart Specialisation Strategy (S3). These mechanisms create a shared macro-policy environment within which NIS operate, rendering the comparative assessment of EU countries analytically coherent and empirically meaningful.

In this context, examining intra-EU collaboration patterns, thematic clustering, and knowledge flows becomes essential for understanding NIS performance within a harmonised regulatory space. The EU-specific co-authorship network is generated using VOSviewer, a software tool designed for creating and visualizing bibliometric networks. It supports analyses such as co-authorship, keyword co-occurrence, citation, co-citation, and bibliographic coupling, making it suitable for mapping the intellectual and collaborative structure of research fields. The software uses the VOS mapping technique to produce clear, distance-based visualisations in which the proximity of nodes reflects the strength of relationships. Its built-in clustering algorithm identifies thematic groups and research communities, enabling users to detect patterns and emerging topics in large datasets.

Figure 4 provides a structural representation of these dynamics, revealing the distribution of collaborative ties, the presence of regional sub-communities, and the role of central actors that function as integrators within the European research ecosystem. Such network features offer insights into how research capacity, institutional cooperation, and knowledge diffusion processes contribute to the evaluation of NIS performance in the EU. Accordingly, the global analysis serves as a necessary structural reference point, while the subsequent EU-focused examination enables a more granular investigation of how European countries are positioned within the broader scientific landscape and how their collaborative configurations influence patterns of NIS performance assessment.

Figure 4: EU country collaboration by authors



Source: Own

This figure was produced as a VOSviewer network map and depicts international co-authorship and collaboration patterns among EU countries in the field of NIS performance assessment. Each node represents a country, and the size of the node indicates the relative number of publications or the strength of its contribution to the topic. Links (edges) between nodes represent co-authorship or institutional collaboration between researchers in those countries, while colours denote distinct clusters of closely collaborating countries or research communities. At the centre of the network, Italy emerges as the most connected node, serving as a major bridge linking Western, Northern, and Eastern European research networks. The network can be interpreted as comprising five regional collaboration clusters, each reflecting patterns of academic and institutional cooperation within Europe.

Cluster 1 or Yellow Cluster has the following core countries: Italy, Spain, Greece, Portugal, Romania, Lithuania, Slovenia, Ireland. This cluster represents a Mediterranean-centred network where Italy functions as the central hub. The close ties between Italy, Spain, Greece, and Portugal indicate strong collaboration in policy-oriented NIS performance studies, often linked to regional innovation systems, sustainability transitions, and EU cohesion policy frameworks. Researchers from these countries frequently analyse innovation through comparative policy evaluation, emphasizing innovation efficiency, regional disparities, and sustainability outcomes, themes aligned with the EU's Smart Specialisation Strategy (S3).

Cluster 2 or Blue Cluster encompasses Germany, France, Netherlands, Austria, Denmark, Finland, Luxembourg. This cluster forms the core innovation research region of the EU, dominated by advanced economies with strong R&D systems and established NIS assessment frameworks. Collaboration here is characterised by high academic productivity and methodological sophistication, particularly in quantitative performance measurement, DEA, and knowledge economy modelling. Germany and the Netherlands act as anchor countries, often leading comparative empirical studies on R&D efficiency, firm-level innovation measurement, and institutional determinants of innovation performance.

Cluster 3 or Red Cluster is focused on Poland, Czechia, Hungary, Slovakia, Latvia, Croatia, Bulgaria. This cluster reflects emerging innovation systems within Central and Eastern Europe (CEE). Collaborations are less dense but reveal a strong regional integration trend. Poland and Czechia appear as key regional connectors, collaborating both eastward and westward.

Cluster 4 or Purple Cluster is a small one having only Sweden and Malta (small node) as core countries. Although smaller in scale, this cluster represents innovation economies that often contribute to methodological and policy benchmarking. Sweden is one of the EU’s top innovation leaders, frequently collaborating with continental partners (Germany, Denmark, and Finland) to develop NIS performance indices, sustainability-driven policy models, and digital innovation systems.

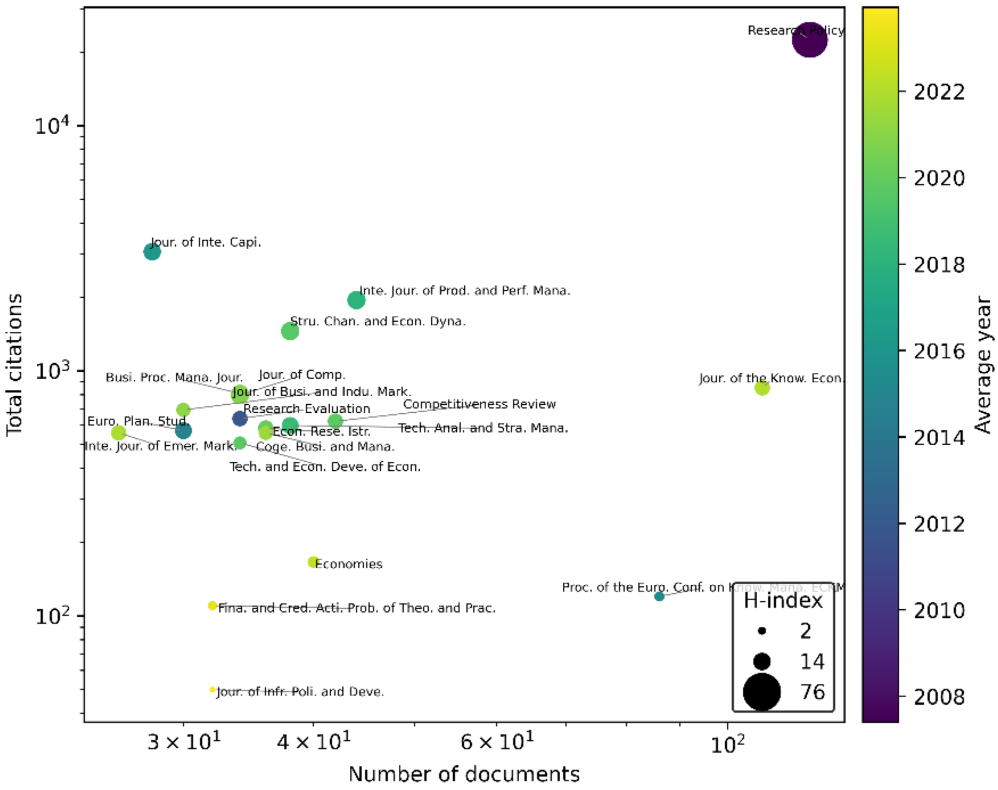
Cluster 5 or Green Cluster has core countries: Belgium, Spain, and connections with France and Portugal. This cluster overlaps with Southern and Western Europe, illustrating cross-regional cooperation between Iberian and Benelux countries. The network structure emphasises knowledge management, innovation networks, and public policy assessment.

Italy is the most central and connected country, bridging Mediterranean, Central, and Eastern European collaborations. Germany, France, and the Netherlands form a core triad of high-impact contributors driving methodological development. Poland and the Czechia play boundary-spanning roles in integrating Eastern European research communities. Sweden and Denmark serve as northern connectors, contributing to innovation policy benchmarking and knowledge efficiency models.

The network exemplifies how European innovation research operates as a distributed system, reflecting the very principles of national and regional innovation systems theory, where performance depends on interconnected institutions, collaborative learning, and policy coherence.

The main journals publishing research on innovation measurement and performance at the national level were identified and presented in Figure 5 below.

Figure 5: Sources publishing research on innovation measurement



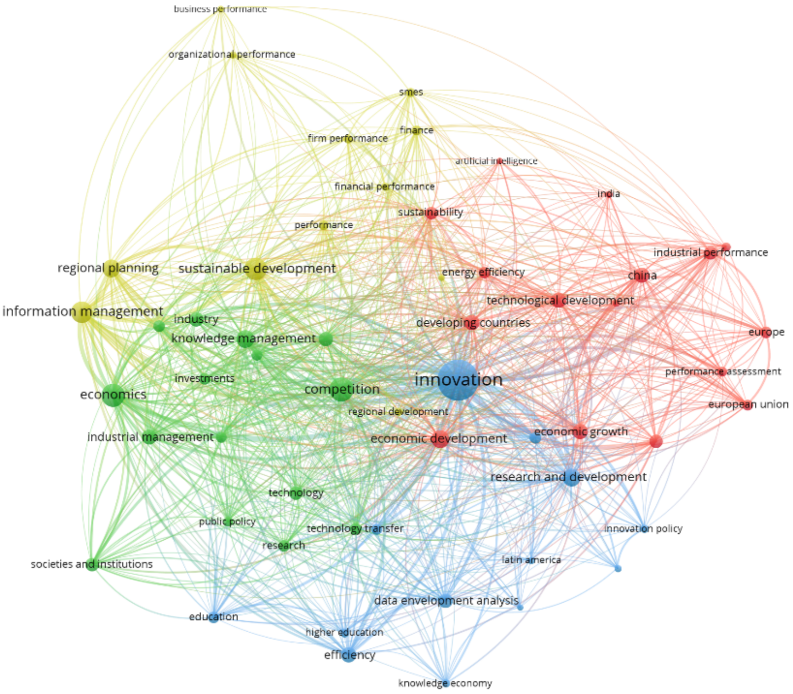
Source: Own

Among them, Research Policy stands out clearly as the most influential outlet, with 120 documents and over 22,000 citations, reflected in an exceptionally high H-index of 76. This places it far ahead of other journals in terms of scholarly impact, even though its average publication year is relatively early (2007), indicating that its influence has been built up over a longer time span. Other established outlets with strong citation impact include the Journal of Intellectual Capital (3,060 citations, H-index 18), Structural Change and Economic Dynamics (1,452 citations, H-index 20), and the International Journal of Productivity and Performance Management (1,942 citations, H-index 20). These journals combine consistent publication activity with considerable influence in the field, even though their citation volumes are much smaller than *Research Policy*.

In contrast, several journals are characterised by more recent publication activity and modest citation counts, reflecting newer contributions to the field. Titles such as Journal of Business and Industrial Marketing, Business Process Management Journal, Journal of Competitiveness, and Cogent Business and Management show average publication years close to 2020-2022, suggesting growing engagement with the topic in the last few years. However, despite substantial numbers of documents (30-44 each), their citation impact remains moderate compared to the more established journals. The Journal of Infrastructure, Policy and Development and Financial and Credit Activity are also very recent outlets, with average years around 2023, but so far with very low citation counts and H-indices. The figure reveals a landscape dominated by a few highly influential and established journals, alongside a broad set of newer outlets that are increasing the volume of publications but have yet to accumulate substantial impact.

Furthermore, by analysing the co-occurrence of keywords in publications addressing NIS performance assessment and related themes the dominant keywords were identified and presented in the network visualisation below Figure 6. Each node represents a keyword, and the size of the node corresponds to its frequency in the dataset. The links show co-occurrence relationships, while the colour clusters denote thematic groupings derived from bibliometric mapping. At the centre of the network lies "innovation" the dominant concept connecting various thematic domains such as economic development, technological change, policy, sustainability, and entrepreneurship.

Figure 6: Author keyword co-occurrence



Source: Own

The visualisation reveals four major clusters, each representing a distinct but interrelated stream of innovation research:

Cluster 1 or Blue Cluster aligns closely with authors who emphasise multi-indicator approaches (R&D, patents, product outputs) for measuring innovative performance. The prominence of R&D, economic development, and innovation policy underscores the role of inputs and outputs in NIS performance assessment, which is key to innovation system efficiency analysis. In this cluster key terms are innovation, research and development, economic development, DEA, knowledge economy, innovation policy. This cluster reflects the quantitative and systemic approaches to measuring NIS performance. The keywords indicate focus areas such as assessing national or regional innovation systems, evaluating efficiency using DEA and linking R&D investment and economic growth.

Cluster 2 or Red Cluster is geographically and industrially oriented, emphasizing comparative NIS performance across nations and regions. The frequent appearance of countries (China, India, Europe) and terms like industrial performance and economic growth highlights research on how innovation drives competitiveness and productivity in different socio-economic contexts. Key terms in this cluster are technological development, industrial performance, China, India, Europe, performance assessment, economic growth.

Cluster 3 or Green Cluster represents the institutional and managerial dimension of innovation. It emphasises knowledge management, information systems, and sustainability, pointing toward innovation as a learning and coordination process within organizations and regions. It integrates regional planning and sustainable development, reflecting growing interest in green innovation and systemic knowledge integration. The key terms in this cluster are knowledge management, information management, sustainable development, regional planning, industrial management, and economics.

Cluster 4 or Yellow Cluster focuses on micro-level innovation performance, particularly at the firm or SME level and the organizational and financial performance, with key terms: business performance, financial performance, SMEs, firm performance, sustainability, finance. The keywords like financial performance and organizational performance suggest a concern with how innovation translates into competitive and financial advantage. The term sustainability bridges firm-level innovation with larger sustainability goals. This cluster resonates with firm-level innovation measurement frameworks linking entrepreneurial capacity to measurable organizational outcomes. It represents the operational layer where innovation policies and entrepreneurship translate into tangible performance.

The node “innovation” is the network’s largest and most connected term, linking economic, institutional, and technological clusters. This reflects innovation’s central theoretical position as both a process (of knowledge generation and application) and an outcome (of systemic and policy interactions). The strong interconnections between clusters suggest that NIS performance assessment is inherently multi-disciplinary:

- Economists measure innovation via efficiency and productivity metrics;
- Policy scholars assess institutional performance and coherence;
- Management researchers study organizational learning and knowledge diffusion; and

- Sustainability studies explore long-term transitions and creative destruction.

This interconnectedness reinforces the need for integrative frameworks that combine quantitative metrics, systemic understanding, and policy evaluation, as proposed collectively by the three referenced studies. From an innovation theory standpoint, the co-occurrence structure illustrates three dominant paradigms:

- Schumpeterian perspective (Schumpeter, 1934, 1942): innovation as a driver of creative destruction and economic transformation;
- Systems of innovation theory: innovation as a systemic process shaped by institutions, networks, and policy frameworks; and
- Measurement and performance theory: innovation as measurable output of R&D, technological development, and commercialisation.

These paradigms correspond directly to the clusters observed in the map and collectively explain how NIS evolve and perform. The cluster analysis confirms that NIS performance assessment spans multiple analytical levels, yet no cluster integrates performance measurement with efficiency analysis within a unified framework.

The four-cluster structure has direct implications for the dissertation's measurement framework. The separation between efficiency-oriented studies (blue cluster) and outcome-focused research (red and yellow clusters) mirrors the performance-efficiency distinction that this dissertation enforces analytically. The absence of a cluster organised around integrated, multi-dimensional assessment, despite extensive work on each component, confirms bibliometrically the structural gap identified through critical literature review in Chapter 2. The framework developed in Chapter 4 responds to this fragmentation by providing an architecture that bridges the clusters' partial perspectives.

3.4 ANALYSIS OF INFLUENTIAL WORKS

The most highly cited publications identified in the bibliometric analysis are presented in Table 8 and Table 9 below. Two complementary citation metrics were used: total citations (lifetime impact) and citations per year (current influence). The comparison highlights overlapping high-impact works, emerging influential papers, and dominant publication venues.

Table 8: Top-cited documents by citations

Authors	Title	Source Title	Year	Cited by
Hagedoorn, John; Cloodt, Myriam M.A.H.	Measuring innovative performance: Is there an advantage in using multiple indicators?	Research Policy	2003	1122

Authors	Title	Source Title	Year	Cited by
Ács, Zoltán J.; Autio, Erkko; Szerb, László.	National Systems of Entrepreneurship: Measurement issues and policy implications	Research Policy	2014	972
Petty, Richard; Guthrie, James.	Intellectual capital literature review: Measurement, reporting and management	Journal of Intellectual Capital	2000	958
Sturgeon, Timothy J.	Modular production networks: A new American model of industrial organization	Industrial and Corporate Change	2002	885
Kivimaa, Paula; Kern, Florian	Creative destruction or mere niche support? Innovation policy mixes for sustainability transitions	Research Policy	2016	826
Antoncic, Bostjan; Hisrich, Robert D.	Intrapreneurship: Construct refinement and cross-cultural validation	Journal of Business Venturing	2001	790
Carlsson, Bo; Jacobsson, Staffan; Holmén, Magnus; Rickne, Annika	Innovation systems: Analytical and methodological issues	Research Policy	2002	771
Beaudry, Catherine; Schiffauerova, Andrea	Who's right, Marshall or Jacobs? The localisation versus urbanization debate	Research Policy	2009	751
Bloom, Nicholas; Draca, Mirko; Van Reenen, John.	Trade induced technical change. The impact of Chinese imports on innovation, IT, and productivity	Review of Economic Studies	2016	742
George, Gerard; McGahan, Anita M.; Prabhu, Jaideep.	Innovation for Inclusive Growth: Towards a Theoretical Framework and a Research Agenda	Journal of Management Studies	2012	676

Source: Own

The following analysis examines the three influential works most relevant to this dissertation's measurement framework: Hagedoorn and Cloodt (2003) on multi-indicator performance measurement, Ács et al. (2014) on systemic entrepreneurship, and Kivimaa and Kern (2016) on policy dynamics.

When further analysing the ten most influential documents per year in Table 9 identified through bibliometric analysis, we identify the emerging research trends that define the modern understanding of NIS performance and NIS itself. Traditional innovation measurement relied heavily on input indicators such as R&D intensity and patent counts. Recent research, however, moves toward systemic and institutional determinants of innovation performance. Ács et al. (2014) propose the National Systems of Entrepreneurship framework as complementary to NIS, emphasizing individual behaviour and institutional quality as systemic drivers of innovation outcomes. Hagedoorn and Cloodt (2003) provide the methodological foundation by demonstrating that multiple indicators such as R&D, patents, and new product announcements should be combined into composite metrics to reflect true innovation capacity. Beaudry and Schiffauerova (2009) survey the localisation-versus-urbanisation debate, showing that whether regional specialisation or diversity better promotes innovation depends critically on industrial classification level and geographical aggregation.

Modern NIS performance research expands from static measurement to dynamic management of innovation systems. Kivimaa and Kern (2016) highlight the role of innovation policy mixes in enabling sustainability transitions, arguing for a balance between niche creation and regime destabilisation, what they term the “motors of creative destruction”. Bloom et al. (2016) link globalisation and trade competition to technological upgrading, showing that external pressures can drive innovation productivity. Together, these works transform performance assessment into a management tool that captures policy coherence, governance effectiveness, and institutional collaboration.

NIS performance is now assessed not only by competitiveness and growth but also by sustainability, inclusivity, and societal well-being. Anu et al. (2023) demonstrate that green innovation and energy efficiency significantly enhance environmental outcomes, making ecological sustainability a core dimension of NIS performance. George et al., (2012) research the concept of inclusive innovation, emphasizing that innovation should generate opportunities for marginalised groups and contribute to social equity. Kivimaa and Kern (2016) further reinforce the sustainability dimension by showing that systemic innovation requires dismantling unsustainable technological regimes. These studies collectively expand NIS performance metrics to include environmental and social impact, aligning NIS performance assessment with global development goals.

Recent literature acknowledges that firm-level innovation capacity aggregates into NIS performance when supported by conducive institutional ecosystems. Lutfi et al. (2023) show that the adoption of big data analytics enhances firm-level performance through organisational readiness and managerial support, illustrating how digital transformation contributes to systemic efficiency. Arsawan et al. (2022) find that knowledge-sharing culture and innovation orientation among SMEs foster sustainable competitiveness, indicating that organisational practices for knowledge management contribute to firm-level innovation capability, which under supportive institutional conditions can aggregate into national innovation capacity. Ács et al. (2014) bridge the firm-system divide by linking entrepreneurial dynamics to institutional structures, reinforcing the ecosystemic view of innovation. Together, these studies show that NIS performance depends on the interdependence between firms, networks, and policy systems.

Table 9: Top-cited documents by citations per year

Authors	Title	Source Title	Year	Citations per year
Kivimaa, Paula; Kern, Florian	Creative destruction or mere niche support. Innovation policy mixes for sustainability transitions	Research Policy	2016	82.6
Ács, Zoltán J.; Autio, Erkko; Szerb, László.	National Systems of Entrepreneurship: Measurement issues and policy implications	Research Policy	2014	81.0
Bloom, Nicholas; Draca, Mirko; Van Reenen, John	Trade induced technical change. The impact of Chinese imports on innovation, IT, and productivity	Review of Economic Studies	2016	74.2
Arsawan, I. Wayan Edi; Koval, Viktor; Rajiani, Ismi; Rustiarini, Ni Wayan; Supartha, I. Wayan Gede; Suryantini, Ni Putu Santi	Leveraging knowledge sharing and innovation culture into SMEs sustainable competitive advantage	International Journal of Productivity and Performance Management	2022	62.25
Anu; Singh, Amit Kumar; Raza, Syed Ali; Nakonieczny, Joanna; Shahzad, Umer	Role of financial inclusion, green innovation, and energy efficiency for environmental performance. Evidence from developed and emerging economies in the lens of sustainable development	Structural Change and Economic Dynamics	2023	56.0
Emerson, Kirk; Nabatchi, Tina	Collaborative governance regimes	Georgetown University Press	2015	53.64
Lutfi, Abdalwali; Alrawad, Mahmaod; AlSyouf, Adi; Almaiah, Mohammed Amin; Khasawneh, Ahmad Y.; Al-Khasawneh, Akif Lutfi; Alshira'h, Ahmad Farhan; Alshirah, Malek Hamed; Saad, Mohamed; Ibrahim, Nahla Mohamad	Drivers and impact of big data analytic adoption in the retail industry: A quantitative investigation applying structural equation modelling	Journal of Retailing and Consumer Services	2023	51.0
Hagedoorn, John; Cloodt, Myriam M.A.H.	Measuring innovative performance: Is there an advantage in using multiple indicators?	Research Policy	2003	48.78
George, Gerard; McGahan, Anita M.; Prabhu, Jaideep C.	Innovation for Inclusive Growth: Towards a	Journal of Management Studies	2012	48.29

Authors	Title	Source Title	Year	Citations per year
	Theoretical Framework and a Research Agenda			
Beaudry, Catherine; Schiffauerova, Andrea	Who's right, Marshall or Jacobs? The localisation versus urbanization debate	Research Policy	2009	44.18

Source: Own

Both datasets share overlapping documents, demonstrating convergence between lifetime and current impact measures. Research Policy dominates as the leading publication venue, followed by multidisciplinary management and economics journals. The 2000-2023 span in both datasets shows that NIS performance research matured after 2000, with notable acceleration in the 2010s. The dual citation analysis approach provides a balanced view: total citations capture historically influential works, while citations per year highlight emerging, high-momentum research directions shaping current debates in NIS performance and NIS itself. The three consistently top-ranked papers: Hagedoorn and Cloodt (2003), Ács et al. (2014), and Kivimaa and Kern (2016) represent the triad of measurement, system, and policy perspectives underpinning the field. Together, these studies offer a comprehensive framework to evaluate how NIS function, perform, and evolve through quantitative indicators, institutional interactions, and policy mechanisms.

Hagedoorn and Cloodt (2003) present one of the most systematic attempts to measure innovative performance using a multi-indicator approach. Their study evaluates nearly 1,200 companies across four high technology industries such as aerospace and defence, pharmaceuticals, electronics, and computing, employing a set of four key indicators: R&D expenditures, patent counts, patent citations, and new product announcements. Each indicator captures a distinct dimension of innovation:

- R&D inputs represent the company’s investment in research and the capacity to generate new knowledge and ideas. They serve as an indicator of innovation effort and potential future output.
- Patents and patent citations serve as proxies for inventive output and the quality of innovation, respectively. Patents indicate the conversion of research efforts into technological inventions, while citations reflect their impact and novelty within the knowledge ecosystem.
- New product announcements represent the commercialisation stage of innovation, capturing how research outputs translate into marketable goods and services.

Using factor analysis, the authors find that all four indicators load strongly on a single common factor, confirming the existence of a latent variable they label as “innovative performance in the broad sense" (Hagedoorn and Cloodt, 2003). This approach acknowledges the non-linear nature of innovation, where inputs, outputs, and outcomes interact dynamically. Moreover, the study underscores that sectoral and national contexts matter: industries vary in their R&D intensity, patenting behaviour, and new product introduction rates, implying that any NIS performance assessment must consider sectoral heterogeneity and technological specialisation. Extending these firm-level findings to NIS performance assessment, this multi-indicator framework offers a methodological foundation for capturing the trajectory from R&D investment to market introduction, while suggesting that researchers can choose indicators based on data availability without major loss of validity in technology-intensive sectors.

While Hagedoorn and Cloodt focus on quantitative indicators at the firm and sectoral levels, Ács et al. (2014) extend the discussion to a systemic, country-level analysis. Their concept of National Systems of Entrepreneurship (NSE) builds on and complements the traditional NIS literature (Freeman, 1987; Lundvall, 1992; Nelson, 1993) by integrating individual-level entrepreneurial behaviour into the system-level analysis of innovation. Ács et al. (2014) define a National System of Entrepreneurship as the institutionally embedded interaction of entrepreneurial attitudes, ability, and aspirations that drives resource allocation through new venture creation. This definition positions entrepreneurship as both an outcome and a driver of innovation system performance. Unlike the NIS framework which emphasises institutional structures such as universities, R&D agencies, and industrial networks, the NSE approach highlights individual agency and institutional interaction as central to innovation outcomes. To operationalise this framework, Ács and colleagues introduce an index, which measures entrepreneurial performance across three dimensions:

- Entrepreneurial attitudes, the cultural and social support for entrepreneurial behaviour, including risk-taking, opportunity perception, and societal values.
- Entrepreneurial ability, the skills, education, and competencies of individuals to recognise and exploit opportunities.
- Entrepreneurial aspirations, the ambition to pursue innovative, high-growth, and international ventures.

The index also includes a penalty for bottleneck mechanisms, which identifies weak system components that constrain overall performance. For instance, even if a country scores highly in entrepreneurial attitudes, deficiencies in financing, education, or regulatory quality can suppress system-level outcomes (Ács et al., 2014). The NSE framework thus integrates institutional quality, individual-level dynamics, and systemic interactions to explain differences in national innovation and entrepreneurship performance. It emphasises that innovation flourishes when entrepreneurial behaviour aligns with supportive institutional frameworks, suggesting that national policies must simultaneously foster individual capabilities and institutional adaptability.

Kivimaa and Kern (2016) contribute to the understanding of innovation system performance by focusing on policy determinants, specifically, how policy mixes for technological innovation systems can foster sustainability-oriented transitions through both creation and destruction mechanisms. Drawing on Schumpeter's concept of creative destruction, they argue that successful innovation policy must not only create innovative technologies and industries but also destabilise and dismantle existing regimes that hinder transformative change (Kivimaa and Kern, 2016). Their framework extends the innovation system approach by introducing the concept of "motors of creative destruction." Traditional innovation system models focus on enabling technological development through functions like knowledge creation, resource mobilisation, and market formation. Kivimaa and Kern expand this to include destabilising functions, such as removing subsidies for dominant regime technologies, tightening regulatory pressure on incumbent industries through control policies, introducing structural reforms to regime rules, and reconfiguring policy networks to displace incumbent actors. In their comparative study of energy efficiency policies in Finland and the United Kingdom, the authors find that both countries exhibit diverse and well-developed policy mixes on the creation side, supporting R&D, market incentives, and network building, but relatively few measures targeting regime destabilisation (Kivimaa and Kern, 2016). The study concludes that the effectiveness of innovation policy mixes for sustainability transitions depends on:

- policy coherence and coordination across domains (innovation, energy, environment, industry);
- the capacity of policy frameworks to undergo restructuring rather than mere layering or drift;
- balanced policy portfolios that simultaneously promote niche innovation and weaken the reproduction of incumbent regimes.

Kivimaa and Kern highlight that the assessment of innovation system performance must move beyond static metrics to consider the dynamic policy interplay between creation and destruction, which they identify as a critical determinant of long-term socio-technical transformation.

The three studies collectively demonstrate that NIS performance emerges from the interplay between quantitative innovation outputs, institutional and entrepreneurial systems, and policy environments. In Table 10 below there is an overview of key contributions of the top-cited documents.

Table 10: Overview of key top cited documents

Dimension	Key Contribution	Reference
Quantitative measurement	Composite indicators capture latent innovative performance across sectors.	Hagedoorn and Cloodt (2003) measuring innovative performance
Systemic and institutional	NSE framework integrates individual entrepreneurship with institutional structures.	Ács, Autio and Szerb (2014) National systems of entrepreneurship
Policy and transition dynamics	Policy mixes must combine innovation creation with regime destabilisation.	Kivimaa and Kern (2016) Creative destruction

Source: Own

Together, these perspectives suggest that NIS performance is multi-layered:

- At the micro level, performance depends on firms' R&D investments and innovative output.
- At the meso (systemic) level, institutional frameworks, entrepreneurship, and network dynamics shape innovation diffusion and sustainability.
- At the macro (policy) level, coherent and dynamic policy mixes determine whether innovation systems evolve or stagnate.

Assessing the NIS performance requires both quantitative rigor and systemic understanding, and Hagedoorn and Cloodt (2003) demonstrate the methodological importance of multi-indicator measurement, highlighting the interconnectedness of R&D, patents, and product innovation. Furthermore, Ács et al. (2014) expand this view by embedding entrepreneurship in a systemic framework that integrates individual entrepreneurial behaviour and institutional context. Kivimaa and Kern (2016) further enrich the analysis by emphasising the dynamic and political nature of innovation policy, where the interplay of niche creation and regime destabilisation drives long-term sustainability transitions. NIS performance is not merely the sum of technological outputs but the product of interactions among policies, institutions, entrepreneurs, and societal goals. Effective assessment frameworks should therefore combine quantitative indicators, systemic diagnostics, and policy evaluation to capture the multifaceted dynamics that underpin innovation and competitiveness.

Among the influential contributions identified in this bibliometric analysis, the dissertation most closely extends the strand represented by studies such as Hagedoorn and Cloodt (2003), which established the principle that innovative performance requires multiple indicators capturing different facets of innovation, a principle this dissertation adopts through the IRPI's multi-indicator structure. However, their framework does not distinguish performance from efficiency, a separation central to this dissertation. Furthermore, Ács et al. (2014) introduced systemic thinking about the relationship between individual-level entrepreneurship and institutional context, anticipating this dissertation's concern with how system conditions shape outcomes. However, their focus on entrepreneurship rather than innovation transformation processes, and their operationalisation of capacity through aggregated observable indicators rather than as a latent interpretive construct, mark points of departure. No work in the top-cited bibliometric corpus integrates result-based performance measurement with two-stage efficiency analysis while treating capacity as a latent interpretive construct, the specific combination this dissertation develops. The bibliometric analysis thus confirms not only that a gap exists, but that addressing it requires synthesising insights from multiple strands rather than extending any single foundational contribution.

3.5 SYNTHESIS OF BIBLIOMETRIC FINDINGS

The bibliometric analysis documents the evolution of NIS performance research from quantitative measurement approaches into a multidimensional discipline integrating systemic, institutional, and policy perspectives. Annual publication output increased from fewer than 50 papers in 2000 to over 300 in 2024, with an average annual growth rate exceeding 17.8% over the past five years. This growth reflects sustained scholarly interest, yet the skewed citation distribution, where a small number of influential works account for most of the influence, indicates that the field's intellectual structure remains concentrated around a few key contributions.

The thematic cluster structure identified through keyword co-occurrence analysis is consistent with the fragmentation documented in Chapter 2. The four clusters: performance measurement, industrial and regional analysis, knowledge management and sustainability, and organisational performance, correspond to distinct research communities employing different indicator sets, analytical methods, and theoretical orientations. The keyword network does not reveal a densely connected core bridging these clusters; instead, each cluster exhibits internal cohesion around its dominant concepts while connections occur primarily through general terms such as "innovation" and "performance".

This pattern aligns with the limitations of the top-cited works analysed in Section 3.4. Hagedoorn and Cloodt (2003) developed multi-indicator approaches for measuring innovative performance but did not integrate efficiency assessment. Ács et al. (2014) proposed a systemic framework but focused on entrepreneurial activity rather than innovation transformation processes. Kivimaa and Kern (2016) examined the dynamic policy conditions under which transformative change occurs but did not operationalise a measurement architecture distinguishing performance from efficiency. Collectively, these contributions address measurement challenges within their respective domains without proposing an integrated framework that simultaneously captures what innovation systems achieve, how efficiently they transform resources, and under which systemic conditions. The policy implications are consequential. As documented in

Chapter 2, composite indices provide benchmarking capacity but obscure transformation processes; efficiency studies assess input-output productivity but abstract from outcome quality, and result-based indicators capture innovation outputs but do not explain how results are generated. The bibliometric mapping confirms that these limitations reflect the broader organisation of the research field itself. The absence of a cluster organised around integrated, multi-dimensional NIS assessment indicates that the diagnostic tools available to policymakers remain partial.

The bibliometric analysis confirms the structural fragmentation identified through critical literature review in Chapter 2. The research field clusters into distinct streams, outcome-oriented composite indicator research and efficiency-oriented DEA research, with limited cross-citation and no dominant cluster organised around integrated, multi-dimensional NIS assessment. The influential works that have shaped the field (Hagedoorn and Cloodt, 2003; Ács et al., 2014; Kivimaa and Kern, 2016) each address measurement challenges within their respective domains but do not propose frameworks that simultaneously capture what innovation systems achieve, how efficiently they transform resources, and under which systemic conditions.

Together, Chapters 2 and 3 have established what is known and what remains unresolved, fulfilling the diagnostic function assigned to them in the dissertation structure (Chapter 1). The evidence justifies a clear conclusion: incremental refinement of existing approaches is insufficient. What is required is a measurement architecture explicitly designed to preserve analytical separation between performance, efficiency, and capacity while enabling their joint interpretation.

Chapter 4 responds directly to this gap. It develops the conceptual framework that structures all subsequent empirical analysis, formally defining the three dimensions, specifying their theoretical relationships, and deriving the methodological requirements that follow from these conceptual distinctions. This framework provides the analytical architecture that guides indicator selection, stage classification, and index construction in Chapter 5.

4 CONCEPTUAL FRAMEWORK: PERFORMANCE, EFFICIENCY, AND CAPACITY

Chapters 2 and 3 established that existing innovation measurement approaches suffer from structural fragmentation: composite indices conflate inputs with outcomes, efficiency studies abstract from result quality, and these research streams have developed largely in parallel. This chapter responds to that gap by developing the conceptual framework that structures the empirical analysis.

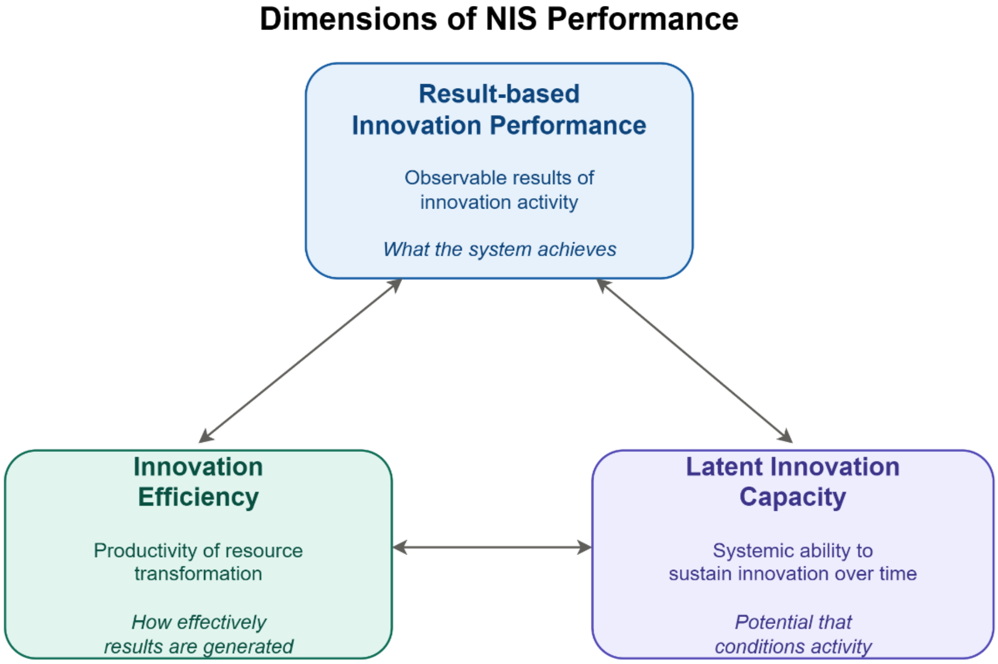
Building on the theoretical foundations established in Chapter 2, which conceptualised innovation as systemic, cumulative, and sequential, this chapter formally defines the three dimensions through which NIS performance and functioning are assessed: result-based innovation performance (observable results), innovation efficiency (transformation processes), and innovation capacity (latent system conditions). The chapter specifies how these dimensions relate theoretically, why only performance and efficiency are directly measured while capacity serves an interpretive function, and what methodological requirements follow from these conceptual distinctions.

The chapter is organised as follows: Section 4.1 develops each dimension in detail, establishing the analytical primacy of performance over efficiency. Section 4.2 examines the determinants that explain cross-country differences in performance and efficiency. Section 4.3 maps the theoretical dimensions to observable indicators, classifying variables into inputs, outputs, outcomes, and impacts. Section 4.4 derives the theoretical expectations and hypotheses that guide empirical testing. This conceptual architecture disciplines the indicator selection, stage classification, and index construction procedures operationalised in Chapter 5.

4.1 DIMENSIONS OF NIS PERFORMANCE ASSESSMENT

NIS are conceptualised along three analytically distinct dimensions: empirically measured result-based innovation performance, empirically measured innovation efficiency, and innovation capacity as a latent systemic property that conditions long-run outcomes (see Figure 7). Recent empirical studies similarly conceptualise NIS through multiple, interrelated dimensions that distinguish between innovation results, the effectiveness or efficiency of innovation processes, and the deeper systemic capacities that support long-term innovation-led development (Bartels et al., 2012; Secka, 2023). Figure 7 illustrates the conceptual structure of the framework, where only performance and efficiency correspond to empirically measured elements, while innovation capacity forms an interpretive layer.

Figure 7: Dimensions of NIS performance



Source: Own

Before developing each dimension in detail, the following definitions establish the framework's analytical structure:

- *Result-based innovation performance* refers to the observable results that a NIS achieves through its activities, specifically, the knowledge and technology outputs produced (Stage 1), and the commercial and economic outcomes realised (Stage 2). Performance captures *what* the system achieves, not the resources it deploys. It is empirically measured through the IRPI.
- *Innovation efficiency* refers to how productively a NIS transforms available resources into results. It captures the effectiveness of transformation processes relative to inputs deployed, assessed separately for knowledge production (Stage 1 efficiency) and commercialisation (Stage 2 efficiency). Efficiency captures *how effectively* results are generated. It is empirically measured through the IEI.
- *Innovation capacity* refers to the latent systemic ability of a NIS to sustain innovation processes over time. It encompasses accumulated capabilities, institutional coherence, and structural connectivity. Capacity conditions innovation activity but does not constitute performance in itself, as it reflects *potential* rather than observed results. Capacity is not directly measured but is inferred from the persistence of performance-efficiency configurations.

These three dimensions are analytically distinct but interrelated: capacity enables innovation activity, efficiency governs how capacity is transformed into results, and performance represents the realised outcomes. This multidimensional perspective aligns with the evolutionary and systemic traditions (Nelson and Winter, 1982; Lundvall, 1992; Nelson, 1993; Metcalfe, 1998), as discussed in Chapter 2. In this framework, performance and efficiency are empirically measured, while capacity is treated as a latent systemic construct that informs interpretation rather than a directly measured dimension. The framework maintains strict separation between dimensions while recognising their sequential relationship. The next sections develop each dimension in full detail. Table 11 below presents the overview of the three dimensions.

Table 11: Core definitions of NIS assessment dimensions

Dimension	Definition	Analytical Focus	Measurement
Performance	Observable results of innovation activity (outputs and outcomes)	*What* the system achieves	IRPI (empirical)
Efficiency	Productivity of resource transformation across two stages	*How effectively* results are generated	IEI (empirical)
Capacity	Latent systemic ability to sustain innovation over time	*Potential* that conditions activity	Inferred from persistence (interpretive)

Source: Own

The choice of three dimensions: performance, efficiency, and capacity, emerges from what innovation systems theory requires us to analytically distinguish. This structure is neither arbitrary nor conventional. Each dimension answers a fundamentally different question about NIS functioning:

- Performance answers: What does the system achieve? This captures observable results: the outputs and outcomes that constitute the ultimate purpose of innovation systems.
- Efficiency answers: How productively are resources transformed? This captures process quality: how close transformation operates to best practice, which cannot be inferred from results alone.
- Capacity answers: What enables sustained innovation over time? This captures latent potential: the systemic conditions that explain why countries with similar resources often achieve markedly different results.

These questions are irreducible: none can be derived from or collapsed into the others without losing essential analytical information. Reducing the framework to two dimensions would force problematic conflations. Dropping capacity would leave the framework unable to explain persistent cross-country differences despite similar observable inputs and efficiency levels, a core puzzle that motivated NIS theory (Freeman, 1987; Lundvall, 1992). Dropping efficiency would conflate resource deployment with transformation quality, obscuring whether high performance reflects productive processes or mere resource abundance. Dropping result-based performance would abandon the outcome-oriented logic that defines NIS performance assessment: a system's value lies in what it achieves, not merely in its potential or process quality. Additional dimensions would introduce either redundancy or unmeasurable constructs. Treating inputs as a fourth dimension would contradict the framework's foundational premise that inputs have no analytical value independent of results, but they are enablers assessed through efficiency, not a separate evaluative criterion (Edquist, 2011). Treating societal impacts as a fourth dimension would extend the framework beyond tractable measurement: impacts materialise over extended time horizons, are shaped by non-innovation factors, and cannot be reliably attributed to system functioning within the scope of NIS assessment. Treating Stage 1 and Stage 2 efficiency as separate dimensions would conflate methodological decomposition with conceptual distinction, as they answer the same analytical question (transformation productivity) at different points in the innovation process.

The three-dimensional structure thus satisfies the principle of analytical parsimony: it is the minimum necessary to capture what NIS theory establishes about how innovation systems function, without introducing redundancy or unmeasurable elements. Fewer dimensions would sacrifice diagnostic capacity, while more would compromise coherence. This structure aligns with recent multidimensional NIS frameworks that similarly distinguish between results, process effectiveness, and systemic conditions (Bartels et al., 2012; Secka, 2023), while making the theoretical rationale for dimensional choice explicit.

Furthermore, the conceptual framework establishes an analytical asymmetry between performance and efficiency that requires explicit justification. Performance is defined as the observable results of innovation activity, and it occupies conceptual primacy because innovation systems exist to generate these results. As Chapter 1 established, drawing on Schumpeter (1934), innovation is defined by its economic application and commercialisation, an invention that is not transformed into marketable outcomes is not, in Schumpeterian terms, an innovation at all. Freeman (1987), Lundvall (1992), and Nelson (1993) each emphasise that innovation systems are ultimately judged by what they produce: scientific knowledge, technological capabilities, and economic value.

This result-oriented logic has found growing recognition in both academic literature and policy practice. As documented in Chapter 2 (Section 2.3.2), a distinct strand of scholarship argues that "NIS performance should be assessed through the results and consequences of innovation processes rather than through inputs or framework conditions" (Edquist et al., 2018; Janger et al., 2017). The IOI reflects this same conceptual commitment, designed explicitly to focus on observable innovation outputs rather than resource endowments or framework conditions (Vértesy and Tarantola, 2014).

However, result-based assessment alone cannot explain how outcomes are generated or whether they reflect efficient transformation processes. This is where efficiency enters as a complementary but conceptually subordinate dimension. Efficiency captures how productively transformation processes operate, but efficiency without outcomes is analytically empty. A system that transforms resources with maximal productivity but generates no valuable outputs has achieved nothing from an innovation policy perspective. Equally, a system that generates strong outcomes with moderate efficiency may be better positioned than one with high efficiency but weak results. As Edquist (2011) argues, innovation systems should be evaluated by their ability to perform essential functions, and the ultimate function is the generation of innovation results.

This asymmetry has direct implications for how performance and efficiency should be conceptually integrated. Performance must serve as the primary ordering criterion; efficiency refines our understanding of performance-based differences but cannot substitute for observed results. A highly efficient system with low performance should not be judged superior to a moderately efficient system with high performance. This "performance-first" principle operationalises the theoretical insight that results are the “reason for existence” of innovation systems, while efficiency is a desirable but subordinate property.

This primacy of performance over efficiency distinguishes the present framework from efficiency-only approaches that, as documented in Chapter 2 (Section 2.3.3), "treat efficiency scores as proxies for overall performance, despite capturing only one aspect of system functioning". It equally distinguishes the framework from input-heavy composite approaches such as the SII, which, as Chapter 2 documented, conflate enabling conditions with realised outcomes. By maintaining the conceptual distinction while establishing their hierarchical relationship, the framework avoids both conflation (treating efficiency as performance) and fragmentation (analysing them in isolation).

4.1.1 Result-based innovation performance

Contemporary measurement frameworks treat NIS performance as the observable outputs and outcomes of innovation activities. This distinction has been central to innovation theory since Schumpeter’s conception of innovation as the creation of new combinations that transform production and markets (Schumpeter, 1934, 1942). The NIS literature subsequently reinforced the idea that the effectiveness of an innovation system must be evaluated based on the results it produces, its knowledge outputs and its commercial and technological outcomes, rather than on the scale of its inputs (Freeman and Soete, 1997; Lundvall, 1992; Nelson, 1993; Edquist, 2011, Edquist et al., 2018).

Empirical analyses confirm this distinction: Barrichello et al. (2020) find that countries' positions in global competitiveness rankings are strongly associated with innovation output factors rather than input intensity alone, underscoring the importance of measuring performance through tangible outputs and outcomes.

A well-performing NIS is characterised first by its capacity to generate valuable, relevant and internationally visible knowledge outputs. These include scientific publications, citations, patents, designs, prototypes and firm-level innovations. The breadth and quality of these outputs reflect the strength of a country’s scientific and technological capabilities. However, the literature emphasises that the volume of outputs alone is insufficient for assessing performance: the relevance, quality and economic potential of knowledge matter equally, if not more (Archibugi and Coco, 2005). Multivariate analyses based on the Global Competitiveness Report show that the quality of scientific research institutions and the number of PCT patent applications are among the most decisive factors for moving countries into higher stages of development, confirming that knowledge and technology outputs are central markers of national performance (Barrichello et al., 2020).

Performance is also reflected in the system’s capacity to translate these outputs into economic and technological outcomes. The Schumpeterian and evolutionary traditions stress that knowledge creation only becomes innovation when it is successfully commercialised or diffused (Schumpeter, 1934; Dosi, 1988). Consequently, outcomes such as the share of turnover from innovative products, the penetration of advanced technologies, the dynamism of entrepreneurial and high-growth firms, and the export of high-tech goods are key indicators of system performance. These measures capture the ability of the NIS to coordinate knowledge flows, selection processes and market dynamics in ways that allow new technologies to enter production and reach users (Metcalfe, 2005). This two-stage understanding of performance is mirrored in the GII innovation output subindex, which groups indicators into knowledge and technology outputs and creative outcomes, thereby treating commercial and economic manifestations of innovation (innovative sales, technology exports, creative goods and services) as the ultimate expression of NIS performance (WIPO et al., 2019).

Within this framework, result-based innovation performance is strictly measured through Stage 1 knowledge and technology outputs and Stage 2 commercial and economic outcomes. Broader societal or macroeconomic impacts remain conceptually relevant but fall outside the empirical scope of this dissertation.

Although this dissertation does not incorporate societal or macroeconomic impacts into the empirical part, the broader innovation literature recognises that innovation activities generate effects that extend far beyond immediate outputs and outcomes. These impacts provide important context for understanding the significance of innovation results and for situating national performance within longer-term developmental trajectories. However, because such impacts typically materialise over extended time horizons and are shaped by complex interactions among economic, institutional and societal forces, they fall outside the operational focus of this dissertation.

At a broader level, a well-performing innovation system contributes to productivity and competitiveness by supporting sectors and activities that embody higher value added and stronger technological capabilities (Fagerberg, 1994, 2000; Castellacci, 2008). Such systems adapt to structural change, harness emerging technological paradigms and respond effectively to disruptions. Resilience is the capacity to maintain or redirect innovation activities during crises and has therefore become an increasingly important facet of system performance (Filippetti and Archibugi, 2011).

4.1.2 Innovation efficiency

Efficiency reflects how close an innovation system operates to best practice (Farrell, 1957; Edquist et al., 2018). When applied to innovation systems, this concept evaluates how well inputs such as R&D expenditure, human capital and research infrastructure are converted into innovation outputs such as patents, trademarks and scientific publications, and ultimately into innovation outcomes such as productivity gains and competitiveness.

Evolutionary and systemic theories further contextualise efficiency by portraying innovation as a distributed, dynamic and uncertain process that unfolds through interactions among heterogeneous actors, institutions and knowledge flows (Lundvall, 1992; Metcalfe, 1995). Within this view, high input levels are neither necessary nor sufficient for strong innovation performance, but what matters is the system’s capacity to translate available resources into valuable results through processes of learning, search, coordination and diffusion. A system may therefore be input-rich but inefficient, or input-poor yet efficient, depending on its structural and functional characteristics (Castellacci and Natera, 2011). A key implication of this theoretical grounding is that inputs have analytical relevance only when viewed relative to the results they produce. Inputs are thus treated in this dissertation not as indicators of performance but as enablers whose effectiveness must be assessed through the lens of efficiency. This resonates with Edquist's (2011) argument that innovation inputs become meaningful only insofar as they contribute to the execution of key innovation system functions.

NIS do not transform inputs into results through a single, unified process. Rather, innovation unfolds through a sequence of interconnected stages, knowledge and technology production (Stage 1) followed by knowledge and technology commercialisation (Stage 2). This cumulative and path-dependent structure has been recognised in innovation research for decades (Kline and Rosenberg, 1986; Metcalfe, 1995). Modern empirical approaches such as two-stage and network DEA model this sequential structure explicitly, emphasising the multiple transformation processes embedded within innovation systems (Guan and Chen, 2012; Liou, 2009; Carayannis et al., 2016; Anouze et al., 2024). These recent applications of multi-stage and network DEA models further confirm that modelling innovation as a sequential system of knowledge production followed by commercialisation yields more realistic efficiency profiles than one-stage approaches.

Based on the efficiency literature reviewed in Chapter 2 (Section 2.3.3), the framework decomposes efficiency into two stages: the knowledge/technological production stage assesses how effectively R&D expenditure, research personnel, and infrastructural resources are transformed into innovation outputs. It aligns with the literature demonstrating that scientific productivity depends not on input magnitude alone but on organisational routines, learning processes, and system coherence (Nelson, 1993; Freeman and Soete, 1997;). On the other side, the knowledge/technological commercialisation stage evaluates how successfully innovation outputs are translated into innovation outcomes. It draws on evidence that technological opportunities, absorptive capacity, and diffusion mechanisms determine how much economic value can be extracted from knowledge (Cohen and Levinthal, 1990; Fagerberg and Verspagen, 2002). This two-stage efficiency dimension of NIS forms the entire analytical basis of the IEI.

By adopting a two-stage perspective, this dissertation conceptualises efficiency in a way that mirrors the functioning of real-world innovation processes. Stage 1 efficiency captures the productivity of knowledge production, while Stage 2 efficiency captures the productivity of commercialisation and diffusion. Together, they allow a granular assessment of the sources of inefficiency within NIS, whether they arise from weak research productivity, bottlenecks in commercialisation or structural misalignments between technological capabilities and market opportunities.

In line with the theoretical foundations discussed above, this dissertation adopts a well result-based conceptualisation of innovation efficiency: inputs matter only to the extent that they are converted into valuable outputs and outcomes. A system that deploys large resources but generates weak results is inefficient, whereas one that produces strong results with limited inputs demonstrates high efficiency. This approach avoids conflating resource intensity with success and aligns with Edquist's (2011) argument that inputs are instruments whose value depends on their contribution to system results. This result-oriented view is also consistent with recent multidimensional assessments of NIS performance where government effectiveness, education spending and labour-force characteristics are treated as predictors of development outcomes rather than performance indicators themselves (Secka, 2023).

The two-stage conceptualisation of efficiency carries inherent temporal implications. The transformation of R&D expenditure, research personnel, and infrastructure investments into knowledge outputs (publications, patents, prototypes) requires time for research projects to mature, experiments to yield results, and findings to be codified and disseminated. The innovation studies literature (Mansfield, 1991; Carayannis et al. 2016; Guan and Chen, 2012) suggests this process typically spans 1-3 years, though considerable variation exists across sectors and technology types. However, Guan and Chen (2012) emphasised that while aggregate effects may appear stable, the appropriate lag structure should reflect the actual temporal dynamics of each sub-process. The OECD (2015) Frascati Manual acknowledges that there is typically a time lag between R&D expenditure and the resulting outputs. Additionally, the transformation of knowledge outputs into commercial outcomes (innovative sales, technology exports, employment in knowledge-intensive activities) involves additional processes: technology transfer, market development, diffusion, and adoption. This stage typically requires 1-4 years depending on market conditions, regulatory environments, and technology complexity. This range reflects the fundamental uncertainty and non-linearity of innovation processes emphasised by evolutionary economics (Nelson and Winter, 1982). However, the precise lag structure cannot be determined purely from theory. It depends on:

- the specific indicators used (patents emerge faster than publications, innovative sales respond faster than structural employment shifts),
- the sectoral composition of national economies,
- institutional factors affecting technology transfer speed, and
- data availability and measurement frequency.

The operationalisation of time lags in the empirical analysis (Chapter 5) therefore reflects both theoretical guidance and practical data constraints.

4.1.3 Innovation capacity

This section develops the concept of innovation capacity as a latent and structural property of NIS. The purpose is not to operationalise capacity as a measurable index, but to clarify the mechanisms through which it conditions observed performance and efficiency outcomes. As established in Chapter 2, innovation capacity refers to underlying potential rather than observed results. It is therefore treated as a latent analytical construct in this dissertation, not as a directly measured dimension. The theoretical basis for treating capacity as latent was established in Chapter 2 (Section 2.2), where the identification problem, that any observable capacity proxy simultaneously functions as an input or output, was documented. Building on this foundation, the present framework defines innovation capacity through three interrelated system properties:

1. Accumulated capabilities: the stock of knowledge, skills, routines, and organisational competencies that enable actors within the system to generate, absorb, and apply new knowledge. As established in Chapter 2, Nelson and

Winter (1982) conceptualise these as firm-level routines and learning processes that aggregate to system-level capability stocks. Capacity in this sense reflects what the system knows how to do.

2. Institutional coherence: the alignment, stability, and coordination quality of formal and informal institutions, including regulatory frameworks, intellectual property regimes, funding mechanisms, and governance structures, that shape incentives and reduce uncertainty for innovation actors. Following Freeman (1987) and Edquist (2011), institutional coherence determines whether knowledge flows smoothly across actor boundaries and whether coordination failures impede transformation processes. Capacity in this sense reflects how well the system is organised.
3. Structural connectivity: the density, quality, and configuration of networks linking firms, universities, research organisations, and public agencies. As Lundvall (1992) emphasises, innovation emerges through interactive learning within user-producer relationships and collaborative networks. Structural connectivity determines whether knowledge generated in one part of the system can be accessed, recombined, and exploited elsewhere. Capacity in this sense reflects how well the system is connected.

Operationally, innovation capacity manifests through the persistence of performance-efficiency configurations over time. Systems with high-capacity exhibit stable or improving performance-efficiency positions across multiple years, demonstrating resilience to shocks and the ability to sustain innovation trajectories. Systems with low-capacity exhibit volatile positioning, declining performance despite maintained inputs, or persistent inability to convert knowledge outputs into commercial outcomes. This temporal signature, with stability versus volatility in the performance-efficiency space, provides the diagnostic basis for inferring capacity from observable patterns in Chapter 7.

The rationale for treating capacity as latent, rather than directly measured, rests on both the identification problem documented in Chapter 2 and on a fundamental stock-flow distinction. The distinction between capacity and inputs is fundamental. Inputs (R&D expenditure, research personnel) are flow variables that can be adjusted in the short term through policy or market decisions. Capacity is a stock variable that accumulates slowly through sustained institutional development, human capital formation, and network building. High inputs do not guarantee high capacity, as Chapter 1 established when noting that countries with similar resource endowments often exhibit markedly different innovation results (Freeman, 1987; Lundvall, 1992; Nelson, 1993). Equally, high capacity may coexist with temporarily constrained inputs yet still manifest in superior transformation efficiency. This asymmetry between stocks and flows is why capacity cannot be directly measured through input indicators without conflation.

The capacity dimension therefore captures systemic readiness, including the ability to accumulate knowledge, adapt to new technological paradigms, and sustain innovation-based growth. It is not quantified as an index in this dissertation, but it serves as a key analytical construct for interpreting both IRPI and IEI patterns in Chapter 7. Case studies illustrate that even when observable performance indicators remain low, underlying institutional reforms, human-capital investments and network-building efforts can gradually strengthen innovation capacity, which only materialises in performance indicators with considerable time lags (Secka, 2023).

A central claim of this conceptual framework is that innovation capacity manifests through the persistence of performance-efficiency configurations over time. The link between capacity and persistence derives from three properties of innovation systems established in Chapter 2:

- Path dependence: Nelson and Winter (1982) demonstrate that innovation trajectories are historically conditioned, current capabilities build on past learning, and organisational routines exhibit strong inertia. Systems with high accumulated capabilities possess deeper repertoires of problem-solving routines that buffer against shocks and sustain performance trajectories. This path dependence implies that capacity-rich systems should exhibit stability in their relative positioning over time.
- Institutional stickiness: As Freeman (1987) establish, institutions change slowly. The formal and informal rules that coordinate innovation activity, such as property rights regimes, funding mechanisms, governance structures, collaborative norms, evolve over years or decades, not months. Systems with high institutional coherence benefit from stable, predictable environments that support long-term investment in innovation. This institutional stickiness implies that capacity differences between systems should persist across observation periods.
- Network cumulation: Lundvall's (1992) emphasis on interactive learning implies that knowledge networks accumulate over time. Relationships between firms, universities, and research organisations deepen through repeated interaction, the trust builds, and knowledge channels become more efficient. Systems with high structural connectivity benefit from these accumulated network assets, which cannot be rapidly created or destroyed. This cumulative property implies that connectivity advantages should endure across multiple years.

Together, these mechanisms generate a testable prediction: systems with high capacity should exhibit persistent positioning in the performance-efficiency space, while systems with low capacity should exhibit volatile positioning. The empirical strategy does not measure capacity directly, which would violate the framework's epistemological commitments, but infers capacity from the observable signature it produces persistence. The inference is diagnostic rather than definitive: persistent high performance-efficiency configurations provide evidence consistent with high capacity, while volatility provides evidence of capacity constraints. This analysis provides the interpretive foundation for understanding why some countries maintain stable innovation trajectories while others exhibit structural volatility despite similar resource endowments.

Beyond persistence, innovation capacity can be indirectly validated through its expected relationship with observable structural and institutional characteristics. If capacity reflects accumulated capabilities, institutional coherence, and structural connectivity as defined above, then countries with high latent capacity should exhibit not only persistent performance-efficiency configurations but also favourable values on variables that theory associates with these properties, such as economic development levels, governance quality, and human capital stocks. These observable characteristics do not constitute capacity itself, but they serve as correlates that should co-vary with the latent construct. A framework that correctly captures innovation system functioning should therefore produce performance-efficiency scores (IE_IRPI) that correlate positively with such capacity-related variables. This theoretical expectation provides a basis for empirical validation: the relationship between measured dimensions and external determinants can be further tested, as developed in Chapter 6. Section 4.2 elaborates on the specific determinants, both internal and external, that shape innovation capacity, and explains how these factors provide the interpretive foundation for understanding cross-country variation in the empirical results.

For these reasons, innovation capacity is treated in this dissertation as a latent conditioning framework that guides interpretation of empirical results, rather than as a separately quantified dimension.

Having defined the three dimensions: performance, efficiency, and capacity, and having established their analytical relationships, the framework must now address why countries differ in their observed outcomes. While dimensions specify *what* is measured, determinants explain *why* systems vary. The following section identifies the structural, institutional, and contextual factors that shape NIS functioning and provide the interpretive foundation for understanding cross-country variation in Chapter 7.

4.2 DETERMINANTS OF NIS PERFORMANCE

As established in Section 4.1, innovation capacity is a latent systemic property that cannot be directly measured but can be indirectly validated through its relationship with observable structural and institutional characteristics. This section identifies and categorises these characteristics, termed determinants, which shape cross-country variation in performance, efficiency, and capacity. Determinants are defined as the underlying structural, institutional, behavioural, and contextual factors that shape the generation, diffusion, and utilisation of knowledge within an NIS. They include scientific and technological capabilities, human capital, institutional quality, network structures, firm-level capabilities, financial and market conditions, and the degree of global integration (Fagerberg and Srholec, 2008; Edquist, 2011; OECD, 2018). Conceptually, they correspond to the "functions" of innovation systems identified by Edquist (2011): those activities and conditions that enable or hinder the creation, dissemination, and exploitation of innovations.

Determinants serve a dual analytical role in this dissertation. First, they provide the interpretive vocabulary for explaining why countries differ in their performance-efficiency configurations, a diagnostic function developed in Chapter 7. Second, they serve as observable correlates of latent capacity, enabling indirect validation of the framework's theoretical structure through correlation and regression analysis in Chapter 6. This dual role requires that determinants remain strictly separated from the measured dimensions: they do not enter the construction of IRPI or IEI, preserving the distinction between what is measured (performance, efficiency) and what explains variation in those measures (determinants reflecting capacity).

Consistent with the systems-of-innovation perspective, it is useful to distinguish between internal and external determinants. Internal determinants refer to features largely located within the national system, such as research capacity, skills, institutions, collaboration networks and innovation culture. External determinants refer to the broader environment in which the NIS operates, including global competition, international knowledge flows, macroeconomic conditions and geopolitical factors. The interaction between internal and external determinants reinforces the systemic nature of innovation: outcomes arise from networks of interdependent actors, institutions and environments, not from isolated variables (Edquist, 2011; Malerba and Nelson, 2011). The following subsections synthesise the main determinant families discussed in the evolutionary and systemic innovation literature.

4.2.1 Internal determinants

A first internal determinant concerns the depth and quality of a country’s scientific and technological capabilities. Schumpeter (1942), Freeman (1987) and Nelson (1993) emphasise that strong research environments including universities, public laboratories, and technologically advanced firms generate more, and more diverse, technological opportunities. Comparative studies show that variations in research excellence, availability of advanced infrastructure and concentration of scientific talent strongly influence the volume and quality of knowledge outputs such as publications, patents and prototypes, as well as the potential for their economic exploitation (Jaffe and Trajtenberg, 2002; Archibugi and Coco, 2005; Hall et al., 2005). These capabilities therefore affect both performance (in terms of results) and the underlying capacity of the system to sustain innovation in the long run. Multivariate analyses based on World

Economic Forum indicators demonstrate that the quality of scientific research institutions and the intensity of international patenting (PCT applications) are among the strongest factors discriminating between more and less developed countries, underlining the centrality of scientific and technological capability for national development (Barrichello et al., 2020).

Human capital is another core determinant. It encompasses not only the number of R&D personnel and STEM graduates, but also the broader skills base necessary to absorb, adapt and apply knowledge. Griliches (1990) shows that the productivity of R&D expenditure depends critically on the quality of the scientific workforce. At the same time, broader education and training systems shape the capacity of firms and organisations to adopt new technologies and reorganise production processes. Secka's (2023) analysis of NIS shows that public spending on education and labour-force characteristics significantly affect development indicators such as GDP, GNI and HDI, highlighting the importance of human capital formation as a foundation for innovation-led development. As such, human capital influences both the creation of knowledge and its diffusion and is a key component of latent innovation capacity (Castellacci and Natera, 2011).

Institutions, including regulatory frameworks, intellectual property regimes, public administration and policy coordination mechanisms, shape the incentive structure and predictability of the innovation environment. Nelson (1993) and Edquist (2011) argue that coherent, credible and stable institutions are fundamental for the long-term development of innovation capabilities. Effective governance reduces uncertainty, facilitates long-term investment and helps align public and private efforts. Conversely, weak or fragmented institutions can limit both result-based innovation performance and efficiency, for example by generating bottlenecks in funding flows, slowing decision-making or creating inconsistent policy signals. Quantitative evidence confirms that government effectiveness is a significant predictor of development performance (Secka, 2023), while recent cross-country work links robust governance structures directly to stronger innovation outcomes and more efficient NIS functioning (Jankowska et al., 2017).

A central insight from the NIS literature is that innovation is an inherently **i**nteractive process (Lundvall, 1992). The density and quality of collaboration networks among firms, universities, research institutes and public agencies affect the intensity of knowledge exchange and collective learning. Strong university-industry linkages, inter-firm cooperation and participation in research consortia enhance opportunities for recombining knowledge, reduce duplication of effort and support cumulative learning (Metcalfe, 1995). These network structures thus influence both the level of outputs and outcomes, and the productivity with which they are generated. Studies emphasise that collaboration among firms, universities and government agencies, often framed in Triple or Quadruple Helix or broader network terms, significantly enhances national innovation capacity by improving knowledge flows and generating positive externalities (Etzkowitz and Leydesdorff, 2000; Carayannis and Campbell, 2009; Jovanović et al., 2022).

At the micro level, firm strategies, routines and capabilities are crucial determinants of NIS performance. Evolutionary theories highlight that firms differ in their search processes, organisational routines and ability to learn (Nelson and Winter, 1982). Cohen and Levinthal (1990) show that absorptive capacity, the ability to recognise, assimilate and exploit external knowledge, is central to firms' innovation performance. Firms with strong absorptive and dynamic capabilities are more likely to generate new products and processes and to adopt external innovations effectively, thereby affecting both result-based performance and the effective use of the knowledge base. Firm-level case studies in emerging and developing economies similarly show that managerial capabilities, organisational learning routines and internal R&D efforts are decisive for transforming national knowledge bases into concrete innovations (Gogodze, 2016).

Although financial systems and product markets are also shaped by external forces, many of their features are determined by national regulations and institutional legacies. Access to credit, equity finance and venture capital significantly affects firms' ability to invest in and scale innovation. Competitive and open markets can strengthen incentives for innovation and accelerate the diffusion of new technologies, while highly concentrated or protected markets may reduce the pressure to innovate (Metcalfe, 2005). These arrangements influence not only the number and type of innovations that reach the market but also the efficiency with which knowledge is transformed into economic outcomes.

4.2.2 External determinants

External knowledge flows through trade, foreign direct investment, international R&D collaboration and mobility of researchers are key determinants of NIS performance and capacity, particularly for catching-up economies. Malerba and Nelson (2011) and Archibugi and Coco (2005) highlight that openness to international knowledge sources facilitates technological upgrading by exposing domestic actors to advanced technologies and practices. Openness can enhance both knowledge production (through imported technologies and collaborative research) and commercial outcomes (through participation in global value chains and export markets). Evidence from developing and emerging economies confirms that integration into global value chains and international production networks is associated with higher innovation output levels and faster technological upgrading (Secka, 2023).

Macroeconomic stability, investment levels and growth patterns affect the resources available for innovation and the willingness of firms and governments to undertake risky, long-term innovation projects. Fagerberg (1994, 2000) and Fagerberg and

Verspagen (2002) show that innovation and growth are mutually reinforcing, and that countries trapped in low-growth regimes often struggle to sustain adequate innovation investments. Economic crises or persistent macroeconomic volatility can reduce innovation efforts and weaken the long-term capacity of the system, even if underlying capabilities exist.

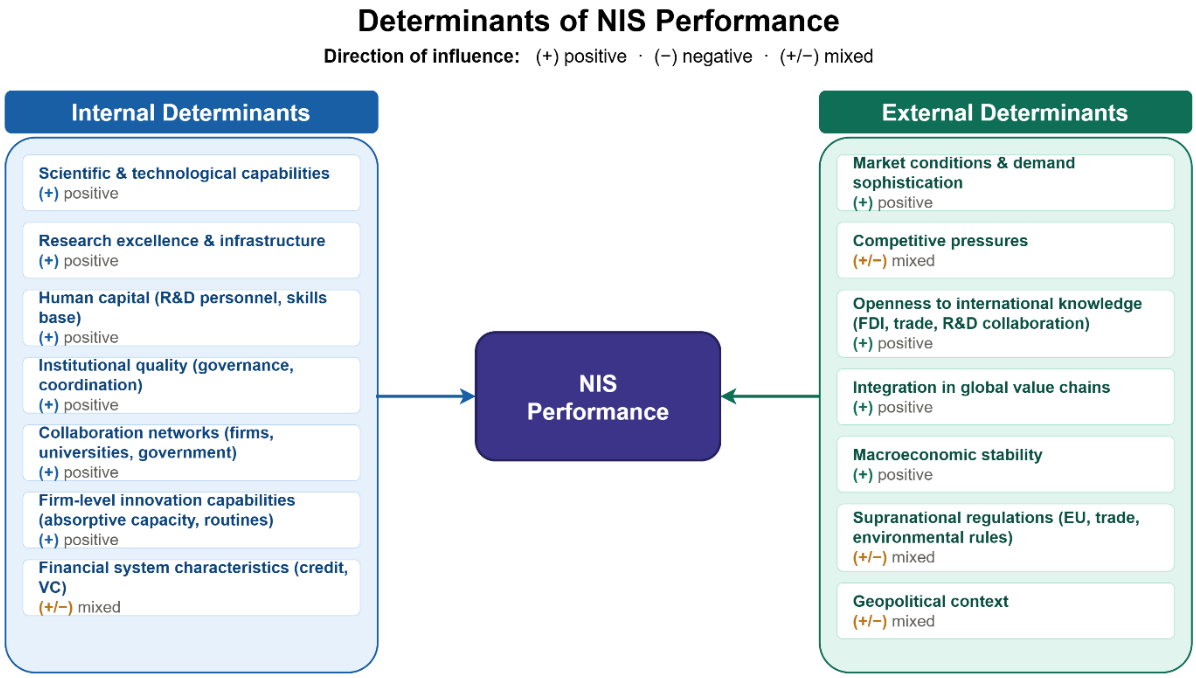
Demand conditions and competitive pressures are also shaped by external factors such as access to foreign markets and integration into global production networks. Porter (1990) and subsequent work on demand-side innovation policy show that sophisticated and demanding customers, both domestically and abroad, push firms to innovate. Exposure to international competition similarly forces firms to improve products and processes, adopt new technologies and explore new markets. These forces intensify selection mechanisms and influence which innovations succeed and diffuse.

Finally, geopolitical developments and supranational regulatory frameworks (e.g. EU competition policy, trade agreements, environmental regulations) form part of the external context in which NIS operate. They create constraints and opportunities that can significantly alter the trajectory of sectors and technologies, thereby affecting performance and capacity at national level.

4.2.3 The role of determinants in the conceptual framework

The overview of key determinants together with the direction of influence on NIS performance (where “+” is positive relationship, and “-” is negative relationship) are presented in Figure 8 below.

Figure 8: Determinants of NIS performance



Source: Own

The determinant categories are not treated as additional dimensions and do not enter the construction of the IRPI and IEI. This analytical use of determinants, rather than their inclusion as index components, is consistent with recent multidimensional NIS studies that separate measurement tools (indices) from explanatory variables used for interpretation and policy design (Secka, 2023; Bartels et al., 2012). Their role in this dissertation is strictly analytical:

- they help to explain why some countries achieve higher or lower values on the dimensions of performance, efficiency and capacity introduced in Section 4.1;
- they provide a theoretically grounded basis for interpreting patterns and typologies of NIS revealed in the empirical analysis; and
- they support the identification of structural and systemic bottlenecks that cannot be detected from index scores alone.

The determinants are not merely explanatory variables for interpreting empirical results, but also provide the basis for indirectly validating the latent capacity construct defined in Section 4.1. This validation mechanism rests on a theoretically grounded expectation: if innovation capacity reflects accumulated capabilities, institutional coherence, and structural connectivity, then observable determinants associated with these properties should correlate systematically with the framework's measured dimensions. This theoretical expectation constitutes a falsification criterion. If IE_IRPI exhibited weak or negative correlations with GDP per capita (proxy for accumulated capabilities) and government effectiveness (proxy for institutional coherence), the framework's construct validity would be called into question. An index claiming to measure innovation system functioning that shows no systematic relationship with the structural foundations theory identifies as determinants would be capturing noise rather than meaningful variation. Chapter 6 evaluates this criterion empirically: positive correlations exceeding r=0.30 with theoretically expected signs would support construct validity; weak or wrongly signed relationships would indicate falsification.

Each component of innovation capacity identified in Section 4.1 corresponds to specific observable determinants:

- Accumulated capabilities are reflected in determinants such as GDP per capita (capturing the stock of economic resources and infrastructure), human capital endowments (the skills and knowledge base available for innovation), and scientific and technological capabilities (the depth of the research system). Countries with high accumulated capabilities possess the foundational resources that enable sustained innovation activity.
- Institutional coherence is reflected in determinants such as government effectiveness (the quality of policy formulation and implementation), regulatory quality (the predictability and consistency of the institutional environment), and the stability of intellectual property regimes. Countries with high institutional coherence provide the governance foundations that reduce uncertainty and support long-term innovation investment.
- Structural connectivity is reflected in determinants such as university-industry collaboration intensity, network density among innovation actors, and participation in international knowledge flows. Countries with high structural connectivity benefit from effective channels for knowledge exchange, recombination, and diffusion.

This mapping generates a testable prediction: countries that score high on the integrated performance-efficiency index (IE_IRPI) should also exhibit favourable values on capacity-related determinants. The theoretical logic is that high capacity enables both strong performance and efficient transformation, manifesting simultaneously in high IE_IRPI scores and favourable determinant values. Therefore, IE_IRPI and capacity-related determinants should correlate positively.

This expected correlation does not imply simple linear causation, as the relationships are complex, reciprocal, and path dependent. Rather, positive correlation would validate that the framework's measured dimensions capture variation systematically related to the deeper systemic properties that innovation theory identifies as foundations of innovation capacity. Chapter 5 specifies how this validation is operationalised, and Chapter 6 reports the empirical results.

The determinants are structural conditions, while the indicators presented in 4.3 are observable measures. Human capital as a determinant refers to the depth, quality, and institutional embeddedness of a country's skill formation system, which is the underlying condition that enables innovation. Human capital indicators (e.g., STEM graduates per 1,000 population) are observable measures that capture one manifestation of this condition. The determinant is the enabling structure while the indicator is a measurable symptom (Edquist, 2011; Lundvall, 1992).

Furthermore, the determinants typically represent stocks while indicators capture flows. Following evolutionary economics (Nelson and Winter, 1982), determinants reflect accumulated system characteristics, such as the stock of capabilities, institutional quality, and network density built over time. Indicators measure flows for example: the annual production of graduates, the current percentage of firms innovating, the year's R&D expenditure. A country's human capital stock (determinant) conditions its annual graduate output (indicator), but they are analytically distinct.

In this conceptual framework indicators enter the construction of IRPI and IEI, and they are the empirical basis for measurement. Determinants do not enter index construction, rather they provide the interpretive vocabulary for explaining why countries achieve different scores (Chapter 7) and for validating the framework through external correlation (Chapter 6). The same empirical phenomenon may therefore appear in both categories but serves different analytical purposes: R&D expenditure is an input indicator for efficiency calculation, while a country's R&D investment culture is a determinant explaining why some systems invest more productively than others.

This separation aligns with Furman et al.'s (2002) distinction between innovative capacity (structural conditions) and innovative output (measurable results), and with Castellacci and Natera's (2011, 2013) distinction between capability stocks and performance flows. Maintaining this separation ensures that the framework does not conflate what it measures (indicators) with what explains variation in those measures (determinants).

4.3 CATEGORISATION OF NIS INDICATORS

The three NIS performance dimensions derive from the sequential model of innovation developed in Chapter 2 and long established in innovation systems research (Kline and Rosenberg, 1986). According to this model, innovation unfolds as a process of resource mobilisation, knowledge creation, and knowledge utilisation, each of which produces distinct, observable results. Performance indicators must therefore correspond directly to these theoretically defined stages.

To operationalise the multidimensional framework, it is necessary to align the indicator logic of this dissertation with established measurement frameworks such as the EIS, GII, and the OECD’s sequential model. In line with the theoretical distinctions established in Chapter 2, indicators are categorised not as additional dimensions, but as distinct roles within the innovation process: Inputs, Outputs, Outcomes, and Impacts. This sequential logic, moving from resource mobilisation to knowledge creation and finally to economic utilisation, forms the conceptual basis for the construction of the IRPI, IEI, and the synthetic efficiency-adjusted performance index (IE_IRPI). Table 12 summarises these categories, their alignment with international statistical standards (Frascati and Oslo Manuals) and composite frameworks such as the EIS, as well as their specific treatment in this dissertation. Within this classification, “Framework Conditions” and “Investments” in the EIS correspond to innovation inputs, “Innovation Activities” largely represent outputs, and “Impacts” (including sales, exports, and employment) correspond to final outcomes. A similar structure is observed in the Global Innovation Index (GII), which distinguishes between innovation input and output subindices.

However, unlike the EIS and GII, which aggregate heterogeneous categories into a single composite score, this dissertation maintains a strict analytical separation between resources, results, and efficiency at the measurement stage. Result-based innovation performance (IRPI) is measured exclusively through outputs and outcomes, with explicit decomposition into knowledge production (IKP) and knowledge commercialisation (IKC), while innovation efficiency (IEI) is assessed through the productivity with which inputs are transformed into these results using a two-stage DEA framework, yielding stage-specific efficiency measures for knowledge production (KPEI) and commercialisation (KCEI). For purposes of ranking, comparison, and robustness analysis, these analytically distinct dimensions are subsequently synthesised through an efficiency-adjusted performance index (IE_IRPI), constructed using a performance-first aggregation in which efficiency enters as a bounded modifier of outcomes. Thus, this synthesis does not replace or obscure the underlying indices: IRPI, IEI, and their stage-level components remain fully observable and are retained for detailed diagnostic analysis of system-specific bottlenecks. This layered structure resolves the methodological limitation of conflating resources with results, ensuring that high input levels are not misinterpreted as high result-based innovation performance while still allowing efficiency to inform the interpretation of observed outcomes.

Table 12: Innovation indicators categories

Category	Definition and theoretical focus	Key indicator examples	Role in dissertation
Inputs	Resources and framework conditions that enable innovation but do not constitute results.	R&D Expenditure (GERD, BERD) Human Capital (Researchers, STEM graduates) Infrastructure (Broadband)	Efficiency enablers: used as Inputs in the Innovation Efficiency Index (IEI). Excluded from the Performance Index (IRPI).

Category	Definition and theoretical focus	Key indicator examples	Role in dissertation
		Institutional incentives	
Outputs	Immediate, measurable results of knowledge and technology production (Stage 1).	Scientific publications & citations Patent applications (PCT) Trademarks & Designs Prototypes	Performance (Stage 1): constitutes the Knowledge Production component of the IRPI. Used as Inputs in IEI (Stage 2).
Outcomes	Commercial and economic utilisation of innovation outputs (Stage 2).	Sales of new-to-market products High-tech & Knowledge-intensive exports Employment in innovative sectors	Performance (Stage 2): constitutes the Commercialisation component of the IRPI.
Impacts	Long-term structural, productivity, and societal effects of sustained innovation.	Productivity growth (TFP) Structural change (high value added) Societal transitions (Green/Digital)	Contextual only: conceptually relevant but empirically excluded from indices due to time lags and causality complexity.
Determinants	Structural and contextual factors explain why performance and efficiency vary.	Governance quality Market openness (FDI, GVC) Institutional coherence	Analytical explanators: excluded from index construction. Used in Chapter 7 to interpret empirical patterns.

Source: Own

Inputs

Inputs, which relate to the capacity dimension, represent the resources and framework conditions that enable innovation activities to take place. They include financial investments (e.g. R&D expenditure), human capital (e.g. researchers, tertiary-educated population), infrastructures (e.g. broadband, laboratories) and institutional or regulatory conditions that shape incentives for innovation. The OECD's Frascati and Oslo manuals conceptualise inputs as expenditures, personnel and structural conditions that precede and enable R&D and innovation activities. In the EIS, many indicators grouped under "Framework conditions" and "Investments" such as human resources, attractive research systems, digitalisation and firm investments function essentially as input or capacity indicators. The EIS measurement framework distinguishes four main blocks: Framework conditions, Investments, Innovation activities, and Impacts and 12 dimensions, summarised through 32 indicators, from which the SII is derived. Empirical work using EIS data often groups indicators into input and output clusters, with human resources, research systems and firm investments treated as innovation inputs (Janger et al., 2017; Onea, 2020). Similarly, the GII separates Innovation input subindex pillars (institutions, human capital and research, infrastructure, market and business sophistication) from output pillars, reflecting the same conceptual distinction.

In this dissertation, input indicators are not used to measure performance. They instead serve as DEA inputs in the IEI, capturing how resources and framework conditions are transformed into Stage 1 outputs and Stage 2 outcomes. This is consistent with Edquist's argument that inputs are instruments whose relevance must be assessed relative to the results they generate, and with recent NIS applications that treat education, government effectiveness and labour-force characteristics as determinants of development outcomes rather than performance items (Edquist, 2011; Secka, 2023).

Outputs

Outputs denote the immediate, measurable results of knowledge and technology production. They correspond to what this dissertation labels Stage 1 results, including scientific publications and their citation impact, patents and patent families, trademarks and designs, prototypes and firm-level innovation activity. These indicators reflect the knowledge generating functions of the NIS. In the EIS/SII, several of the 32 indicators capture this output layer, for example international scientific co-publications, top 10% most-cited publications, PCT patent applications, trademark and design applications, and various intellectual-assets indicators under "Innovation activities" and "Impacts". Studies using the EIS often refer to this group as "innovation output indicators", distinguishing them from inputs (Janger et al., 2017; Onea, 2020). The GII similarly defines knowledge and technology outputs as one of its core performance pillars (patents, scientific publications, high-tech manufacturing and knowledge diffusion), again treating them as results of innovation activities rather than resources.

Outputs are part of result-based performance and form the Stage 1 component of the IRPI. They are not treated as inputs for efficiency analysis but as performance measures in their own right, consistent with the view that innovation systems should be judged on what knowledge they produce and mobilise.

Outcomes

Outcomes capture the commercial and economic utilisation of innovation outputs. They correspond to Stage 2 results in this dissertation and include the share of sales from new or significantly improved products, exports of medium and high technology goods or knowledge intensive services, employment in knowledge intensive activities, and technological upgrading at sectoral level. Outcomes indicate whether and how knowledge is absorbed, commercialised and diffused in the economy.

Within the SII framework, the “Impacts” block is subdivided into employment impacts, sales impacts and, more recently, environmental sustainability, and contains indicators such as employment in knowledge intensive activities, employment in fast growing firms in innovative sectors, exports of medium and high-tech products, exports of knowledge intensive services and sales due to innovation activities. These indicators reflect the economic and labour-market outcomes of innovation and are widely interpreted as capturing the market realisation of innovation results. The EU 2020 “Innovation Output Indicator” designed to monitor how effectively Europe transforms ideas into growth, is built from a similar logic, combining patent-based measures, employment in knowledge-intensive sectors and export performance to approximate the commercial and structural outcomes of innovation systems.

In this dissertation, such indicators belong to the outcome component of the IRPI (Stage 2 results) and capture how far Stage 1 outputs are converted into marketable products, competitive exports and innovation-driven employment. They are treated as performance indicators, not as inputs or determinants.

Impacts

Impacts refer to the longer-term economic, structural and societal effects of sustained innovation activity. They include changes in productivity trajectories, shifts towards higher value-added and knowledge-intensive sectors, progress in digital and green transitions, and broader contributions to societal missions such as sustainability and inclusion. Impacts materialise over longer time horizons and result from complex causal chains that involve not only NIS performance but also macroeconomic, institutional and political factors. In the EIS/SII, the label “Impacts” is used for one of the four main activity blocks, but most indicators in this block still reflect relatively proximate economic and employment outcomes, rather than truly long-run macroeconomic or societal impacts. Other EU initiatives, such as the EU Innovation Output Indicator, have sought to approximate innovation impacts by linking innovation metrics to productivity, growth, sustainability or social outcomes, while acknowledging the methodological challenges involved. The broader innovation and development literature emphasises that persistent differences in innovation outputs and outcomes underpin divergent structural and developmental paths, particularly between advanced and catching up economies (Freeman and Soete, 1997; Fagerberg, 1994; Bartels et al., 2012; Secka, 2023).

In line with Chapter 2, impacts are recognised conceptually, but are excluded from the empirical construction of both IRPI and IEI. The dissertation focuses on more proximate, attributable results (outputs and outcomes), which are more robustly measurable and more tightly linked to specific innovation policies. Impacts are used as a theoretical backdrop for interpreting the importance of observed performance differences, especially when discussing long-run trajectories and the policy relevance of result-based gaps across EU member countries.

The 32 indicators of the SII are all formally labelled “performance indicators”, but they implicitly span the four roles outlined above.

- Framework conditions and Investments primarily capture inputs and capacities (e.g. human resources, R&D expenditure, broadband, firm investment in innovation).
- Innovation activities comprise a mixture of process and output indicators, including SME innovation, innovation collaboration and intellectual assets (patents, trademarks, designs).
- Impacts including employment impacts, sales impacts and environmental sustainability, capture economic and environmental outcomes, not foundational inputs.

The literature on innovation measurement commonly distinguishes between input and output indicators (Janger et al., 2017; Edquist et al., 2018), and Onea (2020) provides evidence of this framing applied to EIS data. However, the choice of which indicators to include and how to combine them is contested, and differences in indicator selection and weighting can lead to divergent country rankings and policy interpretations (Grupp and Schubert, 2010; Alqararah, 2023).

This dissertation builds on that logic but further refines the result side by distinguishing:

- Stage 1 outputs (knowledge and technology outputs) and
- Stage 2 outcomes (commercial and economic results),

while treating inputs as enablers for efficiency analysis and impacts as conceptually relevant but empirically excluded from index construction. In this way, the indicator framework used for the IRPI and IEI remains compatible with the SII/EIS structure and the OECD’s sequential model, while providing a clearer theoretical basis for differentiating what is measured (dimensions) from why it varies across countries (determinants).

Another group of indicators represent the external influences on NIS performance, they include structural, institutional, behavioural and contextual characteristics (Bartels et al., 2012; Malerba and Nelson, 2011). The indicators associated with external determinants must, therefore, capture system features that influence performance, but are not themselves measures of performance. Here, broader institutional quality, governance, regulatory framework, anti-corruption, and institutional coordination shape the coherence, stability and effectiveness of innovation policy (Nelson, 1993; Edquist, 2011). The usual indicators here would be indicators or indices on rule of law, government effectiveness, policy coordination indices, corruption etc. Similarly, market openness shapes opportunities for upgrading and catching up (Archibugi and Coco, 2005). Indicators include: FDI inflows, participation in global value chains etc.

A substantial literature uses DEA or related frontier methods to evaluate the efficiency of NIS. These studies differ in country coverage and model specification, but they converge on a relatively narrow set of input and output indicators that closely align with the conceptual distinctions adopted in this dissertation (Kotsemir, 2013). Edquist and colleagues (Edquist et al., 2018) re-interpret the EIS indicator set from a productivity perspective and construct a Productivity Innovation Index (PII) and an Efficiency Innovation Index (EII) using DEA. They select four indicators as innovation inputs from the EIS 2014-2015:

- Public R&D expenditure (% of GDP)
- Venture capital (% of GDP)
- Business R&D expenditure (% of GDP)
- Non-R&D innovation expenditures (% of turnover)

They classify eight indicators as innovation outputs, drawn mainly from the “Innovators”, “Linkages”, “Intellectual assets”, and “Impacts” dimensions (e.g. share of product/process-innovating SMEs, innovative SMEs collaborating with others, PCT patents per GDP, trademarks, designs, sales of new-to-market and new-to-firm innovations, employment in knowledge-intensive activities). Their approach therefore exemplifies the same logic adopted here:

- Inputs: resource variables intended to stimulate innovation; and
- Outputs: direct innovation results.

Determinants such as institutional quality or demand-side conditions are explicitly recognised as important but left outside the efficiency index due to data and modelling constraints (Edquist et al., 2018). Carayannis et al. (2016) proposes multi-level, multi-stage DEA models in which innovation is represented as a sequence of linked processes. In their national and regional applications, typical inputs and outputs include:

Stage 1 inputs:

- Gross domestic expenditure on R&D (GERD), broken down by sector;
- Number of researchers and R&D personnel;
- Higher-education attainment or tertiary enrolment;
- ICT and physical infrastructure proxies.

Stage 1 outputs / Stage 2 inputs:

- Number of scientific publications;
- Patent applications (e.g. PCT, EPO);

- Technology licences;
- Intermediate innovation indicators from innovation surveys.

Stage 2 outputs:

- High-technology exports and knowledge-intensive services exports;
- Sales of innovative products;
- Productivity or value-added in high-tech sectors.

These models explicitly operationalise the two-stage transformation (inputs into knowledge outputs into commercial outcomes) that underpins the IEI in this dissertation, and they provide empirical support for distinguishing research productivity from commercialisation productivity (Guan and Chen, 2012; Jankowska et al., 2017).

Systematic reviews of NIS efficiency studies (Kotsemir, 2013; Narayanan et al., 2022) show that most cross-country DEA models rely on a relatively standard menu of indicators:

Common inputs:

- GERD or R&D expenditure by sector (business, government, higher education);
- Number of researchers and R&D personnel;
- R&D intensity (R&D/GDP);
- Sometimes education variables (e.g. tertiary students, researchers per million population).

Common outputs:

- Patent counts (PCT, triadic, EPO) and patent families;
- Scientific publications (articles in international journals);
- High technology exports and medium and high-tech manufacturing exports;
- Share of high technology or knowledge intensive sectors in GDP or employment;
- Innovation survey outcomes (e.g. share of innovative firms).

Guan and Chen (2012) explicitly develop a two-stage DEA model in which R&D expenditures and personnel generate patent and publication outputs, which in turn lead to high technology exports and industrial upgrading. These studies collectively reinforce three methodological choices adopted in this dissertation:

- inputs are limited to resource type indicators (R&D, personnel, infrastructure) and are not conflated with performance indicators;

- outputs and outcomes are used to define performance, both in single stage and two stages efficiency models; and
- determinants such as GDP, institutional quality, and openness of the economy are typically treated as explanatory variables outside the DEA model (e.g. in regression or cluster analyses), not as inputs/outputs, which is consistent with this dissertation's strict separation between dimensions (IRPI, IEI indicators) and determinants (used in Chapter 7 for interpretation).

By situating the IRPI and IEI indicator choices within this wider empirical tradition, the chapter demonstrates that the input-output-outcome-impact classification and the three dimension framework (result-based performance, two-stage efficiency, latent capacity) are theoretically grounded, but also consistent with the most influential measurement and efficiency studies on NIS.

4.4 THEORETICAL EXPECTATIONS AND HYPOTHESES

The systemic and evolutionary literature has long emphasised that innovation does not result from isolated inputs or individual actors, but from complex interactions among institutions, firms, knowledge infrastructures and market environments (Freeman, 1987; Lundvall, 1992; Nelson, 1993). Since the foundational contributions of Schumpeter (1934), innovation has been understood as a cumulative, interactive and non-linear process, meaning that measuring NIS performance requires more than counting inputs or single outputs. A central finding in comparative innovation research is that countries with similar levels of investment often generate very different innovation results (Fagerberg, 1994, 2000; Fagerberg and Verspagen, 2002; Bartels et al., 2012; Secka, 2023). These divergences occur because innovation results reflect not only resource endowments, but also how effectively systems transform inputs into outputs, and how outputs are converted into commercial and economic outcomes (Edquist, 2011). Studies show that countries may produce strong scientific outputs yet fail to commercialise them, often due to weak diffusion capacities, missing complementary assets, or insufficient user-producer interactions (Lundvall, 1992; Teece, 1986; Malerba and Nelson, 2011). Multi-stage efficiency analyses further demonstrate that decomposing innovation into sequential stages reveals bottlenecks that remain hidden in aggregated indicators. Multi-stage and network DEA studies consistently show that separating knowledge production from commercialisation uncovers structural inefficiencies within NIS (Guan and Chen, 2012; Liou, 2009). As these studies argue, a system may appear strong in aggregate rankings even while suffering from persistent Stage 2 commercialisation gaps or inefficient use of R&D resources.

Given this accumulated evidence, a unified framework that combines result-based performance, two-stage innovation efficiency, and innovation capacity offers a more realistic diagnostic perspective on NIS functioning. Each dimension corresponds to a distinct theoretical component of the innovation process and collectively reflects the multidimensional nature of innovation described in Chapter 2. This section develops the theoretical expectations that follow from this framework and formulates the hypotheses that guide empirical testing.

4.4.1 Expected performance-efficiency configurations

The two-dimensional performance-efficiency framework implies that four distinct system configurations are theoretically possible:

- High performance and high efficiency: these systems achieve strong innovation outcomes while operating close to the transformation frontier. They represent the benchmark against which other systems are implicitly measured. Theoretically, such systems combine strong institutional coherence, well-developed knowledge networks, and effective commercialisation mechanisms. They are expected to exhibit balanced Stage 1 and Stage 2 efficiency.
- High performance and low efficiency: these systems achieve strong innovation outcomes despite operating below the efficiency frontier. They succeed through resource abundance rather than transformation productivity, by deploying sufficient inputs to generate results even with substantial "waste". Such configurations may characterise large, wealthy economies with extensive R&D infrastructure but coordination failures or institutional rigidities that prevent efficient resource utilisation. Policy implications differ markedly from configuration 1 where the challenge is not expanding outcomes but improving transformation productivity.
- Low performance and high efficiency: these systems operate close to the efficiency frontier but achieve limited outcomes. They transform available resources productively but lack sufficient resources or structural capacity to generate strong results. Such configurations may characterise smaller or less wealthy economies that have developed effective transformation mechanisms but remain constrained by scale or input availability. Policy implications focus on relaxing constraints such as expanding resources, building absorptive capacity, or enhancing access to external knowledge, rather than improving efficiency.
- Low performance and low efficiency: these systems achieve weak outcomes while operating below the efficiency frontier. They face compound challenges: both insufficient results and unproductive transformation. Such configurations may characterise systems with weak institutional coherence, fragmented knowledge networks, or misaligned incentive structures. Policy interventions must address both dimensions simultaneously.

Beyond these four quadrants, the two-stage structure of efficiency reveals additional asymmetries. A system may exhibit high Stage 1 efficiency (knowledge production) but low Stage 2 efficiency (commercialisation), thus indicating that scientific and technological outputs are generated productively but commercial translation is blocked. Alternatively, high Stage 2 with low Stage 1 efficiency suggests that whatever knowledge is produced is commercialised effectively, but knowledge production itself is unproductive. These stage-specific patterns, invisible in aggregate measures, enable diagnosis of where in the innovation process bottlenecks occur.

Certain country types are expected to exhibit characteristic configurations. Small, open economies integrated into global value chains may achieve high commercialisation performance (strong Stage 2 outcomes) through foreign direct investment and multinational enterprise activity, even if domestic knowledge production (Stage 1) is limited. Such economies, potentially including FDI-intensive systems, would appear as high performers driven by outcome (IKC) dominance rather than balanced output-outcome profiles. Equally, large economies with extensive university and public research systems may excel at knowledge production but struggle with commercialisation, appearing as strong Stage 1 but weak Stage 2 performers. These expected patterns provide testable predictions for the empirical analysis in Chapter 6. If the framework successfully captures system heterogeneity, the observed configurations should align with theoretical expectations based on country characteristics.

4.4.2 Expected relationship between performance and efficiency

The analytical distinction between performance and efficiency raises an empirical question: how strongly should these dimensions correlate across innovation systems? Three scenarios are theoretically plausible:

1. Strong positive correlation: If high efficiency necessarily produces high performance, and low efficiency necessarily constrains performance, the dimensions would be strongly correlated. This scenario implies that efficiency is the binding constraint, and systems cannot achieve strong outcomes without efficient transformation.
2. Strong negative correlation: If high-performing systems achieve results through resource abundance rather than transformation productivity, efficiency and performance might be inversely related. This scenario implies diminishing returns where the most successful systems succeed despite, not because of, efficiency.
3. Weak or zero correlation: If performance and efficiency capture genuinely independent aspects of system functioning, the dimensions would be uncorrelated. This scenario implies that multiple pathways to innovation success exist: some systems achieve high performance through high efficiency,

others through resource abundance, and the two properties vary independently across the population of systems.

The theoretical framework developed in this chapter is most consistent with Scenario 3. The evolutionary and systemic traditions emphasise that innovation outcomes emerge from complex interactions among actors, institutions, and learning processes, and not from a single production function that mechanically links inputs to outputs. As Castellacci and Natera (2011) demonstrate, systems may be input-rich but inefficient, or input-poor yet efficient, depending on their structural and functional characteristics. This heterogeneity implies that efficiency and performance capture different properties of system functioning that need not co-vary.

If empirical analysis reveals weak correlation between performance and efficiency indices, this will confirm that the dimensions are not merely analytically distinct but empirically independent, thus validating the framework's core premise that both dimensions must be assessed to understand NIS functioning. This constitutes an explicit falsification test: strong correlation ($r>0.50$) would indicate that one dimension largely subsumes the other, undermining the rationale for separate measurement and invalidating the dual-index architecture. Chapter 6 examines this relationship empirically and reports whether the falsification criterion is met.

4.4.3 Peer-group logic and policy relevance

Innovation policy learning depends on meaningful comparisons among structurally similar systems, a point emphasised throughout the NIS literature (Lundvall, 1992; Nelson, 1993; Malerba and Nelson, 2011). Countries rarely benefit from benchmarking against the highest performing systems, because top performers often possess institutional, historical or structural conditions that cannot be replicated elsewhere (Fagerberg and Srholec, 2008). Instead, policy learning is most effective when countries identify peer systems with comparable capabilities, institutional settings and innovation pathways. Empirical analyses strongly support this argument. Cross-country studies show that composite indices, whether the SII, GII or GCR, often mask heterogeneity among countries with similar aggregate scores (Bartels et al., 2012; Barrichello et al., 2020). NIS that appears similar in overall rank frequently diverge sharply in their underlying knowledge production patterns, efficiency levels or long run capacity. Recent multidimensional NIS frameworks emphasise that grouping countries according to multiple innovation dimensions yield more realistic and policy relevant clusters (Secka, 2023). Research using efficiency classifications also shows that countries can share similar performance levels while differing fundamentally in efficiency and systemic functioning (Guan and Chen, 2012). This reinforces the idea that peer identification requires more than output comparison: it must account for how systems achieve those outputs and at what resource cost.

Because the proposed IRPI and IEI framework jointly captures what innovation systems achieve (outputs and outcomes) and how effectively they achieve it (two-stage efficiency), it is theoretically expected to produce peer groupings that reflect deeper structural similarities. This multidimensional perspective allows policymakers to identify countries that share comparable system dynamics, whether they are high performing but inefficient, efficient but resource constrained, or balanced across both dimensions.

Three mechanisms link structural similarity in the performance-efficiency space to policy transferability. *First*, similar configurations imply similar policy challenges. A country with high performance but low efficiency faces a fundamentally different policy problem than one with low performance but high efficiency. The former must address transformation bottlenecks: why are abundant resources not being converted productively? The latter must address outcome constraints: why does efficient transformation yield limited results? Policy instruments appropriate to one challenge (e.g., improving university-industry linkages) may be irrelevant to the other (e.g., expanding R&D investment). Countries occupying similar positions in the performance-efficiency space are therefore more likely to face analogous challenges requiring analogous interventions. *Second*, structural similarity increases the probability of policy transfer success. As Dolowitz and Marsh (2000) establish in the policy transfer literature, successful transfer depends on contextual compatibility. Countries with similar performance-efficiency profiles are more likely to share underlying structural features, such as comparable institutional configurations, similar sectoral compositions, analogous governance arrangements, that condition policy effectiveness. Benchmarking against structurally dissimilar systems risks emulating policies that succeeded under conditions that cannot be replicated. *Third*, the performance-efficiency space captures policy-relevant variation. Unlike purely outcome-based rankings, which conflate resource endowments with system functioning, the joint performance-efficiency configuration reveals how systems operate. Two countries with identical IRPI scores may differ dramatically in IEI, one achieving results through massive resource deployment, the other through superior transformation. These systems require different policy responses, a distinction invisible in single-dimension rankings but explicit in the two-dimensional framework.

The peer-group logic thus rests not on arbitrary similarity but on the premise that countries facing similar transformation challenges, operating with similar efficiency profiles, and producing similar outcome patterns constitute a policy-relevant comparison set. This premise is testable: if the framework successfully identifies peers, countries grouped together should exhibit greater similarity in policy-relevant characteristics (institutional quality, sectoral structure, governance capacity) than countries grouped by single-dimension rankings alone. Chapter 7 examines whether this expectation is empirically supported.

4.4.4 Hypotheses

The theoretical expectations developed in Sections 4.4.1-4.4.3 generate two testable hypotheses. Chapter 1 presented these hypotheses with their full support criteria; this section clarifies their derivation from the conceptual framework.

Hypothesis 1 predicts that a multidimensional framework integrating result-based innovation performance and innovation efficiency, interpreted through an innovation capacity perspective, provides more comprehensive and diagnostically informative assessment than existing frameworks.

This hypothesis derives from three theoretical expectations:

- From Section 4.4.1 (Configurations): The four-quadrant typology predicts that countries will distribute across performance-efficiency configurations in theoretically meaningful patterns, not cluster in a single quadrant.
- From Section 4.4.2 (Independence): The conceptual separation between "what systems achieve" and "how productively they achieve it" predicts weak empirical correlation between IRPI and IEI.
- From Section 4.1.2 (Two-stage structure): The sequential model of innovation predicts systematic differences between knowledge production efficiency (KPEI) and commercialisation efficiency (KCEI).

The specific support criteria and falsification thresholds are established in Chapter 1 (Section 1.3) and operationalised in Chapter 5 (Section 5.1.4).

Hypothesis 2 predicts that the innovation indices based on the multidimensional framework enable more accurate identification of peer countries for benchmarking and policy learning. This hypothesis derives from:

- *From Section 4.4.3 (Peer-group logic):* Countries in the same performance-efficiency quadrant should share structural characteristics, enabling policy learning based on system similarity rather than rank proximity.
- *From Section 4.2.3 (Determinants):* IE_IRPI should correlate with external determinants of innovation capacity, validating its construct validity.

The specific support criteria and falsification thresholds are established in Chapter 1 and operationalised in Chapter 5.

Chapter 5 specifies the methodological procedures for testing these hypotheses; Chapter 6 presents the empirical results.

4.5 METHODOLOGICAL REQUIREMENTS DERIVED FROM THE CONCEPTUAL FRAMEWORK

The conceptual framework developed in this chapter carries specific implications for methodological design. The research design responds to the theoretical structure of the framework. This section makes explicit the methodological requirements that follow from the conceptual distinctions established above.

Requirement 1: Strict separation between performance and efficiency measurement

The conceptual framework distinguishes result-based innovation performance (what NIS achieve) from innovation efficiency (how productively they transform resources). This distinction is not merely taxonomic, it reflects fundamentally different analytical questions. Performance assessment asks: what outputs and outcomes does the system produce? Efficiency assessment asks: given the inputs deployed, how close does the system operate to best practice? These questions require different measurement approaches. Performance can be assessed through composite indexing of observable outputs and outcomes, aggregated according to their contribution to innovation results. Efficiency requires a frontier-based method that benchmarks each system against the best-observed practice in the sample. Conflating these approaches, for example, by including inputs alongside outputs in a single composite index, would violate the conceptual separation and reproduce the diagnostic limitations of existing frameworks such as the EIS. The methodological implication is that two distinct indices must be constructed: one capturing result-based performance, and one capturing transformation efficiency.

Requirement 2: Two-stage structure reflecting the sequential nature of innovation

The theoretical framework conceptualises innovation as a sequential process encompassing knowledge production (Stage 1) and commercialisation (Stage 2). This structure implies that efficiency cannot be meaningfully assessed in a single step. A system may be efficient at generating knowledge outputs (patents, publications) yet inefficient at converting these outputs into economic outcomes (exports, employment in innovative enterprises), or vice versa. Single-stage efficiency models would obscure such asymmetries. The methodological implication is that efficiency analysis must adopt a two-stage structure, assessing separately the productivity of knowledge production and the productivity of commercialisation. This requires a network or two-stage DEA specification in which outputs from Stage 1 serve as inputs to Stage 2, preserving the sequential logic of the innovation process.

Requirement 3: Data-driven, non-redundant indicator selection

Section 4.3 mapped the theoretical dimensions to observable indicators, revealing that many available indicators capture overlapping phenomena. The methodological implication is that indicator selection cannot rely solely on expert judgement or theoretical mapping, it must be complemented by a data-driven procedure that identifies indicators with high explanatory relevance and low redundancy. Feature selection methods, specifically approaches that assess each indicator's contribution to cross-country differentiation while penalising redundancy, are therefore required.

This requirement raises an apparent epistemological tension: the conceptual framework is theory-driven, yet indicator selection is to be data-driven. This tension is more apparent than real. The conceptual framework specifies what should be measured (Stage 1 outputs, Stage 2 outcomes, transformation efficiency) and why these dimensions matter, it does not specify which particular indicators best capture these dimensions within available data, a question that is empirical, not theoretical. Data-driven selection therefore operates under theoretical constraint, not independently of it: indicators are pre-classified into theoretically defined categories and selection occurs within each category. Section 5.1.2 develops the epistemological justification for this approach in full detail, demonstrating that data-driven indicator selection, when properly constrained by theoretical structure, strengthens rather than undermines the framework's conceptual foundations.

Requirement 4: Treatment of capacity as interpretive, not measured

Innovation capacity is defined in this framework as a latent systemic property that conditions long-run performance and efficiency but is not directly observable. Attempting to measure capacity through a third composite index would require selecting observable proxies (such as institutional quality, R&D infrastructure, or educational attainment), which risks conflating enabling conditions with inputs or outcomes. The methodological implication is that capacity must not be operationalised as a standalone empirical index. Instead, capacity is inferred interpretively from persistent patterns in performance-efficiency configurations and from the external determinants analysed in Section 4.2. This preserves analytical clarity and avoids the contamination problems that afflict existing composite indicators.

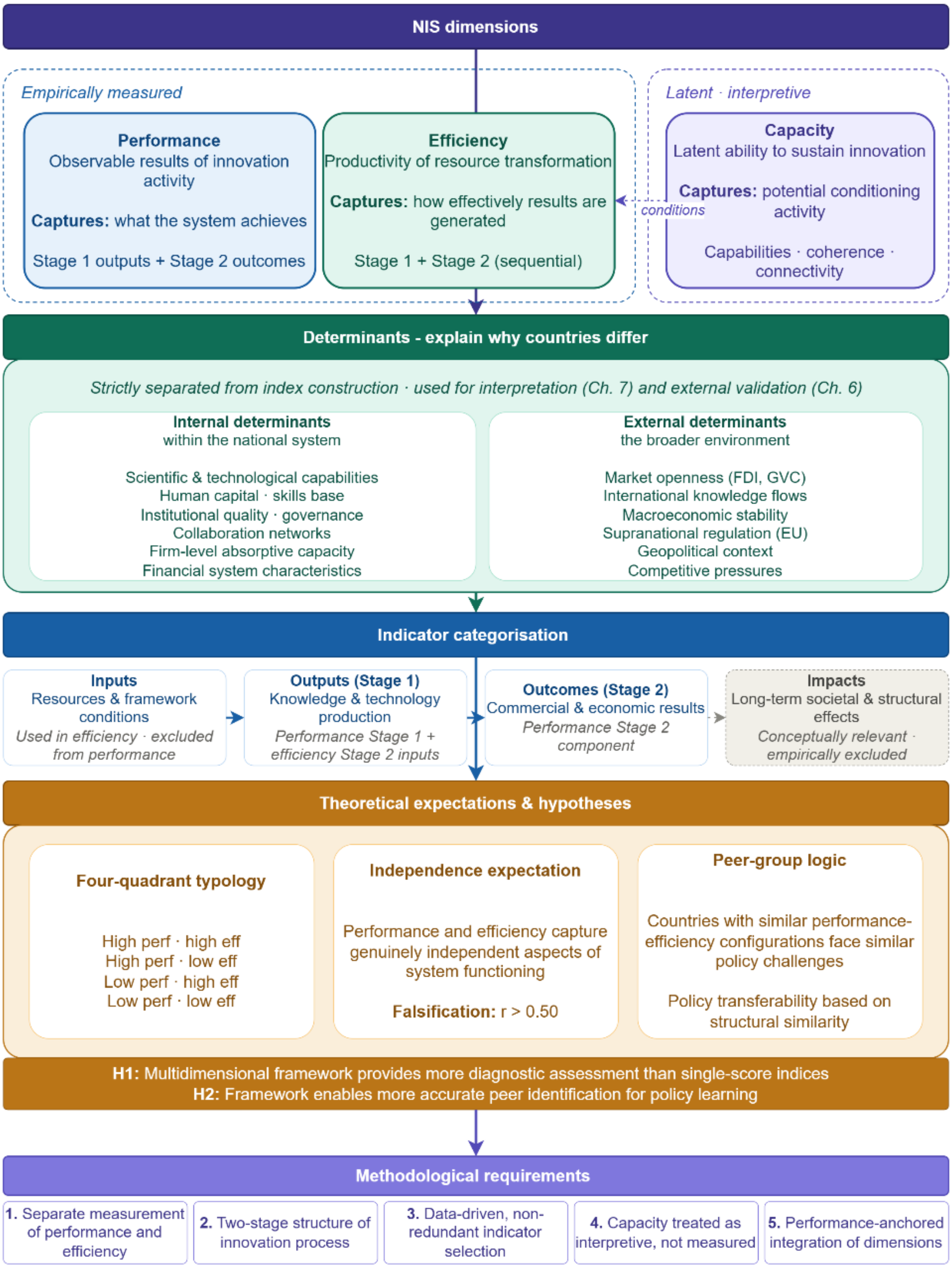
Requirement 5: Performance-anchored integration

The conceptual framework establishes an asymmetry between performance and efficiency: performance represents what innovation systems ultimately exist to achieve, while efficiency represents the productivity of the process through which achievements are generated. This asymmetry has implications for how the two dimensions should be integrated. If efficiency were weighted equally with performance or allowed to dominate, the resulting measure could reward systems that are efficient but achieve little, undermining the outcome-oriented logic of NIS performance assessment. The methodological implication is that any integrated index must preserve performance as the primary ordering criterion, with efficiency entering as a bounded adjustment that refines but does not overturn performance-based rankings. This performance-first aggregation rule operationalises the conceptual primacy of results over processes.

This chapter has developed the conceptual framework that addresses the measurement gap documented in Chapters 2 and 3. It formally defined three analytically distinct dimensions: result-based innovation performance (observable results), innovation efficiency (transformation processes), and innovation capacity (latent system conditions), and specified their theoretical relationships. The framework established performance primacy (results are what innovation systems exist to achieve), the two-stage sequential structure (knowledge production followed by commercialisation), and the interpretive treatment of capacity (inferred from persistent configurations rather than directly measured).

From these conceptual distinctions, five methodological requirements follow necessarily: (1) separate measurement of performance and efficiency, (2) two-stage structure reflecting sequential innovation processes, (3) data-driven non-redundant indicator selection, (4) interpretive treatment of capacity, and (5) performance-anchored integration. These requirements are not arbitrary preferences but direct implications of the theoretical framework. Figure 9 below presents the summary of the conceptual architecture.

Figure 9: Conceptual architecture of NIS performance measurement



Source: Own

Chapter 5 operationalises each requirement through specific measurement instruments. It constructs the Innovation Result-based Performance Index to capture observable outputs and outcomes, the Innovation Efficiency Index using two-stage DEA to assess transformation productivity, and the Efficiency-Adjusted Result-based Performance Index to integrate these dimensions while preserving performance as the primary ordering criterion. Each instrument is explicitly linked to the conceptual distinctions established in this chapter and to the research questions formulated in Chapter 1.

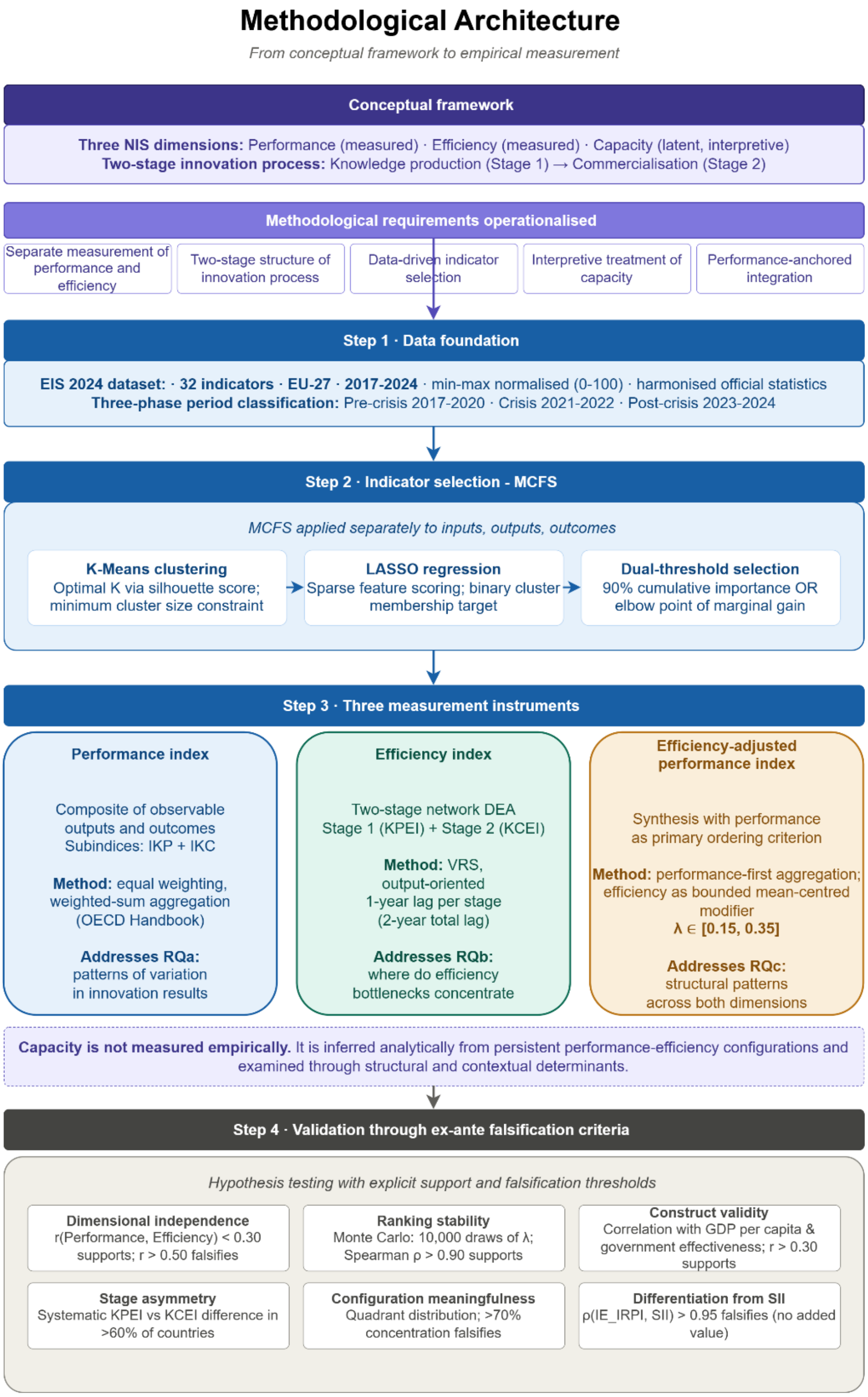
5 RESEARCH METHODOLOGY

Chapter 4 developed the conceptual framework distinguishing result-based innovation performance, efficiency, and capacity, and specified five methodological requirements: (1) separate measurement of performance and efficiency, (2) two-stage structure reflecting the sequential nature of innovation, (3) data-driven indicator selection, (4) interpretive treatment of capacity, and (5) performance-anchored integration. This chapter operationalises these requirements through specific measurement instruments.

The chapter constructs three indices. The Innovation Result-based Performance Index captures observable outputs and outcomes, addressing RQa (what patterns of variation exist in innovation results). The Innovation Efficiency Index derived from two-stage DEA, assesses transformation productivity at the knowledge production and commercialisation stages, addressing RQb (where do efficiency bottlenecks concentrate). The Efficiency-Adjusted Result-based Performance Index integrates these dimensions while preserving performance as the primary ordering criterion, addressing RQc (what structural characteristics emerge when both dimensions are considered together).

The chapter proceeds as follows: Section 5.1 establishes the measurement scope and epistemological foundations. Section 5.2 describes the data sources and treatment procedures. Section 5.3 details the MCFS procedure for indicator selection. Section 5.4 operationalises the IRPI construction. Section 5.5 specifies the two-stage DEA model for IEI. Section 5.6 presents the IE_IRPI aggregation formula. Section 5.7 establishes the validation and robustness analysis protocols applied in Chapter 6. Figure 10 presents the methodological architecture in detail.

Figure 10: Methodological architecture



Source: Own

5.1 OPERATIONALISING THE CONCEPTUAL FRAMEWORK

The methodological design of this thesis is grounded in the conceptualisation of the NIS as a sequential, multi-stage process through which resources are transformed into technological knowledge and, subsequently, into economic value. Schumpeter (1934) introduced the notion of stages, later elaborated by other contemporary authors argued that innovation involves separate but interlinked stages, each characterised by different resource requirements, institutional actors, and economic mechanisms (Lundvall, 1992; Nelson, 1993). Empirical studies widely confirm that innovation should be modelled as a two-stage system, with distinct production technologies and performance determinants (Cai and Hanley, 2012; Guan and Chen, 2012) where Stage 1 transforms inputs into outputs, while Stage 2 transforms outputs into outcomes. Outputs function as both results of Stage 1 and the necessary inputs to Stage 2. This two-stage conceptualisation forms the foundation for both indices constructed in this thesis.

5.1.1 Measurement scope

This dissertation adopts a deliberately restrictive approach to empirical measurement to ensure conceptual clarity and methodological rigor in the assessment of NIS. A central methodological decision concerns the distinction between phenomena that can be directly observed and measured and system characteristics that are theoretically relevant but not directly observable and are therefore unsuitable for direct operationalisation within a composite measurement framework. The analytical framework is broader than the measurement framework: it specifies how NIS are understood conceptually, while the measurement framework is limited to elements that are empirically observable. However, only result-based innovation performance and innovation efficiency are operationalised as empirical indices. Result-based innovation performance is defined strictly in terms of realised results and is captured through observable knowledge and technological outputs and commercial and economic outcomes. Innovation efficiency is defined as productivity with which innovation-related resources are transformed into these results across two sequential stages of the innovation process. Both performance and efficiency satisfy the requirements of empirical observability, reproducibility, and analytical separability, and can therefore be operationalised using available indicators and established quantitative methods.

Consistent with the conceptual framework developed in Chapter 4, Innovation capacity is understood as a latent system-level property rather than a directly measurable dimension and it is not operationalised empirically in this chapter. Capacity-related variables are excluded to avoid conflating structural conditions with performance outputs or efficiency relationships. This preserves analytical separability and prevents circularity in DEA-based estimation. In the NIS literature, capacity refers to deep structural and institutional characteristics, such as learning processes, absorptive ability, governance quality, and systemic coherence, that condition long-run innovation trajectories. These characteristics evolve slowly, operate through complex interactions, and manifest indirectly through persistent patterns of performance and efficiency over time. As such, innovation capacity does not represent an observable outcome or a transformation process at a given point in time. From the perspective of measurement theory, treating innovation capacity as a measured dimension would require strong and contestable assumptions regarding indicator selection, aggregation rules, and temporal alignment. It would also risk conflating enabling conditions and explanatory factors with outcomes, thereby introducing conceptual contamination into the measurement framework. This issue is well documented in innovation measurement practice, where composite indices frequently aggregate inputs, framework conditions, and results into single scores, obscuring causal relationships and reducing diagnostic usefulness. To avoid this pitfall, the present framework restricts empirical measurement to elements that directly reflect either what innovation systems achieve (performance) or how effectively they achieve it (efficiency). Innovation capacity is retained as a theoretical and interpretive element that informs the analysis rather than as a third empirical index. Capacity is inferred analytically from persistent performance and efficiency patterns and is examined indirectly through the analysis of structural and contextual determinants in later chapters, rather than being quantified through a standalone composite indicator.

This distinction between empirically measured dimensions and latent system properties strengthens the internal coherence of the framework. It ensures that the indices developed in this dissertation, IRPI and IEI, capture only observable phenomena that are fully aligned with their measurement objectives. At the same time, it preserves the explanatory role of innovation capacity within the systemic and evolutionary tradition of innovation studies, where underlying capabilities are understood as shaping outcomes over time rather than being directly observable themselves. By clearly separating empirical measurement from theoretical interpretation, the framework enhances conceptual precision, avoids methodological overreach, and provides a transparent and defensible basis for empirical analysis and policy interpretation. The capacity lens is applied interpretively in Chapter 6 and Chapter 7, where persistent performance-efficiency configurations are examined through the three components of capacity defined in Chapter 4: accumulated capabilities, institutional coherence, and structural connectivity. Countries exhibiting stable high-performance, high-efficiency positions across the analysis period provide evidence consistent with high innovation capacity, while volatile or declining configurations suggest capacity constraints. This methodological choice underpins the revised hypotheses and ensures that the empirical contribution of the dissertation rests on analytically precise and scientifically robust foundations.

5.1.2 Epistemological foundations of indicator selection

The conceptual framework developed in Chapter 4 derives its structure from NIS theory and evolutionary economics, establishing theoretically grounded dimensions of innovation performance, efficiency, and capacity. Yet the methodological approach adopted in this dissertation employs data-driven indicator selection through MCFS. This combination raises an apparent epistemological tension that requires explicit resolution: how can a theory-driven conceptual framework be reconciled with data-driven indicator selection without compromising analytical coherence? This tension is more apparent than real, and understanding its resolution is essential to appreciating why MCFS is not merely a convenient statistical technique but an epistemologically appropriate method for operationalising the conceptual framework. The reconciliation rests on a clear division of labour between theoretical specification and empirical identification. The conceptual framework performs three functions that are inherently theoretical:

- Dimensional specification: The framework specifies what should be measured: Stage 1 knowledge and technology outputs, Stage 2 commercial and economic outcomes, and the efficiency of transformation between stages. These dimensions derive from NIS theory's conceptualisation of innovation as a sequential, cumulative process (Lundvall, 1992; Nelson, 1993; Kline and Rosenberg, 1986).
- Categorical classification: The framework prescribes how available indicators should be categorised as inputs, outputs, or outcomes, based on their theoretical role in the innovation process. Table 12 in Chapter 4 provides this classification, grounded in the sequential logic of knowledge production followed by commercialisation.
- Relational structure: The framework establishes why these dimensions matter and how they relate: performance captures what systems achieve, efficiency captures how productively they transform resources, and capacity conditions both as a latent system property. The conceptual primacy of performance over efficiency, justified in Section 4.1, reflects theoretical reasoning about the purpose of innovation systems.

What the conceptual framework does not specify and cannot specify on purely theoretical grounds is which indicators within each category best capture the theoretically defined dimensions given the available data. This question is empirical rather than theoretical. Consider the output category: theory establishes that Stage 1 outputs include scientific publications, patents, trademarks, and designs. But whether PCT patent applications or trademark applications better differentiate EU innovation systems, and whether one is more informative than the other for diagnostic purposes, cannot be determined by theoretical reasoning alone. It requires empirical investigation of how these indicators vary across countries and whether they capture distinct or redundant information. This is where data-driven selection enters, not as a replacement for theory but as its necessary empirical complement. MCFS addresses the question that theory cannot: given the theoretically specified categories, which subset of indicators most effectively captures meaningful cross-country variation with minimal redundancy?

MCFS does not operate in a theoretical vacuum. The data-driven selection is bounded by theoretical constraints at multiple levels:

- MCFS selects indicators within theoretically defined categories, not across them. The algorithm is applied separately to input indicators, output indicators, and outcome indicators as classified in Table 12. An output indicator cannot be reclassified as an outcome indicator based on its statistical properties, categorical membership is theoretically fixed. This ensures that the sequential logic of the innovation process where inputs are enabling outputs, outputs are enabling outcomes, is preserved regardless of which specific indicators are selected.
- The selection process maintains the two-stage structure of the framework. Indicators selected for Stage 1 (knowledge production) remain analytically separate from those selected for Stage 2 (commercialisation). MCFS cannot collapse these stages or blur their boundaries. The theoretical architecture of sequential transformation is thus maintained throughout the empirical operationalisation.
- Unlike dimensionality reduction techniques such as Principal Component Analysis (PCA), which create latent factors that are linear combinations of original variables, MCFS selects original indicators without transformation. As Chapter 2 documented, composite indices often suffer from interpretability problems when aggregation obscures what is being measured. By preserving original indicators, MCFS ensures that each selected variable retains its theoretical meaning. For example, "PCT patent applications per billion GDP" remains interpretable as a measure of knowledge output, it is not transformed into an abstract "Factor 1" that conflates multiple phenomena. This preservation of interpretability is epistemologically crucial: it allows the empirical results to be read back through the theoretical framework without loss of meaning.
- Only indicators that satisfy theoretical relevance criteria enter the selection pool. The EIS indicator set from which selection occurs, has been pre-screened for conceptual alignment with innovation measurement (as documented in Section 4.3). MCFS selects among theoretically relevant indicators, and it does not determine theoretical relevance itself.
- MCFS selects indicators based on their discriminatory power, their ability to differentiate between country clusters. This selection criterion aligns directly with the diagnostic purpose of the conceptual framework.

As Chapter 1 established, the purpose of the framework is not ranking per se but diagnostic capacity: "identifying where system functioning breaks down and which interventions are likely to improve innovation results." Diagnosis requires identifying meaningful differences across innovation systems. Indicators that effectively discriminate between country clusters are precisely those that capture such meaningful differences. An indicator on which all countries score similarly, regardless of its theoretical importance, provides no diagnostic information. It cannot reveal structural heterogeneity or identify system-specific bottlenecks.

A legitimate concern with data-driven selection is that it might reduce the framework to pure empiricism, letting data determine not just which indicators to include but what NIS performance means. This concern must be addressed directly. The safeguard against empiricist reduction lies in the strict separation of functions between theory and data. Theory determines meaning and data determines measurement. The conceptual framework defines what NIS performance is: the observable results of knowledge production and commercialisation. This definition is not subject to empirical revision. What MCFS determines is which available indicators best measure this theoretically defined construct. If MCFS selected indicators that did not align with the theoretical definition, for example, if it selected only input indicators to measure performance, this would indicate a failure of the selection process, not a revision of the concept.

In practice, the selected indicators with MCFS align with theoretical expectations. Output indicators capture knowledge and technology production, outcome indicators capture commercial and economic results. This alignment provides empirical validation that the data-driven selection has operated appropriately within its theoretical constraints. Furthermore, the selection is transparent and replicable. The MCFS algorithm specifies exactly how indicators are evaluated and selected. Unlike expert judgement, which may reflect unstated assumptions or disciplinary biases, MCFS provides an auditable procedure. Researchers can examine whether the selection criteria are appropriate, whether the constraints are correctly specified, and whether alternative specifications would yield different results. This transparency strengthens rather than weakens the epistemological foundations of the framework.

Overall, the combination of theory-driven framework and data-driven indicator selection is not a compromise but a principled methodological design. Each component performs the function it is suited to perform. Theory specifies dimensions, categories, relations, and meaning, that are questions requiring conceptual reasoning grounded in NIS literature and evolutionary economics. Data identifies which specific indicators, among theoretically relevant candidates, most effectively capture cross-country variation with minimal redundancy: this is a question that requires empirical investigation. The division is clean: theory sets the boundaries within which data-driven selection operates and data informs which indicators best serve the theoretically defined purposes. Neither component encroaches on the other's domain. The result is a measurement framework that is simultaneously theoretically grounded and empirically robust, avoiding both the arbitrariness of purely expert-driven selection and the conceptual emptiness of purely statistical approaches. This epistemological foundation underpins the specific MCFS procedure detailed in Section 5.3.

5.1.3 Method selection

The empirical analysis applies methods suited to sequential innovation processes. As innovation systems exist at the national level, the underlying data must be internationally comparable and derived from harmonised official statistical sources. Therefore, the empirical work in this dissertation relies exclusively on secondary data drawn from EIS. The choice to draw indicators from the EIS dataset reflects considerations of data quality, temporal consistency, and cross-country comparability, rather than endorsement of the EIS composite index methodology. The EIS provides one of the few harmonised, longitudinal, and policy-validated indicator sets covering all EU member countries. Using this indicator universe ensures comparability with existing policy benchmarks while enabling the present study to implement a conceptually distinct aggregation strategy.

The methodological challenge is not merely selecting data sources but determining which features (indicators) meaningfully capture the dimensions defined in the conceptual framework.

To ensure the performance index captures distinct dimensions rather than redundant measures, indicator selection employs MCFS. Indicator selection prioritises relevance and non-redundancy. A central methodological task in this dissertation is the selection of indicators for constructing the IRPI and IEI with MCFS. The indicator selection process followed a structured, multi-step approach:

- Mapping observable indicators per each dimension. Based on the conceptual analysis of NIS determinants (Chapter 4), each dimension was translated into measurable, internationally recognised indicators. This mapping relied on established EIS framework. This ensured that indicator selection was grounded in widely accepted measurement practice that could allow for testing of the robustness of the composite indicator and the comparability to other indices such as EIS.
- Feature screening based on relevance and non-redundancy. Indicators were assessed for: conceptual relevance to Stage 1 or Stage 2 and non-duplication (e.g., avoiding two indicators that measure similar constructs). This reduced redundancy in the indicator set.
- Practical screening based on data quality. Indicators were retained only if data met the following criteria: all years coverage for 2017-2024, coverage across all observed NIS (EU member countries), harmonised definitions from official

sources. Indicators failing these criteria were excluded to ensure methodological robustness.

- Construct-level coverage. The final indicator set ensures balanced representation of inputs, outputs, and outcomes. This enables construction of a result-based innovation performance index (IRPI) that captures innovation results, an efficiency index (IEI) that captures productive ability without conceptual overlap and the synthetic index combining the two (IE_IRPI) which is the result-based performance index (IRPI) adjusted for the efficiency dimension (IEI).

Three complementary methodological tools were selected in accordance with the multidimensional framework:

- to test whether high performance implies high efficiency, two-stage DEA is applied. DEA is chosen to model efficiency because it evaluates productivity in transforming inputs into outputs without imposing parametric assumptions. This approach is widely used in innovation efficiency research and enables identification of distinct weaknesses in each stage.
- to examine whether efficiency-adjustment alters country rankings, the integrated index combines results and efficiency measures. Composite index construction for IRPI and IE_IRPI is appropriate given the need to integrate diverse result-based indicators into a coherent performance measure. The IRPI provides a normalised, comparable measure of innovation results, decomposition across outputs and outcomes, transparency in weighting and aggregation decisions, alignment with established guidelines for composite indicator design (Nardo et al., 2008). The process includes indicator normalisation, correlation analysis, weighting, aggregation and sensitivity testing. IE_IRPI uses a performance-adjusted aggregation approach in which result-based innovation performance (IRPI) constitutes the primary ordering criterion, while innovation efficiency (IEI) enters as a bounded, mean-centred modifier that refines but cannot substitute for observed innovation outcomes.

The chapter specifies the procedures used to construct IRPI, IEI and the efficiency-adjusted performance index (IE_IRPI). IEI captures efficiency, enabling diagnosis of Stage 1 and Stage 2 weaknesses. Together, they allow identification of system deficiencies and best peers for policymaking. This dual method design ensures that the empirical results meaningfully test the hypotheses and contribute to innovation policy analysis.

5.1.4 Operationalisation of Research Questions and Hypothesis testing

Chapter 1 established the research questions and hypotheses that guide this dissertation. The conceptual framework (Chapter 4) specified three analytically separable dimensions: performance, efficiency, and capacity, and their theoretical relationships. This section operationalises these constructs for empirical testing, specifying how each research question will be addressed and how each hypothesis will be tested.

The research questions (RQa-RQc) are exploratory: they aim to discover patterns in NIS functioning across EU member countries. The hypotheses (H1 and H2) are confirmatory: they test whether theoretically predicted relationships hold. This section details the specific procedures, variables, and criteria used to address each.

RQa operationalisation: innovation results are measured through IRPI, decomposed into: IKP (Innovation Knowledge Production): captures scientific and technological outputs and IKC (Innovation Knowledge Commercialisation): captures economic and market outcomes. Patterns are identified through descriptive analysis of IKP-IKC configurations across countries, and Section 5.4 details the IRPI construction procedure.

RQb operationalisation: transformation productivity is measured through IEI, decomposed into KPEI (Knowledge Production Efficiency Index) for Stage 1 efficiency and KCEI (Knowledge Commercialisation Efficiency Index) for Stage 2 efficiency. Bottleneck patterns are identified by comparing KPEI and KCEI scores, and Section 5.5 details the two-stage DEA procedure.

RQc operationalisation: joint classification uses the IE_IRPI framework, and countries are mapped onto four quadrants based on median splits of IRPI and IEI. Structural characteristics of quadrant groupings are compared with SII-based groupings and Section 5.6 details the IE_IRPI integration procedure.

Hypothesis testing framework

This dissertation tests whether the measurement framework exhibits the theoretically predicted properties established in Chapter 1. The validation strategy employs correlation analysis, distributional analysis, and robustness testing.

Testing H1: Dimensional distinctiveness and diagnostic value

Chapter 1 established that H1 is supported if: (a) weak correlation between IRPI and IEI; (b) systematic stage-specific efficiency differences; and (c) meaningful performance-efficiency configurations. The operationalisation is as follows:

Test 1a - Dimensional independence:

- Calculate Pearson correlation r(IRPI, IEI) across EU-27.
- Support criterion: $r<0.30$ indicates empirical distinctiveness.
- Falsification criterion: $r>0.50$ would indicate redundancy.

Test 1b - Stage asymmetry:

- Compare KPEI and KCEI for each country.
- Calculate proportion where KPEI > KCEI.
- Support criterion: Systematic asymmetry (>60% of countries showing same pattern).
- Falsification criterion: No systematic pattern (<50% either direction).

Test 1c - Configuration meaningfulness:

- Classify countries into four quadrants using median IRPI and median IEI.
- Support criterion: Countries distributed across all quadrants with theoretically coherent patterns.
- Falsification criterion: >70% concentration in single quadrant.

Testing H2: Peer identification and ranking stability

Chapter 1 established that H2 is supported if: (a) stable rankings under parameter variation; (b) structural coherence within quadrants; and (c) substantive differentiation from aggregate rankings. The operationalisation is the following:

Test 2a - Ranking robustness:

- Monte Carlo simulation: 10,000 draws of $\lambda \in [0.15, 0.35]$
- Calculate Spearman ρ between baseline ($\lambda=0.25$) and simulated rankings
- Support criterion: Mean $\rho>0.90$
- Falsification criterion: Mean $\rho<0.80$ or substantial rank reversals

Test 2b - External validity:

- Correlate IE_IRPI with GDP per capita and Government Effectiveness.
- Support criterion: $r>0.30$ with theoretically expected positive sign.
- Falsification criterion: $r<0.20$ or negative correlation.

Test 2c - Differentiation from SII:

- Compare IE_IRPI and SII rankings.
- Identify countries with rank changes > 5 positions.
- Support criterion: Systematic differences indicating added diagnostic value.
- Falsification criterion: ρ(IE_IRPI, SII) > 0.95 indicating no differentiation.

Scientific credibility requires specifying conditions under which the framework would be disconfirmed in Table 13 below the falsification criteria are presented.

Table 13: Falsification criteria summary

Criterion	Test	Falsification threshold	Implication if met
Performance-efficiency independence	r(IRPI, IEI)	r>0.50	Dimensions redundant, separate measurement unjustified
Stage decomposition value	KPEI-KCEI pattern	No systematic difference	Two-stage adds no diagnostic value
Configuration meaningfulness	Quadrant distribution	>80% in single quadrant	Joint assessment uninformative
Ranking stability	Monte Carlo ρ	ρ<0.90	Rankings are methodological artefacts
External validity	r (IE_IRPI, GDP/GE)	r<0.30 or wrong sign	Construct validity questionable
Differentiation	IE_IRPI vs SII	Identical groupings	No added value over existing index

Source: Own

These thresholds are specified ex ante to ensure that hypothesis testing is not subject to post-hoc interpretation.

This section has operationalised the research questions and hypotheses established in Chapter 1 by specifying:

- the measurement instruments required for each research question,
- the testing procedures for each hypothesis,
- the explicit thresholds for support and falsification, and
- the mapping between methods and research objectives.

The following sections detail the specific methodological procedures used to implement this testing framework.

5.1.5 Research design: performance and efficiency measurement

Two complementary indices are developed to assess national result-based innovation performance and efficiency in the EU member countries. The IRPI is a composite indicator assessing innovation performance based on two sub-dimensions:

- Outputs (knowledge/technological production);
- Outcomes (knowledge/technological commercialisation and economic returns).

The IRPI follows the methodology outlined in the OECD Handbook on Constructing Composite Indicators (Nardo et al., 2008), which prescribes a structured process involving conceptualisation, indicator selection, data treatment, normalisation, weighting, aggregation, and robustness analysis.

The IEI is based on a two-stage network DEA model evaluating the efficiency with which:

- Inputs are transformed into outputs (Stage 1: knowledge/technological production efficiency), and
- Outputs are transformed into outcomes (Stage 2: knowledge/technological commercialisation efficiency).

This modelling approach is consistent with the other DEA studies and applications in innovation efficiency research (Liou, 2009; Guan and Chen, 2012). It captures the sequential nature of innovation production and explicitly accounts for time lags. The transition between stages often involves a time lag, reflecting the reality that it may take years for scientific or technological achievements to generate measurable economic outcomes (ex. patents commercialised into export revenue).

The two indices, IRPI and IEI, thus jointly offer a comprehensive view of innovation system performance. IRPI and IEI rankings identify:

- Top performing innovation systems;
- efficient vs. inefficient systems;
- structural bottlenecks (strong outputs but weak outcomes); and
- countries with strong commercialisation but lagging knowledge production.

Based on the two indices an IE_IRPI is constructed by applying a performance-first composite indicator methodology in which efficiency enters as a bounded diagnostic adjustment, preserving outcome primacy, interpretability, and ranking robustness.

5.2 DATA AND DATA TREATMENT

All indicators used in the IRPI and IEI are derived from the EIS 2024, the most comprehensive and internationally harmonised dataset on innovation in the EU. The dataset of EIS 2024 for 32 indicators for 2017-2024 has been used for the indicator selection, index construction and DEA analysis in order to ensure consistency and comparability with existing benchmarks for the newly developed index. EIS 2024 calculates the SII for all EU-27 countries for 2024 (t+1) based on data collected in 2023 (t). Some of the indicators are collected biannually. EIS provides detailed explanation in their methodology report on the sources of data per indicator, and the description of indicators. As stated in the EIS methodology report for each indicator, a reference year is identified for all countries based on data availability for all those countries for which data availability is at least 75% and for most indicators, this reference year lags two years behind the year in which the EIS is published.

A data limitation concerns the temporal frequency of certain EIS indicators. Several firm-level innovation indicators are collected biennially rather than annually, with intermediate years estimated through interpolation or carried forward from the most recent survey. The annual efficiency and performance scores reported in this dissertation therefore incorporate these imputed values, implicitly assuming smooth inter-survey transitions. This assumption may understate year-to-year volatility in innovation activity, particularly during periods of rapid change such as the COVID-19 crisis (Lundvall, 2023).

The EIS methodology report further provides detailed explanation on the completed data treatment on the raw innovation indicators that includes imputation of missing values per indicator, treatment of outliers and skewness and normalisation and re-scaling procedure. The dataset contains normalised data for all indicators and EU-27 countries for the period 2017-2024, where the only variable missing data after imputation is the 3.2.2 Job-to-job mobility of Human Resources in Science & Technology. Therefore, this indicator is excluded from the indicator selection, DEA and index construction steps. The final dataset thus allows for consistent cross-country comparison and reliable DEA modelling, though DEA results remain sensitive to specification choices, the indicator set was therefore kept parsimonious to preserve discriminatory power.

For the purposes of creating a two-stage model, the EIS indicators are reclassified into the four core categories:

- Inputs (resources);
- Outputs (knowledge/technological production results);
- Outcomes (knowledge/technological commercialisation results); and
- Impact (long-term influences on society and environment).

The impact indicators are excluded from this research because they lie beyond the scope of internal innovation performance. A limitation to this approach is that many key indicators are measured biennially (e.g., digital skills, non-R&D innovation expenditures, SME innovation activities, employment in innovative enterprises, and sales of new-to-market products).

EIS indicators are published in min-max normalised form (0-100 scale). This dissertation uses these pre-normalised values directly for composite index construction and no additional normalisation is applied to avoid double-transformation artifacts. For DEA analysis, the normalised indicators are used as received, ensuring comparability with EIS methodology while maintaining the ratio properties required for efficiency estimation. The normalisation allows each indicator to contribute proportionally to the composite index and is a key step in constructing composite indicators (Nardo et al., 2008). Normalisation converts indicators measured in different units to a common scale without distorting their relative differences. The min-max approach rescales each indicator value between 0 and 1, preserving distributional shape and ensuring interpretability. The minimum re-scaled score equals 0, and the maximum equals 1. EIS uses min-max normalisation:

$$x_i^{\text{norm}} = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}$$

where $x_{\min}$ and $x_{\max}$ represent the lowest and highest non-outlier values, respectively. This approach is scale-invariant, and preserves ordinal relationships (Nardo et al., 2008).

The analysis period 2017-2024 encompasses the COVID-19 pandemic and its aftermath, requiring explicit attention to the timing convention governing innovation indices. Following EIS methodology, indices published in year t+1 reflect indicator data from year t.

The timing framework enables classification of the observation period into three analytically distinct phases, as shown in Table 14.

Table 14: Period classification for crisis analysis

Phase	IRPI years	Data years	Characteristics	Methodological treatment
Pre-crisis	2017-2020	2016-2019	Normal economic conditions	Clean baseline for structural patterns
Crisis	2021-2022	2020-2021	Pandemic shock and recovery	Flag for potential temporary effects
Post-crisis	2023-2024	2022-2023	Towards normal economic conditions	Verify persistence of patterns

Source: Own

This classification informs the interpretation of empirical results in Chapter 6, where patterns consistent across all three phases are classified as structural, while patterns showing significant deviation during the crisis phase are flagged for cautious interpretation.

- The COVID-19 pandemic (2020-2021) represents the most significant exogenous shock to innovation systems in the observation period. Its effects on measured innovation indicators were substantial but heterogeneous.
- Negative effects are disruption of R&D activities, particularly laboratory-based and collaborative research, postponement of firm-level innovation investments due to uncertainty, collapse of global value chains and demand for certain innovative products and services, disruption of international knowledge flows and researcher mobility; and survey response disruptions affecting firm-level innovation statistics.
- Positive effects are acceleration of digital innovation and adoption, emergency R&D in health-related sectors (vaccines, diagnostics, treatments), and increased public R&D expenditure in EU countries.
- The net effect varied substantially across countries, sectors, and indicator types. Countries with stronger digital infrastructure and health-sector specialisation experienced relatively smaller disruptions, while those dependent on physical collaboration and traditional manufacturing faced larger shocks.

The EU's policy response introduced potential structural breaks in innovation financing and activity:

- NextGenerationEU (2021-2026): €750 billion recovery instrument with significant innovation and digitalisation components, potentially elevating innovation investment in recipient countries beyond pre-crisis trends;
- SURE instrument (2020-2022): short-time work support that maintained employment in innovative enterprises;
- Horizon Europe transition (2021): new framework programme with altered funding structures; and
- Green Deal and digital transition policies: accelerated policy push affecting innovation priorities.

These interventions may have altered the structural relationships between inputs, outputs, and outcomes, potentially affecting efficiency measurements.

DEA efficiency scores are particularly sensitive to short-term shocks for several reasons:

- The efficiency frontier is defined by best-performing units in each year. If crisis effects are asymmetric, frontier countries may shift, affecting all relative efficiency scores.

- The implemented time lags mean that: 2020 efficiency scores reflect 2018 inputs/2019 outputs/2020 outcomes, while 2021 efficiency scores reflect 2019 inputs/2020 outputs/2021 outcomes. Crisis disruptions thus propagate through the lag structure with complex timing.
- Several EIS indicators are based on surveys conducted during crisis periods, potentially affected by response rates, firm survival bias, and measurement discontinuities.

This dissertation addresses crisis effects through four mechanisms:

- *Explicit period flagging*: Results are reported with clear identification of pre-crisis, crisis, and recovery periods where relevant.
- *Persistence analysis*: Emphasis on rank persistence across the full period distinguishes structural patterns from temporary fluctuations.
- *Cautious interpretation*: Discussion explicitly flags where observed patterns may reflect crisis artefacts rather than structural differences.

The year-specific results, particularly for 2020-2022, should be interpreted with awareness that they may reflect temporary crisis effects rather than underlying innovation system characteristics. The empirical application of this temporal framework, including rank stability analysis, pre-crisis versus post-crisis comparisons, and classification of structural versus crisis-sensitive findings, is presented in sections 6.5 and 6.7.

5.3 INDICATOR SELECTION PROCEDURE

Variable selection critically shapes indices' results: too many variables reduce discrimination, while omitting key ones biases scores. Researchers have applied methods such as Principal Component Analysis (PCA), Efficiency Contribution Measure (ECM), regression-based selection, bootstrapping, and stepwise approaches. PCA is widely used to reduce dimensionality (Didenko et al., 2019; Adler and Yazhemsky, 2010) and mitigate multicollinearity. Adler and Yazhemsky (2010) demonstrated its efficacy, particularly when input correlations exceed 0.8. However, PCA's transformation of variables compromises interpretability and may exclude temporally relevant features. The ECM method, as analysed by Nataraja and Johnson (2011), evaluates each variable's marginal impact on DEA efficiency. It performs well under low inter variable correlations (<0.2) with sufficiently large datasets (N≥300). While ECM is computationally efficient and intuitive, it lacks the capacity to model underlying structural groupings or temporal changes. Regression-based variable selection techniques estimate the relationship between inputs/outputs and DEA scores, commonly employing OLS or logistic models. These approaches assume linearity and often falter with small samples or multicollinearity. Bootstrapping methods (e.g., Simar and Wilson, 1998) resample data to assess the stability and statistical significance of DEA results under varying model specifications. While statistically rigorous, bootstrapping is computationally intensive and impractical for high dimensional, time dependent analyses. The stepwise approach (Wagner and Shimshak, 2007) sequentially adds or removes variables based on changes in DEA efficiency scores. Although simple and interpretable, this method is prone to overfitting and may yield unstable results with correlated or high dimensional variables. Jenkins and Anderson (2002) proposed statistical correlation-based reduction strategies to eliminate redundant variables. These methods, while reducing dimensionality, may inadvertently discard features with significant explanatory power.

In the research on NISs' efficiency and performance, only a few articles incorporate statistical methods to improve indicator selection and model robustness, but mainly to address multicollinearity and correlations, such as correlation analysis or PCA. To address some of the limitations of these methods, a feature selection rather than feature creation method is applied. This study applies MCFS, which preserves interpretability while capturing latent structural patterns and consistency over time, advantages that distinguish it from traditional methods. To the author's knowledge, MCFS has not previously been applied to indicator selection in innovation index construction.

The aim is to identify input, output and outcome indicators that best differentiate countries year by year. Critically, indicator selection operates within the theoretically defined categories established in Chapter 4: MCFS is applied separately to inputs, outputs, and outcomes, preserving the sequential logic of the innovation process. An indicator classified as an output cannot be reclassified as an outcome based on statistical properties alone. The process first determines the optimal number of clusters using K-Means and silhouette scores (Rousseeuw, 1987), with a minimum cluster size constraint. This score evaluates how similar each sample is to its own cluster versus other clusters, providing a heuristic for the best partitioning. A minimum cluster size constraint is enforced to avoid unstable or meaningless clusters, a practical adaptation for real world datasets where country-level data might be sparse or uneven. Different values of the number of clusters were then examined, and the one with the highest silhouette score was selected, that also respected the cluster size constraint. Once clusters are assigned for a given year, the script employs LASSO regression (Tibshirani, 1996) to perform feature selection. For each cluster, a binary target vector is created (countries in cluster=1, others=0), and Lasso is used to model this classification using all features. We use LASSO and not ordinary least squares regression as it excludes some of the features, for example it sets regression coefficients to exactly zero. This results in sparse solutions where unimportant features are effectively excluded. By aggregating the absolute value of coefficients across all clusters and years, the script produces cumulative importance scores for each feature.

This method combines the local interpretability of LASSO with the global structure provided by clustering, adapting the principles of MCFS as described by Cai et al. (2010).

The applied MCFS enhances the empirical foundation of the IRPI and IEI relative to expert-driven selection, as unlike traditional dimensionality reduction techniques (e.g., PCA) which compromise interpretability by transforming variables, MCFS selects the most discriminatory original features while preserving latent structural patterns. However, the stability of MCFS-based selection in small samples (N=27 countries) should be acknowledged as a limitation; the temporal aggregation across years mitigates but does not fully resolve this constraint. This research implements a time-variant MCFS pipeline designed to identify the input, output, and outcome indicators that best differentiate NIS across the observed period (2017-2024). The selection process follows a three-stage algorithm:

1. Cluster identification (K-Means): for each feature group (Inputs, Outputs, Outcomes) and each year the optimal number of clusters is determined using K-Means clustering. The selection is guided by the silhouette score (Rousseeuw, 1987), subject to a minimum cluster size constraint to ensure stability.
2. Sparse feature scoring (LASSO): once clusters are assigned, the algorithm evaluates the discriminatory power of each feature using LASSO regression (Tibshirani, 1996). For each cluster a binary target vector is created, and a regularised regression is fitted. The absolute coefficients serve as importance scores, effectively excluding irrelevant features where coefficients shrink to zero.
3. Feature selection (dual threshold heuristic): importance scores are aggregated across all years and clusters. The final feature set is selected using a dual threshold heuristic that retains the minimum number of features required to explain 90% of the cumulative importance or the "elbow point" of marginal gain.

This unsupervised approach ensures that the selected indicators are not chosen arbitrarily but are empirically identified as contributing to structural differentiation between EU innovation systems within the observed sample and time period. However, the method's stability may be constrained by the relatively small sample of 27 EU countries. The temporal aggregation across years (2017-2024) partially mitigates this limitation by providing 216 country-year observations for clustering, but the cross-sectional sample remains modest. A sophisticated pipeline for MCFS over time has been implemented, designed to extract meaningful variables from complex international datasets, with indicator selection based on cumulative importance scores aggregated across all years to enhance stability. The detailed implemented MCFS algorithm is presented below.

Algorithm: Multi-Cluster Feature Selection Over Time

Dataset:
$D = \{X_y \mid y \in \{2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024\}\}$, where $X_y \in \mathbb{R}^{27*d}$ for each year y; and d is the number of features.
Feature groups:
$G = \{G_1, G_2, G_3\}$, where each group G_k contains d_k features (columns). These groups are inputs, outputs and outcomes.
Parameters:
- Minimum cluster size: $N_{min} = 3$
- Cluster range: $K_{min} = 2$, $K_{max} = 10$
- Lasso regularization parameter: $\lambda > 0$ (default: $\lambda = 1.0$)
- Cumulative explained importance threshold: $\alpha \in (0.1]$ (default: $\alpha = 0.90$)

Process

```
1. For each feature group $G_k \in G$ (with $d_k = | G_k |$, i. e. number of features):

2. For each year $y \in \{2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024\}$:

  a. Extract group-specific feature matrix: $X_y (G_k) \in \mathbb{R}^{27*dk}$ (select columns for $G_k$ only)
  b. Determine optimal number of clusters $K^*_y$:
    - For each $K \in [K_{min}, K_{max}]$:
        - Fit KMeans: $C_K$ = KMeans($X_y (G_k)$, K)
        - Compute silhouette score $s_K$
        - If any cluster has fewer than $N_{min}$ samples, set $s_K := -\infty$
    - Choose $K^*_y = \text{argmax}_K\, s_K$
  c. Assign clusters:
    - Obtain cluster labels $C_y$ for $K^*_y$ clusters
  d. For each cluster l = 1, ..., $K^*_y$:
    - Construct binary target $z_{y,l}$ (indicator for cluster membership)
    - Fit Lasso regression:
        $\hat{\beta}_{y,l} = arg\,\min_{\beta} \frac{1}{2*N\,y} \left\| z_{y,l} - X_y(G_k) \cdot \beta \right\|_2 + \lambda \cdot \|\beta\|_1 = arg\,\min_{\beta} \frac{1}{54} \left\| z_{y,l} - X_y(G_k) \cdot \beta \right\|_2 + \lambda \cdot \|\beta\|_1$
        (since $N_y = 27$, denominator is $2*N_y = 54$). Here $||_2$ stands for Eucledean (sum of squares) and $||_1$ for Manhattan norm (sum of absolute values)
    - Store absolute coefficients $a_{y,l} = |\hat{\beta}_{y,l}| \in \mathbb{R}^{dk}$

3. Aggregate feature importance across all years and clusters:
   For each feature f = 1, ..., $d_k$:
     $I_f = \sum_{y \in \{2017,2018,2019,2020,2021,2022,2023,2024\}} \sum_{l=1}^{K^*y} a_{y,l}$
```

4. Feature selection with dual-threshold heuristic:
   Let I_f be the importances sorted in decreasing order: $I_1 \geq I_2 \geq \ldots \geq I_{dk}$.
   a. Elbow method:
      $f_{elbow} = \text{argmax}_{f=1,\ldots,dk-1} (I_f \geq I_{f+1})$
   b. Cumulative explained importance (α):
      Find smallest n_α such that $\frac{\sum_{F=1}^{n_\alpha} I_f}{\sum_{f=1}^{d_k} I_f} \geq \alpha$

c. Final feature set:
      Let $n^* = \min(f_{elbow}, n_\alpha)$
      Select $F^* = \{f_1, \ldots, f_{n^*}\}$

Results

For each group G_k: selected features $G_k^* \subseteq \{1, \ldots, d_k\}$

5.4 OPERATIONALISING RESULT-BASED INNOVATION PERFORMANCE

5.4.1 Weighting

After MCFS indicator selection, the IRPI is calculated as an arithmetic mean of two subindices: the Output subindex focused on knowledge/technological production (IKP) and the Outcome subindex focused on knowledge/technological commercialisation (IKC). This approach follows established precedent in policy-relevant composite indicators, including the Human Development Index (United Nations Development Programme [UNDP], 1990; Mariano et al., 2021) and the UN Index of Technological Advancements.

Weighting assigns relative importance to indicators in aggregation and represents one of the most influential and methodologically sensitive steps in constructing composite indicators. Different weighting schemes can substantially alter country rankings, interpretative meaning, and policy conclusions. Critically, all weighting choices, including equal weighting, embody normative judgments about relative indicator importance (Nardo et al., 2008; Greco et al., 2019). There is no "neutral" weighting scheme; equal weights implicitly assume that each indicator contributes equivalently to the underlying construct, while data-driven weights assume that statistical variance corresponds to conceptual importance. This dissertation explicitly acknowledges that the adopted equal weighting scheme constitutes a value-based methodological decision rather than an empirically derived parameter.

The literature identifies four major classes of weighting approaches:

- Equal weighting remains the most widely used approach in policy practice and official scoreboards, including the EIS and the Human Development Index (UNDP, 1990). Its popularity derives from transparency and ease of communication. According to OECD (Nardo et al., 2008), equal weighting is a defensible normative choice when indicators are selected through rigorous screening and when expert consensus on relative importance is lacking. Equal weights reduce the risk of overstating methodological precision and provide a neutral baseline from which sensitivity analysis can be conducted. However, as Greco et al. (2019) highlight, equal weighting may implicitly assume compensability and may overlook potential redundancies between indicators.
- Data-driven weighting methods introduce empirical structure into the weighting process. PCA and factor analysis allocate weights according to variance explained by each indicator (Nardo et al., 2008). Entropy weighting (Zeleny, 1982), which applies Shannon's (1948) information-theoretic entropy measure, assigns larger weights to indicators with greater informational content (e.g. greater dispersion across units). Although analytically sophisticated, these methods can produce counterintuitive results, give excessive influence to statistical noise, or obscure the conceptual meaning of indicator importance (Nardo et al., 2008).
- Expert weighting, operationalised through Budget Allocation Process, or Delphi method, incorporates normative judgments and stakeholder priorities. Munda and Nardo (2005) note that expert weighting enhances legitimacy in policy contexts but introduces subjectivity, depends on expert panel representativeness, and may yield unstable rankings.
- Uncertainty-based and robustness weighting approaches recognise that different weighting schemes reflect different normative assumptions. Weight-uncertainty analysis (OECD, 2005) generates thousands of weighting vectors through Monte Carlo simulations to evaluate sensitivity of composite scores to alternative schemes. This approach is now considered best practice by the OECD for policy-informing composite indices.

The adoption of equal weighting in IRPI construction is justified on the following grounds aligned with the research objectives. *First*, the research objective is to construct a result-based performance measure that captures observable innovation outputs and outcomes without imposing assumptions about which results matter more. Innovation systems theory does not provide unambiguous guidance on whether, for example, patent applications should be weighted more heavily than SME innovation rates. In the absence of such theoretical consensus, equal weighting represents a principled position of agnosticism rather than arbitrary assignment (Nardo et al., 2008). *Second*, the MCFS procedure already identifies indicators based on their empirical contribution to cross-country differentiation over time. Indicators that do not consistently discriminate across countries and years are excluded. Consequently, all retained indicators have demonstrated empirical relevance, reducing the need for further differential weighting. Equal weighting of MCFS-selected indicators avoids double-counting the selection criterion and preserves the interpretability of the composite. *Third*, a core research objective is to develop a framework suitable for policy-oriented comparative assessment. Equal weighting maximises transparency: users can understand exactly how the index is constructed without requiring access to proprietary weighting algorithms or expert panel deliberations. This transparency supports the framework's intended use for benchmarking and policy dialogue.

The two subindices (IKP and IKC) receive equal weight (0.5 each) in the final IRPI aggregation. This reflects the theoretical position, developed in Chapter 4, that result-based innovation performance requires both strong knowledge production capabilities and effective commercialisation. Assigning unequal weights would imply that one stage of the innovation process is inherently more important than the other, a normative judgment for which no theoretical consensus exists. The equal weighting of stages operationalises the conceptual framework's treatment of innovation as a sequential process where both stages are necessary conditions for system-level performance.

Equal weighting is not without limitations. It assumes full compensability (strong performance on one indicator can offset weak performance on another) and treats all indicators as equally important despite potential differences in measurement precision, policy relevance, or causal proximity to ultimate innovation outcomes. These limitations are inherent to any aggregation scheme and are addressed through reporting subindices separately to preserve stage-specific diagnostics, conducting robustness analysis to verify that conclusions are not weighting-dependent, and interpreting IRPI as one component of a multidimensional framework rather than as a definitive single-number summary.

5.4.2 Aggregation

The IRPI formula for country i is:

$$IRPI_i = \frac{1}{2} IKP_i + \frac{1}{2} IKC_i$$

Where $IRPI_i$represents the result-based innovation performance score of country i. The output subindex IKP_i captures performance in knowledge production, such as top-cited scientific publications, design applications, and innovation-active SMEs. The outcome subindex captures commercialisation performance, such as high-tech exports and knowledge-intensive services exports. Both dimensions receive an equal weight (0.5), reflecting the theoretical position that result-based innovation performance requires both strong technological capabilities and strong market uptake.

The subindices are calculated as weighted sums of the selected indicators appropriately for each subindex. The subindices are calculated by using equal weights as well. Output subindex, denoted IKP_i, is the average of all output indicators for country i. Outcome subindex, denoted IKC_i, is the average of all outcome indicators for country i. The subindices (IKP) and (IKC) are calculated by the following formulae:

For country i:

$$IKP_i = \frac{1}{n_{IKP}} \sum_{j=1}^{n_O} X_{ij}$$

$$IKC_i = \frac{1}{n_{IKC}} \sum_{k=1}^{n_{OC}} X_{ik}$$

where $n_{IKP}$and n_{IKC} are the number of output indicators and outcome indicators, respectively.

5.5 OPERATIONALISING STAGE-SPECIFIC EFFICIENCY

5.5.1 Time-lag specification

The conceptual framework in Chapter 4 establishes that innovation transformation involves temporal processes spanning multiple years. The theoretical literature suggests lags of 1-3 years for knowledge production (Stage 1) and 1 years for commercialisation (Stage 2), with aggregate lags of 2-4 years from initial investment to commercial outcomes. The theory overview on lag implementation is presented in Table 15. However, operationalising these conceptual lags in empirical models requires choices constrained by data availability, model parsimony, and comparability with prior research.

Table 15: Lag implementation

Stage	Conceptual lag (Theory)	Implemented lag (Model)	Justification
Stage 1: Inputs into Outputs	1-3 years	1 year	Parsimony, EIS indicator timing, prior DEA studies.
Stage 2: Outputs into Outcomes	1 year	1 year	Parsimony, EIS indicator timing, prior DEA studies.
Total	2-4 years	2 years	Lower bound of theoretical range.

Source: Own

The one-year lag per stage (two-year total lag) is selected based on four considerations: *First*, most two-stage DEA studies on NIS implement one-year lags per stage (Carayannis et al., 2016) and similarly implement one-year lags in multi-level innovation efficiency analysis. Adopting this convention ensures comparability with the established literature. *Second*, EIS indicators are constructed with implicit one-year lags built into the data collection process. Several indicators (particularly firm-level survey data) are collected biennially, creating additional timing constraints. The implemented one-year lag per stage aligns with this data structure. *Third*, longer lags would reduce the effective panel length from 6 years (2019-2024 with efficiency scores) to fewer observations, potentially compromising the statistical reliability of DEA frontier estimation. The 27-country sample with one-year lags provides sufficient observations for robust efficiency discrimination. *Fourth*, the implemented two-year total lag represents the lower bound of the 2-4-year theoretical range. This conservative choice is appropriate for a cross-sectional EU comparison where member countries share relatively similar institutional frameworks and technology transfer mechanisms, potentially compressing lag structures relative to global comparisons.

The empirical results reported in Chapter 6 are conditional on this lag specification. Specifically:

- Efficiency scores (KPEI, KCEI, IEI) measure how productively inputs from year t are transformed into outputs in year t+1 and outcomes in year t+2;
- Countries with faster innovation cycles may appear relatively more efficient under this specification;
- Countries with longer-gestation innovation processes (e.g., pharmaceuticals, basic research-intensive) may be disadvantaged; and
- The persistence patterns observed in efficiency rankings reflect both genuine system characteristics and the imposed lag structure.

This operationalisation represents a conservative lower-bound estimate, and longer lags would reduce the panel substantially. Results should therefore be interpreted as conditional on this modelling choice.

5.5.2 Data Envelopment Analysis

For the IEI calculation this study applies a DEA, following the foundational DEA framework (Charnes et al., 1978) and its extensive subsequent applications expanding from the early 2000s until today (Cooper et al., 2007; Simar and Wilson, 2007; Cullmann et al., 2012; Guan and Chen, 2012; Aristovnik, 2014; Carayannis et al., 2016; Roszko-Wójtowicz and Białek, 2016; Jurickova et al., 2019; Kalapouti et al., 2020; Klevenhusen et al., 2021; Jovanović et al., 2022). The selected model is VRS and output-oriented, as VRS is advantageous over CRS when evaluating entities of different sizes and where doubling inputs does not necessarily double outputs due to inefficiencies, coordination costs, or structural constraints. An output-oriented model was selected to evaluate how effectively EU countries maximise innovation outputs given their resources. Table 16 summarises methodological choices in comparable DEA studies on NIS efficiency.

Table 16: Selected scientific articles for comparison of DEA methodology

Scientific articles	**This research**	**Edquist et al. (2018)**	**Carayannis et al. (2016)**	**Anouze et al. (2024)**	**Guan and Chen (2012)**	**Hudec and Procházková (2013)**	**Liou (2009)**
Indicators Selection method	Yes, MCFS	No	No	No	No	No	No
Data sources	IUS	IUS	IUS, WIPO	GII	OECD, IMD	OECD, Eurostat	OECD
Period	2017-2024	2014-2015	2007-2011	2016-2021	1999-2003	2004-2010	2000-2005
Two-stage DEA	Yes	No, one	Yes	Yes	Yes	Yes	Yes
Applied Method	VRS output - oriented DEA	Input-oriented CRS DEA	MOLP-DEA VRS, input oriented - stage 1 output oriented - stage 2	DNDEA-VRS non-oriented	PLSR VRS and CRS	CRS and VRS input-oriented	Malmquist DEA
Analysed NISs	EU-27 countries	EU-28 countries	EU-23 countries	33 oil countries	OECD-22 countries	EU-19 countries	15 OECD countries

Source: Own

In our output-oriented DEA, DMUs are NISs of the 27 EU countries, where NISk, k = 1, 2, …, 27; for each NIS we perform DEA for 2017-2019, 2018-2020, 2019-2021, 2020-2022, 2021-2023 and 2022-2024 where each DMU in each year is treated as if it were a distinct DMU. The DEA considers the time lag of one year for inputs to transform in outputs, and one year for outputs to transform in outcomes, similar to the previous studies that apply two-stage DEA. The stage 1 efficiency or knowledge/technological production efficiency is calculated with 7 input indicators selected with MCFS, on one side, and one output variable, the output subindex IKP. The stage 2 efficiency or knowledge/technological commercialisation efficiency is calculated with 7 output indicators selected with MCFS (in this stage used as inputs), on one side, and one output variable in this stage, the outcome subindex IKC.

The DEA specification uses the constructed subindices IKP and IKC as output variables rather than raw indicators. This approach follows established practice in the DEA and composite indicator literature. El Gibari, Gómez, and Ruiz (2022) explicitly combine reference point-based composite indicators with DEA for university assessment, using composite indicators as efficiency outputs.

The use of aggregated subindices satisfies the key requirements for DEA variables established by Dyson et al. (2001): both IKP and IKC are measured on ratio scales (0-1 normalised), exhibit isotonic relationships with inputs, and maintain favourable DMU-to-variable ratios. Using 7 raw output indicators as DEA outputs would yield a ratio of 27 DMUs to 14 variables (7 inputs and 7 outputs) per year, falling below the recommended threshold of $n \geq 2\times$(inputs + outputs) for adequate discriminatory power.

The aggregated specification embeds the equal-weighting assumption from subindex construction into efficiency measurement. This is a methodological choice that prioritises theoretical alignment with the two-stage innovation framework and interpretability of efficiency scores over the flexibility of endogenous DEA weighting. The independence of IRPI and IEI is confirmed empirically in Chapter 6, consistent with the conceptual framework developed in Chapter 4. Below we present the two-Stage DEA Efficiency Analysis for 27 EU member countries, for the period 2017-2024.

Dataset

Years: Y = {2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024}
Countries: $C = \{c_1, c_2, ..., c_{27}\}$
For each year y, built matrices:
  - Inputs: $X_y \in \mathbb{R}^{27\times 7}$ (countries × inputs)
  - Outputs: $M_y \in \mathbb{R}^{27\times 7}$ (countries × outputs)

  - Outputs subindex: $IKP_y \in \mathbb{R}^{27\times 1}$ (countries × output subindex)
  - Outcomes subindex: $IKC_y \in \mathbb{R}^{27\times 1}$ (countries × outcomes subindex)

Process

DEA Stage 1 - Technological/Knowledge Production Process (Inputs to Output subindex)

For each year y in {2018, 2019, 2020, 2021, 2022, 2023, 2024}:

Applied output-oriented DEA using X_y as inputs and IKP_{y+1} as output. To account for time lags, inputs from 2017 correspond to output subindex for 2018, inputs from 2018 align with output subindex for 2019, etc.

Obtained efficiency scores $KPEI_{c,y}(1)$ for all c in C.

DEA Stage 2 - Technological/Knowledge Commercialisation Process (Outputs to Outcomes subindex)

For each year y in {2019, 2020, 2021, 2022, 2023, 2024}:

1. Applied output-oriented DEA using M_y as outputs (here inputs to the stage) and IKC_{y+1} as outcomes subindex (here output to the stage). To account for time lags, outputs from 2018 correspond to outcomes subindex for 2019, outputs from 2019 correspond to outcomes subindex for 2020, etc.

2. Obtained efficiency scores $KCEI_{c,y}(2)$ for all c in C.

Results

For each country c and each year y:

Stored and reported pair of efficiency scores: ($KPEI_{c,y}(1)$, $KCEI_{c,y}(2)$), based on the scores for each stage we calculate the overall NIS efficiency index (IEI) as a product of the scores in every stage (KPEI x KCEI).

The dataset used for the DEA analysis was constructed using variables identified by the MCFS algorithm, ensuring the most relevant features were included in each stage. The program used for calculating the relative efficiency scores is the Frontier Analyst 4.0. An IEI score equal to 1.00 indicates that a country is efficient relative to the observed sample and the constructed DEA frontier, while values below 1.00 indicate relative inefficiency in at least one stage of the innovation process. Therefore, policymakers could identify the needs for structural changes and focus their actions.

Because NIS operate as serial production processes in which knowledge creation (Stage 1) must precede and enable knowledge commercialisation (Stage 2), the appropriate aggregation of stage efficiencies follows the logic of two-stage network DEA. In a serial network, the overall system efficiency is mathematically equal to the product of the component stage efficiencies, reflecting proportional propagation of inefficiency through the innovation pipeline. This multiplicative form arises from the structure of the two-stage relational network DEA model (Kao and Hwang, 2008; Kao, 2009), building on the network DEA framework introduced by Färe and Grosskopf (2000). Therefore, given that:

- Stage 1: knowledge/technological production efficiency (KPEI) with value between 0 and 1, and
- Stage 2: knowledge/technological commercialisation efficiency (KCEI) with value between 0 and 1,

the overall NIS efficiency IEI = KPEI × KCEI, that means a country is fully efficient only if both stages operate at the frontier and the weakness in either knowledge production or commercialisation reduces total system performance. IEI shows how efficiently the NIS transforms original inputs into final innovation outcomes, taking both stages jointly as a production chain. A key methodological limitation concerns the scale sensitivity of DEA when applied to highly heterogeneous NIS. Because DEA evaluates efficiency in relative terms and constructs the frontier endogenously, countries with very low input levels may attain frontier efficiency despite modest absolute innovation outputs. In such contexts, efficiency scores capture input-output proportionality rather than innovation strength or system maturity. Accordingly, efficiency estimates in this dissertation are interpreted as conditional measures of input-output proportionality rather than as indicators of innovation strength or maturity. This limitation reinforces the analytical necessity of maintaining a strict separation between efficiency and result-based performance and motivates the use of efficiency metrics as complementary diagnostics rather than as standalone evaluative measures.

Because DEA evaluates efficiency in relative terms and constructs the frontier endogenously from observed best-practice units, countries with very low input levels may attain frontier efficiency despite modest absolute innovation outputs. This occurs because such countries are effectively benchmarked against similarly resource-constrained peers rather than against high-performing innovation leaders. In such contexts, efficiency scores capture input-output proportionality, the absence of measurable waste given available resources, rather than innovation strength or system maturity (Cooper et al., 2007; Bogetoft and Otto, 2011). This limitation has three implications for the interpretation of results in Chapter 6:

- efficiency scores for low-input countries must be interpreted diagnostically and jointly with performance measures. A country achieving efficiency score of 1.00 with very low IRPI should be understood as operating proportionally, not wasting its limited resources, rather than as an innovation leader. The DEA frontier in such cases reflects the best-observed practice among similarly constrained systems, not a global performance standard.
- the analytical necessity of maintaining strict separation between efficiency and result-based performance is reinforced. Conflating these dimensions, as some composite indices do, would allow high efficiency to compensate for weak absolute outcomes, producing misleading country assessments. The framework developed in this dissertation explicitly preserves this separation.
- the efficiency-adjusted performance index is designed to mitigate scale bias by anchoring rankings in absolute performance while using efficiency scores only as a bounded modifier. This ensures that countries appearing efficient solely due to low scale cannot rank highly in the integrated assessment. Thus, efficiency adjustments are constrained to a narrow band around performance scores, preventing frontier artefacts from dominating country rankings.

Accordingly, the efficiency estimates reported in Chapter 6 are interpreted as conditional measures of input-output proportionality within the observed EU-27 sample, not as indicators of absolute innovation strength or system maturity. This interpretive framework should guide the reading of all efficiency results that follow.

5.6 EFFICIENCY-ADJUSTED INNOVATION PERFORMANCE INDEX

The result-based innovation performance reflects the level of realised innovation outputs and outcomes and therefore constitutes the primary criterion for evaluating NIS, whereas innovation efficiency captures the conditional effectiveness with which available inputs are transformed into outcomes and serves a diagnostic role (Edquist, 2011). Thus, its interpretation is bounded by scale heterogeneity and the relative nature of DEA-based efficiency frontiers. Given this conceptual asymmetry, symmetric aggregation methods, such as arithmetic or geometric means or simple products, are methodologically inappropriate, as they implicitly assign equal importance to outcomes and efficiency (Nardo et al., 2008). To preserve outcome primacy while incorporating efficiency information, this study adopts a performance-adjusted aggregation approach, in which efficiency enters as a bounded modifier of result-based innovation performance rather than as a co-equal component. This approach ensures that countries are ranked primarily according to observed innovation results, while efficiency serves as a diagnostic correction reflecting how effectively those results are achieved.

By construction, efficiency cannot compensate for weak performance, nor can inefficiency fully negate strong outcomes. The resulting ranking therefore balances outcome relevance with system functioning, avoiding the distortions inherent in efficiency or frontier-based rankings. The adjusted ranking must satisfy the following three conditions: IRPI is the primary ordering criterion, IEI acts as a penalty or bonus around performance, not as a substitute and adjustment must be bounded and monotonic. This rules out multiplicative indices (pure products) that excessively penalise partial inefficiency and amplify DEA frontier bias, particularly in small or low-input systems. The geometric mean implies equal conceptual importance of dimensions, which is inappropriate when one dimension is diagnostically secondary. Arithmetic mean allows full compensation between performance and efficiency, enabling highly efficient but low-performing systems to rank favourably.

The efficiency-adjusted performance index is computed based on the following formula:

$$\boxed{IE\_IRPI_i \;=\; IRPI_i \times [1 + \lambda \cdot (IEI_i - I\bar{E}I)]}$$

where:

- IE_IRPI_i is the composite indicator for country i,
- $IRPI_i$ is innovation performance,
- IEI_i is innovation efficiency,
- $I\bar{E}I$ is the EU-27 mean efficiency, and

- λ is a scalar parameter controlling the strength of the efficiency adjustment.

The adjustment parameter λ determines the maximum influence of efficiency on the composite index. It is not a statistically estimated coefficient but a normative-methodological parameter, constrained by theoretical considerations and robustness requirements (OECD, 2018).

The efficiency adjustment parameter λ governs the strength of efficiency's influence on performance-adjusted rankings. Its specification requires balancing two competing theoretical considerations: λ must be sufficiently large to allow efficiency information to meaningfully differentiate countries with similar performance levels, yet sufficiently small to preserve the conceptual primacy of realised outcomes over transformation processes established in innovation systems theory. Three criteria guide the specification of λ bounds:

- The conceptual framework positions result-based performance as the primary evaluation criterion, with efficiency serving a diagnostic rather than evaluative role. This asymmetry, grounded in innovation systems theory's emphasis on what systems ultimately achieve rather than how productively they operate (Edquist, 2011), requires that efficiency adjustments remain bounded and unable to override substantial performance differences. Values of λ exceeding 0.30, risk allowing efficiency to dominate performance locally, enabling high-efficiency, low-outcome systems to overtake higher-performing peers. This result would contradict the outcome-primacy principle central to the framework's theoretical foundations.
- The lower bound of λ must ensure that efficiency information retains meaningful discriminatory power among countries with comparable performance levels. Values below approximately 0.15 render efficiency adjustments negligible, eliminating the diagnostic value that motivates incorporating efficiency into an integrated assessment. The parameter must be large enough that efficiency differences within performance bands translate into observable ranking refinements, enabling the identification of systems that achieve similar outcomes through more or less productive transformation processes.
- A fundamental requirement for interpretable composite indicators is that rankings remain monotonic in the primary ordering criterion for units with similar secondary characteristics (Nardo et al., 2008). Applied to IE_IRPI, this means that countries with substantially higher result-based innovation performance should not rank below lower-performing countries solely due to efficiency differences. The upper bound of λ is therefore constrained by the requirement that performance monotonicity is preserved across the observed range of efficiency values in the EU-27 sample.

The bounded parameter design embeds an implicit falsification test. If country rankings proved highly sensitive to the efficiency adjustment parameter, exhibiting rank correlations below ρ=0.90 across the policy-plausible range or systematic tier reclassifications, the aggregation rule would be arbitrary rather than structurally grounded. Section 6.7 evaluates this criterion through Monte Carlo simulation across 10,000 parameter draws.

These theoretical criteria establish a bounded interval within which the efficiency adjustment parameter must fall to satisfy the framework's conceptual requirements. The specific calibration of λ within this interval, and empirical verification of ranking stability across alternative specifications, is addressed through Monte Carlo simulation in Section 6.7.

5.7 VALIDATION AND ROBUSTNESS ANALYSIS

Section 4.2.3 established the theoretical expectation that IE_IRPI should correlate positively with observable determinants associated with innovation capacity. This section specifies how this validation is operationalised. Two external variables are selected to test the capacity-determinants relationship:

- GDP per capita: serving as a composite proxy for accumulated capabilities and economic development, and
- Government effectiveness: serving as a proxy for institutional coherence and governance quality.

These variables are selected based on four criteria:

- Theoretical relevance: Both are explicitly linked to capacity components identified in Section 4.1, GDP per capita reflects accumulated capabilities, while government effectiveness reflects institutional coherence.
- Data availability: Both are available across all EU-27 countries for the analysis period.
- Independence from index construction: Neither variable is derived from EIS indicators used to construct IRPI and IEI, ensuring that observed correlations cannot be attributed to circular construction.
- Precedent in literature: Both are widely used in comparative innovation research as correlates of system functioning (Furman et al., 2002; Secka, 2023).

The validation employs correlation and regression analysis examining the relationship between IE_IRPI and the two determinants. Specifically:

- Pearson correlation coefficients assess bivariate relationships, and

- Multiple regression estimates the joint predictive relationship.

A positive and statistically significant relationship between IE_IRPI and these determinants would support the framework's construct validity, confirming that countries achieving high integrated performance-efficiency scores also possess the structural and institutional foundations associated with high capacity in the innovation systems theory. The absence of significant relationships would raise questions about whether the measured dimensions capture theoretically meaningful variation, suggesting either measurement problems or misspecification of the capacity concept. This validation strategy maintains strict analytical separation: GDP per capita and government effectiveness serve exclusively as external validation criteria, not as components of index construction. Positive relationships therefore reflect genuine co-variation between measured innovation outcomes and deeper system properties.

Regression analysis remains a central statistical tool for modelling relationships among innovation indicators and external variables. The regression analysis that is employed to assess the external validity and interpretability of the proposed IE_IRPI does not focus on establishing causal relationships. Following innovation systems theory, innovation outcomes are understood to reflect long-term structural conditions, including economic development and institutional quality, while efficiency captures relative performance conditional on inputs (Edquist, 2011). GDP per capita and government effectiveness are therefore introduced as contextual correlates, not as determinants incorporated into the index itself, in order to avoid circularity and conceptual overlap. GDP per capita is used here as a summary indicator of long-run economic prosperity rather than as a direct measure of innovation performance. While income levels reflect a wide range of historical, structural, and institutional factors, innovation outcomes are widely recognised as a key contributor to sustained income differences across countries. A positive association between IE_IRPI and GDP per capita is therefore interpreted as evidence of external alignment, not as proof of proportionality or short-run causation. Government effectiveness is included as an institutional reference variable reflecting the quality of public administration, policy implementation, and regulatory capacity. It is not considered part of result-based innovation performance itself, but a contextual factor that conditions innovation outcomes over time. The relationship with IE_IRPI is examined to assess whether realised innovation results tend to co-vary with institutional quality, consistent with the NIS literature.

Linear regression models the relationship between a dependent innovation variable Y(innovation index) and one or more explanatory variables X(ex. R&D intensity, Government effectiveness etc.). The simple regression model takes the general form:

$$Yi=\beta 0+\beta 1\ Xi1+\cdots+\beta k\ Xik+\varepsilon i,$$

The Classical Linear Regression Model (CLRM) provides the theoretical framework for OLS estimation. In a purely cross-sectional setting, the Gauss-Markov assumptions ensure unbiased and efficient estimates under homoskedasticity and exogeneity. However, cross-country innovation indicators frequently exhibit heteroskedastic error structures due to scale effects and structural heterogeneity across economies (Castellacci, 2008). Accordingly, inference is supplemented with heteroskedasticity-robust standard errors to ensure the reliability of statistical conclusions.

In addition to econometric validation, Monte Carlo simulation and scenario analysis play an increasingly important role in testing the robustness of innovation indicators and composite indices. Monte Carlo simulation is widely used to evaluate the sensitivity of result-based innovation performance rankings to uncertainty in data, weighting, and methodological assumptions. The approach is comprehensively treated in Glasserman (2003) for modern risk analysis, and in innovation contexts it is also applied to measure the stability of composite indicators under shocks, missing data, alternative normalisation methods, or different weighting schemes. Recent studies use Monte Carlo simulation to assess robustness of innovation indices such as the GII (Alqararah, 2023). Monte Carlo simulation proceeds through the following steps:

1. Generate data according to a defined distribution or empirical re-sampling approach.
2. Estimate the model and compute relevant statistics (ex. IE_IRPI score, rankings).
3. Store the results.
4. Repeat the process many times (ex. 10,000 simulations).

By generating probability distributions of outcomes, Monte Carlo simulation provides a robustness envelope around country rankings, highlighting stable versus unstable cases. The methodological framework established in this chapter rests on several assumptions that require empirical validation in Chapter 6:

- One-year transformation lags: the assumption that inputs transform into outputs in approximately one year, and outputs into outcomes in approximately one year, is operationalised but requires validation through temporal pattern analysis.
- MCFS indicator stability: the assumption that MCFS yields stable, non-redundant indicators is tested by examining whether selected indicators contribute consistently to cross-country differentiation across the 2017-2024 period.
- Two-stage decomposition value: the assumption that two-stage DEA reveals diagnostic information invisible to single-stage models is validated if systematic KPEI-KCEI asymmetries are observed.

- Performance-efficiency independence: the assumption that performance and efficiency capture distinct dimensions is validated if the correlation between IRPI and IEI is weak ($r<0.50$).
- Aggregation robustness: the assumption that IE_IRPI rankings are not parameter-dependent is validated if Monte Carlo simulation across the bounded λ range produces rank correlations exceeding $\rho=0.90$.

This chapter has operationalised the conceptual framework developed in Chapter 4 through three measurement instruments: IRPI for result-based performance, IEI for stage-specific efficiency, and IE_IRPI for their integration. The MCFS procedure provides transparent, data-driven indicator selection that addresses redundancy and input dominance in existing frameworks. The two-stage DEA model implements the sequential structure (knowledge production into commercialisation) that innovation systems theory implies. The performance-first aggregation rule ensures that efficiency refines but does not overturn outcome-based assessments. The methodological framework rests on several assumptions that require empirical validation: one-year transformation lags produce interpretable temporal patterns, MCFS yields stable, non-redundant indicators; two-stage decomposition reveals information invisible to aggregate models; performance and efficiency capture empirically distinct dimensions ($r<0.50$), and IE_IRPI rankings are robust to parameter variation ($\rho>0.90$ across Monte Carlo draws).

Chapter 6 applies these instruments to the EU-27 over 2017-2024 and reports results against each validation criterion. It addresses the three research questions formulated in Chapter 1: RQa (what patterns of variation exist in innovation results), RQb (where efficiency bottlenecks concentrate), and RQc (what structural characteristics emerge when both dimensions are considered together). The chapter also tests the hypotheses derived in Chapter 4, providing the evidence base for the interpretation and contribution assessment undertaken in Chapter 7.

6 EMPIRICAL RESULTS

This chapter presents the empirical results of applying the measurement framework developed in Chapters 4-5 to the EU-27 over the period 2017-2024. The analysis addresses the three research questions formulated in Chapter 1 and tests the hypotheses derived in Chapter 4.

Due to the one-year lag structure in the two-stage DEA model, the indices have different coverage periods. IRPI (result-based performance) is computed for 2017-2024, while IEI (efficiency) and IE_IRPI (efficiency-adjusted performance) cover 2019-2024, as efficiency calculation requires lagged input data.

The chapter is structured sequentially, following the logic of the conceptual framework. Section 6.1 provides empirical justification for new index construction through correlation analysis of EIS indicators. Section 6.2 reports the MCFS indicator selection results. Section 6.3 presents IRPI results, addressing RQa (result-based innovation performance patterns). Section 6.4 reports IEI results, addressing RQb (stage-specific efficiency and bottlenecks). Section 6.5 presents the integrated IE_IRPI assessment, addressing RQc (performance-efficiency configurations). Section 6.6 examines external criterion validity through regression analysis. Section 6.7 reports robustness and sensitivity analysis including Monte Carlo simulation. Section 6.8 synthesises findings against the hypotheses. Any references to innovation capacity in this chapter are explicitly interpretive and explanatory in nature. Capacity is not operationalised or measured empirically, as it is rather invoked solely to contextualise observed performance-efficiency patterns within broader structural and institutional conditions.

6.1 EMPIRICAL EVIDENCE OF EUROPEAN INNOVATION SCOREBOARD LIMITATIONS

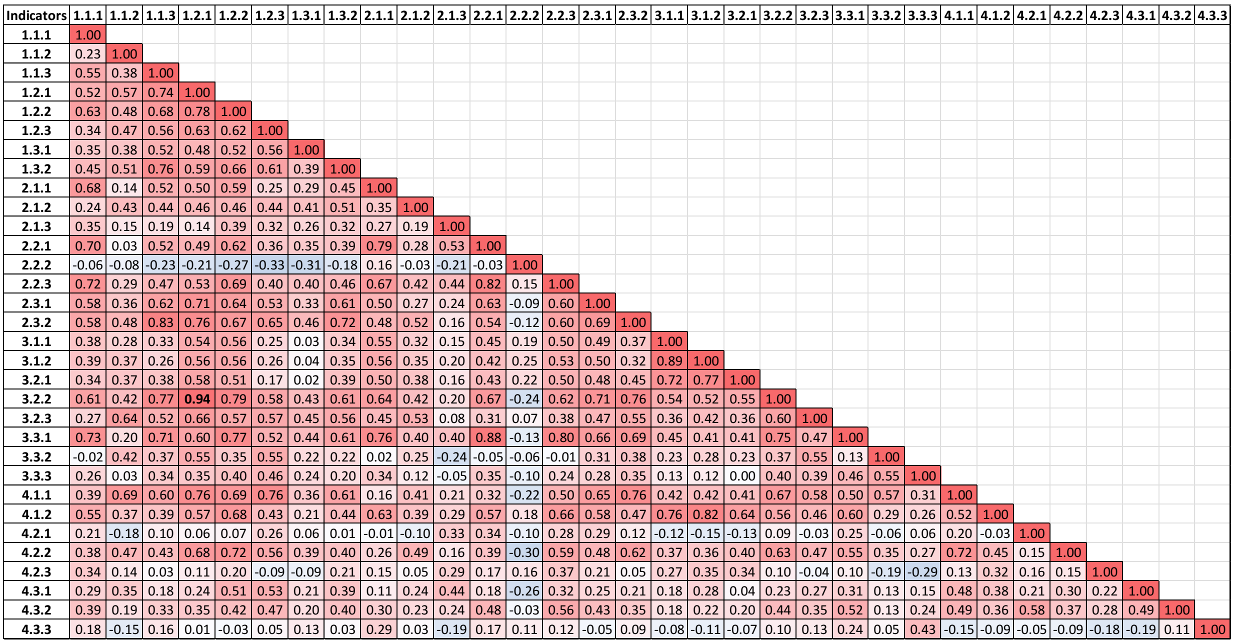
The correlation analysis of all 32 EIS indicators was conducted in SPSS on the normalised and fully treated dataset for the period 2017-2024. Definitions and data sources are provided in EIS 2024 Methodology Report (Hollanders et al., 2024)[2] published by the Directorate-General for Research and Innovation, EC on 8th of July 2024. The Pearson correlation matrix (N=216 observations) shows positive and statistically significant associations across almost all indicator pairs, with no substantial negative correlations emerging in the dataset. This pattern indicates that innovation-relevant inputs, framework conditions, firm activities, outputs and impacts tend to move

---

[2] https://research-and-innovation.ec.europa.eu/document/download/074d5495-433a-440f-bcf9-dc620fce7af1_en?filename=ec_rtd_eis-2024-methodology-report.pdf accessed on 15.12.2025

jointly across European countries. This correlation analysis results in Table 17 confirms that the EIS indicator set is highly internally correlated and structurally redundant.

Table 17: Correlation of 32 EIS 2024 indicators for EU-27 (2017-2024)



Indicators	1.1.1	1.1.2	1.1.3	1.2.1	1.2.2	1.2.3	1.3.1	1.3.2	2.1.1	2.1.2	2.1.3	2.2.1	2.2.2	2.2.3	2.3.1	2.3.2	3.1.1	3.1.2	3.2.1	3.2.2	3.2.3	3.3.1	3.3.2	3.3.3	4.1.1	4.1.2	4.2.1	4.2.2	4.2.3	4.3.1	4.3.2	4.3.3
1.1.1	1.00																															
1.1.2	0.23	1.00																														
1.1.3	0.55	0.38	1.00																													
1.2.1	0.52	0.57	0.74	1.00																												
1.2.2	0.63	0.48	0.68	0.78	1.00																											
1.2.3	0.34	0.47	0.56	0.63	0.62	1.00																										
1.3.1	0.35	0.38	0.52	0.48	0.52	0.56	1.00																									
1.3.2	0.45	0.51	0.76	0.59	0.66	0.61	0.39	1.00																								
2.1.1	0.68	0.14	0.52	0.50	0.59	0.25	0.29	0.45	1.00																							
2.1.2	0.24	0.43	0.44	0.46	0.46	0.44	0.41	0.51	0.35	1.00																						
2.1.3	0.35	0.15	0.19	0.14	0.39	0.32	0.26	0.32	0.27	0.19	1.00																					
2.2.1	0.70	0.03	0.52	0.49	0.62	0.36	0.35	0.39	0.79	0.28	0.53	1.00																				
2.2.2	-0.06	-0.08	-0.23	-0.21	-0.27	-0.33	-0.31	-0.18	0.16	-0.03	-0.21	-0.03	1.00																			
2.2.3	0.72	0.29	0.47	0.53	0.69	0.40	0.40	0.46	0.67	0.42	0.44	0.82	0.15	1.00																		
2.3.1	0.58	0.36	0.62	0.71	0.64	0.53	0.33	0.61	0.50	0.27	0.24	0.63	-0.09	0.60	1.00																	
2.3.2	0.58	0.48	0.83	0.76	0.67	0.65	0.46	0.72	0.48	0.52	0.16	0.54	-0.12	0.60	0.69	1.00																
3.1.1	0.38	0.28	0.33	0.54	0.56	0.25	0.03	0.34	0.55	0.32	0.15	0.45	0.19	0.50	0.49	0.37	1.00															
3.1.2	0.39	0.37	0.26	0.56	0.56	0.26	0.04	0.35	0.56	0.35	0.20	0.42	0.25	0.53	0.50	0.32	0.89	1.00														
3.2.1	0.34	0.37	0.38	0.58	0.51	0.17	0.02	0.39	0.50	0.38	0.16	0.43	0.22	0.50	0.48	0.45	0.72	0.77	1.00													
3.2.2	0.61	0.42	0.77	**0.94**	0.79	0.58	0.43	0.61	0.64	0.42	0.20	0.67	-0.24	0.62	0.71	0.76	0.54	0.52	0.55	1.00												
3.2.3	0.27	0.64	0.52	0.66	0.57	0.57	0.45	0.56	0.45	0.53	0.08	0.31	0.07	0.38	0.47	0.55	0.36	0.42	0.36	0.60	1.00											
3.3.1	0.73	0.20	0.71	0.60	0.77	0.52	0.44	0.61	0.76	0.40	0.40	0.88	-0.13	0.80	0.66	0.69	0.45	0.41	0.41	0.75	0.47	1.00										
3.3.2	-0.02	0.42	0.37	0.55	0.35	0.55	0.22	0.22	0.02	0.25	-0.24	-0.05	-0.06	-0.01	0.31	0.38	0.23	0.28	0.23	0.37	0.55	0.13	1.00									
3.3.3	0.26	0.03	0.34	0.35	0.40	0.46	0.24	0.20	0.34	0.12	-0.05	0.35	-0.10	0.24	0.28	0.35	0.13	0.12	0.00	0.40	0.39	0.46	0.55	1.00								
4.1.1	0.39	0.69	0.60	0.76	0.69	0.76	0.36	0.61	0.16	0.41	0.21	0.32	-0.22	0.50	0.65	0.76	0.42	0.42	0.41	0.67	0.58	0.50	0.57	0.31	1.00							
4.1.2	0.55	0.37	0.39	0.57	0.68	0.43	0.21	0.44	0.63	0.39	0.29	0.57	0.18	0.66	0.58	0.47	0.76	0.82	0.64	0.56	0.46	0.60	0.29	0.26	0.52	1.00						
4.2.1	0.21	-0.18	0.10	0.06	0.07	0.26	0.06	0.01	-0.01	-0.10	0.33	0.34	-0.10	0.28	0.29	0.12	-0.12	-0.15	-0.13	0.09	-0.03	0.25	-0.06	0.06	0.20	-0.03	1.00					
4.2.2	0.38	0.47	0.43	0.68	0.72	0.56	0.39	0.40	0.26	0.49	0.16	0.39	-0.30	0.59	0.48	0.62	0.37	0.36	0.40	0.63	0.47	0.55	0.35	0.27	0.72	0.45	0.15	1.00				
4.2.3	0.34	0.14	0.03	0.11	0.20	-0.09	-0.09	0.21	0.15	0.05	0.29	0.17	0.16	0.37	0.21	0.05	0.27	0.35	0.34	0.10	-0.04	0.10	-0.19	-0.29	0.13	0.32	0.16	0.15	1.00			
4.3.1	0.29	0.35	0.18	0.24	0.51	0.53	0.21	0.39	0.11	0.24	0.44	0.18	-0.26	0.32	0.25	0.21	0.18	0.28	0.04	0.23	0.27	0.31	0.13	0.15	0.48	0.38	0.21	0.30	0.22	1.00		
4.3.2	0.39	0.19	0.33	0.35	0.42	0.47	0.20	0.40	0.30	0.23	0.24	0.48	-0.03	0.56	0.43	0.35	0.18	0.22	0.20	0.44	0.35	0.52	0.13	0.24	0.49	0.36	0.58	0.37	0.28	0.49	1.00	
4.3.3	0.18	-0.15	0.16	0.01	-0.03	0.05	0.13	0.03	0.29	0.03	-0.19	0.17	0.11	0.12	-0.05	0.09	-0.08	-0.11	-0.07	0.10	0.13	0.24	0.05	0.43	-0.15	-0.09	-0.05	-0.09	-0.18	-0.19	0.11	1.00

Source: Own

The Pearson correlation matrix exhibits a very dense pattern of high positive correlations, both within and across conceptual dimensions. Many coefficients are in the range r≈0.60-0.90, and one pair shows high correlation exceeding r=0.90, such as co-publications indicators 1.2.1 and 3.2.2 (with r=0.94), other correlated pairs above r=0.80 are the indicators related to SME introducing innovations 3.1.1 and 3.1.2 (r=0.89) and R&D expenditure in business sector and patents (r=0.88). The almost complete absence of strong negative correlations (and the explicit absence of “highly negative” coefficients) confirms that the problem is not offsetting relationships but rather overlapping positive information across indicators.

The systematic and statistically significant nature of these correlations, where many are significant at the 0.001 level, suggests a highly interconnected innovation system in which resource endowments, capabilities, and performance outputs co-evolve. Indicators belonging to the same conceptual group ex. human capital and skills (1.1.1, 1.1.2, 1.1.3, 1.2.3, 1.3.2, 2.3.2, 3.2.3, 4.1.1, 4.1.2) tend to demonstrate particularly high levels of correlation often with r between 0.50-0.80. Beyond within-group cohesion, the matrix shows substantial cross-dimensional coupling, for example, 1.2.2 (scientific publications) correlates strongly with 4.2.2 (knowledge-intensive services exports) at r=0.72. The SPSS output confirms explicitly that no highly negative correlations were identified across the dataset. Negative correlations that do appear are very small (typically r≈-0.01 to -0.33), sporadic, and not systematic across dimensions. This suggests that none of the EIS indicators act as counter-performing metrics, instead, higher performance in one area generally corresponds to higher performance in others.

Table 17 displays a correlation heatmap as well for all EIS 2024 indicators. Darker red colours indicate strong positive correlations, whereas blue shades would indicate negative relationships. The heatmap reveals a pronounced block structure along the main diagonal: indicators belonging to the same conceptual dimension (e.g. human capital, firm investments, intellectual assets) form dense red blocks, reflecting strong within-group associations. Off-diagonal red blocks capture cross-dimensional couplings, such as the close alignment between human capital, intellectual assets and innovation sales. The virtual absence of blue cells confirms that there are no substantial negative correlations in the dataset.

This analysis of the EIS dataset shows substantial multicollinearity, with many indicators correlated at r>0.70 and some exceeding r=0.90. This redundancy is not unexpected: several authors have criticised composite innovation indicators for insufficient attention to indicator independence. Grupp and Mogee (2004) point out that innovation metrics often capture overlapping dimensions of innovation capacity while Archibugi and Coco (2005) argue that composite indices risk rewarding bundles of structurally related inputs. Such overlaps in the construction of SII cause double counting of structurally similar indicators, inflation of composite scores for already advantaged countries, and reduced discriminative power of the index. These shortcomings confirm critiques who caution that indicator redundancy can compromise the interpretability of innovation scoreboards.

The presence of strong correlations among indicators implies that the effective dimensionality of the dataset is substantially lower than the nominal number of observed variables. A limited set of latent factors, such as human capital and skills, innovation investment, intellectual assets and outputs, and economic impacts, likely accounts for a large share of the total variance. This correlation structure provides a strong empirical justification for the application of dimensionality-reduction techniques, including PCA, and feature selection methods such as MCFS. These approaches mitigate multicollinearity, enhance model parsimony, and yield clearer structural interpretation compared to including a large set of highly correlated indicators directly in regression models.

The observed correlation structure has direct implications for regression analyses that employ EIS indicators as explanatory variables for NIS performance or broader macroeconomic outcomes, such as GDP per capita or productivity. In the presence of strong multicollinearity, regression models that include a large number of raw indicators may still exhibit satisfactory predictive performance, as reflected in high coefficients of determination or acceptable out-of-sample accuracy. However, the interpretation of individual regression coefficients becomes problematic. Parameter estimates are highly sensitive to model specification, often accompanied by inflated standard errors, and do not support reliable inference on the marginal contribution of individual indicators. Instead, the data are informative about broader bundles of innovation-related capabilities rather than isolated measures.

In both cross-sectional and panel regressions, including the full set of EIS indicators may result in over-parameterised models relative to the effective degrees of freedom, undermining statistical efficiency and robustness. Given these challenges, econometrically robust strategies include the use of composite indices in place of large sets of individual indicators, which directly mitigates multicollinearity and aligns with the latent structure revealed by the correlation analysis. Pillar-wise modelling approaches could be applied, as well, whereby only one representative indicator or factor score per conceptual dimension is included, ensuring that regressors remain both conceptually meaningful and statistically identifiable. The correlation analysis confirms that the EIS indicators function as a tightly interrelated system. Consequently, econometric analysis should account for this structure either through the construction of higher-level indices, as implemented with the IRPI and IEI in subsequent sections, or through explicit modelling of latent factors, rather than interpreting individual indicators in isolation.

Furthermore, the EIS relies on a broad set of 32 indicators spanning human capital, R&D investment, innovation activities, knowledge creation, and socio-economic effects. A detailed assessment of the indicator structure shows that the EIS is heavily input-oriented, it includes 14 input indicators, 12 output indicators, 3 outcome indicators, and 3 impact indicators. Looking from different stages perspective, knowledge/technological production and commercialisation, this shows quite a focus on the production (inputs-outputs) and less on commercial results (outputs-outcomes). This structural imbalance has been noted by several scholars. For instance, EC in EIS report 2020 emphasises that the EIS predominantly reflects countries' innovation capacities and not innovation performance, while Godin (2003) argues that European innovation statistics historically overrepresent R&D and human capital inputs at the expense of market and societal outcomes. As a result, the EIS rewards structural conditions more than the translation of innovation into economic or societal value.

Archibugi et al. (2009) note that composite indices often struggle to reflect real innovation performance, which depends more on market adoption and diffusion than on structural capabilities. Castellacci and Natera (2013) demonstrate empirically that outcomes and impacts behave differently from inputs, justifying their separate treatment.

The comparative analysis highlights that while the EIS provides a broad descriptive overview of innovation conditions, it suffers from structural and methodological limitations including input dominance and indicator redundancy.

Therefore, in this research we construct the new innovation result-based index, IRPI, that addresses these shortcomings by reducing the indicator set to nine (9) empirically selected variables using multi cluster feature selection, focusing exclusively on innovation outputs and outcomes, and creating two distinct subindices aligned with the two stages of the innovation process. The IRPI is further complemented by the IEI, an innovation efficiency-based index built using two-stage DEA and only 16 indicators (7 inputs, 7 outputs and 2 outcomes). Together, IRPI and IEI address the main conceptual and statistical limitations of existing composite indices.

6.2 SELECTION OF INDICATORS

MCFS is applied independently to each year, then feature importance scores are aggregated across all years. The following key parameters and choices were set:

- The clustering method uses K-Means clustering with n_init=10 with a set random seed for reproducibility;
- Optimal cluster selection is determined via silhouette score maximization (testing k=2 to 10 clusters). For each year, the k with the highest silhouette score is selected.
- Feature scoring mechanism: LASSO regression with alpha=0.001 (regularisation parameter). For each cluster, a binary classification target is created (1 if observation belongs to cluster, 0 otherwise), then LASSO coefficients are summed across all clusters to produce feature importance scores.
- Feature scores are computed independently for each year (2017-2024), then summed across years. This ensures features are selected based on consistent discriminatory power over time, not just a single cross-section.
- Final feature selection threshold includes two methods that are combined: Elbow method identifies the first large drop in sorted importance scores, and Cumulative contribution selects minimum features needed to reach 90% of total importance. The smaller of these two values determines the final number of features selected.
- Cumulative scores are normalised by dividing by the maximum score to avoid scale effects.
- The code processes indicators by subdimension (grouping related EIS indicators), applying MCFS within each conceptual group rather than to all 32 indicators simultaneously.

As a result of the MCFS, a new smaller set of indicators is identified from the full dataset of 32 EIS indicators covering 27 EU countries for the period 2017-2024. The MCFS analysis produced a ranked list of 16 statistically relevant indicators. These were subsequently evaluated for conceptual relevance, stage alignment, and redundancy.

The MCFS analysis adopts the indicator definitions and statistics from the EIS 2024 to ensure comparability. All indicators were included in the feature selection except impact indicators (4.3.1, 4.3.2, 4.3.3), as the study focuses on inputs and outputs and assessing impact lies beyond its scope; 3.2.3 (job-to-job mobility) due to missing values; and 3.2.2 (public-private co-publications) was excluded because it represents a subset of indicator 1.2.1 (international scientific co-publications), capturing narrower institutional collaboration rather than broader scientific output. Retaining both would introduce near-deterministic redundancy (r=0.94) that could distort cluster structures. Table 18 presents the indicators, their categorisation into inputs, outputs, outcomes, and impacts, and the subset selected through MCFS.

Table 18: EIS Indicators list and selected indicators with MCFS

Code	Indicator	Category	Selected by MCFS
1.1.1	New doctorate graduates in science, technology, engineering, and mathematics (STEM) per 1000	Input	YES
1.1.2	Percentage of population aged 25-34 having completed tertiary education	Input	YES
1.1.3	Percentage population aged 25-64 participating in lifelong learning	Input	NO
1.2.1	International scientific co-publications per million population	Output	NO
1.2.2	Scientific publications among the top 10% most cited publications worldwide as percentage of total scientific publications of the country	Output	YES
1.2.3	Foreign doctorate students as a percentage of all doctorate students	Input	NO
1.3.1	Broadband penetration	Input	YES
1.3.2	Individuals who have above basic overall digital skills (% share)	Input	NO
2.1.1	R&D expenditure in the public sector (percentage of GDP)	Input	NO
2.1.2	Venture capital expenditures (percentage of GDP)	Output	YES
2.1.3	Direct government funding and government tax support for business R&D (percentage of GDP)	Input	YES
2.2.1	R&D expenditure in the business sector (percentage of GDP)	Input	YES
2.2.2	Non-R&D innovation expenditures (percentage of turnover)	Input	YES
2.2.3	Innovation expenditures per person employed	Input	YES
2.3.1	Enterprises providing training to develop or upgrade ICT skills of their personnel	Input	NO
2.3.2	ICT specialists (as a percentage of total employment)	Input	NO
3.1.1	SMEs introducing product innovations (percentage of SMEs)	Output	YES

Code	Indicator	Categor y	Selected by MCFS
3.1.2	SMEs introducing business process innovations (percentage of SMEs)	Output	YES
3.2.1	Innovative SMEs collaborating with others (percentage of SMEs)	Output	YES
3.2.2	Public-private co-publications per million population	Output	NO
3.2.3	Job-to-job mobility of Human Resources in Science & Technology	Input	NO, excluded
3.3.1	PCT patent applications per billion GDP (in PPS)	Output	NO
3.3.2	Trademark applications per billion GDP (in PPS)	Output	NO
3.3.3	Design applications per billion GDP (in PPS)	Output	YES
4.1.1	Employment in knowledge-intensive activities (percentage of total employment)	Output	NO
4.1.2	Employment in innovative enterprises	Output	YES
4.2.1	Exports of medium and high technology products as a share of total product exports	Outcome	YES
4.2.2	Knowledge-intensive services exports as percentage of total services exports	Outcome	YES
4.2.3	Sales of new-to-market and new-to-enterprise innovations as percentage of turnover	Outcome	NO
4.3.1	Resource productivity	Impact	NO, excluded
4.3.2	Air emissions by fine particulate matter (PM2.5) in Industry	Impact	NO, excluded
4.3.3	Development of environment-related technologies, percentage of all technologies	Impact	NO, excluded

Source: Own analysis and Hollanders et al., 2024

Inputs

Input indicators are intended to capture the capacity and resources that enable innovation activities across countries. Thirteen input indicators initially captured education and human capital (e.g., STEM doctorate graduates, tertiary education, lifelong learning), digital infrastructure and skills (broadband, digital literacy, ICT training, ICT specialists), R&D funding (public and business expenditures, tax support). MCFS retained *seven*: STEM doctorate graduates, graduates from tertiary education, broadband penetration, direct government support for business R&D, business R&D expenditure, non-R&D innovation expenditures and innovation expenditures per person employed. The retained features emphasise countries' financial, digital and human investments in innovation, suggesting that these indicators were more effective in distinguishing clusters of countries over time. This reflects their strong correlation with innovation system performance and their ability to capture systemic innovation capacity differences across countries. More behavioural or enabling indicators like digital skills and ICT employment may have shown lower variability or discriminative power in relation to innovation clustering. Public sector R&D funding (2.1.1) was not selected by MCFS, suggesting it contributes limited additional discriminatory power once business-side investment measures are included. This is consistent with the high correlation among financing indicators, which implies overlapping information content.

Outputs

Output indicators are conceptualised as outcomes of the knowledge production stage and, simultaneously, as inputs into the commercialisation stage of the innovation process. While the EIS does not explicitly distinguish between these two stages, this study adopts a two-phase innovation framework, consistent with the innovation systems and knowledge production literature, and categorises indicators accordingly.

Initially, eleven output indicators were considered, capturing diverse aspects of research output and knowledge-related activity. These included few indicators on scientific publications, SME innovation activities and collaboration, patents, trademarks, design applications, venture capital investments, and employment in innovative enterprises. Applying MCFS reduced this set to *seven* indicators: top 10% most cited scientific publications, venture capital expenditures, SMEs with product innovations, SMEs with process innovations, collaborating innovative SMEs, design applications, and employment in innovative enterprises. MCFS did not select PCT patent applications (3.3.1), indicating that this indicator provides limited additional discriminatory power for distinguishing country clusters once other firm-level innovation indicators are included. The high correlation with business R&D expenditure (r=0.88) is consistent with this finding, as both capture similar underlying innovation intensity. In the MCFS framework, which prioritises features that preserve the intrinsic cluster structure of the data, such redundancy reduces the marginal contribution of patents to distinguishing country groupings. Consequently, these indicators provide limited additional discriminatory power once other commercialisation-related outputs are included.

Venture capital expenditures (2.1.2) is classified as an output rather than an input. This reflects the staging of innovation finance: venture capital typically enters at later Technology Readiness Levels (TRL 6 and beyond) when technological feasibility and market potential have been demonstrated, while early-stage research is predominantly funded by public grants (EIB, 2020). High venture capital activity therefore signals that a country's innovation system has successfully produced commercially viable outputs, representing market validation of knowledge production rather than a resource input that precedes it.

Outcomes

Three outcome indicators captured economic returns: exports of high-tech products, exports of knowledge-intensive services, and sales from innovation. MCFS retained two related to exports. These outcomes showed stronger discriminatory power across countries, highlighting diverse strategies in capitalising on innovation investments. The exclusion of Sales of new-to-market and new-to-enterprise innovations as percentage of turnover might suggest that this variable either overlaps with other export measures or that its variance across countries was not well aligned with clustering patterns. The retained indicators reflect both service-oriented knowledge economies and high-tech innovation outcomes, signalling distinct innovation models and economic impacts across national systems.

The correlation matrix for the selected indicators shows high but not extreme correlations with r over 0.90 that are almost perfect, and the structure is more coherent with the innovation-process logic. Overall, the feature selection step substantially reduces redundancy, although some systemic dependencies remain. The strongest correlations are between:

- SMEs introducing product (3.1.1) and SMEs introducing business process (3.1.2) innovations with r=0.89, and SMEs introducing business process (3.1.2) innovations and employment in innovative enterprises (4.1.2) with r=0.82; and
- Business R&D expenditure (2.2.1) and innovation expenditure per person employed (2.2.3) with r=0.82.

The removal of patents/trademarks and reducing the number of indicators for publications and human capital and skills to minimum representative indicators has compressed the covariance structure, reducing the risk of inflation in composite index models. The selected indicators represent a substantially improved, more statistically disciplined innovation measurement framework compared to the full EIS set. The remaining correlations reflect genuine functional relationships in the innovation system rather than redundancy.

The correlation of r=0.89 between SMEs introducing product innovations (3.1.1) and SMEs introducing business process innovations (3.1.2) requires explicit justification, given that indicator 1.2.1 (international co-publications) was removed due to near-perfect correlation with 3.2.2 (r=0.94).

Four considerations, methodological, statistical, conceptual, and empirical, support retaining both SME innovation indicators. *First*, correlation does not imply informational redundancy for cluster discrimination. High correlation indicates that two variables co-vary across observations but does not establish that they contain identical information for distinguishing between country clusters. Two indicators may be highly correlated in aggregate yet capture different structural aspects of innovation systems that matter for differentiating system types. MCFS selects features that jointly preserve the multi-cluster manifold structure of the data and collectively cover all cluster distinctions, rather than scoring features independently, thus explicitly addressing the limitation that correlated features may jointly carry discriminatory information missed by independent evaluation (Cai et al., 2010). Both 3.1.1 and 3.1.2 were retained by MCFS because each contributes discriminatory power to the multi-cluster structure of EU innovation systems that is not fully captured by the other. *Second*, the near-perfect correlation threshold (r=0.95) serves as a safeguard against deterministic relationships, not as the primary selection criterion. Following established guidelines in multivariate analysis, correlations above r=0.90 indicate very high association, but not near-perfect linear dependence where one variable can be almost entirely predicted from another (Hair et al., 2019; Tabachnick and Fidell, 2019). The correlation of r=0.89 between 3.1.1 and 3.1.2, while classified as "very high", remains below this threshold and leaves 20.8% of variance unexplained ($1-r^2=0.208$), representing substantial independent variation. This threshold is not a substitute for MCFS-based selection but a complementary check ensuring that retained indicators are not arithmetically redundant. *Third*, the distinction between product and process innovation is well-established in innovation measurement theory. The Oslo Manual (OECD, 2018) explicitly distinguishes product innovation, the introduction of new or significantly improved goods and services, from process innovation (termed "business process innovation" in the 2018 revision), the implementation of new or significantly improved production, delivery, or support processes. Empirical research confirms that these innovation types follow different determinants and trajectories: Reichstein and Salter (2006) demonstrate that product and process innovation respond differently to firm size, R&D intensity, and external knowledge sources. *Fourth*, country-level patterns reveal meaningful divergence between product and process innovation. While the indicators are correlated at r=0.89 across the pooled EU-27 sample, examination of country positions reveals cases where innovation systems exhibit asymmetric strengths. Countries may score highly on product innovation (reflecting strong technological capabilities and market orientation) while lagging on process innovation (reflecting weaker organisational modernisation or digitalisation), or vice versa. This asymmetry provides diagnostic value for identifying system-specific bottlenecks that would be obscured if only one indicator were retained.

The retention of both indicators is therefore grounded in the data-driven MCFS methodology, supported by established correlation thresholds, and consistent with the Oslo Manual framework for innovation measurement. No correlation among the nine IRPI indicators exceeds the near-perfect threshold (r>0.95) that warranted indicator removal, and MCFS retained all selected indicators based on their contribution to cluster discrimination. This approach directly addresses the methodological critique in Chapter 2 regarding ad hoc indicator selection in existing frameworks, while ensuring that the final indicator set captures the conceptually and empirically distinct dimensions of firm-level innovation activity.

6.3 PERFORMANCE PATTERNS ACROSS NATIONAL INNOVATION SYSTEMS

This section addresses RQa (What patterns of variation exist in innovation results?). The IRPI results for EU-27 countries over 2017-2024 reveal the empirical landscape of result-based innovation performance.

To address the SII weaknesses identified in Chapter 2, specifically, the conflation of inputs with outcomes and the resulting inability to distinguish capacity from performance, the IRPI is designed as a result-based composite index that focuses exclusively on outputs and outcomes of NIS. IRPI does not incorporate inputs or structural conditions representing the innovation capacity, which are instead evaluated through the complementary IEI. This dual-index approach enables a conceptually coherent and empirically robust distinction between the production of innovation results and the efficiency with which countries transform inputs into those results. This structure ensures that the IRPI captures both knowledge outputs and final commercialisation outcomes, reflecting the multi-stage nature of innovation processes emphasised in the systems literature (Edquist, 2011; Kline and Rosenberg, 1986) while accommodating the interactive learning dynamics that Lundvall (1992, 2007) identifies as cumulative across stages.

All input indicators, public and business R&D expenditure, human capital measures, ICT specialists, government support for R&D, are deliberately excluded from the IRPI. Their presence would artificially inflate correlations with outputs and bias the index toward countries with strong structural preconditions, rather than countries that generate strong performance outcomes. This removal directly addresses the identified issue of input dominance in the EIS (Grupp and Mogee, 2004). Inputs are incorporated instead into the IEI, a two-stage DEA model that evaluates:

- Stage 1: Efficiency of Knowledge/technological production, and
- Stage 2: Efficiency of Knowledge/technological commercialisation,

using the IRPI indicators as outputs. This separation ensures conceptual clarity between performance levels (IRPI) and efficiency levels (IEI).

To reflect the sequential nature of innovation processes, the IRPI is composed of two equally weighted subindices:

- Stage 1 subindex: Index for Knowledge/technological Production (IKP) that captures systemic ability to generate new knowledge and intellectual assets.
- Stage 2 subindex: Index for Knowledge/technological Commercialisation (IKC) that captures the extent to which innovation outputs translate into market and employment outcomes.

All indicators are used from EIS for consistency and comparation, they have been normalised using min-max rescaling to the (0,1) interval, based on the observed minimum and maximum across all EU countries and all years (excluding outliers). The aggregation of selected indicators into subindices and the final IRPI score follows a simple arithmetic mean, which avoids overweighting strongly correlated indicators and ensures transparency. The nine output and outcome indicators included in IRPI are 7 outputs: 1.2.2. Scientific publications among the top 10% most cited publications worldwide as percentage of total scientific publications of the country, 2.1.2. Venture capital expenditures (percentage of GDP), 3.1.1. SMEs introducing product innovations (percentage of SMEs), 3.1.2. SMEs introducing business process innovations (percentage of SMEs), 3.2.1. Innovative SMEs collaborating with others (percentage of SMEs), 3.3.3. Design applications per billion GDP (in PPS) and 4.1.2. Employment in innovative enterprises, and 2 outcomes: 4.2.1. Exports of medium and high technology products as a share of total product exports and 4.2.2. Knowledge-intensive services exports as percentage of total services exports. No correlation among the nine IRPI indicators exceeds the near-perfect threshold (r>0.95) applied to remove deterministically related indicators, and MCFS retained all selected indicators based on their contribution to cluster discrimination rather than correlation thresholds alone (Hair et al., 2019).

The IRPI is constructed as the arithmetic mean of two result-based subindices: IKP and IKC. IRPI therefore measures what a country produces and commercialises in terms of innovation output, not what it invests. All values are normalised to a 0-1 scale, where “1” is the highest performance observed in the dataset and “0” is the lowest performance observed in the dataset. The empirical results of IRPI, IKP and IKC for 2024 are presented in Table 19 below, while the results for 2017-2024 for each EU-27 country are presented in Annex 2.

Table 19: IRPI, IKP and IKC for EU-27 member countries, 2024

Country Code	IKP	IKC	IRPI	Country Code	IKP	IKC	IRPI
AT	0.66	0.59	0.63	**IE**	0.53	0.93	0.73
BE	0.76	0.69	0.73	**IT**	0.66	0.57	0.61
BG	0.28	0.45	0.36	**LT**	0.52	0.35	0.43
CY	0.73	0.71	0.72	**LU**	0.50	0.70	0.60
CZ	0.48	0.70	0.59	**LV**	0.25	0.45	0.35
DE	0.62	0.84	0.73	**MT**	0.34	0.60	0.47
DK	0.69	0.72	0.70	**NL**	0.68	0.72	0.70
EE	0.69	0.57	0.63	**PL**	0.32	0.52	0.42
EL	0.62	0.38	0.50	**PT**	0.45	0.41	0.43
ES	0.39	0.42	0.41	**RO**	0.11	0.60	0.35
FI	0.75	0.70	0.73	**SE**	0.76	0.78	0.77
FR	0.59	0.68	0.64	**SI**	0.49	0.61	0.55
HR	0.53	0.28	0.41	**SK**	0.25	0.66	0.46

Country Code	IKP	IKC	IRPI	Country Code	IKP	IKC	IRPI
HU	0.27	0.75	0.51	**EU-27**	0.51	0.61	0.56

Source: Own

IKP captures the breadth and depth of innovation outputs generated by a country, including scientific excellence (measured by top-cited publications, indicator 1.2.2), firm-level innovation activity (SME product innovations 3.1.1, SME process innovations 3.1.2, innovative SME collaboration 3.2.1, design applications 3.3.3, and venture capital 2.1.2), and innovation-related employment (employment in innovative enterprises, indicator 4.1.2). Over the period 2017-2024, the upper bound of IKP lies around 0.80-0.82, indicating that no country fully reaches a theoretical frontier, but several operate close to it. Finland consistently records the highest IKP values, rising from around 0.70 in 2017 to above 0.80 in 2022-2023, before declining to 0.75 in 2024. Sweden follows a similar trajectory, with IKP increasing steadily and reaching approximately 0.76 in 2024. Belgium, Denmark, Germany, Estonia, and Cyprus also show high and improving IKP, particularly after 2020, reflecting strengthened research output and firm-level innovation activity. At the lower end of the distribution, Romania records persistently very low IKP values (around 0.06-0.15), indicating weak scientific output (low top-cited publications 1.2.2), limited SME innovation activity (low values on indicators 3.1.1, 3.1.2, and 3.2.1), minimal design applications (3.3.3), constrained venture capital investment (2.1.2), and low employment in innovative enterprises (4.1.2). Slovakia, Hungary, Bulgaria, Latvia, Croatia, and Poland also occupy the lower tail, typically with IKP values below 0.18-0.35.

These patterns are consistent with structural weaknesses in innovation inputs and outputs, though the descriptive analysis cannot establish whether low scores reflect insufficient resources, ineffective policies, or other factors. Overall, IKP exhibits high temporal persistence, consistent with its interpretation as a structural capability shaped by long-term investments in education, research systems, and innovation ecosystems. These patterns are consistent with NIS theory's emphasis on cumulative, path-dependent capability formation (Nelson and Winter, 1982; Lundvall, 1992). The persistence of IKP rankings confirms that knowledge production capacity reflects long-term institutional investments rather than short-term policy interventions, a core premise of the systems approach (Freeman, 1987). Romania's persistently low IKP exemplifies what Radosevic (2017) discussed as a lagging CEE innovation system, characterised by limited absorptive capacity, weak university-industry linkages, and insufficient critical mass in research infrastructure.

IKC reflects the market uptake and economic value generated from innovation outputs, measured through exports of medium and high-technology products (indicator 4.2.1) and knowledge-intensive services exports (indicator 4.2.2). Note that sales of new-to-market innovations (4.2.3) was excluded by MCFS due to limited discriminatory power across country clusters. In contrast to IKP, IKC displays greater cross-country and temporal variability, with observed values ranging from 0.24 to 0.93. The highest IKC levels are consistently recorded by Ireland, which reaches values between 0.87-0.93 in most years, driven by strong performance in knowledge-intensive services exports and the commercial exploitation of innovation. Germany, Cyprus, Sweden, Finland, Luxembourg, and the Netherlands also exhibit high IKC (typically 0.70-0.85), indicating mature diffusion systems and strong integration into international markets. Repeatedly low IKC scores are observed for Bulgaria, Lithuania, Greece, and Croatia, often below 0.35. These values point to weak innovation adoption, limited scale-up and diffusion, and low penetration of innovative goods and services in domestic and export markets. Compared with IKP, IKC responds more strongly to economic cycles, sectoral specialisation, export demand, and structural shocks, explaining its higher volatility over time.

IRPI, defined as the average of IKP and IKC, captures balanced performance across both stages of the innovation process. Across the dataset, values approaching 0.75-0.80 define the empirical innovation frontier. Countries consistently operating close to this frontier include Sweden, Finland, Germany, Ireland, Denmark, Belgium, and, in peak years, Cyprus. These countries combine strong knowledge production with effective commercialisation, albeit through different structural profiles (e.g. high IKC dominance in Ireland versus more balanced profiles in Sweden and Finland). The lowest IRPI values, typically in the 0.28-0.35 range, are observed for Romania, Bulgaria, Latvia, Lithuania and Croatia, reflecting simultaneous weaknesses in both knowledge production and commercialisation. Greece, Poland, Slovakia, and Hungary also remain persistently below the EU-27 average, although some show gradual improvement over time. Countries with mid-range IRPI values (approximately 0.45-0.60) include Austria, France, Italy, Portugal, Spain, Czechia, and Slovenia. These innovation systems typically generate steady innovation outputs but differ substantially in their ability to translate them into market outcomes, resulting in heterogeneous IKP and IKC balances. Across all country and year observations, the lowest values observed on the scale are:

- IKP minimum ≈ 0.06-0.15 (Romania),
- IKC minimum ≈ 0.24-0.30 (Croatia and Lithuania),
- IRPI minimum ≈ 0.28-0.30 (Croatia, Bulgaria, Romania).

These values define the performance floor of innovation results in the EU over the period studied.

In 2024, IKP values range from 0.11 to 0.76, highlighting substantial heterogeneity in the knowledge production capacity of EU innovation systems. Sweden (0.76) and Belgium (0.76) record the highest IKP levels, closely followed by Finland (0.75) and Cyprus (0.73). These countries represent the knowledge-production frontier, characterised by strong research outputs, high SME innovation activity, and robust innovation-related employment. A group of countries including Denmark (0.69), Estonia (0.69), Netherlands (0.68), Austria (0.66), Italy (0.66), and Germany (0.62) also display solid knowledge production capacity, indicating mature innovation ecosystems with sustained research and innovation outputs. At the lower end of the distribution, Romania (0.11) exhibits the weakest IKP performance, reflecting severe structural constraints in innovation output generation. Bulgaria (0.28), Hungary (0.27), Latvia (0.25), and Slovakia (0.25), also record low IKP values, signalling limited technological creation, weak firm-level innovation activity, and low integration into innovation networks. Overall, the 2024 IKP distribution confirms that knowledge production is persistent and structurally uneven dimension of NIS performance across the EU.

IKC displays a wider range and stronger dispersion than IKP in 2024, with values spanning from 0.28 to 0.93, underscoring the sensitivity of commercialisation outcomes to market conditions, sectoral structure, and international integration. Ireland (0.93) stands out as a clear outlier at the top of the distribution, driven by exceptional performance in knowledge-intensive services exports and innovation-driven market outcomes. Germany (0.84) and Sweden (0.78) follow, indicating strong diffusion and absorption of innovation into competitive markets. High IKC scores are also observed for Hungary (0.75), Denmark (0.72), Netherlands (0.72), Cyprus (0.71), Finland (0.70), Luxembourg (0.70), and Czechia (0.70), reflecting mature commercialisation channels and effective scale-up mechanisms. In contrast, Croatia (0.28), Lithuania (0.35) and Greece (0.38) register the lowest IKC values, pointing to weak innovation diffusion, limited high-tech export intensity, and constrained market uptake. Other countries with relatively low IKC include Portugal (0.41), Spain (0.42), Bulgaria (0.45) and Latvia (0.45). These patterns reinforce the view that commercialisation remains volatile and uneven stage of the innovation process across the EU.

IRPI, defined as the average of IKP and IKC, ranges from 0.35 to 0.77 in 2024 and provides a balanced assessment of innovation outcomes across both stages:

- The innovation frontier is formed by Sweden (0.77), Finland (0.73), Germany (0.73), Belgium (0.73), Ireland (0.73), Cyprus (0.72), Denmark (0.70) and Netherlands (0.70). These countries combine strong or very strong knowledge production with effective commercialisation, although the underlying balance differs markedly. For example, Ireland’s high IRPI is driven predominantly by exceptional IKC, whereas Sweden and Finland exhibit more balanced profiles.
- A second tier of strong performers includes France (0.64), Austria (0.63), Estonia (0.63), Italy (0.61) and Luxembourg (0.60). These countries demonstrate solid innovation outcomes but exhibit some imbalance between IKP and IKC that prevents them from fully converging to the frontier.
- The moderate performers, with IRPI values between roughly 0.45 and 0.59, include Czechia (0.59), Slovenia (0.55), Hungary (0.51), Greece (0.50), Malta

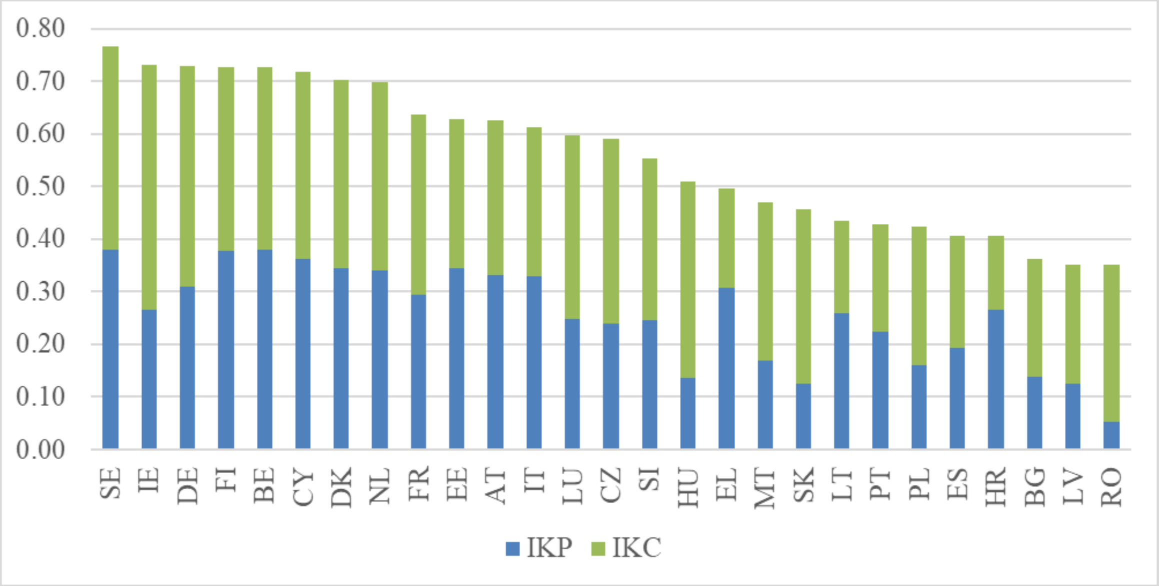
(0.47), and Slovakia (0.46). In this group, innovation systems often display either reasonable commercialisation with weaker knowledge production (e.g. Hungary) or the opposite pattern.

- The lowest IRPI values are recorded by Romania (0.35), Latvia (0.35), Bulgaria (0.36), Spain (0.41), Croatia (0.41), Poland (0.42), Lithuania (0.43), and Portugal (0.43). These countries fall into the group of weak performers, because they face persistent challenges in both stages of the innovation process, resulting in limited overall innovation results despite some improvements in individual dimensions.

The 2024 decomposition of IRPI into IKP and IKC highlights that similar overall innovation outcomes can arise from very different structural configurations. Some countries achieve high IRPI through balanced strength across both stages (e.g. Sweden, Finland), while others rely heavily on commercialisation despite weaker knowledge production (e.g. Ireland, Hungary). Conversely, several countries with moderate IKP fail to translate knowledge into economic value due to weak IKC (e.g. Greece, Croatia). Overall, the 2024 results reinforce the importance of a two-stage interpretation of innovation performance, consistent with the chain-linked model of innovation (Kline and Rosenberg, 1986) and its empirical operationalisation in network DEA (Guan and Chen, 2012). Policy bottlenecks differ markedly depending on whether weaknesses arise in knowledge generation, market diffusion, or the interaction between the two stages, a diagnostic distinction that aggregate indices cannot provide.

The decomposition chart in Figure 11 below clarifies why countries occupy particular positions in the ranking and whether they are strong producers, strong commercialisers, or weak in both stages. Each bar represents the total IRPI score for a country and is split into an IKP segment (knowledge production) and an IKC segment (commercialisation). This visualisation highlights the dual nature of innovation performance, identifies stage-specific bottlenecks, and illustrates that similar IRPI scores can arise from very different structural configurations of NIS.

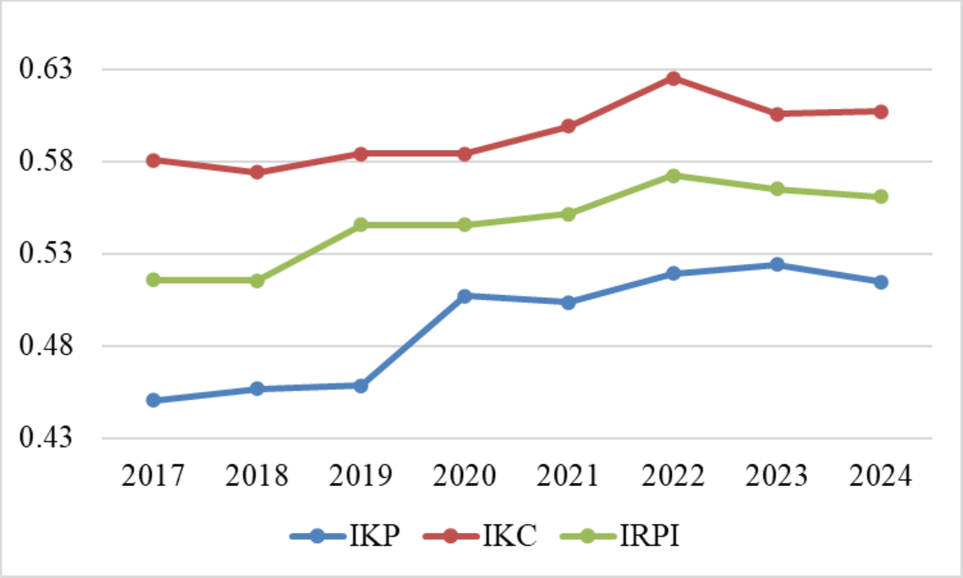
Figure 11: Decomposition of IRPI by country, 2024



Source: Own

Figure 12 presents the evolution of the IRPI and its two constituent subindices, IKP and IKC, for the EU average over the period 2017-2024. When interpreting IRPI, IKP, and IKC over time, it is essential to recognise that the underlying innovation indicators refer to year t, while the composite index, same as in SII, is calculated for year t+1. Accordingly, the increases observed in the indices published in 2021 correspond to COVID-19 period from 2020 when the COVID-19 pandemic resulted in an unprecedented economic contraction in 2020, with EU real GDP falling by 6.1% (De Santis and Stoevsky; 2023), more than during the global financial crisis. While the strong upward jump reported in 2022 reflects the post-pandemic rebound in 2021 when growth resumed forcefully in the EU thanks to strong EC liquidity support to business via SURE instrument. The downturn in the 2023 and 2024 index is attributable to weakened innovation outcomes in 2022 due to the highest peak with the pandemic (December 2021 - March 2022). Covid lockdown officially ended in EU in May 2022. Accounting for this timing convention clarifies that innovation-system dynamics respond to external shocks with both structural inertia (IKP) and cyclical sensitivity (IKC), and that IRPI captures these differentiated effects.

Figure 12: IRPI, IKP and IKC for EU average, 2017-2024



Source: Own

Table 20 assesses whether country rankings represent structural characteristics or crisis artefacts by examining rank correlations across periods.

Table 20: IRPI rank stability across crisis period

Comparison	Spearman	p-value	Interpretation
IRPI 2019 vs IRPI 2020	0.96	<0.001	Very high stability
IRPI 2020 vs IRPI 2021	0.99	<0.001	Very high stability through pandemic onset
IRPI 2021 vs IRPI 2022	0.97	<0.001	Very high stability through recovery
IRPI 2022 vs IRPI 2023	0.97	<0.001	Very high stability post-crisis
IRPI 2023 vs IRPI 2024	0.97	<0.001	Very high stability post-crisis
IRPI 2019 vs IRPI 2024	0.89	<0.001	High structural persistence

Source: Own

The Spearman correlation of 0.89 between the last pre-crisis index (2019, reflecting 2018 data) and the most recent index (2024, reflecting 2023 data) shows that country rankings exhibit high structural persistence across the entire crisis period. This finding supports the interpretation that the IRPI captures underlying innovation system characteristics rather than transient crisis effects.

The IRPI offers a conceptually coherent result-based innovation performance measure that focuses exclusively on outputs and outcomes, it is derived from data-driven feature selection, avoids the weaknesses of input-dominated frameworks like the EIS, incorporates a clear two-stage innovation process, and provides a result-based complement to the input-efficiency assessment of the IEI. Together, IRPI and IEI form a unified measurement system that captures both what innovation systems achieve and how efficiently they achieve it, providing a more nuanced and policy-relevant assessment of NIS performance. In Table 21 the key differences between SII and IRPI are presented.

Table 21: EIS weaknesses addressed by IRPI

EIS Weakness	Addressed by IRPI	Explanation
Input dominance	Yes	Inputs excluded; only outputs/outcomes retained.
Redundant indicators	Yes	MCFS-based reduction eliminates overlap.
Blurring of innovation stages	Yes	Subindices IKP and IKC map cleanly onto Stage 1 and Stage 2.
Excessive multicollinearity	Yes	Nearly perfect correlations removed; deterministic chains removed.
Overweighting of scientific publication indicators	Yes	Only one indicator related to publications (of total 3) kept in the outputs: top 10% most cited scientific publications
Weak representation of outcomes	Yes	IKC subindex focuses exclusively on commercialisation outcomes.
Lack of clarity in linking outputs to commercialisation	Yes	IRPI explicitly separates production and exploitation.

Source: Own

Table 21 demonstrates how each SII limitation identified in Chapter 2 is systematically addressed by the IRPI-IEI framework. The heavy reliance on composite indicators is retained for comparability but restructured: IRPI focuses on results only, while IEI captures transformation efficiency separately. Input dominance is eliminated by excluding all 7 input indicators from performance measurement. The limited use of efficiency analysis is addressed through comprehensive two-stage DEA applied to all 27 EU countries over 6 years. Weak process orientation is replaced by explicit modelling of sequential transformation (inputs into outputs into outcomes). Stage differentiation is achieved through both performance decomposition (IKP/IKC) and efficiency decomposition (KPEI/KCEI). The empirical results presented below provide evidence for whether this framework delivers on its conceptual promise.

6.4 STAGE-SPECIFIC EFFICIENCY DIFFERENCES

This section addresses RQb (Where do efficiency bottlenecks concentrate?). The IEI results, decomposed into KPEI (knowledge production efficiency) and KCEI (commercialisation efficiency), reveal systematic stage-specific patterns across EU-27 countries.

For calculation of the innovation efficiency scores per NIS per year, we apply a Variable Returns to Scale (VRS) output-oriented DEA. An output-oriented model was selected to evaluate how effectively EU countries maximise innovation outputs given their resources. In our output-oriented DEA, DMUs are NISs of the 27 EU countries, where NISk, k = 1, 2, …, 27, as for each NIS we perform DEA for 2017-2024 where each DMU in each year is treated as if it were a distinct DMU. The DEA considers the time lag of one year for inputs to transform in outputs, and one year for outputs to transform in outcomes, similar to the previous studies that apply two-stage DEA. The inputs from 2017 are analysed in the knowledge production efficiency for 2018, and the outputs from 2018 are used for the analysis of commercialisation efficiency in 2019. Below the two-Stage DEA efficiency analysis for 27 EU member countries are presented, based on a dataset for the period 2017-2024. However, the scores per stage and overall efficiency scores are calculated for the period 2019-2024 for all EU-27 countries, as there is a lag of two years. Therefore, there are no efficiency scores for 2017 and 2018. The variables (inputs, outputs, outcomes) used in the analysis were selected based on the results of the MCFS algorithm and the two subindices of IRPI: the output subindex (IKP) and the outcome subindex (IKC). The program used for calculating the relative efficiency scores is the Frontier Analyst 4.0.

In a serial network, the overall system efficiency is mathematically equal to the product of the component stage efficiencies, reflecting proportional propagation of inefficiency through the innovation pipeline. Therefore, given that:

- Stage 1: knowledge/technological production efficiency (KPEI) with value between 0 and 1.
- Stage 2: knowledge/technological commercialisation efficiency (KCEI) with value between 0 and 1.

the overall NIS efficiency IEI=KPEI×KCEI, that means a country is fully efficient only if both stages operate at the frontier and the weakness in either knowledge production or commercialisation reduces total system performance. IEI values are between 0 and 1, where 1.00 means full overall efficiency of NIS. IEI is showing how efficiently does the NIS transform original inputs into final innovation outcomes, taking both stages jointly as a production chain. The efficiency scores calculated with output-oriented DEA with VRS for the two-stages and time lag of one year are presented for each EU member country for 2019 (Stage 1 efficiency based on inputs from 2017, and output subindex IKP for 2018, while Stage 2 efficiency based on outputs for 2018 and outcome subindex IKC for 2019), 2020 (Stage 1 efficiency based on inputs from 2018, and output subindex IKP for 2019, while Stage 2 efficiency based on outputs for 2019 and outcome subindex IKC for 2020) until 2024. The efficiency scores for 2024 for every EU country are presented in Table 22. In Annex 3 the full table with IEI scores for the EU-27 countries for the period 2019-2024 is presented.

Table 22: Efficiency scores per country per stage and overall efficiency for EU-27 countries, 2024

Country code	KPEI	KCEI	IEI	Country code	KPEI	KCEI	IEI
AT	1.00	0.66	0.66	**IE**	0.80	1.00	0.80
BE	0.94	0.75	0.70	**IT**	1.00	0.68	0.68
BG	1.00	1.00	1.00	**LT**	0.72	0.48	0.34
CY	1.00	0.80	0.80	**LU**	1.00	0.86	0.86
CZ	0.91	1.00	0.91	**LV**	1.00	0.72	0.72
DE	0.99	0.94	0.94	**MT**	1.00	1.00	1.00
DK	0.91	0.79	0.72	**NL**	1.00	0.80	0.80
EE	1.00	0.66	0.66	**PL**	1.00	0.84	0.84
EL	1.00	0.54	0.54	**PT**	1.00	0.58	0.58
ES	0.59	0.58	0.34	**RO**	1.00	1.00	1.00
FI	1.00	0.76	0.76	**SE**	0.90	0.84	0.75
FR	0.78	0.81	0.63	**SI**	1.00	0.88	0.88
HR	1.00	1.00	1.00	**SK**	0.49	1.00	0.49
HU	1.00	1.00	1.00				

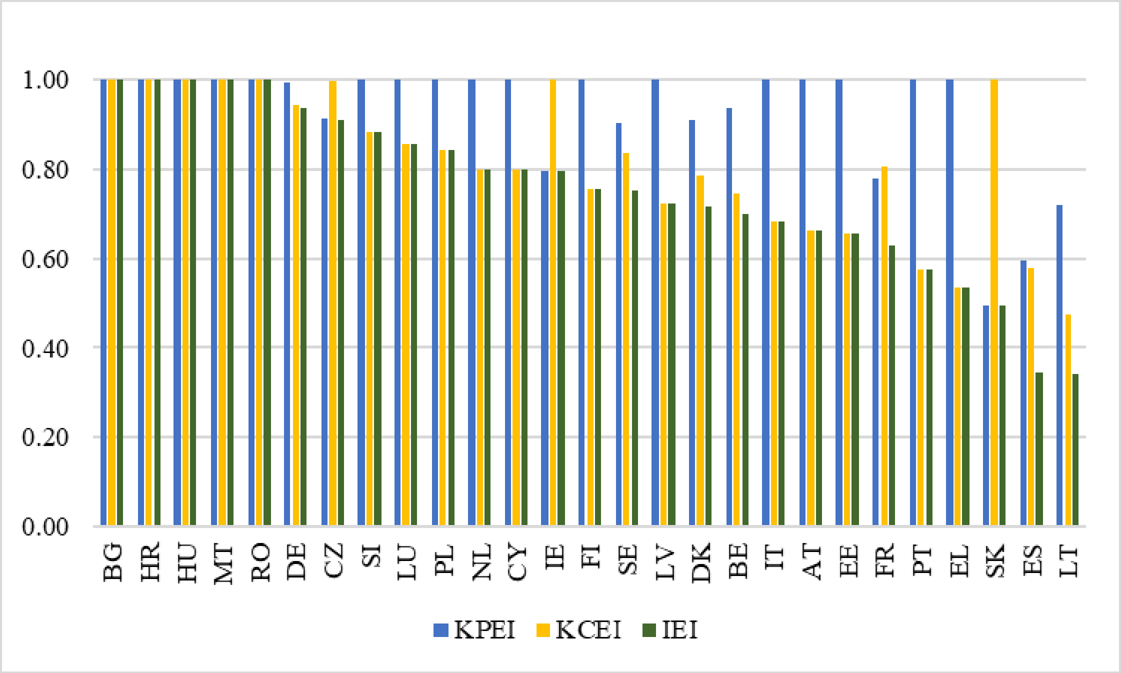
Source: Own

Innovation efficiency is not expected to shift considerably in the short term, as national policies, and mid-term strategies (e.g., R&D roadmaps, digital agendas, smart specialisation strategies) are stable. Innovation systems show structural inertia due to the long-term nature of R&D investments, human capital formation, and institutional learning. As Edquist (2011) notes, the effectiveness of innovation policy is “not instantaneous” but shaped by medium to long-term feedback loops. Consequently, national performance rankings tend to remain stable. The EIS 2024 (Hollanders et al., 2024) similarly observes that performance differences persist over time, with countries rarely shifting categories quickly. Over the last seven years, only a few countries (e.g., Estonia’s move to Strong Innovator) have changed tiers, and these shifts reflect sustained policy-driven growth rather than short-term fluctuations.

The IEI, defined multiplicatively as the product of KPEI and KCEI, displays substantial cross-country heterogeneity in 2024. Overall NIS efficiency ranges from 0.34 (in Lithuania and Spain) to full overall efficiency (IEI=1.00) in countries such as Bulgaria, Croatia, Hungary, Malta and Romania.

High IEI values characterise innovation systems that are simultaneously close to the best-practice frontier in both knowledge generation and commercialisation efficiency or at least do not exhibit a pronounced bottleneck in either stage. By contrast, IEI<1.00 values arise where inefficiency in one stage, most commonly commercialisation, strongly constrains overall system performance, even when knowledge production operates near the frontier. This pattern is evident in 2024 in countries such as AT, CY, EE, EL, FI, IT, LU, LV, NL, PL, PT, and SI, where there is full efficiency in Stage 1 but with weaknesses in Stage 2, leading to markedly lower overall efficiency. Figure 13 showcases overall innovation efficiency (IEI), knowledge production (KPEI) and commercialisation (KCEI) efficiency. Many EU member countries operate close to the frontier in knowledge production, and cross-country variation in overall efficiency is driven almost entirely by differences in commercialisation performance. A few countries such as Ireland, Slovakia and Czechia, have higher commercialisation efficiency compared to production efficiency. This highlights the central role of downstream innovation mechanisms in shaping NIS efficiency.

Figure 13: IEI, KPEI and KCEI for EU-27 countries, 2024



Source: Own

A first group of countries operates at or extremely close to the joint two-stage frontier in 2024, achieving IEI=1.00 or close to 1.00 through full overall efficiency in both knowledge production and commercialisation. In large and research-intensive systems, such frontier positions are consistent with strong research capacity and effective downstream translation; however, frontier efficiency should be interpreted as a relative measure of proportional input-output performance rather than as a direct indicator of innovation system maturity.

The coexistence of high innovation efficiency (IEI) and very low result-based performance (IRPI) in few EU countries reflects a structural property of DEA rather than an anomaly in the results. This distribution confirms a key theoretical prediction: efficiency and performance are analytically distinct dimensions that need not co-vary (Edquist, 2011). From a NIS perspective, efficiency captures the productivity of transformation processes, how well the system converts available resources into results, while performance captures the absolute level of results achieved. Key empirical finding: performance and efficiency are empirically independent dimensions.

The correlation between IRPI and IEI across 27 EU countries in 2024 is r = 0.03, effectively zero. This result is critical for the framework's validity: had the correlation been substantial (e.g., r>0.50), the analytical separation between performance and efficiency would have been empirically unjustified, and the dual-index architecture would have been redundant. The near-zero correlation confirms that the two dimensions capture genuinely distinct properties of NIS, validating the framework's foundational premise. This finding directly resolves the literature gap identified in Chapters 2 and 3: the outcome-oriented and efficiency-oriented streams, which bibliometric analysis showed to have developed in parallel with limited integration, capture genuinely distinct properties of NIS. Performance and efficiency are not merely conceptually separable, as they are empirically independent. This finding supports the argument advanced in subsequent NIS literature that system functioning cannot be reduced to input-output ratios alone (Edquist, 2011; building on Lundvall, 1992), and contradicts composite index approaches that conflate these dimensions into single scores.

As noted in Chapter 5, DEA efficiency scores are scale-sensitive, and in the two-stage DEA framework, efficiency is a relative, scale-dependent measure of input-output proportionality, not an indicator of absolute innovation strength. Given their exceptionally low innovation input levels, these countries are benchmarked against similarly resource-constrained peers and therefore attain frontier efficiency when outputs are commensurate with inputs. This classification indicates limited input waste, not strong or mature innovation systems. The consistently low IRPI values confirm that absolute innovation outcomes remain weak, underscoring that high IEI in this context signals small but proportionally efficient systems rather than innovation leaders. Accordingly, efficiency scores for low-input countries must be interpreted diagnostically and jointly with performance measures, as DEA efficiency rankings are sensitive to scale heterogeneity and do not provide a reliable standalone assessment of innovation success (Cooper et al., 2007; Edquist, 2011). Therefore, despite its scale sensitivity, the IEI is a valuable diagnostic tool for policymakers because it identifies how effectively existing innovation inputs are transformed into outputs and outcomes. In low-input countries, high IEI values indicate that weak result-based innovation performance is driven primarily by insufficient scale rather than by inefficient use of resources. This distinction may help inform policy priorities by suggesting whether weak performance is associated with insufficient scale or with transformation inefficiency. This diagnostic role is explicitly embedded in the IE_IRPI, where efficiency enters as a bounded adjustment to result-based performance, preserving outcome primacy and preventing high efficiency from compensating for weak innovation results.

A smaller group displays the inverse asymmetry, with commercialisation at or near the frontier but weaker knowledge production. CZ, IE and SK are fully efficient in Stage 2.

These cases suggest systems where applied innovation, adoption, or external technology integration compensate partly for limited domestic research capacity.

The 2024 results show that cross-country differences in overall innovation efficiency are driven primarily by variation in commercialisation efficiency, not by knowledge production efficiency. Second, asymmetric stage performance is the norm rather than the exception, revealing clearly identifiable bottlenecks within NIS. Third, while several cohesion economies are classified as efficient relative to their limited input base, a number of larger and research-intensive EU member countries continue to underperform in translating research strength into market outcomes. *Key empirical finding is that commercialisation is the binding constraint across the EU.* This systematic stage asymmetry validates the two-stage DEA methodology established in Chapter 5. The finding that 18 of 27 countries exhibit stronger knowledge production than commercialisation efficiency would be invisible in a single-stage model, confirming that the stage decomposition reveals diagnostic information that aggregate measures obscure.

These findings align with broader EU-level assessments and empirically resolve the lack of stage differentiation identified in Chapter 2. The stage-specific efficiency decomposition shows a measurable gap between knowledge production and commercialisation across the EU. This pattern is consistent with concerns about insufficient knowledge-to-market translation that have animated EU innovation policy debate (Pavitt, 1991; European Commission, 1995; Draghi, 2024), though the strength and interpretation of the so-called 'European Paradox' has been contested in the literature (Dosi et al., 2006). Mean KPEI (0.93) exceeds mean KCEI (0.81), with 18 of 27 countries showing stronger production than commercialisation efficiency. This finding extends the efficiency-oriented stream identified in Chapter 3 by providing systematic EU-wide evidence using consistent methodology.

As established in the temporal framework presented in Section 5.2, the analysis period 2017-2024 encompasses the COVID-19 pandemic and its aftermath, representing major exogenous shocks to European innovation systems. Because the DEA efficiency frontier is defined by best-performing units in each period, frontier position depends on the relative performance of comparator countries and can shift between periods due to technical change or asymmetric performance movements among comparator units (Färe et al., 1994; Zabala-Iturriagagoitia et al., 2021). Exogenous shocks such as economic crises can therefore alter both the frontier position and individual countries' relative efficiency rankings, making efficiency measures more sensitive to short-term disruptions than absolute performance measures. This section explicitly examines crisis effects on efficiency scores to distinguish structural patterns from temporary fluctuations.

The two-stage DEA model implements one-year lags per stage: Stage 1 transforms inputs (t) into outputs (t+1), and Stage 2 transforms outputs (t+1) into outcomes (t+2). Consequently, efficiency scores for IEI year t reflect a two-year data span. Table 23 maps each efficiency index to its underlying data years and the corresponding economic context.

Table 23: IEI timing and COVID-19 impact mapping

IEI year	Inputs (t)	Outputs (t+1)	Outcomes (t+2)	Economic context
2019	2017	2018	2019	Pre-pandemic baseline
2020	2018	2019	2020	Outcomes in pandemic onset (-6.1% GDP)
2021	2019	2020	2021	Outputs in pandemic; Outcomes in SURE recovery
2022	2020	2021	2022	Maximum crisis exposure (all components affected)
2023	2021	2022	2023	Recovery inputs; Omicron-affected outputs
2024	2022	2023	2024	Post-crisis throughout

Source: Own

This timing structure operationalises the methodological framework established in Section 5.2, where the three-period classification (pre-crisis, crisis, recovery) was defined. IEI 2019 provides a clean pre-pandemic baseline. IEI 2021 first incorporates pandemic-year outputs (2020). IEI 2022 represents maximum crisis exposure, incorporating pandemic inputs (2020), recovery outputs (2021), and outcomes from the Omicron peak period (December 2021-March 2022). Index 2024, with data spanning 2022-2024, provides a near-clean post-crisis reference point, as EU lockdowns officially ended in May 2022. Table 24 reports Spearman rank correlations for IEI rankings between adjacent years and across the full period.

Table 24: IEI Rank stability

Comparison	Spearman ρ	Interpretation
2019 vs 2020	0.79	High stability
2020 vs 2021	0.95	Very high stability
2021 vs 2022	0.92	High stability
2022 vs 2023	0.84	High stability
2023 vs 2024	0.70	Moderate stability (recovery-period propagation)
2019 vs 2024	0.68	Moderate structural persistence

Source: Own

The lowest adjacent-year correlation (ρ=0.70) occurs between IEI 2023 and IEI 2024. Under the two-year lag structure, this transition captures the propagation of 2021-2022 recovery-period inputs through to 2023-2024 outcomes, indicating that the largest reordering of efficiency rankings arose from lagged recovery effects rather than from peak-crisis outcomes themselves. The pre-crisis to post-crisis correlation (ρ=0.68) indicates moderate structural persistence, substantially below IRPI ranking persistence (ρ=0.89), confirming that efficiency measures are more sensitive to short-term economic and policy disruptions than result-based performance measures. This validates the conceptual framework's distinction between accumulated capabilities (performance) and transformation processes (efficiency). Table 25 compares efficiency scores between IEI 2019 (pre-pandemic) and IEI 2024 (post-crisis) to identify structural changes versus crisis artefacts.

Table 25: Pre-crisis vs Post-crisis efficiency scores comparison, EU-27

Country	IEI 2019	IEI 2024	Change	Classification
BG	1.00	1.00	0.00	Stable at frontier
HR	1.00	1.00	0.00	Stable at frontier
RO	1.00	1.00	0.00	Stable at frontier
HU	1.00	1.00	0.00	Stable at frontier
MT	0.91	1.00	+0.09	Improved to frontier
DE	0.93	0.94	+0.01	Stable high
LU	0.88	0.86	-0.02	Stable high
IE	0.85	0.80	-0.05	Stable high
NL	0.86	0.80	-0.06	Stable high
PL	0.80	0.84	+0.04	Stable
FI	0.78	0.76	-0.02	Stable
IT	0.77	0.68	-0.09	Moderate decline
SE	0.73	0.75	+0.02	Stable
DK	0.73	0.72	-0.01	Stable
BE	0.73	0.70	-0.03	Stable
AT	0.69	0.66	-0.03	Stable
LV	0.70	0.72	+0.02	Stable
FR	0.70	0.63	-0.07	Moderate decline
EE	0.59	0.66	+0.07	Slight improvement
PT	0.50	0.58	+0.08	Slight improvement
EL	0.51	0.54	+0.03	Stable low
ES	0.34	0.34	0.00	Stable low
LT	0.33	0.34	+0.01	Stable low
CZ	0.62	0.91	+0.29	Structural improvement
SI	0.62	0.88	+0.26	Structural improvement
CY	1.00	0.80	-0.20	Post-crisis normalisation
SK	1.00	0.49	-0.51	Structural decline

Source: Own

Twenty-three countries (85%) exhibit IEI changes of less than ±0.10 between pre-crisis and post-crisis periods, indicating that structural efficiency patterns predominate. Four countries show changes equal to or exceeding ±0.20, requiring specific interpretation. Two additional countries (Greece, Poland) exhibit small endpoint changes but pronounced within-crisis volatility, also requiring contextual interpretation. Each case is discussed below.

- Quantitatively, four countries (SK, CZ, SI, CY) account for over 90% of the squared rank-difference contributions to the 2019 vs 2024 Spearman correlation. When these cases are interpreted separately, the IEI exhibits substantially higher stability for the remaining 23 countries (ρ>0.95). This pattern indicates that the index reliably captures persistent structural features for most countries, with reduced reliability concentrated in cases where actual structural transitions occurred.

- *Czechia and Slovenia* improved KPEI substantially (CZ: 0.62 to 0.91; SI: 0.62 to 1.00), with gains materialising in IEI 2023-2024 (reflecting 2021-2023 data). These may represent structural improvements during the recovery period. While the timing coincides with EU recovery fund investments, the efficiency analysis cannot establish a causal link between policy interventions and efficiency gains.
- *Slovakia* experienced KPEI collapse (1.00 to 0.49) beginning in IEI 2020, which reflects 2018 inputs transforming into 2019 outputs, entirely pre-pandemic data. This is a structural break in knowledge production efficiency unrelated to COVID-19.
- *Cyprus* declined from frontier (1.00 to 0.80) in Index 2024, representing post-crisis normalisation after maintaining artificially high efficiency during the crisis period when the DEA frontier shifted.
- *Greece and Poland* exhibited transient DEA-frontier positions during crisis years (EL reaching IEI=1.00 in 2022-2023; PL reaching IEI=1.00 in 2021-2022) that subsequently reversed. Greece moved from IEI=0.51 (2019) through IEI=1.00 (2022-2023) back to IEI=0.54 (2024), and Poland moved from IEI=0.80 (2019) through IEI=1.00 (2021-2022) to IEI=0.84 (2024). Because their pre-crisis and post-crisis IEI scores are similar, these countries do not appear in Table 25 as cases of structural change. However, their crisis-period frontier positions reflect asymmetric pandemic shocks affecting comparator countries' inputs and outputs rather than absolute efficiency gains, consistent with the DEA-frontier shift mechanism discussed at the beginning of this section. Single-year IEI scores for these countries during 2021-2023 should be interpreted as crisis-period artefacts rather than evidence of sustained efficiency improvement.

Table 26 examines which countries reached the efficiency frontier (IEI=1.00) in each year to identify crisis-period frontier shifts.

Table 26: Countries at DEA frontier (IEI=1.00) by year

Country	2019	2020	2021	2022	2023	2024	Pattern
BG	✓	✓	✓	✓	✓	✓	Permanent frontier
HR	✓	✓	✓	✓	✓	✓	Permanent frontier
RO	✓	✓	✓	✓	✓	✓	Permanent frontier
HU	✓	✓	✓	-	✓	✓	Near permanent
CY	✓	✓	✓	✓	-	-	Crisis-period, post-crisis exit
SK	✓	-	-	-	-	-	Pre-crisis only
LV	-	✓	✓	✓	✓	-	Crisis-period frontier

Country	2019	2020	2021	2022	2023	2024	Pattern
DE	-	✓	✓	✓	-	-	Crisis-period frontier
MT	-	-	-	✓	✓	✓	Post-crisis frontier
PL	-	-	✓	✓	-	-	Crisis-only frontier
EL	-	-	-	✓	✓	-	Crisis-only frontier

Source: Own

Bulgaria, Croatia, and Romania remain at the frontier throughout the period, a DEA artefact reflecting their position as efficient relative to their low input levels rather than high absolute performance. Germany, Latvia, Poland, and Greece reached the frontier specifically during crisis years (2020-2023) but not in the pre-crisis baseline (2019) or post-crisis period (2024). These crisis-period frontier positions should not be over-interpreted as structural efficiency gains, as they may reflect temporary frontier shifts when other countries experienced asymmetric disruptions. Table 27 identifies countries exhibiting KCEI peaks during the crisis period (2021-2022) that subsequently reversed.

Table 27: Countries with crisis-period KCEI peaks

Country	KCEI 2019	KCEI Peak (Year)	KCEI 2024	Reversal
EL	0.51	1.00 (2022-2023)	0.54	Full reversal
LV	0.70	1.00 (2020-2023)	0.72	Full reversal
PL	0.80	1.00 (2021-2023)	0.84	Full reversal
NL	0.86	0.95 (2021)	0.80	Below baseline
SE	0.89	0.95 (2021)	0.84	Below baseline
DK	0.85	0.91 (2021-2022)	0.79	Below baseline
AT	0.69	0.80 (2022)	0.66	Below baseline

Source: Own

Seven countries (26%) exhibited KCEI peaks during 2021-2022 that subsequently declined. The timing is significant: KCEI captures pandemic-depressed outputs (2020-2021) transforming into recovery-boosted outcomes (2021-2022), when the SURE instrument supported demand recovery. Greece, Latvia, and Poland reached the KCEI frontier exclusively during crisis years, these represent crisis artefacts rather than structural efficiency gains and should not be over-interpreted as long-term commercialisation strengths. Based on the temporal analysis, Table 28 classifies the dissertation's efficiency findings by crisis sensitivity.

Table 28: Classification of efficiency findings

Classification	Findings	Confidence
Structural (high confidence)	BG, HR, RO, at frontier (all years); ES, LT stable low (all years); FI, IT, NL KPEI at frontier (all years); IE KCEI at frontier (all years); no similar Stage 1 and Stage 2 efficiency patterns; and IRPI and IEI keep weak correlation.	High
Structural change (not crisis-related)	SK KPEI collapse (pre-pandemic timing: 2018 to 2019 data); CZ, SI KPEI improvement (post-crisis timing: 2021-2023 data).	High

Classification	Findings	Confidence
Crisis-sensitive (interpret cautiously)	KCEI peaks for EL, LV, PL, NL, SE, DK, AT during 2021-2023; DE, LV, PL, EL frontier status limited to crisis years.	Medium
Post-crisis normalisation	CY decline from frontier in 2024.	High

Source: Own

The efficiency analysis suggests that structural patterns predominate despite the COVID-19 shock. Twenty-three countries (85%) show IEI changes below ±0.10 between pre-crisis IEI 2019 and post-crisis IEI 2024. The pre-crisis to post-crisis rank correlation ($\rho=0.68$) indicates moderate structural persistence, with the lowest adjacent-year correlation ($\rho=0.70$) occurring between IEI 2023 and IEI 2024. Under the two-year lag structure, this captures the propagation of 2021-2022 recovery-period inputs through to 2023-2024 outcomes, indicating that the largest reordering of efficiency rankings arose from lagged recovery effects rather than from crisis outcomes themselves.

The main efficiency findings, such as stage-specific patterns, the KPEI-KCEI distinction, and the identification of structurally efficient versus inefficient systems, are robust to crisis effects. However, KCEI scores and frontier positions for 2021-2022 should be interpreted cautiously, as seven countries exhibited crisis-period peaks that subsequently reversed, and four countries reached the overall efficiency frontier exclusively during crisis years. Country-specific efficiency diagnostics should emphasise patterns observed across multiple years rather than single-year scores from the 2021-2022 period. These interpretive guidelines align with the cautious interpretation principle established in Section 5.2.

The alignment between observed efficiency patterns and the implemented lag structure validates the one-year transformation assumption established in Chapter 5. The temporal pattern of efficiency-ranking disruptions is consistent with the implemented lag structure. Crisis-period IEI scores (2021-2022) show moderate disruption (ρ adjacent ≥ 0.92), while the largest reshuffle occurs at IEI 2023-2024 ($\rho=0.70$), reflecting the propagation of 2021-2022 recovery-period inputs through to 2023-2024 outcomes. This suggests that policy-driven structural change (NextGenerationEU absorption) generates larger efficiency reordering than the pandemic itself, providing partial validation that the index detects policy-relevant transformations rather than statistical noise. However, this also means that 2023-2024 IEI scores should be interpreted with awareness that they capture an active structural transition, not a settled post-crisis steady state.

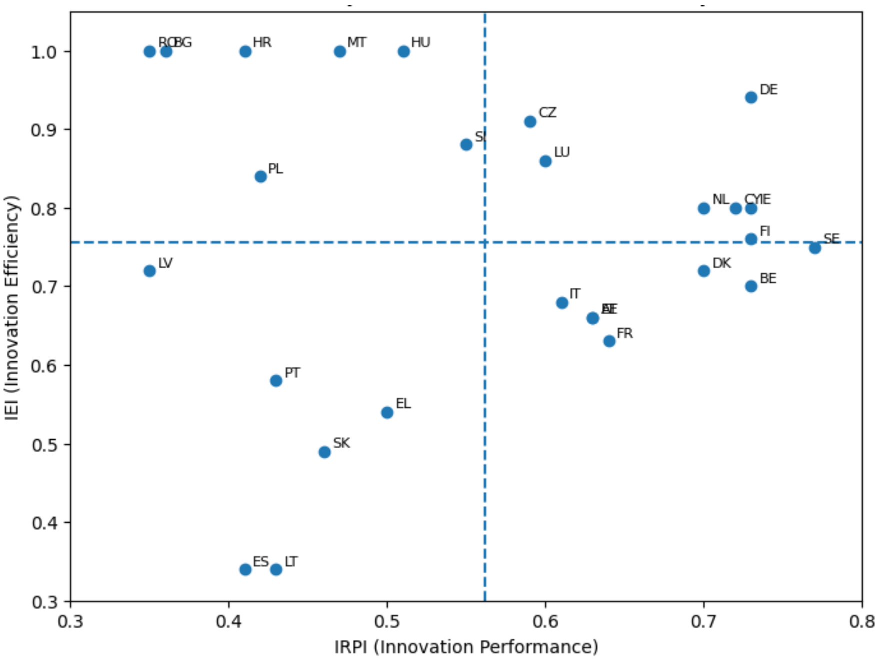
The DEA-based efficiency scores are relative measures that compare each country to the best-observed practice within the EU-27 sample. Key limitations apply efficiency is scale-sensitive, low-input countries may appear efficient due to benchmarking against similarly constrained peers, the efficiency frontier shifts if sample composition changes, the analysis identifies inefficiency but not its causes, and results are conditional on the implemented lag structure and indicator selection. Efficiency scores should be interpreted diagnostically and in conjunction with performance measures, not as standalone evaluations of innovation system quality.

6.5 INTEGRATED ASSESSMENT: PERFORMANCE-EFFICIENCY CONFIGURATIONS

This section addresses RQc (What structural characteristics emerge from joint assessment?). By combining IRPI and IEI into the IE_IRPI framework, the analysis reveals four distinct country configurations with differentiated policy implications. The quadrant analysis examines whether this joint consideration yields more meaningful peer groupings than single-dimension approaches.

Result-based innovation performance and innovation efficiency capture fundamentally different dimensions of NIS. In this research result-based innovation performance is measured by the IRPI while innovation efficiency is captured by the IEI. IEI combines efficiency in knowledge production and commercialisation stages using a multiplicative aggregation. This construction reflects the sequential nature of innovation processes and ensures that inefficiency in either stage constrains overall efficiency. Efficiency is not a substitute for performance. A country may be highly efficient yet generate modest absolute innovation outcomes if its input base is limited. On the other hand, high performance may coexist with inefficiency if strong outcomes are achieved only through disproportionately large inputs. This conceptual distinction motivates a joint analysis of IRPI and IEI. Figure 14 maps EU-27 countries in 2024 according to result-based innovation performance (IRPI) on the horizontal axis and innovation efficiency (IEI) on the vertical axis. The figure provides a joint, non-aggregated view of outcomes and efficiency, allowing countries to be classified according to both the level of innovation results they achieve and how efficiently they transform available inputs into those results.

Figure 14: IRPI vs. IEI across EU-27 Countries, 2024



Source: Own

The quadrant boundaries are defined using EU-27 benchmark values where the vertical dashed line is set at the EU-27 mean IRPI (0.56), so, countries to the right of this line exhibit above-average result-based innovation performance, while those to the left display below-average performance in terms of observable innovation outputs and outcomes. The horizontal dashed line is set at the EU-27 mean IEI (0.76), where countries above this line demonstrate above-average innovation efficiency, while those below it, exhibit below-average efficiency in converting inputs into outputs and outcomes. These benchmarks provide a relative, cross-sectional classification of countries within the EU in 2024. They are intended to support descriptive typology, not to establish normative efficiency or performance thresholds.

Countries in the upper-right quadrant such as Germany, Netherlands, Finland and Ireland, combine above-average innovation outcomes with above-average efficiency. They represent the most balanced innovation systems in the EU, achieving strong results while also making relatively effective use of available resources. These countries can be considered benchmark cases in terms of jointly high innovation outcomes and above-average efficiency, as neither limited scale nor major process inefficiencies constitute a binding constraint.

In the lower-right quadrant, countries such as Sweden, Denmark, Belgium, achieve strong innovation outcomes, but do so with comparatively lower efficiency. This pattern is characteristic of NIS where high input intensity supports performance, but where the transformation of resources into outputs and outcomes could be improved. For these countries, the results suggest that the challenge may lie in enhancing system effectiveness rather than expanding activity, a prescription that follows from capacity theory, which predicts that high-capacity systems face functional rather than structural constraints (Edquist, 2011). However, identifying specific sources of inefficiency and appropriate policy responses requires further diagnostic analysis.

The upper-left quadrant contains countries such as Romania, Bulgaria, Croatia, that exhibit above-average efficiency despite below-average performance, a configuration that NIS theory would describe as underdeveloped systems with limited capacity but proportionate transformation (Radosevic, 2017; Furman et al., 2002). These NIS are able to transform limited innovation inputs into outputs relatively effectively, yet the absolute level of innovation outcomes remains modest. This configuration highlights the key distinction between efficiency and performance: high efficiency reflects effective use of scarce resources rather than strong innovation performance. These cases illustrate why efficiency measures should be interpreted diagnostically and not as substitutes for performance indicators.

Countries in the lower-left quadrant such as Spain, Lithuania, Slovakia, underperform on both dimensions, displaying below-average innovation outcomes and below-average efficiency. This suggests the presence of more fundamental structural challenges, affecting both the scale of innovation activity and the functioning of the innovation system. The results suggest that these countries face compound challenges affecting both outcomes and transformation processes. Specific policy recommendations would require analysis of the underlying causes of underperformance, which the framework describes but does not explain.

“Romania Paradox”

Romania achieves frontier efficiency (IEI=1.00) while recording the lowest performance scores in the EU-27 (IRPI=0.35). This apparent paradox resolves when efficiency and performance are understood as theoretically distinct concepts. The DEA efficiency assesses proportionality, not level. Romania's innovation inputs (R&D expenditure, human capital, infrastructure) are among the lowest in the EU. Given these limited inputs, Romania produces outputs commensurate with what similarly resource-constrained systems achieve, hence its frontier efficiency. But “efficient use of very little” still produces “very little”. Romania lacks the accumulated innovative capacity, such as research infrastructure, skilled workforce, institutional quality, dense firm networks, that would enable higher absolute performance. High efficiency indicates that the binding constraint is insufficient capacity (scale), not inefficient transformation (process). From an NIS perspective (Lundvall, 1992; Furman et al., 2002), Romania's innovation system is “thin” rather than “dysfunctional”. Efficiency-oriented reforms (reducing waste, improving coordination) would yield limited gains because there is minimal inefficiency to eliminate. The priority is capacity-building: expanding the innovation input base through investments in education, research infrastructure, and institutional development. Only after capacity reaches critical mass would efficiency improvements become the binding constraint. This case illustrates why the framework treats efficiency as a diagnostic complement to performance rather than a substitute: high IEI without high IRPI signals a scale problem, not a success.

The figure provides visual confirmation of a central empirical finding: performance ≠ efficiency. Result-based innovation performance and innovation efficiency are not merely weakly related, they are effectively orthogonal (r=0.03). This independence, documented quantitatively and visualised in Figure 14, directly addresses the weak integration between outcome-oriented and efficiency-oriented research streams identified in Chapter 3. The two streams have developed in parallel precisely because they capture distinct phenomena, the present framework integrates them while preserving their analytical separability. This visual separation is reinforced by the correlation analysis, which shows that the contemporaneous correlation between IRPI and IEI across the 27 EU countries in 2024 is extremely weak (r=0.0286), confirming that the two dimensions capture largely independent aspects of NIS. Countries with similar performance levels can differ markedly in efficiency, and vice versa.

Although IEI efficiency scores are calculated using IKP and IKC as outputs, efficiency assesses how proportionally countries convert inputs into those outputs, not the level of outputs achieved. Two countries with identical IKP scores can have very different KPEI scores if one uses substantially fewer inputs. The near-zero correlation between IRPI and IEI (r=0.03 in 2024) confirms that knowing a country's performance level provides almost no information about its efficiency, despite the shared indicator basis. This independence reflects the genuine distinctiveness of the two dimensions: performance captures what systems achieve; efficiency captures how productively they achieve it. The independence between IRPI and IEI is empirical, not constructed by design.

The matrix therefore complements ranking-based analyses by revealing structural heterogeneity in NIS and by clarifying whether underperformance is primarily a matter of insufficient results, inefficient resource use, or both. Thus, it demonstrates that these two dimensions are largely independent and capture distinct properties of NIS. While this joint representation is analytically informative, it does not by itself yield a single summary measure suitable for ranking or comparative assessment without an explicit aggregation rule that preserves the primacy of realised innovation outcomes.

In this section the empirical results of a performance-first assessment of NIS are presented, in which realised innovation outcomes constitute the primary evaluation criterion, and efficiency serves a diagnostic role. Result-based innovation performance and efficiency are combined using a performance-adjusted aggregation approach, whereby efficiency enters as a bounded modifier of outcomes rather than as a co-equal component, ensuring that rankings remain driven by observable results. This framework allows the analysis to distinguish clearly between countries that achieve strong innovation outcomes and those that use existing resources effectively, while preserving robustness and interpretability for policy analysis.

The efficiency-adjusted performance index is computed based on the following formula:

$$\boxed{IE\_IRPI_i \ = \ IRPI_i \times [1 + \lambda \cdot (IEI_i - I\bar{E}I)]}$$

where:

- IE_IRPI_i is the composite indicator for country i,
- $IRPI_i$ is innovation performance,
- IEI_i is innovation efficiency,
- $I\bar{E}I$ is the EU-27 mean efficiency, and
- λ (0.25) is a scalar parameter controlling the strength of the efficiency adjustment.

Following the theoretical criteria established in Chapter 5, the efficiency adjustment parameter λ is calibrated empirically within the bounded interval [0.15, 0.30]. Three considerations guide the selection of the baseline specification:

- Applying λ values across the bounded interval to the observed EU-27 data confirms that values above 0.30 allow efficiency to overturn substantial performance differences, enabling high-efficiency, low-outcome countries to rank above higher-performing peers. This violates the outcome-primacy principle established in the conceptual framework. At λ=0.30, rank reversals between adjacent performance tiers begin to emerge; at λ=0.25, performance monotonicity is preserved across all country pairs with meaningful IRPI differences.
- At λ=0.25, the adjustment mechanism operates transparently: a country with IEI 0.10 above the EU-27 mean receives approximately 2.5% upward adjustment to its composite score, while a country 0.10 below the mean receives a corresponding downward adjustment. Countries at mean efficiency experience no adjustment. This magnitude is sufficient to differentiate countries within comparable performance bands while remaining moderate enough for straightforward policy interpretation.

Accordingly, λ=0.25 is adopted as the baseline specification. This value constitutes a normative modelling assumption reflecting the judgment that efficiency should meaningfully refine, but not override, performance-based rankings. Because λ is not empirically estimated, results should be interpreted conditional on this choice. Section

6.7 presents a sensitivity analysis demonstrating that the main conclusions are robust to alternative values within the bounded interval.

The resulting IE_IRPI values are normalised to a 0-1 scale, where 1 represents the highest efficiency-adjusted performance observed in the dataset and 0 represents the lowest. This normalisation preserves the relative positioning of countries while facilitating cross-year comparability and transparent interpretation of performance tiers.

In Table 29 below the IE_IRPI scores are presented for EU-27 countries for 2024 and the ranking of EU-27 countries based on those scores. In Annex 4 the full table with IE_IRPI scores and annual rankings for the EU-27 countries for the period 2019-2024 is presented. Due to the implemented 2-year time lag in the two-Stage DEA efficiency analysis based on a dataset for the period 2017-2024, there are no efficiency and IE_IRPI scores for 2017 and 2018.

Table 29: IE_IRPI results and rankings of EU-27 countries, 2024

Country code	IE_IRPI	Rank	Country code	IE_IRPI	Rank
SE	0.7670	1	**SI**	0.5710	15
DE	0.7620	2	**HU**	0.5417	16
IE	0.7374	3	**MT**	0.4982	17
FI	0.7277	4	**EL**	0.4688	18
CY	0.7262	5	**PL**	0.4321	19
BE	0.7158	6	**HR**	0.4302	20
NL	0.7054	7	**SK**	0.4279	21
DK	0.6953	8	**PT**	0.4089	22
FR	0.6164	9	**LT**	0.3889	23
CZ	0.6142	10	**BG**	0.3850	24
LU	0.6133	11	**RO**	0.3726	25
EE	0.6130	12	**ES**	0.3644	26
AT	0.6113	13	**LV**	0.3486	27
IT	0.6025	14			

Source: Own

The IE_IRPI values should be interpreted as presented in Table 30 below.

Table 30: IE_IRPI ranges and countries' groups

IE_IRPI range	Group	Interpretation
≥ 0.70	Innovation frontier	Countries with very strong innovation outcomes and leading efficiency-adjusted performance within the EU.
0.57 - 0.69	Strong performer	Countries with consistently above-average innovation outcomes and scope for further performance or efficiency gains.

0.52 - 0.56	Moderate performer	Countries with innovation outcomes broadly in line with the EU-27 average.
< 0.52	Developing performer	Countries with innovation outcomes below the EU-27 reference range, with innovation outcomes reflecting an early-stage or catching-up NIS.

Source: Own

IE_IRPI cut-offs are defined relative to the observed EU-27 performance distribution, using the EU-27 average band as the central reference and the empirical range to identify frontier and weak performers in a transparent and non-arbitrary manner. The EU-27 interval [0.52-0.56] for the period 2019-2024 represents a core performance band around the European average. Cut-offs are defined relative to this band to distinguish frontier, above-average, average, and weak performers in a transparent and interpretable way. The observed IE_IRPI range (for all EU-27 countries for 2019-2024) is min.=0.30 and max.=0.84. Based on this EU-27 countries are grouped according to IRPI thresholds that reflect empirically observed performance clusters and meaningful differences in innovation system maturity, distinguishing frontier innovation systems from strong, moderate, and weak performers. Below is the definition per each group of countries according to the IRPI score:

- *Innovation frontier*: countries in this group combine very strong innovation outcomes with favourable efficiency-adjusted performance. They represent the leading edge of result-based innovation performance within the EU and serve as empirical benchmarks for both outcome generation and system effectiveness. These systems are characterised by mature innovation ecosystems and sustained capacity to translate knowledge into economic and societal value.
- *Strong performers*: this group comprises countries with consistently above-average result-based innovation performance relative to the EU-27 benchmark. Their innovation systems generate solid outcomes and display broadly effective functioning, while still retaining scope for further improvement, either through scaling successful activities or enhancing efficiency in specific stages of the innovation process.
- *Moderate performers*: countries in this category cluster around the EU-27 core performance range, reflecting innovation outcomes broadly in line with the European average. These systems neither exhibit pronounced strengths nor pronounced weaknesses and represent stable, mid-range innovation performers, where incremental improvements in performance or system coordination could shift them into higher-performing groups.
- *Developing performers*: countries in this group display result-based innovation performance below the EU-27 reference range, even after adjusting for efficiency. This positioning reflects early-stage NIS that are still in a development

or catch-up phase. These systems often show significant potential for improvement, particularly through targeted investments, institutional strengthening, and policies aimed at scaling innovation activities and reinforcing system linkages.

The IE_IRPI results for 2024 reveal a highly differentiated and polarised innovation landscape across the EU-27 once innovation outcomes are adjusted for efficiency. At the upper end of the distribution, a small group of countries occupies the innovation frontier (IE_IRPI≥0.70), led by Sweden, Germany, Ireland, Finland, Cyprus, Belgium, the Netherlands, and Denmark. These systems combine strong realised innovation outcomes with favourable efficiency-adjusted performance and therefore represent the most advanced innovation systems in the EU context. Below the frontier, a group of strong performers (IE_IRPI between 0.57 and 0.69) includes France, Czechia, Luxembourg, Estonia, Austria, Italy and Slovenia. These countries perform above the EU-27 benchmark but remain clearly separated from the frontier group, indicating solid innovation outcomes alongside scope for further improvement through either scaling successful activities or enhancing system effectiveness in specific stages of the innovation process.

The moderate performer band (IE_IRPI between 0.52 and 0.56), defined narrowly around the EU-27 average (0.561 in 2024), contains only one country, Hungary (0.54). This reflects the fact that few EU countries operate close to the European mean in performance-adjusted terms. Rather than constituting a broad middle group, this band represents an empirical transition zone between above-average and below-average performance, highlighting limited convergence around the EU average in 2024. The remaining countries fall into the emerging performers category (IE_IRPI<0.52), including Bulgaria, Romania, Latvia, Spain, Lithuania, Portugal, Croatia, Poland, Slovakia, Malta, and Greece. Despite heterogeneity within this group, their IE_IRPI values indicate innovation systems that are still in a development or catching-up phase, with performance levels below the EU-27 reference range even after accounting for efficiency. This positioning is consistent with innovation systems that are still developing. While the results suggest that both scale and structural conditions may require attention, specific policy priorities depend on country-specific factors that the aggregate analysis does not capture.

The 2024 IE_IRPI results point to limited clustering around the EU average and pronounced differentiation between high-performing and emerging innovation systems. Efficiency-only or performance-only measures obscure important structural heterogeneity when they are not jointly interpreted within a framework that preserves outcome primacy and stage-level diagnostics.

The quadrant classification and country groupings presented above are descriptive typologies based on the constructed indices. They identify structural patterns but do not explain why countries occupy particular positions. The framework can inform where to look for policy-relevant differences but cannot prescribe what policies would improve performance. Country-specific policy recommendations require complementary analysis of institutional context, sectoral composition, and policy history that lies beyond the scope of cross-national indicator-based assessment.

As established in Chapter 4, innovation capacity is conceptualised as a latent system property that conditions both performance levels and transformation efficiency but is not directly measured. The empirical patterns observed in Sections 6.3-6.5 can be interpreted through this capacity lens to generate diagnostic insights beyond what performance or efficiency scores reveal individually.

High-capacity systems are characterised by the combination of strong performance (high IRPI) and efficient transformation (high or moderate IEI). Sweden, Finland, Denmark, and the Netherlands exemplify this configuration. From a capacity perspective, these systems possess the institutional coherence, dense actor networks, and complementary assets that enable sustained knowledge production and effective commercialisation (Lundvall, 1992, 2007). Their position in the upper-right quadrant reflects not merely favourable indicator values but underlying systemic capabilities that theory predicts should generate such outcomes.

Constrained capacity systems exhibit the opposite pattern: weak performance despite apparent efficiency. Romania, Bulgaria, and Croatia achieve frontier efficiency (IEI=1.00) while recording the lowest IRPI values in the EU-27. This configuration reveals a critical insight: these systems lack the scale, institutional depth, and accumulated capabilities that would enable absolute innovation strength, as they are efficient only in the limited sense of not wasting their meagre resources. Following Furman et al.'s (2002) framework, their 'innovative capacity' remains underdeveloped despite proportionate input-output conversion. The policy implication, derived from capacity theory rather than from efficiency scores alone, is that these systems require fundamental capability building, such as investments in human capital, research infrastructure, and institutional quality, rather than efficiency-oriented reforms.

Inefficient high-capacity systems such as France and Austria demonstrate that accumulated capabilities do not automatically translate into efficient transformation. These countries possess substantial research infrastructure and human capital (indicators of capacity) but achieve performance levels below what their resource endowments would predict. From a capacity perspective, the binding constraint is not the stock of capabilities but the flow of knowledge through the system, what Edquist (2011) terms 'functional' rather than 'structural' weaknesses. Policy priorities for these systems should focus on coordination mechanisms, technology transfer institutions, and incentive structures that improve system functioning rather than on expanding the input base.

Emerging capacity systems in the middle of the distribution (e.g., Czechia, Estonia, Portugal) exhibit mixed profiles that suggest capacity formation in progress. Their improving trajectories on both IRPI and IEI over 2017-2024 indicate that sustained investment in innovation enablers is beginning to translate into measurable outcomes, a pattern consistent with the catch-up dynamics predicted by NIS theory (Fagerberg and Srholec, 2008).

This capacity-based interpretation demonstrates that the three-dimensional framework (performance, efficiency, capacity) generates richer diagnostics than any single index. By treating capacity as an interpretive lens rather than a measured variable, the framework avoids the circularity that would arise from indexing latent capabilities while still leveraging the concept for policy-relevant analysis.

To further assess whether IE_IRPI captures economically meaningful variation, the following subsection examines its relationship with external income and institutional indicators.

6.6 EXTERNAL CRITERION VALIDATION

6.6.1 Correlation analysis: efficiency-adjusted performance index and the external variables

The three variables are initially analysed using pairwise correlation analysis in SPSS to examine the association between IE_IRPI and two key external structural variables, GDP per capita (PPS, log) and the Government Effectiveness estimate (GE_EST) has been examined. As an exploratory descriptive step, panel-pooled Pearson correlations (N=189, all country-year observations 2018-2024) reveal consistently strong and statistically robust relationships: IE_IRPI shows a strong positive correlation with GDP_PPS_log ($r=0.746$, $p<0.001$) and with GE_EST ($r=0.727$, $p<0.001$). However, because NIS exhibit strong year-on-year persistence, these panel correlations should be interpreted as upper bounds. The methodologically authoritative validation sample, as elaborated in Section 6.6.2, is the 2024 cross-section that eliminates serial dependence by construction. For this cross-section (N=27), the bivariate Pearson correlations are r(IE_IRPI, GDP_log)=0.58 and r(IE_IRPI, GE)=0.72. These correlations address a key falsification criterion specified in Chapter 5: if IE_IRPI had exhibited weak correlations ($r<0.30$) with GDP per capita and government effectiveness, or correlations with signs contrary to theoretical expectations, the framework's construct validity would have been questionable. The observed cross-sectional correlations ($r=0.58$ for GDP, $r=0.72$ for government effectiveness) substantially exceed the falsification threshold and exhibit the theoretically expected positive signs. This confirms that IE_IRPI captures meaningful variation systematically related to the structural and institutional foundations that innovation systems theory identifies as determinants of NIS functioning.

Government effectiveness measures the quality of public administration, policy implementation, and institutional performance. The magnitude of this correlation suggests that institutional capability is a closely associated factor to national innovation results, consistent with mechanisms such as regulatory quality, policy stability, administrative efficiency, and strategic public investment in R&D. Countries with higher government effectiveness are associated with stronger innovation outcomes in this sample. This correlation is consistent with theoretical expectations but does not establish that improving governance would cause improved innovation performance. The external variables GDP_PPS_log and GE_EST themselves are strongly correlated ($r=0.697$, $p<0.001$). This implies that institutional quality and economic development are tightly intertwined, and likely form mutually reinforcing structural conditions that shape NIS performance.

The correlations indicate that IE_IRPI is strongly associated with macroeconomic and institutional context. While IE_IRPI is constructed as efficiency-adjusted result-based innovation performance index, the observed results co-vary systematically with income levels and governance quality, consistent with the NIS literature. Accordingly, cross-country comparisons of IE_IRPI should be interpreted with awareness that realised innovation outcomes are embedded in broader socioeconomic conditions, without implying that these associations identify causal effects.

Robustness analysis is designed to assess whether IE_IRPI results are driven by the dissertation's key modelling choices, rather than to exhaust all possible statistical diagnostics. Accordingly, robustness is focused on sensitivity to the normative efficiency-adjustment parameter λ, evaluated through Monte Carlo simulation over a bounded policy-plausible range, and the rank stability. External criterion validation is conducted descriptively using panel GEE with robust covariance and cross-sectional models.

6.6.2 Regression analysis results and interpretation

To assess external criterion validity, the analysis focuses on cross-sectional evidence rather than panel-year inference. Over the short horizon 2019-2024, NIS exhibit strong persistence, and the very high serial dependence observed in exploratory panel specifications implies that country-year observations do not provide independent information for statistical inference. Accordingly, the main validation evidence is based on cross-sectional regressions, which avoid pseudo-replication and align with the intended use of IE_IRPI for annual benchmarking.

In the OLS model, IE_IRPI was regressed on two theoretically grounded contextual variables: economic development (GDP per capita in PPS, log-transformed) and institutional quality (Government Effectiveness estimates). These variables represent core dimensions of NIS theory and are used in comparative innovation studies. The logarithmic transformation of GDP per capita reflects diminishing marginal returns of income on innovation outcomes and is standard in innovation and growth regressions (Rodrik, 2008).

The expected signs of the coefficients reflect well-established structural relationships in the innovation systems literature rather than causal mechanisms. Higher income levels are expected to be positively associated with innovation performance, as economic development is linked to greater absorptive capacity, market depth, and cumulative investment in human capital. Similarly, higher government effectiveness is expected to be positively associated with innovation outcomes through more effective policy coordination, regulatory quality, and implementation capacity. Both expectations are well grounded in the NIS and growth literature:

- $\beta_1>0$: Higher income levels are associated with stronger innovation performance.
- $\beta_2>0$: Higher government effectiveness is associated with stronger innovation outcomes.

The data for GDP per capita (PPS, indicator code: nama_10_pc) are sourced from Eurostat database and are for the period 2017-2023 to relate to the fact that the indices (IE_IRPI and SII) for 2024 are based on data for 2023 (see SII methodology report). Similarly, the period overlaps for Government Effectiveness Estimates index (code: GE.EST source: www.govindicators.org) sourced from World bank database. Government effectiveness estimates, sourced from the World Bank's Worldwide Governance Indicators, measures administrative capacity, policy quality, and bureaucratic professionalism. Institutional capability is expected to be positively associated with innovation results by ensuring predictable regulatory environments, reducing transaction costs, and supporting effective policy implementation.

The cross-sectional OLS model for 2024 is focusing exclusively on between-country differences at a single point in time, thereby eliminating serial dependence by construction. The model specification for 2024 is as follows:

$$IE\_IRPI_{i,2024} = \alpha + \beta_1 \ln(GDP\_PPS_{i,2024}) + \beta_2 GE\_EST_{i,2024} + u_i$$

The model accounts for a substantial share of cross-country variation in IE_IRPI ($R^2=0.525$). Government effectiveness is positively and statistically associated with IE_IRPI, while GDP per capita remains positive but is no longer precisely estimated once governance is included. This pattern is consistent with the strong empirical correlation between income levels and institutional quality and suggests that, in a cross-sectional perspective, differences in governance effectiveness are particularly informative for distinguishing performance-adjusted innovation outcomes across countries. The estimated model summary is reported in Table 31.

Table 31: Estimated OLS summary

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.725	0.525	0.486	0.10

Coefficients

Model		Unstandardised Coefficients		Standardised Coefficients	t	Sig.
		β	Std. Error	Beta		
1	(Constant)	-0.884	0.947		-0.933	0.360
	GDP_PPS_log	0.126	0.093	0.284	1.356	0.188
	GE_EST	0.119	0.051	0.489	2.338	0.028

Coefficients

Model		95.0% Confidence Interval for B		Collinearity Statistics	
		Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	-2.838	1.070		
	GDP_PPS_log	-0.066	0.319	0.452	2.214
	GE_EST	0.014	0.224	0.452	2.214

Source: Own

The regression analyses provide convergent external validation evidence for IE_IRPI. IE_IRPI exhibits theoretically coherent associations with income levels and, in particular, with government effectiveness. Multicollinearity diagnostics indicate that the estimated cross-sectional specification is not adversely affected by collinearity concerns. Variance inflation factors for log GDP per capita and government effectiveness are identical and remain modest (VIF=2.21; tolerance=0.45), well below conventional thresholds associated with problematic multicollinearity. While the two variables are correlated, given the close structural relationship between income levels and institutional quality, this correlation primarily affects the precision rather than the consistency of coefficient estimates. Accordingly, the attenuation of statistical significance for GDP per capita once governance is included reflects shared explanatory power rather than model misspecification and does not undermine the interpretability or robustness of the estimated associations. These findings support the interpretation of IE_IRPI as a credible performance-adjusted indicator for comparative and policy-oriented analysis.

Accordingly, the regression results support the use of IE_IRPI for comparative and policy-oriented analysis of NIS. At the same time, the validation exercise is intentionally descriptive rather than causal and is based on a limited number of country observations; the reported associations should therefore be interpreted as long-run structural correlations rather than evidence of direct policy effects.

The positive relationships between IE_IRPI and both GDP per capita and government effectiveness, confirmed with the correlation analysis, provide empirical support for the framework's theoretical structure. These findings are interpretable through the lens of innovation capacity developed in Chapter 4. GDP per capita serves as a proxy for accumulated economic capabilities, such as the stock of resources, infrastructure, and market development that enables innovation activity. Government effectiveness captures institutional coherence, such as the quality of governance, regulatory predictability, and policy implementation capacity that reduces uncertainty for innovation actors. Both variables are theoretically associated with the latent capacity dimension defined in Section 4.1. The significant positive coefficients confirm that countries achieving high integrated performance-efficiency scores (IE_IRPI) also tend to possess the structural and institutional characteristics that innovation systems theory identifies as foundations of innovation capacity. This alignment between measured outcomes (IE_IRPI) and theoretical determinants (GDP, government effectiveness) validates the framework's construct validity: the indices capture variation that is systematically related to the deeper system properties the framework aims to assess.

Importantly, these determinants are not included in the construction of IRPI or IEI, as they serve as external validation criteria. The positive relationships therefore cannot be attributed to circular construction, as they reflect genuine co-variation between what the framework measures and what theory predicts should correlate with innovation system functioning.

6.7 ROBUSTNESS AND SENSITIVITY ANALYSIS

When composite indicators include normative modelling choices, robustness analysis is required to assess whether results are driven by arbitrary parameter selection or reflect structural features of the data. In the present study, this issue arises through the efficiency adjustment parameter λ used in the construction of the performance-adjusted index (IE_IRPI). Monte Carlo simulation is therefore employed as a transparent and systematic tool to evaluate the stability of country rankings and group classifications under plausible alternative parameter specifications. Monte Carlo methods are widely recommended in this context because they allow the analyst to propagate uncertainty through the index construction process and to evaluate the stability of results across a large number of plausible alternative scenarios (Nardo et al., 2008; Saltelli et al., 2008).

The Monte Carlo analysis focuses on parameter uncertainty, which represents the most substantively relevant source of potential instability in a policy-oriented composite indicator. The efficiency adjustment parameter λ was allowed to vary within a bounded interval:

$$\lambda \sim U(0.15, 0.30)$$

This range reflects a moderate and policy-plausible influence of efficiency on performance, ensuring that efficiency refines innovation outcomes without dominating them. By repeatedly recalculating the index under alternative λ values, the analysis directly tests whether country rankings depend critically on a specific parameter choice or remain stable across reasonable alternatives. Given the already aggregated nature of the indicators, the robustness analysis therefore emphasises parameter sensitivity and ranking stability rather than fine-grained stochastic noise. The Monte Carlo analysis was conducted in Excel using 10,000 simulation draws. The first step of the procedure was to define parameter uncertainty. The efficiency adjustment parameter λ was allowed to vary within a bounded interval:

$$\lambda \sim U(\lambda_{\min}, \lambda_{\max}), \lambda_{\min} = 0.15, \quad \lambda_{\max} = 0.30$$

As a next step, in each simulation draw, efficiency was incorporated into the composite indicator through a centred and bounded adjustment factor. Thus, ensuring that efficiency influences performance only in relative terms. Country-level efficiency values were expressed as deviations from the EU-27 mean efficiency in that draw, and scaled by the adjustment parameter λ. This centring prevents countries with frontier efficiency scores from receiving mechanically large upward adjustments and ensures that efficiency acts as a modifier of performance rather than a substitute for outcomes. As a result, countries with above-average efficiency receive a moderate positive adjustment to their performance score, while those with below-average efficiency are adjusted downward, without allowing efficiency differences alone to dominate the composite. This step operationalises the conceptual asymmetry between performance and efficiency and preserves the performance-first logic of the IE_IRPI framework. Efficiency enters the composite relative to the EU-27 mean, preventing mechanical dominance of frontier effects:

$$AdjFactor_i = 1 + \lambda \cdot (IEI_i^* - I\bar{E}I^*)$$

In each simulation, countries were ranked according to their recalculated IE_IRPI values. Across 10,000 simulation draws, these rankings were aggregated to produce a set of interpretable robustness metrics, reported in Table 32. These include the median rank, the 5th and 95th percentile ranks, the corresponding rank interval width, and the probability of appearing in the top ten or bottom five positions. These indicators provide direct evidence on whether a country's position is structurally stable or sensitive to parameter variation. Narrow rank intervals and high placement probabilities indicate robust positions, while wider intervals typically occur only for countries located near performance thresholds.

Table 32: Monte Carlo analysis results

Country	Median Rank	P05 Rank	P95 Rank	Interval Width	Pr Top10	Pr Bottom 5	MC Rank
AT	16	15	17	2	0.00	0.00	16
BE	5	3	8	5	1.00	0.00	5
BG	25	24	26	2	0.00	1.00	25
CY	5	3	7	4	1.00	0.00	5
CZ	12	9	13	4	0.26	0.00	12
DE	2	1	4	3	1.00	0.00	2
DK	7	5	8	3	1.00	0.00	7
EE	11	9	13	4	0.49	0.00	10
EL	18	17	18	1	0.00	0.00	18
ES	24	24	26	2	0.00	0.99	24
FI	4	2	7	5	1.00	0.00	3
FR	9	9	12	3	0.81	0.00	9
HR	21	19	23	4	0.00	0.06	21
HU	15	15	16	1	0.00	0.00	15
IE	4	2	6	4	1.00	0.00	3
IT	12	10	13	3	0.16	0.00	12
LT	23	21	24	3	0.00	0.70	23
LU	11	9	13	4	0.27	0.00	11
LV	27	26	27	1	0.00	1.00	27
MT	17	16	18	2	0.00	0.00	17
NL	7	5	8	3	1.00	0.00	7
PL	20	19	22	3	0.00	0.01	20
PT	22	20	23	3	0.00	0.23	22
RO	26	25	27	2	0.00	1.00	26
SE	1	1	3	2	1.00	0.00	1
SI	14	14	15	1	0.00	0.00	14
SK	19	19	21	2	0.00	0.00	19

Source: Own

The Monte Carlo analysis reveals a high degree of overall ranking stability in the IE_IRPI across 10,000 simulations. Median ranks span the full distribution from 1 (Sweden) to 27 (Latvia), providing a clear and interpretable ordering of countries. Median ranks closely align with deterministic rankings, indicating that the central tendency of country positions is not sensitive to parameter variation. Countries at both ends of the distribution show particularly strong stability: leading systems such as Sweden (median rank=1), Germany (2), Finland and Ireland (4) and systems such as Latvia (27), Romania (26), Bulgaria (25) maintain their relative positions on the lower end across almost all simulations.

Rank uncertainty, as measured by the interval width (P95-P05), is generally limited. Across all countries, interval widths range from 1 to 5 ranks, with the majority of countries exhibiting widths between 1 and 3. Very narrow intervals (width=1) are observed for several mid-range countries (e.g. Hungary, Slovenia, Greece, Latvia), indicating highly stable placements. Wider intervals (up to 5 ranks) occur primarily among countries located near performance thresholds, such as Belgium, Finland, Cyprus, and Czechia, where small changes in adjusted performance plausibly affect relative ordering. Notably, no country displays excessive rank dispersion, confirming the absence of rank instability or model-induced volatility.

Probability-based indicators further reinforce this conclusion. Countries with PrTop10=1.00 (e.g. Sweden, Germany, Finland, Ireland, Belgium, Cyprus, Netherlands, Denmark) consistently occupy the upper tail of the distribution, while countries with PrBottom5≈1.00 (e.g. Latvia, Romania, Bulgaria, Spain) are persistently positioned at the lower end. Intermediate countries typically show near-zero probabilities of appearing in extreme positions, reflecting structural mid-range positions rather than uncertainty. The narrow interval widths, stable median ranks, and concentrated tail probabilities demonstrate that the IE_IRPI rankings are robust, well identified, and resilient to uncertainty in efficiency adjustment and measurement noise. The Monte Carlo results confirm that IE-IRPI rankings are stable, interpretable, and resilient to modelling uncertainty, providing strong support for the credibility of the performance-adjusted composite indicator.

The IE_IRPI was recalculated under alternative values of the efficiency adjustment parameter ($\lambda=0.15$ and $\lambda=0.30$) to assess parameter sensitivity. Rank correlations between the baseline specification ($\lambda=0.25$) and these alternative scenarios remain high, indicating strong ordinal stability. Allowing the efficiency adjustment parameter to vary within the bounded interval:

$$\lambda \sim U(0.15, 0.30)$$

produced no substantive reordering of country rankings. Across 10,000 simulation draws, rank correlations between the baseline specification ($\lambda=0.25$) and alternative draws consistently exceeded Spearman $\rho>0.95$, indicating very strong ordinal stability.

This stability validates the performance-first aggregation design: efficiency adjustments refine performance rankings without overturning them. The bounded parameter range [0.15, 0.35] tested in Monte Carlo simulation spans the policy-plausible space identified in Chapter 5, and the consistent $\rho>0.95$ across all draws confirms that the aggregation rule is structurally robust rather than parameter-dependent. This result addresses as well a key falsification concern: had rankings been highly sensitive to parameter choice ($\rho<0.90$), the aggregation rule would have been arbitrary and policy conclusions would have depended on unjustified modelling decisions. The observed stability confirms that IE_IRPI rankings reflect structural features of the data rather than artefacts of parameter selection. The rank changes were limited to ±1-3 positions for the vast majority of countries and occurred almost exclusively among countries located near performance thresholds. Countries at the top and bottom of the distribution never switched performance tiers under any λ draw. This indicates that IE_IRPI rankings do not depend critically on a specific parameter choice but are robust across a policy-plausible range of efficiency adjustment strengths. The centred adjustment factor behaved in a stable and bounded manner across simulations. Observed AdjFactor values typically ranged between approximately 1.12 and 1.27.

Even for countries with frontier efficiency scores (IEI≈1), the upward adjustment remained moderate, confirming that centring on the EU-27 mean successfully prevents mechanical dominance of efficiency. Countries with below-average efficiency experienced proportional downward adjustments, but never of a magnitude sufficient to overturn large performance differences. Thus, efficiency modifies performance outcomes rather than replacing them, exactly as intended in the conceptual design of IE_IRPI.

Recomputing IE_IRPI under alternative λ draws produced narrow score distributions for each country, variability in IE_IRPI values across simulations was small relative to the cross-country dispersion of scores. Countries with strong result-based innovation performance retained high IE_IRPI values in all simulations, while countries with weak performance remained clustered at the lower end. No country crossed from the bottom group into the top group (or vice versa) under any simulation draw. This confirms that the performance-first structure dominates the composite, with efficiency acting only as a bounded refinement.

The purpose of robustness analysis in this study was to assess whether the results of the proposed IE_IRPI are structurally stable or products of modelling choices, rather than to exhaustively characterise all possible sources of statistical uncertainty. In this context, Monte Carlo simulation is sufficient because it directly targets the dominant source of potential instability, namely the presence of a normative efficiency-adjustment parameter (λ) in the construction of IE_IRPI. Monte Carlo simulation systematically varies this parameter over a bounded, policy-plausible range and evaluates the resulting distribution of country rankings. This approach explicitly addresses the core methodological concern relevant to composite indicator construction: whether rankings and group classifications depend on arbitrary parameter selection. The results show high rank stability, narrow uncertainty intervals, and consistent positioning of countries across simulations. As such, the Monte Carlo analysis fulfils its primary role of validating the robustness and interpretability of the index.

In order to check for stability of IE_IRPI across the analysis period and COVID-crises, the Spearman correlations are calculated and presented in Table 33.

Table 33: IE_IRPI Rank stability

Comparison	Spearman ρ	Interpretation
2019 vs 2020	0.96	Very high stability
2020 vs 2021	0.97	Very high stability
2021 vs 2022	0.93	High stability
2022 vs 2023	0.96	Very high stability
2023 vs 2024	0.95	Very high stability
2019 vs 2024	0.88	High structural persistence

Source: Own

IE_IRPI rankings exhibit high structural persistence (ρ=0.88 between pre-crisis 2019 and post-crisis 2024), intermediate between IRPI (ρ=0.89) and IEI (ρ=0.68). This reflects IE_IRPI's design as a performance-anchored index where efficiency adjustments are bounded and centred, preventing efficiency volatility from dominating the integrated measure. Table 34 compares IE_IRPI scores and rankings between 2019 (pre-crisis) and 2024 (post-crisis).

Table 34: IE_IRPI Pre-crisis vs Post-crisis

Country	IE_IRPI 2019	Rank 2019	IE_IRPI 2024	Rank 2024	Rank change	Classification
DE	0.762	1	0.762	2	-1	Stable leader
SE	0.644	7	0.767	1	+6	Structural improver
IE	0.728	3	0.737	3	0	Stable top tier
FI	0.673	5	0.728	4	+1	Stable top tier
CY	0.630	8	0.726	5	+3	Improved
BE	0.654	6	0.716	6	0	Stable
NL	0.705	4	0.705	7	-3	Slight decline
DK	0.629	9	0.695	8	+1	Stable
LU	0.736	2	0.613	11	-9	Structural decliner
FR	0.619	11	0.616	9	+2	Stable
CZ	0.498	14	0.614	10	+4	Improved
AT	0.622	10	0.611	13	-3	Slight decline
IT	0.555	12	0.602	14	-2	Stable
SI	0.433	18	0.571	15	+3	Improved
HU	0.476	15	0.542	16	-1	Stable
EL	0.388	20	0.469	18	+2	Stable
PL	0.361	21	0.432	19	+2	Stable
HR	0.302	27	0.430	20	+7	Improved from bottom
SK	0.431	19	0.428	21	-2	Stable
PT	0.458	17	0.409	22	-5	Declined
LT	0.310	25	0.389	23	+2	Stable low
BG	0.308	26	0.385	24	+2	Stable low
RO	0.361	22	0.373	25	-3	Stable low
ES	0.328	24	0.364	26	-2	Stable low
LV	0.337	23	0.349	27	-4	Stable low

Country	IE_IRPI 2019	Rank 2019	IE_IRPI 2024	Rank 2024	Rank change	Classification
EE	0.464	16	0.613	12	+4	Improved
MT	0.548	13	0.498	17	-4	Declined

Source: Own

Twenty-three countries (85%) exhibit rank changes of ±4 or less, confirming high structural stability. Four countries show larger movements requiring interpretation:

- Structural improvers: Sweden (+6 ranks) rose from rank 7 to rank 1, reflecting sustained improvement in both IRPI and IEI components throughout the period, and Croatia (+7 ranks) improved from rank 27 to rank 20, driven by gradual gains in result-based innovation performance from a low base.
- Structural decliner: Luxembourg (-9 ranks) declined from rank 2 to rank 11, reflecting the IKP decline documented in Section 6.3 that began with pre-pandemic data.
- Notable decliner: Portugal (−5 ranks) declined from rank 17 to rank 22, reflecting weak commercialisation efficiency trends throughout the period.
- Stable extremes: top tier (DE, IE, FI) maintained positions within top 8 throughout the period, while bottom tier (ES, LV, LT, BG, RO) remained in bottom 7 throughout the period.

Cyprus exhibited notable crisis-period gains, rising to rank 1 in 2022-2023 before moderating to rank 5 in 2024. This trajectory reflects the KCEI crisis-peak pattern identified in Section 6.4, where commercialisation efficiency temporarily surged during the recovery period. The partial reversal in 2024 confirms that Cyprus's top ranking during 2022-2023 was partly crisis-influenced rather than purely structural.

IE_IRPI rankings demonstrate high structural persistence across the COVID-19 crisis period. The performance-anchored design successfully buffers against efficiency volatility while preserving the diagnostic value of efficiency adjustments. Country movements primarily reflect structural changes (Sweden's improvement, Luxembourg's decline) rather than crisis artefacts, with Cyprus representing the main case where crisis-period effects temporarily elevated rankings. The stability of top-tier and bottom-tier positions confirms that IE_IRPI captures underlying innovation system characteristics robust to short-term macroeconomic disruptions. This empirical finding validates the methodological expectation articulated in Section 5.2 that cross-period comparisons and persistence patterns provide more reliable evidence of structural differences than single-year observations.

The empirical results reported in this dissertation are robust to crisis effects. The main contributions, such as the multidimensional framework, the IRPI-IEI methodology, and the identification of structural patterns, rest on findings that persist across pre-crisis,

crisis, and post-crisis periods. The robustness analysis confirms that main findings are stable to reasonable parameter variation but does not eliminate all sources of uncertainty.

6.8 EMPIRICAL FINDINGS AND HYPOTHESIS TESTING

This section synthesises the empirical findings presented in Chapter 6 in relation to the research questions and the hypotheses formulated in Chapter 1. The empirical analysis covers the EU-27 over the period 2017-2024 using three indices: IRPI, IEI, and IE_IRPI.

The research questions were designed to discover facts about NIS functioning across EU member countries. The empirical analysis reveals the following:

RQa: What patterns of variation exist in innovation results? NIS exhibit substantial and persistent heterogeneity in result-based performance. Key findings (Section 6.3):

- IRPI ranges from 0.28 (Croatia, 2019) to 0.80 (Cyprus, 2022).
- Knowledge production (IKP) and commercialisation (IKC) vary independently.
- Some countries excel at both stages (balanced); others show pronounced imbalances.
- Limited convergence over 2017-2024 despite cohesion policy.

RQb: Where do efficiency bottlenecks concentrate? Efficiency bottlenecks concentrate systematically in commercialisation. Key findings (Section 6.4):

- Mean KPEI (0.93) substantially exceeds mean KCEI (0.81).
- 18 of 27 countries (67%) show Stage 1 > Stage 2 efficiency.
- This confirms the "European Paradox" at the efficiency level.
- Knowledge production is not the binding constraint; it is commercialisation.

RQc: What structural characteristics emerge from joint assessment? Four distinct configurations emerge with differentiated characteristics. Key findings (Section 6.5) reveal: based on the quadrant analysis countries are distributed across all four performance-efficiency configurations. Countries such as Sweden and Finland achieve high performance through balanced, efficient transformation; Ireland achieves similar performance through IKC-dominant pathways; Hungary and other catching-up economies achieve moderate performance primarily through high efficiency despite limited scale. This discovery demonstrates that multidimensional classification yields substantively different and more structurally meaningful country groupings than single-dimension approaches. The configurations differ substantively from SII-based groupings (44% tier reclassification).

The main research question asked what cross-country variation in results and productivity reveals about NIS heterogeneity. The answer: heterogeneity is structural, not superficial. Countries differ not only in what they achieve but in how they achieve it, and these differences are systematically related to underlying capacity configurations. Single-dimension rankings obscure this structural heterogeneity; the multidimensional framework reveals it.

Testing the Hypotheses

This section evaluates the hypotheses formulated in Chapter 1 against the empirical evidence presented in Sections 6.2-6.7. The testing follows the operationalisation established in Chapter 5 (Section 5.1.4), applying the pre-specified support and falsification criteria.

Testing Hypothesis 1

H1 predicted that the multidimensional framework provides more comprehensive and diagnostically informative assessment than existing frameworks. Chapter 1 specified three support criteria:

Criterion 1a - Weak correlation between performance and efficiency:

Result: r(IRPI, IEI)=0.03 (Section 6.5).

Threshold: Support if r<0.30; Falsify if r>0.50.

Assessment: Criterion satisfied. The near-zero correlation confirms that performance and efficiency capture distinct dimensions.

Criterion 1b - Systematic stage-specific efficiency differences:

- Result: 18 of 27 countries (67%) show KPEI > KCEI (Section 6.4).
- Threshold: Support if >60% show same pattern.
- Assessment: Criterion satisfied. Commercialisation is systematically the binding constraint.

Criterion 1c - Meaningful performance-efficiency configurations:

- Result: Countries distributed across all four quadrants with theoretically coherent patterns (Section 6.5, Table 28).
- Threshold: Support if all quadrants populated; Falsify if >70% in one quadrant.
- Assessment: Criterion satisfied. The four-quadrant typology reveals structural heterogeneity.

Conclusion: Hypothesis 1 is supported. All three criteria are satisfied; no falsification *threshold is triggered.*

Testing Hypothesis 2

H2 predicted that the framework enables more accurate peer identification for benchmarking and policy learning. Chapter 1 specified three support criteria:

Criterion 2a - Stable rankings under parameter variation:

- Result: Monte Carlo simulation yields $\rho>0.95$ across $\lambda\in[0.15, 0.35]$ (Section 6.7).
- Threshold: Support if $\rho>0.90$; Falsify if $\rho<0.80$.
- Assessment: Criterion satisfied. Rankings are robust to methodological choices.

Criterion 2b - External validity with capacity determinants:

- Result: IE_IRPI correlates positively with GDP per capita ($r=0.58$) and Government Effectiveness ($r=0.72$) (Section 6.6).
- Threshold: Support if $r > 0.30$ with positive sign.
- Assessment: Criterion satisfied. IE_IRPI exhibits construct validity.

Criterion 2c - Substantive differentiation from SII:

- Result: 44% tier reclassification rate; $\rho(\text{IE\_IRPI}, \text{SII})=0.85$ (Section 6.5).
- Threshold: Support if systematic differences; Falsify if $\rho>0.95$.
- Assessment: Criterion satisfied. IE_IRPI provides diagnostic value beyond SII.

Conclusion: Hypothesis 2 is supported. All three criteria are satisfied; no falsification threshold is triggered.

Both hypotheses are empirically supported. The framework satisfies all pre-specified criteria without triggering any falsification threshold. For each hypothesis, the table identifies the specific empirical evidence, the relevant results sections, and the assessment of support.

Table 35: Hypothesis-results mapping

Hypothesis	Empirical Test	Key Evidence	Section	Assessment
H1: A multidimensional framework integrating performance and efficiency provides more informative assessment than single indices	IRPI-IEI independence; stage-specific variation	$r(\text{IRPI},\text{IEI})=0.03$; KPEI≠KCEI for 18/27 countries; rankings differ from SII	6.3, 6.4, 6.5	Supported

Hypothesis	Empirical Test	Key Evidence	Section	Assessment
H1a: Performance and efficiency capture distinct dimensions	Correlation analysis; quadrant distribution	Near-zero correlation; four distinct quadrants populated	6.5, 6.6	Supported
H1b: Two-stage efficiency reveals bottlenecks invisible to aggregate measures	KPEI vs KCEI comparison	Mean KPEI (0.93) > Mean KCEI (0.81); commercialisation as binding constraint	6.4	Supported
H2: The framework enables more accurate peer identification	IE_IRPI stability; peer group differentiation	Monte Carlo rank stability >0.95; structurally similar peers differ from rank-adjacent peers	6.6, 6.7	Supported
H2a: Similar performance can arise from different efficiency profiles	Quadrant analysis	SE, FI (balanced) vs IE (IKC-dominant) vs HU (efficiency-driven) achieve similar IRPI	6.3, 6.5	Supported
H2b: High efficiency does not guarantee high performance	IEI-IRPI joint distribution	RO, BG, HR: IEI=1.00 but IRPI<0.41	6.4, 6.5	Supported

Source: Own

The complementary structure of exploratory questions and confirmatory hypotheses strengthens the overall validity of the findings:

- the research questions discovered patterns without presupposing outcomes, reducing confirmation bias,
- the Hypotheses tested specific predictions derived from theory, enabling falsification, and
- together, they provide both descriptive mapping and theoretical validation.

The empirical support for both hypotheses validates the conceptual framework developed in Chapter 4. Three theoretical implications follow:

1. *Dimensional independence confirmed:* the near-zero correlation between performance and efficiency confirms that NIS theory's distinction between "what systems achieve" and "how they achieve it" is not merely conceptual but empirically observable. This supports Edquist's (2011) call for analytical separation between outcomes and processes.
2. *Sequential bottlenecks revealed:* the prevalence of stage-specific asymmetries (high KPEI, low KCEI) confirms the sequential model of innovation articulated

by Kline and Rosenberg (1986) and operationalised in two-stage DEA by Guan and Chen (2012). Aggregate efficiency measures would obscure these bottlenecks.

3. *Efficiency as necessary but insufficient:* the existence of "efficient but weak" systems (Romania, Bulgaria, Croatia) demonstrates that efficiency is a necessary but not sufficient condition for innovation success, a finding that aligns with capacity-based interpretations of NIS functioning (Furman et al., 2002).

The empirical results provide convergent validation of the methodological framework established in Chapter 5:

- the one-year lag structure produces interpretable temporal patterns consistent with innovation transformation timelines,
- the two-stage DEA decomposition reveals systematic stage asymmetries invisible to aggregate models, and
- The performance-first aggregation rule is empirically justified by the near-zero IRPI-IEI correlation and confirmed robust by Monte Carlo simulation.

No methodological assumption was empirically challenged or required substantive revision.

This chapter has interpreted the empirical findings presented in Chapter 6 and articulated the dissertation's contributions. The discussion demonstrated that the results address the literature gaps identified in Chapters 2 and 3: the weak integration between outcome-oriented and efficiency-oriented research streams is resolved through the IRPI-IEI framework, the input dominance in existing indices is contradicted by the IRPI's exclusive focus on outputs and outcomes, and the lack of stage differentiation is addressed through the IKP/IKC and KPEI/KCEI decompositions.

The contributions span four domains. Conceptually, the dissertation enforces strict separation between result-based performance, efficiency, and capacity, as NIS performance dimensions frequently conflated in existing frameworks. Methodologically, it introduces MCFS for transparent indicator selection, implements two-stage DEA aligned with innovation theory, and demonstrates performance-first aggregation. Empirically, it reveals that commercialisation rather than knowledge production constitutes Europe's structural bottleneck, and that performance and efficiency are independent dimensions requiring joint but non-conflated assessment. For measurement theory more broadly, it demonstrates how latent constructs can inform interpretation without being directly measured, and how outcome-based and frontier-based measures can be integrated without one displacing the other.

Chapter 8 concludes the dissertation by synthesising its purpose, findings, and contributions. It restates the key empirical results, draws out implications for EU innovation policy, particularly for benchmarking instruments and funding allocation, acknowledges the limitations of the analysis, and identifies directions for future research. The chapter reflects on the dissertation's broader significance for evidence-based innovation policy and for measurement methodology in complex, NIS multi-dimensional assessment contexts.

7 DISCUSSION AND CONTRIBUTIONS

Chapter 6 presented the empirical results, while this chapter interprets their significance. The discussion explicitly demonstrates how the findings address the literature gaps identified in Chapters 2 and 3, validate or refine the conceptual distinctions developed in Chapter 4, and contribute to innovation measurement theory and policy.

The chapter proceeds in four steps. First, Section 7.1 maps the empirical findings against the dominant research streams identified through literature review and bibliometric analysis, specifying which claims the results extend, contradict, or resolve. Second, Section 7.2 interprets the hypothesis testing results and their implications for the research questions. Third, Section 7.3 articulates the dissertation's contributions across four domains: conceptual contributions to innovation systems theory, methodological contributions to measurement practice, empirical and policy contributions to EU NIS performance assessment, and contributions to measurement theory more broadly.

The key finding requiring interpretation is the near-zero correlation between IRPI and IEI ($r=0.03$), which empirically confirms the conceptual independence of result-based performance and efficiency that the framework was designed to preserve. This finding, combined with systematic stage-specific asymmetries (KPEI>KCEI) and the emergence of distinct performance-efficiency configurations across countries, validates the three-dimensional framework and demonstrates its diagnostic value over existing single-dimension approaches.

7.1 DISCUSSION

This section interprets the empirical findings presented in Chapter 6 by explicitly mapping them against the literature gaps identified in Chapters 2 and 3. The discussion is organised around three key findings that address the research questions, followed by methodological validation and comparison with existing indices.

7.1.1 Mapping empirical findings to literature gaps

The bibliometric analysis (Chapter 3) identified two dominant but weakly integrated research streams in NIS performance measurement: an outcome-oriented stream focused on composite indicators, and an efficiency-oriented stream using DEA-based approaches. Chapter 2 documented five structural limitations in existing frameworks: heavy reliance on composite indicators, dominance of input-based measures, limited use of efficiency analysis, weak process orientation, and lack of stage differentiation. Table 36 maps the empirical findings of this dissertation against these literature streams and gaps.

Table 36: Mapping empirical results to literature streams

Literature Stream/Gap	Empirical Finding	Relationship	Evidence
Outcome-oriented stream (composite indices)	IRPI captures outputs/outcomes only	Extends	IRPI excludes all inputs; focuses on 7 outputs + 2 outcomes
Efficiency-oriented stream (DEA studies)	IEI provides two-stage efficiency	Extends	KPEI/KCEI decomposition with MCFS-selected indicators
Weak integration between streams	Result-based performance ≠ Efficiency	Resolves	r(IRPI, IEI) = 0.03; dimensions are empirically independent
Input dominance in existing indices	Inputs excluded from performance	Contradicts EIS practice	0 input indicators in IRPI vs. 14 in SII
Lack of stage differentiation	Two-stage decomposition	Resolves	IKP/IKC and KPEI/KCEI reveal stage-specific bottlenecks
Single-stage efficiency models	Commercialisation as binding constraint	Extends	Mean KPEI (0.93) > Mean KCEI (0.81)

Source: Own

The mapping reveals that this dissertation's findings:

- extend the outcome-oriented stream by constructing IRPI with exclusive focus on outputs and outcomes (0 input indicators versus 14 in SII);
- extend the efficiency-oriented stream by providing two-stage decomposition (KPEI/KCEI) with MCFS-selected indicators;
- resolve the weak integration between streams by demonstrating that result-based performance and efficiency are empirically independent (r=0.03); and
- resolve the lack of stage differentiation by revealing that commercialisation (KCEI) constitutes the binding constraint (mean KPEI 0.93 > mean KCEI 0.81).

These findings collectively close the bibliometric loop by providing empirical evidence for claims that have remained primarily theoretical in the innovation measurement literature.

7.1.2 Independence of result-based performance and efficiency

The correlation between IRPI and IEI across 27 EU countries in 2024 is r=0.03 (p>0.05). This near-zero correlation constitutes the dissertation's central empirical finding. It demonstrates that result-based performance and efficiency are not merely conceptually distinct but empirically independent dimensions of NIS functioning. Knowing a country's NIS performance provides essentially no information about its efficiency, and vice versa.

The joint IRPI-IEI mapping for the EU-27 in 2024 shows that efficiency is not a substitute for result-based performance: countries may achieve high efficiency with modest outcomes when inputs are limited, or strong outcomes with relatively low efficiency when NIS performance is driven by high resource intensity. Using EU-27 average benchmarks, countries cluster into four analytically meaningful groups: balanced high performers with above-average outcomes and efficiency (e.g. Germany, the Netherlands, Finland, Ireland), high-result performance but low-efficiency systems (e.g. Sweden, Denmark, Belgium), efficient but low-result performance systems (e.g. Romania, Bulgaria, Croatia), and countries underperforming on both dimensions (e.g. Spain, Lithuania, Slovakia).

From a capacity perspective, these configurations reveal distinct system types: high-capacity balanced systems, high-capacity systems with coordination inefficiencies, constrained-capacity systems efficiently using limited resources, and systems facing compound capacity-efficiency challenges. This typology demonstrates that innovation capacity, while not directly measured, emerges interpretively from the joint performance-efficiency distribution and provides the diagnostic bridge between observed outcomes and underlying system properties. The very weak contemporaneous correlation between IRPI and IEI in 2024 (r=0.0286) confirms that the two measures capture different structural properties of innovation systems. This finding provides direct empirical validation of the performance-first aggregation assumption established in Chapter 5. The assumption that efficiency should enter as a bounded modifier rather than a co-equal component rested on the theoretical expectation that result-based performance and efficiency capture distinct dimensions. The near-zero correlation confirms this expectation empirically: the dimensions are not merely conceptually separable but statistically independent, strengthening the justification for treating efficiency diagnostically rather than evaluatively. This joint representation therefore provides a powerful diagnostic tool for identifying whether innovation weaknesses stem from insufficient results, inefficient resource use, or both, while also highlighting the need for a performance-anchored aggregation, such as IE_IRPI, when a single comparative index is required.

Summary theoretical implications of IEI-IRPI:

1. contradicts composite index approaches that conflate inputs, activities, and outcomes into single scores. If result-based performance and efficiency were related, such conflation might be defensible; their independence demonstrates it is not.
2. validates the theoretical separation argued by Edquist (2011) and operationalised in this framework. The conceptual distinction between "what systems achieve" and "how they achieve it" is empirically grounded.

3. extends DEA-based studies by demonstrating that efficiency rankings cannot substitute for result-based performance assessment. Prior efficiency studies (Guan and Chen, 2012; Carayannis et al., 2016) implicitly treat efficiency as a NIS performance proxy and this finding challenges that assumption.

The independence of result-based performance and efficiency supports the framework's treatment of innovation capacity as a latent construct. Countries with similar resources achieve different outcomes through different transformation processes, suggesting that underlying system properties, such as institutional coherence, network density, accumulated capabilities, condition both dimensions without determining either one directly.

7.1.3 Two-stages and commercialisation as the binding constraint

The two-stage DEA framework used in this dissertation reveals nuances that are obscured in one-stage models. For example, Denmark appears efficient in Edquist et al. (2018) but shows significant inefficiency in KCEI here. This discrepancy highlights the importance of explicitly separating stages: aggregate efficiency scores may conceal downstream bottlenecks by averaging across fundamentally different processes. Across both stages, the comparison confirms that relative efficiency classifications are highly sensitive to indicator selection, stage definition, and model structure, a point repeatedly emphasised in the literature but rarely demonstrated systematically. Countries such as Sweden and the Netherlands alternate between efficient and inefficient classifications across studies, underscoring that DEA results should be interpreted comparatively rather than as absolute performance judgements. At the same time, several core regularities persist. First, knowledge production is generally more efficient than commercialisation across EU countries. Second, a small group of countries consistently appears close to the frontier in both stages, while many others exhibit asymmetric efficiency profiles. Third, stage-specific inefficiencies are more informative for policy design than aggregate efficiency scores. By combining MCFS-based indicator selection with a two-stage DEA-VRS model, the present study reduces arbitrariness in variable choice and improves the interpretability of efficiency scores. Compared with earlier studies, the results place greater weight on transformation effectiveness rather than input intensity, leading to more conservative efficiency classifications for high-input systems. This methodological refinement explains why some countries appear less efficient here than in prior research, while the overall pattern of relative strengths and weaknesses remains broadly consistent.

When analysing the result-based innovation performance of EU-27 countries based on IRPI, we notice that knowledge production (IKP) and commercialisation (IKC) outcomes in 2024 reveal a clear structural asymmetry across EU innovation systems. In 2024, IKP ranges from very low values around 0.11 to frontier levels close to 0.76, confirming that knowledge production capacity is a highly persistent and uneven dimension of NIS performance. Sweden and Belgium followed closely by Finland and Cyprus, define the knowledge-production frontier, reflecting strong research excellence, high SME innovation activity, and dense innovation-related employment. A second group, including Denmark, Estonia, the Netherlands, Austria, Italy, and Germany, also displays solid IKP values above 0.60, indicative of mature but not frontier-level knowledge systems. At the lower end, Romania, Bulgaria, Hungary, Latvia, Slovakia, and Croatia record very weak IKP scores, pointing to deep structural constraints in generating innovation outputs that evolve only slowly over time. By contrast, IKC exhibits much wider dispersion in 2024, ranging from 0.28 to 0.93, and is considerably more sensitive to market conditions and sectoral specialisation. Ireland stands out as a clear outlier with an IKC of 0.93, driven by exceptional NIS performance in knowledge-intensive services exports, while Germany, Sweden, the Netherlands, Denmark, Finland, Cyprus, Luxembourg, and Czechia also record high commercialisation capacity. Persistently low IKC values in Croatia, Lithuania, Greece, Portugal, Spain, Bulgaria, and Latvia highlight weak diffusion, limited scale-up, and constrained market uptake of innovation outcomes.

These stage-specific patterns provide the analytical foundation for transitioning from IKP and IKC to IRPI. By averaging IKP and IKC, IRPI captures balanced innovation outcomes across both stages and shows that similar overall result-based innovation performance levels can arise from very different structural configurations. In 2024, the innovation frontier, Sweden, Finland, Germany, Belgium, Ireland, Cyprus, Denmark, and the Netherlands, combines strong results in both dimensions, albeit with markedly different profiles (balanced systems such as Sweden and Finland versus IKC-dominated systems such as Ireland). Conversely, countries at the bottom of the IRPI distribution suffer from simultaneous weaknesses in knowledge production and commercialisation, while several mid-ranking countries exhibit pronounced imbalances between the two stages.

Across the EU-27, mean KPEI (0.93) substantially exceeds mean KCEI (0.81), with 18 of 27 countries showing higher Stage 1 than Stage 2 efficiency. This systematic stage asymmetry confirms that Europe's innovation challenge lies primarily in commercialisation rather than knowledge production. The "European Paradox", strong research but weak market translation, operates not only at the output level but at the efficiency level: EU countries transform inputs into knowledge more productively than they transform knowledge into economic value. This finding confirms the European Paradox hypothesis (Pavitt, 1991; European Commission, 1995) with efficiency-level evidence, extends Guan and Chen (2012) and Carayannis et al. (2016) by providing EU-wide evidence with consistent methodology, and aligns with the Draghi Report (2024): EU is failing to translate innovation into commercialisation. The findings show stage-specific patterns:

- Knowledge production frontier: Germany, Finland, Netherlands, Sweden, Italy operate at or near KPEI = 1.00;
- Commercialisation laggards: Finland (KCEI=0.76), Denmark (KCEI=0.79) show high KPEI but lower KCEI, exemplifying the paradox at country level; and
- Balanced performers: Germany, Netherlands, France achieve relatively high efficiency in both stages.

The aggregate efficiency scores obscure where bottlenecks arise. Stage-specific decomposition reveals that policy priorities should focus on commercialisation mechanisms, scale-up support, market access, regulatory simplification, rather than additional R&D funding for countries already efficient in knowledge production.

7.1.4 Result-based performance-efficiency configurations

The joint IRPI-IEI distribution reveals four distinct country configurations that align with theoretically predicted capacity types. When IRPI-IEI are interpreted through the capacity lens the following could be implied:

1. High-High countries possess institutional coherence, dense actor networks, and complementary assets enabling both strong outcomes and efficient transformation. These are the structural benchmarks.
2. High-Low countries have substantial accumulated capabilities but face coordination failures or diminishing returns. Policy priority: improve system functioning rather than expand inputs.
3. Low-High countries exemplify "thin" rather than "dysfunctional" systems (Radosevic, 2017). They transform limited resources proportionately but lack scale. Policy priority: expand resource base and absorptive capacity.
4. Low-Low countries face compound challenges requiring coordinated, system-wide intervention across education, finance, firm capabilities, and diffusion mechanisms.

One critical finding in the analysis is that high efficiency can coexist with low result-based performance. For example, Romania, Bulgaria, and Croatia achieve frontier efficiency (IEI=1.00) while recording the lowest IRPI values in the EU-27 (<0.40). This configuration:

- contradicts the implicit assumption that efficiency implies success;
- validates the capacity-based interpretation: these systems maximise output from minimal inputs; and
- demonstrates why efficiency must be a diagnostic complement, not a substitute for result-based performance assessment.

The IE_IRPI results for 2024 reveal a sharply differentiated and polarised innovation landscape across the EU-27 once innovation outcomes are adjusted for efficiency. A small group of countries, led by Sweden, Germany, Ireland, Finland, Cyprus, Belgium, the Netherlands, and Denmark, forms the innovation frontier (IE_IRPI≥0.70), combining strong realised outcomes with favourable efficiency-adjusted performance. A second tier of strong performers, including France, Czechia, Luxembourg, Estonia, Austria, Italy, and Slovenia, performs above the EU average but remains clearly separated from the frontier, indicating scope for further improvement through scaling or targeted efficiency gains. Very few countries cluster around the EU-27 average, with Hungary alone occupying this narrow middle range, underscoring limited convergence in performance-adjusted terms. The remaining member countries fall into an emerging-performer group with IE_IRPI values below 0.52, reflecting innovation systems that are still in a catching-up phase despite efficiency adjustments. The 2024 results highlight pronounced stratification rather than convergence and demonstrate that joint performance-efficiency measures such as IE_IRPI are essential for revealing structural heterogeneity that would be obscured by result-based innovation performance-only or efficiency-only indicators.

The temporal stability analysis yields a clear finding: efficiency measures exhibit robust structural persistence despite the COVID-19 crisis. IEI rankings show a Spearman correlation of $\rho=0.68$ between pre-crisis (2019) and post-crisis (2024) periods, with 85% of countries demonstrating score changes below ±0.10. This moderate-to-high stability indicates that DEA-based efficiency measurement captures genuine system characteristics rather than transient crisis effects. The integrated index IE_IRPI shows even stronger persistence ($\rho=0.88$), with 85% of countries maintaining rank positions within ±4 places across the six-year period. The stability hierarchy, where IRPI ($\rho=0.89$) > IE_IRPI ($\rho=0.88$) > IEI ($\rho=0.68$), confirms the conceptual framework's design logic: result-based performance reflects accumulated capabilities with high structural inertia, efficiency captures transformation processes with moderate sensitivity to external conditions, and the bounded efficiency adjustment in IE_IRPI successfully integrates both dimensions without allowing short-term efficiency fluctuations to dominate the composite measure. The crisis-sensitive patterns identified, where KCEI peaks in seven countries during 2021-2022, crisis-only frontier positions for four countries, represent localised effects that do not undermine the aggregate structural findings. Accordingly, the efficiency and integrated performance results reported in this dissertation can be interpreted as reflecting underlying innovation system characteristics robust to macroeconomic shocks, provided that country-specific scores from the index 2021-2022 are contextualised within the temporal framework established in this analysis.

The lower temporal stability of IEI relative to IRPI ($\rho=0.68$ vs 0.89) raises a question about index reliability. Nonetheless, the rank disagreement is highly concentrated: four countries (SK, CZ, SI, CY) account for over 90% of the squared rank-difference contributions, while the remaining 23 countries show ρ exceeding 0.95. Additionally, every high-volatility case is structurally interpretable: SK reflects a pre-pandemic capability breakdown; CZ and SI reflect post-2022 capability gains coinciding with NextGenerationEU implementation; CY reflects a DEA-frontier artefact normalising post-crisis. Lastly, the conceptual framework predicts this pattern: IRPI captures accumulated capabilities with high inertia, while IEI captures active transformation processes that respond to structural change. An IEI displaying ρ near 1 would indicate the index fails to detect genuine system change, undermining its diagnostic value. The observed pattern, where IEI moves precisely when and where structural transitions occurred, supports the index's reliability for its intended purpose: diagnostic identification of efficiency shifts in NIS.

7.1.5 Methodological validation: indicators selection

The MCFS procedure selected 16 indicators that differ systematically from those used in prior DEA studies and the EIS. At the indicator level, the MCFS procedure introduced in Chapter 6 represents a major empirical advance. Unlike the EIS and GII, which rely on expert-driven indicator selection, MCFS identifies variables that consistently contribute to cross-country differentiation over time. The resulting IRPI indicator set overlaps only partially with those used in EIS and GII, and this divergence is theoretically meaningful. Particularly, several indicators heavily weighted in the EIS, such as scientific publication counts and citation-based measures, are reduced in numbers in IRPI. This contrasts with studies that equate scientific output with NIS performance. MCFS prioritises indicators capturing technological outputs and commercialisation results, producing stronger alignment between measured NIS performance and economic relevance. Comparison with existing studies shows partial overlap with outcome-focused indices used in GII and OECD analyses, but the MCFS-based selection avoids redundancy and excessive multicollinearity. This explains why correlations between IRPI and EIS are positive but far from perfect: the indices capture related but conceptually distinct phenomena.

The MCFS-selected indicators are used for calculating the innovation efficiency in knowledge/technological production and commercialisation stages, and when comparing our results with previous DEA studies, our results align with earlier work and theory by Edquist et al. (2018). Yet, prior research shows no consensus on which indicators to include or how to classify them as inputs or outputs. By applying a systematic feature selection method, this study provides new clarity on indicator choice and categorisation. Recent DEA applications (Liou, 2009; Guan and Chen, 2012; Hudec and Prochádzková, 2013; Carayannis et al., 2016; Edquist et al., 2018; Anouze et al., 2024) rely on ad hoc indicator selection, drawing from prior studies or authors' assumptions without systematic testing. In contrast, our use of MCFS offers an empirically grounded approach. In Chapter 6 the indicators used across studies are compared. The comparison of articles regarding selected indicators is presented in Table 37 below.

Table 37: Comparison of articles regarding selected EIS indicators

Code	Categ.	This research MCFS	Edquist et al. (2018)	Carayannis et al. (2016)	Anouze et al. (2024)	Guan and Chen (2012)	Hudec and Procházková (2013)	Liou (2009)
1.1.1	Input	Yes		Yes		Similar	Similar	Similar
1.1.2	Input	Yes						
1.1.3	Input			Yes				
1.2.1	Output			Yes				
1.2.2	Output	Yes		Yes	Yes	Similar	Similar	Similar
1.2.3	Input							
1.3.1	Input	Yes						
1.3.2	Input							
2.1.1	Input		Yes	Yes	Similar	Similar	Similar	Similar
2.1.2	Output	Yes	Yes, but input	Yes		Similar, but external	Similar, but external	
2.1.3	Input	Yes	Yes					
2.2.1	Input	Yes	Yes		Similar	Similar, but external	Similar, but external	
2.2.2	Input	Yes						
2.2.3	Input	Yes						
2.3.1	Input							
2.3.2	Input							
3.1.1	Output	Yes	Yes					
3.1.2	Output	Yes	Yes					
3.2.1	Output	Yes		Yes		Similar, but external	Similar, but external	Similar

Code	Categ.	This research MCFS	Edquist et al. (2018)	Carayannis et al. (2016)	Anouze et al. (2024)	Guan and Chen (2012)	Hudec and Prochádzková (2013)	Liou (2009)
3.2.2	Output							
3.2.3	Input							
3.3.1	Output			Yes	Similar	Similar	Similar	Similar
3.3.2	Output		Yes	Yes, but output				
3.3.3	Output	Yes	Yes					
4.1.1	Output			Yes	Similar	Similar	Similar	
4.1.2	Output	Yes						
4.2.1	Outcome	Yes	Yes	Yes	Similar	Similar	Similar	
4.2.2	Outcome	Yes	Yes		Similar			
4.2.3	Outcome		Yes	Yes				
4.3.1	Impact							
4.3.2	Impact							
4.3.3	Impact							
Total number		16	11	12	7	9	9	5
Overlapping		N/A	8	7	4	6	6	3

Source: Own

The comparison between the MCFS-selected indicators and those used in previous DEA-based innovation studies reveals both areas of convergence and important points of departure, shedding light on how methodological choices shape empirical findings and policy interpretation. None of the reviewed studies employ an identical set of indicators, except Hudec and Prochádzková (2013), who replicated Guan and Chen (2012) on a different country sample. The applied MCFS results with 16 selected indicators and the overlap is highest with Edquist et al. (2018) where 8 indicators are the same, followed by Carayannis et al. (2016) with 7. MCFS retains new doctorate graduates in STEM (1.1.1) and tertiary education attainment (1.1.2), while excluding lifelong learning participation (1.1.3). Earlier studies, such as Carayannis et al. (2016) includes 1.1.1 and 1.1.3, while Guan and Chen (2012), Hudec and Prochádzková (2013) and Liou (2009), include 1.1.1 or similar indicator. The MCFS results indicate that early-stage human capital formation is empirically more discriminating for cross-country innovation efficiency than continuous training indicators, and this has been the standing of few of the authors that focus on STEM education, rather than lifelong learning. Among publication-related measures, MCFS retains top 10% most cited scientific publications (1.2.2) but excludes international co-publications (1.2.1) and foreign doctorate students (1.2.3). Most of the authors have as well selected 1.2.2 as the most representative scientific result.

MCFS retains broadband penetration (1.3.1) but excludes advanced digital skills (1.3.2). This is in line with previous research that links broadband diffusion to higher firm productivity, increased innovation activity, and improved efficiency. Studies show that firms with reliable high-speed internet are more likely to adopt digital technologies, engage in process and product innovation, and collaborate across organisational and geographic boundaries. At the macro level, broadband penetration has been associated with faster knowledge spillovers and more efficient matching between innovative ideas and market opportunities, particularly in knowledge-intensive and service-based sectors. Within the context of the present framework, broadband penetration is classified as an input because it represents a foundational infrastructure that conditions innovation activity without constituting result-based innovation performance per se. The other authors have not included neither of the two indicators.

The strongest convergence across studies occurs in the domain of financial and R&D-related indicators. Among input indicators, MCFS retains government support for business R&D (2.1.3), business R&D expenditure (2.2.1), non-R&D innovation expenditures (2.2.2), and innovation expenditures per employee (2.2.3). This selection closely mirrors the core input sets used by Guan and Chen (2012), Carayannis et al. (2016), and Anouze et al. (2024), confirming a broad consensus that financial and investment-related inputs are central drivers of innovation efficiency.

Venture capital expenditures (2.1.2) indicator represents a departure from prior studies that classify it as an input. In the European innovation finance landscape, venture capital typically enters at later Technology Readiness Levels (TRL 6+) after technological feasibility and market need have been established, while early-stage research relies on public funding instruments such as Horizon Europe grants (EIB, 2020). Accordingly, VC investment is treated as an output indicator reflecting market validation of innovation quality, as evidence that knowledge production has generated commercially promising opportunities, rather than as a resource input that enables innovation activity.

MCFS retains both SMEs introducing product innovations (3.1.1) and SMEs introducing business process innovations (3.1.2), as well as venture capital expenditures (2.1.2) and innovative SMEs collaborating with others (3.2.1). The other authors include at least one of the three indicators, as they are widely recognised in the literature as core manifestations of innovation activity and knowledge diffusion. A notable divergence from much of the literature is the exclusion of PCT patent applications (3.3.1) and trademark applications (3.3.2), while design applications

(3.3.3) are retained. Many earlier studies rely heavily on patent counts as proxies for inventive activity. The MCFS results suggest that, in the EU context, patents offer limited incremental discriminatory power once other firm-level innovation indicators are included, especially the highly correlated business R&D expenditure. MCFS excludes employment in knowledge-intensive activities (4.1.1) but retains employment in innovative enterprises (4.1.2).

Both exports of medium and high technology products (4.2.1) and knowledge-intensive services exports (4.2.2) are retained by MCFS, while sales of new-to-market and new-to-enterprise innovations (4.2.3) are excluded. This contrasts with several studies that emphasise innovation sales. The MCFS outcome suggests that export-based indicators provide a more stable and comparable measure of commercialisation success across countries, whereas innovation sales shares may be more sensitive to reporting practices and sectoral composition.

All impact indicators, such as resource productivity, air emissions, and environment-related technologies, are excluded by MCFS, in line with all reviewed DEA-based studies. This convergence reflects a shared view that long-term societal and environmental impacts are poorly suited to static efficiency models and are better analysed using longitudinal or causal approaches.

The comparison shows that MCFS produces an indicator set that is more selective, more stage-consistent, and more closely aligned with the transformation logic of innovation systems than those typically employed in prior DEA studies. While there is strong convergence on core financial and firm-level innovation indicators, MCFS departs from the literature by de-emphasising patents, broad human capital proxies, and impact measures, in favour of indicators that more directly capture innovation activity and market exploitation. This provides a clearer empirical basis for stage-specific efficiency analysis and enhances the diagnostic value of the resulting innovation efficiency measures for both researchers and policymakers.

Cross-study comparison is difficult: most assessments cover different country groups, periods, variables, or DEA variants. To illustrate the impact of indicator and methodological choices, we reviewed country rankings across recent DEA studies and presented the comparative analysis in Table 38. This comparison is not intended to validate our model but to demonstrate how rankings shift when indicators or samples vary. Edquist et al. (2018) used one stage model, however due to the relevance of their analysis they have been included as well. Countries appearing in multiple studies are often classified inconsistently, efficient in one, inefficient in another, highlighting the decisive role of methodological design. This study approach aligns most closely with Carayannis et al. (2016), yet even here differences in indicator selection lead to divergent rankings, confirming that choice of variables strongly shapes conclusions about national innovation efficiency.

Table 38: Innovation efficiency results for countries, compared articles

Stage 1: Technological development / knowledge production stage efficiency

Country	KPEI 2024	Edquist et al., 2018	Carayannis et al., 2016	Anouze et al., 2024	Guan & Chen, 2012	Hudec & Prochádzková, 2013	Liou, 2009
DE	0.99		0.95	0.85	0.92	Efficient	Efficient
DK	0.91	Efficient	Efficient	Inefficient	Inefficient	0.92	
ES	0.59	Efficient	0.89	Inefficient	Efficient	Efficient	Efficient
FI	1.00		Efficient	0.90	Efficient	0.98	0.99
FR	0.78	Efficient	0.80		Inefficient	Inefficient	Efficient
IT	1.00	Efficient	0.96		0.83	0.80	Efficient
NL	1.00	Efficient	Inefficient	Inefficient	0.84	Efficient	Efficient
SE	0.90	Inefficient	Inefficient	Efficient	0.85	Efficient	

Stage 2: Technological development / knowledge commercialisation stage efficiency

Country	KCEI 2024	Edquist et al., 2018	Carayannis et al., 2016	Anouze et al., 2024	Guan & Chen, 2012	Hudec & Prochádzková, 2013	Liou, 2009
DE	0.94		Efficient	Inefficient	Efficient	Efficient	Efficient
DK	0.79	Efficient	Inefficient	Efficient	Inefficient	Inefficient	
ES	0.58	Efficient	Inefficient	Efficient	Efficient	Efficient	Efficient
FI	0.76		Inefficient	0.85	Inefficient	Inefficient	0.99
FR	0.81	Efficient	Efficient		Efficient	Efficient	Efficient
IT	0.68	Efficient	Inefficient		Efficient	Efficient	Efficient
NL	0.80	Efficient	Efficient	Inefficient	Efficient	Efficient	Efficient
SE	0.84	Inefficient	Efficient	Efficient	Inefficient	0.85	

Source: Own

Our results mirror Carayannis et al. (2016) and other studies in showing that fewer countries achieve efficiency in commercialisation than in knowledge production, underscoring the EU's persistent difficulty in translating R&D into economic gains. This finding echoes the Draghi Report (2024), which highlights Europe's structural barriers to scaling innovation: "The problem is not that Europe lacks ideas or ambition. We have many talented researchers and entrepreneurs filing patents. But innovation is blocked at the next stage: we are failing to translate innovation into commercialisation, and innovative companies that want to scale up in Europe are hindered at every stage by inconsistent and restrictive regulations" (Draghi, 2024, p. 6). By contrast, Hudec and Prochádzková (2013), building on Guan and Chen (2012), produced unexpected efficiency rankings across the two stages, raising concerns about indicator alignment and modelling assumptions.

7.1.6 Comparison with existing indices

In this section IE_IRPI is compared to SII with focus on countries rankings. The SII constitutes the EU's primary policy benchmark for cross-country innovation comparison. Despite extensive methodological debate, the SII remains deeply embedded in European policy practice and therefore represents an unavoidable reference point for any alternative framework seeking policy relevance (Nardo et al., 2008; Edquist, 2011). For this reason, SII is included in the *present analysis strictly as an external comparator*, not as a validation target or optimisation criterion. This positioning follows established best practice in composite indicator construction, which recommends benchmarking new indices against dominant existing ones in order to clarify conceptual scope and analytical added value (OECD, 2018). Similar comparative strategies have been adopted in prior studies proposing revisions or alternatives to the EIS, where divergence from SII is interpreted as evidence of different measurement priorities rather than methodological inconsistency (Hollanders et al., 2017). In Table 39 below the rankings according to IE_IRPI and SII are presented.

Table 39: IE_IRPI and SII rankings of EU-27 countries in 2019-2024

Country code	Year	IE_IRPI Rank	SII Rank	Country code	Year	IE_IRPI Rank	SII Rank
AT	2019	10	7	**AT**	2022	10	7
BE	2019	6	8	**BE**	2022	8	5
BG	2019	26	26	**BG**	2022	23	26
CY	2019	8	18	**CY**	2022	1	12
CZ	2019	14	17	**CZ**	2022	15	17
DE	2019	1	9	**DE**	2022	2	6
DK	2019	9	3	**DK**	2022	3	1
EE	2019	16	14	**EE**	2022	13	10
EL	2019	20	20	**EL**	2022	14	20
ES	2019	24	12	**ES**	2022	22	16
FI	2019	5	5	**FI**	2022	5	3
FR	2019	11	10	**FR**	2022	11	11
HR	2019	27	24	**HR**	2022	20	21
HU	2019	15	21	**HU**	2022	16	22
IE	2019	3	6	**IE**	2022	6	9
IT	2019	12	16	**IT**	2022	12	14
LT	2019	25	19	**LT**	2022	26	19
LU	2019	2	4	**LU**	2022	9	8
LV	2019	23	25	**LV**	2022	27	25
MT	2019	13	11	**MT**	2022	17	13
NL	2019	4	2	**NL**	2022	7	4
PL	2019	21	23	**PL**	2022	19	24
PT	2019	17	15	**PT**	2022	24	18

Country code	Year	IE_IRPI Rank	SII Rank	Country code	Year	IE_IRPI Rank	SII Rank
RO	2019	22	27	**RO**	2022	25	27
SE	2019	7	1	**SE**	2022	4	2
SI	2019	18	13	**SI**	2022	18	15
SK	2019	19	22	**SK**	2022	21	23
AT	2020	10	8	**AT**	2023	9	6
BE	2020	8	5	**BE**	2023	7	5
BG	2020	27	26	**BG**	2023	25	26
CY	2020	2	13	**CY**	2023	1	12
CZ	2020	15	18	**CZ**	2023	10	17
DE	2020	1	7	**DE**	2023	2	7
DK	2020	9	2	**DK**	2023	4	1
EE	2020	14	11	**EE**	2023	13	10
EL	2020	18	20	**EL**	2023	15	20
ES	2020	26	15	**ES**	2023	24	16
FI	2020	7	4	**FI**	2023	3	3
FR	2020	11	10	**FR**	2023	12	11
HR	2020	24	23	**HR**	2023	20	21
HU	2020	16	21	**HU**	2023	17	22
IE	2020	3	9	**IE**	2023	6	9
IT	2020	12	14	**IT**	2023	14	15
LT	2020	25	19	**LT**	2023	22	19
LU	2020	4	6	**LU**	2023	11	8
LV	2020	22	25	**LV**	2023	27	25
MT	2020	13	12	**MT**	2023	16	13
NL	2020	6	3	**NL**	2023	8	4
PL	2020	23	24	**PL**	2023	21	24
PT	2020	19	17	**PT**	2023	23	18
RO	2020	21	27	**RO**	2023	26	27
SE	2020	5	1	**SE**	2023	5	2
SI	2020	17	16	**SI**	2023	18	14
SK	2020	20	22	**SK**	2023	19	23
AT	2021	10	9	**AT**	2024	13	6
BE	2021	7	5	**BE**	2024	6	5
BG	2021	26	26	**BG**	2024	24	26
CY	2021	2	12	**CY**	2024	5	10
CZ	2021	16	18	**CZ**	2024	10	15
DE	2021	1	7	**DE**	2024	2	9
DK	2021	9	1	**DK**	2024	8	1
EE	2021	12	11	**EE**	2024	12	11
EL	2021	17	20	**EL**	2024	18	20
ES	2021	27	17	**ES**	2024	26	14
FI	2021	8	4	**FI**	2024	4	3
FR	2021	11	10	**FR**	2024	9	12
HR	2021	23	22	**HR**	2024	20	22
HU	2021	15	21	**HU**	2024	16	21
IE	2021	3	8	**IE**	2024	3	7
IT	2021	13	15	**IT**	2024	14	16

Country code	Year	IE_IRPI Rank	SII Rank	Country code	Year	IE_IRPI Rank	SII Rank
LT	2021	25	19	**LT**	2024	23	18
LU	2021	6	6	**LU**	2024	11	8
LV	2021	24	25	**LV**	2024	27	25
MT	2021	14	13	**MT**	2024	17	17
NL	2021	5	3	**NL**	2024	7	4
PL	2021	22	24	**PL**	2024	19	23
PT	2021	19	16	**PT**	2024	22	19
RO	2021	21	27	**RO**	2024	25	27
SE	2021	4	2	**SE**	2024	1	2
SI	2021	18	14	**SI**	2024	15	13
SK	2021	20	23	**SK**	2024	21	24

Source: Own

To empirically demonstrate that IE_IRPI is not merely a relabelling of the EIS hierarchy, rank correlations and tier reclassifications are examined. The Spearman rank correlation between SII and IE_IRPI for 2024 is $\rho=0.85$ ($p<0.001$), indicating strong but far from perfect correspondence, approximately 28% of ranking variance is not shared between the indices. More substantively, 12 of 27 countries (44%) would be assigned to a different NIS performance tier under IE_IRPI compared to their SII classification. Seven countries move upward, achieving higher tier positions under IE_IRPI due to strong outcome realisation relative to their input base. Five countries move downward, receiving lower IE_IRPI rankings because their SII scores are inflated by strong framework conditions and inputs that do not translate proportionately into innovation outcomes. The largest displacements, Germany rising from rank 9 to rank 2, Spain falling from rank 14 to rank 26, reflect systematic differences in how the indices treat input intensity versus outcome realisation. This pattern of theoretically interpretable displacement confirms that IE_IRPI captures genuinely distinct structural information about NIS functioning, not a re-weighted version of the same hierarchy.

Countries such as Belgium, Ireland, Sweden, Finland, the Netherlands, rank highly across the two measures, indicating that the proposed framework preserves the core information content captured by the SII at the upper end of the distribution. This pattern is consistent with findings from other analysis of the EIS, which show that alternative aggregation rules tend to reproduce leadership groups while affecting differentiation among mid-ranking countries. Similar alignment has been reported in studies constructing alternative innovation indices based on reduced indicator sets or revised weighting schemes, where high-income innovation leaders remain broadly stable across specifications (Grupp and Schubert, 2010).

Beyond aggregate alignment, the comparison reveals systematic divergences that are analytically informative. These divergences reflect fundamental differences in indicator design. The SII aggregates a large number of indicators spanning framework conditions, inputs, outputs, and outcomes. As a result, it conflates innovation resources and realised NIS performance into a single headline figure, a limitation noted explicitly in the innovation systems literature (Edquist, 2011). Several studies modifying or decomposing the EIS demonstrate that country rankings are sensitive to this broad aggregation logic. Grupp and Schubert (2010) argue that composite innovation indices heavily influenced by inputs tend to overstate NIS performance in countries with strong research infrastructures but weaker commercialisation outcomes. The present results are consistent with these findings. Countries such as Austria, France, and Sweden achieve very high SII scores due to strong input and framework indicators yet display more moderate NIS performance under IE_IRPI because commercialisation outcomes lag behind knowledge production. Conversely, countries such as Germany, Cyprus and Ireland achieve higher NIS performance rankings under the proposed framework than under SII, reflecting strong outcome realisation relative to their input base.

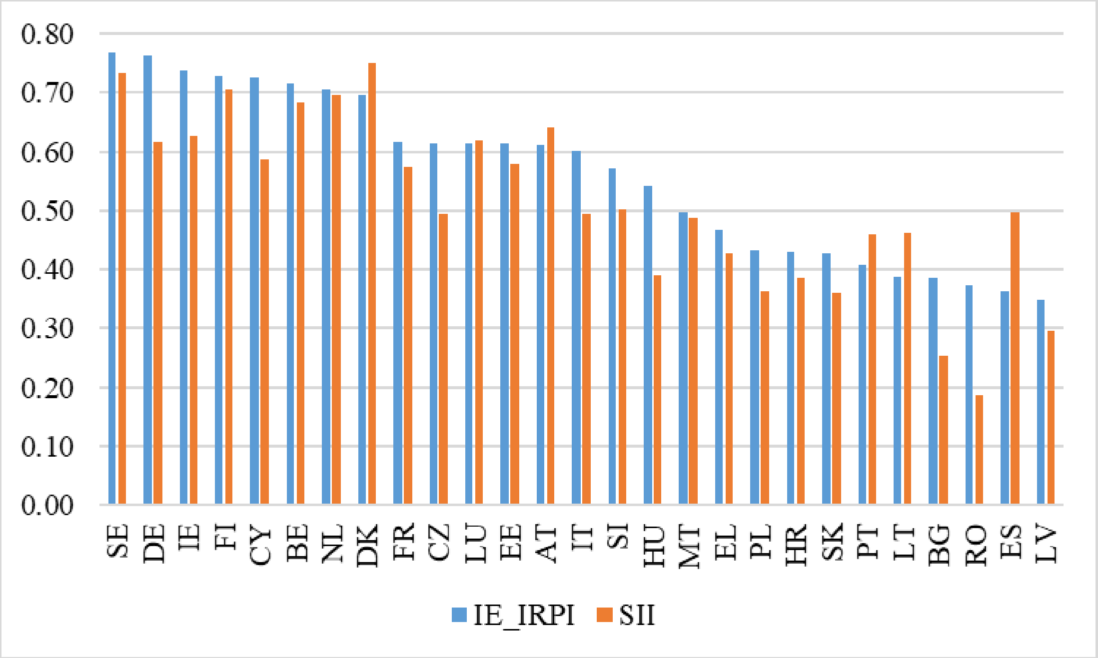
A further source of divergence arises from the explicit treatment of efficiency in the proposed framework. The SII does not distinguish between outcome levels and the efficiency with which those outcomes are generated. By contrast, the separation between IRPI and IEI reveals cases where countries appear efficient despite low result-based performance levels, notably several cohesion economies. This pattern illustrates a key limitation of the SII headline indicator as it cannot distinguish between low NIS performance due to insufficient scale and low NIS performance due to systemic inefficiency. This critique aligns with earlier methodological assessments of composite indicators, which caution that headline indices that aggregate heterogeneous inputs, processes, and outcomes into a single score may obscure important structural differences across countries when underlying dimensions are not separately observed.

In summary, the comparison confirms that the outcome-oriented indices reproduce broad leadership patterns identified by the EIS, while also revealing systematic and theoretically meaningful divergences. These divergences are consistent with published critiques of the SII and with prior studies demonstrating the sensitivity of innovation rankings to indicator choice and aggregation rules. Accordingly, differences between SII and the proposed measures are interpreted not as inconsistencies but as evidence that the proposed framework provides complementary and more diagnostically precise insights into NIS.

The bar chart for 2024 in Figure 15 below shows IE_IRPI and SII scores across EU countries and reveals both broad alignment and important systematic deviations between the two indices. Innovation leaders such as Finland, Denmark, Sweden, Belgium, and the Netherlands score highly on both metrics, confirming that IE_IRPI is consistent with established assessments of frontier NIS performance. However, the figure also shows notable divergences as countries such as Germany, Ireland, and Cyprus exhibit significantly higher IE_IRPI than SII values, indicating that they achieve stronger innovation results than their structural input conditions would predict.

Furthermore, at the lower end of the NIS performance spectrum countries such as Bulgaria, Latvia and Romania stand out for their low IE_IRPI scores, significantly below the EU average. Their low IE_IRPI values reflect weaknesses in both knowledge production and commercialisation. While countries including Denmark, Lithuania, Portugal, and Spain display higher SII than IE_IRPI scores, reflecting high R&D resources and investments but comparatively weaker commercialisation outcomes. These deviations demonstrate that IE_IRPI captures NIS performance differences that remain invisible in the input-focused SII, providing a more direct measure of realised innovation outcomes and system efficiency.

Figure 15: IE_IRPI vs SII for EU-27 countries, 2024



Source: Own

The comparison with SII therefore serves a limited but important validation purpose. It implies that IE_IRPI aligns with the dominant EU benchmarking tool at the level of realised outcomes, ensuring continuity and interpretability for policymakers. At the same time, it highlights the added value of IE_IRPI in correcting the distortions inherent in composite indices that conflate inputs, conditions, and results. Accordingly, correlation with SII is not used as proof of validity in an absolute sense, but as evidence that IE_IRPI captures a purified version of the same outcome-related dimension that the SII measures imperfectly. The empirical validation of IE_IRPI rests not on endorsement by SII, but on its conceptual consistency, transparent construction, and superior diagnostic capacity when interpreted jointly with innovation efficiency results. In this sense, the strong correlation between IE_IRPI and SII reinforces rather than contradicts the critique developed earlier: it shows that meaningful NIS performance signals can be extracted from existing data once inputs and structural redundancies are analytically removed.

7.2 INTERPRETATION OF HYPOTHESES TESTING AND RESEARCH QUESTIONS

This section interprets the empirical findings presented in Chapter 6 in relation to the hypotheses and research questions formulated in Chapter 1. Rather than restating results, the discussion focuses on what the evidence implies for understanding NIS and how it advances existing measurement approaches.

Interpretation of Hypothesis 1

Section 6.8 confirmed that H1 is supported: all three criteria are satisfied. The interpretive significance extends beyond statistical confirmation:

The near-zero correlation ($r=0.03$) between IRPI and IEI is not merely a methodological finding but a substantive discovery about NIS functioning. It confirms that "what systems achieve" and "how productively they achieve it" are genuinely independent dimensions, not merely conceptually separable but empirically distinct. This independence validates the framework's core premise: single composite indicators that conflate outcomes with processes obscure fundamental structural differences.

The systematic KPEI-KCEI asymmetry (67% of countries showing Stage 1 > Stage 2 efficiency) confirms the "European Paradox" at the efficiency level. EU countries transform inputs into knowledge more productively than they transform knowledge into economic value. This stage-specific diagnosis is invisible in aggregate efficiency measures and has direct policy implications: the binding constraint lies in commercialisation mechanisms, not in research capacity.

The four-quadrant distribution aligns with theoretically predicted capacity configurations. High-High countries exhibit systemic coherence; High-Low countries show coordination failures despite accumulated capabilities; Low-High countries face scale constraints despite efficient transformation; Low-Low countries require system-wide intervention. This typology provides a qualitatively different basis for policy learning than one-dimensional rankings.

Interpretation of Hypothesis 2

Section 6.8 confirmed that H2 is supported: all three criteria are satisfied. The interpretive significance concerns the framework's practical value for policy:

The robustness finding ($\rho>0.95$ under Monte Carlo simulation) demonstrates that IE_IRPI rankings reflect persistent structural relationships, not artefacts of specific parameter choices. Policymakers can use the rankings with confidence that small methodological variations would not substantially alter country positions.

The external validity finding (positive correlations with GDP per capita and Government Effectiveness) confirms that IE_IRPI captures meaningful dimensions of innovation-system functioning. The stronger correlation with institutional quality (r=0.72) than with income level (r=0.58), reinforced by the joint OLS specification (Table 31) in which Government Effectiveness has nearly twice the standardised regression weight of GDP per capita (β=0.489 vs β=0.284), suggests that governance capacity is more strongly associated with efficiency-adjusted performance than resource abundance, a finding consistent with the efficiency results showing that some lower-income countries achieve high IEI.

The differentiation from SII (44% tier reclassification, ρ=0.85) confirms that IE_IRPI provides diagnostic value beyond existing benchmarks. Countries like Germany (+7 ranks) and Spain (-12 ranks) illustrate how efficiency adjustment reveals structural characteristics that aggregate rankings obscure.

Interpretation of Research Questions

The answers to the three research questions address a central gap in the innovation measurement literature: the dominance of composite indices that conflate inputs, framework conditions, and results, thereby limiting diagnostic usefulness.

RQa interpretation: Result-based performance can be measured as results alone

The IRPI demonstrates that result-based innovation performance can be meaningfully conceptualised and measured as realised outputs and outcomes alone, without relying on inputs or institutional proxies. This resolves a long-standing conceptual ambiguity: existing indices conflate what systems achieve with what they possess, making it impossible to distinguish capacity from performance. The IRPI's focus on outputs (IKP) and outcomes (IKC) provides a cleaner measure of what NIS actually produce. The substantial cross-country variation discovered through RQa is not artefactual, it reflects genuine differences in innovation system functioning that persist over time (ρ=0.89 between 2019 and 2024 IRPI rankings).

RQb interpretation: The European Paradox confirmed at efficiency level

The IEI demonstrates that innovation efficiency can be operationalised as a sequential transformation process, revealing bottlenecks invisible in aggregate measures. The systematic KPEI-KCEI asymmetry (67% of countries showing Stage 1 > Stage 2) confirms the European Paradox not merely as a descriptive observation but as a structural efficiency constraint. This finding has direct policy implications: the binding constraint lies in commercialisation mechanisms, not in research capacity. Additional R&D funding without addressing commercialisation bottlenecks will not resolve the underlying structural problem.

RQc interpretation: Joint assessment reveals structural heterogeneity

The four-quadrant typology demonstrates that combining result-based performance and efficiency, without collapsing them, yields qualitatively different and more policy-relevant country groupings than single-dimension rankings. Countries at similar aggregate performance levels may face fundamentally different challenges:

- High-High countries need to maintain frontier position
- High-Low countries need to improve system efficiency
- Low-High countries need to expand resource base
- Low-Low countries need system-wide intervention

This structural diagnosis is unavailable from aggregate rankings and directly supports evidence-based policy learning.

Together, the RQ findings confirm that NIS heterogeneity is multidimensional. The dissertation's framework makes this heterogeneity visible and interpretable, moving innovation assessment from descriptive ranking toward structural diagnosis. The exploratory discoveries and confirmatory hypothesis tests provide convergent validation of the conceptual framework developed in Chapter 4. In interpretive terms, the empirical evidence confirms that result-based innovation performance and innovation efficiency are analytically distinct yet interdependent dimensions of NIS functioning. Their joint assessment provides superior explanatory and diagnostic power compared to existing one-stage and one-dimensional approaches. By operationalising this distinction empirically and validating it through robustness and external criteria, the dissertation advances both the conceptual clarity and practical relevance of NIS performance measurement. This interpretive synthesis provides the foundation for the subsequent discussion of contributions, limitations, and implications for innovation policy and future research.

7.3 CONTRIBUTIONS OF THE DISSERTATION

7.3.1 Conceptual contributions

The primary conceptual contribution of this dissertation lies in the explicit and systematic differentiation between innovation performance, innovation efficiency, and innovation capacity as analytically distinct yet interrelated dimensions of NIS functioning. This contribution operates at two levels: within innovation studies, it resolves long-standing conceptual ambiguities regarding what innovation indices actually measure; within measurement theory more broadly, it demonstrates how theoretically grounded dimensional separation can be operationally enforced in composite indicator design. While each of these concepts appears in the innovation literature, they are rarely integrated into a coherent analytical framework that enforces strict boundaries at the level of empirical measurement.

The theoretical foundations for this distinction are well established. The innovation systems literature has long emphasised that innovation is a cumulative, multi-stage process shaped by institutional interactions, learning dynamics, and systemic coherence (Freeman and Soete, 1997; Lundvall, 1992; Nelson, 1993). More recent contributions explicitly argue that innovation systems should be assessed along multiple dimensions, including results, processes, and underlying capacities (Edquist, 2011; Secka, 2023). However, as discussed in Chapter 2, most empirical applications collapse these dimensions into single composite indices, thereby conflating what systems achieve with the conditions that enable those achievements.

This dissertation advances the literature by translating these conceptual distinctions into a disciplined analytical framework. Result-based innovation performance is defined strictly as the observable outputs and outcomes of innovation activities, corresponding to the realised results of knowledge creation and economic exploitation. This definition is consistent with Schumpeter's distinction between invention and innovation (Schumpeter, 1934) and with result-oriented interpretations of NIS performance in empirical literature (Fagerberg, 2003; Edquist, 2011). Crucially, the framework excludes inputs, framework conditions, and institutional determinants from the NIS performance dimension, thereby avoiding the circularity inherent in indices that treat resources as outcomes.

Innovation efficiency is conceptualised as the productivity with which innovation systems transform inputs into outputs and outcomes across sequential stages. While efficiency analysis has been applied in innovation studies, particularly through DEA-based approaches (Guan and Chen, 2012; Jovanović et al., 2022), it is typically employed either in isolation or as a substitute for broader NIS performance measurement. This dissertation departs from that practice by treating efficiency as a complementary dimension rather than as a proxy for NIS performance. By explicitly aligning efficiency measurement with the two-stage innovation process articulated in Chapter 4, the framework captures how effectively systems convert resources into results without conflating efficiency with success.

The third dimension, innovation capacity, represents the most conceptually novel element of the framework. Drawing on the notion of "national innovative capacity" (Furman et al., 2002) and subsequent work on systemic capabilities and absorptive capacity (Bartels et al., 2012; Edquist, 2011), innovation capacity is defined as a latent system-level property reflecting the ability of a country to mobilise and recombine knowledge, skills, institutions, and networks over time. Unlike existing indices that operationalise capacity through input proxies or institutional indicators, this dissertation deliberately refrains from measuring capacity directly. Instead, capacity is treated as an interpretive construct that explains persistent patterns in result-based performance and efficiency rather than constituting a measurable outcome itself. This distinction, developed in Chapters 3 and 4, resolves a key conceptual ambiguity in the literature between determinants of innovation and innovation results.

The novelty of the framework therefore lies not in introducing new concepts per se, but in enforcing conceptual discipline between dimensions that are frequently conflated in empirical work. By maintaining analytical separation between result-based performance, efficiency, and capacity at the measurement stage, while allowing these dimensions to interact at the level of interpretation, the dissertation offers a more coherent representation of NIS functioning. This approach responds directly to critiques in the literature that existing innovation metrics lack theoretical grounding and diagnostic clarity (Edquist, 2011; Barrichello et al., 2020; Secka, 2023).

To conclude, the conceptual contribution of this dissertation consists in operationalising a multidimensional NIS performance assessment framework that is firmly rooted in innovation systems theory, consistent with the sequential nature of innovation processes, and explicitly designed to avoid the conflation of inputs, processes, and outcomes. By doing so, it advances the theoretical foundations of NIS performance measurement and provides a clearer basis for both empirical analysis and policy interpretation.

7.3.2 Methodological contributions

The methodological innovations in this dissertation address challenges that extend beyond innovation measurement to the general theory of composite indicator construction (Nardo et al., 2008; Greco et al., 2019). The problems of dimensional conflation, indicator redundancy, aggregation bias, and latent construct identification arise wherever complex multi-dimensional phenomena are synthesised into comparable metrics. The solutions developed here, such as enforced separation, data-driven selection, performance-first aggregation, and interpretive treatment of latent properties, therefore constitute contributions to measurement methodology applicable across policy domains. This dissertation's methodological novelty lies in five principles that contribute to both innovation measurement and composite indicator theory more broadly:

1. Enforced dimensional separation: Result-based performance and efficiency are measured through independent indices rather than collapsed into a single score. The near-zero correlation (r=0.03) validates this design: conflated indices would obscure structural independence. This principle is generalisable to any multi-dimensional assessment where theory predicts distinct constructs.
2. Theoretically anchored stages: The two-stage structure (knowledge production to commercialisation) derives from NIS theory, not ad hoc data groupings. This

demonstrates how domain theory should drive indicator architecture rather than emerge post hoc from statistical patterns.

3. Data-driven indicator selection: MCFS replaces expert judgment, directly addressing multicollinearity and input dominance. This offers a replicable alternative to the PCA-based or expert-driven selection that dominates composite indicator practice.
4. Capacity as latent construct: Unlike indices that proxy capacity through inputs, this framework avoids the identification problem by treating capacity interpretively. This resolves a general measurement-theoretic challenge: how to incorporate theoretically important but empirically elusive constructs without conflating determinants with outcomes.
5. Performance-first aggregation: IE_IRPI is not a weighted average; efficiency adjusts but cannot override result-based performance, reflecting theoretical outcome primacy. This demonstrates how frontier-based and outcome-based measures can be integrated without one displacing the other, a contribution to the literature on hybrid performance assessment.

The framework generates insights structurally unavailable from existing approaches: stage-specific bottleneck identification, structural peer grouping, and the empirical independence of result-based performance and efficiency.

The dissertation contributes by integrating composite index construction with two-stage DEA in a coherent framework that preserves analytical separability while allowing synthesis. The combination of IRPI and IEI enables both detailed diagnostic analysis and aggregate comparison through IE_IRPI. The extensive robustness and sensitivity analysis presented in Chapter 6, including Monte Carlo simulation and external criterion validation, addresses a key critique in the composite indicator literature concerning sensitivity to normative assumptions. The results demonstrate that the findings are stable and not driven by arbitrary modelling choices.

One of the key methodological contributions lies in the introduction and application of MCFS as a transparent and data-driven procedure for selecting NIS performance indicators. To the author's knowledge, this represents a novel application of MCFS in the context of NIS performance measurement. This approach directly addresses long-standing methodological weaknesses of existing composite indicators, most notably the SII, which relies on expert-driven indicator selection that often result in input dominance, redundancy, and conceptual ambiguity. The MCFS procedure is applied longitudinally to the candidate indicator pool, allowing feature relevance to be evaluated not only across countries but also consistently over time. This temporal dimension is critical in the context of innovation systems, where structural relationships are expected to persist rather than fluctuate arbitrarily from year to year. By identifying variables that consistently contribute to clustering structure across multiple periods, MCFS provides an empirically grounded basis for indicator selection that is both reproducible and theoretically defensible. This data-driven selection process directly responds to the critique articulated in Chapter 2 regarding the construction of the SII, where large numbers of correlated indicators are retained despite limited incremental information content. MCFS explicitly penalises redundancy by favouring variables that contribute unique discriminatory power across clusters, thereby reducing multicollinearity and eliminating deterministic indicator chains. As a result, the final indicator set underlying the IRPI and IEI is parsimonious, non-redundant, and empirically justified.

Building on this selection procedure, the IRPI offers a conceptually coherent measure of result-based innovation performance that focuses exclusively on outputs and outcomes, in line with the theoretical framework developed in Chapter 4. Inputs, framework conditions, and enabling factors are deliberately excluded from the performance measure, preventing the conflation of resources with results that characterises input-dominated frameworks such as the EIS. This design choice ensures that high resource intensity is not misinterpreted as strong NIS performance and directly addresses the first research gap identified in Chapter 1. Moreover, the IRPI explicitly incorporates the two-stage structure of the innovation process, distinguishing between knowledge and technological production (Stage 1) and commercialisation and diffusion (Stage 2). This is operationalised through the IKP and IKC subindices, which map cleanly onto the sequential stages of innovation activity. By contrast, most existing indices collapse outputs and outcomes into a single dimension, obscuring whether system weaknesses originate in research performance, valorisation, or diffusion. The IRPI therefore provides a level of diagnostic granularity that is absent from one-stage composite indicators.

Essentially, the IRPI is not designed as a standalone replacement for efficiency analysis but as a result-based complement to the IEI. While IRPI captures what innovation systems achieve, IEI assesses how efficiently they transform inputs into those achievements across the same two stages. Together, IRPI and IEI form a unified measurement system that preserves analytical separability while enabling integrated interpretation. This dual structure resolves a core methodological limitation of the EIS, which embeds inputs, activities, and outcomes into a single score and thereby blurs the distinction between result-based performance and efficiency.

In methodological terms, the integration of MCFS-based indicator selection with a two-stage performance-efficiency framework represents a significant advance over existing innovation measurement approaches. It replaces expert-driven aggregation with a transparent, data-driven procedure, and enforces conceptual discipline between inputs, outputs, and outcomes. Additionally, this approach embeds NIS performance within a sequential process logic consistent with innovation systems theory. As a result, the combined IRPI-IEI framework provides a more nuanced, robust, and policy-relevant assessment of NIS performance than prevailing benchmarking instruments.

7.3.3 Empirical and policy contributions

A central empirical and policy contribution of this dissertation lies in providing policymakers with a diagnostically meaningful and theoretically grounded measurement framework for assessing NIS. The positive influence of innovation on economic growth, competitiveness, and societal welfare has long motivated public intervention, but the increasing scale and complexity of innovation policy has intensified the need for reliable tools to evaluate outcomes and guide strategic decision-making.

The multidimensional framework developed in this dissertation directly addresses identified SII challenges by integrating result-based innovation performance and two-stage innovation efficiency into a coherent analytical system. Empirically, this framework enables policymakers to distinguish between the level of innovation results achieved and the processes through which these results are generated. Result-based indicators (IRPI, IKP, IKC) capture scientific, technological, and economic achievements, while efficiency indicators (IEI, KPEI, KCEI) reveal how productively innovation systems transform existing inputs into outputs and outcomes. When combined through the efficiency-adjusted performance index these dimensions provide a structured diagnostic tool that aligns closely with the functional reality of innovation systems.

From a policy perspective, this distinction is critical. Governments must allocate public resources, design interventions, and prioritise reforms under conditions of fiscal constraint and increasing uncertainty. Without analytically sound NIS performance measures, there is a substantial risk of misidentifying bottlenecks, misallocating funds, or focusing on symptoms rather than systemic causes. The empirical results presented in Chapter 6 demonstrate that countries with similar innovation outcomes often differ substantially in efficiency, and conversely, that high efficiency may coexist with weak result-based performance due to scale constraints. These patterns cannot be detected using one-dimensional rankings but are directly observable within the proposed framework.

The diagnostic value of the IEI is particularly relevant in this context. Despite its scale sensitivity, IEI allows policymakers to identify whether weak innovation outcomes are primarily driven by insufficient resource endowment or by ineffective transformation processes. In low-input countries, high efficiency scores indicate that NIS performance constraints stem mainly from limited scale rather than from inefficiency, suggesting that policy priorities should focus on expanding the innovation input base. In contrast, low efficiency in high-input systems signals process-level bottlenecks that call for institutional reform, improved coordination, or better incentives rather than additional funding. This distinction is essential for designing proportionate and effective policy responses. This diagnostic logic is explicitly embedded in the construction of IE_IRPI, where efficiency enters as a bounded adjustment to result-based performance. By preserving outcome primacy, the index prevents high efficiency from compensating for weak innovation results, while still incorporating valuable information on transformation processes. As a result, IE_IRPI supports coherent ranking and comparison without sacrificing interpretability or policy relevance. It enables policymakers to assess not only where countries stand in terms of innovation outcomes, but also why they occupy those positions.

Beyond ranking, the framework enhances policy learning and peer-group identification, which are central objectives of EU-level coordination mechanisms. Current EIS groupings cluster countries based on overall scores, without differentiating between result-based performance levels and efficiency patterns. The empirical application of the new framework suggests that such groupings often mask fundamental structural differences. By contrast, the combined IRPI-IEI perspective allows policymakers to identify peer countries that face similar structural challenges, whether related to scale, efficiency, or stage-specific weaknesses in knowledge production or commercialisation. This improves the relevance of benchmarking exercises and strengthens the foundations for horizontal policy learning.

More broadly, the framework responds to calls in the literature for innovation measurement tools that support both policy formulation and evaluation. As noted by Patel and Pavitt (1994), and Edquist et al. (2018), innovation policy requires measurement systems that go beyond descriptive monitoring and provide actionable insights into system functioning. By combining empirically grounded result-based performance measures with efficiency diagnostics, the proposed framework equips policymakers with analytical instruments that are better aligned with long-term objectives such as digital transformation, technological sovereignty, and the green transition, while remaining responsive to short-term shocks and structural change.

In empirical terms, the dissertation contributes by demonstrating that innovation systems differ not only in their observable results, but also in the mechanisms through which these results are generated. In policy terms, it offers a practical, evidence-based tool that enables governments to move from ranking to diagnosis, from generic benchmarking to targeted intervention, and from static comparison to informed policy learning.

Finally, these empirical and policy contributions are underpinned by methodological innovations that extend beyond innovation measurement. The following section explicates these contributions to measurement theory as a discipline.

7.3.4 Contributions to measurement theory

Beyond innovation-specific applications, this dissertation contributes to the broader literature on composite indicator construction and multi-dimensional NIS performance assessment. The methodological challenges addressed here, such as dimensional conflation, indicator redundancy, aggregation bias, and the integration of level-based and frontier-based measures, arise across diverse policy domains where complex phenomena must be synthesised into comparable metrics (Nardo et al., 2008; Greco et al., 2019).

First, the dissertation demonstrates how theoretically grounded dimensional separation can be operationally enforced in composite indicator design. The standard approach in measurement theory treats composite indices as weighted aggregations of correlated indicators, with techniques such as PCA or equal weighting used to manage redundancy (Nardo et al., 2008). This dissertation proposes an alternative architecture: dimensions that theory defines as conceptually distinct (result-based performance vs. efficiency) should be measured through separate indices rather than collapsed into a single score. The empirical finding that IRPI and IEI are effectively uncorrelated (r=0.03) validates this design principle: enforced separation reveals structural independence that conflated indices would obscure. This approach offers a replicable template for other domains, such as governance quality, sustainability, or human development, where multi-dimensional constructs risk being flattened into unidimensional rankings.

Second, the dissertation advances the integration of composite indices with frontier-based efficiency analysis. Composite indicators and DEA represent distinct traditions in NIS performance assessment: the former aggregates observable outcomes, the latter benchmarks units against best-practice frontiers (Cherchye et al., 2007). These approaches are rarely combined within a unified framework. The IE_IRPI demonstrates how efficiency can enter as a bounded modifier of outcome-based performance rather than as a substitute or competing measure. The performance-first aggregation rule, where efficiency adjusts but cannot override realised outcomes, embodies the measurement-theoretic principle that diagnostic information should complement rather than displace primary constructs. This integration strategy is generalisable to any assessment context where both outcome levels and transformation productivity are policy relevant.

Third, the MCFS-based indicator selection procedure contributes to the literature on variable selection for composite indicators. The OECD Handbook (Nardo et al., 2008) identifies indicator selection as a critical yet under-theorised step, noting that expert judgment often dominates despite risks of redundancy and bias. By applying a data-driven algorithm that evaluates indicators based on their contribution to cross-sectional clustering structure over time, the dissertation provides a reproducible and theoretically defensible selection procedure. Unlike PCA, which transforms variables and sacrifices interpretability, MCFS retains original indicators while eliminating redundancy. This approach addresses the multicollinearity problem endemic to composite indicators without the information loss associated with dimension-reduction techniques.

Fourth, the treatment of latent constructs offers a methodological contribution to validity in composite indicator design. Measurement theory distinguishes between formative indicators, which constitute a construct, and reflective indicators, which manifest an underlying latent variable (Diamantopoulos and Winklhofer, 2001). Innovation capacity, as conceptualised here, is neither: it is treated as a latent system property that conditions observable outcomes but is not directly measured. This interpretive approach avoids the identification problem that arises when latent constructs are proxied through correlated indicators, which risks conflating determinants with outcomes. The dissertation demonstrates that latent constructs can retain explanatory power in a measurement framework without being operationalised as separate indices, a design principle applicable wherever theoretical constructs resist direct measurement.

Fifth, the robustness analysis contributes to best practice in composite indicator validation. The OECD Handbook recommends sensitivity analysis but provides limited guidance on implementation. This dissertation operationalises robustness testing through Monte Carlo simulation across the policy-plausible parameter space, reporting not only point estimates but rank distributions, interval widths, and tail probabilities. The finding that IE_IRPI rankings exhibit Spearman $\rho>0.95$ across 10,000 parameter draws demonstrates that the index is structurally stable rather than an artefact of specific modelling choices. This validation protocol, combining parameter perturbation with distributional reporting, offers a template for other composite indicator studies seeking to establish credibility.

To summarise, while the empirical application concerns innovation systems, the measurement-theoretic contributions extend to any domain requiring multi-dimensional assessment: the principle of enforced dimensional separation, the integration of outcome-based and frontier-based measures, data-driven indicator selection, the interpretive treatment of latent constructs, and systematic robustness validation. These contributions position the dissertation within the broader literature on index construction methodology, not merely as an application to innovation policy.

Returning to the dissertation's central objective, the measurement-theoretic contributions summarised above serve a specific analytical purpose: to construct a comprehensive measurement framework for assessing NIS performance in EU countries. The comprehensiveness of this framework is now empirically established. At the conceptual level, the framework maintains strict separation between result-based performance, efficiency, and capacity, dimensions that existing indices conflate.

At the methodological level, it integrates result-based composite indices with frontier-based efficiency analysis through MCFS-selected indicators and a transparent aggregation rule. At the empirical level, it has been validated across all 27 EU member states over 2017-2024, producing evidence on performance patterns (RQa), efficiency bottlenecks (RQb), and structural configurations (RQc) while confirming both hypotheses.

The near-zero correlation between IRPI and IEI constitutes one of the framework's key empirical validations: it demonstrates that result-based performance and transformation efficiency are not merely conceptually distinct but statistically independent dimensions of NIS functioning. This finding confirms that the comprehensive framework captures structural information that undifferentiated composite indices, such as the SII, systematically obscure.

Chapter 8 concludes the dissertation by synthesising these achievements, drawing out implications for EU innovation policy, and identifying directions for future research.

8 CONCLUSION

This concluding chapter synthesises the dissertation's purpose, findings, and contributions, and reflects on their implications for innovation policy and future research.

The dissertation addressed a central limitation of existing innovation benchmarks: the conflation of what innovation systems achieve with the conditions and processes through which achievements are generated. In response, it developed and applied a multidimensional measurement framework that distinguishes result-based innovation performance (IRPI), stage-specific innovation efficiency (IEI), and their integration (IE_IRPI), with innovation capacity serving as an interpretive lens rather than a measured dimension. The framework was empirically implemented for the EU-27 over 2017-2024.

The chapter is organised as follows: Section 8.1 restates the dissertation's purpose and scope. Section 8.2 summarises the key empirical findings, including the confirmation that commercialisation represents Europe's primary structural bottleneck and that result-based performance and efficiency are empirically independent dimensions. Section 8.3 draws out implications for innovation policy, particularly for EU benchmarking and funding instruments. Section 8.4 acknowledges limitations and identifies directions for future research. Section 8.5 provides concluding remarks on the dissertation's broader significance for innovation measurement and evidence-based policy design.

The objective of this dissertation was to develop a conceptually coherent and empirically robust framework for assessing the NIS performance in the EU. Motivated by long-standing critiques of existing innovation benchmarking tools, particularly the EIS, the research addressed the absence of a clear distinction between innovation outcomes and the processes through which these outcomes are generated. The central premise was that meaningful assessment of innovation systems requires simultaneous consideration of what innovation systems achieve and how efficiently they transform resources into those achievements. To address this challenge, the dissertation proposed and operationalised a multidimensional measurement framework integrating result-based innovation performance, two-stage innovation efficiency, and a performance-anchored aggregation mechanism. The framework was empirically implemented for the EU-27 over the period 2017-2024, enabling both cross-sectional and dynamic analysis of innovation systems.

Conceptually, the dissertation advances innovation systems research by enforcing a strict analytical separation between result-based innovation performance, innovation efficiency, and innovation capacity. While these concepts are well established in literature, they are rarely distinguished consistently in empirical measurement. Innovation performance was defined exclusively as result-based and in terms of observable outputs and outcomes, innovation efficiency as the productivity of transforming inputs into results across sequential stages, and innovation capacity as a latent structural property used for interpretation rather than direct measurement. This distinction resolves a central ambiguity in existing composite indicators, which often conflate determinants with results.

Methodologically, the dissertation makes three main contributions. First, it introduces MCFS as a transparent, data-driven approach to indicator selection in innovation measurement. This represents a methodological departure from expert-driven or precedent-based indicator choice and directly addresses redundancy, multicollinearity, and input dominance in existing frameworks. Second, it applies a two-stage DEA model to innovation systems, explicitly distinguishing between knowledge and technological production and commercialisation and diffusion. Third, it proposes the Efficiency-Adjusted Result-based Performance Index which integrates performance and efficiency using a performance-first aggregation rule that preserves outcome primacy while incorporating efficiency as a bounded modifier. Together, these elements form a unified measurement system that is theoretically grounded, empirically validated, and suitable for comparative and policy-relevant analysis.

The empirical analysis (Chapter 6) and discussion (Chapter 7) addressed the research questions and hypotheses formulated in Chapter 1. This section synthesises how each was resolved.

Research Questions

The main research question asked how National Innovation Systems differ in their innovation results and in the productivity with which they transform resources into those results, and what this variation reveals about cross-country heterogeneity. The three specific research questions explored distinct dimensions of this overarching inquiry.

RQa asked what patterns of variation exist in innovation results across EU member states. The IRPI results reveal substantial and persistent heterogeneity in result-based innovation performance. Countries range from frontier performers such as Sweden, Germany, Finland, and Ireland to lagging systems including Romania, Bulgaria, and Latvia, with limited convergence observed over the 2017-2024 period. The decomposition into knowledge production and commercialisation outcomes shows that these two stages vary independently: some countries achieve balanced performance across both stages, while others exhibit pronounced imbalances. This confirms that result-based innovation performance can be meaningfully measured as realised outputs and outcomes alone, without relying on inputs or institutional proxies (see

Section 7.2 for full interpretation).

RQb asked where efficiency bottlenecks concentrate in the sequential transformation of resources into innovation results. The IEI results demonstrate that commercialisation efficiency constitutes the binding constraint across most EU countries. Mean knowledge production efficiency (0.93) substantially exceeds mean commercialisation efficiency (0.81), with 18 of 27 countries showing higher Stage 1 than Stage 2 efficiency. This pattern confirms the European Paradox at the efficiency level: EU countries transform inputs into knowledge more productively than they transform knowledge into economic value. The two-stage framework reveals structural bottlenecks that remain systematically hidden in one-stage assessments (see Section 7.2 for full interpretation).

RQc asked what structural characteristics emerge when countries are grouped by both innovation results and transformation productivity together. The joint IRPI-IEI distribution reveals four distinct configurations aligned with theoretically predicted capacity types. High-performing and highly efficient systems, exemplified by Germany, the Netherlands, Finland, and Ireland, exhibit systemic coherence enabling both strong results and productive transformation. High-performing but less efficient systems, including Sweden, Denmark, and Belgium, achieve strong outcomes despite coordination challenges that limit transformation productivity. Lower-performing but efficient systems, such as Romania, Bulgaria, and Croatia, demonstrate productive resource use constrained by limited scale. Lower-performing and less efficient systems, including Spain, Lithuania, and Slovakia, face compound challenges requiring system-wide intervention. These groupings differ substantively from peer groups based on aggregate rankings alone, confirming that joint assessment yields more relevant benchmarking and peer identification (see Section 7.2 for full interpretation).

The joint IRPI-IEI distribution reveals four distinct configurations aligned with theoretically predicted capacity types in Table 40.

Table 40: Configurations and characteristics

Configuration	Characteristics	Countries (2024)
High-High	Strong outcomes, efficient transformation	DE, NL, FI, IE
High-Low	Strong outcomes, inefficient transformation	SE, DK, BE
Low-High	Modest outcomes, efficient transformation	RO, BG, HR
Low-Low	Weak outcomes, inefficient transformation	ES, LT, SK

Source: Own

These groupings differ substantively from peer groups based on aggregate rankings alone, confirming that combining result-based performance and efficiency without collapsing them yields more relevant benchmarking and peer identification (see Section 7.2 for full interpretation).

The main research question is thus answered: cross-country variation in results and productivity reveals that NIS heterogeneity is structural rather than superficial. Countries differ systematically in both what they achieve and how they achieve it, and these differences align with underlying capacity configurations that single-dimension rankings obscure.

Hypotheses

Hypothesis 1 proposed that a multidimensional framework integrating result-based innovation performance and innovation efficiency, interpreted through an innovation capacity perspective, provides a more comprehensive and diagnostically informative assessment of NIS than existing frameworks. The empirical evidence strongly supports this hypothesis through three validation criteria established in Chapter 1. The near-zero correlation between IRPI and IEI (r=0.03) confirms that performance and efficiency capture empirically distinct dimensions of NIS functioning. The systematic stage-specific asymmetry, with 67% of countries showing higher knowledge production than commercialisation efficiency, confirms that two-stage decomposition reveals bottleneck patterns invisible in aggregate measures. The four-quadrant performance-efficiency distribution shows meaningful country configurations aligned with capacity-related characteristics, confirming that joint assessment is diagnostically informative. No falsification threshold was triggered, and Hypothesis 1 is supported.

Hypothesis 2 proposed that innovation indices based on the multidimensional framework enable more accurate identification of peer countries for benchmarking and policy learning. The empirical evidence supports this hypothesis through three validation criteria. Monte Carlo simulation confirms ranking robustness, with Spearman correlations exceeding 0.95 across the policy-plausible parameter range, demonstrating that IE_IRPI rankings reflect persistent structural relationships rather than arbitrary parameter choices. External validation confirms construct validity, with IE_IRPI correlating positively with GDP per capita and government effectiveness as theoretically expected. Comparison with SII rankings reveals substantive differentiation, with 44% of countries experiencing tier reclassification, indicating that IE_IRPI provides diagnostic value beyond existing benchmarks. No falsification threshold was triggered, and Hypothesis 2 is supported.

The support for both hypotheses validates the conceptual framework developed in Chapter 4 and confirms that the measurement architecture achieves its intended purpose: providing a comprehensive and diagnostically informative basis for assessing NIS performance across EU member states.

The following findings synthesise the conceptual, methodological, and empirical insights developed across the preceding chapters.

- *Result-based performance and efficiency are empirically independent.* The near-zero correlation between IRPI and IEI provides strong empirical confirmation that these dimensions capture distinct structural properties of NIS, validating the theoretical separation and supporting calls for joint, but non-conflated, assessment (see Section 7.1.2).
- *Commercialisation is the binding constraint.* Across the EU-27, mean KPEI exceeds mean KCEI, confirming the "European Paradox" at the efficiency level and demonstrating that aggregate efficiency scores obscure where bottlenecks arise (see Section 7.1.3).
- *Four distinct performance-efficiency configurations emerge.* The joint distribution reveals systematic country groupings aligned with capacity theory, enabling peer identification based on structural similarity rather than rank proximity (see Section 7.1.4).
- *Knowledge production is structurally stronger than commercialisation in the EU.* Across the EU-27, the 2024 results consistently show higher levels and efficiency in knowledge production (IKP, KPEI) than in commercialisation (IKC, KCEI). This pattern confirms long-standing findings in the innovation systems literature that Europe's core weakness lies not in research generation but in diffusion, scale-up, and market exploitation (Pavitt, 1991; European Commission, 1995; Draghi 2024). Empirical DEA studies similarly report that fewer countries reach the frontier in downstream innovation stages than in upstream knowledge creation (Guan and Chen, 2012; Carayannis et al., 2016; Anouze et al., 2024). The results reinforce this diagnosis using a full EU-27 sample and a stage-consistent indicator framework.
- *Limited convergence persists despite decades of cohesion policy.* The pronounced polarisation observed across IRPI, IEI, and IE_IRPI aligns with earlier evidence that innovation convergence within the EU has been limited, despite decades of cohesion and innovation policy (Hollanders et al., 2024; Radosevic, 2017). Bibliometric and efficiency-based studies repeatedly identify stable leader and laggard groups (Filippetti and Peyrache, 2011; Castellacci and Natera, 2013). The research results extend this literature by showing that limited convergence holds even after adjusting outcomes for efficiency, suggesting that structural differences in innovation systems remain deeply entrenched.
- *Frontier countries exhibit systemic coherence rather than a single success model.* Countries at the innovation frontier in 2024 (e.g. Sweden, Germany, Finland, the Netherlands, Ireland, Denmark) share the common feature that neither insufficient scale nor severe inefficiency constrains outcomes. This finding resonates with the concept of systemic coherence in innovation systems theory, where strong institutions, dense networks, and complementary policies allow multiple innovation functions to operate effectively (Lundvall, 1992, 2007).

The coexistence of balanced systems (e.g. Sweden, Finland) and commercially dominant systems (e.g. Ireland) further support the view that there is no single model of innovation success, only different viable structural configurations (Nelson, 1993; Fagerberg and Srholec, 2008).

- *High-performing countries can sustain outcomes through input intensity despite inefficiency.* Several high-performing countries exhibit strong IRPI despite below-average IEI, indicating that innovation outcomes are sustained through high input intensity rather than particularly efficient transformation processes. This pattern has been documented in earlier studies, which show that large, resource-rich systems can compensate for inefficiencies through scale (Guan and Chen, 2012; Edquist et al., 2018). The analysis results corroborate this insight and demonstrate that even innovation leaders retain substantial scope for improving efficiency, particularly in commercialisation stages.
- *Efficient but low-performing systems illustrate the limits of relative efficiency metrics.* Countries with high IEI but low IRPI illustrate that efficiency reflects relative productivity, not absolute success. This distinction has been emphasised in DEA-based innovation studies, where catching-up economies often appear efficient because they operate close to a frontier defined by peers with similar resource endowments (Bogetoft and Otto, 2011; Hudec and Prochádzková, 2013). The findings reinforce the argument that efficiency metrics must be interpreted diagnostically and anchored in outcome measures to avoid misleading policy conclusions.
- *Compound challenges require system-wide intervention.* Countries that score below average on both result-based performance and efficiency face compound structural challenges affecting education systems, firm capabilities, finance, and diffusion mechanisms. The literature stresses that such cases require coordinated, system-wide interventions rather than isolated policy measures (Rodríguez-Pose and Crescenzi, 2008; Radosevic, 2017). The joint IRPI-IEI framework makes this condition empirically visible and distinguishes it from cases where problems are limited to either scale or efficiency.
- *IE_IRPI successfully integrates efficiency without overturning outcome-based rankings.* The IE_IRPI results demonstrate that efficiency can be incorporated into innovation assessment without overturning outcome-based rankings. This addresses a long-standing critique of composite indicators, which either ignore efficiency or embed it implicitly through inputs (Nardo et al., 2008; Edquist, 2011). By preserving the primacy of realised outcomes while penalising persistent inefficiency, IE_IRPI advances both methodological rigor and policy interpretability relative to the SII and similar indices.

Supported by both innovation systems theory and comparative empirical research, the 2024 results confirm that Europe's innovation challenge is structural, stage-specific, and unevenly distributed across countries. Knowledge creation is not the primary constraint; rather, the translation of knowledge into scalable economic value remains fragmented and inefficient. By jointly analysing result-based performance and efficiency across stages, and by integrating these dimensions through IE_IRPI, the dissertation provides an empirically grounded and theoretically consistent framework that moves innovation assessment beyond headline rankings toward diagnosis of systemic bottlenecks, in line with the recommendations of the innovation policy literature (Patel and Pavitt, 1994; Edquist, 2011).

The findings have important implications for innovation policy and benchmarking in the EU. They demonstrate that headline rankings based on composite scores are insufficient for diagnosing innovation system weaknesses and may lead to inappropriate policy conclusions. By contrast, the multidimensional framework developed in this dissertation enables policymakers to distinguish between scale problems, efficiency problems, and stage-specific bottlenecks.

For high-performing but inefficient systems, policy priorities should focus on improving system effectiveness and reducing inefficiencies, particularly in commercialisation. For efficient but low-performing systems, the central challenge lies in expanding the innovation input base and scaling successful activities. For countries underperforming on both dimensions, coordinated, system-wide reforms are required across education, finance, firm capabilities, and diffusion mechanisms. By enabling peer-group identification based on structural similarity rather than rank proximity, the framework also enhances the potential for meaningful policy learning at both national and EU levels.

The empirical findings suggest that EU innovation policy would benefit from a clearer differentiation between scale-oriented capacity building and efficiency-oriented performance improvement:

- *Horizon Europe* programmes should be assessed based on efficiency-adjusted outcomes rather than absolute outputs alone. For EIT RIS countries that rank highly on IE_IRPI but modestly on conventional indicators, evaluation criteria could reward them through differentiated co-funding rates or targeted support for scaling proven research outputs.

- *Cohesion Policy and Structural Funds* should explicitly incorporate efficiency diagnostics when allocating innovation-related funding. In systems where low efficiency rather than resource scarcity is the dominant constraint, conditionalities linked to institutional quality and coordination capacity could be prioritised over further input expansion.

The results support a modular redesign of the EIS in which result-based performance and efficiency are reported as analytically distinct but complementary dimensions.

Incorporating an efficiency-adjusted performance layer would improve diagnostic value by enabling policymakers to distinguish between innovation systems that underperform due to insufficient resources and those that underperform due to systemic inefficiencies. Such a redesign would not replace the SII but would contextualise it, reducing the risk of policy prescriptions that conflate scale with effectiveness.

Despite the robustness of the proposed framework, the following key limitations delimit the scope of the empirical results and point to promising directions for future research.

First, innovation efficiency is inherently relative and dependent on the selected sample. The IEI is based on DEA, which evaluates efficiency relative to the best-performing units within the observed sample. Consequently, efficiency scores depend on the composition of countries included and may change if additional units or alternative comparators are introduced (Bogetoft and Otto, 2011). This relative nature explains why some low-performing countries may appear efficient when compared to peers with similarly constrained input structures. While the performance-adjusted aggregation adopted in Chapter 6 mitigates this issue by anchoring the composite index in observed outcomes, the sample dependence of frontier-based efficiency remains a structural limitation. Future studies could extend the framework by employing meta-frontier or global frontier approaches that allow efficiency comparisons across heterogeneous groups, or by constructing time-consistent frontiers using dynamic DEA or Malmquist productivity indices to assess efficiency changes over time.

Second, the composite indicator construction involves normative weighting choices at two levels. The equal weighting of indicators within subindices and of subindices within IRPI constitutes a methodological decision justified by theoretical agnosticism, MCFS-based pre-selection, and transparency objectives, but alternative schemes reflecting different value judgments would produce different, though likely correlated, rankings. Further, the efficiency adjustment parameter ($\lambda=0.25$) in IE_IRPI embodies normative assumptions about the relative importance of efficiency versus realised result-based performance. While λ is theoretically bounded and subjected to extensive robustness testing (Monte Carlo simulation confirms rank correlations exceeding $\rho>0.95$ across the policy-plausible range), it is not empirically estimated. This subjectivity is inherent to all composite indicator construction involving weighting or aggregation rules and cannot be fully eliminated (Nardo et al., 2008). Future research could explore data-driven approaches to parameter calibration, such as anchoring weights to policy objectives, stakeholder preferences, or empirical loss functions, or develop scenario-based indices that explicitly report results under alternative normative assumptions.

Third, aggregation entails unavoidable information loss. Both the IRPI and IEI compress complex, multidimensional innovation processes into scalar indices. Even with a hierarchical, performance-first design, aggregation obscures heterogeneity across indicators, sectors, and institutional contexts. As a result, the composite index should be interpreted as a synthetic summary rather than a standalone diagnostic tool, consistent with Edquist's (2011) caution regarding overly aggregated innovation metrics. Future work could integrate the composite framework with interactive dashboards, cluster-based typologies, or machine-learning techniques that preserve multidimensional structure while maintaining comparability, enabling richer diagnostics without sacrificing transparency.

Fourth, cross-country comparability is constrained by structural heterogeneity. NIS differ substantially in size, sectoral composition, institutional architecture, and integration into global value chains. Although normalisation and relative efficiency methods improve comparability, the framework cannot fully control for these deep structural differences. Consequently, rankings and peer-group classifications should be interpreted as indicative rather than definitive, particularly when comparing countries with fundamentally different innovation models (OECD, 2018). Future studies could apply the framework to more homogeneous sub-samples, such as regional innovation systems, sector-specific innovation systems, or country clusters defined by development stage, and could explicitly incorporate structural variables as conditioning factors in comparative analysis.

Fifth, a methodological limitation concerns the inclusion of the COVID-19 crisis period (2020-2021) within the analysis timeframe. While temporal stability analysis confirms that aggregate findings are robust, seven countries exhibited crisis-period KCEI peaks that subsequently reversed, and four countries reached the efficiency frontier exclusively during crisis years. Future research extending the observation period beyond 2024 would enable more definitive separation of structural efficiency characteristics from residual crisis effects.

Sixth, the two-stage DEA model implements a fixed one-year lag per stage, operationalising a two-year aggregate transformation period. While this specification aligns with prior DEA-based innovation studies and ensures model parsimony, the innovation systems literature documents considerable heterogeneity in actual transformation lags across firms, sectors, and institutional contexts (Griliches, 1990; Mansfield, 1991; Markman et al., 2005). The implemented two-year lag represents a lower-bound estimate that may understate the full transformation period for basic research and complex technologies, while potentially overstating lags for incremental innovations. Future research could examine sensitivity to alternative lag specifications or implement sector-differentiated lag structures where data permit, and with explicit sensitivity analysis across lag lengths, orientations, and frontier groupings.

Seventh, constructed indices as DEA outputs. The efficiency analysis uses the aggregated subindices IKP and IKC as DEA output variables rather than raw indicators. This specification ensures favourable DMU-to-variable ratios and maintains theoretical alignment with the two-stage framework, but it embeds the equal-weighting assumption from subindex construction into efficiency measurement. Future research could explore alternative aggregation approaches, including benefit-of-the-doubt weighting within DEA or multi-output specifications with endogenous indicator weights.

Future research should avoid: (1) interpreting DEA efficiency scores as absolute measures or comparing them across different country samples, since the frontier is sample-specific; (2) calibrating the efficiency parameter λ to maximise statistical fit or achieve desired rankings, which would undermine outcome primacy; (3) over-interpreting small composite score differences that fall within Monte Carlo confidence intervals; and (4) using rankings to prescribe policy convergence toward top-ranked country profiles without accounting for legitimate structural differences. These boundaries ensure the framework serves as a diagnostic tool rather than a deterministic ranking mechanism.

These limitations underscore that the proposed composite indicator should not be interpreted as a definitive measure of innovation "success." Rather, its primary value lies in its role as a structured, performance-anchored analytical tool that supports comparative diagnosis. By revealing systematic patterns, stage-specific bottlenecks, and performance-efficiency mismatches, the framework enhances understanding of NIS functioning while avoiding claims of mechanical precision or exhaustive evaluation.

This dissertation set out to develop a conceptually coherent and empirically robust framework for assessing the NIS performance in the EU. The research was motivated by a fundamental critique: existing benchmarks conflate what innovation systems achieve with the conditions and processes through which achievements are generated, limiting their diagnostic value for policy.

The framework developed in response maintains strict separation between result-based performance, efficiency, and capacity while enabling their joint interpretation. The empirical application to the EU-27 over 2017-2024 addressed three research questions and confirmed both hypotheses:

- RQa: Result-based innovation performance can be meaningfully measured as realised outputs and outcomes, revealing substantial and persistent cross-country heterogeneity.
- RQb: Commercialisation efficiency constitutes the binding constraint, confirming the European Paradox at the efficiency level.
- RQc: Joint performance-efficiency assessment yields four distinct configurations with differentiated policy implications.

- H1: The multidimensional framework provides diagnostically superior assessment through dimensional independence, stage-specific decomposition, and capacity-aligned configurations.
- H2: IE_IRPI enables accurate peer identification through stable rankings and structurally coherent groupings.

The central message is that Europe's innovation challenge is structural, stage-specific, and heterogeneous. Knowledge creation is not the primary constraint, rather, the translation of knowledge into scalable economic value remains fragmented and inefficient. By jointly analysing result-based performance and efficiency across stages, and by integrating these dimensions through IE_IRPI as a NIS performance measure, the dissertation advances innovation assessment beyond one-dimensional benchmarking toward structured diagnosis of systemic strengths and weaknesses.

The proposed framework provides a theoretically consistent and empirically validated basis for comparative analysis and evidence-based innovation policy. More broadly, it contributes to ongoing efforts in innovation studies to move from descriptive rankings toward analytically grounded tools capable of informing strategic decision-making in increasingly complex and heterogeneous innovation environments.

EXTENDED ABSTRACT

This dissertation develops a conceptually coherent and empirically robust framework for assessing the performance of National Innovation Systems (NIS) in the European Union (EU). Innovation has long been recognised as a central driver of long-run growth, productivity, and structural renewal, and innovation policy has accordingly become a strategic priority of the EU, articulated through successive frameworks such as the Lisbon Agenda, Innovation Union, Horizon 2020, and Horizon Europe. As the public budgets devoted to research and innovation continue to expand, the question facing policymakers is no longer whether to intervene, but how to intervene effectively. This shift requires measurement instruments capable of providing diagnostic insight into where innovation systems succeed, where they break down, and which interventions are most likely to improve results. The European Innovation Scoreboard (EIS) and similar composite tools have become the dominant references for cross-country assessment, but the literature has repeatedly documented their conceptual and methodological limitations. Their central weakness is the conflation of inputs, framework conditions, institutional determinants, and realised innovation outcomes within a single aggregated score, which obscures whether observed differences reflect favourable structural endowments or genuinely effective transformation processes. As a result, existing benchmarks support ranking but offer limited diagnostic and policy relevance.

The research gap addressed by this dissertation is therefore not the absence of benchmarking tools, but the absence of an integrated, multidimensional measurement framework that mirrors the internal logic of innovation systems. Innovation systems theory and evolutionary economics conceptualise innovation as a sequential, cumulative, and institutionally embedded process in which inputs are converted into knowledge and technological outputs, and these outputs are subsequently transformed into commercial and economic outcomes. Yet prevailing measurement practice collapses these distinct phenomena into composite indices or analyses them in isolation through input-output efficiency models. The dissertation argues that meaningful assessment requires the simultaneous but analytically separated consideration of three dimensions: innovation performance understood strictly as observable results, innovation efficiency understood as productivity of transformation across stages, and innovation capacity understood as a latent systemic property that conditions both. Maintaining strict boundaries between these dimensions at the measurement stage, while allowing them to interact at the level of interpretation, is the central conceptual move of the dissertation.

Building on this premise, the dissertation pursues three specific research questions and tests two hypotheses. The research questions ask: (RQa) what patterns of variation exist in innovation results across EU countries and whether systems balance

knowledge production and commercialisation symmetrically; (RQb) where bottlenecks concentrate in the sequential transformation of resources into innovation results, and whether the so-called European Paradox holds at the efficiency level; and (RQc) what structural configurations emerge when countries are grouped jointly by results and transformation productivity, and whether these configurations differ from peer groups based on aggregate rankings. The first hypothesis (H1) predicts that a multidimensional framework integrating result-based performance and efficiency, interpreted through a capacity perspective, provides a more comprehensive and diagnostically informative assessment of NIS than existing frameworks. The second hypothesis (H2) predicts that indices built on this framework enable more accurate identification of peer countries for benchmarking and policy learning. Each hypothesis is paired with explicit falsification criteria, ensuring that the framework is tested against conditions under which it could be disconfirmed.

Methodologically, the dissertation operationalises the conceptual framework through three integrated measurement instruments. First, indicator selection is conducted through a Multi-Cluster Feature Selection (MCFS) procedure, which provides a transparent, replicable, and data-driven alternative to expert-driven or precedent-based indicator choice. MCFS operates strictly within theoretically defined categories of inputs, outputs, and outcomes, addressing redundancy and multicollinearity without sacrificing interpretability. Second, result-based innovation performance is measured through the Innovation Result-based Performance Index (IRPI), which captures observable knowledge and technological outputs and commercial outcomes only, deliberately excluding inputs and institutional determinants. Third, innovation efficiency is assessed through the Innovation Efficiency Index (IEI), derived from a two-stage Data Envelopment Analysis (DEA) model that mirrors the sequential structure of innovation processes by separately evaluating efficiency in knowledge production and in commercialisation. The two indices are subsequently integrated through the Efficiency-Adjusted Result-based Performance Index (IE_IRPI), built on a performance-first aggregation rule in which efficiency acts as a bounded modifier rather than an equal-weight component. The framework is empirically applied to all twenty-seven EU member states over the period 2017-2024, enabling both cross-sectional and dynamic analysis. Innovation capacity is treated throughout as an interpretive lens that informs the explanation of persistent patterns rather than as a directly measured dimension, thereby preserving analytical separation and avoiding the identification problem associated with capacity proxies.

The empirical analysis confirms substantial and persistent structural heterogeneity across European NIS. With respect to RQa, the IRPI reveals pronounced variation in result-based performance, ranging from frontier countries such as Sweden, Germany, Finland, and Ireland to lagging systems such as Romania, Bulgaria, and Latvia, with limited convergence over the observation period. With respect to RQb, mean knowledge production efficiency (0.93) substantially exceeds mean commercialisation efficiency (0.81), and 18 of 27 countries display higher Stage 1 than Stage 2 efficiency. This pattern confirms the European Paradox at the efficiency level: EU countries are more productive at converting inputs into knowledge than at converting knowledge into economic value. With respect to RQc, the joint IRPI-IEI distribution generates four distinct configurations aligned with theoretically anticipated capacity types: high-performing and efficient systems (e.g. Germany, Netherlands, Finland, Ireland), high-performing but less efficient systems (e.g. Sweden, Denmark, Belgium), lower-performing but efficient systems (e.g. Romania, Bulgaria, Croatia), and lower-performing and inefficient systems (e.g. Spain, Lithuania, Slovakia). These groupings differ substantively from peer groups based on the EIS Summary Innovation Index, with roughly 44% of countries reclassified across tiers.

These results provide robust support for both hypotheses. Hypothesis 1 is confirmed by three converging lines of evidence: the near-zero correlation between IRPI and IEI ($r = 0.03$) indicates that performance and efficiency capture empirically distinct rather than redundant aspects of NIS functioning; the systematic asymmetry between knowledge production and commercialisation efficiency demonstrates that two-stage decomposition exposes bottleneck patterns invisible in aggregate measures; and the four-quadrant distribution reveals theoretically meaningful country configurations. Hypothesis 2 is confirmed by Monte Carlo simulation, which shows ranking robustness with Spearman correlations exceeding 0.95 across the policy-plausible parameter space, by external validation against GDP per capita and government effectiveness, and by the substantive differentiation of IE_IRPI peer groups from those produced by aggregate rankings. No falsification threshold was triggered for either hypothesis.

The dissertation contributes on three connected levels. Conceptually, it advances innovation systems research by enforcing strict analytical separation between performance, efficiency, and capacity, dimensions that are widely acknowledged in theory but rarely distinguished in empirical measurement. Methodologically, it offers a unified architecture combining data-driven indicator selection through MCFS, theoretically anchored two-stage DEA, and a performance-first aggregation rule that integrates outcome-based and frontier-based measures without allowing one to override the other. These methodological principles also contribute to the broader literature on composite indicator construction, demonstrating how dimensional separation, latent-construct treatment, and systematic robustness validation can be applied across policy domains. Empirically, the application to the EU-27 yields evidence that is not available from existing benchmarking instruments, including the empirical independence of performance and efficiency, the localisation of structural bottlenecks in commercialisation, and the existence of stable structural configurations of NIS rather than convergence around the EU average. However, the persistence of stage-specific asymmetries and the structural stability of country configurations

highlight the need for further research into the institutional and capacity-related mechanisms that sustain commercialisation bottlenecks across different innovation system types.

In policy terms, the framework moves innovation assessment beyond ranking toward structural diagnosis. It enables policymakers to distinguish scale problems from efficiency problems and to identify stage-specific bottlenecks. For high-performing but inefficient systems, policy priorities should focus on improving transformation processes, particularly in commercialisation. For efficient but lower-performing systems, the central challenge is expanding the innovation input base and scaling successful activities. For systems underperforming on both dimensions, coordinated, system-wide reforms across education, finance, firm capabilities, and diffusion mechanisms are required. The framework also supports a modular redesign of the EIS in which result-based performance and efficiency are reported as analytically distinct but complementary dimensions and provides an evidence base for differentiating Horizon Europe and Cohesion Policy instruments according to system type rather than aggregate rank. The central message is that Europe's innovation challenge is structural, stage-specific, and unevenly distributed: knowledge creation is not the binding constraint, but the translation of knowledge into scalable economic value remains fragmented and inefficient. By jointly analysing result-based performance and efficiency across stages and integrating them through a performance-anchored composite index, the dissertation provides a theoretically grounded and empirically validated foundation for evidence-based innovation policy and structurally meaningful peer learning.

Keywords: Innovation policy, innovation performance, innovation efficiency, innovation capacity, national innovation system, innovation index, Data Envelopment Analysis.

DALJŠI POVZETEK

CELOVIT OKVIR MERJENJA ZA OCENJEVANJE USPEŠNOSTI NACIONALNIH INOVACIJSKIH SISTEMOV V DRŽAVAH EU

Ta disertacija razvija konceptualno koherenten in empirično robusten okvir za ocenjevanje uspešnosti nacionalnih inovacijskih sistemov (NIS) v Evropski uniji (EU). Inovacije so že dolgo prepoznane kot osrednji dejavnik dolgoročne rasti, produktivnosti in strukturne obnove, inovacijska politika pa je zato postala strateška prednostna naloga EU, izražena v zaporednih okvirih, kot so Lizbonska agenda, Unija inovacij, Horizon 2020 in Horizon Europe. Ker se javni proračuni, namenjeni raziskavam in inovacijam, še naprej širijo, se oblikovalci politik ne soočajo več s vprašanjem, ali naj posredujejo, temveč kako učinkovito posredovati. Ta premik zahteva merilne instrumente, ki lahko zagotovijo diagnostični vpogled v to, kje inovacijski sistemi uspevajo, kje se porušijo in kateri posegi najverjetneje izboljšajo rezultate. Evropski sistem inovacijskih kazalnikov (EIS) in podobna sestavljena orodja so postala prevladujoča referenca za meddržavno ocenjevanje, vendar je literatura večkrat dokumentirala njihove konceptualne in metodološke omejitve. Njihova osrednja slabost je združevanje vložkov, okvirnih pogojev, institucionalnih determinant in doseženih inovacijskih rezultatov v enem samem agregiranem rezultatu, kar zakriva, ali opažene razlike odražajo ugodne strukturne danosti ali resnično učinkovite procese preobrazbe. Posledično obstoječa merila podpirajo razvrščanje, vendar ponujajo omejeno diagnostično in politično relevantnost.

Raziskovalna vrzel, ki jo obravnava ta disertacija, torej ni odsotnost orodij za primerjalno analizo, temveč odsotnost integriranega, večdimenzionalnega okvira za merjenje, ki bi odražal notranjo logiko inovacijskih sistemov. Teorija inovacijskih sistemov in evolucijska ekonomija konceptualizirata inovacije kot zaporedni, kumulativni in institucionalno vgrajeni proces, v katerem se vhodni podatki pretvorijo v znanje in tehnološke rezultate, ti rezultati pa se nato pretvorijo v komercialne in gospodarske rezultate. Vendar pa prevladujoča praksa merjenja te različne pojave strne v sestavljene indekse ali jih analizira ločeno z modeli učinkovitosti vhodnih in izhodnih podatkov. Disertacija trdi, da smiselna ocena zahteva sočasno, a analitično ločeno upoštevanje treh dimenzij: inovacijske uspešnosti, razumljene strogo kot opazovani rezultati, inovacijske učinkovitosti, razumljene kot produktivnost preobrazbe v različnih fazah, in inovacijske zmogljivosti, razumljene kot latentne sistemske lastnosti, ki pogojuje obe. Osrednja konceptualna poteza disertacije je ohranjanje strogih meja med temi dimenzijami v fazi merjenja, hkrati pa jim omogočanje interakcije na ravni interpretacije.

Na podlagi te predpostavke disertacija zasleduje tri specifična raziskovalna vprašanja in preizkuša dve hipotezi. Raziskovalna vprašanja sprašujejo: (RQa) kakšni vzorci variacij obstajajo v rezultatih inovacij v državah EU in ali sistemi simetrično uravnavajo proizvodnjo in komercializacijo znanja; (RQb) kje se ozka grla koncentrirajo pri zaporedni transformaciji virov v rezultate inovacij in ali tako imenovani evropski paradoks velja na ravni učinkovitosti; in (RQc) kakšne strukturne konfiguracije se pojavijo, ko so države združene glede na rezultate in produktivnost transformacije, in ali se te konfiguracije razlikujejo od skupin enakovrednih držav na podlagi skupnih uvrstitev. Prva hipoteza (H1) napoveduje, da večdimenzionalni okvir, ki združuje uspešnost in učinkovitost, ki temeljita na rezultatih, interpretiran z vidika zmogljivosti, zagotavlja bolj celovito in diagnostično informativno oceno NIS kot obstoječi okviri. Druga hipoteza (H2) napoveduje, da indeksi, zgrajeni na tem okviru, omogočajo natančnejšo identifikacijo enakovrednih držav za primerjalno analizo in učenje politik. Vsaka hipoteza je povezana z eksplicitnimi merili za ponarejanje, kar zagotavlja, da se okvir preizkusi glede na pogoje, pod katerimi bi ga lahko ovrgli.

Metodološko disertacija operacionalizira konceptualni okvir s tremi integriranimi merilnimi instrumenti. Prvič, izbira kazalnikov se izvaja s postopkom izbire večklasterskih značilnosti (MCFS), ki zagotavlja pregledno, ponovljivo in na podatkih temelječo alternativo izbiri kazalnikov, ki jo vodijo strokovnjaki ali precedensi. MCFS deluje strogo znotraj teoretično opredeljenih kategorij vhodnih podatkov, izhodnih podatkov in rezultatov, pri čemer obravnava redundanco in multikolinearnost, ne da bi pri tem žrtvoval možnost interpretacije. Drugič, inovacijska uspešnost, ki temelji na rezultatih, se meri z indeksom inovacijske uspešnosti, ki temelji na rezultatih (IRPI), ki zajema le opazno znanje in tehnološke izhodne podatke ter komercialne rezultate, pri čemer namerno izključuje vhodne podatke in institucionalne determinante. Tretjič, inovacijska učinkovitost se ocenjuje z indeksom inovacijske učinkovitosti (IEI), ki izhaja iz dvostopenjskega modela analize ovojnosti podatkov (DEA), ki odraža zaporedno strukturo inovacijskih procesov z ločenim ocenjevanjem učinkovitosti pri proizvodnji znanja in komercializaciji. Oba indeksa se nato integrirata z indeksom uspešnosti, prilagojenim rezultatom (IE_IRPI), ki temelji na pravilu agregacije, ki najprej upošteva uspešnost, pri čemer učinkovitost deluje kot omejeni modifikator in ne kot komponenta z enako težo. Okvir se empirično uporablja za vseh sedemindvajset držav članic EU v obdobju 2017-2024, kar omogoča tako presečno kot dinamično analizo. Inovacijska zmogljivost se ves čas obravnava kot interpretativna leča, ki vpliva na razlago trajnih vzorcev, in ne kot neposredno izmerjena dimenzija, s čimer se ohranja analitično ločevanje in se izognemo problemu identifikacije, povezanemu s približki zmogljivosti.

Empirična analiza potrjuje znatno in vztrajno strukturno heterogenost med evropskimi NIS. Kar zadeva RQa, IRPI razkriva izrazite razlike v uspešnosti, ki temelji na rezultatih, od držav na robu razvoja, kot so Švedska, Nemčija, Finska in Irska, do zaostalih sistemov, kot so Romunija, Bolgarija in Latvija, z omejeno konvergenco v opazovanem obdobju. Kar zadeva RQb, povprečna učinkovitost proizvodnje znanja (0,93) bistveno presega povprečno učinkovitost komercializacije (0,81), 18 od 27 držav pa kaže višjo učinkovitost na 1. stopnji kot na 2. stopnji. Ta vzorec potrjuje evropski paradoks na ravni učinkovitosti: države EU so bolj produktivne pri pretvarjanju vložkov v znanje kot pri pretvarjanju znanja v ekonomsko vrednost. Glede na RQc skupna porazdelitev IRPI-IEI ustvari štiri različne konfiguracije, usklajene s teoretično pričakovanimi tipi zmogljivosti: visoko zmogljivi in učinkoviti sistemi (npr. Nemčija, Nizozemska, Finska, Irska), visoko zmogljivi, a manj učinkoviti sistemi (npr. Švedska, Danska, Belgija), manj zmogljivi, a učinkoviti sistemi (npr. Romunija, Bolgarija, Hrvaška) in manj zmogljivi in neučinkoviti sistemi (npr. Španija, Litva, Slovaška). Te skupine se bistveno razlikujejo od skupin primerljivih sistemov na podlagi indeksa inovacij EIS Summary Innovation Index, saj je bilo približno štiriinštirideset odstotkov držav prerazvrščenih med stopnjami.

Ti rezultati zagotavljajo trdno podporo obema hipotezama. Hipotezo 1 potrjujejo trije konvergentni dokazi: skoraj ničelna korelacija med IRPI in IEI (r = 0,03) kaže, da uspešnost in učinkovitost zajemata empirično različne in ne odvečne vidike delovanja NIS; sistematična asimetrija med proizvodnjo znanja in učinkovitostjo komercializacije kaže, da dvostopenjska razgradnja razkriva vzorce ozkih grl, ki so nevidni v agregatnih merah; štirikvadrantna porazdelitev pa razkriva teoretično smiselne konfiguracije držav. Hipotezo 2 potrjujejo Monte Carlo simulacija, ki kaže robustnost uvrstitve s Spearmanovimi korelacijami, ki presegajo 0,95 v celotnem prostoru parametrov, ki so politično verjetni, zunanja validacija glede na BDP na prebivalca in učinkovitost vlade ter vsebinska diferenciacija skupin primerljivk IE_IRPI od tistih, ki jih tvorijo agregatne uvrstitve. Za nobeno od hipotez ni bil sprožen prag ponarejanja.

Disertacija prispeva na treh povezanih ravneh. Konceptualno napreduje pri raziskavah inovacijskih sistemov z uveljavljanjem stroge analitične ločitve med uspešnostjo, učinkovitostjo in zmogljivostjo, dimenzijami, ki so v teoriji splošno priznane, a se le redko razlikujejo v empiričnih meritvah. Metodološko ponuja enotno arhitekturo, ki združuje izbiro kazalnikov, ki temeljijo na podatkih, prek MCFS, teoretično zasidranega dvostopenjskega DEA in pravilo agregacije, ki je na prvem mestu uspešnost, ki združuje meritve, ki temeljijo na rezultatih in mejah, ne da bi eno preglasilo drugo. Ta metodološka načela prispevajo tudi k širši literaturi o konstrukciji sestavljenih kazalnikov in prikazujejo, kako se lahko dimenzijska ločitev, obravnava latentnih konstruktov in sistematično preverjanje robustnosti uporabljajo na različnih področjih politike. Empirično uporaba za EU-27 daje dokaze, ki niso na voljo iz obstoječih instrumentov za primerjalno analizo, vključno z empirično neodvisnostjo uspešnosti in učinkovitosti, lokalizacijo strukturnih ozkih grl pri komercializaciji in obstojem stabilnih strukturnih konfiguracij NIS namesto konvergence okoli povprečja EU. Vendar pa vztrajanje asimetrij, specifičnih za posamezne faze, in strukturna stabilnost konfiguracij držav poudarjata potrebo po nadaljnjih raziskavah institucionalnih in z zmogljivostjo

povezanih mehanizmov, ki vzdržujejo ozka grla pri komercializaciji v različnih vrstah inovacijskih sistemov.

V političnem smislu okvir premika ocenjevanje inovacij od razvrščanja k strukturni diagnozi. Omogoča oblikovalcem politik, da ločijo težave z obsegom od težav z učinkovitostjo in prepoznajo ozka grla, specifična za posamezne faze. Za visoko učinkovite, a neučinkovite sisteme bi se morale prednostne naloge politike osredotočiti na izboljšanje procesov preoblikovanja, zlasti pri komercializaciji. Za učinkovite, a manj učinkovite sisteme je osrednji izziv širitev baze inovacijskih vložkov in skaliranje uspešnih dejavnosti. Za sisteme, ki v obeh dimenzijah ne dosegajo želenih rezultatov, so potrebne usklajene, sistemske reforme na področju izobraževanja, financ, zmogljivosti podjetij in mehanizmov difuzije. Okvir podpira tudi modularno preoblikovanje EIS, v katerem se o uspešnosti in učinkovitosti, ki temeljita na rezultatih, poroča kot o analitično ločenih, a dopolnilnih dimenzijah, ter zagotavlja dokazno podlago za razlikovanje instrumentov programa Horizon Europe in kohezijske politike glede na vrsto sistema in ne glede na skupno uvrstitev. Osrednje sporočilo je, da je evropski inovacijski izziv strukturen, specifičen za posamezne faze in neenakomerno porazdeljen: ustvarjanje znanja ni zavezujoča omejitev, vendar je pretvarjanje znanja v skalabilno ekonomsko vrednost še vedno razdrobljeno in neučinkovito. Z analizo uspešnosti in učinkovitosti, ki temeljita na rezultatih, v različnih fazah in njihovim povezovanjem prek sestavljenega indeksa, ki temelji na uspešnosti, disertacija zagotavlja teoretično utemeljeno in empirično potrjeno podlago za inovacijsko politiko, ki temelji na dokazih, in strukturno smiselno medsebojno učenje.

Ključne besede: Inovacijska politika, inovacijska uspešnost, inovacijska učinkovitost, inovacijska zmogljivost, nacionalni inovacijski sistem, inovacijski indeks, Analiza podatkovne učinkovitosti.

PROŠIRENI SAŽETAK

SVEOBUHVATNI OKVIR MJERENJA ZA PROCJENU USPJEŠNOSTI NACIONALNIH INOVACIJSKIH SUSTAVA U ZEMLJAMA EU

Ova disertacija razvija konceptualno koherentan i empirijski robustan okvir za procjenu uspješnosti nacionalnih inovacijskih sustava (NIS) u Europskoj uniji (EU). Inovacije su odavno prepoznate kao središnji pokretač dugoročnog rasta, produktivnosti i strukturne obnove, a inovacijska politika je stoga postala strateški prioritet EU, artikuliran kroz uzastopne okvire kao što su Lisabonska agenda, Unija inovacija, Horizon 2020 i Horizon Europe. Kako se javni proračuni namijenjeni istraživanju i inovacijama nastavljaju širiti, pitanje s kojim se suočavaju kreatori politika više nije treba li intervenirati, već kako učinkovito intervenirati. Ova promjena zahtijeva mjerne instrumente sposobne pružiti dijagnostički uvid u to gdje inovacijski sustavi uspijevaju, gdje se raspadaju i koje intervencije najvjerojatnije poboljšavaju rezultate. Europska ploča inovacijskih rezultata (EIS) i slični složeni alati postali su dominantne reference za procjenu među zemljama, ali literatura je više puta dokumentirala njihova konceptualna i metodološka ograničenja. Njihova središnja slabost je spajanje ulaznih podataka, okvirnih uvjeta, institucionalnih odrednica i ostvarenih inovacijskih ishoda unutar jednog agregiranog rezultata, što prikriva odražavaju li uočene razlike povoljne strukturne resurse ili istinski učinkovite procese transformacije. Kao rezultat toga, postojeći kriteriji podržavaju rangiranje, ali nude ograničenu dijagnostičku i političku relevantnost.

Istraživačka praznina kojom se bavi ova disertacija stoga nije nedostatak alata za mjerenje, već nedostatak integriranog, višedimenzionalnog okvira mjerenja koji odražava unutarnju logiku inovacijskih sustava. Teorija inovacijskih sustava i evolucijska ekonomija konceptualiziraju inovaciju kao sekvencijalni, kumulativni i institucionalno ugrađeni proces u kojem se ulazi pretvaraju u znanje i tehnološke rezultate, a ti se rezultati potom transformiraju u komercijalne i ekonomske ishode. Ipak, prevladavajuća praksa mjerenja sažima ove različite pojave u složene indekse ili ih analizira izolirano putem modela učinkovitosti ulaza i izlaza. Disertacija tvrdi da smislena procjena zahtijeva istovremeno, ali analitički odvojeno razmatranje tri dimenzije: inovacijske uspješnosti shvaćene strogo kao uočljivi rezultati, inovacijske učinkovitosti shvaćene kao produktivnost transformacije kroz faze i inovacijskog kapaciteta shvaćenog kao latentno sistemsko svojstvo koje uvjetuje oboje. Održavanje strogih granica između ovih dimenzija u fazi mjerenja, uz istovremeno dopuštanje njihove interakcije na razini interpretacije, središnji je konceptualni potez disertacije.

Nadovezujući se na ovu pretpostavku, disertacija se bavi trima specifičnim istraživačkim pitanjima i testira dvije hipoteze. Istraživačka pitanja postavljaju: (RQa) koji obrasci varijacija postoje u rezultatima inovacija u zemljama EU-a i uravnotežuju li sustavi proizvodnju znanja i komercijalizaciju simetrično; (RQb) gdje se uska grla koncentriraju u sekvencijalnoj transformaciji resursa u rezultate inovacija i vrijedi li takozvani europski paradoks na razini učinkovitosti; i (RQc) koje strukturne konfiguracije nastaju kada se zemlje grupiraju zajedno prema rezultatima i produktivnosti transformacije te razlikuju li se te konfiguracije od skupina usporedivih zemalja na temelju agregatnih rangova. Prva hipoteza (H1) predviđa da višedimenzionalni okvir koji integrira performanse i učinkovitost temeljene na rezultatima, interpretiran kroz perspektivu kapaciteta, pruža sveobuhvatniju i dijagnostički informativniju procjenu NIS-a od postojećih okvira. Druga hipoteza (H2) predviđa da indeksi izgrađeni na ovom okviru omogućuju točniju identifikaciju usporedivih zemalja za mjerenje i učenje politika. Svaka hipoteza uparena je s eksplicitnim kriterijima opovrgavanja, osiguravajući da se okvir testira u uvjetima pod kojima bi se mogao opovrgnuti.

Metodološki, disertacija operacionalizira konceptualni okvir putem tri integrirana instrumenta mjerenja. Prvo, odabir pokazatelja provodi se putem postupka odabira više klasterskih značajki (MCFS), koji pruža transparentnu, ponovljivu i alternativu odabiru pokazatelja vođenom stručnjacima ili na temelju presedana. MCFS djeluje strogo unutar teoretski definiranih kategorija ulaza, izlaza i ishoda, adresirajući redundanciju i multikolinearnost bez žrtvovanja interpretabilnosti. Drugo, inovacijska učinkovitost temeljena na rezultatima mjeri se putem Indeksa uspješnosti inovacija temeljenog na rezultatima (IRPI), koji obuhvaća samo uočljivo znanje i tehnološke izlaze te komercijalne ishode, namjerno isključujući ulaze i institucionalne odrednice. Treće, učinkovitost inovacija procjenjuje se putem Indeksa inovacijske učinkovitosti (IEI), izvedenog iz dvostupanjskog modela analize obuhvata podataka (DEA) koji odražava sekvencijalnu strukturu inovacijskih procesa odvojeno procjenjujući učinkovitost u proizvodnji znanja i u komercijalizaciji. Dva indeksa se potom integriraju putem Indeksa uspješnosti temeljenog na rezultatima prilagođenog učinkovitosti (IE_IRPI), izgrađenog na pravilu agregacije koje prvo stavlja na učinak, u kojem učinkovitost djeluje kao ograničeni modifikator, a ne kao komponenta jednake težine. Okvir se empirijski primjenjuje na svih dvadeset sedam država članica EU-a u razdoblju 2017-2024, omogućujući i presječnu i dinamičku analizu. Inovacijski kapacitet se u cijelosti tretira kao interpretativna leća koja utječe na objašnjenje trajnih obrazaca, a ne kao izravno izmjerena dimenzija, čime se čuva analitička odvojenost i izbjegava problem identifikacije povezan s pokazateljima kapaciteta.

Empirijska analiza potvrđuje značajnu i trajnu strukturnu heterogenost u europskim NIS-ovima. Što se tiče RQa, IRPI otkriva izražene varijacije u učinku temeljenom na rezultatima, u rasponu od zemalja na granici razvoja poput Švedske, Njemačke, Finske i Irske do sustava u zaostajanju poput Rumunjske, Bugarske i Latvije, s ograničenom konvergencijom tijekom razdoblja promatranja. Što se tiče RQb, prosječna učinkovitost proizvodnje znanja (0,93) znatno premašuje prosječnu učinkovitost komercijalizacije (0,81), a 18 od 27 zemalja pokazuje veću učinkovitost u 1. fazi nego u 2. fazi. Ovaj obrazac potvrđuje europski paradoks na razini učinkovitosti: zemlje EU produktivnije su u pretvaranju inputa u znanje nego u pretvaranju znanja u ekonomsku vrijednost. Što se tiče RQc-a, zajednička IRPI-IEI distribucija generira četiri različite konfiguracije usklađene s teoretski predviđenim tipovima kapaciteta: visokoučinkoviti i učinkoviti sustavi (npr. Njemačka, Nizozemska, Finska, Irska), visokoučinkoviti, ali manje učinkoviti sustavi (npr. Švedska, Danska, Belgija), sustavi s nižim performansama, ali učinkoviti (npr. Rumunjska, Bugarska, Hrvatska) i sustavi s nižim performansama i neučinkoviti sustavi (npr. Španjolska, Litva, Slovačka). Ove se grupacije značajno razlikuju od skupina usporedivih zemalja na temelju EIS Summary Innovation Indexa, s otprilike 44% zemalja reklasificiranih po razinama.

Ovi rezultati pružaju snažnu podršku za obje hipoteze. Hipotezu 1 potvrđuju tri konvergentna dokaza: gotovo nulta korelacija između IRPI-ja i IEI-ja (r = 0,03) ukazuje na to da performanse i učinkovitost obuhvaćaju empirijski različite, a ne redundantne aspekte funkcioniranja NIS-a; sustavna asimetrija između proizvodnje znanja i učinkovitosti komercijalizacije pokazuje da dvostupanjska dekompozicija otkriva obrasce uskih grla nevidljive u agregatnim mjerama; a distribucija s četiri kvadranta otkriva teoretski značajne konfiguracije zemalja. Hipoteza 2 potvrđena je Monte Carlo simulacijom, koja pokazuje robusnost rangiranja sa Spearmanovim korelacijama većim od 0,95 u prostoru parametara vjerodostojnih za politiku, vanjskom validacijom u odnosu na BDP po glavi stanovnika i učinkovitost vlade te suštinskom diferencijacijom IE_IRPI skupina srodnih skupina od onih dobivenih agregiranim rangiranjem. Ni za jednu hipotezu nije aktiviran prag krivotvorenja.

Disertacija doprinosi na tri povezane razine. Konceptualno, ona unapređuje istraživanje inovacijskih sustava provođenjem stroge analitičke odvojenosti između performansi, učinkovitosti i kapaciteta, dimenzija koje su široko priznate u teoriji, ali rijetko se razlikuju u empirijskim mjerenjima. Metodološki, nudi jedinstvenu arhitekturu koja kombinira odabir pokazatelja temeljen na podacima putem MCFS-a, teoretski usidrene dvofazne DEA i pravilo agregacije usmjerene na performanse koje integrira mjere temeljene na rezultatima i mjere temeljene na granicama bez dopuštanja da jedna nadjača drugu. Ovi metodološki principi također doprinose široj literaturi o konstrukciji kompozitnih pokazatelja, pokazujući kako se dimenzionalno odvajanje, tretman latentnim konstruktom i sustavna validacija robusnosti mogu primijeniti u svim područjima politike. Empirijski, primjena na EU-27 daje dokaze koji nisu dostupni iz postojećih instrumenata za mjerenje, uključujući empirijsku neovisnost performansi i učinkovitosti, lokalizaciju strukturnih uskih grla u komercijalizaciji i postojanje stabilnih strukturnih konfiguracija NIS-a, a ne konvergenciju oko prosjeka EU. Međutim,

postojanost asimetrija specifičnih za faze i strukturna stabilnost konfiguracija zemalja naglašavaju potrebu za daljnjim istraživanjem institucionalnih i mehanizama povezanih s kapacitetom koji održavaju uska grla u komercijalizaciji u različitim vrstama inovacijskih sustava.

U smislu politike, okvir pomiče procjenu inovacija dalje od rangiranja prema strukturnoj dijagnozi. Omogućuje kreatorima politika da razlikuju probleme razmjera od problema učinkovitosti i da identificiraju uska grla specifična za pojedinu fazu. Za visokoučinkovite, ali neučinkovite sustave, prioriteti politike trebali bi se usredotočiti na poboljšanje procesa transformacije, posebno u komercijalizaciji. Za učinkovite, ali manje učinkovite sustave, središnji izazov je proširenje baze inovacijskih ulaganja i skaliranje uspješnih aktivnosti. Za sustave koji ne postižu dobre rezultate u obje dimenzije, potrebne su koordinirane reforme na razini cijelog sustava u obrazovanju, financijama, sposobnostima tvrtki i mehanizmima difuzije. Okvir također podržava modularni redizajn EIS-a u kojem se rezultati i učinkovitost temeljeni na rezultatima prikazuju kao analitički različite, ali komplementarne dimenzije, te pruža bazu dokaza za razlikovanje instrumenata programa Obzor Europa i kohezijske politike prema vrsti sustava, a ne prema agregatnom rangu. Središnja poruka je da je europski inovacijski izazov strukturiran, specifičan za pojedinu fazu i neravnomjerno raspoređen: stvaranje znanja nije obvezujuće ograničenje, ali prevođenje znanja u skalabilnu ekonomsku vrijednost ostaje fragmentirano i neučinkovito. Zajedničkom analizom rezultata i učinkovitosti u svim fazama, te njihovom integracijom putem kompozitnog indeksa utemeljenog na učinku, disertacija pruža teoretski utemeljenu i empirijski validiranu osnovu za inovacijsku politiku utemeljenu na dokazima i strukturno smisleno učenje među kolegama.

Ključne riječi: Politika inovacija, učinak inovacija, učinkovitost inovacija, kapacitet inovacija, nacionalni inovacijski sustav, indeks inovacija, Analiza omota podataka.

REFERENCES

Ács, Z. J., Autio, E., and Szerb, L. (2014). National Systems of Entrepreneurship: Measurement issues and policy implications. *Research Policy*, *43*(3), 476-494. https://doi.org/10.1016/j.respol.2013.08.016

Adler, N., & Yazhemsky, E. (2010). Improving discrimination in data envelopment analysis: PCA-DEA or variable reduction. *European Journal of Operational Research*, 202(1), 273-284. https://doi.org/10.1016/j.ejor.2009.03.050

Alnafrah, I. (2021). Efficiency evaluation of BRICS's national innovation systems based on bias-corrected network data envelopment analysis. *Journal of Innovation and Entrepreneurship, 10*(1), Article 26. https://doi.org/10.1186/s13731-021-00159-3

Alqararah, K. (2023). Assessing the robustness of composite indicators: the case of the Global Innovation Index. *Journal of Innovation and Entrepreneurship*, *12*(1), 1-22. https://doi.org/10.1186/s13731-023-00332-w

Anderson, H. J., and Stejskal, J. (2019). Diffusion efficiency of innovation among EU member states: A data envelopment analysis. *Economies, 7*(2), Article 34. https://doi.org/10.3390/economies7020034

Anouze, A. L., al Khalifa, M. M., and Al-Jayyousi, O. R. (2024). Reevaluating national innovation systems: An index based on dynamic-network data envelopment analysis. *Socio-Economic Planning Sciences*, *95*, 1-13. https://doi.org/10.1016/j.seps.2024.102003

Ansaloni, V., Barbero, M., Bridgford, M., Giacomotto, S., Bloemers, D., Licciardello, A., and Hsia, M. (2019). *European Innovation Scoreboard-Exploratory Report A: What to measure and why? Final Report 2 European Innovation Scoreboard-Exploratory Report.* European Commission, Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs.

Antoncic, B., and Hisrich, R. D. (2001). Intrapreneurship: Construct refinement and cross-cultural validation. *Journal of Business Venturing*, *16*(5), 495-527. https://doi.org/10.1016/S0883-9026(99)00054-3

Anu, Singh, A. K., Raza, S. A., Nakonieczny, J., and Shahzad, U. (2023). Role of financial inclusion, green innovation, and energy efficiency for environmental performance? Evidence from developed and emerging economies in the lens of sustainable development. *Structural Change and Economic Dynamics*, *64*, 213-224. https://doi.org/10.1016/j.strueco.2022.12.008

Aparicio, J., Ortiz, L., Pastor, J. T., and Zabala-Iturriagagoitia, J. M. (2020). Introducing cross-productivity: A new approach for ranking productive units over time in Data Envelopment Analysis. *Computers and Industrial Engineering*, *144*, 1-15. https://doi.org/10.1016/j.cie.2020.106456

Archibugi, D., and Coco, A. (2004). A new indicator of technological capabilities for developed and developing countries (ArCo). *World Development*, *32*(4), 629-654. https://doi.org/10.1016/j.worlddev.2003.10.008

Archibugi, D., and Coco, A. (2005). Measuring technological capabilities at the country level: A survey and a menu for choice. *Research Policy*, *34*(2), 175-194. https://doi.org/10.1016/j.respol.2004.12.002

Archibugi, D., Denni, M., and Filippetti, A. (2009). The technological capabilities of nations: The state of the art of synthetic indicators. *Technological Forecasting and Social Change*, *76*(7), 917-931. https://doi.org/10.1016/j.techfore.2009.01.002

Aristovnik, A. (2014). Efficiency of the R&D sector in the EU-27 at the regional level: An application of DEA. *Lex Localis – Journal of Local Self-Government, 12*(3), 519–531. https://doi.org/10.4335/12.3.519-531(2014)

Arsawan, I. W. E., Koval, V., Rajiani, I., Rustiarini, N. W., Supartha, W. G., and Suryantini, N. P. S. (2022). Leveraging knowledge sharing and innovation culture into SMEs sustainable competitive advantage. *International Journal of Productivity and Performance Management*, *71*(2), 405-428. https://doi.org/10.1108/IJPPM-04-2020-0192

Barbero, J., Zabala-Iturriagagoitia, J. M., and Zofío, J. L. (2021). Is more always better? On the relevance of decreasing returns to scale on innovation. *Technovation*, *107*. https://doi.org/10.1016/j.technovation.2021.102314

Barrichello, A., Santos, E. G. dos, and Morano, R. S. (2020). Determinant and priority factors of innovation for the development of nations*. Innovation and Management Review*, *17*(3), 307-320. https://doi.org/10.1108/INMR-04-2019-0040

Bartels, F. L., Voss, H., Lederer, S., and Bachtrog, C. (2012). Determinants of National Innovation Systems: Policy implications for developing countries. *Innovation*, *14*(1), 2-18. https://doi.org/10.5172/impp.2012.14.1.2

Beaudry, C., and Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy*, *38*(2), 318-337. https://doi.org/10.1016/j.respol.2008.11.010

Bloom, N., Draca, M., and Van Reenen, J. (2016). Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity. *The Review of Economic Studies*, *83*(1), 87-117. https://doi.org/10.1093/restud/rdv039

Bogetoft, P., and Otto, L. (2011). *Benchmarking with DEA, SFA, and R*. Springer. https://doi.org/10.1007/978-1-4419-7961-2

Cai, D., Zhang, C., and He, X. (2010). Unsupervised feature selection for Multi-Cluster data. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 333-342. https://doi.org/10.1145/1835804.1835848

Cai, Y., and Hanley, A. (2012). *Building BRICS: 2-Stage DEA analysis of R&D efficiency* (No. 1788; Kiel Working Paper). https://hdl.handle.net/10419/64832

Carayannis, E. G., and Campbell, D. F. J. (2009). 'Mode 3' and 'Quadruple Helix': Toward a 21st century fractal innovation ecosystem. *International Journal of Technology Management*, *46*(3/4), 201-234. https://doi.org/10.1504/IJTM.2009.023374

Carayannis, E. G., Grigoroudis, E., and Goletsis, Y. (2016). A multilevel and multistage efficiency evaluation of innovation systems: A multiobjective DEA approach. *Expert Systems with Applications*, *62*, 63-80. https://doi.org/10.1016/j.eswa.2016.06.017

Carlsson, B., Jacobsson, S., Holmén, M., and Rickne, A. (2002). Innovation Systems: Analytical and methodological issues. *Research Policy*, *31*(2), 233-245. https://doi.org/10.1016/S0048-7333(01)00138-X

Castellacci, F. (2008). Innovation and the competitiveness of industries: Comparing the mainstream and the evolutionary approaches. *Technological Forecasting and Social Change*, *75*(7), 984-1006. https://doi.org/10.1016/j.techfore.2007.09.002

Castellacci, F., and Natera, J. M. (2011). A new panel dataset for cross-country analyses of national systems, growth and development (CANA). *Innovation and Development*, *1*(2), 205-226. https://doi.org/10.1080/2157930X.2011.605871

Castellacci, F., and Natera, J. M. (2013). The dynamics of national innovation systems: A panel cointegration analysis of the coevolution between innovative capability and absorptive capacity. *Research Policy*, *42*(3), 579-594, https://doi.org/10.1016/j.respol.2012.10.006

Charnes, A., Cooper, W., and Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, *2*(6), 429-444.

Cherchye, L., Moesen, W., Rogge, N., and Puyenbroeck, T. van. (2007). An introduction to “benefit of the doubt” composite indicators. *Social Indicators Research*, *82*(1), 111-145. https://doi.org/10.1007/s11205-006-9029-7

Cohen, W. M., and Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, *35*(1), 128-152.

https://doi.org/10.2307/2393553

Cooper, W., Seiford, L. and Tone, K. (2007). *Data Envelopment Analysis: A comprehensive text with models, applications, references and DEA-solver software.* Springer. https://doi.org/10.1007/978-0-387-45283-8

Cullmann, A., Schmidt-Ehmcke, J., and Zloczysti, P. (2012). R&D efficiency and barriers to entry: a two stage semi-parametric DEA approach. *Oxford Economic Papers, 64*(1), 176-196. https://doi.org/10.1093/oep/gpr015

Desai, M., Fukuda-Parr, S., Johansson, C., and Sagasti, F. (2002). Measuring the Technology Achievement of Nations and the Capacity to Participate in the Network Age. *Journal of Human Development*, *3*(1), 95-122. https://doi.org/10.1080/14649880120105399

Diamantopoulos, A., and Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research*, *38*(2), 269-277.

Didenko, A., Loseva, O., and Abdikeev, N. (2019). Measuring efficiency of regional innovation system with DEA and PCA. 11th IEEE International Conference on Application of Information and Communication Technologies, *AICT 2017 - Proceedings*. https://doi.org/10.1109/ICAICT.2017.8687279

Dosi, G. (1988). Sources, procedures, and microeconomic effects of innovation. *Journal of Economic Literature, 26*(3), 1120-1171. https://www.jstor.org/stable/2726526

Dosi, G., Llerena, P., and Labini, M.,S. (2006). The relationships between science, technologies and their industrial exploitation: An illustration through the myths and realities of the so-called ‘European Paradox’. *Research Policy*, *35*(10), 1450-1464. https://doi.org/10.1016/j.respol.2006.09.012

Dolowitz, D. P., and Marsh, D. (2000). Learning from abroad: The role of policy transfer in contemporary policymaking. *Governance*, *13*(1), 5-23. https://doi.org/10.1111/0952-1895.00121

Draghi, M. (2024). *The future of European competitiveness, Part A: A competitiveness strategy for Europe.* European Commission. https://doi.org/10.2872/9356120

Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C. S., and Shale, E. A. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research*, *132*(2), 245-259. https://doi.org/10.1016/S0377-2217(00)00149-1

Edquist, C. (2011). Design of innovation policy through diagnostic analysis: Identification of systemic problems (or failures). *Industrial and Corporate Change*, *20*(6), 1725-1753. https://doi.org/10.1093/icc/dtr060

Edquist, C., and Johnson, B. (1997). Institutions and organisations in systems of innovation. In C. Edquist (ed.), *Systems of innovation: Technologies, institutions and organizations* (pp. 41-60). Pinter.

Edquist, C., Zabala-Iturriagagoitia, J. M., Barbero, J., and Zofío, J. L. (2018). On the meaning of innovation performance: Is the synthetic indicator of the Innovation Union Scoreboard flawed? *Research Evaluation, 27*(3), 196-211. https://doi.org/10.1093/reseval/rvy011

El Gibari, S., Gómez, T., and Ruiz, F. (2022). Combining reference point based composite indicators with data envelopment analysis: application to the assessment of universities. *Scientometrics*, *127*(8), 4363-4395.

Emerson, K., and Nabatchi, T. (2015). *Collaborative Governance Regimes.* Georgetown University Press. http://www.jstor.org/stable/j.ctt19dzcvf

Etzkowitz, H., and Leydesdorff, L. (2000). The dynamics of innovation: From National Systems and "Mode 2" to a Triple Helix of university–industry–government relations. *Research Policy*, *29*(2), 109-123. https://doi.org/10.1016/S0048-7333(99)00055-4

European Commission. (1995). *Green Paper on Innovation* (COM(95) 688 final). European Commission.

European Commission. (2005). *Working together for growth and jobs: A new start for the Lisbon Strategy* [Communication to the Spring European Council] (COM(2005) 24 final). Office for Official Publications of the European Communities. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52005DC0024

European Commission. (2013). *Measuring innovation output in Europe: Towards a new indicator* (COM(2013) 624 final). European Commission.

European Commission. (2022). *A new European innovation agenda* (Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, COM(2022) 332 final). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022DC0332

European Commission. (2025). *A dynamic EU budget for the priorities of the future: The Multiannual Financial Framework 2028-2034* [Communication]. https://commission.europa.eu/strategy-and-policy/eu-budget/long-term-eu-budget/eu-budget-2028-2034_en

European Investment Bank. (2020). *Financing the Deep Tech Revolution.* In EIB Investment Report 2019/2020: Accelerating Europe's Transformation (Chapter 5). European Investment Bank.

Fagerberg, J. (1994). Technology and International Differences in Growth Rates. *Journal of Economic Literature*, *32*(3), 1147-1175.

Fagerberg, J. (2000). Technological progress, structural change and productivity growth: a comparative study. *Structural Change and Economic Dynamics*, *11*(4), 393-411. https://doi.org/10.1016/S0954-349X(00)00025-4

Fagerberg, J., and Verspagen, B. (2002). Technology-gaps, innovation-diffusion and transformation: An evolutionary interpretation. *Research Policy*, *31*(8-9), 1291-1304. https://doi.org/10.1016/S0048-7333(02)00064-1

Fagerberg, J. (2003). Innovation: A guide to the literature [Paper presented at the workshop "The Many Guises of Innovation: What we have learnt and where we are heading", Ottawa, October 23-24, 2003]. Centre for Technology, Innovation and Culture, University of Oslo.

Fagerberg, J., and Srholec, M. (2008). National innovation systems, capabilities and economic development. *Research Policy*, *37*(9), 1417-1435. https://doi.org/10.1016/j.respol.2008.06.003

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *120*(3), 253-281. https://doi.org/10.2307/2343100

Färe, R., Grosskopf, S., Norris, M., and Zhang, Z. (1994). Productivity growth, technical progress, and efficiency change in industrialized countries. *The American Economic Review*, *84*(1), 66-83. http://www.jstor.org/stable/2117971

Färe, R., and Grosskopf, S. (2000). Network DEA. *Socio-Economic Planning Sciences*, *34*(1), 35-49. https://doi.org/10.1016/S0038-0121(99)00012-9

Filippetti, A., and Archibugi, D. (2011). Innovation in times of crisis: National Systems of Innovation, structure, and demand. *Research Policy*, *40*(2), 179–192. https://doi.org/10.1016/j.respol.2010.09.001

Filippetti, A., and Peyrache, A. (2011). The Patterns of Technological Capabilities of Countries: A Dual Approach using Composite Indicators and Data Envelopment Analysis. *World Development*, *39*(7), 1108-1121. https://doi.org/10.1016/j.worlddev.2010.12.009

Freeman, C. (1987). *Technology policy and economic performance: Lessons from*

Japan. Pinter.

Freeman, C., and Louçã, F. (2001). *As time goes by: From the industrial revolutions to the information revolution*. Oxford University Press.

Freeman, C., and Soete, L. (1997). *The economics of industrial innovation* (3rd ed.). Pinter.

Furman, J. L., Porter, M. E., and Stern, S. (2002). The determinants of national innovative capacity. *Research Policy, 31*(6), 899-933. https://doi.org/10.1016/S0048-7333(01)00152-4

George, G., McGahan, A. M., and Prabhu, J. (2012). Innovation for Inclusive Growth: Towards a Theoretical Framework and a Research Agenda. *Journal of Management Studies*, *49*(4), 661-683. https://doi.org/10.1111/j.1467-6486.2012.01048.x

Glasserman, P. (2003). *Monte Carlo methods in financial engineering*. Springer-Verlag.

Godin, B. (2003). The emergence of S&T indicators: why did governments supplement statistics with indicators? *Research Policy*, *32*(1), 679-691.

Gogodze, J. (2016). Mechanisms and functions within a National Innovation System. *Journal of Technology Management and Innovation*, *11*(4), 12-21. http://jotmi.org

Greco, S., Ishizaka, A., Tasiou, M., and Torrisi, G. (2019). On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. Social Indicators Research, *141*(1), 61-94.

Griliches, Z. (1990). Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, *28*(4), 1661-1707. http://www.jstor.org/stable/2727442

Groenewegen, J., and van der Steen, M. (2006). The evolution of national innovation systems. *Journal of Economic Issues, 40*(2), 277-285.

Grupp, H., and Mogee, M. E. (2004). Indicators for national science and technology policy: How robust are composite indicators? *Research Policy*, *33*(9), 1373-1384. https://doi.org/10.1016/j.respol.2004.09.007

Grupp, H., and Schubert, T. (2010). Review and new evidence on composite innovation indicators for evaluating national performance. *Research Policy*, *39*(1), 67-78. https://doi.org/10.1016/j.respol.2009.10.002

Guan, J., and Chen, K. (2012). Modeling the relative efficiency of national innovation systems. *Research Policy*, *41*(1), 102-115. https://doi.org/10.1016/j.respol.2011.07.001

Gutiérrez, C., Heijs, J., and Baumert, T. (2017). Are we spending our scarce R&D resources adequately? Analyzing the efficiency of EU's regional innovation systems [Conference paper]. *XVII Congreso Latino-Iberoamericano de Gestión Tecnológica (ALTEC 2017),* Mexico City, Mexico.

Hagedoorn, J., and Cloodt, M. (2003). Measuring innovative performance: Is there an advantage in using multiple indicators? *Research Policy, 32*(8), 1365-1379. https://doi.org/10.1016/S0048-7333(02)00137-3

Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.

Hall, B. H., Jaffe, A., and Trajtenberg, M. (2005). Market value and patent citations. *The RAND Journal of Economics*, *36*(1), 16-38. https://www.jstor.org/stable/1593752

Hollanders, H., and Celikel Esser, F. (2007). *Measuring innovation efficiency* (INNO-Metrics Thematic Paper). European Commission, DG Enterprise, and Industry.

Hollanders, H., Es-Sadki, N., and Kanerva, M. (2017). *European innovation scoreboard 2017*. European Commission.

Hollanders, H., Es-Sadki, N., Merkelbach, I., and Khalilova, A. (2020). *European innovation scoreboard 2020*. Publications Office of the European Union.

Hollanders, H., Es-Sadki, N., and Khalilova, A. (2024). *European innovation scoreboard 2024*. European Commission.

Huang, H.-C., Shih, H.-Y., and Wu, Y.-C. (2011). Contagion effects of national innovative capacity: Comparing structural equivalence and cohesion models. *Technological Forecasting and Social Change, 78*(2), 244-255. https://doi.org/10.1016/j.techfore.2010.07.017

Hudec, O., and Prochádzková, M. (2013). The relative efficiency of knowledge innovation processes in EU countries. *Studies in Regional Science*, *43*(1), 145-162.

Izsak, K., Markianidou, P., and Radošević, S. (2015). Convergence of national innovation policy mixes in Europe - Has it gone too far? An Analysis of Research and Innovation Policy Measures in the Period 2004-12. *Journal of Common Market Studies*, *53*(1), 786-802. https://doi.org/10.1111/jcms.12221

Jaffe, A. B., and Trajtenberg, M. (2002*). Patents, citations, and innovations: A window on the knowledge economy.* The MIT Press. https://doi.org/10.7551/mitpress/5263.001.0001

Jaklič, M. (2017*). Poslovno okolje in trajnostni razvoj*. Univerza v Ljubljani: Ekonomska fakulteta,196-200.

Janger, J., Schubert, T., Andries, P., Rammer, C., and Hoskens, M. (2017). The EU 2020 innovation indicator: A step forward in measuring innovation outputs and outcomes? *Research Policy*, *46*(1), 30-42. https://doi.org/10.1016/j.respol.2016.10.001

Jankowska, B., Matysek-Jędrych, A., and Mroczek-Dabrowska, K. (2017). Efficiency of National Innovation Systems - Poland and Bulgaria in the context of the Global Innovation Index. *Comparative Economic Research*, 20(3), 77-94. https://doi.org/10.1515/cer-2017-0021

Jansen, H. J. (2025). Innovation systems: a bibliometric review of the contemporary research employing natural language processing. *International Journal of Innovation Studies*, *9*(2), 144-164. https://doi.org/10.1016/j.ijis.2025.04.001

Jenkins, L., and Anderson, M. (2002). A multivariate statistical approach to reducing the number of variables in data envelopment analysis. *European Journal of Operational Research*, *147*(1), 51-61. https://doi.org/10.1016/S0377-2217(02)00243-6

Jovanović, M., Savić, G., Cai, Y., and Levi-Jakšić, M. (2022). Towards a Triple Helix based efficiency index of innovation systems. *Scientometrics*, *127*(5), 2577-2609. https://doi.org/10.1007/s11192-022-04304-x

Jurickova, E., Pilik, M., and Kwarteng, M. A. (2019). Efficiency measurement of National Innovation Systems of the European Union countries: DEA Model Application. *Journal of International Studies*, *12*(4), 286-299. doi: 10.14254/2071-8330.2019/12-4/19

Kalapouti, K., Petridis, K., Malesios, C., and Dey, P. K. (2020). Measuring efficiency of innovation using combined Data Envelopment Analysis and Structural Equation Modeling: empirical study in EU regions. *Annals of Operations Research*, *294*(1-2), 297-320. https://doi.org/10.1007/s10479-017-2728-4

Kao, C. (2009). Efficiency decomposition in network data envelopment analysis: A relational model. *European Journal of Operational Research*, *192*(3), 949–962. https://doi.org/10.1016/j.ejor.2007.10.008

Kao, C., and Hwang, S.-N. (2008). Efficiency decomposition in two-stage data envelopment analysis: An application to non-life insurance companies in Taiwan. *European Journal of Operational Research*, *185*(1), 418-429. https://doi.org/10.1016/j.ejor.2006.11.041

Kivimaa, P., and Kern, F. (2016). Creative destruction or mere niche support?

Innovation policy mixes for sustainability transitions. *Research Policy*, *45*(1), 205-217. https://doi.org/10.1016/j.respol.2015.09.008

Klevenhusen, A., Coelho, J., Warszawski, L., Moreira, J., Wanke, P., and Ferreira, J. J. (2021). Innovation Efficiency in OECD Countries: a Non-parametric Approach. *Journal of the Knowledge Economy*, *12*(3), 1064-1078. https://doi.org/10.1007/s13132-020-00652-2

Kline, S. J. and Rosenberg, N. (1986). "*An Overview of Innovation*", in R. Landau and N. Rosenberg (eds). The Positive Sum Strategy: Harnessing Technology for Economic Growth, National Academy Press, (pp. 275-304).

Kotsemir, M. (2013). Measuring national innovation systems efficiency - a review of DEA approach. *Higher School of Economics Research Paper No. WP BRP 16/STI/2013*.

Liou, D. Y. (2009). Two-stage R&D efficiency evaluation from the perspective of national innovation system. *In PICMET'09-2009 Portland International Conference on Management of Engineering and Technology*, pp. 180-186. IEEE.

López-Rubio, P., Roig-Tierno, N., and Mas-Verdú, F. (2022). Assessing the Origins, Evolution and Prospects of National Innovation Systems. *Journal of the Knowledge Economy*, *13*(1), 161-184. https://doi.org/10.1007/s13132-020-00712-7

Lundvall, B.-Å. (1988). Innovation as an interactive process: From user-producer interaction to the national system of innovation. In G. Dosi, C. Freeman, R. R. Nelson, G. Silverberg, and L. Soete (eds.), *Technical change and economic theory* (pp. 349-369). Pinter Publishers.

Lundvall, B.-Å. (ed.). (1992). *National systems of innovation: Towards a theory of innovation and interactive learning*. Pinter Publishers.

Lundvall, B.-Å. (2007). *Innovation system research and policy: Where it came from and where it might go* [Paper presentation]. CAS Seminar, Centre for Advanced Study, Oslo, Norway.

Lundvall, B.-Å. (2023). The COVID-19 crisis, national innovation systems, and world development. *Science, Technology and Society*, *28*(3), 359-369. https://doi.org/10.1177/09717218231178201

Lundvall, B.-Å. (2024). Transformative innovation policy–lessons from the innovation system literature. *Innovation and Development*, *14*(2), 297-314. https://doi.org/10.1080/2157930X.2022.2158996

Lutfi, A., Alrawad, M., Alsyouf, A., Almaiah, M. A., Al-Khasawneh, A., Al-Khasawneh, A.

L., Alshira'h, A. F., Alshirah, M. H., Saad, M., and Ibrahim, N. (2023). Drivers and impact of big data analytic adoption in the retail industry: A quantitative investigation applying structural equation modeling. *Journal of Retailing and Consumer Services*, *70*. https://doi.org/10.1016/j.jretconser.2022.103129

Mansfield, E. (1991). Academic research and industrial innovation, *Research Policy*, *20*(1), 1-12. https://doi.org/10.1016/0048-7333(91)90080-A

Malerba, F. and Nelson, R. R. (2011). Learning and catching up in different sectoral systems: Evidence from six industries. *Industrial and Corporate Change, 20*, 1645-1675. https://doi.org/10.1093/icc/dtr062

Maloney, W. F. (2017). *Revisiting the national innovation system in developing countries* (Policy Research Working Paper No. 8219). World Bank.

Mariano, E. B., Ferraz, D., and de Oliveira Gobbo, S. C. (2021). The Human Development Index with multiple data envelopment analysis approaches: A comparative evaluation using social network analysis. *Social Indicators Research*, *157*(2), 443-500. https://doi.org/10.1007/s11205-021-02660-4

Markman, G. D., Gianiodis, P. T., Phan, P. H., and Balkin, D. B. (2005). Innovation speed: Transferring university technology to market. *Research Policy*, *34*(7), 1058-1075. https://doi.org/10.1016/j.respol.2005.05.007

Metcalfe, J. S. (1995). Technology systems and technology policy in an evolutionary framework. *Cambridge Journal of Economics*, *19*(1), 25-46.

Metcalfe, J. S. (1998). *Evolutionary Economics and Creative Destruction.* Routledge.

Metcalfe, J. S. (2005). *Systems Failure and the Case for Innovation Policy.* In: Llerena, P., Matt, M. (eds.) Innovation Policy in a Knowledge-Based Economy. Springer. https://doi.org/10.1007/3-540-26452-3_3

Munda, G., and Nardo, M. (2005). Constructing Consistent Composite Indicators: the Issue of Weights. Institute for the Protection and Security of the Citizen. https://publications.jrc.ec.europa.eu/repository/handle/JRC32434

Narayanan, E., binti Ismail, W. R., and bin Mustafa, Z. (2022). A data-envelopment analysis-based systematic review of the literature on innovation performance. *Heliyon*, *8*(12), 1-15. Elsevier Ltd. https://doi.org/10.1016/j.heliyon.2022.e11925

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffmann, A., and Giovannini, E. (2008). *Handbook on constructing composite indicators: Methodology and user guide*. OECD Publishing. https://doi.org/10.1787/9789264043466-en

Nataraja, N. R., and Johnson, A. L. (2011). Guidelines for using variable selection techniques in data envelopment analysis. *European Journal of Operational Research*, 215(3), 662-669. https://doi.org/10.1016/j.ejor.2011.06.045

Nelson, R. R. (ed.) (1993). *National Systems of Innovation: A Comparative Study*, Oxford University Press.

Nelson, R. R., and Winter, S. G. (1982). *An evolutionary theory of economic change*. Harvard University Press.

OECD, and Eurostat. (2005). *Oslo manual: Guidelines for collecting and interpreting innovation data* (3rd ed.). OECD Publishing. https://doi.org/10.1787/9789264013100-en

OECD. (2015). *Frascati manual 2015: Guidelines for collecting and reporting data on research and experimental development*. OECD Publishing. https://doi.org/10.1787/9789264239012-en

OECD, and Eurostat. (2018). *Oslo manual 2018: Guidelines for collecting, reporting and using data on innovation* (4th ed.). OECD Publishing. https://doi.org/10.1787/9789264304604-en

Onea, I. A. (2020). Innovation Indicators and the Innovation Process-Evidence from the European Innovation Scoreboard. *Management and Marketing*, *15*(4), 605-620. https://doi.org/10.2478/mmcks-2020-0035

Patel, P. and Pavitt, K. (1994). National Innovation Systems: Why they are important, and how they might be measured and compared. *Economics of Innovation and New Technology*, *3*, 77-95. https://doi.org/10.1080/10438599400000004

Pavitt, K. (1991). What makes basic research economically useful? *Research Policy*, 20(2), 109-119. https://doi.org/10.1016/0048-7333(91)90074-Z

Petty, R., and Guthrie, J. (2000). Intellectual capital literature review: Measurement, reporting and management. *Journal of Intellectual Capital*, *1*(2), 155-176. https://doi.org/10.1108/14691930010348731

Porter, M. E. (1990). *The competitive advantage of nations*. Macmillan/Free Press

Radosevic, S. (2017). Upgrading technology in Central and Eastern European economies. *IZA World of Labor 2017, 338*. https://doi.org/10.15185/izawol.338

Reichstein, T., and Salter, A. (2006). Investigating the sources of process innovation among UK manufacturing firms. *Industrial and Corporate Change*, *15*(4), 653-682. https://doi.org/10.1093/icc/dtl014

Rodríguez-Pose, A., and Crescenzi, R. (2008). Research and development, spillovers, innovation systems, and the genesis of regional growth in Europe. *Regional Studies*, *42*(1), 51-67. https://doi.org/10.1080/00343400701654186

Rodrik, D. (2008). The real exchange rate and economic growth. *Brookings Papers on Economic Activity, 2008*(2), 365-439.

Rogers, E. (1995). *Diffusion of innovations* (4th.ed.), The Free Press.

Romer, P. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, *94*, 1002-1037. https://doi.org/10.1086/261420

Roper, S., and Hewitt-Dundas, N. (2006). International innovation comparisons: Insight or illusion? *Environment and Planning C: Government and Policy*, *24*(3), 385-401. https://doi.org/10.1068/c0452

Roszko-Wójtowicz, E., and Białek, J. (2016). A multivariate approach in measuring innovation performance. *Zbornik Radova Ekonomskog Fakulteta u Rijeci / Proceedings of Rijeka Faculty of Economics*, *34*(2), 443-479. https://doi.org/10.18045/zbefri.2016.2.443

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53-65. https://doi.org/10.1016/0377-0427(87)90125-7

Saisana, M., and Tarantola, S. (2002). *State-of-the-art report on current methodologies and practices for composite indicator development* (EUR 20408 EN). European Commission, Joint Research Centre, Institute for the Protection and Security of the Citizen.

Saltelli, A. (2007). Composite indicators between analysis and advocacy. *Social Indicators Research*, *81*(1), 65-77. https://doi.org/10.1007/s11205-006-0024-9

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S. (2008). *Global sensitivity analysis: The primer*. John Wiley and Sons. https://doi.org/10.1002/9780470725184

Schot, J., and Steinmueller, W. E. (2018). Three frames for innovation policy: R&D, systems of innovation and transformative change. *Research Policy*, *47*(9), 1554-1567. https://doi.org/10.1016/j.respol.2018.08.011

Schumpeter, J. A. (1934). *The theory of economic development*. (R. Opie, Trans.). Harvard University Press.

Schumpeter, J. A. (1942). *Capitalism, socialism and democracy*. Harper and Brothers.

Schumpeter, J. A. (1954). *History of economic analysis* (E. B. Schumpeter, Ed.). Oxford University Press.

Secka, E. G. (2023). *Measurement of national innovations systems performance trends in Sub-Saharan Africa: A multidimensional study [Doctoral dissertation].* University of North Texas. https://digital.library.unt.edu

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379-423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Simar, L., and Wilson, P. W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, *44*(1), 49-61.

Simar, L., and Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, *136*(1), 31-64. https://doi.org/10.1016/j.jeconom.2005.07.009

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, *70*(1), 65-94. https://doi.org/10.2307/1884513

Sturgeon, T. J. (2002). Modular production networks: A new American model of industrial organization. *Industrial and Corporate Change*, *11*(3), 451-496. https://doi.org/10.1093/icc/11.3.451

Tabachnick, B. G., and Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.

Teece, D. J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy*, *15*(6), 285-305. https://doi.org/10.1016/0048-7333(86)90027-2

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society*: Series B (Methodological), *58*(1), 267-288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Tziogkidis, P., Philippas, D., Leontitsis, A., and Sickles, R. C. (2020). A data envelopment analysis and local partial least squares approach for identifying the optimal innovation policy direction. *European Journal of Operational Research*, *285*(3), 1011-1024. https://doi.org/10.1016/j.ejor.2020.02.023

Umek, L., and Ravšelj, D. (2025). Biblium: An advanced Python library for bibliometric and scientometric analysis. *Proceedings of the 20th International Conference on Scientometrics and Informetrics (ISSI 2025)*. https://doi.org/10.51408/issi2025_134

United Nations Development Programme. (1990). *Human development report 1990: Concept and measurement of human development*. Oxford University Press.

Vértésy, D., and Tarantola, S. (2014). *The innovation output indicator 2014: Methodology report*. Publications Office of the European Union. https://doi.org/10.2788/836690

Wagner, J. M., and Shimshak, D. G. (2007). Stepwise selection of variables in data envelopment analysis: Procedures and managerial perspectives. *European Journal of Operational Research*, *180*(1), 57-67. https://doi.org/10.1016/j.ejor.2006.02.048

Wintjes, R., and Nauwelaers, C. (2008). Innovation policy, innovation in policy: policy learning within and across systems and clusters. *Innovation policy in Europe* (pp. 225-269). Edward Elgar Publishing.

WIPO, Cornell University, and INSEAD. (2019). *The Global Innovation Index 2019: Creating healthy lives-The future of medical innovation.* GII Efficiency Index. Geneva.

Zabala-Iturriagagoitia, J. M., Aparicio, J., Ortiz, L., Carayannis, E. G., and Grigoroudis, E. (2021). The productivity of national innovation systems in Europe: Catching up or falling behind? *Technovation*, *102*, 102215. https://doi.org/10.1016/j.technovation.2020.102215

Zabala-Iturriagagoitia, J. M., Voigt, P., Gutiérrez-Gracia, A., and Jiménez-Sáez, F. (2007). Regional innovation systems: How to assess performance. *Regional Studies*, *41*(5), 661-672. https://doi.org/10.1080/00343400601120270

Zeleny, M. (1982). *Multiple criteria decision making*. McGraw-Hill.

Zofio, J. L., Aparicio, J., Barbero, J., and Zabala-Iturriagagoitia, J. M. (2023). The influence of bottlenecks on innovation systems performance: Put the slowest climber first. *Technological Forecasting and Social Change*, *193*. Article 122607. https://doi.org/10.1016/j.techfore.2023.122607

ANNEX 1: DATASET FOR 32 EIS INDICATORS FOR EU-27 COUNTRIES (2017-2024)

Code	Year	1.1.1	1.1.2	1.1.3	1.2.1	1.2.2	1.2.3	1.3.1	1.3.2	2.1.1	2.1.2	2.1.3	2.2.1	2.2.2	2.2.3	2.3.1	2.3.2	3.1.1	3.1.2	3.2.1	3.2.2	3.2.3	3.3.1	3.3.2	3.3.3	4.1.1	4.1.2	4.2.1	4.2.2	4.2.3	4.3.1	4.3.2	4.3.3
AT	2017	0.59	0.48	0.51	0.48	0.72	0.46	0.38	0.63	0.76	0.36	0.80	0.93	0.41	0.52	0.86	0.50	0.64	0.73	0.78	0.60	0.44	0.85	0.59	0.92	0.52	0.77	0.78	0.37	0.49	0.42	0.89	0.67
BE	2017	0.59	0.71	0.34	0.50	0.83	0.73	0.81	0.48	0.61	0.59	0.66	0.72	0.45	0.61	0.83	0.68	0.67	0.58	0.78	0.46	0.51	0.71	0.44	0.35	0.65	0.80	0.64	0.69	0.30	0.59	0.77	0.35
BG	2017	0.32	0.25	0.02	0.06	0.11	0.07	0.34	0.08	0.13	0.31	0.03	0.29	0.40	0.13	0.09	0.39	0.15	0.12	0.09	0.05	0.08	0.28	0.50	1.00	0.33	0.12	0.34	0.36	0.18	0.03	0.23	0.81
CY	2017	0.11	0.91	0.32	0.46	0.65	0.18	0.25	0.36	0.19	0.41	0.02	0.03	0.24	0.08	0.56	0.40	0.31	0.41	0.30	0.26	0.53	0.31	1.00	0.26	0.75	0.43	0.81	0.74	0.16	0.30	0.55	0.22
CZ	2017	0.52	0.28	0.18	0.24	0.26	0.24	0.29	0.43	0.77	0.18	0.44	0.44	0.60	0.35	0.51	0.52	0.43	0.46	0.43	0.22	0.15	0.40	0.34	0.37	0.46	0.49	0.90	0.34	0.60	0.30	0.79	0.45
DE	2017	0.86	0.33	0.25	0.26	0.70	0.39	0.45	0.32	0.86	0.42	0.19	0.86	0.71	0.84	0.76	0.56	0.58	0.54	0.28	0.33	0.71	0.98	0.49	0.81	0.49	0.85	0.95	0.76	0.55	0.48	0.92	0.66
DK	2017	0.93	0.66	0.80	0.79	0.93	0.54	0.93	0.71	1.00	0.41	0.17	0.83	0.30	0.49	0.74	0.73	0.43	0.44	0.44	0.93	1.00	0.96	0.54	0.98	0.61	0.53	0.63	0.75	0.27	0.31	0.94	1.00
EE	2017	0.38	0.50	0.65	0.37	0.46	0.17	0.32	0.51	0.67	0.65	0.19	0.28	0.57	0.22	0.26	0.77	0.33	0.49	0.87	0.28	0.51	0.43	0.68	0.39	0.53	0.13	0.52	0.41	0.42	0.06	0.20	1.00
EL	2017	0.32	0.53	0.09	0.22	0.51	0.01	0.20	0.38	0.54	0.20	0.10	0.12	0.54	0.25	0.30	0.16	0.58	0.73	0.80	0.19	0.13	0.29	0.33	0.12	0.39	0.58	0.17	0.46	0.52	0.25	0.50	0.64
ES	2017	0.66	0.64	0.49	0.24	0.57	0.25	0.71	0.73	0.46	0.53	0.46	0.26	0.35	0.28	0.51	0.44	0.11	0.33	0.20	0.17	0.30	0.48	0.53	0.37	0.39	0.33	0.61	0.22	0.66	0.72	0.75	0.66
FI	2017	0.86	0.42	1.00	0.58	0.73	0.33	0.49	0.94	0.84	0.58	0.20	0.82	0.32	0.59	0.99	0.97	0.84	0.68	0.75	0.70	0.54	1.00	0.50	0.65	0.61	0.65	0.52	0.78	0.37	0.04	0.71	0.69
FR	2017	0.73	0.69	0.37	0.24	0.59	0.65	0.69	0.58	0.65	0.74	1.00	0.61	0.42	0.64	0.31	0.50	0.49	0.63	0.46	0.22	0.33	0.78	0.35	0.38	0.55	0.68	0.80	0.61	0.40	0.59	0.81	0.60
HR	2017	0.38	0.38	0.20	0.15	0.11	0.03	0.00	0.58	0.29	0.42	0.11	0.17	0.69	0.19	0.60	0.47	0.32	0.56	0.32	0.18	0.27	0.28	0.20	0.11	0.37	0.37	0.46	0.08	0.18	0.37	0.50	0.63
HU	2017	0.18	0.22	0.18	0.14	0.23	0.10	0.31	0.38	0.24	0.48	0.91	0.48	0.53	0.29	0.32	0.40	0.18	0.12	0.18	0.16	0.24	0.45	0.28	0.10	0.44	0.10	0.98	0.44	0.51	0.26	0.75	0.47
IE	2017	0.66	1.00	0.47	0.44	0.72	0.43	0.51	0.76	0.22	0.62	1.00	0.35	0.42	0.73	0.78	0.79	0.52	0.63	0.39	0.42		0.61	0.43	0.21	1.00	0.75	0.70	0.97	1.00	0.42	0.83	0.27
IT	2017	0.45	0.10	0.33	0.21	0.66	0.21	0.27	0.40	0.41	0.44	0.13	0.34	0.45	0.31	0.22	0.39	0.59	0.56	0.17	0.20	0.19	0.56	0.45	0.74	0.48	0.57	0.69	0.45	0.40	0.81	0.77	0.46
LT	2017	0.32	0.89	0.28	0.13	0.16	0.05	0.65	0.41	0.66	0.45	0.06	0.11	0.92	0.30	0.17	0.39	0.47	0.59	0.58	0.09	0.43	0.33	0.36	0.17	0.40	0.54	0.39	0.10	0.34	0.25	0.69	0.51
LU	2017	0.45	1.00	0.63	0.68	0.71	1.00	0.73	0.60	0.48	0.85	0.13	0.27	0.18	0.28	0.62	0.85	0.57	0.61	0.32	0.57	0.72	0.48	0.99	1.00	1.00	0.75	0.65	0.95	0.25	0.91	0.78	0.69
LV	2017	0.25	0.56	0.28	0.10	0.14	0.13	0.41	0.42	0.36	0.59	0.01	0.05	0.46	0.07	0.22	0.39	0.16	0.22	0.17	0.10	0.36	0.39	0.38	0.31	0.33	0.12	0.38	0.45	0.20	0.26	0.21	0.54
MT	2017	0.11	0.49	0.48	0.23	0.41	1.00	0.64	0.67	0.24	0.15	0.10	0.14	0.34	0.13	0.61	0.55	0.22	0.27	0.08	0.15	0.44	0.35	1.00	1.00	0.78	0.48	0.96	0.51	0.15	0.47	0.94	0.33
NL	2017	0.45	0.84	0.95	0.54	0.98	0.62	0.72	1.00	0.67	0.60	0.68	0.59	0.21	0.40	0.40	0.85	0.70	0.53	0.50	0.57	0.81	0.92	0.50	0.41	0.79	0.70	0.64	0.80	0.44	0.89	0.81	0.47
PL	2017	0.18	0.46	0.17	0.08	0.19	0.01	0.42	0.36	0.42	0.31	0.15	0.19	0.70	0.29	0.20	0.37	0.07	0.09	0.12	0.07	0.41	0.29	0.39	0.76	0.30	0.05	0.64	0.32	0.25	0.16	0.52	0.63
PT	2017	0.52	0.56	0.42	0.31	0.61	0.35	0.78	0.53	0.55	0.43	0.44	0.24	0.49	0.22	0.54	0.50	0.79	0.83	0.32	0.18	0.29	0.33	0.45	0.56	0.38	0.63	0.44	0.35	0.24	0.23	0.00	0.56
RO	2017	0.32	0.00	0.15	0.05	0.15	0.02	0.65	0.10	0.16	0.21	0.10	0.07	0.26	0.04	0.01	0.19	0.00	0.00	0.02	0.05	0.04	0.20	0.22	0.08	0.11	0.00	0.72	0.37	0.25	0.07	0.49	0.67
SE	2017	1.00	0.67	1.00	0.69	0.83	0.56	0.95	0.68	0.86	0.45	0.40	0.96	0.67	1.00	0.84	1.00	0.54	0.44	0.45	0.72	0.43	1.00	0.54	0.62	0.84	0.64	0.73	0.76	0.27	0.27	0.76	0.46
SI	2017	0.93	0.43	0.70	0.39	0.39	0.13	0.40	0.34	0.41	0.17	0.70	0.71	0.56	0.34	0.72	0.55	0.31	0.47	0.41	0.39	0.28	0.71	0.49	0.42	0.63	0.55	0.75	0.27	0.51	0.35	0.63	0.31
SK	2017	0.59	0.40	0.44	0.16	0.15	0.14	0.26	0.36	0.73	0.23	0.04	0.12	0.46	0.16	0.41	0.47	0.16	0.18	0.26	0.12	0.08	0.30	0.29	0.19	0.33	0.19	0.95	0.25	0.80	0.34	0.74	0.78
AT	2018	0.59	0.48	0.51	0.49	0.69	0.48	0.38	0.63	0.81	0.41	0.75	0.94	0.41	0.52	0.97	0.50	0.64	0.73	0.78	0.62	0.44	0.85	0.59	0.89	0.52	0.77	0.78	0.36	0.49	0.46	0.89	0.57
BE	2018	0.59	0.71	0.34	0.51	0.85	0.76	0.81	0.48	0.67	0.59	0.68	0.74	0.45	0.61	0.87	0.68	0.67	0.58	0.78	0.46	0.51	0.70	0.45	0.34	0.65	0.80	0.63	0.71	0.30	0.63	0.76	0.38
BG	2018	0.25	0.25	0.02	0.07	0.08	0.09	0.34	0.08	0.08	0.35	0.03	0.23	0.40	0.13	0.09	0.39	0.15	0.12	0.09	0.05	0.08	0.32	0.49	0.87	0.33	0.12	0.35	0.36	0.18	0.01	0.06	0.81
CY	2018	0.18	0.91	0.32	0.52	0.52	0.23	0.25	0.36	0.15	0.41	0.02	0.07	0.24	0.08	0.51	0.40	0.31	0.41	0.30	0.31	0.53	0.35	1.00	0.55	0.75	0.43	0.73	0.75	0.16	0.33	0.58	0.11
CZ	2018	0.52	0.28	0.18	0.26	0.23	0.26	0.29	0.43	0.54	0.15	0.39	0.43	0.60	0.35	0.51	0.52	0.43	0.46	0.43	0.23	0.15	0.40	0.35	0.32	0.46	0.49	0.91	0.37	0.60	0.30	0.83	0.44
DE	2018	0.79	0.33	0.25	0.27	0.70	0.39	0.45	0.32	0.86	0.45	0.19	0.86	0.71	0.84	0.74	0.56	0.58	0.54	0.28	0.34	0.71	0.97	0.50	0.79	0.49	0.85	0.95	0.76	0.55	0.52	0.92	0.62
DK	2018	0.93	0.66	0.80	0.82	0.90	0.57	0.93	0.71	0.98	0.42	0.21	0.86	0.30	0.49	0.72	0.73	0.43	0.44	0.44	0.93	1.00	0.94	0.57	1.00	0.61	0.53	0.62	0.71	0.27	0.30	0.94	0.94
EE	2018	0.45	0.50	0.65	0.39	0.45	0.19	0.32	0.51	0.47	0.57	0.16	0.26	0.57	0.22	0.26	0.77	0.33	0.49	0.87	0.32	0.51	0.44	0.71	0.48	0.53	0.13	0.48	0.42	0.42	0.07	0.43	0.65
EL	2018	0.32	0.53	0.09	0.22	0.54	0.00	0.20	0.38	0.46	0.21	0.10	0.17	0.54	0.25	0.30	0.16	0.58	0.73	0.80	0.18	0.13	0.30	0.35	0.16	0.39	0.58	0.15	0.47	0.52	0.27	0.49	0.51
ES	2018	0.79	0.64	0.49	0.24	0.56	0.25	0.71	0.73	0.44	0.53	0.46	0.26	0.35	0.28	0.54	0.44	0.11	0.33	0.20	0.17	0.30	0.49	0.53	0.38	0.39	0.33	0.58	0.23	0.66	0.71	0.73	0.66
FI	2018	0.79	0.42	1.00	0.60	0.71	0.35	0.49	0.94	0.80	0.59	0.20	0.76	0.32	0.59	0.90	0.97	0.84	0.68	0.75	0.68	0.54	1.00	0.56	0.54	0.61	0.65	0.55	0.79	0.37	0.05	0.72	0.62
FR	2018	0.66	0.69	0.37	0.25	0.58	0.65	0.69	0.58	0.64	0.77	1.00	0.62	0.42	0.64	0.31	0.50	0.49	0.63	0.46	0.22	0.33	0.79	0.35	0.37	0.55	0.68	0.79	0.62	0.40	0.65	0.81	0.61
HR	2018	0.25	0.38	0.20	0.17	0.15	0.13	0.00	0.58	0.35	0.27	0.01	0.15	0.69	0.19	0.52	0.47	0.32	0.56	0.32	0.21	0.27	0.30	0.22	0.10	0.37	0.37	0.47	0.08	0.18	0.36	0.52	0.62
HU	2018	0.18	0.22	0.18	0.14	0.24	0.18	0.31	0.38	0.24	0.47	0.97	0.48	0.53	0.29	0.33	0.40	0.18	0.12	0.18	0.17	0.24	0.45	0.28	0.12	0.44	0.10	0.95	0.44	0.51	0.29	0.78	0.37
IE	2018	0.66	1.00	0.47	0.46	0.77	0.45	0.51	0.76	0.22	0.62	0.89	0.35	0.42	0.73	0.78	0.79	0.52	0.63	0.39	0.43		0.60	0.40	0.15	1.00	0.75	0.75	0.99	1.00	0.60	0.90	0.24
IT	2018	0.45	0.10	0.33	0.22	0.67	0.23	0.27	0.40	0.39	0.40	0.27	0.34	0.45	0.31	0.21	0.39	0.59	0.56	0.17	0.21	0.19	0.57	0.47	0.82	0.48	0.57	0.68	0.46	0.40	0.87	0.77	0.44
LT	2018	0.18	0.89	0.28	0.15	0.22	0.06	0.65	0.41	0.44	0.35	0.09	0.11	0.92	0.30	0.14	0.39	0.47	0.59	0.58	0.10	0.43	0.34	0.39	0.18	0.40	0.54	0.41	0.12	0.34	0.25	0.70	0.52
LU	2018	0.45	1.00	0.63	0.72	0.79	1.00	0.73	0.60	0.45	0.85	0.13	0.29	0.18	0.28	0.72	0.85	0.57	0.61	0.32	0.51	0.72	0.53	1.00	1.00	1.00	0.75	0.56	0.95	0.25	0.82	0.82	0.46
LV	2018	0.18	0.56	0.28	0.12	0.16	0.18	0.41	0.42	0.22	0.83	0.01	0.03	0.46	0.07	0.20	0.39	0.16	0.22	0.17	0.10	0.36	0.23	0.38	0.24	0.33	0.12	0.39	0.47	0.20	0.25	0.23	0.67
MT	2018	0.11	0.49	0.48	0.26	0.54	1.00	0.64	0.67	0.09	0.97	0.08	0.13	0.34	0.13	0.54	0.55	0.22	0.27	0.08	0.14	0.44	0.45	1.00	1.00	0.78	0.48	0.74	0.51	0.15	0.43	0.93	0.53
NL	2018	0.45	0.84	0.95	0.56	0.98	0.69	0.72	1.00	0.64	0.67	0.63	0.60	0.21	0.40	0.53	0.85	0.70	0.53	0.50	0.58	0.81	0.94	0.52	0.44	0.79	0.70	0.64	0.79	0.44	0.89	0.81	0.44
PL	2018	0.11	0.46	0.17	0.09	0.22	0.01	0.42	0.36	0.21	0.34	0.13	0.26	0.70	0.29	0.20	0.37	0.07	0.09	0.12	0.07	0.41	0.30	0.40	0.75	0.30	0.05	0.63	0.33	0.25	0.19	0.55	0.67
PT	2018	0.52	0.56	0.42	0.33	0.57	0.43	0.78	0.53	0.54	0.36	0.41	0.25	0.49	0.22	0.55	0.50	0.79	0.83	0.32	0.18	0.29	0.33	0.46	0.56	0.38	0.63	0.45	0.32	0.24	0.25	0.02	0.49
RO	2018	0.18	0.00	0.15	0.05	0.15	0.04	0.65	0.10	0.10	0.32	0.10	0.10	0.26	0.04	0.00	0.19	0.00	0.00	0.02	0.05	0.04	0.21	0.23	0.10	0.11	0.00	0.73	0.39	0.25	0.04	0.50	0.78
SE	2018	0.93	0.67	1.00	0.71	0.82	0.59	0.95	0.68	0.86	0.48	0.38	0.97	0.67	1.00	0.84	1.00	0.54	0.44	0.45	0.71	0.43	1.00	0.57	0.61	0.84	0.64	0.73	0.74	0.27	0.30	0.79	0.51
SI	2018	1.00	0.43	0.70	0.40	0.38	0.15	0.40	0.34	0.38	0.15	0.55	0.65	0.56	0.34	0.69	0.55	0.31	0.47	0.41	0.38	0.28	0.67	0.48	0.38	0.63	0.55	0.77	0.27	0.51	0.36	0.65	0.35
SK	2018	0.52	0.40	0.44	0.17	0.15	0.14	0.26	0.36	0.28	0.22	0.07	0.16	0.46	0.16	0.46	0.47	0.16	0.18	0.26	0.13	0.08	0.26	0.32	0.14	0.33	0.19	0.94	0.26	0.80	0.36	0.77	0.46
AT	2019	0.59	0.48	0.51	0.52	0.72	0.51	0.38	0.63	0.81	0.36	0.71	0.92	0.41	0.52	0.78	0.50	0.64	0.73	0.78	0.65	0.46	0.85	0.60	0.87	0.52	0.77	0.77	0.37	0.49	0.47	0.91	0.49
BE	2019	0.45	0.71	0.34	0.54	0.80	0.34	0.81	0.48	0.68	0.60	0.67	0.80	0.45	0.61	0.91	0.68	0.67	0.58	0.78	0.50	0.51	0.69	0.46	0.33	0.65	0.80	0.62	0.71	0.30	0.64	0.76	0.41
BG	2019	0.32	0.25	0.02	0.08	0.09	0.09	0.34	0.08	0.09	0.31	0.03	0.21	0.40	0.13	0.12	0.39	0.15	0.12	0.09	0.05	0.08	0.29	0.49	0.68	0.33	0.12	0.38	0.34	0.18	0.05	0.46	0.81
CY	2019	0.11	0.91	0.32	0.60	0.53	0.25	0.25	0.36	0.15	0.35	0.01	0.07	0.24	0.08	0.65	0.40	0.31	0.41	0.30	0.36	0.56	0.28	1.00	0.57	0.75	0.43	0.81	0.74	0.16	0.32	0.60	0.38
CZ	2019	0.59	0.28	0.18	0.28	0.25	0.28	0.29	0.43	0.55	0.15	0.32	0.47	0.60	0.35	0.55	0.52	0.43	0.46	0.43	0.24	0.20	0.38	0.35	0.48	0.46	0.49	0.94	0.36	0.60	0.32	0.85	0.47
DE	2019	0.79	0.33	0.25	0.28	0.68	0.39	0.45	0.32	0.86	0.49	0.19	0.90	0.71	0.84	0.70	0.56	0.58	0.54	0.28	0.35	0.71	0.96	0.50	0.83	0.49	0.85	0.95	0.76	0.55	0.54	0.92	0.59
DK	2019	0.86	0.66	0.80	0.88	0.94	0.60	0.93	0.71	0.97	0.45	0.19	0.79	0.30	0.49	0.66	0.73	0.43	0.44	0.44	0.99	1.00	0.95	0.60	0.94	0.61	0.53	0.63	0.72	0.27	0.30	0.95	0.89
EE	2019	0.45	0.50	0.65	0.44	0.51	0.23	0.32	0.51	0.56	0.62	0.09	0.24	0.57	0.22	0.25	0.77	0.33	0.49	0.87	0.33	0.48	0.38	0.76	0.71	0.53	0.13	0.45	0.45	0.42	0.08	0.35	0.46
EL	2019	0.38	0.53	0.09	0.24	0.48	0.00	0.20	0.38	0.48	0.24	0.06	0.23	0.54	0.25	0.20	0.16	0.58	0.73	0.80	0.22	0.16	0.27	0.37	0.15	0.39	0.58	0.16	0.48	0.52	0.29	0.51	0.50
ES	2019	1.00	0.64	0.49	0.26	0.55	0.30	0.71	0.73	0.43	0.55	0.46	0.27	0.35	0.28	0.54	0.44	0.11	0.33	0.20	0.19	0.44	0.47	0.53	0.37	0.39	0.33	0.57	0.21	0.66	0.77	0.72	0.57
FI	2019	0.73	0.42	1.00	0.63	0.74	0.37	0.49	0.94	0.82	0.65	0.18	0.76	0.32	0.59	0.99	0.97	0.84	0.68	0.75	0.70	0.53	1.00	0.58	0.50	0.61	0.65	0.54	0.77	0.37	0.04	0.71	0.60
FR	2019	0.73	0.69	0.37	0.26	0.59	0.65	0.69	0.58	0.62	0.78	1.00	0.61	0.42	0.64	0.31	0.50	0.49	0.63	0.46	0.22	0.42	0.77	0.35	0.38	0.55	0.68	0.78	0.59	0.40	0.68	0.81	0.57
HR	2019	0.32	0.38	0.20	0.20	0.17	0.14	0.00	0.58	0.33	0.26	0.01	0.16	0.69	0.19	0.56	0.47	0.32	0.56	0.32	0.24	0.29	0.25	0.24	0.10	0.37	0.37	0.46	0.08	0.18	0.37	0.56	0.46
HU	2019	0.18	0.22	0.18	0.15	0.28	0.24	0.31	0.38	0.24	0.46	0.47	0.48	0.53	0.29	0.36	0.40	0.18	0.12	0.18	0.18	0.23	0.46	0.30	0.14	0.44	0.10	0.92	0.46	0.51	0.31	0.77	0.30
IE	2019	0.59	1.00	0.47	0.50	0.75	0.49	0.51	0.76	0.21	0.68	0.84	0.39	0.42	0.73	0.77	0.79	0.52	0.63	0.39	0.45		0.52	0.40	0.17	1.00	0.75	0.75	0.98	1.00	0.55	0.90	0.22
IT	2019	0.45	0.10	0.33	0.23	0.67	0.24	0.27	0.40	0.38	0.41	0.34	0.35	0.45	0.31	0.24	0.39	0.59	0.56	0.17	0.23	0.19	0.57	0.48	0.78	0.48	0.57	0.68	0.46	0.40	0.91	0.79	0.39
LT	2019	0.25	0.89	0.28	0.18	0.22	0.06	0.65	0.41	0.46	0.38	0.10	0.13	0.92	0.30	0.18	0.39	0.47	0.59	0.58	0.13	0.53	0.24	0.43	0.21	0.40	0.54	0.42	0.12	0.34	0.25	0.75	0.44
LU	2019	0.66	1.00	0.63	0.75	0.85	1.00	0.73	0.60	0.44	0.88	0.13	0.28	0.18	0.28	0.70	0.85	0.57	0.61	0.32	0.68	0.72	0.50	0.99	0.82	1.00	0.75	0.54	0.95	0.25	0.85	0.84	0.46
LV	2019	0.11	0.56	0.28	0.13	0.15	0.15	0.41	0.42	0.26	0.78	0.02	0.04	0.46	0.07	0.14	0.39	0.16	0.22	0.17	0.12	0.41	0.35	0.40	0.16	0.33	0.12	0.40	0.46	0.20	0.32	0.28	0.66
MT	2019	0.11	0.49	0.48	0.29	0.30	1.00	0.64	0.67	0.08	0.97	0.06	0.14	0.34	0.13	0.64	0.55	0.22	0.27	0.08	0.19	0.53	0.37	1.00	1.00	0.78	0.48	0.68	0.48	0.15	0.46	0.94	0.63
NL	2019	0.45	0.84	0.95	0.60	1.00	0.74	0.72	1.00	0.63	0.69	0.73	0.62	0.21	0.40	0.57	0.85	0.70	0.53	0.50	0.61	0.81	0.93	0.52	0.51	0.79	0.70	0.64	0.79	0.44	1.00	0.82	0.40
PL	2019	0.11	0.46	0.17	0.10	0.22	0.02	0.42	0.36	0.25	0.38	0.31	0.28	0.70	0.29	0.22	0.37	0.07	0.09	0.12	0.08	0.50	0.33	0.39	0.71	0.30	0.05	0.62	0.34	0.25	0.18	0.58	0.61
PT	2019	0.45	0.56	0.42	0.36	0.60	0.46	0.78	0.53	0.53	0.49	0.40	0.28	0.49	0.22	0.48	0.50	0.79	0.83	0.32	0.20	0.43	0.38	0.50	0.50	0.38	0.63	0.48	0.29	0.24	0.28	0.04	0.30
RO	2019	0.11	0.00	0.15	0.05	0.16	0.05	0.65	0.10	0.09	0.42	0.11	0.11	0.26	0.04	0.00	0.19	0.00	0.00	0.02	0.06	0.04	0.19	0.24	0.16	0.11	0.00	0.77	0.38	0.25	0.05	0.50	0.43
SE	2019	0.86	0.67	1.00	0.74	0.85	0.60	0.95	0.68	0.86	0.52	0.37	1.00	0.67	1.00	0.84	1.00	0.54	0.44	0.45	0.75	0.43	1.00	0.59	0.60	0.84	0.64	0.73	0.72	0.27	0.29	0.80	0.58
SI	2019	0.52	0.43	0.70	0.44	0.40	0.14	0.40	0.34	0.36	0.13	0.46	0.59	0.56	0.34	0.66	0.55	0.31	0.47	0.41	0.41	0.40	0.51	0.50	0.37	0.63	0.55	0.77	0.27	0.51	0.40	0.67	0.37
SK	2019	0.45	0.40	0.44	0.17	0.14	0.15	0.26	0.36	0.29	0.18	0.06	0.19	0.46	0.16	0.36	0.47	0.16	0.18	0.26	0.15	0.14	0.28	0.34	0.20	0.33	0.19	0.95	0.30	0.80	0.34	0.77	0.48

Code	Year	1.1.1	1.1.2	1.1.3	1.2.1	1.2.2	1.2.3	1.3.1	1.3.2	2.1.1	2.1.2	2.1.3	2.2.1	2.2.2	2.2.3	2.3.1	2.3.2	3.1.1	3.1.2	3.2.1	3.2.2	3.2.3	3.3.1	3.3.2	3.3.3	4.1.1	4.1.2	4.2.1	4.2.2	4.2.3	4.3.1	4.3.2	4.3.3
AT	2020	0.59	0.48	0.51	0.55	0.71	0.57	0.38	0.63	0.81	0.38	0.69	0.93	0.44	0.55	0.66	0.50	0.64	0.82	0.55	0.70	0.47	0.83	0.61	0.83	0.52	0.80	0.79	0.38	0.51	0.48	0.91	0.52
BE	2020	0.59	0.71	0.34	0.54	0.82	0.71	0.81	0.48	0.69	0.62	0.69	0.88	0.45	0.94	0.94	0.68	0.54	0.87	0.94	0.49	0.51	0.71	0.47	0.33	0.65	0.88	0.66	0.71	0.56	0.65	0.78	0.41
BG	2020	0.25	0.25	0.02	0.07	0.11	0.11	0.34	0.08	0.09	0.25	0.03	0.22	0.40	0.09	0.12	0.39	0.28	0.15	0.12	0.06	0.11	0.26	0.49	0.60	0.33	0.19	0.40	0.34	0.23	0.05	0.48	1.00
CY	2020	0.18	0.91	0.32	0.69	0.52	0.24	0.25	0.36	0.19	0.70	0.01	0.09	0.49	0.15	0.66	0.40	1.00	1.00	1.00	0.46	0.69	0.28	1.00	0.26	0.75	0.32	0.74	0.74	0.50	0.27	0.59	0.42
CZ	2020	0.59	0.28	0.18	0.30	0.24	0.30	0.29	0.43	0.62	0.28	0.37	0.50	0.53	0.29	0.61	0.52	0.46	0.53	0.39	0.25	0.20	0.35	0.35	0.50	0.46	0.49	0.95	0.37	0.53	0.35	0.86	0.37
DE	2020	0.79	0.33	0.25	0.29	0.66	0.39	0.45	0.32	0.86	0.55	0.19	0.92	0.73	0.90	0.76	0.56	0.78	0.82	0.51	0.36	0.71	0.97	0.50	0.78	0.49	0.84	0.94	0.75	0.58	0.56	0.92	0.58
DK	2020	0.79	0.66	0.80	0.91	0.85	0.62	0.93	0.71	1.00	0.64	0.17	0.80	0.34	0.83	0.71	0.73	0.61	0.67	0.46	1.00	1.00	0.95	0.60	0.88	0.61	0.51	0.67	0.70	0.27	0.31	0.96	0.88
EE	2020	0.45	0.50	0.65	0.52	0.48	0.28	0.32	0.51	0.69	0.53	0.08	0.24	0.89	0.44	0.25	0.77	1.00	0.79	0.96	0.43	0.51	0.44	0.86	0.78	0.53	0.64	0.46	0.46	0.45	0.06	0.36	0.77
EL	2020	0.25	0.53	0.09	0.26	0.54	0.00	0.20	0.38	0.51	0.30	0.08	0.24	0.59	0.33	0.28	0.16	0.85	0.83	0.72	0.23	0.17	0.29	0.38	0.13	0.39	0.68	0.19	0.49	0.69	0.31	0.53	0.37
ES	2020	0.86	0.64	0.49	0.27	0.53	0.28	0.71	0.73	0.43	0.55	0.46	0.29	0.38	0.28	0.50	0.44	0.19	0.22	0.22	0.20	0.44	0.46	0.53	0.33	0.39	0.35	0.57	0.20	0.81	0.79	0.72	0.54
FI	2020	0.73	0.42	1.00	0.68	0.72	0.38	0.49	0.94	0.84	0.71	0.16	0.77	0.53	0.64	0.94	0.97	0.70	0.68	0.98	0.73	0.56	1.00	0.56	0.49	0.61	0.84	0.58	0.75	0.46	0.06	0.74	0.58
FR	2020	0.59	0.69	0.37	0.26	0.54	0.65	0.69	0.58	0.62	0.80	1.00	0.61	0.42	0.64	0.31	0.50	0.61	0.53	0.45	0.23	0.46	0.74	0.36	0.36	0.55	0.68	0.79	0.61	0.40	0.64	0.81	0.55
HR	2020	0.25	0.38	0.20	0.22	0.15	0.14	0.00	0.58	0.38	0.23	0.02	0.18	0.75	0.20	0.57	0.47	0.75	0.65	0.42	0.28	0.36	0.31	0.23	0.10	0.37	0.48	0.49	0.09	0.32	0.40	0.55	0.36
HU	2020	0.18	0.22	0.18	0.16	0.26	0.32	0.31	0.38	0.24	0.47	0.56	0.48	0.59	0.29	0.35	0.40	0.33	0.15	0.31	0.20	0.31	0.45	0.30	0.14	0.44	0.16	0.96	0.46	0.30	0.28	0.76	0.40
IE	2020	0.66	1.00	0.47	0.53	0.73	0.51	0.51	0.76	0.14	0.67	0.56	0.35	0.42	0.73	0.77	0.79	0.52	0.51	0.76	0.45		0.54	0.37	0.20	1.00	0.70	0.77	0.98	1.00	0.58	0.90	0.23
IT	2020	0.38	0.10	0.33	0.25	0.67	0.26	0.27	0.40	0.39	0.43	0.61	0.38	0.51	0.41	0.36	0.39	0.71	0.81	0.49	0.24	0.21	0.56	0.49	0.71	0.48	0.66	0.67	0.44	0.51	0.95	0.79	0.39
LT	2020	0.25	0.89	0.28	0.20	0.23	0.10	0.65	0.41	0.44	0.59	0.08	0.15	0.92	0.32	0.12	0.39	0.58	0.61	0.46	0.16	0.95	0.29	0.45	0.20	0.40	0.60	0.44	0.14	0.61	0.23	0.75	0.66
LU	2020	0.66	1.00	0.63	0.81	0.70	1.00	0.73	0.60	0.44	0.63	0.13	0.25	0.26	0.37	0.68	0.85	0.54	0.54	0.40	0.62	0.72	0.54	0.97	0.81	1.00	0.78	0.61	0.96	0.18	0.82	0.78	0.48
LV	2020	0.11	0.56	0.28	0.15	0.17	0.16	0.41	0.42	0.37	0.75	0.03	0.05	0.53	0.11	0.17	0.39	0.26	0.26	0.19	0.14	0.27	0.30	0.42	0.22	0.33	0.22	0.36	0.49	0.32	0.29	0.29	0.37
MT	2020	0.11	0.49	0.48	0.29	0.28	1.00	0.64	0.67	0.10	0.97	0.10	0.14	0.61	0.24	0.63	0.55	0.58	0.55	0.26	0.21	0.65	0.50	1.00	1.00	0.78	0.26	0.68	0.48	0.32	0.64	0.95	0.86
NL	2020	0.45	0.84	0.95	0.61	0.98	0.76	0.72	1.00	0.62	0.71	0.70	0.60	0.21	0.51	0.63	0.85	0.50	0.54	0.47	0.62	0.81	0.90	0.51	0.48	0.79	0.74	0.66	0.80	0.42	0.94	0.82	0.40
PL	2020	0.11	0.46	0.17	0.11	0.23	0.02	0.42	0.36	0.29	0.36	0.29	0.33	0.66	0.28	0.26	0.37	0.15	0.12	0.10	0.09	0.55	0.28	0.39	0.64	0.30	0.07	0.64	0.36	0.24	0.18	0.58	0.61
PT	2020	0.52	0.56	0.42	0.39	0.56	0.49	0.78	0.53	0.53	0.51	0.43	0.28	0.63	0.29	0.44	0.50	0.51	0.39	0.22	0.22	0.44	0.37	0.52	0.46	0.38	0.84	0.52	0.28	0.39	0.25	0.06	0.34
RO	2020	0.11	0.00	0.15	0.06	0.22	0.06	0.65	0.10	0.08	0.53	0.11	0.11	0.26	0.01	0.01	0.19	0.09	0.00	0.03	0.07	0.00	0.19	0.24	0.16	0.11	0.00	0.77	0.40	0.17	0.09	0.54	0.17
SE	2020	0.73	0.67	1.00	0.76	0.82	0.61	0.95	0.68	0.86	0.56	0.36	1.00	0.54	0.94	0.84	1.00	0.85	0.69	0.49	0.76	0.43	1.00	0.59	0.54	0.84	0.64	0.75	0.76	0.35	0.28	0.80	0.58
SI	2020	0.52	0.43	0.70	0.46	0.37	0.15	0.40	0.34	0.38	0.20	0.52	0.61	0.51	0.29	0.72	0.55	0.69	0.47	0.43	0.42	0.27	0.52	0.49	0.33	0.63	0.43	0.81	0.25	0.34	0.42	0.68	0.39
SK	2020	0.45	0.40	0.44	0.20	0.17	0.15	0.26	0.36	0.27	0.20	0.06	0.18	0.53	0.20	0.38	0.47	0.20	0.19	0.26	0.16	0.14	0.31	0.34	0.24	0.33	0.22	0.97	0.31	0.85	0.33	0.78	0.70
AT	2021	0.59	0.48	0.51	0.59	0.66	0.61	0.38	0.63	0.81	0.37	0.75	0.94	0.44	0.55	0.40	0.50	0.64	0.82	0.55	0.72	0.54	0.84	0.62	0.79	0.52	0.80	0.78	0.40	0.51	0.51	0.92	0.50
BE	2021	0.59	0.71	0.34	0.59	0.77	0.39	0.81	0.48	0.71	0.64	0.81	1.00	0.45	0.94	0.96	0.68	0.54	0.87	0.94	0.51	0.51	0.68	0.46	0.34	0.65	0.88	0.70	0.71	0.56	0.71	0.79	0.41
BG	2021	0.25	0.25	0.02	0.09	0.10	0.11	0.34	0.08	0.15	0.21	0.02	0.23	0.40	0.09	0.16	0.39	0.28	0.15	0.12	0.06	0.13	0.27	0.49	0.62	0.33	0.19	0.42	0.39	0.23	0.06	0.39	0.99
CY	2021	0.25	0.91	0.32	0.84	0.60	0.36	0.25	0.36	0.21	0.79	0.02	0.11	0.49	0.15	0.78	0.40	1.00	1.00	1.00	0.52	0.71	0.28	1.00	0.33	0.75	0.32	0.78	0.79	0.50	0.30	0.60	0.08
CZ	2021	0.52	0.28	0.18	0.34	0.25	0.33	0.29	0.43	0.63	0.29	0.37	0.50	0.53	0.29	0.60	0.52	0.46	0.53	0.39	0.25	0.30	0.34	0.36	0.28	0.46	0.49	0.95	0.38	0.53	0.37	0.87	0.43
DE	2021	0.79	0.33	0.25	0.31	0.65	0.39	0.45	0.32	0.86	0.54	0.19	0.93	0.73	0.90	0.81	0.56	0.78	0.82	0.51	0.37	0.71	0.96	0.49	0.68	0.49	0.84	0.94	0.75	0.58	0.61	0.93	0.57
DK	2021	0.79	0.66	0.80	1.00	0.85	0.64	0.93	0.71	1.00	0.68	0.21	0.79	0.34	0.83	0.78	0.73	0.61	0.67	0.46	1.00	1.00	0.95	0.59	0.93	0.61	0.51	0.70	0.74	0.27	0.33	0.96	0.97
EE	2021	0.45	0.50	0.65	0.54	0.57	0.36	0.32	0.51	0.64	0.79	0.10	0.36	0.89	0.44	0.37	0.77	1.00	0.79	0.96	0.39	0.68	0.47	0.90	0.53	0.53	0.64	0.48	0.48	0.45	0.06	0.45	0.73
EL	2021	0.38	0.53	0.09	0.27	0.53	0.04	0.20	0.38	0.58	0.33	0.12	0.24	0.59	0.33	0.31	0.16	0.85	0.83	0.72	0.25	0.19	0.29	0.40	0.09	0.39	0.68	0.29	0.47	0.69	0.34	0.55	0.44
ES	2021	0.45	0.64	0.49	0.30	0.54	0.30	0.71	0.73	0.43	0.62	0.49	0.29	0.38	0.28	0.51	0.44	0.19	0.22	0.22	0.21	0.48	0.44	0.53	0.32	0.39	0.35	0.58	0.22	0.81	0.73	0.71	0.52
FI	2021	0.66	0.42	1.00	0.72	0.74	0.40	0.49	0.94	0.85	0.83	0.15	0.79	0.53	0.64	0.97	0.97	0.70	0.68	0.98	0.75	0.60	1.00	0.56	0.48	0.61	0.84	0.58	0.78	0.46	0.04	0.74	0.56
FR	2021	0.59	0.69	0.37	0.27	0.52	0.65	0.69	0.58	0.61	0.84	1.00	0.61	0.42	0.64	0.31	0.50	0.61	0.53	0.45	0.23	0.51	0.74	0.35	0.33	0.55	0.68	0.77	0.61	0.40	0.68	0.83	0.52
HR	2021	0.25	0.38	0.20	0.25	0.21	0.11	0.00	0.58	0.44	0.49	0.02	0.21	0.75	0.20	0.55	0.47	0.75	0.65	0.42	0.28	0.46	0.27	0.25	0.06	0.37	0.48	0.47	0.10	0.32	0.41	0.57	0.31
HU	2021	0.18	0.22	0.18	0.17	0.29	0.39	0.31	0.38	0.25	0.52	0.70	0.47	0.59	0.29	0.35	0.40	0.33	0.15	0.31	0.21	0.31	0.42	0.31	0.13	0.44	0.16	0.98	0.46	0.30	0.25	0.73	0.38
IE	2021	0.66	1.00	0.47	0.59	0.71	0.57	0.51	0.76	0.13	0.67	0.42	0.38	0.42	0.73	0.79	0.79	0.52	0.51	0.76	0.48		0.50	0.36	0.19	1.00	0.70	0.83	0.98	1.00	0.59	0.90	0.19
IT	2021	0.38	0.10	0.33	0.28	0.66	0.25	0.27	0.40	0.40	0.44	0.71	0.38	0.51	0.41	0.44	0.39	0.71	0.81	0.49	0.27	0.25	0.55	0.50	0.72	0.48	0.66	0.68	0.43	0.51	0.96	0.80	0.40
LT	2021	0.25	0.89	0.28	0.23	0.25	0.16	0.65	0.41	0.45	0.62	0.08	0.17	0.92	0.32	0.18	0.39	0.58	0.61	0.46	0.15	0.65	0.28	0.49	0.16	0.40	0.60	0.48	0.15	0.61	0.26	0.75	0.79
LU	2021	0.52	1.00	0.63	0.83	0.69	1.00	0.73	0.60	0.43	0.76	0.12	0.26	0.26	0.37	0.68	0.85	0.54	0.54	0.40	0.63	0.57	0.50	0.95	0.91	1.00	0.78	0.52	0.96	0.18	0.90	0.79	0.40
LV	2021	0.18	0.56	0.28	0.19	0.20	0.17	0.41	0.42	0.36	0.25	0.02	0.06	0.53	0.11	0.39	0.39	0.26	0.26	0.19	0.16	0.34	0.35	0.46	0.22	0.33	0.22	0.37	0.49	0.32	0.27	0.30	0.00
MT	2021	0.11	0.49	0.48	0.31	0.36	1.00	0.64	0.67	0.10	0.13	0.12	0.14	0.61	0.24	0.65	0.55	0.58	0.55	0.26	0.18	0.79	0.41	1.00	0.78	0.78	0.26	0.82	0.45	0.32	0.57	0.97	1.00
NL	2021	0.52	0.84	0.95	0.66	0.97	0.79	0.72	1.00	0.62	0.72	0.66	0.62	0.21	0.51	0.63	0.85	0.50	0.54	0.47	0.63	0.81	0.85	0.52	0.50	0.79	0.74	0.72	0.82	0.42	0.93	0.83	0.43
PL	2021	0.18	0.46	0.17	0.13	0.25	0.03	0.42	0.36	0.38	0.37	0.37	0.34	0.66	0.28	0.26	0.37	0.15	0.12	0.10	0.10	0.54	0.26	0.39	0.63	0.30	0.07	0.63	0.38	0.24	0.19	0.60	0.50
PT	2021	0.45	0.56	0.42	0.44	0.55	0.53	0.78	0.53	0.53	0.54	0.60	0.30	0.63	0.29	0.70	0.50	0.51	0.39	0.22	0.25	0.48	0.37	0.52	0.43	0.38	0.84	0.52	0.28	0.39	0.27	0.08	0.49
RO	2021	0.11	0.00	0.15	0.07	0.24	0.06	0.65	0.10	0.08	0.48	0.11	0.10	0.26	0.01	0.04	0.19	0.09	0.00	0.03	0.07	0.05	0.17	0.24	0.12	0.11	0.00	0.80	0.42	0.17	0.09	0.56	0.22
SE	2021	0.79	0.67	1.00	0.80	0.81	0.60	0.95	0.68	0.86	0.74	0.35	1.00	0.54	0.94	0.84	1.00	0.85	0.69	0.49	0.76	0.43	1.00	0.58	0.52	0.84	0.64	0.75	0.80	0.35	0.28	0.81	0.57
SI	2021	0.52	0.43	0.70	0.50	0.45	0.31	0.40	0.34	0.41	0.24	0.54	0.64	0.51	0.29	0.72	0.55	0.69	0.47	0.43	0.52	0.49	0.62	0.51	0.28	0.63	0.43	0.86	0.25	0.34	0.40	0.70	0.45
SK	2021	0.38	0.40	0.44	0.23	0.17	0.16	0.26	0.36	0.26	0.24	0.11	0.18	0.53	0.20	0.40	0.47	0.20	0.19	0.26	0.16	0.17	0.29	0.34	0.18	0.33	0.22	1.00	0.30	0.85	0.31	0.80	0.75
AT	2022	0.59	0.48	0.51	0.66	0.66	0.63	0.38	0.63	0.87	0.53	0.79	0.96	0.34	0.55	0.39	0.50	0.58	0.75	0.57	0.83	0.52	0.82	0.65	0.82	0.52	0.78	0.75	0.46	0.62	0.52	0.92	0.51
BE	2022	0.59	0.71	0.34	0.65	0.77	0.55	0.81	0.48	0.77	0.75	0.98	1.00	0.56	1.00	0.84	0.68	0.68	1.00	0.86	0.57	0.51	0.68	0.48	0.33	0.65	0.86	0.69	0.74	0.65	0.71	0.80	0.43
BG	2022	0.18	0.25	0.02	0.09	0.09	0.13	0.37	0.08	0.15	0.25	0.02	0.23	0.34	0.08	0.05	0.39	0.39	0.25	0.23	0.07	0.04	0.26	0.53	0.55	0.33	0.23	0.42	0.54	0.24	0.07	0.37	0.66
CY	2022	0.25	0.91	0.32	0.97	0.48	0.40	0.45	0.36	0.24	0.82	0.05	0.14	0.63	0.23	0.62	0.40	0.78	1.00	0.99	0.64	0.86	0.28	1.00	0.29	0.75	0.84	0.74	0.98	0.47	0.29	0.61	0.21
CZ	2022	0.45	0.28	0.18	0.38	0.26	0.37	0.35	0.43	0.67	0.59	0.37	0.51	0.63	0.34	0.60	0.52	0.69	0.79	0.50	0.29	0.33	0.32	0.37	0.30	0.46	0.50	0.91	0.47	0.52	0.39	0.89	0.53
DE	2022	0.73	0.33	0.25	0.34	0.66	0.39	0.55	0.32	0.86	0.64	0.20	0.90	0.77	0.94	0.57	0.56	0.67	0.84	0.46	0.40	0.71	0.94	0.52	0.65	0.49	0.91	0.92	0.80	0.61	0.65	0.93	0.59
DK	2022	0.66	0.66	0.80	1.00	0.88	0.62	1.00	0.71	1.00	0.93	0.22	0.78	0.42	0.88	0.77	0.73	0.61	0.66	0.50	1.00	0.90	0.92	0.59	1.00	0.61	0.64	0.66	0.83	0.42	0.33	0.97	1.00
EE	2022	0.38	0.50	0.65	0.62	0.51	0.43	0.37	0.51	0.66	1.00	0.13	0.40	1.00	0.49	0.37	0.77	0.50	0.80	0.89	0.43	0.85	0.44	0.89	0.53	0.53	0.99	0.43	0.60	0.54	0.11	0.59	0.66
EL	2022	0.32	0.53	0.09	0.31	0.52	0.00	0.20	0.38	0.70	0.35	0.13	0.29	0.62	0.40	0.21	0.16	0.99	1.00	0.68	0.29	0.24	0.32	0.42	0.10	0.39	0.72	0.22	0.75	1.00	0.38	0.56	0.45
ES	2022	0.45	0.64	0.49	0.32	0.55	0.32	0.84	0.73	0.53	0.71	0.53	0.32	0.46	0.35	0.46	0.44	0.29	0.28	0.23	0.24	0.59	0.46	0.55	0.32	0.39	0.22	0.55	0.48	0.67	0.79	0.70	0.52
FI	2022	0.73	0.42	1.00	0.78	0.79	0.42	0.71	0.94	0.85	0.94	0.13	0.83	0.45	0.63	1.00	0.97	0.75	0.82	0.99	0.82	0.65	1.00	0.57	0.59	0.61	0.80	0.55	0.86	0.59	0.06	0.77	0.55
FR	2022	0.52	0.69	0.37	0.28	0.52	0.65	0.69	0.58	0.64	0.92	1.00	0.63	0.40	0.62	0.31	0.50	0.49	0.63	0.53	0.24	0.59	0.73	0.34	0.30	0.55	0.60	0.72	0.68	0.35	0.70	0.84	0.51
HR	2022	0.25	0.38	0.20	0.28	0.23	0.12	0.19	0.58	0.55	0.70	0.03	0.24	0.49	0.15	0.55	0.47	0.68	0.69	0.42	0.32	0.59	0.28	0.29	0.08	0.37	0.57	0.43	0.21	0.47	0.42	0.57	0.28
HU	2022	0.18	0.22	0.18	0.19	0.36	0.43	0.35	0.38	0.26	0.57	0.67	0.52	0.41	0.26	0.34	0.40	0.34	0.24	0.33	0.24	0.31	0.46	0.32	0.12	0.44	0.15	0.93	0.51	0.35	0.26	0.74	0.33
IE	2022	0.66	1.00	0.47	0.68	0.73	0.61	0.51	0.76	0.12	0.64	0.60	0.38	0.42	0.78	0.67	0.79	0.55	0.70	0.51	0.52		0.45	0.36	0.21	1.00	0.36	0.87	1.00	1.00	0.60	0.92	0.14
IT	2022	0.38	0.10	0.33	0.31	0.68	0.26	0.36	0.40	0.44	0.46	0.61	0.39	0.62	0.58	0.32	0.39	0.57	0.67	0.45	0.30	0.23	0.55	0.52	0.70	0.48	0.79	0.65	0.60	0.70	0.96	0.79	0.40
LT	2022	0.25	0.89	0.28	0.26	0.29	0.09	0.67	0.41	0.49	0.83	0.09	0.21	0.78	0.28	0.27	0.39	0.58	0.64	0.49	0.15	0.91	0.26	0.54	0.18	0.40	0.62	0.46	0.21	0.38	0.26	0.81	0.43
LU	2022	0.45	1.00	0.63	0.92	0.91	1.00	0.72	0.60	0.44	0.67	0.11	0.22	0.28	0.39	0.49	0.85	0.41	0.46	0.35	0.87	0.86	0.47	0.92	0.71	1.00	0.44	0.51	0.89	0.24	0.90	0.81	0.34
LV	2022	0.11	0.56	0.28	0.22	0.14	0.18	0.36	0.42	0.37	0.31	0.02	0.10	0.40	0.07	0.36	0.39	0.19	0.26	0.19	0.22	0.35	0.35	0.49	0.25	0.33	0.27	0.34	0.47	0.33	0.28	0.31	0.07
MT	2022	0.11	0.49	0.48	0.40	0.28	1.00	0.72	0.67	0.12	0.13	0.08	0.16	0.55	0.23	0.69	0.55	0.28	0.46	0.24	0.25	0.91	0.40	1.00	0.50	0.78	0.55	0.76	0.45	0.38	0.68	0.99	1.00
NL	2022	0.38	0.84	0.95	0.71	0.97	0.82	0.80	1.00	0.68	0.90	0.68	0.66	0.21	0.51	0.57	0.85	0.52	0.61	0.63	0.70	0.81	0.85	0.54	0.57	0.79	0.55	0.66	0.83	0.32	0.98	0.84	0.45
PL	2022	0.11	0.46	0.17	0.15	0.24	0.12	0.51	0.36	0.40	0.34	0.39	0.36	0.49	0.22	0.39	0.37	0.20	0.27	0.21	0.12	0.56	0.27	0.41	0.66	0.30	0.09	0.63	0.43	0.25	0.23	0.63	0.49
PT	2022	0.45	0.56	0.42	0.48	0.55	0.56	0.85	0.53	0.56	0.38	0.84	0.38	0.34	0.16	0.55	0.50	0.45	0.62	0.20	0.29	0.58	0.38	0.52	0.37	0.38	0.38	0.49	0.41	0.50	0.28	0.07	0.42
RO	2022	0.11	0.00	0.15	0.08	0.29	0.06	0.79	0.10	0.07	0.41	0.11	0.10	0.15	0.05	0.03	0.19	0.02	0.00	0.01	0.08	0.00	0.16	0.26	0.11	0.11	0.00	0.73	0.50	0.35	0.07	0.55	0.49
SE	2022	0.66	0.67	1.00	0.85	0.81	0.61	0.95	0.68	0.87	0.89	0.35	1.00	0.44	0.95	0.83	1.00	0.76	0.77	0.52	0.82	0.43	1.00	0.59	0.52	0.84	0.78	0.70	0.84	0.56	0.27	0.81	0.58
SI	2022	0.45	0.43	0.70	0.57	0.45	0.33	0.55	0.34	0.44	0.26	0.62	0.67	0.19	0.30	0.63	0.55	0.68	0.58	0.45	0.58	0.42	0.49	0.55	0.26	0.63	0.52	0.83	0.34	0.50	0.46	0.71	0.40
SK	2022	0.38	0.40	0.44	0.23	0.19	0.18	0.33	0.36	0.31	0.31	0.12	0.20	0.56	0.25	0.34	0.47	0.20	0.28	0.24	0.18	0.22	0.27	0.34	0.16	0.33	0.21	0.94	0.40	0.45	0.38	0.82	0.69

Code	**Year**	**1.1.1**	**1.1.2**	**1.1.3**	**1.2.1**	**1.2.2**	**1.2.3**	**1.3.1**	**1.3.2**	**2.1.1**	**2.1.2**	**2.1.3**	**2.2.1**	**2.2.2**	**2.2.3**	**2.3.1**	**2.3.2**	**3.1.1**	**3.1.2**	**3.2.1**	**3.2.2**	**3.2.3**	**3.3.1**	**3.3.2**	**3.3.3**	**4.1.1**	**4.1.2**	**4.2.1**	**4.2.2**	**4.2.3**	**4.3.1**	**4.3.2**	**4.3.3**
AT	2023	0.59	0.50	0.55	0.64	0.63	0.67	0.37	0.63	0.91	0.55	1.00	0.97	0.34	0.55	0.39	0.58	0.58	0.75	0.57	0.80	0.67	0.81	0.68	0.92	0.56	0.78	0.73	0.51	0.62	0.48	0.91	0.56
BE	2023	0.59	0.72	0.35	0.63	0.74	0.44	0.79	0.48	0.75	0.76	0.93	1.00	0.56	1.00	0.84	0.68	0.68	1.00	0.86	0.56	0.53	0.66	0.48	0.33	0.66	0.86	0.60	0.76	0.65	0.77	0.79	0.44
BG	2023	0.18	0.25	0.02	0.10	0.08	0.14	0.47	0.08	0.14	0.30	0.02	0.21	0.34	0.08	0.05	0.39	0.39	0.25	0.23	0.08	0.08	0.27	0.57	0.40	0.33	0.23	0.35	0.53	0.24	0.07	0.43	0.53
CY	2023	0.18	0.93	0.35	1.00	0.63	0.45	0.71	0.36	0.25	0.57	0.08	0.13	0.63	0.23	0.62	0.52	0.78	1.00	0.99	0.62	0.85	0.31	1.00	0.43	0.80	0.84	0.69	0.91	0.47	0.29	0.58	0.06
CZ	2023	0.52	0.27	0.31	0.36	0.27	0.41	0.42	0.43	0.64	0.62	0.34	0.53	0.63	0.34	0.60	0.50	0.69	0.79	0.50	0.29	0.28	0.34	0.42	0.32	0.44	0.50	0.91	0.50	0.52	0.41	0.88	0.50
DE	2023	0.79	0.33	0.26	0.33	0.64	0.38	0.66	0.32	0.86	0.66	0.19	0.90	0.77	0.94	0.57	0.58	0.67	0.84	0.46	0.40	0.71	0.92	0.54	0.62	0.51	0.91	0.90	0.82	0.61	0.65	0.92	0.61
DK	2023	0.73	0.66	1.00	1.00	0.79	0.62	1.00	0.71	0.94	0.96	0.40	0.73	0.42	0.88	0.77	0.73	0.61	0.66	0.50	1.00	0.91	0.95	0.58	0.91	0.61	0.64	0.63	0.87	0.42	0.35	0.96	1.00
EE	2023	0.45	0.52	0.75	0.60	0.57	0.49	0.42	0.51	0.65	1.00	0.16	0.41	1.00	0.49	0.37	0.84	0.50	0.80	0.89	0.41	0.90	0.44	0.95	0.60	0.52	0.99	0.38	0.64	0.54	0.12	0.55	0.55
EL	2023	0.45	0.56	0.09	0.32	0.54	0.01	0.14	0.38	0.66	0.37	0.23	0.28	0.62	0.40	0.21	0.18	0.99	1.00	0.68	0.31	0.34	0.30	0.46	0.13	0.37	0.72	0.16	0.64	1.00	0.39	0.58	0.36
ES	2023	0.52	0.69	0.53	0.32	0.55	0.32	0.93	0.73	0.52	0.77	0.59	0.33	0.46	0.35	0.46	0.47	0.29	0.28	0.23	0.24	0.60	0.45	0.56	0.31	0.41	0.22	0.53	0.41	0.67	0.68	0.67	0.51
FI	2023	0.59	0.43	0.90	0.78	0.75	0.45	0.86	0.94	0.81	1.00	0.21	0.88	0.45	0.63	1.00	1.00	0.75	0.82	0.99	0.79	0.69	1.00	0.58	0.61	0.65	0.80	0.51	0.88	0.59	0.06	0.77	0.54
FR	2023	0.52	0.70	0.46	0.27	0.53	0.63	0.69	0.58	0.61	1.00	1.00	0.62	0.40	0.62	0.31	0.47	0.49	0.63	0.53	0.24	0.68	0.70	0.35	0.31	0.55	0.60	0.69	0.69	0.35	0.76	0.84	0.53
HR	2023	0.32	0.38	0.20	0.29	0.23	0.12	0.27	0.58	0.56	1.00	0.08	0.24	0.49	0.15	0.55	0.47	0.68	0.69	0.42	0.31	0.46	0.28	0.32	0.12	0.37	0.57	0.35	0.13	0.47	0.37	0.57	0.32
HU	2023	0.25	0.20	0.26	0.20	0.33	0.47	0.43	0.38	0.29	0.55	0.66	0.52	0.41	0.26	0.34	0.44	0.34	0.24	0.33	0.24	0.45	0.40	0.34	0.09	0.45	0.15	0.90	0.52	0.35	0.31	0.74	0.35
IE	2023	0.66	1.00	0.40	0.66	0.71	0.64	0.62	0.76	0.10	0.55	0.60	0.37	0.42	0.78	0.67	0.79	0.55	0.70	0.51	0.51		0.47	0.37	0.22	0.97	0.36	0.84	1.00	1.00	0.71	0.93	0.18
IT	2023	0.45	0.12	0.32	0.31	0.73	0.15	0.47	0.40	0.43	0.44	0.29	0.36	0.62	0.58	0.32	0.40	0.57	0.67	0.45	0.31	0.25	0.55	0.54	0.67	0.48	0.79	0.63	0.59	0.70	0.97	0.78	0.38
LT	2023	0.32	0.91	0.28	0.24	0.31	0.12	0.66	0.41	0.45	0.83	0.12	0.22	0.78	0.28	0.27	0.48	0.58	0.64	0.49	0.14	0.90	0.28	0.63	0.25	0.43	0.62	0.39	0.30	0.38	0.23	0.81	0.45
LU	2023	0.66	0.98	0.64	0.99	0.82	1.00	0.82	0.60	0.40	0.77	0.12	0.21	0.28	0.39	0.49	1.00	0.41	0.46	0.35	0.91	0.83	0.49	0.87	0.53	1.00	0.44	0.50	0.89	0.24	0.92	0.83	0.38
LV	2023	0.18	0.57	0.32	0.21	0.22	0.21	0.36	0.42	0.36	0.36	0.04	0.10	0.40	0.07	0.36	0.48	0.19	0.26	0.19	0.21	0.32	0.33	0.51	0.28	0.38	0.27	0.30	0.52	0.33	0.28	0.31	0.16
MT	2023	0.11	0.48	0.45	0.45	0.50	1.00	0.81	0.67	0.11	0.15	0.01	0.17	0.55	0.23	0.69	0.58	0.28	0.46	0.24	0.29	0.73	0.43	1.00	0.64	0.83	0.55	0.80	0.40	0.38	0.55	1.00	0.31
NL	2023	0.45	0.86	0.94	0.69	0.92	0.83	0.86	1.00	0.67	0.94	0.74	0.64	0.21	0.51	0.57	0.94	0.52	0.61	0.63	0.68	0.81	0.84	0.54	0.54	0.83	0.55	0.59	0.81	0.32	1.00	0.83	0.43
PL	2023	0.11	0.46	0.25	0.15	0.28	0.15	0.55	0.36	0.42	0.36	0.42	0.38	0.49	0.22	0.39	0.37	0.20	0.27	0.21	0.12	0.59	0.26	0.45	0.68	0.30	0.09	0.61	0.44	0.25	0.24	0.64	0.41
PT	2023	0.45	0.48	0.46	0.48	0.52	0.56	0.94	0.53	0.54	0.36	1.00	0.42	0.34	0.16	0.55	0.47	0.45	0.62	0.20	0.31	0.74	0.38	0.54	0.37	0.37	0.38	0.47	0.42	0.50	0.30	0.05	0.32
RO	2023	0.18	0.00	0.16	0.09	0.34	0.06	0.95	0.10	0.06	0.27	0.07	0.11	0.15	0.05	0.03	0.23	0.02	0.00	0.01	0.09	0.00	0.14	0.29	0.10	0.12	0.00	0.67	0.48	0.35	0.05	0.56	0.53
SE	2023	0.66	0.75	1.00	0.84	0.75	0.61	0.95	0.68	0.84	1.00	0.37	1.00	0.44	0.95	0.83	1.00	0.76	0.77	0.52	0.82	0.46	1.00	0.61	0.46	0.92	0.78	0.67	0.81	0.56	0.29	0.81	0.54
SI	2023	0.45	0.43	0.70	0.54	0.48	0.35	0.60	0.34	0.44	0.31	0.56	0.66	0.19	0.30	0.63	0.50	0.68	0.58	0.45	0.55	0.48	0.49	0.57	0.26	0.60	0.52	0.82	0.36	0.50	0.46	0.72	0.31
SK	2023	0.45	0.39	0.44	0.22	0.21	0.20	0.38	0.36	0.29	0.29	0.17	0.21	0.56	0.25	0.34	0.47	0.20	0.28	0.24	0.17	0.14	0.27	0.38	0.20	0.34	0.21	0.91	0.40	0.45	0.39	0.83	0.63
AT	2024	0.52	0.51	0.60	0.63	0.64	0.69	0.47	0.60	0.89	0.54	0.92	0.94	0.36	0.59	0.46	0.63	0.58	0.75	0.57	0.81	0.55	0.83	0.65	0.78	0.56	0.78	0.75	0.43	0.53	0.50	0.93	0.57
BE	2024	0.59	0.69	0.38	0.62	0.73	0.43	0.76	0.52	0.80	0.80	0.93	1.00	0.48	1.00	0.86	0.65	0.68	1.00	0.86	0.55	0.47	0.70	0.45	0.30	0.65	0.95	0.65	0.74	0.62	0.87	0.79	0.37
BG	2024	0.25	0.30	0.02	0.10	0.14	0.16	0.48	0.08	0.12	0.36	0.01	0.21	0.34	0.09	0.13	0.47	0.21	0.10	0.17	0.09	0.08	0.27	0.58	0.64	0.35	0.32	0.39	0.50	0.29	0.07	0.30	0.52
CY	2024	0.18	1.00	0.37	1.00	0.71	0.45	0.88	0.45	0.25	0.40	0.08	0.12	0.49	0.19	0.72	0.65	0.78	1.00	0.99	0.65	0.89	0.36	1.00	0.38	0.87	0.82	0.56	0.86	0.57	0.29	0.53	0.31
CZ	2024	0.52	0.24	0.33	0.35	0.30	0.44	0.45	0.67	0.59	0.62	0.33	0.53	0.87	0.63	0.55	0.47	0.47	0.53	0.46	0.27	0.24	0.34	0.43	0.30	0.47	0.68	0.95	0.45	0.59	0.42	0.88	0.49
DE	2024	0.73	0.37	0.27	0.32	0.64	0.38	0.69	0.34	0.86	0.67	0.21	0.90	0.74	0.85	0.68	0.56	0.49	0.75	0.39	0.38	0.71	0.91	0.52	0.52	0.54	0.89	0.92	0.75	0.58	0.68	0.93	0.58
DK	2024	0.66	0.66	1.00	1.00	0.79	0.64	1.00	0.76	1.00	1.00	0.40	0.76	0.56	0.72	0.87	0.73	0.61	0.66	0.50	1.00	0.91	0.98	0.54	0.68	0.61	0.59	0.63	0.80	0.62	0.38	0.96	1.00
EE	2024	0.52	0.51	0.83	0.59	0.64	0.55	0.48	0.66	0.67	1.00	0.19	0.42	0.80	0.44	0.42	0.85	0.55	0.59	0.69	0.39	0.78	0.42	0.97	0.54	0.57	0.83	0.50	0.63	0.42	0.14	0.70	0.25
EL	2024	0.45	0.54	0.09	0.32	0.55	0.03	0.26	0.34	0.65	0.38	0.33	0.30	0.59	0.40	0.26	0.16	0.95	0.96	0.44	0.31	0.29	0.30	0.45	0.11	0.37	0.93	0.20	0.56	0.85	0.44	0.62	0.29
ES	2024	0.52	0.74	0.55	0.32	0.53	0.33	0.98	0.74	0.52	0.76	0.62	0.34	0.45	0.31	0.48	0.48	0.27	0.29	0.29	0.24	0.51	0.48	0.53	0.30	0.42	0.27	0.57	0.28	0.91	0.73	0.69	0.43
FI	2024	0.66	0.39	0.93	0.76	0.75	0.47	0.93	1.00	0.84	1.00	0.20	0.86	0.40	0.66	1.00	1.00	0.65	0.65	0.90	0.75	0.67	1.00	0.54	0.46	0.67	0.86	0.57	0.83	0.81	0.08	0.79	0.53
FR	2024	0.73	0.74	0.52	0.26	0.52	0.62	0.74	0.57	0.60	1.00	1.00	0.61	0.27	0.69	0.31	0.53	0.51	0.65	0.51	0.22	0.56	0.69	0.34	0.30	0.57	0.63	0.72	0.64	0.24	0.74	0.84	0.52
HR	2024	0.45	0.38	0.20	0.28	0.25	0.12	0.35	0.45	0.55	0.90	0.04	0.32	0.35	0.12	0.48	0.47	0.68	0.69	0.42	0.32	0.51	0.26	0.34	0.15	0.37	0.63	0.45	0.11	0.54	0.44	0.61	0.10
HU	2024	0.25	0.13	0.32	0.21	0.32	0.49	0.52	0.52	0.27	0.49	0.73	0.42	0.49	0.33	0.40	0.45	0.28	0.19	0.31	0.24	0.49	0.41	0.32	0.08	0.45	0.23	0.97	0.53	0.30	0.36	0.73	0.24
IE	2024	0.73	1.00	0.42	0.65	0.70	0.66	0.68	0.85	0.08	0.51	0.58	0.32	0.22	0.99	0.56	0.77	0.46	0.52	0.75	0.51		0.50	0.34	0.15	0.98	0.63	0.88	0.98	1.00	0.84	0.95	0.19
IT	2024	0.52	0.16	0.39	0.31	0.74	0.19	0.51	0.39	0.42	0.46	0.30	0.32	0.47	0.44	0.44	0.44	0.71	0.86	0.52	0.32	0.25	0.58	0.51	0.65	0.51	0.67	0.67	0.47	0.55	0.96	0.79	0.35
LT	2024	0.32	0.89	0.36	0.25	0.37	0.15	0.70	0.47	0.41	0.91	0.17	0.20	0.94	0.39	0.25	0.56	0.40	0.61	0.39	0.15	0.92	0.27	0.66	0.25	0.51	0.69	0.42	0.28	0.47	0.26	0.82	0.36
LU	2024	0.93	0.96	0.57	0.98	0.69	1.00	0.86	0.51	0.37	0.69	0.12	0.20	0.27	0.30	0.51	1.00	0.41	0.46	0.35	0.90	0.77	0.56	0.80	0.44	1.00	0.43	0.53	0.87	0.13	0.93	0.83	0.44
LV	2024	0.18	0.55	0.36	0.21	0.23	0.20	0.39	0.27	0.37	0.36	0.02	0.10	0.31	0.06	0.31	0.48	0.20	0.29	0.21	0.22	0.44	0.37	0.52	0.23	0.38	0.22	0.35	0.56	0.24	0.27	0.34	0.24
MT	2024	0.11	0.59	0.58	0.43	0.48	1.00	0.93	0.71	0.11	0.15	0.01	0.18	0.40	0.21	0.72	0.53	0.21	0.33	0.17	0.28	0.58	0.52	1.00	0.63	0.83	0.39	0.86	0.34	0.25	0.67	1.00	0.35
NL	2024	0.45	0.81	0.94	0.68	0.92	0.83	0.84	1.00	0.64	0.94	0.73	0.66	0.21	0.51	0.74	0.89	0.52	0.61	0.63	0.66	0.70	0.86	0.51	0.47	0.81	0.68	0.65	0.78	0.36	1.00	0.83	0.37
PL	2024	0.18	0.59	0.29	0.15	0.31	0.03	0.61	0.34	0.39	0.31	0.43	0.40	0.46	0.24	0.60	0.47	0.20	0.27	0.21	0.12	0.50	0.28	0.45	0.66	0.38	0.29	0.64	0.41	0.29	0.26	0.62	0.27
PT	2024	0.52	0.44	0.46	0.47	0.49	0.57	0.98	0.56	0.49	0.34	0.94	0.45	0.35	0.19	0.57	0.50	0.49	0.57	0.28	0.31	0.52	0.38	0.51	0.35	0.39	0.61	0.51	0.31	0.60	0.30	0.05	0.25
RO	2024	0.18	0.00	0.21	0.09	0.35	0.05	1.00	0.11	0.05	0.26	0.07	0.10	0.08	0.07	0.12	0.19	0.03	0.00	0.00	0.08	0.00	0.16	0.31	0.11	0.10	0.00	0.72	0.47	0.20	0.07	0.55	0.29
SE	2024	0.73	0.80	1.00	0.83	0.75	0.61	0.95	0.70	0.79	0.95	0.36	1.00	0.40	0.92	0.89	1.00	0.73	0.65	1.00	0.80	0.21	1.00	0.58	0.41	0.92	0.81	0.74	0.81	0.52	0.28	0.85	0.49
SI	2024	0.52	0.43	0.70	0.53	0.45	0.37	0.61	0.32	0.49	0.30	0.56	0.63	0.15	0.27	0.73	0.39	0.71	0.54	0.43	0.50	0.55	0.53	0.57	0.33	0.53	0.68	0.91	0.32	0.50	0.47	0.74	0.31
SK	2024	0.38	0.41	0.35	0.22	0.23	0.22	0.42	0.38	0.31	0.30	0.25	0.23	0.55	0.26	0.32	0.45	0.23	0.25	0.30	0.17	0.15	0.28	0.39	0.16	0.34	0.30	0.99	0.34	0.62	0.42	0.81	0.51

ANNEX 2: IRPI, IKP AND IKC SCORES FOR EU-27 COUNTRIES (2017-2024)

Country Code	Year	IKP	IKC	IRPI	Country Code	Year	IKP	IKC	IRPI
AT	2017	0.70	0.58	0.64	IE	2017	0.55	0.84	0.69
AT	2018	0.70	0.57	0.64	IE	2018	0.55	0.87	0.71
AT	2019	0.69	0.57	0.63	IE	2019	0.56	0.87	0.71
AT	2020	0.68	0.58	0.63	IE	2020	0.58	0.87	0.73
AT	2021	0.66	0.59	0.62	IE	2021	0.58	0.91	0.74
AT	2022	0.67	0.60	0.64	IE	2022	0.53	0.93	0.73
AT	2023	0.68	0.62	0.65	IE	2023	0.51	0.92	0.72
AT	2024	0.66	0.59	0.63	IE	2024	0.53	0.93	0.73
BE	2017	0.66	0.67	0.66	IT	2017	0.53	0.57	0.55
BE	2018	0.66	0.67	0.66	IT	2018	0.54	0.57	0.56
BE	2019	0.65	0.67	0.66	IT	2019	0.54	0.57	0.55
BE	2020	0.71	0.69	0.70	IT	2020	0.64	0.55	0.60
BE	2021	0.71	0.70	0.71	IT	2021	0.64	0.55	0.60
BE	2022	0.75	0.72	0.73	IT	2022	0.61	0.62	0.62
BE	2023	0.75	0.68	0.71	IT	2023	0.62	0.61	0.61
BE	2024	0.76	0.69	0.73	IT	2024	0.66	0.57	0.61
BG	2017	0.27	0.35	0.31	LT	2017	0.42	0.25	0.33
BG	2018	0.25	0.35	0.30	LT	2018	0.42	0.26	0.34
BG	2019	0.22	0.36	0.29	LT	2019	0.43	0.27	0.35
BG	2020	0.24	0.37	0.31	LT	2020	0.47	0.29	0.38
BG	2021	0.24	0.40	0.32	LT	2021	0.47	0.31	0.39
BG	2022	0.29	0.48	0.38	LT	2022	0.52	0.34	0.43
BG	2023	0.27	0.44	0.36	LT	2023	0.53	0.35	0.44
BG	2024	0.28	0.45	0.36	LT	2024	0.52	0.35	0.43
CY	2017	0.40	0.78	0.59	LU	2017	0.69	0.80	0.74
CY	2018	0.42	0.74	0.58	LU	2018	0.70	0.75	0.73
CY	2019	0.41	0.77	0.59	LU	2019	0.69	0.74	0.72
CY	2020	0.69	0.74	0.71	LU	2020	0.63	0.79	0.71
CY	2021	0.72	0.78	0.75	LU	2021	0.66	0.74	0.70
CY	2022	0.74	0.86	0.80	LU	2022	0.56	0.70	0.63
CY	2023	0.75	0.80	0.77	LU	2023	0.54	0.69	0.62
CY	2024	0.73	0.71	0.72	LU	2024	0.50	0.70	0.60
CZ	2017	0.37	0.62	0.50	LV	2017	0.24	0.42	0.33
CZ	2018	0.36	0.64	0.50	LV	2018	0.27	0.43	0.35
CZ	2019	0.38	0.65	0.52	LV	2019	0.25	0.43	0.34
CZ	2020	0.41	0.66	0.54	LV	2020	0.30	0.43	0.36
CZ	2021	0.39	0.67	0.53	LV	2021	0.23	0.43	0.33
CZ	2022	0.52	0.69	0.60	LV	2022	0.23	0.40	0.32
CZ	2023	0.53	0.70	0.62	LV	2023	0.25	0.41	0.33
CZ	2024	0.48	0.70	0.59	LV	2024	0.25	0.45	0.35
DE	2017	0.60	0.85	0.72	MT	2017	0.37	0.73	0.55

Country Code	Year	IKP	IKC	IRPI	Country Code	Year	IKP	IKC	IRPI
DE	2018	0.60	0.86	0.73	MT	2018	0.51	0.63	0.57
DE	2019	0.61	0.86	0.73	MT	2019	0.47	0.58	0.53
DE	2020	0.71	0.85	0.78	MT	2020	0.56	0.58	0.57
DE	2021	0.69	0.84	0.77	MT	2021	0.42	0.63	0.53
DE	2022	0.69	0.86	0.77	MT	2022	0.35	0.60	0.48
DE	2023	0.68	0.86	0.77	MT	2023	0.40	0.60	0.50
DE	2024	0.62	0.84	0.73	MT	2024	0.34	0.60	0.47
DK	2017	0.59	0.69	0.64	NL	2017	0.63	0.72	0.68
DK	2018	0.59	0.66	0.63	NL	2018	0.64	0.71	0.68
DK	2019	0.59	0.67	0.63	NL	2019	0.66	0.72	0.69
DK	2020	0.66	0.68	0.67	NL	2020	0.63	0.73	0.68
DK	2021	0.67	0.72	0.70	NL	2021	0.63	0.77	0.70
DK	2022	0.74	0.74	0.74	NL	2022	0.68	0.74	0.71
DK	2023	0.72	0.75	0.73	NL	2023	0.67	0.70	0.69
DK	2024	0.69	0.72	0.70	NL	2024	0.68	0.72	0.70
EE	2017	0.47	0.46	0.47	PL	2017	0.23	0.48	0.35
EE	2018	0.47	0.45	0.46	PL	2018	0.23	0.48	0.36
EE	2019	0.52	0.45	0.48	PL	2019	0.23	0.48	0.36
EE	2020	0.74	0.46	0.60	PL	2020	0.24	0.50	0.37
EE	2021	0.76	0.48	0.62	PL	2021	0.24	0.51	0.37
EE	2022	0.75	0.51	0.63	PL	2022	0.29	0.53	0.41
EE	2023	0.76	0.51	0.64	PL	2023	0.30	0.52	0.41
EE	2024	0.69	0.57	0.63	PL	2024	0.32	0.52	0.42
EL	2017	0.50	0.32	0.41	PT	2017	0.60	0.39	0.50
EL	2018	0.51	0.31	0.41	PT	2018	0.58	0.38	0.48
EL	2019	0.51	0.32	0.41	PT	2019	0.60	0.39	0.49
EL	2020	0.58	0.34	0.46	PT	2020	0.50	0.40	0.45
EL	2021	0.58	0.38	0.48	PT	2021	0.50	0.40	0.45
EL	2022	0.62	0.49	0.56	PT	2022	0.42	0.45	0.44
EL	2023	0.63	0.40	0.52	PT	2023	0.42	0.44	0.43
EL	2024	0.62	0.38	0.50	PT	2024	0.45	0.41	0.43
ES	2017	0.35	0.42	0.38	RO	2017	0.06	0.54	0.30
ES	2018	0.35	0.40	0.37	RO	2018	0.08	0.56	0.32
ES	2019	0.35	0.39	0.37	RO	2019	0.11	0.57	0.34
ES	2020	0.34	0.39	0.36	RO	2020	0.15	0.59	0.37
ES	2021	0.35	0.40	0.37	RO	2021	0.14	0.61	0.37
ES	2022	0.37	0.51	0.44	RO	2022	0.12	0.62	0.37
ES	2023	0.38	0.47	0.43	RO	2023	0.11	0.58	0.34
ES	2024	0.39	0.42	0.41	RO	2024	0.11	0.60	0.35
FI	2017	0.70	0.65	0.68	SE	2017	0.57	0.75	0.66
FI	2018	0.68	0.67	0.67	SE	2018	0.57	0.73	0.65
FI	2019	0.69	0.65	0.67	SE	2019	0.58	0.72	0.65
FI	2020	0.73	0.66	0.70	SE	2020	0.66	0.75	0.70
FI	2021	0.75	0.68	0.71	SE	2021	0.68	0.78	0.73
FI	2022	0.81	0.70	0.76	SE	2022	0.72	0.77	0.74

Country Code	Year	IKP	IKC	IRPI	Country Code	Year	IKP	IKC	IRPI
FI	2023	0.82	0.69	0.76	SE	2023	0.72	0.74	0.73
FI	2024	0.75	0.70	0.73	SE	2024	0.76	0.78	0.77
FR	2017	0.57	0.71	0.64	SI	2017	0.39	0.51	0.45
FR	2018	0.57	0.71	0.64	SI	2018	0.38	0.52	0.45
FR	2019	0.57	0.69	0.63	SI	2019	0.38	0.52	0.45
FR	2020	0.57	0.70	0.64	SI	2020	0.42	0.53	0.48
FR	2021	0.57	0.69	0.63	SI	2021	0.43	0.56	0.49
FR	2022	0.57	0.70	0.64	SI	2022	0.46	0.58	0.52
FR	2023	0.59	0.69	0.64	SI	2023	0.47	0.59	0.53
FR	2024	0.59	0.68	0.64	SI	2024	0.49	0.61	0.55
HR	2017	0.32	0.27	0.29	SK	2017	0.19	0.60	0.40
HR	2018	0.30	0.27	0.29	SK	2018	0.18	0.60	0.39
HR	2019	0.30	0.27	0.28	SK	2019	0.19	0.63	0.41
HR	2020	0.40	0.29	0.34	SK	2020	0.21	0.64	0.43
HR	2021	0.44	0.28	0.36	SK	2021	0.21	0.65	0.43
HR	2022	0.48	0.32	0.40	SK	2022	0.23	0.67	0.45
HR	2023	0.53	0.24	0.38	SK	2023	0.23	0.65	0.44
HR	2024	0.53	0.28	0.41	SK	2024	0.25	0.66	0.46
HU	2017	0.20	0.71	0.45	EU-27	2017	0.45	0.58	0.52
HU	2018	0.20	0.69	0.45	EU-27	2018	0.46	0.57	0.52
HU	2019	0.21	0.69	0.45	EU-27	2019	0.46	0.58	0.55
HU	2020	0.26	0.71	0.49	EU-27	2020	0.51	0.58	0.55
HU	2021	0.27	0.72	0.49	EU-27	2021	0.50	0.60	0.55
HU	2022	0.30	0.72	0.51	EU-27	2022	0.52	0.63	0.57
HU	2023	0.29	0.71	0.50	EU-27	2023	0.52	0.61	0.56
HU	2024	0.27	0.75	0.51	EU-27	2024	0.51	0.61	0.56

ANNEX 3: IEI, KPEI AND KCEI SCORES FOR EU-27 COUNTRIES (2019-2024)

Year	Country code	KPEI	KCEI	IEI	Country code	KPEI	KCEI	IEI
2019	AT	1.00	0.69	0.69	IE	0.85	1.00	0.85
2020	AT	1.00	0.76	0.76	IE	0.88	1.00	0.88
2021	AT	0.94	0.78	0.73	IE	0.81	1.00	0.81
2022	AT	1.00	0.80	0.80	IE	0.82	1.00	0.82
2023	AT	1.00	0.73	0.73	IE	0.72	1.00	0.72
2024	AT	1.00	0.66	0.66	IE	0.80	1.00	0.80
2019	BE	0.94	0.78	0.73	IT	1.00	0.77	0.77
2020	BE	0.94	0.81	0.76	IT	1.00	0.78	0.78
2021	BE	0.98	0.80	0.79	IT	1.00	0.71	0.71
2022	BE	0.99	0.78	0.77	IT	1.00	0.79	0.79
2023	BE	0.97	0.74	0.71	IT	1.00	0.76	0.76
2024	BE	0.94	0.75	0.70	IT	1.00	0.68	0.68
2019	BG	1.00	1.00	1.00	LT	0.82	0.40	0.33
2020	BG	1.00	1.00	1.00	LT	0.93	0.42	0.39
2021	BG	1.00	1.00	1.00	LT	0.67	0.45	0.30
2022	BG	1.00	1.00	1.00	LT	0.65	0.48	0.31
2023	BG	1.00	1.00	1.00	LT	0.71	0.49	0.35
2024	BG	1.00	1.00	1.00	LT	0.72	0.48	0.34
2019	CY	1.00	1.00	1.00	LU	1.00	0.88	0.88
2020	CY	1.00	1.00	1.00	LU	1.00	0.93	0.93
2021	CY	1.00	1.00	1.00	LU	1.00	0.95	0.95
2022	CY	1.00	1.00	1.00	LU	1.00	0.88	0.88
2023	CY	1.00	0.99	0.99	LU	1.00	0.86	0.86
2024	CY	1.00	0.80	0.80	LU	1.00	0.86	0.86
2019	CZ	0.62	1.00	0.62	LV	1.00	0.70	0.70
2020	CZ	0.68	1.00	0.68	LV	1.00	1.00	1.00
2021	CZ	0.61	0.98	0.59	LV	1.00	1.00	1.00
2022	CZ	0.67	0.98	0.65	LV	1.00	1.00	1.00
2023	CZ	0.93	1.00	0.93	LV	1.00	1.00	1.00
2024	CZ	0.91	1.00	0.91	LV	1.00	0.72	0.72
2019	DE	0.93	1.00	0.93	MT	1.00	0.91	0.91
2020	DE	1.00	1.00	1.00	MT	1.00	0.90	0.90
2021	DE	1.00	1.00	1.00	MT	1.00	0.90	0.90
2022	DE	1.00	1.00	1.00	MT	1.00	1.00	1.00
2023	DE	1.00	0.98	0.98	MT	1.00	1.00	1.00
2024	DE	0.99	0.94	0.94	MT	1.00	1.00	1.00
2019	DK	0.86	0.85	0.73	NL	1.00	0.86	0.86
2020	DK	0.86	0.87	0.76	NL	1.00	0.87	0.87
2021	DK	0.92	0.91	0.83	NL	1.00	0.95	0.95
2022	DK	1.00	0.91	0.91	NL	1.00	0.90	0.90
2023	DK	1.00	0.83	0.83	NL	1.00	0.80	0.80
2024	DK	0.91	0.79	0.72	NL	1.00	0.80	0.80

Year	Country code	KPEI	KCEI	IEI	Country code	KPEI	KCEI	IEI
2019	EE	0.92	0.64	0.59	PL	1.00	0.80	0.80
2020	EE	1.00	0.64	0.64	PL	1.00	0.81	0.81
2021	EE	1.00	0.60	0.60	PL	1.00	1.00	1.00
2022	EE	1.00	0.60	0.60	PL	1.00	1.00	1.00
2023	EE	1.00	0.63	0.63	PL	0.66	1.00	0.66
2024	EE	1.00	0.66	0.66	PL	1.00	0.84	0.84
2019	EL	1.00	0.51	0.51	PT	1.00	0.50	0.50
2020	EL	1.00	0.70	0.70	PT	1.00	0.50	0.50
2021	EL	0.94	0.84	0.79	PT	0.69	0.58	0.40
2022	EL	1.00	1.00	1.00	PT	0.71	0.64	0.45
2023	EL	1.00	1.00	1.00	PT	0.61	0.67	0.40
2024	EL	1.00	0.54	0.54	PT	1.00	0.58	0.58
2019	ES	0.56	0.60	0.34	RO	1.00	1.00	1.00
2020	ES	0.58	0.58	0.34	RO	1.00	1.00	1.00
2021	ES	0.49	0.58	0.28	RO	1.00	1.00	1.00
2022	ES	0.58	0.74	0.43	RO	1.00	1.00	1.00
2023	ES	0.61	0.66	0.40	RO	1.00	1.00	1.00
2024	ES	0.59	0.58	0.34	RO	1.00	1.00	1.00
2019	FI	1.00	0.78	0.78	SE	0.82	0.89	0.73
2020	FI	1.00	0.79	0.79	SE	0.84	0.92	0.77
2021	FI	1.00	0.75	0.75	SE	0.88	0.95	0.84
2022	FI	1.00	0.75	0.75	SE	0.90	0.92	0.83
2023	FI	1.00	0.76	0.76	SE	0.89	0.81	0.71
2024	FI	1.00	0.76	0.76	SE	0.90	0.84	0.75
2019	FR	0.82	0.85	0.70	SI	0.62	1.00	0.62
2020	FR	0.83	0.86	0.71	SI	0.63	1.00	0.63
2021	FR	0.78	0.87	0.67	SI	0.58	1.00	0.58
2022	FR	0.80	0.86	0.69	SI	0.65	0.87	0.57
2023	FR	0.75	0.84	0.63	SI	0.70	0.91	0.64
2024	FR	0.78	0.81	0.63	SI	1.00	0.88	0.88
2019	HR	1.00	1.00	1.00	SK	1.00	1.00	1.00
2020	HR	1.00	1.00	1.00	SK	0.54	1.00	0.54
2021	HR	1.00	1.00	1.00	SK	0.48	1.00	0.48
2022	HR	1.00	1.00	1.00	SK	0.47	1.00	0.47
2023	HR	1.00	1.00	1.00	SK	0.48	1.00	0.48
2024	HR	1.00	1.00	1.00	SK	0.49	1.00	0.49
2019	HU	1.00	1.00	1.00				
2020	HU	1.00	1.00	1.00				
2021	HU	1.00	1.00	1.00				
2022	HU	0.90	1.00	0.90				
2023	HU	1.00	1.00	1.00				
2024	HU	1.00	1.00	1.00				

ANNEX 4: IE_IRPI SCORES AND RANKING FOR EU-27 COUNTRIES (2019-2024)

Country code	Year	IE_IRPI	IE_IRPI Rank	Country code	Year	IE_IRPI	IE_IRPI Rank
AT	2019	0.622	10	AT	2022	0.638	10
BE	2019	0.654	6	BE	2022	0.729	8
BG	2019	0.308	26	BG	2022	0.401	23
CY	2019	0.630	8	CY	2022	0.841	1
CZ	2019	0.498	14	CZ	2022	0.583	15
DE	2019	0.762	1	DE	2022	0.813	2
DK	2019	0.629	9	DK	2022	0.766	3
EE	2019	0.464	16	EE	2022	0.598	13
EL	2019	0.388	20	EL	2022	0.584	14
ES	2019	0.328	24	ES	2022	0.402	22
FI	2019	0.673	5	FI	2022	0.750	5
FR	2019	0.619	11	FR	2022	0.619	11
HR	2019	0.302	27	HR	2022	0.422	20
HU	2019	0.476	15	HU	2022	0.523	16
IE	2019	0.728	3	IE	2022	0.736	6
IT	2019	0.555	12	IT	2022	0.618	12
LT	2019	0.310	25	LT	2022	0.375	26
LU	2019	0.736	2	LU	2022	0.645	9
LV	2019	0.337	23	LV	2022	0.332	27
MT	2019	0.548	13	MT	2022	0.499	17
NL	2019	0.705	4	NL	2022	0.730	7
PL	2019	0.361	21	PL	2022	0.428	19
PT	2019	0.458	17	PT	2022	0.400	24
RO	2019	0.361	22	RO	2022	0.388	25
SE	2019	0.644	7	SE	2022	0.751	4
SI	2019	0.433	18	SI	2022	0.491	18
SK	2019	0.431	19	SK	2022	0.412	21
AT	2020	0.626	10	AT	2023	0.643	9
BE	2020	0.697	8	BE	2023	0.701	7
BG	2020	0.323	27	BG	2023	0.375	25
CY	2020	0.753	2	CY	2023	0.815	1
CZ	2020	0.523	15	CZ	2023	0.639	10
DE	2020	0.819	1	DE	2023	0.813	2
DK	2020	0.667	9	DK	2023	0.745	4
EE	2020	0.579	14	EE	2023	0.614	13
EL	2020	0.450	18	EL	2023	0.545	15
ES	2020	0.324	26	ES	2023	0.385	24
FI	2020	0.698	7	FI	2023	0.752	3
FR	2020	0.624	11	FR	2023	0.615	12
HR	2020	0.362	24	HR	2023	0.405	20
HU	2020	0.513	16	HU	2023	0.525	17

Country code	Year	IE_IRPI	IE_IRPI Rank	Country code	Year	IE_IRPI	IE_IRPI Rank
IE	2020	0.745	3	IE	2023	0.706	6
IT	2020	0.595	12	IT	2023	0.610	14
LT	2020	0.341	25	LT	2023	0.392	22
LU	2020	0.733	4	LU	2023	0.628	11
LV	2020	0.380	22	LV	2023	0.349	27
MT	2020	0.585	13	MT	2023	0.528	16
NL	2020	0.698	6	NL	2023	0.689	8
PL	2020	0.370	23	PL	2023	0.398	21
PT	2020	0.418	19	PT	2023	0.390	23
RO	2020	0.387	21	RO	2023	0.362	26
SE	2020	0.703	5	SE	2023	0.719	5
SI	2020	0.457	17	SI	2023	0.511	18
SK	2020	0.401	20	SK	2023	0.410	19
AT	2021	0.617	10	AT	2024	0.611	13
BE	2021	0.710	7	BE	2024	0.716	6
BG	2021	0.338	26	BG	2024	0.385	24
CY	2021	0.791	2	CY	2024	0.726	5
CZ	2021	0.502	16	CZ	2024	0.614	10
DE	2021	0.808	1	DE	2024	0.762	2
DK	2021	0.707	9	DK	2024	0.695	8
EE	2021	0.590	12	EE	2024	0.613	12
EL	2021	0.478	17	EL	2024	0.469	18
ES	2021	0.328	27	ES	2024	0.364	26
FI	2021	0.709	8	FI	2024	0.728	4
FR	2021	0.612	11	FR	2024	0.616	9
HR	2021	0.381	23	HR	2024	0.430	20
HU	2021	0.521	15	HU	2024	0.542	16
IE	2021	0.749	3	IE	2024	0.737	3
IT	2021	0.587	13	IT	2024	0.602	14
LT	2021	0.346	25	LT	2024	0.389	23
LU	2021	0.731	6	LU	2024	0.613	11
LV	2021	0.348	24	LV	2024	0.349	27
MT	2021	0.541	14	MT	2024	0.498	17
NL	2021	0.732	5	NL	2024	0.705	7
PL	2021	0.395	22	PL	2024	0.432	19
PT	2021	0.407	19	PT	2024	0.409	22
RO	2021	0.395	21	RO	2024	0.373	25
SE	2021	0.737	4	SE	2024	0.767	1
SI	2021	0.468	18	SI	2024	0.571	15
SK	2021	0.397	20	SK	2024	0.428	21