Universidade de Brasília
Faculdade de Administração, Economia, Contabilidade e Gestão Pública

Unpacking Economic Complexity and Relatedness for Industrial Policy: Filtering Noise, Mapping Traps, and Framing Developmental Pathways

Renan Abrantes de Sousa

TESE DE DOUTORADO
ECONOMIA

Brasília
2025

Universidade de Brasília
Faculdade de Administração, Economia, Contabilidade e Gestão Pública

Desvelando Complexidade Econômica e Conexões Produtivas para a Política Industrial: Filtrando Ruídos, Mapeando Armadilhas e Estruturando Caminhos de Desenvolvimento

Renan Abrantes de Sousa

Tese de doutorado submetida como requisito parcial para obtenção do grau de doutor em Economia.

Orientador: Prof. Dr. Bernardo Mueller

Brasília
2025

FICHA CATALOGRÁFICA

Abrantes de Sousa, Renan.

Unpacking Economic Complexity and Relatedness for Industrial Policy: Filtering Noise, Mapping Traps, and Framing Developmental Pathways / Renan Abrantes de Sousa; orientador Bernardo Mueller. -- Brasília, 2025.

137 p.

Tese de Doutorado (Economia) -- Universidade de Brasília, 2025.

1. Economic complexity. 2. Relatedness. 3. Industrial policy. 4. Diversification. 5. Product space. I. Mueller, Bernardo, orient. II. Título.

Universidade de Brasília
Faculdade de Administração, Economia, Contabilidade e Gestão Pública

Unpacking Economic Complexity and Relatedness for Industrial Policy: Filtering Noise, Mapping Traps, and Framing Developmental Pathways

Renan Abrantes de Sousa

Tese de doutorado submetida como requisito parcial para obtenção do grau de doutor em Economia.

Trabalho aprovado. Brasília, 21 de agosto de 2025:

Prof. Dr. Bernardo Mueller,
UnB/FACE/ECO
Orientador

Prof. Dr. Daniel Oliveira Cajueiro,
UnB/FACE/ECO
Examinador interno

Prof. Dr. João Prates Romero,
UFMG/FACE/CEDEPLAR
Examinador externo

Prof. Dr. Pedro Fernando Nery,
IDP
Examinador externo

À Estela.

Agradecimentos

Ao fim desta jornada, quero expressar minha profunda gratidão a todas as pessoas que, de diferentes formas, tornaram este doutorado possível — e mais leve.

À Natália, minha companheira de vida, por ter estado ao meu lado em cada etapa deste percurso. Obrigado por dividir comigo o dia a dia e também os meandros invisíveis que sustentam, suavizam e dão sentido a uma caminhada longa como a do doutorado.

Ao meu orientador, Bernardo Mueller, que mais do que um orientador, foi e é meu mentor intelectual. Aprendi com ele a escutar ideias novas, a acolher o que é disruptivo e a cultivar o pensamento com abertura e rigor. Sua generosidade intelectual e sua confiança foram fundamentais para que este trabalho pudesse tomar forma.

Agradeço também a César Hidalgo, que me recebeu em Toulouse para o período de doutorado-sanduíche e me aproximou da fronteira do conhecimento na literatura de complexidade econômica.

Aos amigos que fiz ao longo do caminho, na UnB e no *Center for Collective Learning*, meu sincero agradecimento pelo companheirismo, pelas trocas e pelo apoio mútuo. Em especial: Cristina, Deise, Débora, Eva, João Victor, Lorena, Manuel, Mariana, Ole, Pedro, Raphael e Sérgio, além de tantos outros que, embora não mencionados aqui, deixaram marcas importantes nesta trajetória.

Cada avanço deste trabalho carrega também o gesto, o tempo e a generosidade de quem esteve por perto.

Financiamento

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

“Creía en infinitas series de tiempos, en uma red creciente y vertiginosa de tiempos divergentes, convergentes y paralelos. Esa trama de tiempos que se aproximan, se bifurcan, se cortan o que secularmente se ignoran, abarca todas la posibilidades. No existimos en la mayoría de esos tiempos; en algunos existe usted y no yo; en otros, yo, no usted; en otros, los dos. En éste, que un favorable azar me depara, usted ha llegado a mi casa; en otro, usted, al atravesar el jardín, me ha encontrado muerto; en otro, yo digo estas mismas palabras, pero soy un error, un fantasma.” (El jardín de senderos que se bifurcan, Jorge Luis Borges, 1941)

Abstract

Industrial policy has re-emerged as a central pillar of development strategy, renewing long-standing debates about its design, implementation, and effectiveness. This dissertation contributes to that debate by advancing the Economic Complexity and Relatedness framework to better inform structural transformation and policy targeting. The first paper, *Complexity Traps in the Product Space: Why Some Countries Get Stuck in Local Maxima*, investigates how economies can become structurally constrained when certain products offer returns that are disproportionately high relative to their neighbors in the product space. These localized incentives make it more attractive to exploit products that function as local maxima—but with low complexity—than to explore nearby opportunities that would promote capability accumulation. While appealing in the short term, this pattern ultimately constrains diversification, limiting the range of activities a country can develop over time. Over the long run, such constraints give rise to persistent regions of structural stasis - complexity traps - where countries struggle to transition into more complex activities. These findings highlight the need for industrial strategies that deliberately counteract the distorting effects of local optima and promote broader diversification. The second paper, *Less is More: How Relatedness Filtering Enhances Productive Upgrading Predictions*, demonstrates that statistical noise in the product space can hinder accurate identification of viable diversification paths. By filtering weak and spurious connections, we significantly improve the ability to predict future productive transitions—especially for less diversified economies—offering a more precise empirical foundation for strategic industrial policy. The third paper, *From Capabilities to Economic Convergence: A Structural Growth Framework Linking Economic Complexity, Institutions, and Human Capital*, proposes an integrated model that explains both current income levels and future growth using a multidimensional view of capabilities. It introduces a new complexity measure based on input–output data, which captures the sophistication of production networks beyond trade flows. The results show that multidimensional complexity, institutional quality, and human capital jointly shape development trajectories, and that countries with unexpectedly high complexity relative to their income tend to grow faster. Together, the three studies offer both diagnostic and prescriptive contributions to the Economic Complexity literature, helping to identify structural bottlenecks, improve targeting of policy tools, and reframe long-run development strategies.

Keywords: economic complexity; relatedness; industrial policy; diversification; product space.

Resumo

A política industrial voltou a ocupar um papel central nas estratégias de desenvolvimento, reacendendo debates clássicos sobre seu desenho, implementação e efetividade. Esta tese contribui para essa discussão ao aprimorar o arcabouço de Complexidade Econômica e *Relatedness*, com o objetivo de oferecer melhores instrumentos para orientar a transformação estrutural e o direcionamento de políticas públicas. O primeiro artigo, *Complexity Traps in the Product Space: Why Some Countries Get Stuck in Local Maxima*, investiga como certas estruturas produtivas podem restringir o avanço econômico quando produtos específicos oferecem retornos desproporcionalmente altos em relação aos seus vizinhos no espaço-produto. Esses incentivos localizados tornam mais atraente explorar produtos que funcionam como máximos locais, contudo de baixa complexidade, em detrimento de alternativas próximas que poderiam favorecer o acúmulo de capacidades. Embora vantajosa no curto prazo, essa lógica acaba por prejudicar a diversificação e limita o leque de atividades que um país é capaz de desenvolver ao longo do tempo. No longo prazo, esse tipo de configuração dá origem a armadilhas de complexidade, em que países permanecem presos em trajetórias de baixo dinamismo estrutural. Os resultados reforçam a necessidade de estratégias industriais que enfrentem essas distorções e ampliem o horizonte de diversificação produtiva. O segundo artigo, *Less is More: How Relatedness Filtering Enhances Productive Upgrading Predictions*, mostra que o excesso de ruído no espaço-produto pode obscurecer caminhos relevantes de diversificação. Ao aplicar técnicas de filtragem para remover conexões espúrias ou pouco informativas, o estudo melhora significativamente a capacidade de prever quais atividades produtivas um país tende a desenvolver — sobretudo em economias menos diversificadas —, oferecendo uma base empírica mais robusta para o desenho de políticas industriais mais precisas. O terceiro artigo, *From Capabilities to Economic Convergence: A Structural Growth Framework Linking Economic Complexity, Institutions, and Human Capital*, propõe um modelo integrado que busca explicar tanto os níveis atuais de renda quanto o crescimento de longo prazo a partir de uma abordagem multidimensional das capacidades. O trabalho introduz uma nova medida de complexidade baseada em dados de insumo-produto, que permite capturar a sofisticação das redes produtivas internas, além das exportações. Os resultados indicam que a complexidade produtiva, o capital humano e a qualidade institucional interagem de forma decisiva na definição das trajetórias de desenvolvimento — e que países com níveis de complexidade acima do esperado, dado seu nível de renda, tendem a apresentar maior dinamismo econômico ao longo do tempo. Em conjunto, os três estudos oferecem contribuições analíticas e aplicadas à literatura de Complexidade Econômica, ao mesmo tempo em que propõem ferramentas úteis para diagnosticar gargalos estruturais, qualificar o uso de instrumentos de política industrial e repensar estratégias de desenvolvimento no longo prazo.

Palavras-chave: complexidade econômica; relatedness; política industrial; diversificação; espaço-produto.

List of figures

List of tables

Contents

1 Introduction

Industrial policy has experienced a resurgence in recent years. Previously marginalized in economic discourse, its revival has ignited intense debate among academics and policymakers alike. Moving away from Gary Becker's 1985 assertion that *"the best industrial policy is none at all"*[1], the global debate has shifted in favor of such policies. As Ricardo Hausmann noted in 2023, *"we now know that the real question is not whether such policies should exist, but how to manage them"*[2].

The ongoing debate around industrial policy doesn't center on its underlying rationale, but rather on two practical objections (Juhász; Lane; Rodrik, 2023). The first relates to informational limitations, and the second to political capture. The informational critique argues that even if market failures exist, real-world governments may lack sufficient knowledge to accurately identify these failures. The political critique contends that even with access to relevant information, industrial policy risks encouraging self-interested lobbying and political influence, steering government efforts toward private gains rather than public benefit.

An emerging body of literature on Economic Complexity and Relatedness offers promising solutions to the informational challenges of industrial policy. By applying network theory and machine learning techniques, this field helps identify which sectors to prioritize along two key dimensions: those that can introduce greater embedded knowledge — complexity — into an economy, and those sectors closely related to existing productive capabilities. This dual approach not only aims to make industrial policy more targeted and less risky but also provides a guiding framework for skeptics by offering a *via negativa*, revealing paths that national governments should avoid.

As multilateralism wanes, advanced economies have taken the lead in the resurgence of industrial policy (Evenett *et al.*, 2024). Yet, the need for effective industrial policy design is more pressing in emerging and developing countries, where fiscal space is often limited and the opportunity costs, relative to other government initiatives, are high. In Brazil, for example, past challenges with industrial policy have left the current debate deeply polarized.

To address these challenges and advance the field, this thesis is organized around three interconnected essays, each tackling a distinct but related aspect of economic complexity and industrial policy design. This dissertation advances this framework by exploring its boundaries and potential refinements across three dimensions: (i) how pecuniary incentives

1 https://www.nytimes.com/1985/09/15/business/l-industrial-policy-004530.html

2 https://www.project-syndicate.org/commentary/why-economists-have-rediscovered-industrial-policy-by-ricardo-hausmann-2023-01

over the product space can hinder diversification and induce complexity traps, (ii) how network filtering can improve the predictive power of relatedness, and (iii) how a multidimensional view of capabilities — combining trade and input–output linkages to capture multidimensional economic complexity, and embedding this structure with institutions and human capital — can better account for differences in income levels and long-term growth dynamics.

In the first essay, I investigate why some economies remain trapped in narrow productive patterns, consistently failing to diversify into new products. Adopting an exploration-exploitation perspective within the Product Space framework, I introduce the Product Peak Index (PPI) to identify products representing local maxima—products whose pecuniary returns significantly exceed those of their immediate neighbors. Aggregating these product-level measures at the country level through the Country Peak Index (CPI), I show that economies overly focused on exploiting such high-peak yet low-complexity products substantially diminish their incentives for exploring more complex and capability-enhancing economic activities. Consequently, this persistent exploitation constrains long-term capability accumulation and limits opportunities for structural upgrading and economic diversification.

Empirical analysis of the proposed mechanism reveals distinct structural traps within a CPI-Diversity phase space, including regions of resource-based stasis and middle-development inertia. Findings also show a hump-shaped relationship between CPI and the likelihood of diversifying into new products, emphasizing diminishing returns to diversification efforts beyond a certain threshold of CPI. This highlights important policy implications, cautioning against premature specialization in high-return, low-complexity products. Strategic industrial policies, therefore, should focus on carefully timing diversification strategies to avoid these structural traps, promoting targeted interventions that encourage the exploration of complex, capability-enhancing economic activities.

The second essay tackles critical shortcomings inherent in traditional relatedness measures commonly applied to predict productive diversification in the Product Space. Recognizing that standard metrics often incorporate spurious correlations and informational noise, I propose integrating network backbone extraction methods to filter these weak or irrelevant product relationships. By systematically applying various filtering techniques—including naive heuristics, statistical backbone extraction, and directional adaptations—I show that removing noisy connections significantly enhances predictive accuracy. Among these methods, the Directed Disparity Filter In-Degree (DDF-In-Degree) proves most effective, clearly outperforming the traditional, unfiltered relatedness density.

These findings hold substantial implications for industrial policymaking. Filtering the Product Space refines the precision of relatedness indicators, illuminating clearer and more realistic pathways for productive upgrading. Through detailed case studies of Brazil and the Philippines, I illustrate how filtered relatedness can reshape strategic industrial policy, enabling policymakers to more effectively identify viable diversification targets. Consequently, this approach minimizes the risk of misallocating resources, providing a robust foundation for policies that support sustainable economic upgrading and structural transformation.

The third essay proposes a structural framework to jointly analyze the role of three key capability vectors—economic complexity, institutional quality, and human capital—in shaping long-term economic convergence. Central to this framework is the introduction of a new measure of complexity, the Input-Output Economic Complexity Index (IO ECI), which captures intersectoral linkages beyond traditional trade-based metrics. Together with the established Trade ECI, IO ECI constitutes a multidimensional view of economic complexity. To empirically assess these key development vectors, I develop a two-stage structural growth framework. In the first stage, cross-sectional regressions isolate the portion of per capita income not explained by multidimensional economic complexity, institutional quality, or human capital, capturing structural misalignments for each vector. In the second stage, panel regressions evaluate how these capability-driven residuals, combined with subsequent changes in complexity, institutions, and human capital, predict long-term economic growth.

This structural framework redefines industrial policy as orchestrating multidimensional capability systems built upon three pillars of economic convergence: productive complexity, institutional robustness, and human capital accumulation. Empirical findings highlight the critical importance of integrating industrial policies with complementary measures aimed at strengthening institutional quality and enhancing human capital to promote sustained long-term growth. Moreover, the framework provides policymakers with a diagnostic tool, allowing them to identify which structural vector currently represents the largest constraint—whether productive capabilities, institutional foundations, or human capital—and thus prioritize policy interventions accordingly. By leveraging this targeted approach, countries can systematically address their specific developmental weaknesses and foster more effective and accelerated economic convergence.

Taken together, the three essays offer both theoretical and practical contributions to the Economic Complexity and Relatedness literature. They help identify structural bottlenecks, refine empirical tools for targeting, and broaden the conceptual foundation for industrial policy in diverse development contexts.

2 Complexity Traps in the Product Space: Why Some Countries Get Stuck in Local Maxima

Abstract

The principle of relatedness has become a dominant framework for explaining productive diversification. Yet, some economies remain locked into narrow sets of products, failing to explore adjacent opportunities—a paradox we seek to address. We hypothesize that when certain products offer substantially higher pecuniary returns than their neighbors in the product space, countries may prioritize their exploitation over exploration, ultimately getting stuck in local maxima and limiting long-term diversification. To capture this mechanism, we introduce the Product Peak Index (PPI), which identifies local maxima based on relative income returns, and the Country Peak Index (CPI), which measures a country's exposure to these peaks. We show that specialization in such products is associated with reduced diversification and the emergence of development traps—including resource-based stasis and middle-development inertia—identified through a dynamic phase-space approach. At the micro level, Probit models reveal a hump-shaped effect of CPI on the probability of product activation, indicating diminishing returns to exploration beyond a certain threshold. These findings underscore the importance of pecuniary incentives in shaping development paths and offer actionable insights for industrial policy—particularly regarding the timing and targeting of diversification strategies in economies vulnerable to premature, sub-optimal peak specialization.

Keywords: economic complexity, diversification, product space, economic convergence, development traps

2.1 Introduction

When it comes to diversification, resource-rich economies such as those of Chile, Angola, and Australia, exhibit an interesting pattern: while these economies tend to have relatively high levels of GDP per capita for their levels of economic complexity, they tend to diversify less into other economic activities. This lack of diversification is important because it can limit the accumulation of the capabilities that economies need to stay adaptive in a changing technological environment. Diversification is often explained using the principle of relatedness: the notion that economies tend to diversify into related products. Yet, the reduced diversification of some resource rich countries cannot be explained by this principle alone, if we mechanistically assume that spillovers are constant at the same level of relatedness.

Here we explain this paradox by expanding the principle of relatedness by introducing a product peak index, capturing the idea that the pecuniary return of some products is much larger than that of its related products. We hypothesize that the high returns associated with specializing in peak products reduces the incentive of these economies to explore new activities, contributing to less diversification. We test this hypothesis by showing that countries specialized in peak products tend to diversify less than countries specialized in non-peak products, holding for the same level of relatedness.

This reduced tendency to diversify doesn't necessarily imply an income trap, since countries specialized in high peak index products can achieve high income levels through the intensive exploitation of these activities (e.g. Saudi Arabia, Australia). Yet, when peak products are not enough to reach high-income, economies specialized in such activities may lack the capabilities and incentives needed to bridge into the related and lower income activities needed to reach higher complexity sectors of the product space. These findings contribute to our understanding of diversification and of the role of product level characteristics to the resource curse (Ploeg, 2011) by adding a mechanism that reduces the diversification incentives of poorly diversified, low-complexity, and resource-rich economies.

To investigate this mechanism, we proceed in four steps. First, we construct a novel representation of the product space — the Product-PRODY-Space — to visualize pecuniary diversification incentives and identify local maxima in the product space layout. Next, we structurally derive the Product Peak Index (PPI), a product-level metric that captures how much a product outperforms its neighbors in terms of export-weighted income returns. Second, we aggregate the PPI at the country level to create the Country Peak Index (CPI), and explore how countries' positions in the CPI–Diversity phase space shape their structural development trajectories. Using a dynamic vector field approach, we uncover empirical attractors, or structural traps, that constrain complexity accumulation. Third, we examine the predictive power of the CPI at the country–product level by estimating Probit models of product activation, controlling for relatedness and relative relatedness. We test for nonlinearities and find a hump-shaped effect of CPI on diversification probabilities, consistent with the structural traps observed in the macro analysis. Finally, we discuss the policy implications of these findings, emphasizing the risks of premature specialization and the importance of carefully timing the transition from exploration to exploitation in industrial strategies.

2.2 Literature Review

The study of development traps and convergence are fundamentally interconnected, each representing the flip side of the other's coin. Poverty traps and middle-income traps have been thoroughly examined to pinpoint the primary obstacles preventing countries from ascending the income ladder (Gill; Kharas, 2015; Im; Rosenblatt, 2015; Kraay; McKenzie, 2014; Barrett; Carter, 2013; Rodrik, 2011; Lin, 2011). The Economic Complexity literature contributed to this debate by finding the existence of income convergence conditional to the level of the complexity of a country (Hausmann *et al.*, 2014). Nevertheless, the debate about complexity convergence – instead of income convergence – and complexity traps remains an active and evolving area of debate.

In contrast to earlier methods that focus on pinpointing specific factors for economic convergence, economic complexity has adopted a more open-ended stance. Rather than assuming the nature of the contributing factors a priori, this theory aims to gauge their joint impact Hidalgo (2021). Its main assumption rests on the ability of products, activities, industries, among others, to serve as vehicles for the transmission and accumulation of productive knowledge.

For Hausmann *et al.* (2014), increases in collective knowledge result in an expansion of the range of activities that a country can undertake, venturing into the adjacent possible – new, attainable activities that are related to existing capabilities. The potential for such diversification is, in turn, shaped by the underlying capabilities in the country's current productive structure. Such dynamics create a feedback loop between the productive structure and the complexity level of a country, dictating what is produced in the future and leading to a cycle that can be either virtuous or vicious in terms of economic development. Therefore, explicit path dependency and exploration of the adjacent possible are key aspects of the Economic Complexity methods.

Adjacent possible is an evolutionary concept coined by Kauffman (2000) to explain how complex adaptative systems are shaped by the exploration and the never-ending updates of the set of possibilities available to agents in these systems, with highly path-dependency features. Recent studies used the principle of relatedness to uncover the adjacent possible and model the structural outcomes of geographical entities inscribed in various complex economic systems (Hidalgo *et al.*, 2007; Neffke; Henning; Boschma, 2011; Muneepeerakul *et al.*, 2013; Boschma; Balland; Kogler, 2015; Kogler; Rigby; Tucker, 2015). The principle of relatedness (Hidalgo *et al.*, 2018) states that when two activities share similar productive requisites, they are likely to co-occur. In this sense, a geographical entity (country, region, city) raises its probability of performing a new activity in tandem with the number of related activities that it already makes. Fundamentally, relatedness revolves around productive complementarities, framing economic convergence within the Economic Complexity literature as a challenge of diffusion through the network of relatedness between products, referred to as the product space.

Building on international trade data, Hidalgo *et al.* (2007) explained country economic development as a process of productive upgrading within the adjacent possible of exports.

The interaction between a country's productive structure and its complexity level exacerbates a "Matthew effect" in economic development, whereby more developed countries are better suited than less developed ones to diversify into new products and accumulate more capabilities (Sousa; Mueller, 2025). Countries with low levels of capabilities may lack the necessary productive capabilities to diversify into new products and break free from path dependency. Moreover, acquiring an additional productive capability yields greater returns for highly diversified nations than for those with limited diversification, due to the exponentially increasing number of possible capability combinations.

Economic divergence emerges as the rule rather than the exception when knowledge is difficult to transmit and acquire. This dynamic reveals the risk of a quiescence trap (Hausmann; Hidalgo, 2011), in which countries with limited capabilities find themselves locked into a state of low developmental inertia. It also aligns with the possibility of multiple equilibrium states, echoing Baumol's concept of convergence clubs (Baumol, 1986; Ben-David, 1998), where countries tend to cluster into groups with similar levels of economic development.

To identify such convergence clubs or groups of countries prone to developmental stasis, Quah (1992) and Kremer, Onatski, and Stock (2001) employed transition matrices based on Markov chains to describe the dynamics of income per capita convergence across countries. Similarly, Im and Rosenblatt (2015) applied this approach to investigate middle-income traps. In this context, identifying both macro and micro determinants that lead countries not only into income traps but also into *economic complexity traps* becomes crucial, given that a country's productive structure sophistication is closely linked to its GDP per capita.

Despite its importance, the literature on complexity traps remains limited. Building on the research surrounding middle-income traps, Hartmann *et al.* (2021) introduce the notion of middle-development traps, wherein countries with medium levels of complexity face significant challenges in further advancing their economic sophistication to reach developed status. Their study emphasizes that the transition toward a highly complex productive structure coevolves with increasing relatedness between a country's current capabilities and the complexity level of its potential products. This relationship is captured by the so-called S-curve of economic sophistication, which illustrates a pivotal nonlinear shift at the critical stages of structural transformation toward high complexity.

Balland and Boschma (2024) introduce an evolutionary perspective on development traps, emphasizing the structural inability of European regions to diversify into more complex economic activities. The authors propose a typology of economic traps that captures how regions can be stuck, not just because they are poor, but because their productive structure offers limited opportunities for upgrading. By shifting attention to these structural constraints, the paper offers a dynamic framework to understand why some regions remain stagnant while others continue to evolve.

One way of analyzing development trajectories and how to reach full developmental status is by documenting the relationship between country diversification and income (Imbs; Wacziarg, 2003; Al-Marhubi, 2000; Herzer; D, 2006; Hesse *et al.*, 2009). As income increases, countries tend to diversify their productive structures. At later stages of development, however, specialization re-emerges, leading to an inverted-U relationship commonly referred to as the "diversification hump." The timing of the hump is, thus, important for a country to escape any kind of developmental stasis. Within the complexity framework, Dam and Frenken (2022) introduce a combinatorial model showing that this hump reflects a constraint on the range of product complexities compatible with a country's capabilities. In their model, late-stage specialization occurs through product exits, specifically, the abandonment of less complex products.

Therefore, further deepening the empirical understanding of how countries journey through the diversification hump and their concurrent economic complexity index can aid the identification of complexity convergence clubs. Although diversification is mathematically orthogonal to the ECI (Mealy; Farmer; Teytelboym, 2019), its role in productive upgrading becomes evident when economic systems exhibit nestedness, a structural pattern in which complex and rare products appear only in highly diversified countries (Hidalgo, 2021). This mirrors patterns found in ecological systems, where rare species tend to inhabit more diverse environments. Nestedness implies that complex capabilities are harder to diffuse and remain structurally concentrated in a few locations (Lee, 2016; Bustos *et al.*, 2012), reinforcing the idea that development trajectories are shaped by the joint distribution of diversity and complexity.

Economic Complexity methods gave researchers powerful tools to understand the intricacies of development paths, but is there more to uncover about complexity paths and complexity traps? Is it possible to further understand which productive structures are more akin to become stuck in the complexity journey? In pursuit of answers, we draw upon a statement from (Hidalgo *et al.*, 2007, Supplementary Material, p. 13): "If the structural transformation only moves countries to more sophisticated goods, a local maximum would trap countries."

Let us recall the metaphor of the product space as a forest, as in Hausmann *et al.* (2014). Each tree of the forest represents a different product, with those requiring similar skills clustered together. Firms are monkeys that live in specific trees, exploiting those products. The productive structure variation among countries is captured by the number and distribution of these monkeys within the forest. The journey of economic development, aiming for greater product diversity and complexity, is akin to monkeys expanding their territory across the forest. This expansion is more straightforward when monkeys can easily jump to adjacent trees, which implies a gradual acquisition of new productive capabilities without needing to bridge large gaps. However, when the distance between trees increases, firms may struggle to diversify or innovate, potentially limiting their economic development by confining them to their existing range of activities.

Alongside the influence of increasing capability distances on the diminishing likelihood of transitioning between products, we delve into how a monkey, indulging in a tree laden with exceptionally juicy fruits, might grow overly dependent on this singular bounty, given its superior rewards compared to the offerings of adjacent trees. Such a dependency may deter the monkey from exploring further afield, particularly when the immediate alternatives are less enticing, and only more distant options offer superior benefits. This situation highlights how certain configurations of productive structures can result in the overexploitation of suboptimal local maxima, thereby hindering broader exploration and diversification within the product space. When the incentives of the extensive margin are prematurely outweighed by the benefits of the intensive margin in the economic development journey of a country, traps arise.

This idea is intimately linked to the exploitation versus exploration dilemma in adaptive processes (Schumpeter; Swedberg, 2021; March, 1991; Berger-Tal *et al.*, 2014). In the context of finite and constraining resources, systems must continuously navigate the trade-off between exploiting familiar, reliable opportunities and exploring unfamiliar, potentially superior ones. Exploitation supports efficiency, stability, and incremental gains, but over-reliance on it can entrench systems in routines that become increasingly misaligned with evolving environments. Exploration, on the other hand, enables adaptation, discovery, and long-term resilience, though it often entails risk, uncertainty, and short-term costs. The most adaptive systems are those that sustain a dynamic balance—leveraging the known while probing the unknown. Within the product space metaphor, development traps emerge when countries—like monkeys overly fixated on a single fruitful tree—fail to invest in exploring new branches of the forest. In such cases, the pursuit of immediate rewards undermines the broader journey toward diversification and structural transformation.

There is also a crossover between this idea and the resource curse literature. The term resource curse, first coined by Auty (2002), refers to the phenomenon in which natural resource dependence hampers economic growth. Several empirical studies have confirmed this relationship (Sachs; Warner, 1995; Gylfason, 2001; Mehlum; Moene; Torvik, 2006), while others have challenged it (Lederman, 2007; Cavalcanti; Mohaddes; Raissi, 2011; James, 2015). Economic dependence on natural resources can represent a major source of developmental stagnation. Badeeb, Lean, and Clark (2017) outlines the main channels through which this dependence may hinder growth: the Dutch Disease, commodity price volatility, economic mismanagement, rent-seeking, and corruption and deteriorating institutional quality.

A particularly compelling mechanism through which productive structure dynamics operate is the Dutch Disease, which manifests through two main channels: the spending effect and the resource pull effect. The spending effect arises when a natural resource boom increases national income and, consequently, domestic demand. This surge in demand leads to inflationary pressures and an appreciation of the real exchange rate, making non-resource tradable goods relatively more expensive on the international market. As a result, these sectors lose competitiveness and become less attractive to investment. The resource pull effect, in turn, refers to the reallocation of domestic inputs—such as labor and raw materials—toward the booming resource sector. This reallocation drives up the domestic cost of these inputs, raising production costs for other export-oriented industries, particularly manufacturing and agriculture, thereby contracting them. Connecting back to the core argument of this study, both effects distort the balance between exploration of new productive opportunities and continued exploitation of existing ones in the product space. In this context, products prone to overexploitation and low in complexity are likely to overlap with natural resources and their derivatives, reinforcing structural dependence and limiting diversification.

This perspective invites a closer examination of the product space—not only in terms of capability distances, but also in terms of relative returns and structural inertia. Certain products may offer outsized pecuniary gains compared to their neighbors, forming local maxima that anchor countries into narrow specialization. When the benefits of intensive exploitation outweigh the incentives for extensive diversification, countries may fall into a dual trap: one of quiescence and another of local optimality.

To better understand how such structural traps manifest, we propose two empirical contributions. First, we develop the Product Peak Index (PPI), which captures the extent to which a product represents a local maximum in terms of complexity-adjusted returns relative to its neighbors in the product space. Second, we aggregate this information at the country level to construct the Country Peak Index (CPI), which reflects how concentrated a country's productive structure is around such local peaks. These indices allow us to identify when and where productive configurations are likely to inhibit further upgrading—despite the presence of higher-potential paths within reach. In the following sections, we present the methodology used to construct these measures and empirically examine their relationship with observed development trajectories.

2.3 The Product-PRODY-Space and the Product Peak Index

The product space is a network model that can be used to map an economy's productive frontier or adjacent possible. Since the focus of this work is on understanding country development paths and traps, we make use of international trade data to build the product space, just as in Hidalgo *et al.* (2007). We extracted the HS92 4-digit dataset for period

1995-2020 from the Observatory of Economic Complexity – OEC[1].

The international trade dataset has been cleaned to reduce statistical noise originating from poor statistical quality reporting, variations in the size of the economies, and export breaks caused by war or highly politically unstable situations. The filters include discarding all countries that in any year of the period had a population smaller than 0.016% of the world population, a total yearly trade below 0.0067% of the world trade, and that scored equal or higher than 26 in the Fragile States Index[2] when summing up the dimensions of "Security Apparatus", "Refugees and Internally Displaced Persons", and "External Intervention". We also applied the baseline of OEC's product space to account for the set of products used in this work. After these steps, the HS92 dataset captures the trade of 866 products between 119 countries, which represented 96.9% of global GDP and 94.6% of global trade in 2010.

We follow Balassa's concept (Balassa, 1965) of Revealed Comparative Advantage (RCA) to consider if a country is a competitive exporter of a product or if it is not. It compares the export share of country c for a product p with the world's export share of that product:

$$\mathrm{RCA}_{cp} = \frac{\frac{X_{cp}}{\sum_p X_{cp}}}{\frac{\sum_c X_{cp}}{\sum_c \sum_p X_{cp}}} \tag{2.1}$$

If RCA is equal to or higher than unity, the country is a competitive exporter of that product, and thus has the required capabilities to produce it. The following binary matrix summarizes which country makes what and is the starting point to calculate both the product space and the measures of economic complexity:

$$M_{cp} = \begin{cases} 1, & \text{se } \mathrm{RCA}_{cp} \geq 1; \\ 0, & \text{otherwise.} \end{cases} \tag{2.2}$$

The Economic Complexity Index (ECI) and the Product Complexity Index (PCI) are measures of the productive knowledge embedded in geographical entities and economic activities, respectively. Defined as iterative averages, ECI and PCI measures behave as mirrors of each other. The ECI is the average of the PCI of the products that a country exports, and vice versa. These measures relate deeply to the economic development journey, since ECI explains cross-country differences in GDP per capita and predicts long-term economic growth (Hidalgo; Hausmann, 2009; Hausmann *et al.*, 2014; Stojkoski; Koch; Hidalgo, 2023).

1 http://oec.world

2 http://fragilestatesindex.org/

$$\mathrm{ECI}_c = \frac{1}{k_c} \sum_p M_{cp}\, \mathrm{PCI}_p, \tag{2.3}$$

$$\mathrm{PCI}_p = \frac{1}{k_p} \sum_c M_{cp}\, \mathrm{ECI}_c. \tag{2.4}$$

Replacing the PCI definition in the ECI one, and vice-versa, the measures can be computed by solving the following eigenvalue equations (Hidalgo, 2021):

$$\mathrm{ECI}_c = \sum_p \frac{M_{cp}}{k_p k_c} \sum_c M_{cp}\, \mathrm{ECI}_c, \tag{2.5}$$

$$\mathrm{PCI}_p = \sum_c \frac{M_{cp}}{k_p k_c} \sum_p M_{cp}\, \mathrm{PCI}_p. \tag{2.6}$$

The product space $\phi_{p,p'}$ is a matrix that measures the relative similarity between two products. It uses an adjacency matrix that counts how many times each pair of products is co-exported and divides this co-occurrence by the number of countries that export each of these two products, picking the smallest resulting number. The result is the minimum conditional probability of having RCA in each pair of products:

$$\phi_{p,p'} = \frac{\sum_c M_{cp} M_{cp'}}{\max\left(\sum_c M_{cp}, \sum_c M_{cp'}\right)} \tag{2.7}$$

The value of pairwise relative similarity is called proximity and goes from nil to unity. The greater the value of proximity, the more that pair of products share capabilities. Proximity is the feature that originally captures if firms occupying one product are more likely to jump into another product. The product space is a flattened, two-dimensional representation of the map for productive upgrading. Here, diversification is uniquely constrained by the subset of pairwise product proximities that matches a country's productive structure. Therefore, the density, denoted as ω, indicates the closeness of a productive structure to the required capabilities for producing a potential target:

$$\omega_{c,p} = \frac{\sum_c M_{cp}\, \phi_{p,p'}}{\sum_{p'} \phi_{p,p'}} \tag{2.8}$$

One of the most important applications of economic complexity methods is predicting product appearances by using the density measure. The greater the density of a country over a product, the greater the chances of having the required capabilities of that product, and thus of starting to export it in the future (Hidalgo *et al.*, 2007; Neffke; Henning; Boschma, 2011). Pinheiro *et al.* (2022) further refine this concept by proposing the relative density – the country density of a product compared to the products that the country does not export yet

(RCA<1) – to find that related diversification is more frequent than unrelated diversification for countries at lower levels of ECI. The relative density formula is:

$$\tilde{\omega}_{cp} = \frac{\omega_{cp} - \left(\frac{\sum_{p'} \omega_{cp'}}{N_{O_c}}\right)}{\sigma_{\omega_{cp'}}} \tag{2.9}$$

where $\frac{\sum_{p'} \omega_{cp'}}{N_{O_c}}$ is the average density of set of products that a country does not export yet and $\sigma_{\omega_{cp'}}$ is the standard deviation of the density of this set of products.

However, as firms contemplate their prospective payoffs in the journey of productive upgrading, a second-order problem of diversification emerges – an exploitation versus exploration dilemma. This dilemma arises when the potential payoff from exploiting existing products outweighs the incentives for exploring new possibilities. In such cases, there's a tendency to prioritize the exploitation of known products over the exploration of adjacent possibilities, hindering the recombination of existing knowledge and impeding the venture into the production of new products.

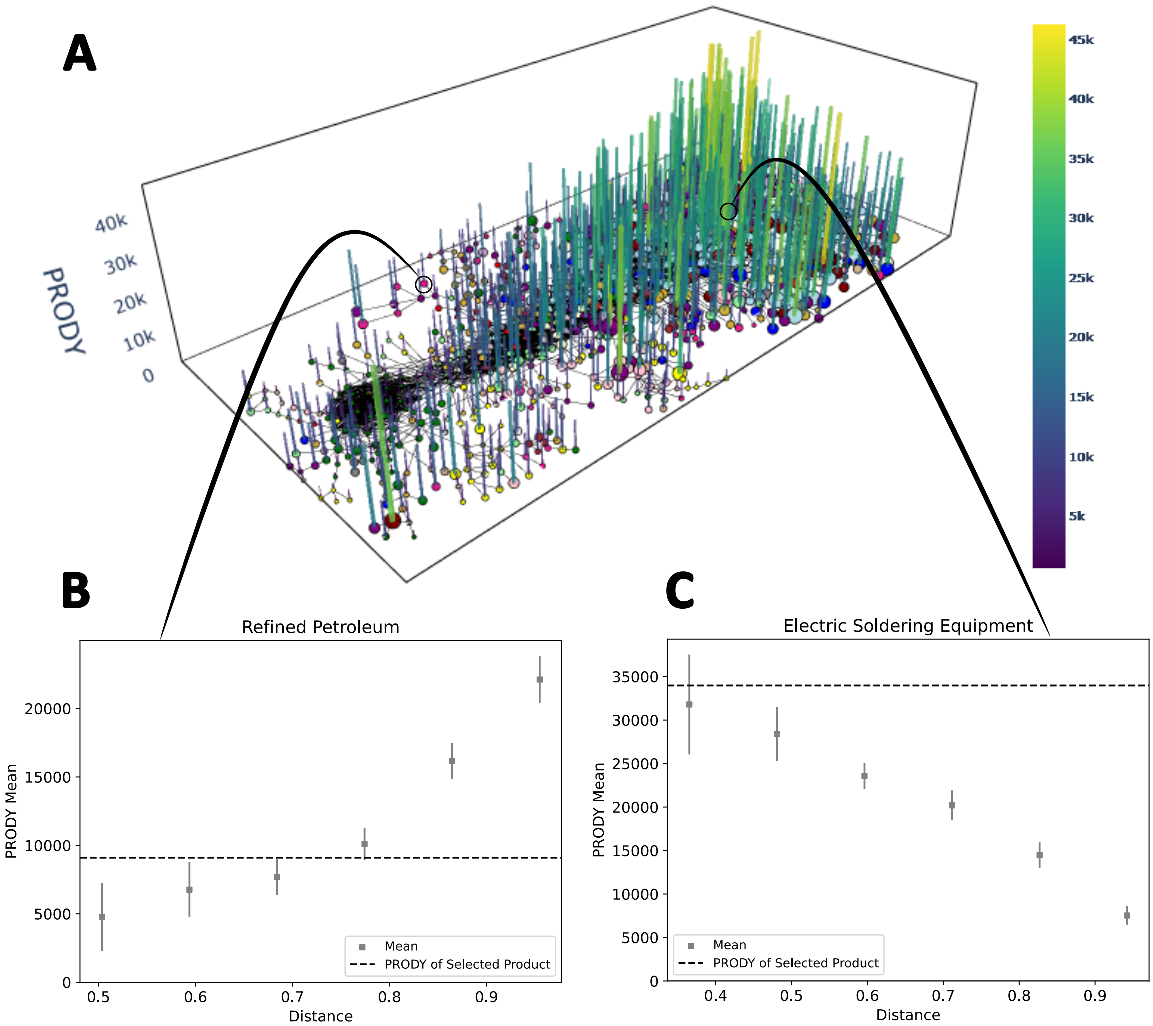
The exploration-exploitation dilemma is elucidated through the application of fitness landscapes, a concept first introduced by geneticist Wright *et al.* (1932) to examine species' adaptive evolution. These landscapes serve as a representation of the solution space for a problem, like topographical maps of physical terrain. Each location on the map corresponds to a potential solution, with its elevation reflecting the solution's payoff. A simple visual observation allows us to discern superior solutions. The use of fitness landscapes has expanded beyond its original purpose over time, becoming a versatile analytical instrument across diverse fields. As a model of more general evolutionary processes, fitness landscapes can depict any evolutionary dynamics, encompassing a wide range of areas such as economic development, organizational structures, belief systems, culture, language, among others, as noted by Mueller (2025).

We chose the PRODY measure (Hausmann; Hwang; Rodrik, 2007) to portray the payoff structure of the fitness landscape of product space. PRODY is the weighted average GDP of countries that exports a product with RCA ≥ 1, where the weights correspond to the RCA of each country c in product p:

$$\mathrm{PRODY}_p = \frac{\sum_c M_{cp}\, \mathrm{RCA}_{cp}\, Y_c}{\sum_c M_{cp}\, \mathrm{RCA}_{cp}} \tag{2.10}$$

In determining PRODY, we chose to use the current GDP per capita rather than Purchasing Power Parity (PPP) GDP to represent financial incentives more accurately, drawing on data from the World Bank. Using current prices allows for a more effective capture of the fluctuations in global prices, particularly in the commodities sector. To provide a detailed perspective on the financial returns landscape within the product space, we developed a three-dimensional visualization network, which we call the Product-PRODY-Space. Figure 2.1A showcases the Product-PRODY-Space for the year 2006.

Figure 2.1 – The Product-PRODY-Space



Panel **A** depicts the Product-PRODY-space, a three-dimensional visualization of the traditional product space layout, based on PRODY values for year 2006. **B** and **C** illustrate two different local maxima: electric soldering equipment (high-PRODY, core) and refined petroleum (low-PRODY, periphery), showing how average PRODY varies with distance from each product.

Initially, it is essential to visually examine the topology of the Product-PRODY-Space. If this landscape mimics Mount Fuji, characterized by one or a few local maxima, the choice between exploitation and exploration would not pose a dilemma. Yet, the actual terrain is generally characterized by a rugged landscape with abundant local maxima. Within the Product-PRODY-Space, a distinct core-periphery pattern emerges, where high-PRODY products cluster at the highly connected core of this three-dimensional network, and numerous low-PRODY local maxima scatter across its periphery. This configuration suggests a greater danger of a country getting trapped in network areas marked by lower payoffs, emphasizing the critical need to carefully navigate development paths in this complex landscape.

The examples of refined petroleum and electric soldering equipment underscore the nuances of local maxima, as detailed in Figures 2.1B and 2.1C. These figures demonstrate how the average PRODY of other products changes as their distance from these two examples increases. Electric soldering equipment is at the core of the network and represents a local maximum with good payoffs, featuring a high PRODY of approximately US$34,000 (compared to the US$46,300 PRODY of the global maximum – hormones). It also has a high PCI, of 4.16. Increasing distances to electric soldering equipment consistently result in a decrease in the average PRODY of products compared.

On the other hand, refined petroleum is in the periphery of the network, featuring a local maximum with a low PRODY, of US$9,100, compared to a PRODY average of US$15,473 and a median of US$14,109. It also has a low PCI, of -1.00. The products in the immediate vicinity have lower PRODYs than refined petroleum, whereas those positioned further away tend to achieve higher PRODY values. A product's average PRODY only surpasses that of refined petroleum when its distance to this product is near 0.8. We are particularly interested in low-PRODY, low-complexity local maxima such as refined petroleum because low-diversified countries that are overly dependent on this kind of products may face little payoff incentives to diversify, curbing their economic complexity enhancing.

While the Product-PRODY-Space provides insights for rugged landscapes and potential development traps, it may overlook some linkages between products. To address this, we suggest a more structured approach by introducing an index at the product level that evaluates the likelihood of a product being a local maximum relative to its closest counterparts. Begin by fixing the k top proximity Φ values of product p in the product space. If the kth value occurs more than once, consider the subset of all products that have proximities within this value range.

$$\mathrm{TOP}_{k,\phi_p} = \left\{p' \left(\phi_{pp'} \mid \mathrm{rank}(\phi_p) \leq k\right)\right\} \tag{2.11}$$

Next, establish the Product Peak Count (PPC) as the number of times that the PRODY of product p is equal to or higher than the PRODY of the products within the top k proximities.

$$\mathrm{PPC}_{p,k} = \sum_{p' \in \mathrm{TOP}_{k,\phi_p}} \begin{cases} 1, & \text{if } \mathrm{PRODY}_p \geq \mathrm{PRODY}_{p'} \\ 0, & \text{if } \mathrm{PRODY}_p < \mathrm{PRODY}_{p'} \end{cases} \tag{2.12}$$

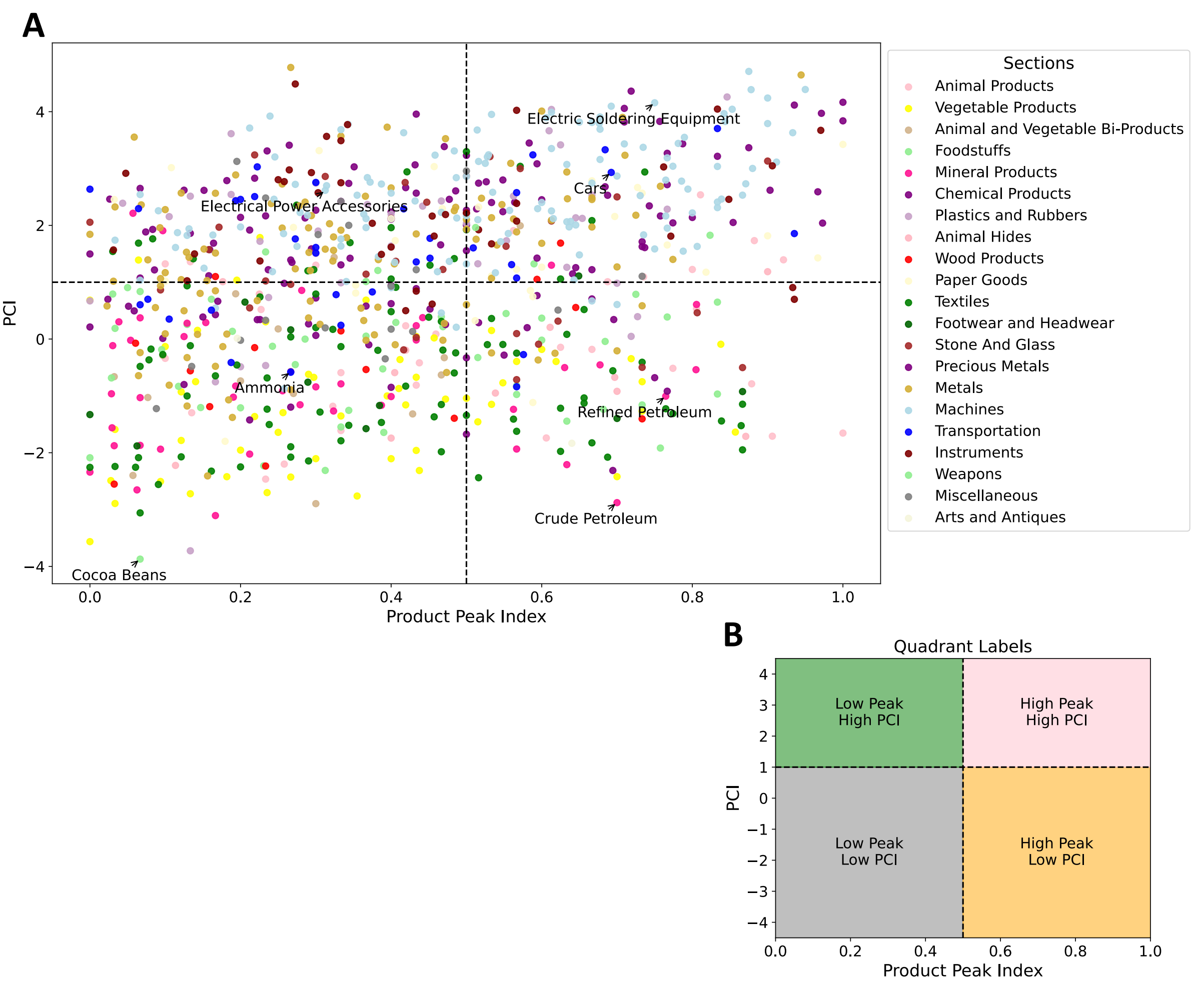
Finally, formulate the Product Peak Index (PPI) as a relative measure of the local maximum feature of a product in comparison to its closest neighbors.

$$\mathrm{PPI}_{p,k} = \frac{\mathrm{PPC}_{p,k}}{\left|\mathrm{TOP}_{k,\phi_p}\right|} \tag{2.13}$$

By setting k=30, we enable the PPI to capture the PRODY fitness of products based on their local vicinity. The bigger k is, the less important will be these local restrictions to the networked process of product diffusion. In this sense, incorporating the PPI into the analysis of the productive upgrading process allows for the inclusion of a layer of productive substitutability, complementing the implied productive complementarity introduced by Hidalgo *et al.* (2007).

In Figure 2.2, we plot the PPI versus the PCI of a product for the year 2006, splitting products into quadrants. The correlation between the PPI and the PCI is 0.3. The high-peak, high-PCI situation is portrayed in the first quadrant. It has products with good local maxima, such as cars and electric soldering equipment. The high complexity of these products means that they require a diverse set of capabilities. Therefore, occupying these local maxima generally does not represent lower chances of further diversification.

Figure 2.2 – Product Peak Index vs Product Complexity Index – Year 2006



Panel **A** plots each product's PPI against its PCI for year 2006, dividing the space into four quadrants. Product dots are colored by their corresponding product section, according to the HS92 (4-digit) classification. Panel **B** summarizes the quadrant classification.

The second quadrant is the high-peak, low-PCI zone. It presents the most relatable products to development traps. Most products in this quadrant belong to a few product sections—Mineral Products, Animal Products, Vegetable Products, and Textiles—which, despite their sectoral differences, share the structural feature of combining low complexity with high current returns. Lack of incentives for diversification coming from the high-peak feature can lead to a low-complexity equilibrium for a productive structure that is highly dependent on this kind of product.

This quadrant resonates strongly with the resource curse literature, which argues that dependence on natural resources, especially extractive commodities like oil and gas, can hamper long-term economic growth. Products located in this region offer strong short-term incentives due to relatively high current returns , but they lie in sparsely connected areas of the product space and embody low complexity. This combination creates an incentive structure that favors persistent exploitation over exploration, potentially trapping countries highly specialized in these products in a low-complexity equilibrium. The local maximum nature of these products, as captured by a high PPI and low PCI, adds a structural dimension to the resource curse hypothesis, relatable to the Dutch Disease: they anchor productive structures in areas of the product space that offer limited paths to future upgrading.

Interestingly, not all products in this quadrant are natural resources. While many paradigmatic cases—such crude and refined petroleum—fit the classical narrative of the resource curse, the identification of non-resource-based products that share similar structural characteristics raises important policy considerations. These are goods that, despite not being commodities in the traditional sense, also exhibit high local payoff and low complexity, and lie in poorly connected regions of the product space. If intensely exploited, they may reproduce the same developmental inertia observed in resource-rich economies, reinforcing specialization patterns that discourage diversification and hinder the accumulation of new capabilities.

The low-peak, low-PCI zone has products suitable to diversification but with an undetermined path for ECI growth. Ammonia, for example, stands in the fourth quadrant. The closest products to ammonia are refined and crude petroleum. Therefore, even having diversification prospects, it does not necessarily imply a virtuous path to economic development. The fourth and last quadrant is the low-peak, high-PCI region, which has products with high impact for diversification and ECI growth. Strategic industrial policy should target products in this zone.

In this section, we presented the Product-PRODY-Space, a rugged fitness landscape far removed from the simplicity of a singular peak, emphasizing the existence of many local maxima products with low payoffs in the less connected areas of this network. The introduction of the PPI further enhances our analysis by quantitatively evaluating the likelihood of a product being a local maximum in its product neighborhood, thereby shedding light on a second-order problem of diversification. It highlights the aspects of product substitutability within the process of productive diffusion, in addition to the productive complementary implied by the original product space. In the next section, we leverage the PPI to establish a country-level metric, aiding in the exploration of country development trajectories and the pinpointing of complexity traps.

2.4 The Country Peak Index, Development Paths, and Complexity Traps

In this section, we analyze development paths and complexity traps through the lens of the intertwined relationship between local maxima and relatedness. We begin by examining how a country's level of diversification, when combined with its degree of specialization in peak products, relates to its overall level of economic complexity. This aggregate analysis enables the identification of distinct forms of low-complexity developmental stasis, including one rooted in overexploitation of high-PPI products. In the second part of the analysis, we investigate whether specialization in such products hinders the exploration of untapped opportunities at the country–product level, curbing diversification even after controlling for relatedness.

Previous literature has highlighted that the relationship between diversification and income is not linear, but rather follows a hump-shaped trajectory: countries tend to diversify as they develop, up to a point where further advancement is associated with re-specialization Imbs and Wacziarg (2003). This same principle can be extended to the structure of economic complexity. When projected into a phase space of diversification and local maxima specialization, countries are expected to follow non-linear paths shaped by the types of products they populate. In particular, countries that populate high-peak but low-complexity products too early in their development may face strong incentives to keep exploiting these local maxima, thereby reducing the returns from adjacent exploration.

Depending on the stage of diversification, populating local maxima products can yield very different outcomes for knowledge accumulation. In the course of diversification, productive upgrading is expected to occur by populating products with better payoffs, thus moving closer to local maxima. However, if a country in the early stages of capability accumulation specializes prematurely in such products—especially those in the high-Peak, low-PCI quadrant of Figure 2.2—the incentives for venturing into nearby, more complex products rapidly decay. This premature hump reinforces the lack of prospects for accumulating capabilities through the means of diversification, launching the country into a harsh mix of quiescence and local maxima developmental trap.

We build on the PPI to create a country-level measure for acknowledging the local maxima stance of development paths. The Country Peak Index (CPI) measures how much a set of products exported by a country occupies local maximum positions in the product space. CPI is the average of the PPI of the products a country exports with RCA, weighted by the share of each product in the country's export basket. The export share s_{cp} works as normalizing factor to account for the country-varying financial incentives of exporting a product.

$$\mathrm{CPI}_c = \frac{\sum_p M_{cp}\, s_{cp}\, \mathrm{PPI}_p}{\sum_p M_{cp}\, s_{cp}} \tag{2.14}$$

Country development trajectories often follow structurally constrained paths, shaped not only by capabilities accumulated over time but also by their surrounding configuration of productive possibilities and diversification incentives. We analyze these dynamics empirically by constructing a phase space composed of two key structural indicators: CPI and Diversity. Figure 2.3 presents the evolution of selected countries within this phase space in the period between 1998 and 2014.

Figure 2.3 – Country Peak Index vs Diversity – 1998-2014

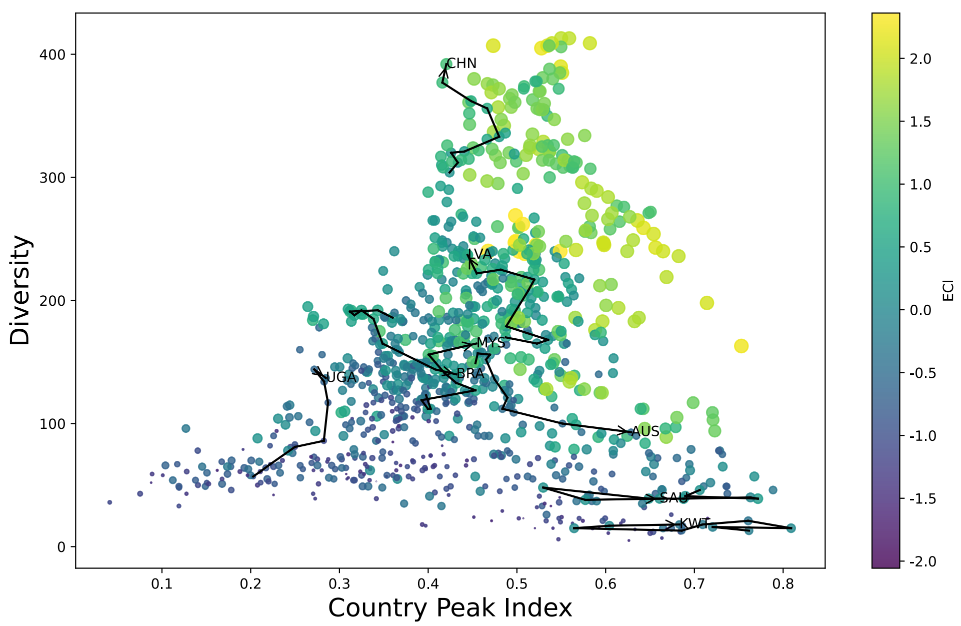


Figure 2.3 displays all country-year observations in the CPI–Diversity space from 1998 to 2014, with 2-year intervals. Color and size reflect the corresponding ECI. Development trajectories are shown for selected countries.

The figure reveals a triangular structure. Countries with low diversity tend to exhibit low complexity, populating the base of the triangle, while the upper-right corner concentrates the most diversified and complex countries. This suggests that different combinations of CPI and Diversity can lead to high levels of complexity, but the viable region is bounded. Although a developmental direction may appear self-evident, it reinforces the notion that countries starting from low levels of both CPI and Diversity must experience a joint increase in these dimensions in order to reach high ECI levels, a path that is far from automatic and often subject to structural constraints.

Country development trajectories exhibit different patterns. Uganda is the typical country escaping from the quiescence trap, but still in the initial stages of development in which country diversification increases in hand with the CPI and the ECI. Latvia is the case of a country that has been successfully climbing the ladder of economic complexity, but with still unclear prospects if it is time to start retreating from diversification with a respective increase in the CPI for becoming a highly complex economy. China, a case by its own, is a highly diversified economy with a clearer challenge of passing through the hump phase and start specializing in products of the high peak, high PCI quadrant of Figure 2.2.

Regardless of their distinct positions in Figure 2.3, Brazil and Australia follow comparable development trajectories. Both nations are significant players in the global commodities markets and have experienced a pronounced concentration of their export portfolios into a more restricted set of commodities. These trends raise concerns regarding a premature occurrence of the hump effect and the subsequent ability of these countries to sustain a beneficial process of productive upgrading. Saudi Arabia and Kuwait, on the other hand, exemplify countries with low diversity and high CPI, where the composite quiescence and local maxima trap takes place. During the 1998-2014 period, both countries have not made considerable advances in diversity nor economic complexity.

While Figure 2.3 provides a representation of the CPI–Diversity space, capturing the variety of developmental trajectories, a more structured analysis is required to uncover deeper regularities and patterns that may govern the evolution of productive structures. To this end, we adopt a discretized approach, segmenting both CPI and Diversity into $Q = 5$ quantiles and generating a $Q \times Q$ grid. Each cell (i, j) in this grid defines a discrete developmental state $S_{i,j}$, characterized by three analytical components:

1. **Average Economic Complexity (ECI):**

$$\mathrm{ECI}_{i,j} = \frac{1}{|C_{i,j}|} \sum_{c \in C_{i,j}} \mathrm{ECI}_c \tag{2.15}$$

   where $C_{i,j}$ is the set of country-year observations falling into cell (i, j).

2. **Average Transition Vector:** For each country-year c, we define the transition vector as:

$$\Delta\vec{x}_c = \left(\mathrm{CPI}_{c,t+\Delta t} - \mathrm{CPI}_{c,t},\ \mathrm{Diversity}_{c,t+\Delta t} - \mathrm{Diversity}_{c,t}\right) \tag{2.16}$$

   where Δt is the length of the time window : four years. The average transition vector in cell (i, j) is then:

$$\vec{v}_{i,j} = \frac{1}{|C_{i,j}|} \sum_{c \in C_{i,j}} \Delta\vec{x}_c \tag{2.17}$$

This vector field captures the average directional movement of countries across developmental states.

3. **Attractor Identification:**
A cell $S_{i,j}$ is classified as an attractor if it satisfies two conditions:
   - *Low local velocity:*
   $$\|\vec{v}_{i,j}\| < \epsilon \tag{2.18}$$
   where $\epsilon = 0.01$, indicating stagnation or directional neutrality.
   - *Local convergence dynamics:* we estimate the Jacobian matrix $J_{i,j}$ of local transitions:
   $$J_{i,j} = \begin{bmatrix} \frac{\partial \Delta \text{CPI}}{\partial \text{CPI}} & \frac{\partial \Delta \text{CPI}}{\partial \text{Diversity}} \\ \frac{\partial \Delta \text{Diversity}}{\partial \text{CPI}} & \frac{\partial \Delta \text{Diversity}}{\partial \text{Diversity}} \end{bmatrix} \tag{2.19}$$
   A cell is classified as a common attractor if all eigenvalues of $J_{i,j}$ have negative real parts:
   $$\text{Re}(\lambda_k(J_{i,j})) < 0, \quad \forall k$$
   and as a spiral attractor if the eigenvalues have negative real parts but non-zero imaginary components:
   $$\text{Re}(\lambda_k(J_{i,j})) < 0 \quad \text{and} \quad \text{Im}(\lambda_k(J_{i,j})) \neq 0$$

Figure 2.4 displays the empirical implementation of this framework. It allows us to uncover convergence patterns, structural inertia, and development traps in a topological fashion, moving beyond individual trajectories to a systemic understanding of productive transformation. The background color gradient denotes the average ECI levels in each cell, while the black arrows represent the average transition vectors $\vec{v}_{i,j}$, i.e country transitions over four-year periods. Red markers indicate attractors, either spiral or common, representing zones of convergence in the development space. Attractor regions can be interpreted as empirical analogs to steady states in dynamical growth models, where structural inertia stabilizes trajectories. Common attractors exhibit minimal motion and direct convergence, while spiral attractors involve rotational dynamics with dampened trajectories. Both types are identified where vector magnitudes fall below 0.01 and the Jacobian has eigenvalues with negative real parts.

Our analysis identifies three attractor regions within this phase space, each associated with distinct development dynamics and structural constraints. Region 1, the Middle-Development Stasis, corresponds to countries with intermediate levels of both CPI and Diversity. This is the most common trap for developing economies — a structural analog to the well-known middle-income trap. In this zone, vector directions are weak and multidirectional, indicating that countries struggle to build the structural momentum necessary to advance. As countries transition from low to medium levels of economic complexity, diversification strategies that follow closely the logic of relatedness tend to become exhausted.

Figure 2.4 – Phase Space of Productive Development: Vector Fields and Structural Traps

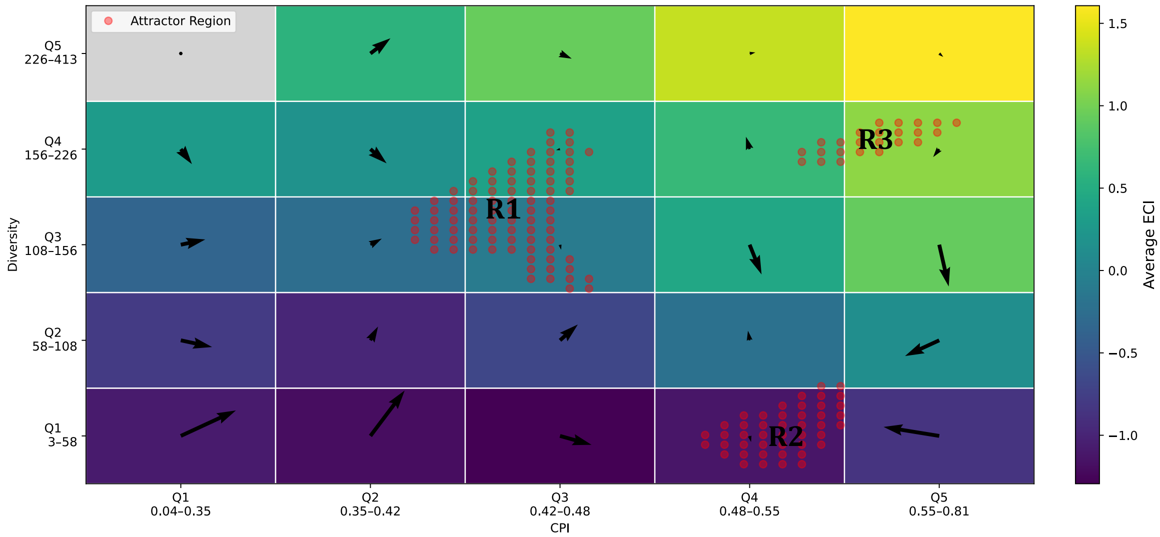


Figure 2.4 depicts a discretized CPI–Diversity phase space in 5×5 grid of quantile-based cells, providing a basis for the analysis of structural development trajectories via vector fields and attractor identification. Background colors represent average ECI; black arrows show average country transition vectors over four-year periods. Red dots indicate attractor regions, defined as zones of low vector magnitude (< 0.01) and local convergence (negative eigenvalues of the Jacobian). Three key attractor regions are identified: R1, R2, and R3.

In other words, while initially effective, these related diversification pathways cease to be sufficiently ECI-enhancing once a country reaches intermediate complexity. Escaping Region 1, therefore, often requires a strategic shift toward more ambitious and structurally transformative moves — including unrelated diversification into products that lie further from the current productive frontier (Pinheiro *et al.*, 2022).

Region 2, the Resource-Based Stasis, is located in the bottom-right quadrant — characterized by high CPI and low Diversity. It typically includes countries whose productive structures are highly concentrated in a few high-peak products, most often natural-resource intensive. This region aligns closely with the logic of the natural resource curse: while initial specialization in commodities may provide high returns and elevate the CPI, it often comes at the cost of long-term structural transformation. The stasis here is of a different kind than in Region 1 — it is not the exhaustion of relatedness, but the entrenchment of early specialization. If resource exploitation is not sufficient to propel the country to high-income status — or if the gains are not reinvested in productive diversification — the country may become trapped in a fragile equilibrium with limited structural resilience and little scope for complexity upgrading.

Region 3, the High-Complexity Equilibrium, occupies the upper-right corner of the phase space. This is the "let it be" zone: countries here combine high levels of CPI and Diversity, forming the most desirable and stable structural configuration. These countries have successfully accumulated a broad set of capabilities and managed to integrate them into complex products. Vector magnitudes are minimal and convergence is strong, indicating that once this region is reached, little structural adjustment is required to sustain high economic performance. While entry into this region is rare and typically nonlinear, it represents the empirical manifestation of long-run development success.

The configuration of attractors and directional vectors highlights the path dependency nature of economic development. An underdeveloped economy begins in the bottom-left, where both CPI and Diversity are low. From there, countries must diversify (move upward) and eventually specialize in increasingly complex activities (move rightward). However, the empirical evidence shows that many fail to complete this trajectory. Some fall prematurely into Region 2, attracted by early gains from resource-based specialization, while others stall in Region 1, unable to generate the structural momentum needed to transition further. Only a few manage to navigate through these zones and reach Region 3, where the productive structure becomes self-reinforcing.

This representation also allows for a reinterpretation of the hump effect in the development literature — the idea that export diversification rises during early and intermediate stages of growth but eventually declines as countries reach advanced stages. In this framework, the timing of the hump becomes crucial. If diversification contracts prematurely, for example, while still in Region 1 or 2, it reinforces structural traps and prevents the transition to a high-complexity equilibrium. In contrast, a late-stage hump, occurring only after a country has attained high levels of diversification and structural readiness, may reflect efficient specialization rather than stasis. Hence, the policy implication is that the hump is not a problem in itself — but mistiming it is.

From a developmental policy standpoint, countries positioned in Regions 1 and 2 may require policies aimed at breaking structural stasis. Incremental or path-following efforts may not suffice; rather, these countries must consider more ambitious strategies that promote unrelated diversification — that is, the deliberate entry into sectors that are not adjacent in the product space but hold high structural potential (Alshamsi; Pinheiro; Hidalgo, 2018; Pinheiro *et al.*, 2022). Thus, policy must carefully calibrate when to shift from related to unrelated strategies, ensuring that the economy is sufficiently prepared to absorb and sustain more complex capabilities. The goal is not merely to follow comparative advantage, but to strategically reshape it in a way that unlocks new developmental trajectories.

In sum, the Diversity–CPI phase space offers a rich empirical landscape of structural transformation, stasis, and divergence. Economic development is not a smooth or automatic process, but one punctuated by zones of resistance and convergence. Understanding where a country is located in this phase space — and how its productive structure responds to policy choices and opportunity — is essential for navigating viable development trajectories. This macro-level mapping helps uncover the structural conditions under which countries may fall into complexity stagnation or advance toward higher levels of economic sophistication.

Yet, while the phase space reveals systemic, aggregate dynamics, it also raises a complementary micro-level question: how does the structure of opportunity operate at the level of individual products? Next, we shift focus from national development trajectories to product-level activation, examining how the CPI interacts with relatedness to explain whether a country is likely to start exporting a new product. This micro-level lens allows us to explore how countries move through the product space and whether the CPI helps refine the predictive power of traditional diffusion measures like density.

We want to test whether there is a hump-shaped effect of the CPI on the probability of exporting a new product, conditional on the density and relative density of a country with respect to that product. This investigation is motivated by the macro-level dynamics revealed in Figure 2.4, where we observed that most countries initially progress by simultaneously increasing both their CPI and export Diversity, a path consistent with a phase of exploratory growth. In this early stage, higher CPI levels may reflect a productive structure that supports entry into a broader range of new products, reinforcing diversification.

However, if the CPI continues to rise without a commensurate expansion of the product basket, diversification incentives may begin to weaken. A hump-shaped effect would suggest that, while increases in CPI initially promote exploration by enhancing the structural base for diversification, there comes a point where the incentive structure shifts. At higher levels of CPI, countries may begin to prioritize the exploitation of existing specializations rather than the exploration of new activities, leading to a decline in the likelihood of diversifying into additional products. In this sense, the hump marks a transition from exploration-led upgrading to a potential narrowing of focus.

We use a Probit model to test CPI's effect on predicting product appearances on the country-product level, as in O'Clery, Yıldırım, and Hausmann (2021). The complete regression model takes the following form:

$$\begin{aligned} P(\mathrm{RCA}_{cp_{t+4}} \geq 1 \mid \mathrm{RCA}_{cp_t} \leq 0.5) = \beta_1 \omega_{cp_t} + \beta_2 \tilde{\omega}_{cp_t} + \beta_3 (\omega_{cp_t} \cdot \tilde{\omega}_{cp_t}) \\ + \beta_4 \mathrm{CPI}_{cp_t} + \beta_5 \mathrm{CPI}^2_{cp_t} + \mu_t \end{aligned} \tag{2.20}$$

where $P(\mathrm{RCA}_{cp_{t+4}} \geq 1 \mid \mathrm{RCA}_{cp_t} \leq 0.5)$ is the probability of becoming a competitive exporter of that product in the next four years, given that the country's RCA of that product was lower or equal than 0.5 in year t. The explanatory variables are, in order, density, relative density, the interaction term between these densities, the CPI and its quadratic term, and a dummy that controls for year fixed effects. We use a 2-year window for t, ranging from 1998 to 2016.

We collect pseudo-R^2, AUC-ROC (Area Under the Receiver Operating Characteristic Curve) and best F1 score statistics to compare the tested Probit models, following O'Clery, Yıldırım, and Hausmann (2021) and Albora *et al.* (2023). The AUC-ROC and the F1 score are both performance indicators used to evaluate the quality of binary classification models. AUC-ROC is a plot of the ratio of true positive predictions to the total number of actual positive instances (True Positive Rate, or Recall) against the ratio of false positives to the total number of actual negative instances (False Positive Rate, or 1 - Specificity) at various threshold settings. Its value can go from nil to unity, with 0.5 representing a random classifier, and 1.0 representing a perfect model. The F1 Score is a harmonic mean of the ratio of true positive predictions to the total number of positive predictions (Precision) and the Recall measure. The Best F1 score is computed by finding the threshold that maximizes the F1 score. F1 scores are interesting to compute when there is an imbalanced distribution, where one class is much more frequent than the other. In this case, out of 773,974 observations with $\mathrm{RCA}_{cp_t} \leq 0.5$, only 11,954 (1.54%) ended up being successful, i.e. $\mathrm{RCA}_{cp_{t+4}} \geq 1$.

Table 2.1 summarizes the results obtained for four different specifications of the model. Model 1 serves as the baseline, incorporating only density and the year dummies. Model 2 builds on the baseline by adding relative density and its interaction term with density. Model 3 adds a linear term for CPI, keeping density, relative density and year dummies. Model 4 presents the complete specification by further including the quadratic term of CPI. An increase in predictive power is evident in the pseudo-R^2 measure, rising from 0.9% in Model 1 to 3.8% in Model 2, to 4.8% in Model 3, and finally to 5.0% in Model 4. Additionally, the AUC-ROC for Model 1 is 0.602, increasing to 0.682 in Model 2, 0.695 in Model 3, and further to 0.699 in Model 4. The Best F1 Score follows a similar trend, with values of 0.0420 for Model 1, 0.0696 for Model 2, and 0.0762 for both Models 3 and 4, indicating improved balance between precision and recall with each subsequent model specification. The improvements of the complete model when compared to the baseline are similar to alternative frameworks for predicting product appearances, such as using the ‘EcoSpace’ (O’Clery; Yıldırım; Hausmann, 2021) or Random Forest approaches (Albora *et al.*, 2023).

In Model 3, the coefficient for CPI is negative and statistically significant, indicating that higher levels of CPI are associated with a lower probability of product appearances. In Model 4, we further account for nonlinearities by including the quadratic term of CPI. The results reveal a concave (inverted-U) relationship: while initial increases in CPI positively affect the likelihood of product activation, this relationship eventually reverses. The negative and significant coefficient on CPI^2 confirms the presence of a hump-shaped effect of CPI on diversification prospects.

In Figure 2.5, we illustrate this non-linearity by plotting the predicted probabilities for different levels of CPI. The peak of the curve occurs at CPI $\approx$ 0.235, where the probability of success is maximized. As CPI increases beyond this threshold, the predicted probability declines steadily, falling back to the baseline level at CPI $\approx$ 0.43, and approaching zero near the upper bound of the sample (CPI = 0.81). Interestingly, the CPI level of 0.43 aligns with Region 1 of the macro-level phase space, where structural inertia dominates and development

Table 2.1 – Probit results for predicting country-product appearances.

	Probit - Predicting appearances in *t+4*							
	Coefficients				Marginal Effects			
	1	2	3	4	1	2	3	4
Density_t	1.0473***	0.6663***	0.9236***	0.7367***	0.0394***	0.0222***	0.0297***	0.0234***
	(0.035)	(0.04)	(0.042)	(0.044)	(0.0013)	(0.0013)	(0.0013)	(0.0014)
$\text{Relative Density}_t$		0.2513***	0.2593***	0.2677***		0.0084***	0.0083***	0.0085***
		(0.006)	(0.006)	(0.007)		(0.0002)	(0.0002)	(0.0002)
$\text{Density Interaction}_t$		-0.1767***	-0.2713***	-0.2912***		-0.0059***	-0.0087***	-0.0092***
		(0.036)	(0.037)	(0.038)		(0.0012)	(0.0012)	(0.0012)
CPI_t			-0.8901***	1.152***			-0.0286***	0.0365***
			(0.029)	(0.147)			(0.0009)	(0.0046)
CPI^2_t				-2.4465***				-0.0776***
				(0.172)				(0.0054)
Pseudo R^2	0.008	0.040	0.048	0.050				
AUC-ROC	0.602	0.682	0.695	0.699				
F1 Score	0.0420	0.0696	0.0762	0.0762				

Table 2.1 reports Probit regression results with period-fixed effects, using intervals of two years between each t from 1998 to 2016, totaling 773,974 observations. Standard errors are in parentheses. ***$p < 0.01$.

trajectories tend to stall. This reinforces the interpretation that, beyond a certain point, rising CPI signals diminishing returns to exploration, as exploitation incentives begin to dominate, echoing the attractor dynamics and complexity traps identified in the macro framework.

Figure 2.5 – Probit Model 4 - Predicted Probability vs. CPI

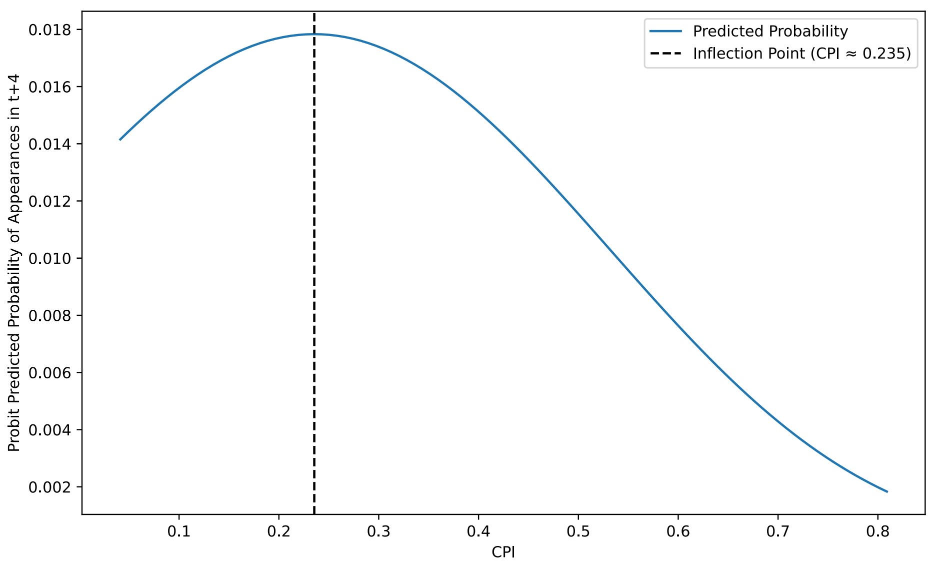


Figure 2.5 illustrates the hump-shaped effect of the Country Peak Index (CPI) on the predicted probability of a product appearance within four years. The dashed line marks the inflection point.

The hump effect of CPI in the country-product microlevel suggests that when a country shows high levels of CPI, incentives to keep exploiting its current set of product exports may supersede incentives to explore new products, even after controlling for density and relative density. It suggests that there is more to gathering all the required productive capabilities to start exporting a product. Diversification cannot be explained by the principle of relatedness alone. An additional layer, of relative pecuniary returns, also plays an important role in defining diversification. Connecting with Dam and Frenken (2022), the hump effect may be the resulting concentration process of the export set not only in the last stages of development but also in any development stage for a country with an excessively high level of CPI.

Linking the findings of the country-product microlevel with the macrolevel of country phase space, a low-diversified country may face a premature diversification unwinding or stall when its CPI spikes without further increases in diversity, impeding the accumulation of productive knowledge at the extensive margin, and thus hampering economic complexity growth. Such dynamics are particularly evident in Region 2 of the Diversity–CPI phase space (Figure 2.4), which we identify as a natural resource-based trap. If the exploitation of high peak products is not sufficient for the economy to achieve high-income status, the country may find itself in an income trap as well.

This section explored the intricate dynamics between local maxima, diversification, and the corresponding economic complexity of countries. We modeled the phase space of development by discretizing the Diversity–CPI space into quantile-defined cells and estimating the average directional transitions and structural attractor zones across countries over time. This approach revealed regions of stasis and convergence, allowing us to identify where development trajectories are likely to stall and under what structural configurations. We also analyzed the non-linear impacts of CPI in product appearances at the country–product level, concluding that high levels of CPI curb diversification probabilities even after controlling for relatedness. These insights underline the delicate balance between diversifying and specializing, whereas premature specialization into products with high PPI but low PCI risks leading countries into a newly identified complexity trap, characterized by the joint effects of quiescence and local maxima developmental stasis. In the next section, we discuss these findings.

2.5 Discussion

Our study ventures into the critical domain of economic convergence within development economics, highlighting complexity convergence and the diverse pathways through which productive structures evolve. We propose a nuanced hypothesis that alongside the inherent complementarity in the product space, there's a concurrent phenomenon of productive substitutability. This theory sheds light on the tendency of economies to prioritize the exploitation of existing products due to their relative higher returns, often neglecting the exploration of novel opportunities in the adjacent possible.

We demonstrate how certain configurations of productive structures can inadvertently lead to the overexploitation of suboptimal peaks, thereby constraining diversification and limiting the integration of new productive knowledge. To capture this mechanism, we introduce a new metric that quantifies exposure to local maxima development traps, extending the discourse beyond the traditional quiescence trap. This tool enables a fine-grained analysis of development trajectories and helps identify both aggregate and individual aspects under which countries are more likely to experience stagnation in complexity upgrading.

Our findings also offer actionable insights for industrial policy, particularly regarding the timing and targeting of interventions. In addition to advising against premature specialization in low-complexity peak products—thereby contributing to the "what" and "when" framework of Hidalgo (2023)—our results suggest that countries situated within or near empirically identified zones of developmental stasis may require more intensive and proactive industrial strategies. Specifically, these situations may justify the pursuit of unrelated diversification (Pinheiro *et al.*, 2022). In such contexts, policy may play a decisive role in overcoming inertia, unlocking new paths for capability accumulation and structural transformation.

Further exploration is warranted to dissect the implications of product overexploitation on diversification at the microlevel of country-product interactions. It would also be insightful to investigate how incentives for diversification and specialization shift during the critical transition from a middle to a high-complexity economy, integrating the hump phenomenon with the economic sophistication's S-curve. Lastly, given the multidimensional nature of economic complexity (Stojkoski; Koch; Hidalgo, 2023), which spans beyond international trade to encompass various economic processes, it's important to examine the applicability of the local maxima framework across different datasets, thus enriching our comprehension of economic complexity in a broader context.

Beyond its role in explaining productive upgrading dynamics, the CPI may serve as a diagnostic tool for broader development challenges. Countries with high CPI but without a corresponding level of economic complexity may not only face structural inertia in their diversification processes, but also exhibit deficits in key dimensions of human development. These may include weaker educational outcomes, lagging health indicators, lower institutional quality, and even heightened political instability—factors that often coevolve with productive structures. Future research could explore these associations systematically, using the CPI to investigate how pecuniary specialization patterns intersect with multidimensional development outcomes. In doing so, the CPI could contribute to a more integrated understanding of structural development traps across both economic and social domains.

3 Less is More: How Relatedness Filtering Enhances Productive Upgrading Predictions

Abstract

Relatedness measures indicate how close a country is to developing a new activity, with implications for industrial policymaking. However, estimating relatedness is challenging, and relying on its traditional, Product Space-based measure to design industrial policies can be misleading. To overcome its limitations, several attempts have been made to completely reformulate relatedness, including the use of machine learning techniques. Yet, simplicity can be just as effective. Rather than proposing an entirely new approach, we suggest using network filtering as an intermediary step in the standard relatedness calculation, testing several network filtering methods for the Product Space. We find that filtered relatedness significantly outperforms baseline relatedness in predicting which new activities a country will develop, providing policymakers with a valuable tool to design more effective and lower-risk industrial strategies.

Keywords: relatedness, product space, economic complexity, network filtering, backbone extraction

3.1 Introduction

Industrial Policy Is Back. Now What?[1]

The resurgence of industrial policy as a key component of economic strategy has sparked debates among economists and policymakers about how best to design effective interventions. Once considered a marginal approach, industrial policy has now become a focal point of policy discourse, with the question shifting from ‘if’ to ‘how’ it should be implemented. One promising framework aiding this shift is the use of Economic Complexity and Relatedness metrics, using network theory to identify strategic industries that match with a country’s current capabilities. In practice, governments, think tanks, and state institutions have already begun employing these metrics as tools for crafting targeted industrial policies.

[1] Authors: Jakob Hafele, Célestin Mong, Tano Santos, Luigi Zingales, Dani Rodrik and James K. Galbraith. Access on: https://www.project-syndicate.org/onpoint/industrial-policy-is-back-now-what.

Despite the appeal of these metrics, relying solely on the traditional relatedness measure presents limitations that could misguide policy decisions. The traditional approach often fails to capture the nuances required for accurate prediction of productive upgrading. This is especially risky for countries with constrained resources and high opportunity costs in policy choices, where missteps could exacerbate economic challenges rather than alleviate them. In response, recent literature has explored alternative, complex approaches. In this study, we explore a simpler, yet powerful alternative: applying network filtering as an intermediary step in the calculation of relatedness density of a country over a product.

Rather than proposing a new relatedness metric entirely, we examine how filtering techniques, particularly those designed to extract the backbone of networks, can enhance the predictive performance of traditional relatedness measures. By focusing only on the most meaningful connections within the Product Space, filtered relatedness reduces informational noise and highlights actionable insights, offering policymakers a streamlined yet effective tool for identifying high-potential targets. This minimalist approach, as the title suggests, operates on the principle that 'less is more', suggesting that, in the context of productive upgrading, simplicity can lead to more accurate and practical results.

While this study applies network filtering techniques to international trade and product diversification, the methodology developed here could also be extended to the analysis of knowledge diffusion and patent networks. Filtering noisy connections in these contexts could help uncover more precise patterns of technological development and innovation dynamics, addressing similar informational challenges to those encountered in trade-based relatedness measures. This perspective resonates with the work of Balland and Rigby (2017) and Balland *et al.* (2019), who demonstrated the importance of accounting for network structure in studying the diffusion of complex knowledge across technologies. Extending filtering techniques to these domains represents a promising avenue for future research.

In the following section, we review the existing literature on relatedness metrics, highlighting the underutilization of network filtering as a tool to improve predictions of productive upgrading. We then present our methodological approach, applying distinct network filtering techniques to the product space. Some of these techniques adapt established statistical backbone extraction methods to incorporate directionality into the filtering process. Among these, one adapted method — the Directed Disparity Filter In-Degree (DDF-In-Degree) — stands out, outperforming the others in predicting which products are likely to become competitive exports for a country. To illustrate the practical implications, we conduct a brief case study on Brazil and the Philippines, comparing industrial policy recommendations derived from DDF-Out-Degree relatedness density with those based on traditional relatedness density. Finally, we conclude with a discussion and propose directions for future research.

3.2 Literature Review

Relatedness was introduced as an empirical description of the probability that a region enters (or exits) an economic activity as a function of the number of related activities present in that location (Hidalgo *et al.*, 2018). It helps identify optimal pathways for countries to diversify their productive structure (Hausmann *et al.*, 2014), ultimately promoting economic development. The core idea behind relatedness is that a country's potential to start producing a specific product can be estimated through colocation — if a country already produces a set of products that are typically produced alongside the target product, it is likely to meet the productive requirements of that new product.

The concept of *relatedness* is gaining traction in industrial policy debates, as it offers a framework to identify strategic directions for countries to expand into new products and industries. In 2019, the World Bank's Global Economic Prospects (World Bank, 2019) applied relatedness and complexity metrics to assess development trajectories across selected nations. The European Commission's Joint Research Center has also used this framework, publishing studies to guide member states in formulating development strategies (Pugliese; Tacchella, 2020; Caldarola *et al.*, 2024). Recently, the Draghi Report (Draghi, 2024) further spotlighted relatedness, discussing optimal approaches for Europe's future economic development. Online platforms such as DataViva (for Brazil) and DataMexico (for Mexico) have also made complexity and relatedness metrics accessible within national contexts, supporting informed policymaking.

Traditional approaches, pioneered by Hidalgo *et al.* (2007), measure relatedness based on the idea of Revealed Comparative Advantage (RCA), as defined by Balassa (1965). A country is defined to be competitive in exporting product p, if the weight of the exports of product i relative to its export portfolio is greater than or equal to the share of that same product in global exports, i.e. $RCA_{c,p} \geq 1$ with $RCA_{c,p}$ defined as

$$RCA_{c,i} = \frac{X_{c,p}}{\sum_p X_{c,p}} \cdot \frac{\sum_c X_{c,p}}{\sum_{c,p} X_{c,p}} = \frac{x_{c,p} \cdot \sum_{c,p} x_{c,p}}{\sum_p x_{c,p} \cdot \sum_c x_{c,p}} \tag{3.1}$$

Based on the RCA, Hidalgo *et al.* (2007) first define an unweighted bipartite graph between countries and products, $M_{c,p}$, in which a country is linked to a product if the country exports the product with a revealed comparative advantage:

$$M_{c,p} = \begin{cases} 1 & \text{if } RCA_{c,p} \geq 1; \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

To measure relative similarity between products, the authors count how many times each pair of products is co-exported and divides this co-occurrence by the number of countries that export each of these two products, picking the smallest resulting number. The result is the minimum conditional probability of having RCA in each pair of products:

$$\varphi_{p,p'} = \frac{\sum_c M_{c,p} M_{c,p'}}{\max\left(\sum_c M_{c,p}, \sum_c M_{c,p'}\right)} \tag{3.3}$$

The resulting network $\varphi_{p,p'}$ is called Product Space, a unipartite graph that translates into a correlation matrix between products. The value of pairwise relative similarity is called proximity and goes from nil to unity. The assumption is the greater the value of proximity, the more productive requirements are expected to be shared between the two products. Baseline relatedness density, denoted as $\omega_{c,p}$, allegedly indicates the closeness of a productive structure — a country — to the required capabilities for producing a potential target:

$$\omega_{c,p} = \frac{\sum_{p'} M_{c,p'} \varphi_{p,p'}}{\sum_{p'} \varphi_{p,p'}} \tag{3.4}$$

What explains the predictability of the presence or absence of specific industries in each country is the nestedness of the international trade network (Bustos *et al.*, 2012). Nestedness is a well-established concept in ecology used to analyze the structure of ecological systems. In these systems, species are categorized as either generalists or rare, much like locations can be diverse or less diverse. Nestedness describes a pattern in which rare species are found only in highly diverse locations, while widespread species inhabit both diverse and less diverse areas. This triangular arrangement within a bipartite matrix suggests that more complex information tends to diffuse with greater difficulty and, as a result, is accessible only in a limited number of highly diverse locations. The nested structure of the international trade network remains stable over time, driven by two key biases: industries that exist despite breaking the nested pattern are more likely to vanish, while those absent but expected within the pattern are more likely to emerge.

This inherent structure of the trade network shapes how countries expand their productive capabilities. Rather than making random leaps, countries tend to develop products that are similar to those they already export — a process that reflects their embedded position in the Product Space. The implications for industrial policy purposes, according to Sousa and Mueller (2025), is that if a country's productive structure has little connection to a target industry (low density), any industrial policy directed toward that industry is unlikely to succeed. However, if a country with a challenging productive structure chooses to pursue products with low densities, it must be prepared to face significant risks and, consequently, build resilience against industrial policy setbacks.

A country's productive structure determines product densities, which in turn define future diversification opportunities and, ultimately, shape its productive structure. This self-reinforcing process creates a path-dependent pattern in economic development. In line with this process, a "Matthew effect" becomes evident, where more developed countries are better positioned than less developed ones to diversify into new products and accumulate productive capabilities. Countries with limited capabilities often lack the density needed to diversify into new products, leaving them constrained by path dependency. Thus, this literature introduces a new hypothesis about the economic development divergence among countries — a gap that could only close if all countries could reach any area of the Product Space (Hidalgo *et al.*, 2007). In this sense, Sousa and Mueller (2025) argue that industrial policy re-emerges as a viable strategy to counteract this Matthew effect in economic diversification and development. For countries unable to access desired areas of the product space, industrial policy may be justified, even when low densities and high risks are involved.

An accurate measure of relatedness is crucial, particularly for countries caught in the "dark side" of this Matthew Effect, where misguided industrial policy could further constrain their already limited fiscal space. However, the traditional relatedness density measure has significant limitations when it comes to predicting which products a country will begin exporting competitively in the near future. These limitations reduce its reliability and risk steering industrial policy in the wrong direction (Albora *et al.*, 2023; Tacchella *et al.*, 2023). One major issue is that traditional relatedness density correlates strongly with a country's overall level of diversification (Hidalgo, 2021), making it more of a global indicator of a country's connection to all products rather than a specific measure of its connection to individual products.

Another potential limitation is the presence of spurious correlations within the Product Space matrix. The traditional measure of relatedness density, calculated as a weighted average of proximities between a country's exported products and the target product, may be influenced by correlations that do not truly reflect shared productive requirements. Depending on how these proximities are distributed, such non-informative links may carry disproportionate weight, introducing noise into the density calculation. A natural solution to this issue lies in applying network filtering and backbone extraction techniques.

Although network backbone extraction and filtering methods have been introduced in the relatedness literature, these methods have not yet been systematically employed as a central approach to improve predictions of product appearances within the trade network. In their seminal work, Hidalgo *et al.* (2007) used a maximum spanning tree to extract the backbone of the product space, but solely for visualization purposes. Later, Alshamsi, Pinheiro, and Hidalgo (2018) incorporated this visualization-filtered product space to model product diffusion and define optimal diversification strategies, without having as a goal the enhancement of relatedness density's predictive accuracy for product appearances.

Several academic efforts have sought to enhance predictive performance within the international trade network by rethinking relatedness measures, with network backbone extraction and filtering typically playing only a secondary role. Zaccaria *et al.* (2014), for instance, proposed the Taxonomy Space, which adjusts the Product Space by using each country's diversification to normalize their co-location contributions in assessing pairwise product relationships. They apply a filtering procedure that retains only the highest proximity value for each product, which is then used to calculate relatedness density.

The Eco Space (O'Clery; Yıldırım; Hausmann, 2021), on the other hand, refines relatedness by identifying which products were active in the period preceding the activation of a specific product, using only the top 25 proximities for each product to calculate a country's relatedness density over a new product. The top-k filtering approach is largely responsible for the improvement in prediction performance of the Eco Space model when compared to the baseline relatedness density, though this detail is noted only in the study's Supplementary Information.

More recently, machine learning approaches (Albora *et al.*, 2023; Tacchella *et al.*, 2023) have demonstrated that algorithms such as Random Forest and XGBoost, when applied directly to the international trade network, offer even stronger predictive power for product appearances.. These models do not rely on explicit notions of relatedness or filtering, but their success underscores the importance of refining input representations for prediction tasks.

Interestingly, studies that do place network filtering at the core of relatedness modeling have mostly focused on multilayer network settings, aiming to uncover structural correlations across dimensions, not to boost predictive accuracy. Pugliese *et al.* (2019) explores the relationships among research, patents, and export products; Cunzo *et al.* (2022) investigates the links between green technologies and exported products; and Barbieri *et al.* (2023) analyzes interactions between green and non-green patent networks in European regions. While these studies contribute to a richer, multidimensional view of industrial ecosystems, they do not directly address the challenge of predicting product appearances.

In summary, while network filtering techniques are present within this literature, they have yet to be fully leveraged as a central tool for enhancing the predictive power of relatedness. In the next section, we outline a straightforward approach that applies filtering directly to the Product Space values, enabling a comparative assessment of various filtering methods based on their effectiveness in predicting product appearances.

3.3 Methods

Instead of introducing complex or disruptive new methods for measuring relatedness, we apply network backbone extraction and filtering techniques to retain only the most essential information within the Product Space. This approach adds a straightforward step to the calculation of relatedness density, potentially enhancing the baseline measure. By doing so, it addresses a key limitation of the baseline relatedness measure: the spurious correlations within the Product Space that are otherwise incorporated into the standard relatedness density calculation.

Following Hidalgo *et al.* (2007), we used international trade data to test the filtering approach for relatedness, drawing the HS92 4-digit dataset for 1995–2020 from the Observatory of Economic Complexity (OEC)[2]. To reduce statistical noise due to data quality issues, variations in economic size, and export disruptions from war or political instability, we applied several filters. Specifically, we excluded countries that, in any year of the period, had a population below 0.016% of the global population, an annual trade volume below 0.0067% of global trade, or a score of 26 or higher on the Fragile States Index[3], considering the dimensions "Security Apparatus," "Refugees and Internally Displaced Persons," and "External Intervention." We also used the OEC's product space baseline to determine the set of products analyzed in this study. After these steps, the HS92 dataset captures trade data for 866 products across 119 countries, representing 96.9% of global GDP and 94.6% of global trade in 2010.

We propose refining the Product Space by introducing an additional step between equations 3.3 and 3.4, where a backbone extraction or filtering method is applied to remove spurious correlations. This filtering process, based on the idea that some correlations may arise from random network properties, ensures that only the most meaningful proximities are preserved, enhancing the accuracy of the relatedness measure.

After applying the filtering method F, some previously non-zero proximities in the Product Space $\varphi^F_{p,p'}$ will be set to zero, resulting in a more accurate representation of product similarities:

$$\varphi^F_{p,p'} = F[\varphi_{p,p'}] \tag{3.5}$$

Finally, we plug the filtered Product Space into the baseline relatedness density 3.4, resulting in the filtered relatedness density:

$$\omega^F_{c,p} = \frac{\sum_c M_{cp}\varphi^F_{p,p'}}{\sum_{p'} \varphi^F_{p,p'}} \tag{3.6}$$

Figure 3.1 presents a schematic workflow of the proposed approach. The ABC flow is the baseline approach, while ABDE is the proposed one.

Depending on the filtering method applied to the Product Space in step D, it becomes possible to introduce directionality into the proximity matrix — that is, the proximity from product i to product j may differ from the proximity from j to i if one of them is zeroed out by the filtering process. This asymmetry is analytically valuable, as it may reflect a directed and non-reciprocal nature of capability accumulation. In particular, it allows the analysis to emphasize the perspective of the target product in identifying necessary capabilities, which resonates with the logic of productive ecosystems introduced by O'Clery, Yıldırım, and

2 http://oec.world/

3 https://fragilestatesindex.org/

Figure 3.1 – Network Filtering Schematic Workflow

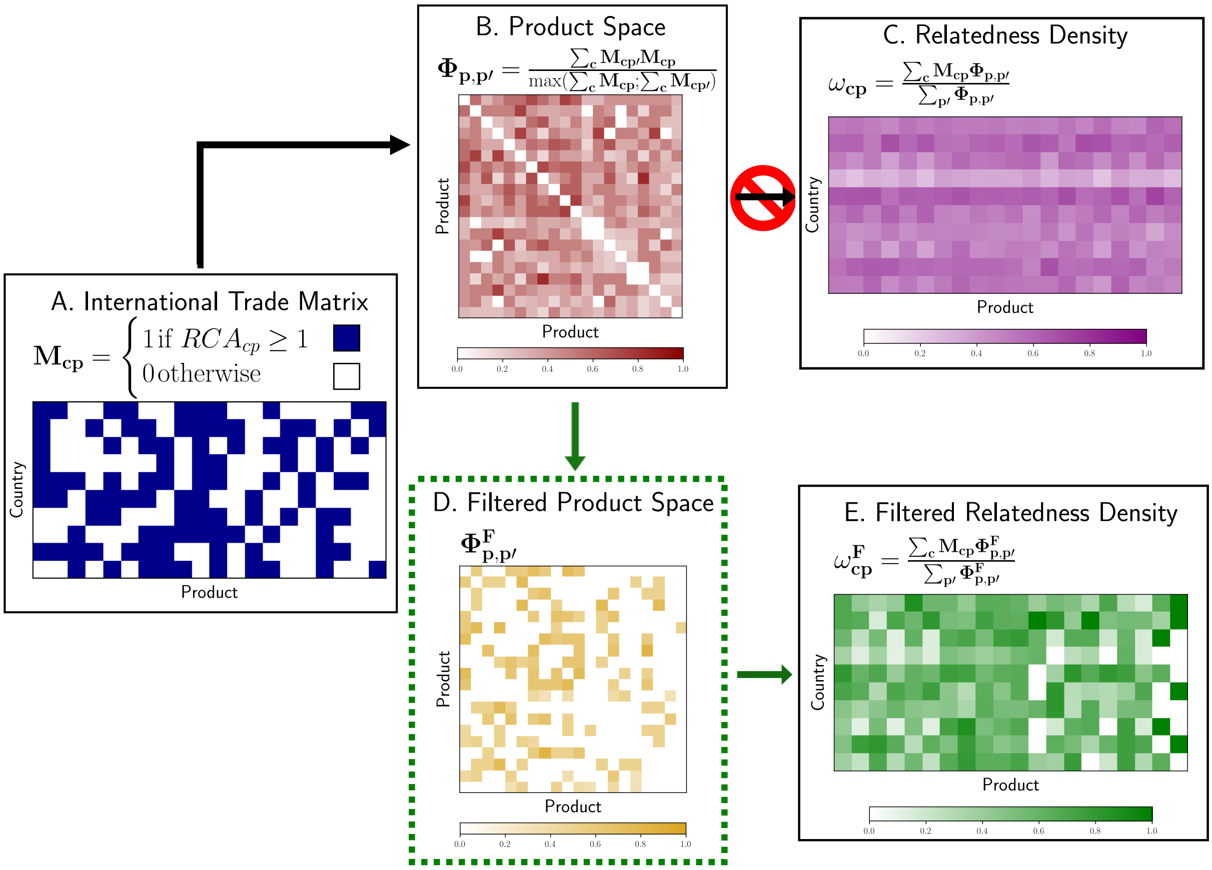


Figure 3.1 presents a schematic workflow illustrating the proposed network filtering approach to refine the relatedness density measure. The traditional method follows the sequence **ABC**, while our approach introduces a refined path through steps **ABDE**.

Hausmann (2021). This reinforces the idea that productive diversification is not reciprocal, and that meaningful upgrading relies on identifying signals that point specifically toward the conditions that enable the emergence of a given product.

There are two ways of applying the final filtered relatedness density formula 3.6. One adopts the perspective of the product as an outgoing signal emitter, and the other as an ingoing signal receiver:

$$\omega^{\text{In}}_{cp'} = \frac{\sum_p M_{cp} \cdot \varphi^{F(p')}_{pp'}}{\sum_p \varphi^{F(p')}_{pp'}} \qquad \omega^{\text{Out}}_{cp'} = \frac{\sum_p M_{cp} \cdot \varphi^{F(p)}_{pp'}}{\sum_p \varphi^{F(p)}_{pp'}} \tag{3.7}$$

The filtered density $\omega^{\text{In}}_{cp'}$ measures how much of the productive structure of country c aligns with the specific capability requirements of product p', based solely on the set of related products that p' itself recognizes as informative. This formulation captures a demand-side view of path-dependent upgrading, in which diversification into a new product is more likely if the country already exports products that constitute its required building blocks. These building blocks represent complementary bundles of productive capabilities. The denominator aggregates all proximities that p' considers relevant, defining the full reference space of inputs or technological affinities it relies upon.

In contrast, the filtered density $\omega_{cp'}^{\text{Out}}$ reflects a supply-side perspective. It evaluates whether the products already exported by country c perceive p' as a natural extension of their embedded capabilities. Here, the path-dependent logic flows from the existing portfolio of activities, capturing the extent to which the country's current productive base tends to emit strong capability signals toward p'. This dual perspective — assessing whether the product seeks the country or whether the country is positioned to reach the product — contributes to a more nuanced understanding of economic relatedness and the structural paths through which countries may diversify.

In this study, whenever directionality is established by the filtering method, we adopt the ingoing, signal-receiver, demand-side $\omega_{cp'}^{\text{In}}$ specification. This choice emphasizes the perspective of the target product and assesses how well a country's current export basket aligns with its capability requirements. The approach is particularly suited for investigating whether countries are structurally positioned to absorb new productive activities, based on what those activities demand, rather than on how the country's current production profile is inclined to evolve.

To assess how different filtering techniques affect predictive accuracy, we apply three classes of methods to compute $\varphi_{p,p'}^{F}$, the filtered product proximity measure: naive heuristics, statistical backbone extraction techniques, and their directional adaptations. Each of these groups reflects a different philosophy about how to reduce noise and extract meaningful signals from the Product Space.

The first class, based on naive heuristics such as top-k thresholds and cumulative proximity cutoffs, introduces directionality by construction, as proximities are filtered separately for each product. In this context, we adopt the ingoing specification $\omega_{cp'}^{\text{In}}$ to interpret the resulting asymmetric matrices from the perspective of the target product. The second class, comprising statistical backbone extraction methods, operates on undirected networks and produces symmetric proximity matrices by design. Despite lacking directional structure, these methods offer a valuable complement to naive filters by introducing statistical criteria to separate signal from noise. The third class combines the strengths of both approaches: it adapts statistical filtering procedures to allow for asymmetric evaluation of proximities, thereby preserving directionality while incorporating statistical rigor. This unified framework enables us to examine how directionality and statistical significance interact to shape the predictive power of relatedness density.

We refer to the first class as naive filters, not because they are simplistic, but because they rely on mechanical thresholds or rankings without any statistical modeling of the underlying distribution. These methods are straightforward, intuitive, and widely used as baseline benchmarks:

- **Top-k Filter:** For each product, retain all proximities corresponding to the top k values, including ties when applicable.
- **Cumulative Proximity Filter:** For each product, retain the highest proximities that together account for at least a fraction p of the total proximity mass.
- **Threshold Filter:** For each product, retain proximities greater than a threshold t. If no proximity exceeds t, retain the maximum value to ensure at least one connection.

The second class of methods for filtering the Product Space comprises statistical backbone extraction techniques, which aim to preserve only those proximities that are statistically significant relative to a node's local distribution or to a global null model. Unlike naive methods, these approaches attempt to distinguish signal from noise by testing each link against a reference distribution.

- **Disparity Filter** (Serrano; Boguná; Vespignani, 2009): Evaluates the significance of each edge by comparing its normalized weight to what would be expected if weights were uniformly distributed across a node's connections. An edge is retained if it is significant for at least one of the two connected nodes.
- **Locally Adaptive Network Sparsification (LANS)** (Foti; Hughes; Rockmore, 2011): Does not assume any theoretical distribution, but compares each edge to the empirical distribution of normalized weights from its node. An edge is preserved if it is unusually strong for at least one of its endpoints.
- **Noise-Corrected Filter (NC)** (Coscia; Neffke, 2017): Evaluates edge significance by estimating its expected weight under a binomial model, given the overall strength of both source and target nodes. A Bayesian framework is used to compute posterior variances and construct confidence intervals around the expected weight.

The statistical backbone extraction techniques discussed above were originally developed for undirected networks, such as the Product Space. In their standard implementation, these methods evaluate proximities symmetrically, as if the informational value of a connection between two products were equivalent in both directions. Even though the original network is undirected, each product has a unique distribution of proximities, and the presence of a connection does not guarantee that it carries meaningful information for both endpoints. In highly nested structures such as the international trade network, its product monopartite projection — the Product Space — has products that appear close to many others simply due to their ubiquity, leading to distorted measures of relatedness.

To address this, we develop directional adaptations of the Disparity Filter and LANS that evaluate proximity from the viewpoint of the receiving product. This logic is consistent with the $\omega^{\mathrm{In}}_{cp'}$ formulation introduced earlier, where the target product plays an active role in selecting which signals are relevant.

- **Directed Disparity Filter In-Degree (DDF-In-Degree):** Applies the Disparity Filter to the set of proximities received by each product. An edge from i to j is retained only if it is significant from the perspective of j, capturing the idea that i carries information relevant to j, but not necessarily the other way around.
- **Directed LANS In-Degree (D-LANS-In-Degree):** Similarly applies LANS only to the ingoing proximities of each product. It evaluates edge significance using the empirical distribution of proximities from each target.

It is important to note that the Noise-Corrected (NC) filter cannot be straightforwardly extended to directed networks. By design, the NC filter assesses the significance of each edge based on the combined strength of both nodes, treating the edge symmetrically. In a directed network, however, edges inherently differentiate a source from a target, which would require testing significance based only on one node's characteristics. Such an adaptation would break the core logic of the NC method, which fundamentally relies on bilateral node properties to model noise and compute z-scores.

In summary, we compare naive filtering, statistical backbone, and directional statistical backbone techniques to investigate how each one shapes the structure of the Product Space and affects the predictive power of relatedness density. All methods were implemented within a Python environment. For the original Disparity and LANS filters, we utilized the NetBone package[4] (Yassin *et al.*, 2023). For the Noise-Corrected filter, we used publicly available code[5] (Coscia; Neffke, 2017).

To evaluate the predictive performance of each filtering method, we use a Probit model to estimate the likelihood of product appearances at the country-product level, following O'Clery, Yıldırım, and Hausmann (2021). The full specification of the regression model is given by:

$$P(RCA_{cp,t+4} + 4 \geq 1 \mid RCA_{cp,t} \leq 0.5) = \beta_1 \omega F_{cp,t} + \mu_t \tag{3.8}$$

where $P(RCA_{cp,t+4} + 4 \geq 1 \mid RCA_{cp,t} \leq 0.5)$ is the probability of becoming a competitive exporter of that product in four years after year t, given that the country's RCA of that product was lower than or equal to 0.5 in t. The explanatory variables are, in order, filtered relatedness density and a dummy that controls for year fixed effects. We apply this model across rolling two-year windows, from 1998 to 2016.

For each filtering technique, we adjust its parameters to maximize the AUC-ROC score. We then compute a suite of performance metrics to compare methods, including

[4] http://gitlab.liris.cnrs.fr/coregraphie/netbone/

[5] https://www.michelecoscia.com/?page_id=287/

Precision, Best F1 Score, AUC-ROC, AUC-PR, and Pseudo-R^2, following Albora *et al.* (2023) and O'Clery, Yıldırım, and Hausmann (2021). These metrics are described below:

- **Precision (TP/(TP+FP))**: The proportion of true positives among all predicted positives. High precision indicates a low false-positive rate, which is crucial in contexts where incorrect recommendations carry significant costs.
- **Best F1 Score**: The maximum F1 value achieved across thresholds, where the F1 Score is the harmonic mean of precision (TP/(TP+FP)) and recall (TP/(TP+FN)). This metric captures the optimal trade-off between correctly identifying positives and minimizing false positives.
- **AUC-ROC (Area Under the Receiver Operating Characteristic Curve)**: Measures the model's ability to discriminate between classes by plotting the true positive rate, which is the recall (TP/(TP+FN)), against the false positive rate (FP/(FP+TN)). A value closer to 1 indicates stronger predictive discrimination.
- **AUC-PR (Area Under the Precision-Recall Curve)**: Especially informative in imbalanced classification settings, this metric reflects how well the model maintains both high precision and recall. A higher AUC-PR signals better identification of positive instances with fewer false alarms.
- **Pseudo $\mathbf{R^2}$**: A goodness-of-fit indicator for models with binary dependent variables. Although not directly comparable to the R^2 of linear regression, it provides a useful approximation of explanatory power in the Probit context.

3.4 Network Filtering Results

Table 3.1 presents the results for each filtering method, with parameters optimized to maximize the AUC-ROC score. All tested approaches outperformed the baseline relatedness density across every evaluation metric. The baseline achieved an AUC-ROC of 0.602. Among the naive filters, performance was relatively similar: global thresholding yielded an AUC-ROC of 0.658, while the top-k and cumulative proximity methods reached 0.666 and 0.667, respectively. Statistical backbone extraction methods performed slightly better, with LANS, Disparity, and Noise-Corrected filters achieving AUC-ROC scores of 0.673, 0.674, and 0.679. The best results were obtained by the directional statistical filters: DLANS-In-Degree and DDF-In-Degree achieved AUC-ROC values of 0.685 and 0.688, respectively.

These results highlight not only the predictive gains from applying statistical and directional filters, but also the value of incorporating asymmetry in proximity evaluation. While naive methods already improve upon the baseline by pruning weak or noisy signals, methods that combine statistical significance with directional filtering achieve the highest predictive accuracy. This suggests that the process of productive upgrading is better captured when proximities are interpreted from the perspective of the target product, reinforcing the

Table 3.1 – Comparative Metrics for Filtering Techniques

Method	Parameter Value	F1 Score	AUC-ROC	AUC-PR	Precision	Pseudo R^2
Baseline	-	.0366	.602	.0196	.0202	.0083
Global Thresholding	t = 0.32	.0566	.658	.0278	.0347	.0256
Top-k	k = 31	.0603	.666	.0300	.0370	.0281
Cumulative Filtering	p = 0.11	.0603	.667	.0303	.0369	.0286
LANS	α = 0.03	.0634	.673	.0311	.0384	.0318
Disparity	α = 0.14	.0643	.674	.0308	.0386	.0327
Noise-Corrected	z = 0.36	.0683	.679	.0320	.0405	.0364
DLANS-In-Degree	α = 0.06	.0722	.685	.0340	.0492	.0395
DDF-In-Degree	α = 0.17	.0749	.688	.0345	.0498	.0415

Table 3.1 presents comparative metrics for tested network backbone extraction and filtering techniques. The table reports goodness-of-fit measures for predicting product appearances in a 4-year window, using intervals of two years between each t from 1998 to 2016, totaling 773,974 observations.

theoretical argument for directional, demand-side relatedness.

Beyond improvements in AUC-ROC, the more advanced filtering techniques also deliver superior performance across all other evaluation metrics. The baseline F1 Score is notably low at 0.0366, reflecting a poor balance between precision and recall. While naive methods provide moderate improvements, the directional statistical filters, DLANS-In-Degree and DDF-In-Degree, raise the F1 Score to 0.0722 and 0.0749, more than doubling the baseline.

This trend continues in the AUC-PR metric, which is particularly informative for imbalanced classification problems. The baseline AUC-PR is just 0.0196, whereas the best-performing methods reach values above 0.034, indicating a stronger ability to correctly identify rare positive cases. Precision also increases consistently with the sophistication of the filtering strategy: from 0.0202 in the baseline to nearly 0.050 in the DDF-In-Degree model. Higher precision implies that a greater share of predicted positive cases actually correspond to true positives, reducing the number of false positives. This improvement is particularly important for industrial policy design, as it helps policymakers prioritize targets with a higher likelihood of successful upgrading, minimizing the risk of misallocating resources to unpromising opportunities. Likewise, Pseudo-R^2 improves from 0.0083 to 0.0415, underscoring a substantial gain in model fit.

Albora *et al.* (2023) compare the performance of relatedness-based predictors using three approaches: a baseline relatedness density, Eco Space (O'Clery; Yıldırım; Hausmann, 2021), and a Random Forest model. The latter represents a machine learning technique that automatically selects and weighs predictive features through an ensemble of decision trees. In their analysis, the baseline density measure achieves an AUC-ROC of 0.637 and an F1

Score of 0.022. Random Forest performs best, with an AUC-ROC of 0.689 and an F1 Score of 0.042, followed by Eco Space, which yields 0.663 and 0.035, respectively. While differences in modeling strategy and sample construction prevent a direct comparison with our study, it is notable that the DDF-In-Degree method achieves an AUC-ROC nearly equivalent to that of Random Forest. Moreover, the performance gain over the baseline is even larger in our framework: DDF-In-Degree increases AUC-ROC by 0.086, whereas Random Forest improves it by only 0.052 in the other study.

Taken together, these results suggest that filtering techniques that incorporate both statistical rigor and directional logic provide a more informative and reliable foundation for predicting productive upgrading and modeling economic diversification paths. The benefits extend beyond improvements in AUC-ROC, encompassing gains across all goodness-of-fit measures, while also offering straightforward interpretability.

Table 3.2 presents the correlation matrix for the year 2010, including diversity, baseline relatedness density, and all optimized filtering methods. A well-documented limitation of the baseline measure is its extremely high correlation with diversity, which, in our analysis, reaches a coefficient of 0.95. When relatedness density is so tightly linked to diversity, the metric ceases to capture a product-specific signal and instead becomes a near-linear proxy for a country's overall productive breadth. In such cases, measuring how many capabilities a country possesses that are relevant to a specific product adds little new information beyond what diversity already conveys. Relative relatedness density (Pinheiro *et al.*, 2022), a form of Z-score normalization at the country level, can partially mitigate this issue by centering measurements within each country. However, while this adjustment reduces the collinearity with diversity, it also removes the broader, cross-country context that baseline density offers, limiting its policy relevance.

In contrast, the filtering methods reduce this entanglement to varying degrees. The correlation between diversity and filtered relatedness is inversely associated with model performance. The DDF-In-Degree method, which achieves the highest AUC-ROC, also exhibits the lowest correlation with diversity — 0.51. This suggests that directional statistical filters are more effective at isolating product-specific signals from aggregate country-level effects.

However, the table also shows that the filtered relatedness densities remain highly correlated with one another, with pairwise correlations never dropping below 0.91. This indicates that, despite methodological differences, all filters still capture a largely overlapping structure of the Product Space. These findings reinforce the importance of moving beyond diversity-correlated baselines while acknowledging that filtering techniques often converge toward a common signal space, which may benefit from further refinement or complementary layers of differentiation.

In the next section, we explore how the application of the DDF-In-Degree filter, the

Table 3.2 – Correlation Table - Year 2010

	Diversity	Baseline	Global Thres.	Top-k	Cum. Filtering	LANS	Disparity	Noise-Corrected	DLANS-In-Degree	DDF-In-Degree
Diversity		0.95	0.68	0.62	0.62	0.60	0.59	0.58	0.55	0.51
Baseline			0.85	0.80	0.81	0.79	0.78	0.77	0.74	0.72
Global Thres.				0.94	0.95	0.95	0.93	0.91	0.92	0.91
Top-k					0.99	0.98	0.96	0.94	0.94	0.93
Cum. Filtering						0.98	0.96	0.94	0.94	0.93
LANS							0.97	0.95	0.96	0.95
Disparity								0.98	0.95	0.96
Noise-Corrected									0.94	0.96
DLANS-In-Degree										0.98
DDF-In-Degree										

Table 3.2 presents the correlation matrix between different filtered relatedness densities for the 2010 data. Number of observations: 105,652.

top-performing method identified in our analysis, reshapes the structure of the Product Space. We compare the resulting filtered network to the original, unfiltered one, assessing the extent to which this backbone extraction method preserves core connectivity while eliminating noisy or redundant links.

3.5 From Noise to Signal: Revisiting the Product Space

Building on the results from the filtering analysis, we now revisit the Product Space through a refined lens. Applying the DDF-In-Degree filter, identified as the top-performing method, we extract the core signals of the network, removing much of the surrounding noise. For this purpose, we focus on the Product Space as it stood in 2010, using this year as a baseline to compare the structure of the original, unfiltered network with the backbone revealed through the filtering process. This comparison allows us to assess how network filtering reshapes the connectivity patterns that underpin opportunities for economic upgrading.

Figure 3.2 presents the network visualization of the DDF-In-Degree filtered Product Space based on year 2010 data. Node colors represent product sections according to the HS4-92 classification, while node sizes are proportional to the total weighted in-degree of each product. The network layout was constructed by first running a maximum spanning tree on the filtered Product Space, then adding all edges with values higher than 0.55, following the methodology originally proposed by Hidalgo *et al.* (2007).

Figure 3.2 – Filtered Product-Space - Year 2010

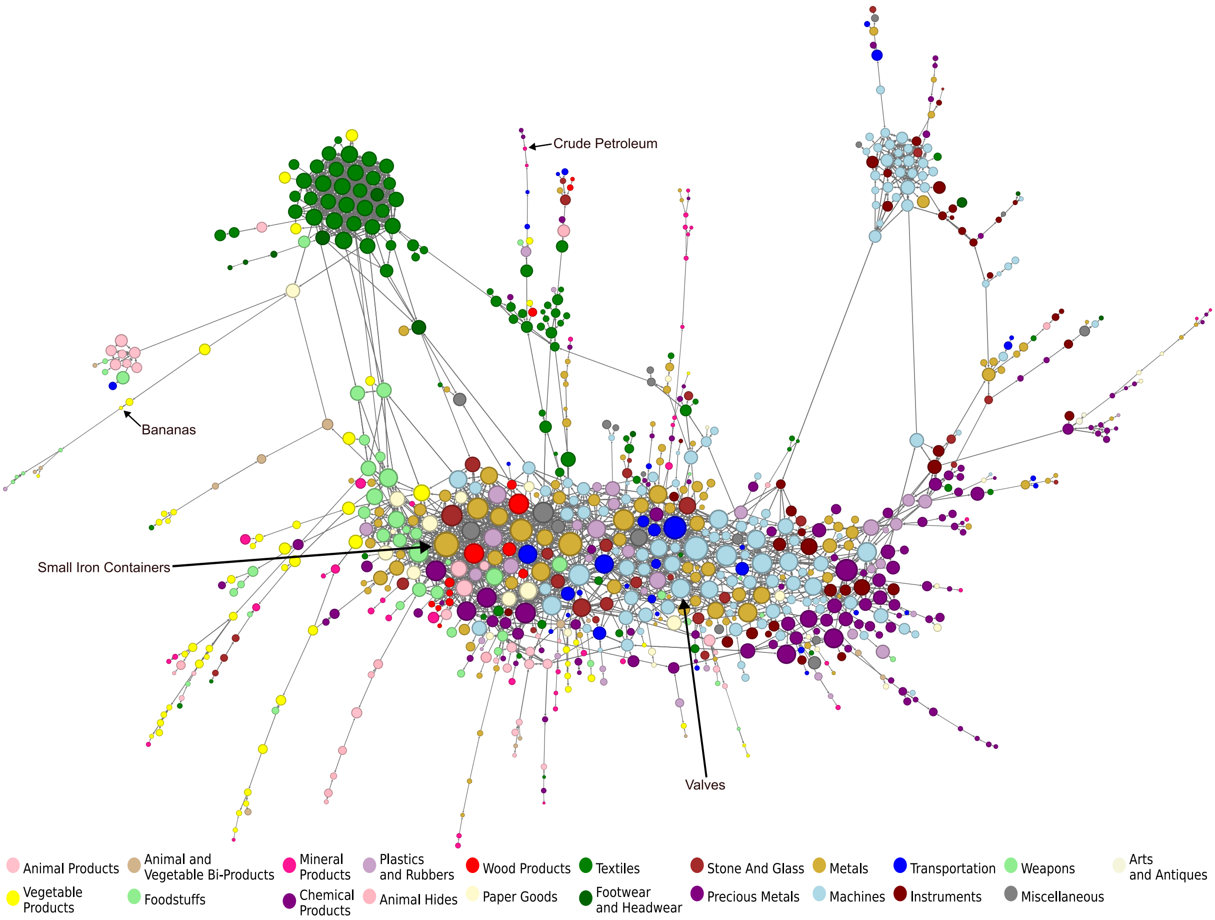


Figure 3.2 presents the network visualization of the DDF-In-Degree filtered Product Space based on year 2010 data. Node colors represent product sections according to the HS4-92 classification, while node sizes are proportional to the total weighted in-degree of each product. The network layout was constructed by first running a maximum spanning tree on the filtered Product Space, then adding all edges with values higher than 0.55.

The filtered Product Space highlights the structural position and connectivity patterns of different product sections, with some emerging as particularly cohesive or peripheral clusters. Textiles products (in dark green) form a dense and cohesive community in the upper-left region, clearly separated from the rest of the network, suggesting strong internal relatedness but limited outward connectivity. Another notable cluster in the filtered Product Space corresponds to what can be described as Electronic and Precision Devices, located in the upper-right periphery of the network. This group blends products from the Machines section, such as Semiconductors, Electric Batteries, and Printed Circuit Boards, with items from Instruments, including Oscilloscopes, Measuring Instruments, and Broadcasting Accessories. Despite its partial detachment from the dense industrial core, this cluster forms a cohesive subnetwork characterized by high internal connectivity and functional relatedness.

Its somewhat peripheral position suggests that, while these products are technologically advanced and economically significant, their productive capabilities are less intertwined with the broader set of industrial goods in the product space.

In contrast, Machines, Metals and Chemical Products (in light blue, mustard and purple, respectively) dominate the central core of the network, with several large nodes indicating products with high weighted in-degree, such as Valves and Small Iron Containers. These sections appear as structural backbones of the network, linking a wide variety of other product categories. Meanwhile, Crude Petroleum, a globally significant export from the Mineral Products section, appears in a highly peripheral position, weakly connected to the rest of the network, illustrating how some economically important commodities may lack dense interconnections within the filtered product space. Similarly, Bananas from the Vegetable Products section occupy a sparsely connected branch, highlighting their limited overlap with the core industrial structure.

The positioning and size of nodes of the filtered layout of the Product Space underscore how filtering with DDF-In-Degree preserves core productive capabilities and interconnectivity. Remarkably, the network remains fully connected after filtering, with the largest component still comprising all 866 products, which is the node size of the original unfiltered Product Space. At the same time, the filtering process significantly reduces the overall density of the network. Out of 717,336 non-zero proximity values in the original matrix, only 55,557 remain after applying the filter, representing a reduction of approximately 92.3% in the number of connections. This highlights the method's ability to retain the structural backbone of the network while eliminating the vast majority of weaker or potentially spurious links.

Figure 3.3 compares the degree and weighted degree distributions of the original and filtered Product Space for the year 2010. As shown in panel A, the baseline network is extremely dense, with a mean degree of 828.3, indicating that each product is, on average, connected to nearly all others. After applying the DDF-In-Degree filter, this number drops substantially to a mean in-degree of 64.2 (the same 92.3% reduction). A similar pattern is observed in panel B, where the average weighted degree decreases from 168.4 in the baseline to just 26.4 in the filtered network, corresponding to a reduction of approximately 84.3%. These shifts reflect the substantial sparsification achieved by the filtering process. The correlation between degrees in the filtered and unfiltered versions remains weak, with Pearson and Spearman correlation coefficients of .19 and .05, respectively, for degree, and values of .12 and .07 for the weighted degree. This suggests that the filter does not simply downscale the original structure, but also redefines the network's topology by retaining only the most statistically significant and structurally meaningful connections.

Figure 3.4 explores the relationship between weighted in-degree and weighted out-degree in the filtered Product Space for 2010. This comparison becomes especially meaningful in the context of our directional filtering approach, as it allows us to detect asymmetries in

Figure 3.3 – Baseline vs Filtered Product Space - Degree Distribution Histogram - Year 2010

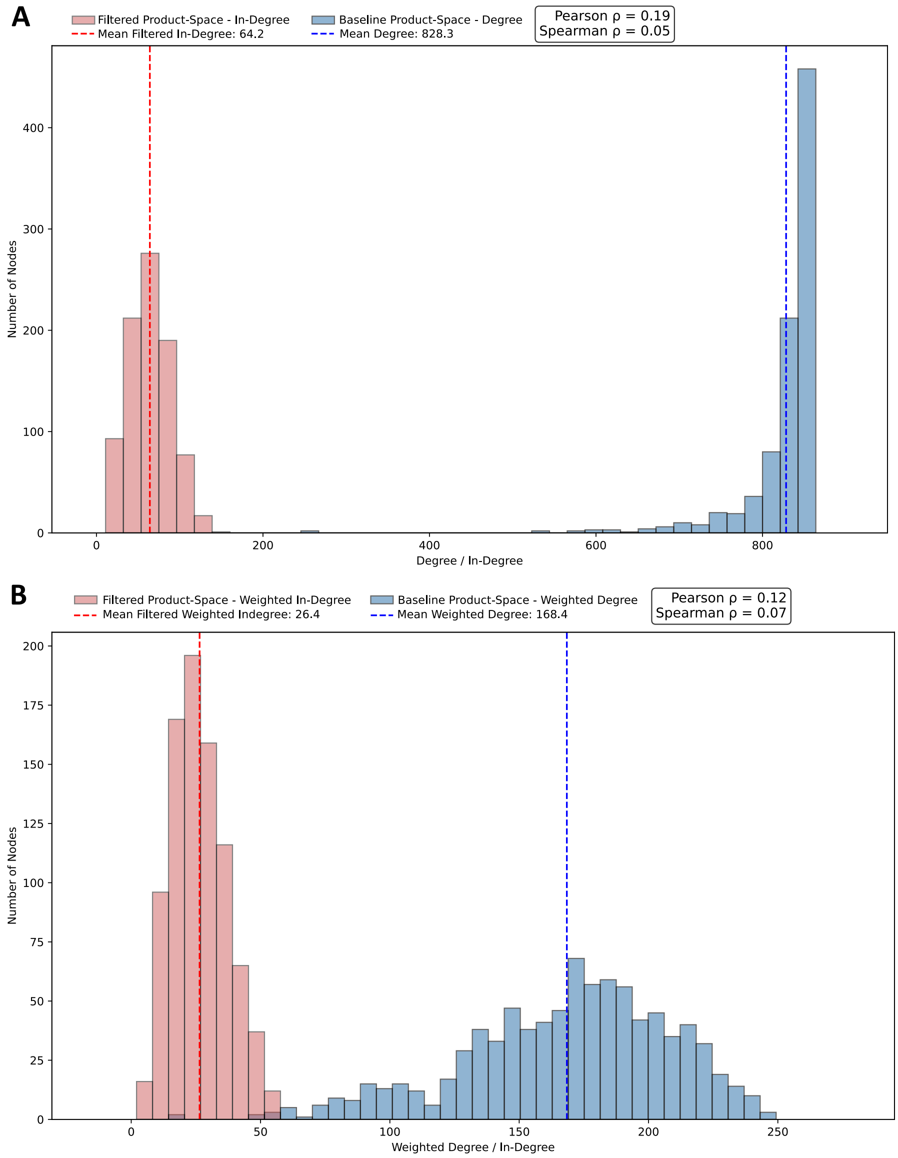
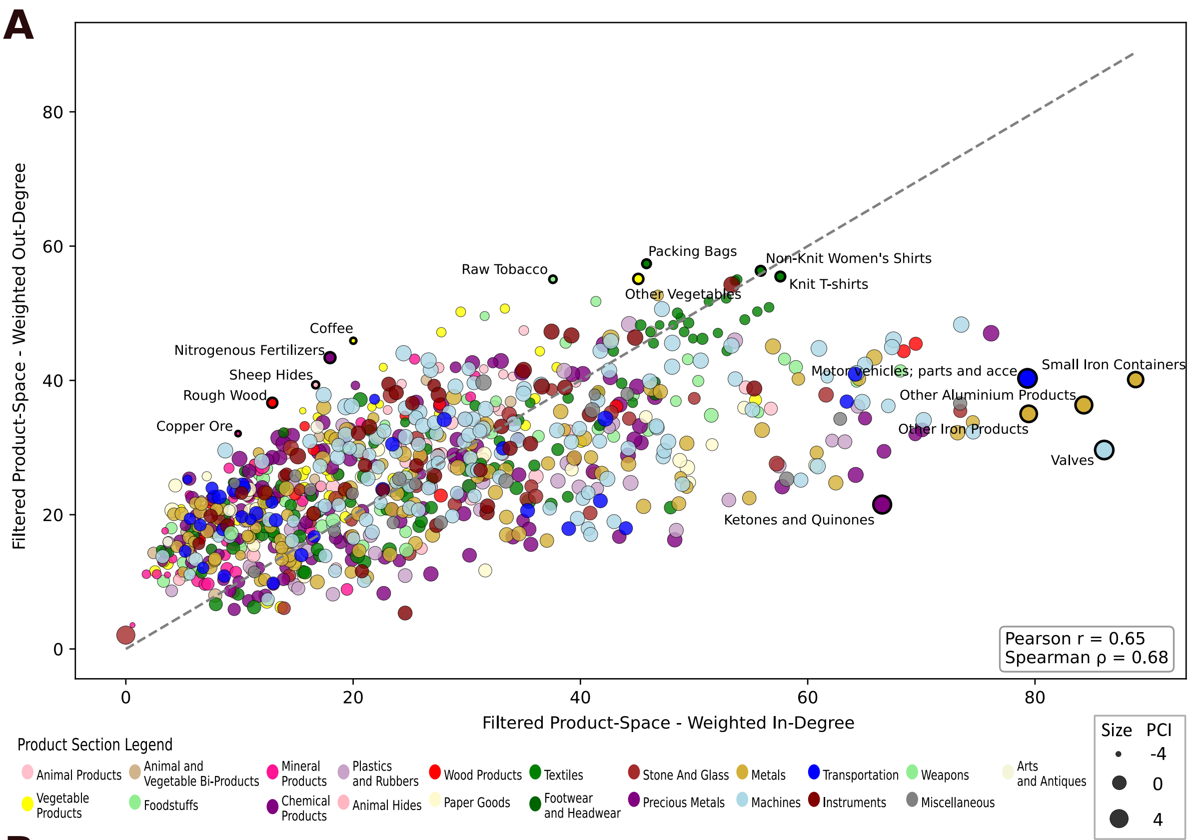


Figure 3.3 shows the degree **(Panel A)** and weighted degree **(Panel B)** distributions of the baseline and filtered Product Space in 2010. Filtering with the DDF-In-Degree method reduces average degree from 828.3 to 64.2 and weighted degree from 168.4 to 26.4, while preserving full network connectivity. Vertical dashed lines indicate the mean of each distribution.

the role each product plays within the network, depending on whether it functions more as a signal emitter (high out-degree) or as a signal receiver (high in-degree). Panel A shows a strong positive correlation between the two measures (Pearson = 0.65, Spearman = 0.68), but also reveals substantial variation along the diagonal, suggesting that many products tend to concentrate more heavily on one directional role than the other.

Figure 3.4 – Weighted In-Degree vs Weighted Out-Degree - Filtered Product Space - Year 2010



B

Top 10 Products by Weighted Degree

Top Out-Degree (1–5)			Top Out-Degree (6–10)			Top In-Degree (1–5)			Top In-Degree (6–10)		
Product	PCI	Value	Product	PCI	Value	Product	PCI	Value	Product	PCI	Value
Packing Bags	-1.85	57.380	Knit Men's Coats	-1.60	55.020	Small Iron Containers	1.19	88.857	Lubricating Products	4.13	76.138
Non-Knit Women's Shirts	-1.47	56.320	Non-Knit Men's Undergarments	-1.90	54.841	Valves	3.28	86.094	Traffic Signals	2.84	74.571
Knit T-shirts	-1.47	55.473	Chemical Analysis Instruments	4.23	54.252	Other Aluminium Products	2.21	84.298	Iron Structures	1.26	74.529
Other Vegetables	-1.29	55.115	Raw Iron Bars	-0.81	52.653	Other Iron Products	1.99	79.446	Transmissions	3.84	73.509
Raw Tobacco	-2.24	55.066	Knit Men's Undergarments	-1.50	52.240	Motor vehicles; parts and acce	3.36	79.341	Cement Articles	1.36	73.423

Figure 3.4 compares weighted in-degree and weighted out-degree in the filtered Product Space for 2010. **A** plots the relationship between the two measures, with node colors representing product sections according to the HS4-92 classification and node sizes proportional to product complexity (PCI). **B** lists the top ten products ranked by weighted in-degree and by weighted out-degree.

Several products from the textile cluster, such as Packing Bags, Non-Knit Women's

Shirts, Knit T-shirts, and Knit Men's Coats, have high weighted out-degree but relatively low in-degree. Although these products emit a large volume of relatedness signals, the visualization reveals that most of these connections remain confined within their own textile cluster. This suggests that their productive capabilities are highly interconnected among themselves, but less integrated into the broader network. In contrast, industrial goods like Valves, Small Iron Containers, and Other Aluminum Products exhibit high in-degree, indicating that they function as convergence points for signals of relatedness coming from a wider range of product sections. Panel B complements this analysis by listing the top ten products according to each measure, reinforcing the contrast between outward-oriented but locally clustered textile products and inward-oriented industrial components embedded across the network. This asymmetry illustrates how the directional filtering method offers a more nuanced interpretation of each product's structural position and its potential strategic relevance.

The relationship between weighted in-degree and out-degree in the filtered Product Space also reveals meaningful patterns when considered alongside product complexity. Products located in the upper-left region of the plot, which exhibit relatively high weighted out-degree compared to their in-degree, are concentrated in lower-complexity sections such as Textiles, Animal Products, and other primary goods. These products tend to emit many relatedness signals but are not widely targeted by others, suggesting that their productive capabilities are relevant within narrowly defined clusters but have limited structural centrality in the broader network. In contrast, products in the lower-right region, with relatively high weighted in-degree compared to their out-degree, are more frequently associated with Machines, Metals, and Chemical Products, which typically display higher levels of complexity. These products attract relatedness signals from a wide range of other goods, indicating that they occupy more central positions in the Product Space and require more sophisticated and diverse capabilities. This pattern suggests that the directionality of connections is not only structurally informative but also aligned with the underlying technological and organizational complexity of products.

The directional analysis of the filtered Product Space reveals not only a clearer structural organization but also new dimensions of asymmetry and complexity that are obscured in the unfiltered network. By disentangling signal from noise, the DDF-In-Degree filtering process exposes the central pathways through which productive capabilities are related and transferred across products. These insights offer a more refined map of diversification opportunities and potential bottlenecks. In the next section, we build on these results to explore their implications for industrial policy design, examining how the structure of the filtered Product Space can inform more strategic and evidence-based targeting decisions.

3.6 Policy Implications

The filtering of the Product Space carries important implications for industrial policy design. At a general level, it reveals that the path toward economic upgrading may be less daunting than what the unfiltered relatedness density metric suggests, with many opportunities becoming more visible once noisy or redundant connections are removed. By exposing a clearer structure of relatedness, the filtering process offers a more optimistic and actionable map for diversification strategies. To further illustrate these practical implications, the following analysis examines two country cases, Brazil and the Philippines, comparing the recommendations derived from the baseline and the filtered versions of relatedness density. This approach highlights how backbone extraction can reshape targeting priorities and uncover overlooked opportunities for productive transformation.

Figure 3.5 provides a comparative overview of baseline and filtered density metrics across five ECI groups in the year 2020. For each country, the analysis focuses on its top-10 density products with RCA ≤ 0.5, as these represent the most immediate and potentially viable opportunities for diversification. The comparison considers two key dimensions: the average density of these products and the average difference between their PCI and the country's ECI. These indicators are evaluated for both the original Product Space and the version filtered through the DDF-In-Degree method, along with their respective variations across complexity groups. In addition to the boxplot representation, each panel includes a red curve depicting the fitted quadratic regression estimated across all data points, accompanied by the corresponding equation and R^2 statistic. This figure serves as a starting point for understanding how filtering reshapes the perceived structure of diversification opportunities at different levels of economic complexity.

The first insight highlights the emergence of two distinct development traps for countries at low and middle levels of economic complexity, as revealed by the unfiltered density metric. In the baseline configuration, countries across almost all ECI ranges, except those at the very high end, tend to face negative average (PCI – ECI) values among their most proximate products, with sharp variations in available density across ECI groups. Middle-ECI countries, in particular, exhibit higher density than low-ECI countries, yet the structure of their nearby opportunities is more detrimental to further complexity upgrading. While low-ECI countries suffer from a low-complexity trap, characterized by scarce and unsophisticated opportunities, middle-ECI countries encounter a middle-complexity trap, where the relative abundance of accessible products masks a deeper structural constraint: most nearby options would deeply lower their economic complexity. The combination of low mean (PCI – ECI) values and the configuration of baseline density thus reveals distinct developmental challenges that permeate countries of both low and middle levels of economic complexity.

Countries at middle levels of complexity may become structurally constrained to target less sophisticated products, reinforcing a self-perpetuating cycle of limited upgrading. This

Figure 3.5 – Baseline vs. DDF-In-Degree Density: Boxplot Results by ECI Group - Year 2020

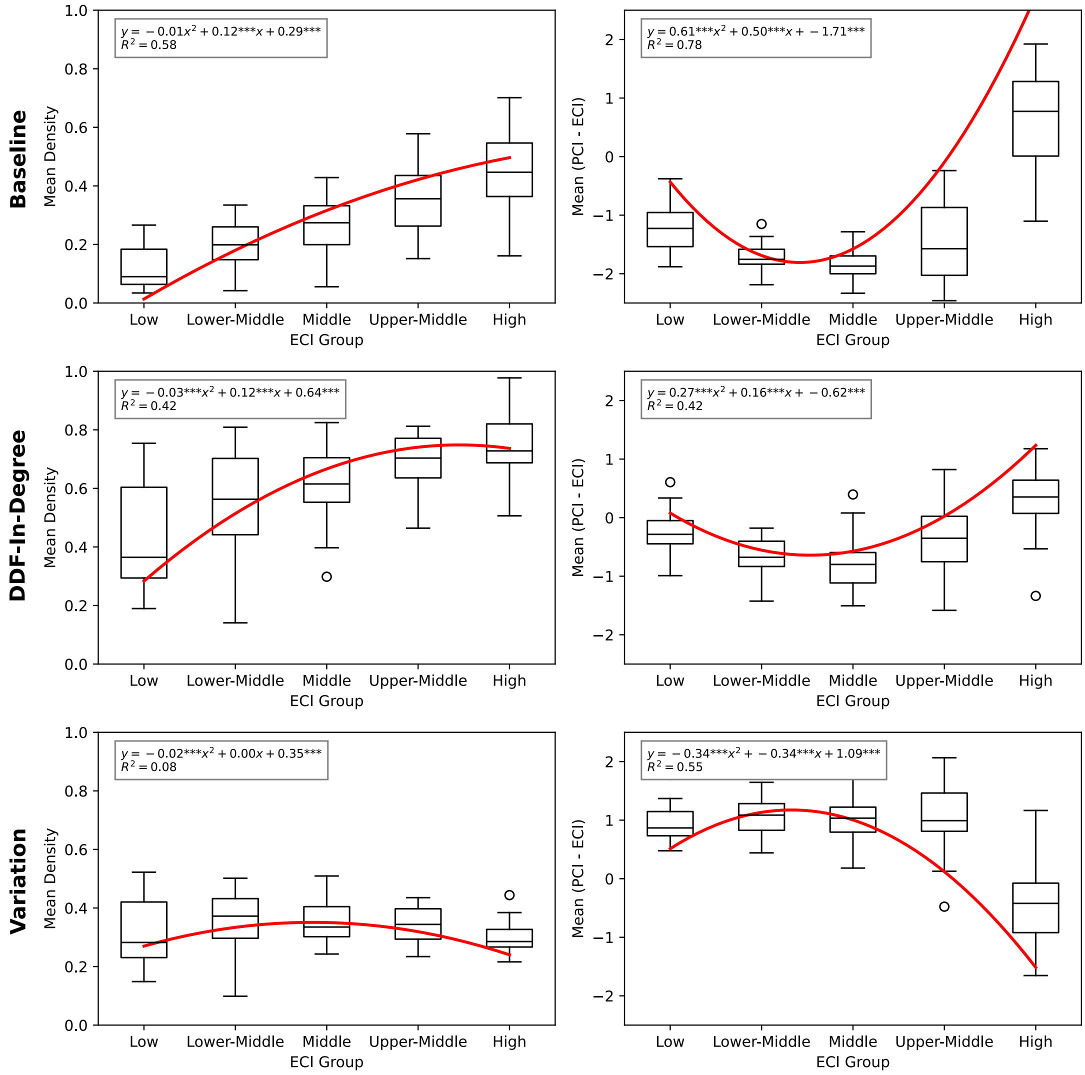


Figure 3.5 presents boxplots comparing, across five ECI groups, the top-10 density products of each country with RCA $\leq$ 0.5 in the year 2020. The analysis considers two dimensions — mean density and mean difference between product PCI and country ECI — for both baseline and DDF-In-Degree methods, as well as their respective variations across ECI groups. In addition to the boxplot representation, each subplot includes a red curve depicting the fitted quadratic regression trend estimated over all country data points. The corresponding regression equation and the R^2 statistic are displayed within each panel. The ECI ranges for the groups are defined as follows: Low: -1.9084 to -0.8720; Lower-Middle: -0.8175 to -0.4403; Middle: -0.4229 to 0.0835; Upper-Middle: 0.1409 to 0.9768; High: 0.9866 to 2.3298.

connects with findings in the economic complexity literature on when countries should target nearby or distant products (Hidalgo, 2023), which often argues that economies in intermediate complexity ranges may need to leap toward more distant opportunities to avoid being trapped in a complexity stasis (Pinheiro *et al.*, 2022; Alshamsi; Pinheiro; Hidalgo, 2018). The strong U-shaped relationship observed between ECI and mean (PCI–ECI), with an R^2 of 0.78, further indicates that this constraint is structural. As a result, pursuing complexity-enhancing diversification would require betting on distant, less accessible products, a strategy often marked by higher risks and lower probabilities of success. For a deeper discussion on related and unrelated diversification, please see Appendix 3.3.

The filtered approach offers a less pessimistic view of the diversification prospects for countries with intermediate levels of economic complexity. First, the average density of the top-10 products for these countries, when measured after applying the DDF-In-Degree filtering, is much closer to that of countries with higher ECI levels. This suggests that, once spurious or redundant links are removed, the proximity landscape faced by middle-complexity economies becomes significantly more favorable. Rather than being overwhelmingly surrounded by marginal or unattainable opportunities, these countries appear to have a set of relatively accessible paths for upgrading, comparable in density magnitude to those available to more advanced economies. Second, although the relationship between ECI and the mean difference between product PCI and country ECI still displays a U-shaped pattern, this curve becomes significantly flatter after filtering. While middle-complexity countries continue to face a slight bias toward less sophisticated products, the severity of this constraint is much reduced compared to the baseline.

Finally, the analysis of variations further reveals that all ECI groups experience gains in filtered density relative to the unfiltered baseline, smoothing the initially strong correlation between density and ECI levels. Moreover, countries at lower and intermediate ECI levels display particularly pronounced improvements in the mean PCI–ECI difference among their most proximate products. This suggests that filtering not only enhances the density landscape but also opens more realistic and complexity-enhancing pathways for countries that were previously seen as structurally disadvantaged.

While the previous analysis focused on the characteristics of the top-10 most proximate products, a comprehensive evaluation of diversification prospects also requires examining the predicted probabilities of product activation. Beyond mere density, the probability of successfully exporting a product depends on how strongly the density signal relates with future activation, as estimated through the Probit models. Specifically, we use the coefficients obtained from the Probit regressions, separately estimated for baseline density and filtered density, to predict the probability of activation for each country-product pair. These regressions control for year fixed effects, thereby accounting for time-specific shocks and global trends that could otherwise bias the estimated relationship between density and activation likelihood. However, when generating the predicted probabilities for the year 2020, which lies outside the estimation sample, we apply only the estimated density coefficients without adjusting for a year-specific intercept, under the assumption that the average propensity for product activation remained stable.

Figure 3.6 presents the results of this exercise for the year 2020. In Panel A, we plot the theoretical predicted probabilities as a function of density, applying the respective Probit coefficients for the baseline and filtered measures. The results reveal that, for density values above approximately 0.18, the predicted probability of activation becomes consistently higher under the filtered specification, with the gap widening at higher density levels. Panels B and C depict the relative frequency histograms of the predicted probabilities across all country-product pairs with RCA ≤ 0.5, using baseline density and filtered density, respectively.

Under the baseline configuration (Panel B), the distribution of predicted probabilities is relatively flat and continuous, with low predicted probabilities even at the right tail (95th percentile). In contrast, the filtered configuration (Panel C) exhibits a highly skewed distribution, where most observations concentrate at low probability levels, but with a substantially higher right tail. This pattern emerges as a direct consequence of the proper filtering of proximities, which refines and reduces the set of products generally required for the production of a given target product. From a policy perspective, this distinction is crucial: although opportunities with high predicted probabilities remain relatively rare, their magnitude under the filtered approach is considerably higher, improving the identification of products that offer realistic prospects for complexity-enhancing diversification.

To illustrate the practical implications of filtering the Product Space, we now turn to a more micro-level analysis through country-specific case studies. Specifically, we apply an adapted version of the Diversification Frontier, also known as the Relatedness-Complexity Diagram, to assess how policy recommendations change when filtered proximity measures are used.

The original Diversification Frontier (Balland *et al.*, 2019; Hidalgo, 2023) plots, for each product that a country does not yet export with revealed comparative advantage, two dimensions: on the horizontal axis, the density, which captures how closely the country's current capabilities are connected to that product; and on the vertical axis, the PCI of the product itself. In this framework, products with higher density values are considered more accessible for future competitive upgrading, while products with higher PCI values contribute more significantly to the country's overall complexity. This tool thus aims to help policymakers prioritize targets for industrial policy based on a balance between feasibility and impact.

In our adapted version, we replace baseline density with the predicted probability of product activation, as estimated through filtered proximity measures. This adjustment allows for a more precise evaluation of future diversification paths, correcting for the noise

Figure 3.6 – Baseline vs DDF-In-Degree - Predicted Probabilities and Histogram - Year 2020

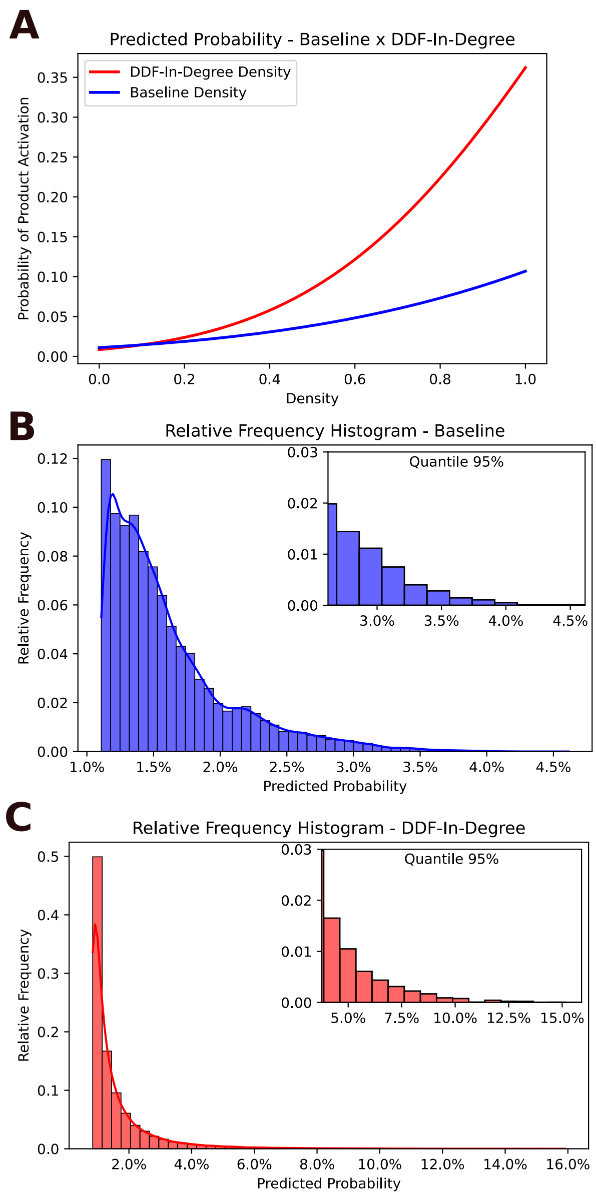


Figure 3.6. **A** shows the predicted probabilities of product activation within four years, calculated from the coefficients of Probit regressions, based on varying theoretical baseline and disparity densities, from nil to one. **B** and **C** present relative frequency histograms of the implied probabilities for all country-product pairs with RCA <= 0.5 in 2020, using the baseline and disparity methods, respectively. The subplots represent the right tail of each distribution, corresponding to the 95% quantile.

and redundancy inherent in the traditional Product Space. By applying this methodology to Brazil and the Philippines, we aim to show how filtering reshapes the identification of strategic products and ultimately leads to different industrial policy recommendations.

Figure 3.7 illustrates the impact of filtering the Product Space on industrial policy recommendations for Brazil and the Philippines, focusing on all products for which each country had an RCA ≤ 0.5 in 2020. Panel A shows the predicted probabilities of achieving RCA ≥ 1 within four years for each product, comparing results based on baseline density and DDF-In-Degree density. Panels B and C present two versions of the Diversification Frontier for each country: Panel B plots baseline-based predicted probabilities against PCI values, while Panel C plots DDF-In-Degree filtered predicted probabilities against PCI values.

Panel A of Figure 3.7 reveals substantial shifts between the baseline and disparity-filtered approaches for both Brazil and the Philippines in terms of predicted probabilities of product activation. The most immediate contrast is the difference in scale: while the maximum predicted probability based on baseline density remains slightly below 2% for both countries, it approaches 12% under the filtered approach. This indicates that, once spurious proximities are corrected, certain products are considered far more realistically accessible than initially suggested by the baseline configuration.

The correlation metrics provide additional context to these findings. For Brazil, the Pearson and Spearman correlations between baseline and filtered probabilities are 0.72 and 0.82, respectively. For the Philippines, they reach 0.76 and 0.77. Although these values indicate a strong positive association, they also suggest that filtering introduces important changes to the prioritization landscape. While the relative ranking of products is broadly preserved, the adjustments made by filtering are substantial enough to reshape strategic decisions regarding diversification opportunities.

The relative rankings of some products shift considerably. For Brazil, wool maintains a predicted probability of approximately 1.9% across both approaches but falls from a top-ranked position under the baseline to a mid-range position under the filtered measure. Sorghum remains highly ranked across both specifications, but its predicted probability jumps from around 1.9% to nearly 12%, reflecting a substantial improvement in perceived accessibility. In the case of the Philippines, products such as flat panel display modules and watch movements emerge as particularly promising under the filtered measure, with watch movements attaining a probability close to 10% compared to much lower values under the baseline. These examples highlight how filtering not only amplifies success probabilities but also reshapes the prioritization among potential diversification targets.

Turning to the Diversification Frontiers in Panels B and C, Brazil's prospects show specific adjustments. Under the baseline configuration shown in Panel B, sorghum and wool appear similarly positioned, both in terms of probability and complexity. However, using filtered probabilities in Panel C, sorghum clearly outperforms wool. Despite this

Figure 3.7 – Diversification Frontiers - Brazil and Phillipines - Year 2020

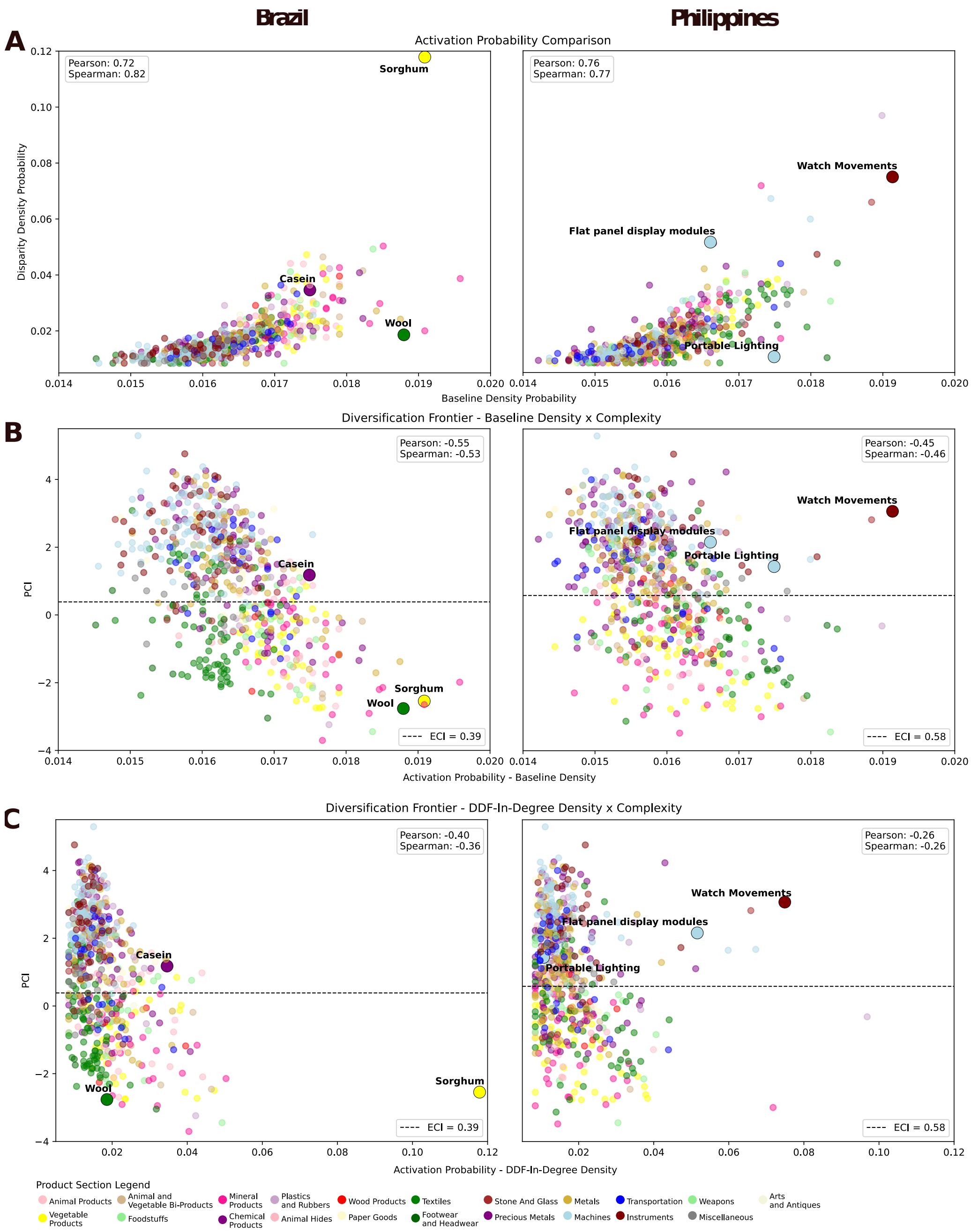


Figure 3.7 illustrates the differences between the baseline and disparity methods for industrial policymaking in Brazil and Philippines, considering all products for which each country had an RCA <= 0.5 in 2020. For each country, **A** displays the predicted probabilities of product activation (RCA >= 1) within four years for both density methods. **B** depicts the diversification frontier of both countries, combining baseline density probabilities with PCI, and **C** presents an alternative diversification frontier, using DDF-In-Degree density probabilities and PCI.

improvement, the PCI values of both products remain below Brazil's current ECI, indicating that neither would be ideal targets if the objective is to enhance complexity. If a choice were necessary, sorghum would represent a more promising option. Casein presents a stronger profile, with a PCI that exceeds Brazil's current ECI. The predicted probability of activation for casein nearly doubles when moving from the baseline to the filtered approach, increasing from approximately 1.9% to 3.8%. This substantial improvement strengthens its case as a viable diversification target. Filtering therefore not only adjusts predicted success rates but also refines the strategic prioritization of opportunities for Brazil.

For the Philippines, Panels B and C reveal an important evolution in the diversification frontier. Under the baseline specification in Panel B, there is a clear trade-off between complexity and feasibility. Flat panel display modules exhibit a higher PCI compared to portable lighting but are associated with a lower baseline density, making the decision between them less straightforward. When considering filtered probabilities in Panel C, this trade-off disappears. Flat panel display modules maintain a high PCI while achieving a substantially higher predicted probability of activation, making them a strictly superior candidate for industrial policy targeting. In addition, watch movements consistently appear among the most attractive targets in both the baseline and filtered configurations, combining relatively high probabilities of activation with high PCI values. This stability reinforces their strategic relevance within the diversification frontier of the Philippines.

Comparing Brazil and the Philippines, the filtered diversification frontier of the Philippines is visibly more favorable. The Philippines shows a larger set of products combining high predicted probabilities and high PCI values, concentrated mainly in the machinery, electronics, and precision instruments sectors. Furthermore, the evolution of correlations from Panels B to C provides additional insights. For both countries, Pearson correlations between activation probability and PCI become less negative after filtering. In Brazil, the Pearson correlation improves from -0.55 to -0.40, while in the Philippines it improves from -0.45 to -0.26. This shift, consistent with the patterns observed in Figure 3.5, indicates that, for countries like Brazil and the Philippines with intermediate levels of economic complexity, correcting for spurious proximities results in a set of diversification opportunities that are less misaligned with complexity upgrading. Although the filtered prospects do not necessarily offer a strong orientation toward complexity enhancement, they represent a notable improvement compared to the baseline, reducing the extent to which diversification paths are biased toward low-complexity activities.

These comparisons illustrate how filtering the Product Space not only adjusts perceived feasibility but also refines the trade-offs faced by policymakers, offering clearer and more robust pathways for structural transformation.

3.7 Discussion

This analysis reinforces the idea that less is more when applying network filtering techniques in the context of economic complexity and relatedness. By selectively amplifying the strongest signals and filtering out the surrounding noise, methods like DDF-In-Degree reveal a clearer map of countries' true productive capacities. Weak or spurious proximities, which often cloud strategic insights, are discarded, sharpening the focus on opportunities that are both feasible and high-impact. This minimalist approach yields more actionable guidance, offering an advantage in addressing the informational challenges noted by Juhász, Lane, and Rodrik (2023) in the development of effective industrial policy.

The principle that less is more underscores a subtle truth: simplifying complex data does not dilute its value, it enhances it by strengthening the reliability of the underlying signals. By foregrounding only the most meaningful relationships, filtering techniques remove informational noise and better align policy recommendations with the actual capabilities of each country. In doing so, they transform the Product Space from a dense and often distorted network into a more navigable and strategically powerful tool, allowing for more precise targeting of sectors where the impact can be profound and sustainable.

Rather than redefining relatedness, filtering enhances its precision. In the traditional Product Space, the excessive density of connections often obscures genuine diversification opportunities. Filtering counteracts this by removing redundant proximities, allowing the density measure to more accurately reflect the feasibility of transitioning into new products relative to a country's existing capabilities.

Empirically, filtering substantially improves prediction performance, with higher AUC-ROC and F1 scores compared to the baseline. It also mitigates the strong correlation between relatedness density and overall diversification, allowing a clearer, product-specific view of feasible upgrading paths. Beyond improving predictive metrics, filtering also softens structural barriers to complexity upgrading: countries at intermediate levels of economic complexity, in particular, face a diversification landscape that becomes less skewed toward low-complexity activities and offers a broader set of more realistic pathways for upgrading than suggested by the unfiltered Product Space.

The country cases of Brazil and the Philippines illustrate this effect: the filtered approach identifies more realistic and higher-probability diversification targets, correcting the distorted prioritization of opportunities that characterizes the traditional framework.

Overall, filtering relatedness measures refines rather than replaces the foundations of economic complexity analysis, enhancing its strategic usefulness for industrial policy-making by clarifying priorities and reducing the risks associated with identifying productive upgrading opportunities.

Beyond its immediate contributions, this approach opens several avenues for future research. Although this study applies network filtering techniques to the context of international trade and product diversification, the methodology is not inherently limited to this domain. Network filtering techniques could be applied to other datasets, such as patent networks, research collaboration matrices, or subnational economic structures, providing insights into productive diversification across different technological and geographic domains. For instance, applying directional filtering to regional innovation systems, such as those underpinning EU smart specialization policies (Balland *et al.*, 2019), could test whether their core relatedness findings persist under stricter statistical thresholds.

In addition, while this chapter focused primarily on improving the targeting of *what* countries should diversify into, filtering can also inform broader strategic dimensions. It may clarify *when* countries should pursue diversification opportunities, *where* productive upgrading is most likely to succeed geographically, and *who* is better positioned to lead these processes, in line with the 4Ws framework proposed by Hidalgo (2023). Expanding the use of filtering across these dimensions offers a promising pathway for refining industrial policy design and enhancing its alignment with the evolving structure of economic opportunities.

Appendix 3.1: Formal Definitions of Network Filtering Methods

This appendix presents the formal equations used in the network filtering methods implemented in this study, including: Disparity Filter, LANS, Directed Disparity Filter In-Degree, and Directed LANS In-Degree. For consistency, all notations refer to nodes (products) as p, p', or p''.

Disparity Filter (Serrano; Boguná; Vespignani, 2009)

For a node p with degree k_p, each edge weight $\varphi_{p,p'}$ is normalized:

$$\alpha_{p,p'} = \frac{\varphi_{p,p'}}{\sum_{p'' \in \mathcal{N}(p)} \varphi_{p,p''}}$$

Assuming a uniform distribution over the interval $[0,1]$, the significance level for the edge is given by:

$$\text{p-value:} \quad p_{p,p'} = 1 - (k_p - 1) \int_0^{\alpha_{p,p'}} (1-x)^{k_p - 2} dx$$

Because the network is undirected, the edge (p,p') is evaluated from both sides. The edge is retained if:

$$\min\left(p_{p,p'}^{(p)}, p_{p',p}^{(p')}\right) < \alpha$$

That is, the edge is kept if it is statistically significant for at least one of the two nodes.

Directed Disparity Filter In-Degree (DDF-In-Degree)

This method adapts the Disparity Filter to capture directionality by evaluating proximities from the perspective of the receiving node. Instead of considering outgoing flows, it focuses on the distribution of incoming proximities for each product p', emphasizing the significance of signals received.

Given a directed network where each node p' receives proximity values $\varphi_{p,p'}$ from other nodes, the normalized inflow is:

$$\alpha_{p,p'} = \frac{\varphi_{p,p'}}{\sum\limits_{p'' \in \mathcal{N}^{\text{in}}(p')} \varphi_{p'',p'}}$$

Assuming these weights are uniformly distributed over $[0,1]$, the significance of the edge $(p \to p')$ is:

$$p_{p,p'} = (1 - \alpha_{p,p'})^{k_{p'}^{\text{in}} - 1}$$

where $k_{p'}^{\text{in}}$ is the indegree of node p'. The edge is retained if $p_{p,p'} < \alpha$.

LANS (Foti; Hughes; Rockmore, 2011)

LANS evaluates the statistical significance of each edge weight $\varphi_{p,p'}$ by comparing it to all other weights from the same node p. The p-value is defined as the proportion of other edges with lower or equal weights:

$$p_{p,p'}^{(p)} = \frac{\left|\left\{p'' \in \mathcal{N}(p) \setminus \{p'\} \,:\, \varphi_{p,p''} \leq \varphi_{p,p'}\right\}\right|}{k_p - 1}$$

$$p_{p,p'} = \max\left(p_{p,p'}^{(p)}, p_{p',p}^{(p')}\right)$$

The edge (p,p') is retained if $p_{p,p'} < \alpha$.

Directed LANS In-Degree (DLANS-In-Degree)

DLANS-In-Degree adapts the original LANS method to directed networks by focusing exclusively on the incoming proximities of each node. For each target node p', the empirical significance of the edge from p to p' is calculated by comparing it to the other incoming proximities to p':

$$p_{p,p'} = \frac{\left|\left\{p'' \in \mathcal{N}^{\text{in}}(p') \setminus \{p\} \,:\, \varphi_{p'',p'} \leq \varphi_{p,p'}\right\}\right|}{k_{p'}^{\text{in}} - 1}$$

This value represents the proportion of other incoming signals to p' that are weaker than or equal to the one from p. As in the original LANS method, edges are retained if $p_{p,p'} < \alpha$.

These statistical methods allow the extraction of a more meaningful and interpretable backbone from the Product Space, minimizing the influence of noisy or spurious proximities and emphasizing the perspective of the product receiving the signal.

Appendix 3.2: Parameter Calibration and AUC-ROC Results

This appendix presents the calibration curves used to identify optimal parameter values for each network filtering method, based on their predictive performance. The figure below shows how AUC-ROC values vary across different parameter settings.

As illustrated in Figure 3A.1, most filtering techniques exhibit a single-peaked or near single-peaked behavior in their AUC-ROC response. In many cases, a relatively broad plateau of high-performing parameter values can be observed, indicating that predictive accuracy remains stable around the global maximum. This robustness simplifies the calibration process and enhances the practical usability of the filtering methods.

Figure 3A.1 – Parameter Calibration and AUC-ROC Results

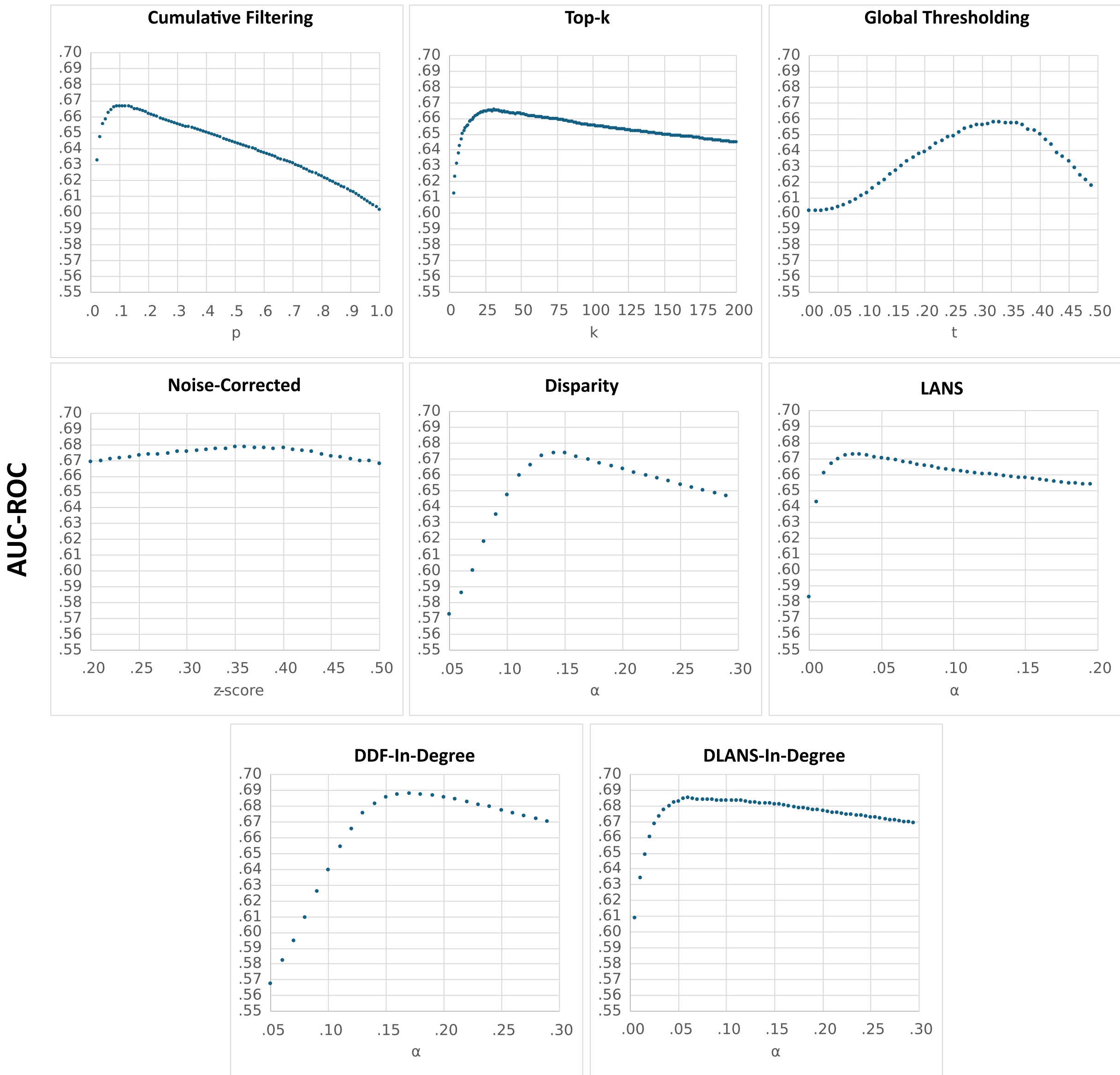


Figure 3A.1. AUC-ROC values obtained across different parameter configurations for each network filtering technique.

Appendix 3.3: Related and Unrelated Diversification

There is a growing literature investigating at which stages of economic complexity it becomes advantageous for countries to target unrelated products. Pinheiro *et al.* (2022) and Alshamsi, Pinheiro, and Hidalgo (2018) argue that unrelated diversification tends to be more prevalent, and necessary, during the intermediate stages of complexity development. This insight is closely complemented by the stylized fact of an S-shaped relationship between a country's ECI and the correlation between density and PCI among its latent products (Hartmann *et al.*, 2021). Specifically, this pattern shows that for most countries, particularly those with medium and upper-medium levels of complexity, the products that appear closest in the Product Space tend to be less sophisticated. As a result, countries at these stages often face a developmental plateau, where nearby opportunities no longer contribute meaningfully to further complexity upgrading. The complementarities between these two findings suggest that economies at intermediate ECI levels are structurally pushed toward seeking complexity-enhancing diversification among less related, and often more distant, products.

However, a fundamental limitation underlies this discussion: the reliance on baseline density to evaluate proximity between countries and products. As previously discussed, baseline density is heavily influenced by spurious correlations embedded in the original Product Space, meaning that judgments about whether a product is relatively near or far for a country are subject to significant bias. Consequently, conclusions drawn solely from the baseline configuration may misrepresent the true diversification landscape available to countries, particularly at intermediate stages of complexity. To address this concern, this appendix section applies the DDF-In-Degree filtering method to reassess whether the patterns and implications described above, particularly regarding the need for unrelated diversification during middle stages of complexity, remain robust after removing redundant and statistically insignificant proximities.

To further investigate the relationship between economic complexity and the relatedness of successfully upgraded products, Figure 3A.2 presents an empirical analysis based on data from 1998 to 2016. Following the relative relatedness methodology proposed by Pinheiro *et al.* (2022), we selected, for each country-year, products with RCA <= 0.5 at time t that achieved RCA >= 1.0 at t+4, and calculated the average density and relative average density of these products. Countries were then grouped into 20 equally sized ECI bins, with the x-axis reporting the average ECI within each bin. The left panels display results using the baseline approach, while the right panels apply the filtered DDF-In-Degree approach. In each case, a quadratic regression is fitted (in red), with the estimated coefficients and R^2 reported alongside. This setup allows us to assess whether the observed patterns of relatedness upgrading vary when controlling for the noise and redundancy inherent in the baseline Product Space structure.

The first thing that stands out from Figure 3A.2 is the strong positive relationship

Figure 3A.2 – Relatedness and Relative Relatedness - Baseline vs Filtered Approach - 1998-2016

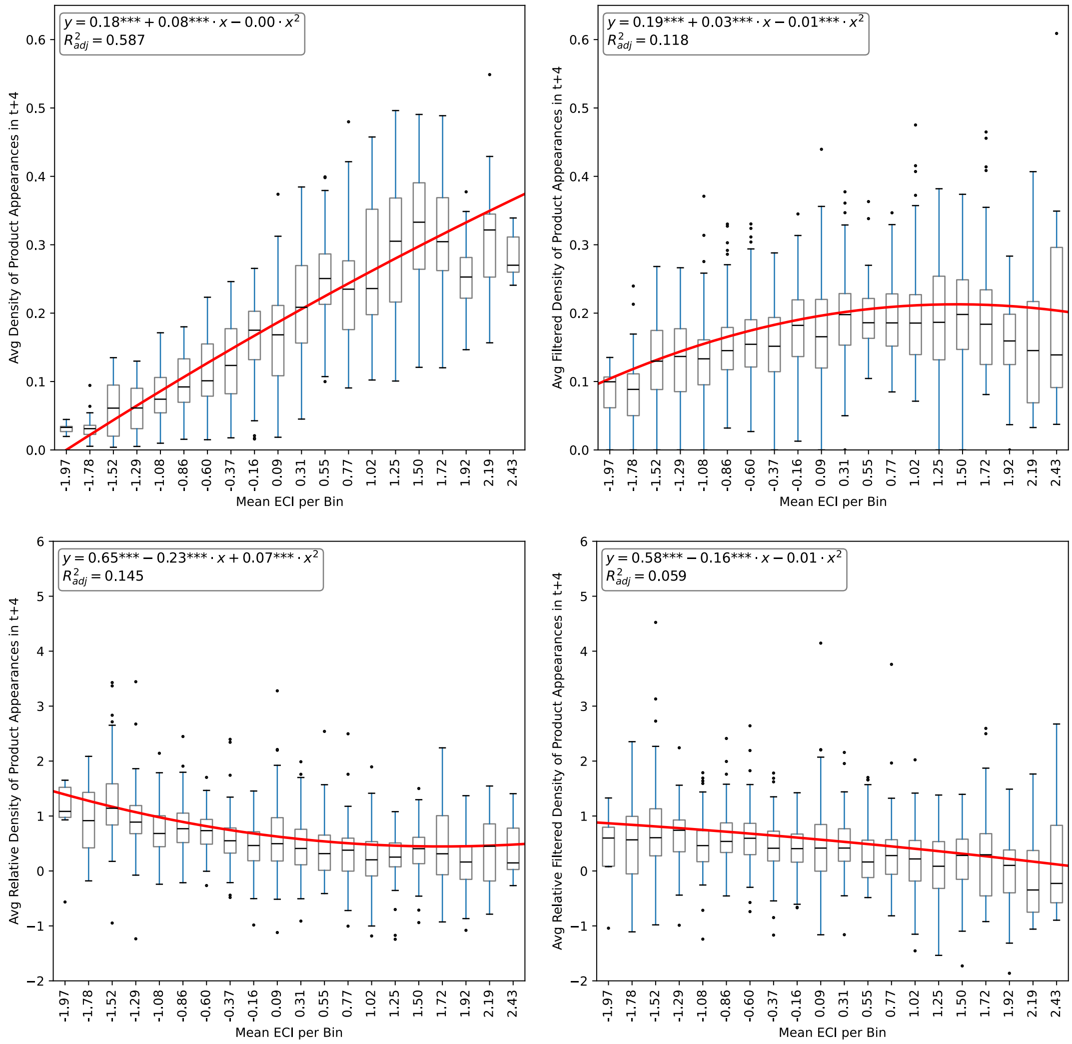


Figure 3A.2 presents the relationship between countries' ECI)and the relatedness of products they successfully upgraded to, based on data from 1998 to 2016. For each country-year, we selected products with RCA 0.5 at time t that achieved RCA > 1.0 at t+4. We then calculated the average density (top panels) and average relative density (bottom panels) of these newly exported products. Countries were grouped into 20 equally sized ECI bins, with the average ECI of each bin displayed on the x-axis. The left panels shows results based on the traditional baseline density measure, while the right one applies the filtered DDF-In-Degree method. In each panel, the red lines represent fitted quadratic regression curves, and the accompanying boxes report the estimated coefficients and the corresponding R2. The construction of the relative relatedness measure and the empirical strategy employed in this figure follow the methodology proposed by Pinheiro *et al.* (2022).

between countries' ECI and the baseline density of the products they successfully upgraded to. The fitted regression yields an R^2 of 0.59. However, this strong correlation likely reflects a substantial burden of spurious proximity effects and high correlation between density and diversity, as previously discussed. It is worth recalling that the predictive performance of baseline density in Probit models was relatively modest, with an AUC-ROC of only 0.602. In contrast, when we turn to the filtered density measure, the relationship between ECI and the density of newly exported products becomes much weaker, with an R^2 of only 0.12, while its AUC-ROC is substantially higher, at 0.688. This suggests that, after correcting for redundant and noisy links, the level of density faced by countries in their diversification efforts, now measured in a more precise way, is much less dependent on their current ECI level.

Regarding the relative relatedness analysis, we observe that the fitted relationship between ECI and relative relatedness is much better defined when using the baseline measure: the baseline regression achieves an R^2 of 0.145, compared to only 0.059 when using the filtered measure. In the baseline, the relationship follows a clear quadratic pattern, whereas in the filtered case, it becomes linear and much weaker, although still slightly decreasing, as indicated by the negative slope.

These combined results suggest that, once some of the spurious correlations embedded in the original Product Space are corrected, relative relatedness no longer reveals strong patterns about the types of products that countries tend to upgrade into, nor does it point to a specific development stage where focusing on unrelated diversification would be particularly advantageous. A plausible interpretation is that the explanatory power of relative relatedness under the baseline configuration stems largely from the strong, association between baseline density and overall diversification levels. Consequently, relying on baseline-based measures of relative relatedness may lead to misleading conclusions about the role of unrelated diversification in the development process. Nevertheless, it remains possible that the DDF-In-Degree filtering has not fully eliminated all sources of noise in the Product Space, which may explain why a small residual correlation persists even after filtering.

Figure 3A.3 further explores the relationship between countries' economic complexity and the structure of their diversification opportunities by plotting the S-curve pattern based on the correlation between density and PCI among latent products (Hartmann *et al.*, 2021; Pinheiro *et al.*, 2022). Compared to the baseline configuration, the filtered version of the S-curve appears notably flatter. This suggests that the overall diversification prospects for countries, particularly those at lower levels of complexity, are not as pessimistic as implied by the baseline measure.

The evidence from Figure 3A.3 complements the previous analysis of the characteristics of newly exported products in Figure 3A.2. While the earlier results highlighted that, after filtering, the density of successfully upgraded products no longer varies strongly with a country's ECI, the S-curve shown here provides a broader perspective on the underlying

Figure 3A.3 – S-Curve - Baseline vs Filtered Approach - 1998-2016

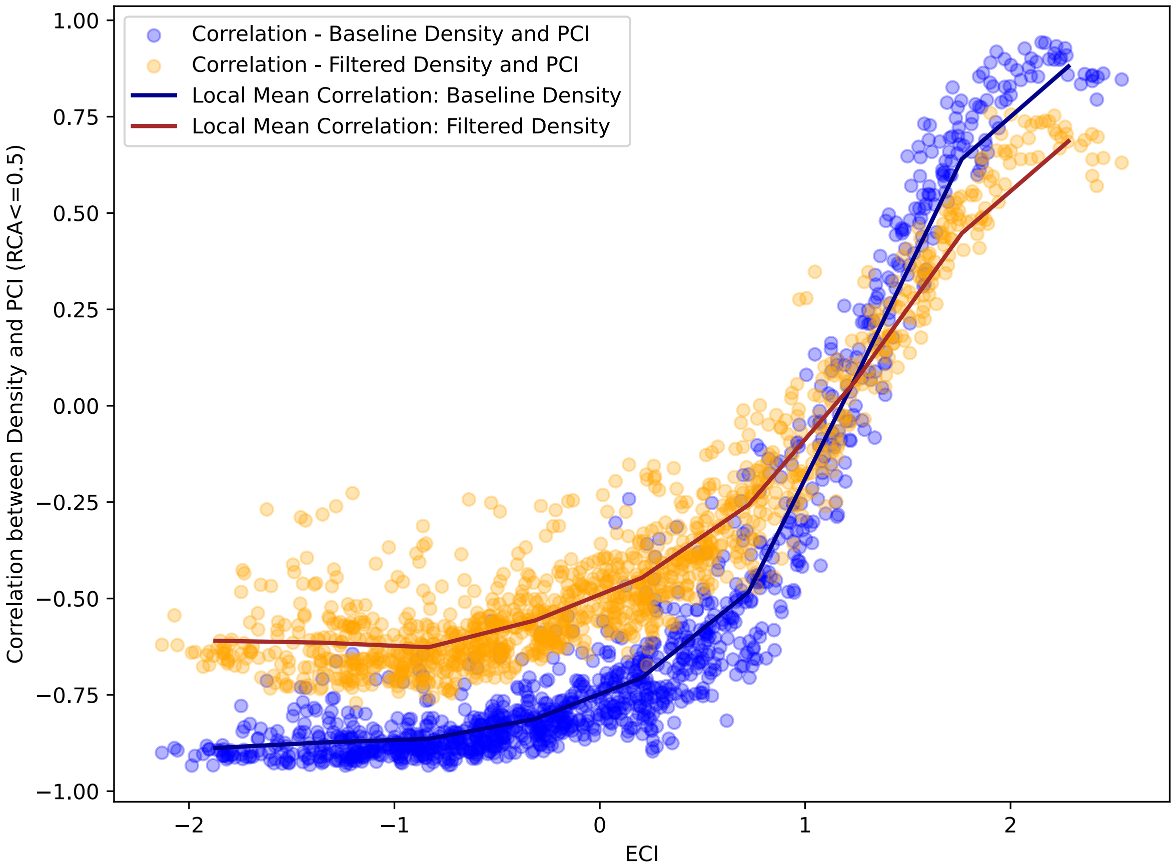


Figure 3A.3 presents the S-shaped relationship between a country's ECI and the correlation between density and PCI among products with RCA <= 0.5, based on data from 1998 to 2016. For each country-year observation, we calculated the Pearson correlation between density and PCI across all latent products and plotted it against the corresponding ECI. Blue dots represent observations based on the traditional baseline density measure, while orange dots correspond to the filtered density measure obtained through the DDF-In-Degree approach. The solid lines show the local mean trends for each method, highlighting how the baseline configuration (blue) produces a sharper S-curve pattern, whereas the filtered configuration (orange) reveals a flatter and smoother relationship.

diversification landscape faced by countries.

In particular, the flattening of the S-curve after filtering suggests that, for most countries, the average quality of latent diversification opportunities is not as systematically biased against low-ECI economies as initially implied by the baseline measure. However, both analyses converge in indicating that countries reaching intermediate complexity levels (around ECI = 1.0) encounter a narrowing funnel of development paths: diversification into complexity-enhancing products becomes significantly more challenging precisely at this stage. Thus, while filtering the Product Space reduces the apparent severity of structural barriers for low-ECI countries, the existence of a development bottleneck at intermediate stages remains a persistent feature across both upgraded products and latent opportunities.

4 From Capabilities to Economic Convergence: A Structural Growth Framework Linking Economic Complexity, Institutions, and Human Capital

Abstract

Why do some countries achieve sustained economic convergence while others remain stuck in low-growth traps? This paper investigates how structural capabilities — encompassing input-output economic complexity, trade-based complexity, institutional quality, and human capital — jointly influence countries' long-run development paths. We introduce a novel measure of economic complexity based on intersectoral input-output linkages and integrate it with existing trade-based metrics to capture the multidimensional structure of productive capabilities. To assess how these capabilities shape economic convergence, we develop a two-stage empirical strategy that explicitly distinguishes the role of baseline structural conditions from their subsequent evolution. In the first stage, we isolate the portion of per capita income not explained by complexity, institutions, or human capital using cross-sectional regressions for each base year. In the second stage, we estimate panel regressions to evaluate how both these residuals and ten-year structural transformations affect subsequent growth. Applying this framework to a panel of over 60 countries from 1999 to 2020, we find robust evidence that economic complexity, institutional quality, and human capital operate as complementary structural drivers of long-term development. Our results highlight the importance of investing in productive capabilities, institutional foundations, and human capital to enable sustained economic convergence.

Keywords: economic complexity, input-output structure, economic growth, industrial policy

4.1 Introduction

Input-output methods are one of the first in Economics — if not the first one — to recognize the interconnectedness of markets and sectors as crucial to understanding economic systems. Formalized by Leontief in the 1930s, these methods anticipated key principles that would later become central to the field of network theory. Fundamental network theory concepts such as 'nodes', 'edges', 'adjacency matrix', 'centrality', and 'contagion', all have their doppelgangers in input-output analysis, respectively matched by ‘industries’, ‘inter-industries flows’, ‘input-output table’, ‘industry multipliers’, and ‘shock propagation analysis’. Input-output analysis is an important tool, used for economic planning, forecasting, and policy analysis. It helps policymakers understand the ripple effects of economic changes and make informed decisions about resource allocation and economic development.

Economic complexity, on the other hand, arises in the 2000s as a field that applies machine learning and network theory techniques to understand economic systems and their processes (Hidalgo, 2021). The conceptual cornerstone of economic complexity is the importance of learning to economic development and the impediments to accumulating and transmitting productive knowledge through a selected economic network. Sharing with input-output analysis network theory as part of its foundations, it also relies deeply on the patterns of interaction between nodes and their spillovers within an economic network. For Hidalgo (2023), economic complexity research represents an example of unbalanced growth theory, as posed by Hirschman (1977) with forward and backward linkages in the use of input-output methods, providing methods to identify tailored diversification strategies based on a region’s current pattern of specialization. After all, both approaches focus on structural relationships and feedback loops that drive economic development.

The toolkits of input-output analysis and economic complexity have great potential to be integrated, enhancing their contributions to understanding economic development and designing better economic policies. However, irrespective of the sheer similarities of both approaches and their potential to be combined, there are not many studies that leverage them simultaneously (Gala *et al.*, 2018; Koch; Schwarzbauer, 2021; Pereira; Silva; Larruscaim, 2023). In this work, we contribute to filling this gap by applying economic complexity methods to the Inter-Country Input-Output (ICIO) tables, extracted from the Organisation for Economic Co-operation and Development (OECD) database.

We show that Input-Output Economic Complexity complements trade-based ECI by capturing structural features of production networks that go beyond the merchandise export basket. While trade ECI reflects the sophistication of what countries export, IO-based complexity incorporates how sectors contribute to value creation through both domestic and international input-output linkages, including services, which play an increasingly important role in modern production.

This broader view of productive capabilities reveals how countries differ not only in what they export but also in how their economies are embedded in complex, multi-sectoral value chains. Taken together, trade and IO-based measures form a multidimensional notion of economic complexity (Stojkoski; Koch; Hidalgo, 2023) that combines the external structure of exports with the internal architecture of production, helping to explain persistent differences in income levels across countries.

To assess how this multidimensional notion of complexity translates into long-term development, we propose a two-stage empirical framework. In the first stage, we estimate the portion of GDP per capita that cannot be accounted for by structural capabilities, namely economic complexity, institutional quality, and human capital, using cross-sectional regressions for each base year. In the second stage, we analyze how changes in these structural capabilities over time, along with the residuals from the first-stage regressions, influence subsequent economic growth. This two-stage approach allows us to separate the effects of initial conditions from those of structural transformation, providing a dynamic perspective on how productive capabilities shape long-term convergence, and serves also as a diagnostic tool for identifying capability-income gaps, highlighting areas where countries possess latent potential for structural upgrading and convergence.

The next section reviews the literature on economic complexity and long-term economic growth. Section 3 presents the methodological approach, detailing both the construction of the Input-Output Economic Complexity Index and the design of the two-stage framework. In Section 4, we analyze the empirical properties of the IO ECI and its relationship with income levels across countries. Section 5 turns to the two-stage framework, evaluating how complexity, institutions, and human capital contribute to long-term growth trajectories. Finally, we conclude by discussing the broader implications of our findings for economic development.

4.2 Literature Review

The economic complexity literature builds on the idea that a country's development path is shaped by the diversity and sophistication of its productive capabilities: those embedded, often tacit, forms of know-how that enable the production of complex goods. Development, in this view, is not neutral with respect to productive structure. What countries make, and how diversified and knowledge-intensive those products are, matters fundamentally. The seminal work by Hidalgo and Hausmann (2009) formalized this intuition into a measurable concept of economic complexity, proposing that the network connecting countries and the products they export can reveal the underlying distribution of capabilities. This framework rests on the insight that capabilities are not easily transferable, and that the accumulation of productive knowledge is a path-dependent, embedded process.

Since its introduction, the ECI has gained prominence as a structural indicator of a country's productive capabilities and development prospects. One of the main reasons for its growing influence is its strong empirical association with income levels and future economic growth. The Atlas of Economic Complexity (Hausmann *et al.*, 2014) popularized the use of ECI in policy and academic circles by showing that more complex economies tend to be richer and grow faster. Subsequent studies have reinforced this relationship and refined the methodology, demonstrating that economic complexity remains a robust predictor of growth even when controlling for human capital and institutional quality (Albeaik *et al.*, 2017; Gala; Rocha; Magacho, 2018; Stojkoski; Koch; Hidalgo, 2023). A common empirical strategy in this literature involves Barro-style growth regressions, where ECI is shown to be a robust predictor of long-run GDP per capita growth, even when controlling for traditional variables like human capital and institutional quality.

In parallel to the ECI framework, an alternative complexity metric known as fitness was proposed by Tacchella *et al.* (2012). Based on a non-linear iterative algorithm, the fitness approach seeks to capture the hidden capabilities of countries and the complexity of products by leveraging the structure of the export network without relying on eigenvector methods. A growing body of research has used this metric to analyze development trajectories and growth potential. Cristelli, Tacchella, and Pietronero (2015), Sbardella *et al.* (2018), and Tacchella, Mazzilli, and Pietronero (2018) showed that fitness is consistently associated with long-term GDP per capita growth, often outperforming conventional predictors in forecasting models. These contributions confirm that, despite methodological differences, both ECI and fitness highlight productive sophistication as a central structural determinant of economic performance.

Economic complexity methods are not restricted to the exchange of goods among countries. A significant portion of the literature has been developed for other applications, providing valuable policy recommendations that extend beyond the scope of international trade. When it comes to international trade, it is important to note that product exports serve as a vehicle for the interactive process between product types and countries. Similarly, other economic complexity applications always define an economic vector through which other two dimensions interact. Patent data, for instance, can reveal how technology and innovation spread through interactions between technological classes and geographical locations (Rigby, 2015; Balland; Rigby, 2017; Balland *et al.*, 2019). Employment variables by type of industry are also useful to be studied as a vehicle to economic processes (Neffke; Henning; Boschma, 2011; Mealy; Farmer; Teytelboym, 2019; Queiroz; Romero; Freitas, 2023).

In this sense, recent work has highlighted the need to move beyond a one-dimensional view of complexity to explain economic development. A multidimensional approach, combining different complexity metrics, offers a more complete picture of the knowledge embedded in an economy. Stojkoski, Koch, and Hidalgo (2023), for example, show that both Trade ECI and Technology ECI (based on patent data) contribute jointly to explaining long-run economic growth. Building on this perspective, recent work by Angelini *et al.* (2024) forecasts GDP growth using a combination of international trade fitness and technological fitness, reinforcing the idea that different facets of complexity can complement each other in capturing the full scope of a country's development potential.

Among alternative applications, input-output data stands out as a particularly promising yet underexplored avenue for measuring economic complexity. By capturing the flows of value added across sectors within and across countries, input-output frameworks provide a structural map of how industries are interconnected, offering a systemic perspective that complements trade-based approaches. While merchandise trade data reflects gross exports and is often biased toward tangible goods, input-output data encompasses the full spectrum of economic activity, including services and sectoral interdependencies. From a statistical standpoint, combining trade and input-output perspectives offers a richer representation of productive capabilities, one that balances domestic and international dimensions and better reflects the organization of modern economies.

Despite this potential, only a handful of studies have extended the complexity framework in this direction. Gala *et al.* (2018) used sectoral employment data from input-output matrices to show that the composition of employment across industries—particularly in manufacturing and sophisticated services—correlates with long-run growth and higher complexity. Koch and Schwarzbauer (2021) constructed value-added trade complexity indices using input-output flows for 56 industries and 43 countries, showing that these metrics complement traditional ones in explaining income levels. Pereira, Silva, and Larruscaim (2023) proposed an innovative mapping between HS product codes and ISIC activities to derive activity-based complexity measures for Brazil, revealing structural bottlenecks linked to the country's specialization in low-complexity value chains.

These extensions demonstrate that economic complexity can be meaningfully captured through a variety of empirical lenses. Yet, even as the literature has evolved to incorporate new data sources and dimensions, most studies continue to treat other key structural drivers of growth, particularly institutions and human capital, as secondary. Typically, these factors are included in growth regressions as control variables, while the primary focus remains on complexity metrics. This empirical treatment stands in contrast to a vast literature in development economics that identifies both institutional quality and human capital accumulation as fundamental engines of long-run growth. As such, there is a clear need for frameworks that go beyond isolated effects and instead consider how these foundational dimensions interact with one another, potentially reinforcing or constraining each other, in shaping development outcomes.

The institutional foundations of economic growth have long been emphasized in development economics. Pioneering works by North (1990) and Acemoglu, Johnson, and Robinson (2001), Acemoglu and Johnson (2005), Acemoglu, Johnson, and Robinson (2005) argue that inclusive institutions — those that ensure secure property rights, political stability, and broad access to opportunity — are fundamental for sustained development. Institutions shape the incentives that govern investment, innovation, and coordination, and thus play a critical role in enabling the accumulation and deployment of productive capabilities. Rodrik, Subramanian, and Trebbi (2004) provide compelling empirical evidence for this argument, showing that institutional quality dominates geography and trade integration as a determinant of long-term income differences across countries. This finding — that "institutions rule" — has since become a cornerstone in growth diagnostics and development strategy.

Building on this tradition, recent research has begun to bridge institutional theory and the complexity framework. Frenken, Neffke, and Dam (2023) propose a synthesis in which institutions are not external constraints but integral to the realization of capabilities. They argue that producing complex goods increasingly depends not only on the availability of technological capabilities but also on institutional arrangements that support knowledge sharing, coordination across firms, and trust within production systems. Further evidence of this enabling or constraining role comes from Boschma and Capone (2015), who demonstrate how different institutional architectures, such as coordinated market economies and liberal market economies, shape the direction of diversification. Coordinated market economies tend to facilitate related diversification through dense networks and collaborative innovation systems, while liberal market economies more often enable jumps into unrelated sectors, though typically at higher risk. This integrative view is further reinforced by Vu (2022), who provides empirical evidence that institutional quality is not only a driver of economic performance in general, but a direct determinant of a country's economic complexity. The study shows that better institutions foster economic complexity through two intertwined mechanisms: by incentivizing innovative entrepreneurship, particularly in the high-risk cost discovery process of identifying viable new products, and by promoting human capital accumulation and directing it toward productive uses.

These institutional effects also shape the practical effectiveness of industrial policy. Sousa and Mueller (2025) argue that the tools of economic complexity, though powerful in identifying viable paths for productive upgrading, are not sufficient in the absence of institutional arrangements that support implementation, learning, and commitment. In their view, institutions influence whether complexity-informed policies lead to sustained transformation or remain trapped by rent-seeking, policy discontinuity, or political capture. These contributions highlight that institutions do not merely "allow" development. They co-determine the paths through which complexity unfolds. Any attempt to model long-term growth trajectories must therefore treat institutions as fundamental structural factors that delimit and shape the possibilities for productive transformation.

The role of human capital in fostering long-term economic growth has been extensively discussed in both classical and endogenous growth theories. These frameworks converge in recognizing that a better-educated and more skilled labor force enhances productivity, facilitates technological adoption, and drives innovation (Jr, 1988; Romer, 1990; Mankiw; Romer; Weil, 1992). Foundational empirical studies (Becker, 1994; Mincer, 1981) further emphasized education and training as critical forms of investment, yielding both private and social returns. More recent reviews confirm that higher levels of educational attainment correlate with increased GDP per capita, although the strength and direction of causality can vary depending on institutional and structural contexts (Osiobe *et al.*, 2019). Benhabib and Spiegel (1994) highlight that human capital not only contributes directly as a factor of production but also enhances a country's ability to absorb and adapt frontier technologies. This view is reinforced by Hanushek and Woessmann (2008), who stress that cognitive skills and educational quality are more predictive of growth than schooling duration alone.

Importantly, Čadil, Petkovová, and Blatná (2014) argue that the returns to education depend on the structure of the economy, particularly its ability to absorb skilled labor. In countries with limited structural sophistication, education may lead to underemployment or skill mismatch, limiting productivity gains. While not framed in terms of complexity, this insight resonates with the core intuition of the complexity literature: human capital yields greater returns when embedded in a production structure capable of utilizing it. Just as complexity relies on the recombination of diverse capabilities, the full value of education emerges in economies with a sufficiently complex and integrated productive system. Human capital is thus necessary, but not always sufficient, for economic growth, with its impact contingent on the broader productive and institutional environment.

Recent empirical evidence reinforces this complementarity between human capital and the productive structure. Drawing on cross-country panel data, Lin (2011) show that educational attainment and economic complexity interact positively in promoting long-term growth. In particular, they find that secondary education plays a pivotal role in enabling countries to benefit from more complex productive activities, especially in developing economies. Their results suggest that human capital and complexity are mutually reinforcing: the effectiveness of education increases when embedded in a more sophisticated economic structure, and productive upgrading becomes more viable when supported by an educated workforce. This perspective aligns with the notion that economic growth depends not only on the accumulation of skills, but also on the capacity of the productive system to absorb and deploy those skills in complex and value-generating activities.

Taken together, this body of evidence supports a multidimensional framework for structural development. Whereas much of the economic complexity literature focuses on productive capabilities alone, our approach treats three structural vectors — productive capabilities, institutions, and human capital — as equally fundamental, while estimating their distinct empirical contributions. We operationalize capabilities through a multidimensional lens of economic complexity, combining the standard Trade ECI with a novel IO ECI. While Trade ECI captures the complexity of exports, IO ECI reflects the intersectoral sophistication of economy, considering both domestic and external dimensions, and including services. These two measures jointly define the productive vector. Institutions provide the enabling or restraining environment for coordination, learning, and credible commitments. Human capital equips the labor force to operate and expand production complexity. Together, these vectors shape both current development levels and future trajectories.

In sum, we build on a rich tradition linking capabilities, institutions, and human capital to economic development. Our contribution lies in proposing a unified framework where these vectors are treated symmetrically and interactively. Central to this is the introduction of the IO ECI, which complements trade-based measures to capture the full scope of a country's production structure. The next section details the construction of this index and presents the empirical strategy — the two-stage structural growth framework — designed to disentangle the role of each vector in explaining both income levels and long-term growth.

4.3 Methods

This section presents the methodological approach we use to examine how structural capabilities shape long-term economic growth. We begin by developing a new Input-Output Economic Complexity Index (IO ECI), which treats each intersectoral flow—from a supplying to a demanding industry—as a distinct economic activity. This formulation allows us to capture the structural sophistication embedded in how production is organized across sectors. Building on this measure, we adopt a multidimensional view of productive complexity by combining the IO ECI with the traditional trade-based Economic Complexity Index (Trade ECI). Together, they form a more comprehensive vector of productive capabilities. Our analysis is structured around three vectors of structural capabilities: (i) productive capabilities, captured by this combination of Trade ECI and IO ECI; (ii) institutional capabilities, measured through six dimensions of governance quality; and (iii) human capital, proxied by the Human Capital Index. We then implement a two-stage empirical strategy that separates the baseline contribution of these capabilities from the impact of their evolution over time. This framework allows us to identify how each structural vector contributes to economic convergence and growth.

To create the Input-Output Complexity Index, we use data from the OECD ICIO tables[1](OECD, 2023), which encompasses input-output data for 76 countries and 45 industries between years 1995 and 2020. This database was updated in 2023 and is unique among other input-output databases because of its comprehensive coverage in terms of country and time dimensions. Although its coverage is still far from resembling international trade databases in both dimensions, the country sample of ICIO is large and diverse enough to account for the process of economic development. We apply the same filter of Hausmann *et al.* (2014), excluding from the calculations all countries with population smaller than 1.2 million in 2010 or with unavailable data, resulting in 70 remaining countries.

1 http://oe.cd/icio/

The limited number of 45 industries constrains the application of economic complexity methods, since these methods extract information from matricial variance. To circumvent this limitation, we redefine the unit of analysis. Rather than treating each industry as an isolated activity, we designate the directed connection between two industries, from supplying sector s_1 to demanding sector s_2, as the basic unit of economic activity. This approach effectively treats the intersectoral flow itself as the analog of a "product" in standard trade-based complexity analysis, allowing us to construct a bipartite country–activity matrix.

The resulting structure yields up to 2,250 (44 x 50) unique Input-Output Activities, corresponding to all feasible sector combinations, including flows to final demand categories (e.g., household consumption, investment), while excluding the sector "Activities of households as employers." This formulation captures not only what is produced, but also how production is integrated across sectors, an essential feature of productive entangling. Not all 2250 activities are included in the calculations, since activities of minor importance produce noise for complexity methods. We filter all activities that have a world share of the final output higher than 0.00568%, which is an eighth of the expected share if all activities contributed equally to the final world output, totaling 1129 activities. Figure 4.1 provides a visual overview of this conceptual transformation from raw input-output data into a structured space of intersectoral activities.

We define the set of Input-Output Activities $\mathcal{A}$ as:

$$\mathcal{A} = \{s_1 \times s_2 \mid s_1, s_2 \in \mathcal{Z}\} \cup \{s_1 \times y \mid s_1 \in \mathcal{Z}, y \in \mathcal{Y}\} \tag{4.1}$$

where $\mathcal{Z}$ denotes the set of intermediate demand, and $\mathcal{Y}$ denotes the set of final demand. In this formulation, s_1 represents the supplying sector, while s_2 or y represent, respectively, the demanding production sector or the final demand category. The unit of analysis is thus a directed flow from s_1 to s_2 or from s_1 to y, capturing both intermediate and final activities.

Let $\varphi_{c,a}$ represent the total value generated by country c in activity a, aggregating over all destination countries c':

$$\varphi_{c,a} = \sum_{c' \in C} F_{c,a;c'} \tag{4.2}$$

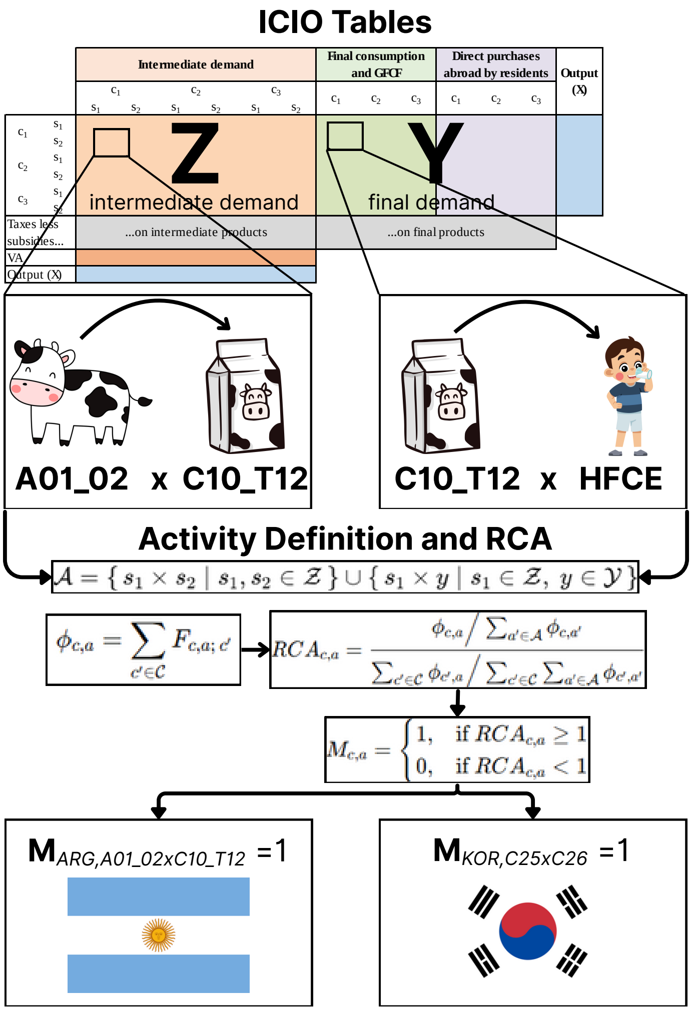
The revealed comparative advantage (RCA) of country c in activity a is given by:

$$RCA_{c,a} = \frac{\varphi_{c,a} / \sum_{a' \in \mathcal{A}} \varphi_{c,a'}}{\sum_{c' \in C} \varphi_{c',a} / \sum_{c' \in C} \sum_{a' \in \mathcal{A}} \varphi_{c',a'}} \tag{4.3}$$

Finally, the binary matrix $M_{c,a}$ is defined as:

$$M_{c,a} = \begin{cases} 1, & \text{if } RCA_{c,a} \geq 1 \\ 0, & \text{otherwise} \end{cases} \tag{4.4}$$

Figure 4.1 – Conceptual Map of Input-Output Data Extraction



Example: **Argentina** has comparative advantage in agriculture-to-food processing; **Korea** in metal products-to-electronics manufacturing.

Figure 4.1 illustrates the conceptual process through which input-output data is transformed into complexity measures. The diagram depicts how intersectoral linkages—such as the flow of raw agricultural inputs into industrial processing—are used to construct a network representation of productive activities. These flows are the basis for measuring Input-Output Complexity and its role in economic development analysis.

A central feature of our approach lies in the setting of the inter-industry flow as an activity, since we allow for a different set of required capabilities when distinguishing between the same industry supplying (or demanding) inputs to different industries. The relationship between 'Chemical and chemical products', for instance, and any other demanding industry, such as 'Agriculture, hunting, forestry' or 'Computer, electronic and optical equipment', carries diverse embedded economic information and capabilities, and may yield completely different levels of sophistication, i.e. different PCIs.

We compute the Input-Output Economic Complexity Index (IO ECI) and the Input-Output Activity Complexity Index (IO ACI) by applying the standard complexity methodology of (Hidalgo; Hausmann, 2009) to the binary matrix $M_{c,a}$. Following the method of reflections, we construct the matrix $\tilde{M}_{c,c'}$:

$$\tilde{M}_{c,c'} = \sum_{a} \frac{M_{c,a} \, M_{c',a}}{k_c \, k_a} \tag{4.5}$$

where $k_c = \sum_a M_{c,a}$ is the diversification of country c, and $k_a = \sum_c M_{c,a}$ is the ubiquity of activity a.

The IO ECI is given by the eigenvector associated with the second largest eigenvalue of the matrix $\tilde{M}_{c,c'}$. The IO ECI and the IO ACI can be expressed as iterative averages that mirror each other:

$$\text{IO ECI}_c = \frac{1}{k_c} \sum_{a} M_{c,a} \cdot \text{IO ACI}_a \tag{4.6}$$

$$\text{IO ACI}_a = \frac{1}{k_a} \sum_{c} M_{c,a} \cdot \text{IO ECI}_c \tag{4.7}$$

To evaluate the contribution of the newly developed IO ECI to long-term economic growth, we implement an econometric strategy designed to isolate its structural role in shaping cross-country growth trajectories. This strategy builds on a two-stage framework that distinguishes between the static effects of baseline embedded capabilities and the dynamic impact of their transformation over time.

A common approach in the economic complexity literature is to run Barro-style growth regressions, which includes the initial GDP per capita together with measures such as ECI to explain future economic performance. However, this raises important econometric concerns. Since ECI and income are strongly correlated, including both variables in the same regression introduces a risk of multicollinearity and complicates the interpretation of coefficients. More critically, ECI may reflect information already embedded in income levels, leading to issues of endogeneity and simultaneity. In other words, complexity and income are jointly determined at the initial time point, making it difficult to identify the independent role of complexity in shaping future growth.

To isolate the true contribution of economic complexity at the starting point to future economic growth, we implement a residualization procedure inspired by the two-stage residual inclusion (2SRI) strategy (Terza; Basu; Rathouz, 2008): first, regress complexity on initial income; second, use residuals to predict economic growth. Figure 4.2 provides a visual summary of this two-stage structural growth framework, illustrating how each capability vector is used to generate income-adjusted residuals that enter the second-stage growth regressions.

Figure 4.2 – Conceptual Map of Two-Stage Structural Growth Framework

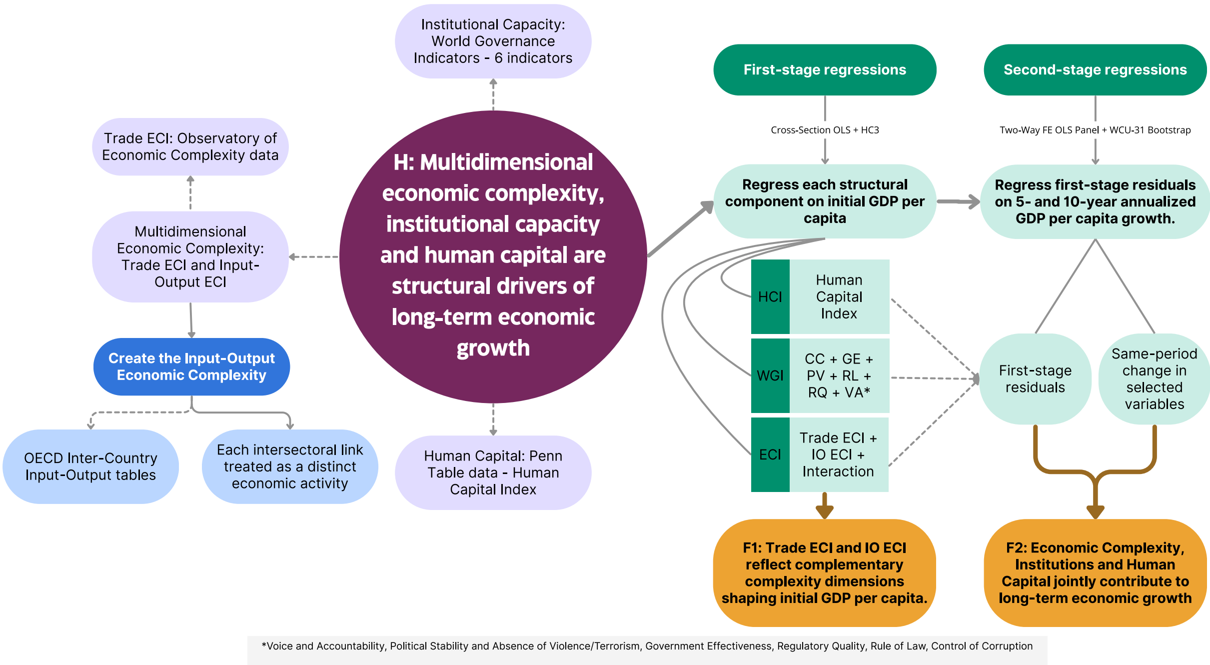


Figure 4.2 presents the two-stage structural growth framework developed in this study. In the first stage, initial GDP per capita is explained by structural variables: multidimensional economic complexity, institutional quality, and human capital. The second stage relates economic growth to both the residuals from stage one and structural transformations over the same period. This approach allows disentangling baseline capabilities from dynamic changes in explaining long-run development.

Specifically, we aim to identify the portion of complexity that diverges from the country's income level at time t. The underlying idea is that economic complexity captures a country's productive capabilities, and such capabilities should, in principle, be associated with a certain level of income. When a country's complexity level is higher than its corresponding income per capita, this may reflect untapped potential — a productive structure capable of supporting a higher level of output. In this case, we expect the country to grow faster in the future, converging toward the income level implied by its capabilities (Hausmann *et al.*, 2014). Conversely, countries with income levels above what their complexity would suggest may grow more slowly. This logic provides a structural interpretation of economic convergence, grounded in productive fundamentals rather than income alone.

Stojkoski, Koch, and Hidalgo (2023) have emphasized the idea that economic complexity may be inherently multidimensional, with different indicators capturing distinct facets of a country's productive capabilities. In this study, we build on that perspective by proposing that the IO ECI complements the traditional Trade ECI in capturing different layers of productive structure. From a statistical standpoint, input-output data captures value-added flows across all stages of production, spanning primary, secondary, and tertiary activities, and reflects both domestic and international linkages. In contrast, international trade data focuses on cross-border merchandise transactions, with a bias toward tradable, industrial products and gross export values. By combining IO ECI and Trade ECI, we aim to capture a broader and more nuanced picture of the productive capabilities embedded in an economy, and the level of income that such capabilities should support.

Beyond testing the role of multidimensional economic complexity, captured jointly by Trade ECI and IO ECI, we extend the analysis to other foundational capabilities that shape long-term development. In particular, we consider institutional capabilities and human capital capabilities as complementary to productive capabilities. Each of these is treated analogously: rather than including them directly as explanatory variables of growth, we isolate their residual component with respect to initial income. This comprehensive approach offers three key advantages. First, it prevents institutional quality and human capital from acting as hidden confounders for residual ECI estimates. Second, it allows for an equal footing comparison of the individual influence of each long-term growth structural driver. Third, it avoids the interpretational paradox of including initial income alongside structural variables that are, by construction, strongly associated with it. By including these residuals in the second stage, we test whether countries whose institutional capacity or human capital levels exceed (or fall short of) what their income would suggest experience systematically different growth paths, extending the logic of capability-driven convergence beyond productive structures alone. Additionally, since the residuals are orthogonal to initial income, multicollinearity is substantially reduced.

To implement the first stage of our empirical strategy, we run separate cross-sectional regressions for each base year, in which the log of GDP per capita is explained by structural capability measures. We report heteroskedasticity-robust standard errors using the HC3 estimator. The goal is to isolate the portion of each vector capability that is not already reflected in a country's income level at time t. For each structural vector $X_{c,t}$, we estimate:

$$\log(\widehat{\mathrm{GDPpc}}_{c,t}) = \alpha + \beta X_{c,t} \tag{4.8}$$

and compute the residual as:

$$\epsilon_{c,t} = \log(\mathrm{GDPpc}_{c,t}) - \log(\widehat{\mathrm{GDPpc}}_{c,t}) \tag{4.9}$$

where:

- $\log(\text{GDPpc}_{c,t})$ is the actual log of GDP per capita (PPP, constant USD) for country c in year t;
- $\log(\widehat{\text{GDPpc}}_{c,t})$ is the predicted income based on the country's capability level;
- $X_{c,t}$ represents one of the structural vectors: productive, institutional, or human capital;
- $\epsilon_{c,t}$ is the residual term, interpreted as the income gap.

The residual $\epsilon_{c,t}$ represents the difference between a country's actual income and the income predicted by its capabilities at time t. A positive residual indicates that the country's GDP per capita is higher than what its capabilities would predict, suggesting that it may be overperforming relative to its structural foundations. In such cases, we expect slower growth going forward, as the country may regress toward the income level implied by its underlying capabilities. Conversely, a negative residual indicates that the country is underperforming, its productive, institutional, or human capital capabilities would support a higher level of income. These countries are expected to grow faster, converging upward toward the level of income compatible with their capabilities. This interpretation provides a structural foundation for economic convergence dynamics, centered on latent development potential rather than on income gaps alone.

We define the three foundational capability vectors, each with its own first-stage specification, as detailed below:

- **Productive capabilities:** The productive vector combines complexity measures from merchandise international trade and input-output data. It includes Trade ECI (sourced from the Observatory of Economic Complexity[2]), our IO ECI, and their interaction term. The specification is:

$$X_{c,t}^{\text{ECI}} = \left[\text{Trade ECI}_{c,t}, \text{IO ECI}_{c,t}, \text{Trade ECI}_{c,t} \times \text{IO ECI}_{c,t}\right] \tag{4.10}$$

- **Institutional capabilities:** The institutional vector includes all six dimensions of the World Bank Worldwide Governance Indicators (WGI) (Kaufmann; Kraay, 2024): Control of Corruption (CC), Government Effectiveness (GE), Political Stability and Absence of Violence (PV), Rule of Law (RL), Regulatory Quality (RQ), and Voice and Accountability (VA).

$$X_{c,t}^{\text{INST}} = \left[\text{CC}_{c,t}, \text{GE}_{c,t}, \text{PV}_{c,t}, \text{RL}_{c,t}, \text{RQ}_{c,t}, \text{VA}_{c,t}\right] \tag{4.11}$$

- **Human capital capabilities:** The human capital vector consists of a single variable: the Human Capital Index from the Penn World Table (Feenstra; Inklaar; Timmer, 2015). This index is based on the average years of schooling and an assumed rate of return to education. The specification is:

2 http://oec.world/en/

$$X_{c,t}^{\mathrm{HC}} = \left[\mathrm{HCI}_{c,t}\right] \tag{4.12}$$

To assess how deviations from capability-implied income levels translate into subsequent economic performance, we implement a second-stage panel regression in which five- and ten-year growth rates are explained by the residuals obtained from the first-stage regressions. These residuals, computed at the initial year t, capture the portion of each capability not reflected in income at baseline and serve as capability-adjusted predictors of convergence. All second-stage models are estimated using two-way fixed effects (country and base-year) with WCU-S (Wild Cluster Unrestricted – Score) bootstrap, clustered at the country level to ensure robust inference under potential heteroskedasticity, within-country serial correlation, and a limited number of clusters.

To ensure that the residuals truly reflect deviations from capability-implied income at the baseline, rather than capturing later structural shifts, we include the variation over the five- or ten-year period in selected components of the capability vectors. We control for changes in IO ECI, Trade ECI, and two key institutional dimensions: Political Stability and Absence of Violence (PV), and Regulatory Quality (RQ). This approach is critical to avoid omitted variable bias that could arise if countries undergo significant reforms during the growth window—transformations that are not captured by initial capability levels but may substantially influence economic performance. By separating the cross-sectional component (residuals) from the intertemporal component (changes), we can identify whether growth stems from the structural conditions at the starting point or from improvements in those conditions over time. Moreover, the estimated coefficients on these capability changes may provide additional insights into the coevolution between structural transformation and economic growth.[3]

We also control for natural endowments using the share of natural resource rents in GDP, sourced from the World Bank World Development Indicators. This control is important because commodity price cycles can significantly affect medium- and long-term growth in resource-dependent countries, independently of their underlying capabilities.

Given the limited availability of complete data between 1999 and 2020 (or up to 2019 when the Human Capital Index is included), we designed our panel strategy to balance the desire for long-term growth analysis with the need for a sufficient number of time observations to estimate two-way fixed effects models. Ideally, we would analyze non-overlapping ten-year growth periods to capture long-run dynamics. However, this would drastically reduce the number of time points available, limiting within-country variation. To address

[3] We do not include the change in human capital directly, as the Human Capital Index from the Penn World Table is only available through 2019. Including this variable would substantially reduce the degrees of freedom in several panel blocks, as it would eliminate the final observation in each panel sequence. Furthermore, when tested in alternative specifications, the change in human capital did not exhibit statistical significance.

this constraint, we construct two distinct sets of panels, each with $t = 4$ periods, which allows us to estimate models with both country and time fixed effects while preserving variation in the explanatory variables. This overlapping structure, while relatively uncommon, has precedent in empirical macroeconomic research, including studies such as Panizza and Presbitero (2014), Bekaert, Harvey, and Lundblad (2005), and Tornell, Westermann, and Martinez (2004), which adopt overlapping windows to study long-term effects under data limitations.

The first panel set focuses on ten-year growth windows with overlapping periods. We construct three separate regressions with different initial base years — 1999, 2000, and 2001 — each forming a moving panel with a three-year step (e.g., 1999, 2002, 2005, 2008). This setup allows us to test convergence over a long horizon while maintaining $t = 4$, although it introduces substantial overlap across observations (e.g., 1999 and 2002 share seven years of economic growth in common). As a complementary robustness check, we build a second panel set using five-year growth windows without overlap. These non-overlapping panels begin in 1999 and 2000, progressing every five years (e.g., 1999, 2004, 2009, 2014). While this design sacrifices the long-term perspective of the ten-year models, it eliminates intertemporal dependence between observations and offers an alternative structure to assess the consistency of our results. Each panel is balanced across time, including the same set of countries within each specification. While the country sample may vary slightly across panel blocks—ranging from 63 to 65 countries—each panel maintains internal consistency over the four time periods considered.

To ensure reliable inference given the relatively small number of clusters in our panels, we employ the WCU-S (Wild Cluster Unrestricted – Score) bootstrap (MacKinnon; Nielsen; Webb, 2023), clustering at the country level. This method addresses heteroskedasticity, within-cluster correlation, and violations of cross-cluster independence that are common in fixed effects panel regressions. When the number of clusters is small, conventional CRV1-based inference often understates uncertainty, leading to over-rejection of null hypotheses (Type I error inflation). The WCU-S bootstrap mitigates this issue by applying wild bootstrap weights to jackknife-modified score contributions while retaining the CRV1 variance estimator. This approach yields finite-sample-robust p-values and is particularly effective in panels with overlapping observations, where time dependence can exacerbate the limitations of standard methods. The jackknife transformation helps accommodate complex dependence structures, enabling more reliable inference than conventional cluster-robust techniques.[4]

The second-stage regression is specified as follows:

[4] We did not use WCU-B because, although it pairs jackknife scores with CRV3, it often suffers from matrix inversion issues in unbalanced panel designs. In contrast, WCU-S avoids these numerical instabilities while still delivering finite-sample-robust p-values.

$$\begin{aligned}\text{Growth}_{c,t,t+h} = \gamma &+ \delta_1 \epsilon^{\text{ECI}}_{c,t} + \delta_2 \epsilon^{\text{INST}}_{c,t} + \delta_3 \epsilon^{\text{HC}}_{c,t} + \theta_1 \Delta\text{IO ECI}_{c,t,t+h} \\ &+ \theta_2 \Delta\text{Trade ECI}_{c,t,t+h} + \theta_3 \Delta\text{PV}_{c,t,t+h} + \theta_4 \Delta\text{RQ}_{c,t,t+h} \\ &+ \lambda \log \text{NRR}_{c,t} + \mu_c + \tau_t + u_{c,t}\end{aligned} \tag{4.13}$$

where:

- $\text{Growth}_{c,t,t+h}$ is the annualized growth rate of GDP per capita between year t and $t + h$, with $h = 5$ or 10;
- $\epsilon^{\text{ECI}}_{c,t}$, $\epsilon^{\text{INST}}_{c,t}$, and $\epsilon^{\text{HC}}_{c,t}$ are the capability-adjusted income residuals from the first-stage regressions of economic complexity, institutional capacity and human capital, respectively;
- $\Delta\text{IO ECI}_{c,t,t+h}$, $\Delta\text{Trade ECI}_{c,t,t+h}$, $\Delta\text{PV}_{c,t,t+h}$, and $\Delta\text{RQ}_{c,t,t+h}$ are the changes in the corresponding capability measures over the growth period;
- $\log \text{NRR}_{c,t}$ denotes the natural logarithm of the share of natural resource rents in GDP;
- μ_c and τ_t denote country and base-year fixed effects;
- $u_{c,t}$ is the error term.

In sum, this methodological approach allows us to examine how different types of structural capabilities contribute to long-term economic growth. We define three foundational capability vectors: productive, institutional, and human capital. The productive vector combines a newly proposed Input-Output Economic Complexity Index (IO ECI) with a traditional trade-based measure, capturing complementary dimensions of economic sophistication. By separating the role of initial conditions from the effects of change over time, the two-stage framework provides a structured way to evaluate how each vector shapes the path of economic convergence.

In the next section, we present the empirical results. We begin by examining key properties and cross-country patterns of the IO ECI, highlighting how it complements traditional trade-based measures. We then assess how the three vectors of capabilities shape future economic growth through the lens of the two-stage structural framework.

4.4 Input-Output Economic Complexity

Development is not a random walk. Countries do not stumble into prosperity by accident, nor do they diversify into new activities by rolling dice. Productive structures evolve along discernible paths, constrained by history, capabilities, and the tacit logic embedded in economic networks. Yet, to unpack these constraints, we must first look beneath the surface of GDP and trade flows, and examine the architecture of what countries actually do.

Figure 4.3 presents the country–economic activity matrix $\mathbf{M}_{ca}$ for the year 2010. This binary matrix indicates whether a given country displays revealed comparative advantage

(RCA $\geq$ 1) in a particular economic activity, based on input-output data. To explore its structure, we apply spectral co-clustering (**k** = **6**), which is a method that identifies latent structural patterns in a bipartite graph by exploiting the duality between row and column entities (Dhillon, 2001); in our case, countries and economic activities. Rather than assigning each row and column to independent clusters, spectral co-clustering models the data as a bipartite graph and uses the second left and right singular vectors of a normalized matrix to simultaneously uncover patterns of co-association. This approach corresponds to a real-valued relaxation of the NP-complete minimum cut problem in bipartite graphs and tends to yield globally meaningful partitions. Unlike traditional clustering, it captures the recursive structure whereby groups of countries are defined by the activities they specialize in, and vice versa.

Figure 4.3 – Country-Economic Activity Matrix - 2010

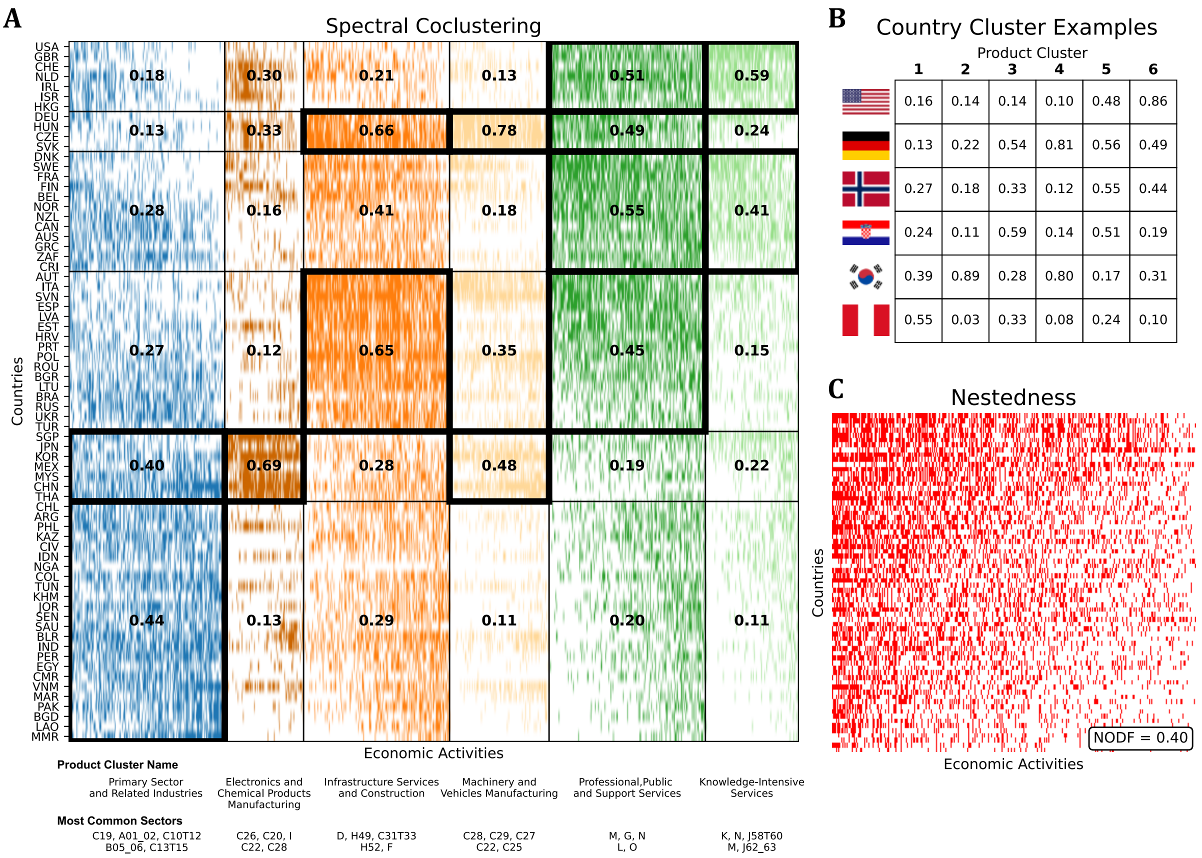


Figure 4.3 presents the country–economic activity binary matrix for 2010. **A** shows the spectral coclustering of countries and economic activities. The matrix is reordered into 6 clusters for each dimension, with average density values shown in each country–activity block. The most common sectors associated with each activity cluster are listed below. **B** depicts illustrative examples of country cluster densities across product clusters for six representative economies. **C** displays the reordered binary matrix used to compute nestedness; the NODF value is reported in the bottom right.

The result is a structured checkerboard of six country clusters and six activity clusters, which we order by the average complexity of their components (measured by IO ECI for countries and IO ACI for activities). Each block displays its concurrent density, i.e the share of activities within a given cluster for which countries in a given group hold revealed comparative advantage. Stronger-than-average density associations (above 0.5 standard deviations) are highlighted with thicker borders, suggesting areas of productive coherence.

The first activity cluster, in blue, captures low-complexity domains — primary sectors and low value-added manufacturing. Clusters two, three, and four (orange hues) span more sophisticated manufacturing and infrastructure-related activities: cluster two concentrates electronics and chemicals; cluster four focuses on vehicles and machinery; cluster three connects construction with broader infrastructure services. Clusters five and six, in dark and light green, represent service activities. The fifth emphasizes professional, administrative, and public services; the sixth is composed of high knowledge-intensity activities, digital services and finance.

Country clusters display striking heterogeneity. At the base sits a broad group of developing economies, with dense engagement in primary sectors but limited presence in manufacturing or services, being the the cluster with the lowest average ECI. Above it, we encounter a classic "factory economy" configuration, composed primarily of East Asian countries such as China, South Korea, and Japan. These nations exhibit high density in core industrial activities, particularly clusters two and four, yet remain anchored in primary sectors and show limited diversification into services.

Climbing further, we reach a hybrid cluster of Southern European and Baltic countries (e.g., Portugal, Estonia), and emerging economies like Brazil and Turkey. Compared to the former country cluster, these countries show declining reliance on raw materials and growing participation in construction, logistics, and administrative services. Although their manufacturing footprint is narrower, they engage moderately in high-value manufacturing segments, especially those linked to vehicles and machinery.

The next cluster reveals two intertwined paths to complexity. On one side, Western European and Nordic economies like Denmark and France rely on advanced manufacturing, professional services, and a dense web of knowledge-based service activities. On the other, resource-rich countries like Norway and Australia, though specialized in primary exports, show unexpectedly high engagement in advanced service sectors. These economies remind us that complexity does not always require industrialization in the classical sense: it can emerge from unusual pairings of natural resource wealth and high-end services activities.

Above them lies a tightly integrated block of highly industrialized Central European economies, centered around Germany. This group exhibits consistently high densities across all manufacturing clusters, especially in machinery and vehicles, reflecting deep industrial specialization. These countries pair technical sophistication with strong administrative service capacity, yet they display a modest presence in the domain of high-end, knowledge-intensive services.

Finally, we find a small group of advanced economies that seem to have left the factory behind, not by abandoning production, but by moving the frontier of value creation elsewhere, to knowledge-intensive services. Their goods are still made, but the real value-added work happens in code, contracts, coordination, and ideas. In these economies, complexity no longer sits on the factory floor; it lives in design studios, research labs, regulatory frameworks, and global service architectures.

Taken together, this 6-by-6 structure illustrates that there is no single path to economic development. Some countries climb through manufacturing; others through services. But the most successful appear to combine both. The real question is how and whether countries stuck in industrial specialization or natural resource dependence can make the leap toward enhanced input-output complexity. For East Asian factory economies, the challenge is to shift from manufacturing dominance to service diversification. For resource-rich countries, the existence of complexity in the almost absence of manufacturing is rare, but not impossible. The structure of the $\mathbf{M}_{ca}$ suggests that such transitions, while difficult, may be within reach for those who know where to look.

To understand what makes a country complex in the input-output, intersectoral linkage sense, one must look at the structure of the activities in which it holds a revealed comparative advantage. The IO-based Economic Complexity Index (IO ECI) of a country is, by construction, an average of the complexity of its activities—measured through the IO-based Activity Complexity Index (IO ACI). In this sense, a country's productive complexity reflects the complexity of the activities in which it engages. Figure 4.4 provides a first look into this mirror by ranking economic sectors according to the average complexity of the activities in which they are involved—either as input providers (supply side) or as input users (demand side). In doing so, it reveals which sectors tend to be embedded in high-complexity interactions, offering a structural perspective on how complexity concentrates across the economy's productive architecture.

By separating the supply and demand perspectives, Figure 4.4 exposes a duality in how sectors relate to complexity. Some sectors exhibit high complexity as input demanders, orchestrating diverse and sophisticated sourcing networks, while others are more complex as input providers, supporting high-value production in other sectors. These asymmetries are not merely descriptive. They evoke Albert Hirschman's classic notions of backwardness and forwardedeness (Hirschman, 1958; Hirschman, 1977), in which the developmental potential of a sector depends on its structural position within the web of intersectoral linkages. Inspired by this intuition, we introduce the notion of forward- and backward-complexity: the idea that the complexity associated with a sector depends on whether it absorbs sophistication from upstream inputs or transmits it downstream through the production of enabling outputs. In this view, complexity is not just a property of sectors in isolation, but a function of their relational embeddedness: whether they anchor upstream capabilities or channel downstream

Figure 4.4 – Average Complexity of Input-Output Activities by Industry - 2010

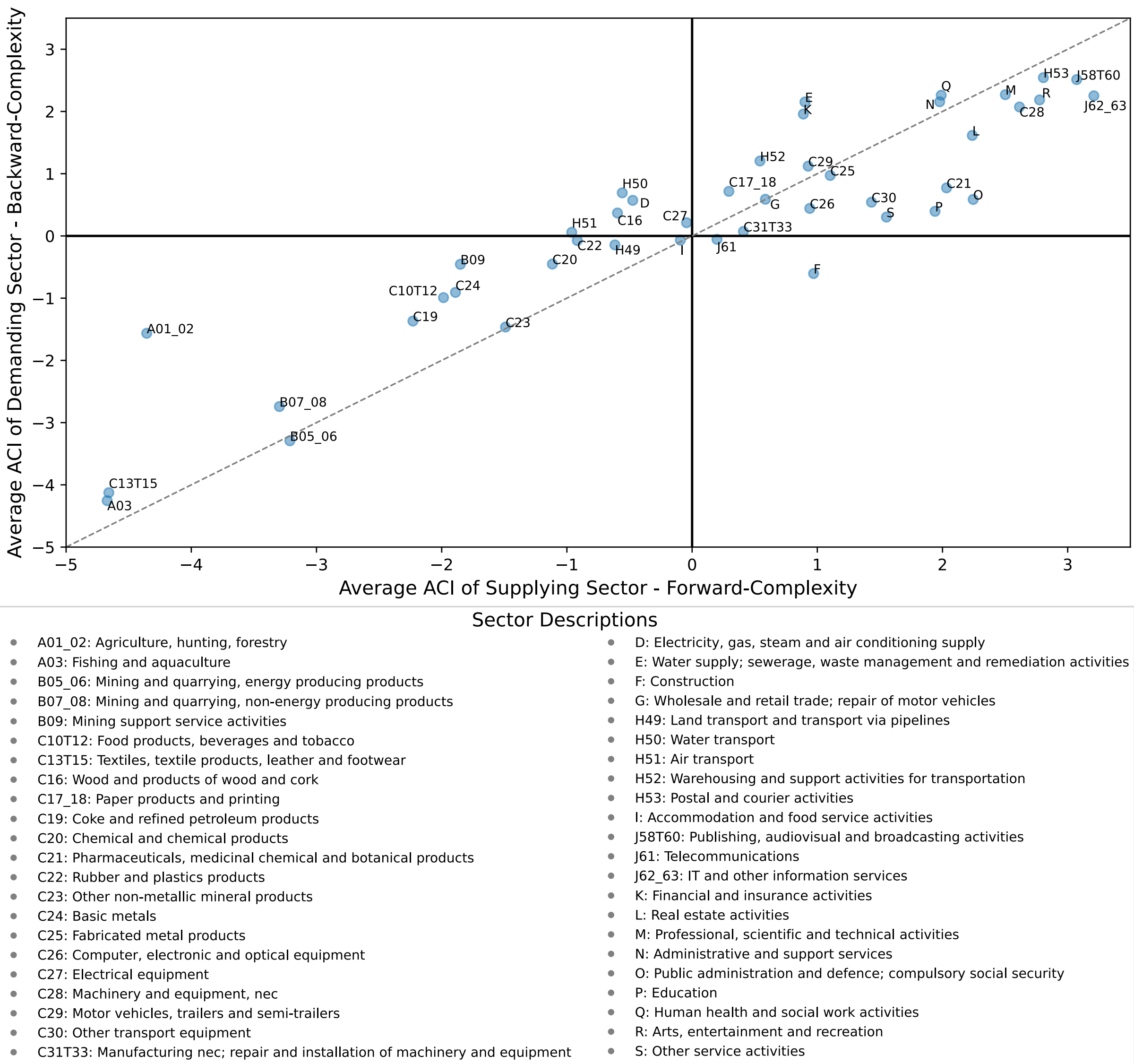


Figure 4.4 presents the average ACI for each sector when acting as an input demander (vertical axis) versus an input supplier (horizontal axis), based on input-output linkages in 2010. Each point represents a sector, labeled by its code, with a full description provided below the plot.

capabilities across the economy.

The structural patterns in Figure 4.4 reveal a clear hierarchy of sectoral complexity. At the bottom end of the spectrum, we find primary sectors and related services, which tend to exhibit the lowest average complexity, given their limited interdependence with other high-complexity activities. Interestingly, the agriculture sector (**A01_02**) displays greater complexity when acting as a demander of inputs rather than as a supplier. Among manufacturing industries, those most closely tied to primary activities, such as food products (**C10T12**), petroleum products (**C19**), non-metallic mineral products (**C23**), and basic metals (**C24**), also rank low in both demand- and supply-side complexity. Additionally, Textiles (**C13T15**) stands out as the least complex manufacturing sector by a wide margin.

On the opposite end, knowledge-intensive service sectors such as publishing and audiovisual (**J58T60**) and information technology (**J62_63**) top the complexity ranking. These sectors show slightly higher complexity on the supply side, indicating their role in feeding capabilities into other high-value-added segments of the economy, possibly linked to advanced manufacturing. Just behind them, we observe other high-end services, such as professional, scientific, and technical activities (**M**), reinforcing the importance of such capabilities in modern productive structures. The financial sector (**K**) illustrates a distinct asymmetry in its complexity profile. As a demander of inputs, it exhibits higher backward-oriented complexity, relying heavily on sophisticated services such as information technology, legal expertise, and professional consulting to deliver financial solutions. In contrast, its role as an input supplier reflects its more traditional function, which is providing capital to other sectors. While essential to the functioning of the broader economy, this upstream role appears less embedded in complex intersectoral structures, reflecting lower forward-complexity in its role as an input provider.

A distinct cluster of sectors positioned in the upper-right quadrant of Figure 4.4 comprises services typically provided publicly or associated with strong externalities, such as public administration and defense (**O**), education (**P**), and human health and social work activities (**Q**). These sectors exhibit positive average complexity both as input suppliers and as demanders, indicating their embeddedness in diverse and interdependent value chains. While not traditionally viewed as drivers of productive upgrading, these services play an essential supporting role for high-value manufacturing and knowledge-intensive activities. At the same time, their prominence in mature economies may also reflect the structural shift described by Baumol (1967), in which sectors with lower productivity growth gain relative importance as economies develop. A similar dynamic may apply to administrative services (**N**) and other service activities (**S**), which also appear in the first quadrant. These sectors may not lead complexity, but they help sustain it, suggesting a nuanced interplay between support functions, systemic externalities, and the evolving fabric of economic sophistication.

In general, manufacturing sectors occupy the middle of the complexity spectrum. Yet this middle is far from uniform. Among them, the motor vehicles industry (**C28**) stands out as the most complex in both its input and output roles. Other manufacturing sectors show marked asymmetries. The pharmaceutical industry (**C21**), for instance, exhibits high average complexity as a supplier of inputs, but relatively lower complexity in its demand structure. A similar pattern emerges for electronics (**C26**) and other transport equipment (**C30**), where complexity is concentrated on the supply side. Perhaps the most pronounced case of sectoral imbalance appears in the construction sector (**F**), which displays a forward-oriented complexity profile: its average complexity as a supplier is close to 1, while its complexity as a demander hovers near -1. This makes construction a candidate for high forward-complexity: it contributes to sophisticated production processes while depending on a relatively uncomplex input base.

Several manufacturing industries occupy the middle range of the complexity spectrum. Sectors such as chemicals (**C20**), plastics (**C22**), wood products (**C16**), electrical equipment (**C27**), paper and printing (**C17_18**), and furniture and other manufacturing (**C31T33**) exemplify this intermediate zone. Rather than acting as complexity attractors, these industries function as connective tissue in the productive structure, bridging upstream and downstream flows with moderate, yet indispensable, complexity. A similar role is played by infrastructure-oriented services such as air transport (**H51**), water transport (**H50**), electricity and gas supply (**D**), and wholesale and retail trade (**G**). These sectors serve as functional enablers of the economy, facilitating the circulation of goods, energy, and information. While not highly complex in themselves, their importance lies in enabling complex activities elsewhere, anchoring the invisible logistics and infrastructure backbone that holds the economic system together. In a system of interdependence, these sectors may not stand at the peak of complexity, but they make the productive core possible.

This perspective challenges a common inference drawn from trade-based complexity data, where manufacturing sectors often appear as the pinnacle of productive sophistication. When we shift the lens to input-output structures, however, a different picture emerges. Here, manufacturing is no longer the final destination of economic upgrading. It becomes a conduit. Its value lies less in what it produces in isolation and more in how it integrates, connects, and enables other sectors to accumulate and recombine knowledge. In this view, complexity is not embedded in manufacturing per se, but in the web of interdependencies it supports. Rather than being the end goal, industry becomes a means, an essential scaffold in the architecture of input-output economic complexity.

To better understand the differences between IO ECI and Trade ECI, as well as how these measures jointly capture the sophistication of a country's productive structure and its relationship with income levels, we present Figure 4.5.

Panels A and B reveal how different notions of economic complexity relate to income. Both IO ECI and Trade ECI are strongly associated with the log of GDP per capita, with

Figure 4.5 – IO ECI, Trade ECI and GDP per capita - 2010

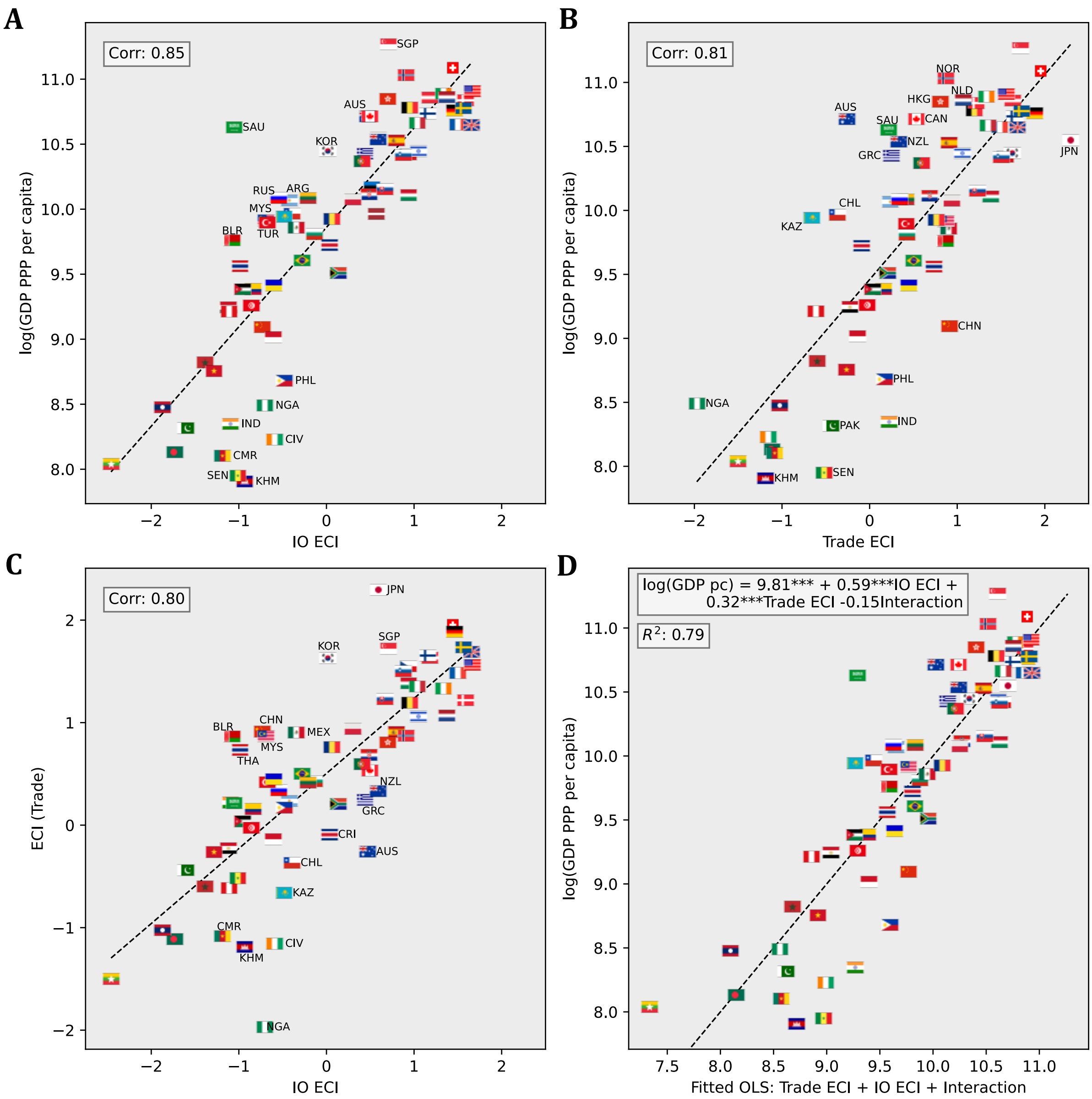


Figure 4.5 explores the relationship between economic complexity and income per capita in 2010. The **A** and **B** panels relate log GDP per capita (PPP, constant 2021 international dollars) to IO ECI and Trade ECI, respectively. **C** shows the relation between the two complexity measures. **D** compares observed and fitted values from an OLS model including IO ECI, Trade ECI, and their interaction, with robust standard errors (HC3). In all panels, dashed lines represent fitted values. Country codes are shown in the first three panels for observations with standardized residuals ≥ 1.

correlations of 0.85 and 0.81. But what matters is not just the strength of the correlation — it's the shape of the deviation. Some countries fall neatly along the fitted line, as if their income levels were simply a function of their productive sophistication. Others resist this expectation.

China and Japan are two such cases. When measured by Trade ECI, they appear richer in productive knowledge than in income, with their merchandise export baskets suggesting a level of prosperity not yet reflected in GDP. But IO ECI tells a different story. Once productive knowledge embedded in the intersectoral linkages of production are accounted for, both countries align more closely with their income, suggesting that value-added complexity, not just merchandise export sophistication, matters for economic development. New Zealand offers another perspective. Its Trade ECI places it well below the fitted line, reflecting a narrow and low-complexity export profile. Yet its IO ECI pulls it back toward the expected path. This adjustment reflects the country's input-output productive capabilities, particularly in high value-added services, that are hidden from merchandise trade data alone.

These contrasting patterns strengthen the case for a multidimensional view of economic complexity. Trade ECI and IO ECI do not always tell the same story, but that is precisely what makes their combination so valuable. In some cases, the two measures balance each other out. In others, they reinforce one another, amplifying the distance between productive structure and observed income. Saudi Arabia is a case in point: both metrics place it far from the expected path, suggesting that its income is shaped by forces that lie beyond the scope of productive complexity. This duality, which is sometimes counterbalancing, sometimes compounding, is what gives the joint use of IO and Trade ECI its analytical power.

Panel C places the two complexity measures side by side. Countries that lie significantly above the fitted line export goods that are more complex than their input-output structures would suggest. These cases represent the most pronounced expressions of the factory economy model — nations deeply integrated into global value chains, yet with domestic production systems that remain relatively shallow, particularly in services. Their growth trajectories rely heavily on external demand, manufacturing specialization, and the dynamics of international production.

Below the line, the logic is reversed. Countries such as Australia, New Zealand, and Chile exhibit domestic production networks that are more complex than their export baskets imply. These economies tend to combine natural resource wealth with high value-added service sectors that are largely invisible to trade-based metrics. Costa Rica also falls into this category, though not because of natural endowments. Its position reflects a strategic choice to pursue development through the expansion of knowledge-intensive services.

What these patterns reveal is that no single measure of complexity can fully account for the structural nuances behind economic development. All complexity metrics are shaped by the data they draw from, and thus inherit their blind spots. Trade-based measures emphasize the physical goods that a country sells to the world, but overlook value-added aspects and economic processes that remain within its borders. Input-output measures capture the architecture of production, including services and domestic linkages, but miss aspects of external competitiveness encoded in trade. The divergence between these two perspectives is not a flaw, but a feature. It highlights the need for a multidimensional approach.

Panel D brings these two perspectives together. By combining IO ECI, Trade ECI, and their interaction in a single model, we are able to explain income levels with higher precision. Countries that deviate under one dimension are often realigned when the other is considered, and the interaction term captures complementarities that neither measure can reveal on its own. In this view, development is not the result of a single dimension of productive complexity, but of how different economic layers work in tandem. It is in the tension between IO ECI and Trade ECI that the deeper logic of their economic potential becomes visible. Crucially, the residuals that emerge from this multidimensional complexity model serve as structural signals. These residuals indicate whether a country income is ahead or behind its productive structure and form the basis for predicting the direction and intensity of future economic growth.

In the next section, we operationalize these residuals within a two-stage empirical framework. By using them as inputs in growth regressions, we assess whether countries with income levels above or below what their productive complexity would predict tend to grow faster or slower over time. We extend this approach to test structural misalignments not only in terms of complexity, but also with respect to institutional quality and human capital. This allows us to evaluate whether gaps between income and foundational underlying capabilities contain predictive power for long-term development trajectories.

4.5 Two-Stage Structural Growth Framework

This section introduces the empirical framework through which we test whether multidimensional complexity, captured by the combination of Trade ECI and IO ECI, contributes to sustained economic growth. The central question is whether countries whose current income falls short of what their economic complexity would predict tend to grow faster over time, as they converge toward the income levels implied by their productive capabilities. To answer this, we place economic complexity on equal footing with two other foundational capabilities widely acknowledged in the development literature: human capital and institutional quality. Rather than treating these factors merely as controls, we conceptualize all three — complexity, institutions, and human capital — as structural conditions that shape a country's growth potential. The framework is designed to assess how each of these dimensions helps explain long-run growth trajectories.

The empirical strategy begins with a first-stage estimation designed to quantify the structural alignment between foundational capabilities and income. For each of the three vectors—multidimensional economic complexity, institutional quality, and human capital—we estimate cross-sectional regressions of log GDP per capita on the respective indicators, using data for a given base year. The residuals from these regressions capture the degree of misalignment between a country's level of income and what would be expected given its capabilities. A positive residual suggests that income exceeds the level predicted by a particular structural dimension, while a negative residual indicates that the country's productive, institutional, or human capital base could support a higher level of income. These residuals are interpreted as structural signals, and serve as the core inputs for the second stage. In that stage, we test whether such misalignments systematically predict future economic growth, assessing the extent to which countries converge toward the income levels implied by their underlying capabilities.

To implement the second stage, we rely on two distinct panel structures that allow us to estimate two-way fixed effects models while preserving temporal and cross-sectional variation. The first set of panels follows ten-year growth periods starting every three years — for example, 1999-2009, 2002-2012, 2005-2015, and 2008-2018. This overlapping design captures long-run growth dynamics while maintaining a reasonable number of periods to estimate two-way fixed effects. As a robustness check, we also estimate models based on five-year, non-overlapping windows, such as 1999-2004, 2004-2009, 2009-2014, and 2014-2019, which avoid intertemporal dependence at the cost of a shorter growth horizon.

In what follows, we focus on the overlapping design using the panel that begins in 2000. Accordingly, the first-stage regressions presented in Table 4.1 correspond to the base years 2000, 2003, 2006, and 2009, which define the temporal anchors for the second-stage estimation. For each year, we regress log GDP per capita on one vector of foundational capabilities at a time — multidimensional economic complexity, institutional quality, and human capital — allowing us to assess how strongly each dimension explains cross-sectional variation in income. The estimated coefficients capture the average structural relationship between capabilities and income at each point in time, while the residuals from these regressions form the basis for the misalignment measures used in the second stage.

The results in Table 1 reveal that all three vectors of foundational capabilities are significantly associated with income per capita, though with varying strength across base years. Within the economic complexity vector, both Trade ECI and IO ECI display positive and statistically significant coefficients in most years. This suggests that each dimension of complexity captures a distinct and relevant channel through which productive structure shapes development. The interaction term between the two indices is also consistently significant and negative, implying diminishing marginal returns. For human capital, the coefficient on log HCI is positive, stable, and highly significant across all years, reinforcing its well-documented role in explaining income variation. Among the institutional indicators,

Table 4.1 – First-Stage Regressions - Foundational Capabilities and Income per Capita

		Dependent variable: GDP per capita (PPP)			
Vector	Variable	2000	2003	2006	2009
ECI	Trade ECI	0.559***	0.449***	0.263	0.258*
		(0.000)	(0.000)	(0.191)	(0.053)
	Input-Output ECI	0.416**	0.543***	0.631***	0.683***
		(0.014)	(0.000)	(0.002)	(0.000)
	Interaction Term	-0.166**	-0.217***	-0.177***	-0.217***
		(0.015)	(0.000)	(0.005)	(0.000)
	Adj. R^2	0.739	0.76	0.703	0.738
Institutional - WGI	Control of Corruption (CC)	0.020	0.007	0.134	-0.200
		(0.943)	(0.979)	(0.448)	(0.426)
	Government Effectiveness	0.417	0.129	0.286	0.262
	(GE)	(0.242)	(0.714)	(0.478)	(0.418)
	Political Stability and	0.267*	0.142	0.250**	0.213**
	Absence of Violence (PV)	(0.089)	(0.219)	(0.019)	(0.032)
	Rule of Law (RL)	-0.219	0.026	-0.322	0.031
		(0.580)	(0.954)	(0.542)	(0.938)
	Regulatory Quality (RQ)	0.409*	0.631	0.535*	0.645
		(0.088)	(0.114)	(0.076)	(0.112)
	Voice and Accountability	0.090	-0.022	0.038	-0.031
	(VA)	(0.758)	(0.921)	(0.881)	(0.889)
	Adj. R^2	0.732	0.763	0.756	0.759
Human Capital	log Human Capital Index	2.971***	3.104***	3.225***	3.250***
		(0.000)	(0.000)	(0.000)	(0.000)
	Adj. R^2	0.616	0.646	0.672	0.708

Table 1 reports the results of cross-sectional OLS regressions for four base years, in which log GDP per capita is regressed separately on each foundational capability vector: multidimensional economic complexity, human capital, and institutional quality. Each coefficient is reported alongside its HC3 p-value in parentheses. Asterisks denote statistical significance: p*** < 0.01, p** < 0.05, and p* < 0.10. Adjusted R^2 values are shown at the bottom of each block. The sample includes 63 countries observed consistently across all four years, forming a balanced panel for the second-stage growth analysis.

regulatory quality and rule of law emerge as the most consistent predictors, while other dimensions such as voice and accountability or political stability show more erratic results.

Adjusted R^2 values remain relatively stable over time within each vector, always exceeding 0.6, which confirms the strong structural alignment between capabilities and income. However, these correlations are far from perfect and substantial variation remains unexplained. It is precisely this residual variation, interpreted as a measure of structural misalignment, that we carry forward to the second-stage analysis in order to test whether such gaps predict future economic growth.

Having established the rationale for structural residuals through first-stage regressions, we now turn to the second stage to examine whether these misalignments predict future economic growth. Table 4.2 presents the set of models estimated using two-way fixed-effects OLS panel regressions, where the dependent variable is the ten-year annualized growth rate of GDP per capita (PPP). Regressions are conducted on a balanced panel of 63 countries, using four partially overlapping base years: 2000, 2003, 2006, and 2009. Standard errors are clustered by country, and p-values (reported in parentheses) are computed using the WCU-31 wild cluster bootstrap. This framework allows us to test whether countries systematically converge toward the income levels implied by their productive, institutional, and human capital structural foundations.

Table 4.2 – Second-Stage Regressions: Foundational Capabilities and GDP Growth

	Dependent variable: 10-year annualized GDP per capita growth (PPP)								
	1	2	3	4	5	6	7	8	9
log NRR	-0.025	-0.061	-0.039	-0.042	-0.089	-0.095	-0.180	-0.233	-0.186
	(0.877)	(0.746)	(0.496)	(0.863)	(0.639)	(0.559)	(0.468)	(0.473)	(0.461)
ε^{ECI}	-2.203***			-1.305**	-4.189***	-3.174***	-2.700***	-2.632***	-2.577***
	(0.005)			(0.038)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ε^{HCI}		-4.744***		-4.083***		-2.964***	-2.491***	-2.675***	-2.445***
		(0.000)		(0.000)		(0.004)	(0.005)	(0.003)	(0.007)
ε^{INST}			-0.808	-0.373		-0.303	-1.831***	-1.894***	-1.844***
			(0.131)	(0.406)		(0.446)	(0.000)	(0.001)	(0.001)
Δ Trade $ECI_{t10,t}$					15.426*	11.612	8.693	8.296	8.289
					(0.071)	(0.192)	(0.146)	(0.156)	(0.170)
Δ IO $ECI_{t10,t}$					18.641***	15.735***	13.972***	13.641***	14.791***
					(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Δ WGI $PV_{t10,t}$							4.490***	4.506***	4.618***
							(0.006)	(0.005)	(0.004)
Δ WGI $RQ_{t10,t}$							13.097***	12.553***	13.625***
							(0.002)	(0.006)	(0.002)
Δ log $HCI_{t10,t}$								42.686	
								(0.136)	
Interaction									-62.309
(Δ Trade ECI, Δ IO ECI)									(0.282)
Adj. R^2	0.790	0.818	0.768	0.826	0.840	0.856	0.883	0.886	0.883
R^2 within	0.109	0.229	0.013	0.267	0.326	0.403	0.521	0.533	0.524

Table 4.2 presents all models tested using two-way fixed-effects OLS panel regressions, where the dependent variable is the ten-year annualized growth rate of GDP per capita (PPP). Regressions are estimated on a balanced panel of 63 countries using four partially overlapping base years: 2000, 2003, 2006, and 2009. P-values, in parentheses, are computed using the WCU-31 wild cluster bootstrap. Significance levels: $p^{***} < 0.01$, $p^{**} < 0.05$, $p^{*} < 0.10$.

The second-stage results begin with a straightforward test of the predictive power of each structural residual in isolation. Models 1 through 3 include one residual at a time — economic complexity, human capital, and institutions — together with a control for the share of natural rents in GDP. All coefficients are negative, but only the residuals for economic complexity and human capital are statistically significant at the 1% level. When all three misalignments are included simultaneously in Model 4, the pattern persists: only complexity and human capital residuals remain significant, and all three coefficients retain their expected negative signs. This is fully consistent with the theoretical logic underpinning the framework. A positive residual indicates that a country's income is higher than what its capabilities would suggest; in such cases, we expect slower subsequent growth. Conversely, a negative residual signals latent potential—a structural foundation that exceeds current income—and is therefore associated with faster convergence.

Models 5 and 6 extend this logic by adding the ten-year variation in complexity measures Trade ECI and IO ECI. These dynamics are critical to include, as residuals may correlate with the evolution of structural conditions over time. Controlling for these changes helps isolate the predictive role of the misalignment itself. It also offers a glimpse into the co-evolutionary process linking capabilities and income growth. In both models, the variation terms enter with positive coefficients, but only IO ECI variation remains significant once human capital and institutional residuals are included (Model 6). The pattern of significance among the residuals remains consistent with Model 4, reaffirming the salience of complexity and human capital in shaping long-run trajectories.

Model 7 introduces institutional dynamics by including the only two components of the Worldwide Governance Indicators that showed statistical relevance in the first-stage regressions: changes in regulatory quality (RQ) and voice and accountability (VA). Once these institutional variations are accounted for, the residual capturing structural misalignment in institutions becomes statistically significant for the first time—suggesting that its previous insignificance may have stemmed from omitted dynamic effects. Both WGI change variables are also significant and carry the expected positive signs, reinforcing the idea that improvements in institutional quality contribute meaningfully to growth. In Model 8, we include the ten-year change in the log of HCI. While the coefficient displays the expected sign, it is not statistically significant. Finally, Model 9 adds an interaction term between the changes in Trade ECI and IO ECI to test for potential non-linear complexity co-movement. However, this term also remains statistically insignificant.

Model 7 delivers one of the most consequential findings of this study. Once institutional dynamics are properly accounted for, all three vectors of foundational capabilities emerge as significant predictors of long-term growth. This result offers compelling evidence that development is not driven by any single structural dimension, but by the joint influence of productive sophistication, human capital formation, and institutional effectiveness.

Multidimensional economic complexity captures the depth and tacit structure of knowledge embedded in a country's productive fabric — what it is capable of building, often without being able to codify it fully. This knowledge enables specialization and differentiation, but its value depends on how well it is mobilized. Human capital, in turn, provides the cognitive foundation and general-purpose skills that enhance productivity, allowing countries to make better use of their existing structures and to adapt more quickly to new opportunities. Institutions shape the environment in which both productive knowledge and human capital are deployed. By reducing transaction costs, providing stability, and aligning incentives, they amplify the effectiveness of the other two pillars, turning capabilities into actual engines of economic transformation.

This configuration constitutes what we call the **three pillars of economic convergence**. While some countries may reach high levels of development despite weaknesses in one of these dimensions, doing so is the exception rather than the rule. For most nations, sustained and inclusive growth becomes far more attainable when economic complexity, human capital, and institutional quality advance together, each reinforcing and stabilizing the others. Public policies that address these three domains in an integrated way offer not just complementary support, but a coherent foundation for long-term economic transformation.

To assess the robustness of our second-stage findings, we replicate the full specification of Model 7 across five alternative panel structures — three based on overlapping ten-year windows and two using non-overlapping five-year periods. These panel designs vary both the base years and the growth horizon, allowing us to evaluate the consistency of our findings across different temporal configurations. This exercise ensures that our conclusions are not an artifact of a particular panel construction, and provides a broader test of the model's stability.

In addition to this robustness check, we use the same five panels to test interaction terms between each pair of structural residuals, with the goal of identifying potential complementarities between the three foundational capabilities. Since only the interaction between the institutional and human capital residuals achieved statistical significance in at least two of the five panels, we add this specific interaction term to Model 7 and estimate it across all five panel designs. Results are reported in Table 4.3.

Just as macroeconomists use the concept of an output gap to capture the short-run distance between actual and potential output, the structural residuals in our framework can be interpreted as long-term income gaps relative to a country's productive, institutional, and cognitive capabilities. Rather than reflecting cyclical fluctuations, these residuals signal whether a country is operating below the income level implied by its foundational characteristics. When the residual is negative, it indicates that the country has more capability than is reflected in its current income—a form of untapped long-run potential that may lead to faster growth as this gap closes.

The results in Table 4.3 provide strong empirical support for this interpretation. Across all five panel specifications — with varying base years and growth horizons — the structural residuals for multidimensional economic complexity, human capital, and institutional quality are consistently negative and statistically significant. In other words, countries whose income lags behind what their capabilities would predict tend to grow faster, as they close the gap between realized and potential income. The magnitude and consistency of these effects across all three vectors reinforce the core premise of the framework: long-term growth is shaped not only by the level of foundational capabilities, but also by the distance that separates a country from fully translating those capabilities into economic outcomes.

The inclusion of an interaction term between the institutional and human capital residuals adds an additional layer of nuance. Although this term is not significant in every specification, it reaches statistical significance in two of the five panels, always with a

Table 4.3 – Panel Comparison - Second-Stage Results

Dependent variable:	10-year growth			5-year growth	
Base years:	1999, 2002, 2005, 2008	2000, 2003, 2006, 2009	2001, 2004, 2007, 2010	1999, 2004, 2009, 2014	2000, 2005, 2010, 2015
log NRR	-0.149	-0.183	-0.052	-0.221	-0.057
	(0.307)	(0.466)	(0.678)	(0.328)	(0.806)
ε^{ECI}	-2.939***	-2.684***	-2.252***	-1.801***	-1.225*
	(0.001)	(0.000)	(0.001)	(0.010)	(0.089)
ε^{HCI}	-2.622***	-2.479***	-2.972**	-2.311**	-1.674*
	(0.004)	(0.003)	(0.014)	(0.033)	(0.053)
ε^{INST}	-1.043**	-1.814***	-1.293*	-2.042**	-2.133**
	(0.015)	(0.001)	(0.065)	(0.040)	(0.020)
Interaction	-1.550***	-1.058	-0.441	-2.845**	-1.452
(ε^{HCI}, ε^{INST})	(0.010)	(0.282)	(0.876)	(0.043)	(0.220)
Δ Trade $ECI_{t10,t}$	7.314	8.356	10.197	5.455	8.038**
	(0.154)	(0.159)	(0.162)	(0.132)	(0.045)
Δ IO $ECI_{t10,t}$	10.87***	13.384***	8.325*	7.233***	7.139**
	(0.009)	(0.000)	(0.072)	(0.001)	(0.04)0
Δ WGI $PV_{t10,t}$	5.021***	4.126**	5.097***	2.748*	5.394***
	(0.01)	(0.012)	(0.002)	(0.058)	(0.002)
Δ WGI $RQ_{t10,t}$	12.654***	13.258***	4.793	12.180***	8.440**
	(0.002)	(0.002)	(0.299)	(0.000)	(0.016)
Adj. R^2	0.892	0.886	0.874	0.613	0.626
R^2 within	0.584	0.534	0.405	0.374	0.284
N (countries)	63	63	65	63	63

Table 4.3 compares the full second-stage specification across different panel constructions. Each column reports the results of two-way fixed-effects regressions using a distinct set of base years and growth periods. The first three columns correspond to overlapping ten-year panels starting in 1999, 2000, and 2001, respectively. The last two columns present non-overlapping five-year panels beginning in 1999 and 2000. Significance levels: $p^{***} < 0.01$, $p^{**} < 0.05$, $p^{*} < 0.10$.

negative coefficient. Since the residuals are typically negative in growing countries, a negative interaction implies that the positive effect of a gap in one domain becomes weaker when a gap also exists in the other. This finding points to diminishing marginal returns across these two capabilities—suggesting that having unrealized potential in both areas does not yield additive growth benefits. Instead, the developmental gains from addressing one constraint may be less pronounced if the other constraint remains unresolved.

Beyond the structural residuals, dynamic changes in capabilities also play an important role. The ten-year variation in IO ECI is statistically significant in all five panel specifications, consistently exhibiting a positive coefficient, reinforcing the idea that complexity improvements in intersectoral productive structures support long-term growth. Although the change in Trade ECI reaches statistical significance in only one panel, its coefficient is positive across all models, suggesting a consistent directional effect despite limited statistical power. Changes in institutional quality also show robust associations with growth: regulatory quality (RQ) and political voice (PV) are statistically significant in most panels and carry positive coefficients throughout. Due to data limitations for the year 2020, it was not possible to include the ten-year change in the log of the Human Capital Index (HCI) in two of the five panel configurations. For consistency across specifications, this variable was therefore omitted from the models reported in Table 4.3. However, as shown in Table 4.2, it was tested using the ten-year growth panel with a base year starting in 2000. In that specification, the coefficient on the HCI variation was positive, as expected, but failed to reach statistical significance.

Taken together, these results suggest that foundational capabilities and economic growth may evolve jointly over time. Even after controlling for structural misalignments — i.e., the gaps between income and what each capability vector would predict — improvements in economic complexity and institutional quality remain positively associated with future income gains. This pattern implies that growth does not depend solely on catching up to structurally expected income levels, but is also reinforced by ongoing transformations in the underlying capabilities themselves. The evidence is consistent with a co-evolutionary dynamic in which countries not only grow because they have untapped potential, but also because they continue to build and strengthen that potential along the way.

The empirical findings presented above underscore the central role of foundational capabilities—productive complexity, human capital, and institutional quality—in shaping long-run economic growth. In the following section, we reflect on the broader implications of these results, considering how they inform our understanding of structural convergence and the design of development strategies.

4.6 Discussion

This study investigated how long-run economic growth is shaped by three foundational capabilities: productive (reflected in a country's economic complexity) institutional and human capital. We argue that these dimensions should be treated as equally fundamental structural pillars of development, rather than as secondary controls. To operationalize this idea, we developed a two-stage empirical framework. In the first stage, we measured how much each capability vector explains contemporary income levels. In the second, we tested whether the residuals from these regressions, interpreted as long-term income gaps relative to each capability, predict future growth. These residuals function as structural signals: when negative, they reveal the existence of untapped potential that may drive convergence over time.

Just as macroeconomists use the concept of an output gap to measure short-term inefficiencies, our framework offers a long-term analogue: a capability-income gap. Rather than reflecting cyclical slack, this gap captures a more persistent misalignment between what a country is and what it could achieve, given its structural endowments. This perspective offers a new diagnostic lens for understanding development, not in terms of singular constraints, but through the lens of multidimensional capacity and misalignment.

This structural interpretation is empirically supported by robust results. Countries with negative residuals — indicating that their income lags behind what their capabilities would suggest — tend to grow faster, consistent with a process of structural convergence. The results of the two-stage model consistently show that each of the three vectors is a significant predictor of long-term growth. Countries tend to converge toward the income levels implied by their structural capabilities, and this convergence is stronger when those capabilities are well-developed across all three dimensions. These findings reinforce the idea that sustained economic development is supported by what we call the three pillars of economic convergence. Efforts focused on only one domain, be it productive upgrading, institutional reform, or education, are unlikely to succeed in isolation. Instead, development policies must be coordinated across these three fronts, reinforcing one another in a mutually supportive process.

In this framework, productive capabilities are expressed through the notion of multidimensional economic complexity, which we capture by combining two complementary measures: the traditional Trade ECI and the newly proposed IO ECI. While Trade ECI reflects the knowledge embedded in a country's export basket, IO ECI complements it by capturing the intersectoral sophistication of also domestic production networks, incorporating both goods and services, and offering a value-added perspective that is increasingly relevant in service-driven economies. One of the key findings of this study is that IO ECI adds explanatory power beyond Trade ECI in accounting for current income levels, confirming its value as a complementary dimension of productive capabilities.

This broadened understanding of productive capabilities also invites a reappraisal of how we conceptualize the role of services in development. While much of the economic complexity literature has focused on manufacturing and tradable goods, recent contributions suggest that certain labor-absorbing services, particularly those embedded in industrial linkages, may offer viable pathways for productive upgrading in developing economies (Rodrik; Sandhu, 2024). These services, such as logistics, industrial maintenance, and technical support, can accumulate organizational and technical capabilities over time, generating both employment and structural transformation when effectively integrated into broader production systems. This perspective aligns with our emphasis on intersectoral complexity and reinforces the idea that development strategies should not be confined to high-tech or export-oriented activities alone. Instead, they should recognize and cultivate the upgrading potential of service sectors that combine absorptive capacity with the possibility of incremental sophistication.

The Input-Output ECI also sheds light on the asymmetric roles that sectors play in complex production systems. Drawing from Hirschman's notions of backward and forward linkages, we introduce the concepts of backward- and forward-complexity to capture whether a sector tends to absorb or transmit sophistication through the value chain. Some sectors, such as construction or finance, exhibit high complexity depending on whether they function primarily as input providers or as demanders. The most complex activities identified in this framework are concentrated in knowledge-intensive service sectors, such as information technology and professional business services, highlighting the growing importance of intangible capabilities and coordination in modern productive architectures.

Beyond explaining growth, this two-stage framework has diagnostic and normative potential. The residuals generated in the first stage provide a structured way to identify where a country's development potential is most latent. They allow policymakers to prioritize areas of structural capability investment, not based on stylized policy prescriptions, but on the actual distance between capabilities and outcomes. Moreover, while the model treats the three pillars symmetrically, the results suggest that their relative importance may vary across countries and over time. This opens a path to future research on the sequencing of interventions and the dynamics of structural interplay.

Several extensions are possible. One is to refine the multidimensional economic complexity vector itself, incorporating measures based on other datasets, such as patents or scientific research. Another is to explicitly model the interactions and co-evolution of capabilities: how institutional improvements enable productive upgrading, or how human capital expansion changes the returns to complexity. Finally, applying this framework at regional and subnational levels could help illuminate internal patterns of structural divergence and convergence.

Appendix 4.1: List of Countries

Figure 4A.1 – List of Countries

Country Code	Country	Country Code	Country
ARG	Argentina	JPN	Japan
AUS	Australia	KAZ	Kazakhstan
AUT	Austria	KHM	Cambodia
BEL	Belgium	KOR	Korea
BGD	Bangladesh	LAO	Lao (People's Democratic Republic)
BGR	Bulgaria	LTU	Lithuania
BLR	Belarus	LUX	Luxembourg
BRA	Brazil	LVA	Latvia
BRN	Brunei Darussalam	MAR	Morocco
CAN	Canada	MEX	Mexico
CHE	Switzerland	MLT	Malta
CHL	Chile	MMR	Myanmar
CHN	China (People's Republic of)	MYS	Malaysia
CIV	Côte d'Ivoire	NGA	Nigeria
CMR	Cameroon	NLD	Netherlands
COL	Colombia	NOR	Norway
CRI	Costa Rica	NZL	New Zealand
CYP	Cyprus	PAK	Pakistan
CZE	Czechia	PER	Peru
DEU	Germany	PHL	Philippines
DNK	Denmark	POL	Poland
EGY	Egypt	PRT	Portugal
ESP	Spain	ROU	Romania
EST	Estonia	RUS	Russian Federation
FIN	Finland	SAU	Saudi Arabia
FRA	France	SEN	Senegal
GBR	United Kingdom	SGP	Singapore
GRC	Greece	SVK	Slovakia
HKG	Hong Kong, China	SVN	Slovenia
HRV	Croatia	SWE	Sweden
HUN	Hungary	THA	Thailand
IDN	Indonesia	TUN	Tunisia
IND	India	TUR	Türkiye
IRL	Ireland	TWN	Chinese Taipei
ISL	Iceland	UKR	Ukraine
ISR	Israel	USA	United States
ITA	Italy	VNM	Viet Nam
JOR	Jordan	ZAF	South Africa

Excluded from complexity calculations after data cleaning.

Figure 4A.1 lists the countries used for the input-output economic complexity exercise.

Appendix 4.2: Supplementary Regressions

This appendix reports supplementary regressions conducted as robustness checks to support the main findings of the paper. The results are organized into four sections, each addressing a specific dimension of model robustness. First, we present regressions that incorporate changes in the Human Capital Index (HCI), testing the sensitivity of results to the inclusion of human capital dynamics. Second, we report models estimated without the Input-Output Economic Complexity Index (IO ECI), using both restricted and expanded country samples. Third, we provide regressions using an aggregated residual vector that combines all capability components into a single structural gap, simplifying the interpretation of misalignment. Finally, we include specifications that test interaction terms between the residuals of economic complexity, human capital, and institutions, exploring potential complementarities between these structural drivers of growth.

A. Regressions with HCI Change Included

This section presents regressions that incorporate changes in the Human Capital Index (HCI), testing the sensitivity of results to the inclusion of human capital dynamics in explaining growth trajectories.

The main second-stage results presented in Table 4.3 compare growth regressions across different panel constructions. However, that specification did not include the change in HCI as a control because two of the panels cover time spans for which HCI data are not available — the HCI dataset (Penn World Table 10.01) extends only up to 2019. As a result, including HCI change in those panels would significantly reduce the sample and distort comparability across the different constructions.

To assess how the model behaves when HCI change is included, we estimate regressions for the three panel configurations where complete data are available. The results on Table 4A.1 confirm that HCI change contributes to explaining growth trajectories, and importantly, the inclusion of this variable does not alter the significance or direction of the coefficients associated with the structural residual vectors. This reinforces the robustness of the main findings and highlights the complementary role of human capital dynamics in the structural convergence process.

B. Regressions without Input-Output ECI

This section reports models estimated without the IO ECI, using both restricted and expanded country samples. These specifications assess the added explanatory power provided by IO ECI relative to trade-based measures alone.

It is important to test the model without using the IO ECI, both in constructing the resid-

Table 4A.1 – Panel Comparison - Second-Stage Results with HCI Change

Dependent variable:	10-year growth		5-year growth
Base years:	1999, 2002, 2005, 2008	2000, 2003, 2006, 2009	1999, 2004, 2009, 2014
log NRR	-0.210	-0.235	-0.258
	(0.188)	(0.476)	(0.355)
ε^{ECI}	-2.805***	-2.618***	-1.813***
	(0.001)	(0.000)	(0.010)
ε^{HCI}	-2.929***	-2.657***	-2.283**
	(0.002)	(0.002)	(0.036)
ε^{INST}	-1.009**	-1.876***	-2.038**
	(0.013)	(0.001)	(0.044)
Interaction	-1.435***	-1.024	-2.798**
$(\varepsilon^{HCI}, \varepsilon^{INST})$	(0.009)	(0.289)	(0.036)
Δ Trade $ECI_{t10,t}$	7.229	7.984	5.717
	(0.119)	(0.161)	(0.115)
Δ IO $ECI_{t10,t}$	11.054***	13.083***	7.318***
	(0.005)	(0.000)	(0.002)
Δ WGI $PV_{t10,t}$	52.842**	41.144	17.666
	(0.037)	(0.125)	(0.554)
Δ WGI $RQ_{t10,t}$	5.071***	4.153***	2.735*
	(0.009)	(0.009)	(0.058)
Δ $HCI_{t10,t}$	12.035***	12.729***	12.199***
	(0.003)	(0.005)	(0.000)
Adj. R^2	0.896	0.888	0.612
R^2 within	0.601	0.546	0.376
N (countries)	63	63	63

Table 4A.1 compares the full second-stage specification across different panel constructions, with HCI change included. Each column reports the results of two-way fixed-effects regressions using a distinct set of base years and growth periods. The first two columns correspond to overlapping ten-year panels starting in 1999 and 2000, respectively. The last column present a non-overlapping five-year panel beginning in 1999. Significance levels: $p^{***} < 0.01$, $p^{**} < 0.05$, $p^{*} < 0.10$.

ual vector for productive complexity and in accounting for IO ECI change in the second-stage regressions. This is essential to evaluate whether the IO ECI indeed provides complementary information to the model or whether it is redundant relative to existing international trade-based complexity measures. To this end, we estimate specifications relying exclusively on Trade ECI, thereby abandoning the multidimensional approach to productive complexity.

First, in Table 4A.2, we keep exactly the same country sample as in Table 4.3, ensuring a precise comparison of the impact of excluding IO ECI. The results show that the model's within R-squared weakens, indicating a loss of explanatory power. Nevertheless, the significance of the structural residuals remains robust overall, including the residual vector based on Trade ECI. Interestingly, Trade ECI change becomes statistically significant in this specification, "taking over" the role played by IO ECI change in the full model. These results reinforce the complementary value of IO ECI in capturing dimensions of productive complexity beyond trade, while also confirming that the model retains reasonable predictive power even when IO ECI is omitted.

Table 4A.2 – Panel Comparison - Second-Stage Results withouth IO ECI - Same Country Sample

Dependent variable:	10-year growth			5-year growth	
Base years:	1999, 2002, 2005, 2008	2000, 2003, 2006, 2009	2001, 2004, 2007, 2010	1999, 2004, 2009, 2014	2000, 2005, 2010, 2015
log NRR	-0.156	-0.185	-0.081	-0.259	-0.052
	(0.220)	(0.356)	(0.469)	(0.231)	(0.822)
ε^{ECI}	-3.11***	-2.815***	-2.573***	-2.275***	-1.648*
	(0.000)	(0.000)	(0.003)	(0.004)	(0.063)
ε^{HCI}	-2.585***	-2.522***	-2.975***	-2.229***	-1.611**
	(0.001)	(0.002)	(0.009)	(0.010)	(0.022)
ε^{INST}	-1.028**	-1.647***	-1.077	-1.798*	-1.937**
	(0.018)	(0.002)	(0.165)	(0.068)	(0.044)
Interaction	-1.110**	-0.684	0.105	-2.084*	-1.177
(ε^{HCI}, ε^{INST})	(0.028)	(0.652)	(0.997)	(0.085)	(0.328)
Δ Trade $ECI_{t10,t}$	19.139***	17.319**	19.802**	10.285**	10.848**
	(0.000)	(0.050)	(0.034)	(0.024)	(0.034)
Δ WGI $PV_{t10,t}$	4.212**	3.659*	4.387**	2.972**	5.444***
	(0.045)	(0.067)	(0.021)	(0.036)	(0.004)
Δ WGI $RQ_{t10,t}$	12.548***	12.723***	3.136	11.541***	8.137**
	(0.001)	(0.004)	(0.471)	(0.000)	(0.016)
Adj. R^2	0.891	0.876	0.868	0.615	0.625
R^2 within	0.575	0.490	0.375	0.374	0.278
N (countries)	63	63	65	63	63

Table 4A.2 reports an alternative second-stage specification estimated across different panel constructions. In this specification, the IO ECI variable is excluded from the productive vector residuals, and the IO ECI change is omitted from the second-stage regressions. The set of countries is kept constant across panels relative to the full specification. Each column reports the results of two-way fixed-effects regressions using a distinct set of base years and growth periods. The first three columns correspond to overlapping ten-year panels starting in 1999, 2000, and 2001, respectively. The last two columns present non-overlapping five-year panels beginning in 1999 and 2000. Significance levels: $p^{***} < 0.01$, $p^{**} < 0.05$, $p^{*} < 0.10$.

In Table 4A.3, we test the same model but without restricting the country sample to those with available IO ECI data. This expands the set of countries to a range between 84 and 89, depending on the panel tested. The residual vector for productive complexity remains statistically significant only for the first two panels, and although it loses significance in the broader panels, the coefficient signs consistently remain negative, aligning with the structural convergence mechanism proposed in the main model. Furthermore, Trade ECI change loses its significance entirely in these broader samples. These findings suggest that the results of the model are sensitive to the composition of the sample, reinforcing the importance of considering structural differences across countries when applying capability-based growth models. At the same time, the findings highlight the importance of exploring heterogeneity in structural convergence processes across different groups of countries, providing a potential avenue for future research.

Table 4A.3 – Panel Comparison - Second-Stage Results withouth IO ECI - Unrestricted Country Sample

Dependent variable:	10-year growth			5-year growth	
Base years:	1999, 2002, 2005, 2008	2000, 2003, 2006, 2009	2001, 2004, 2007, 2010	1999, 2004, 2009, 2014	2000, 2005, 2010, 2015
log NRR	-0.093	-0.032	0.042	-0.060	0.117
	(0.416)	(0.882)	(0.712)	(0.793)	(0.659)
ε^{ECI}	-1.592**	-1.495*	-0.358	-1.514	-1.033
	(0.010)	(0.063)	(0.628)	(0.113)	(0.265)
ε^{HCI}	-5.216***	-4.332***	-4.890***	-2.946**	-3.085**
	(0.001)	(0.002)	(0.000)	(0.030)	(0.012)
ε^{INST}	-0.831**	-1.595***	-1.044**	-2.252***	-1.556**
	(0.021)	(0.008)	(0.023)	(0.003)	(0.033)
Interaction	-0.408	-0.130	-0.090	0.000	-0.733
$(\varepsilon^{HCI}, \varepsilon^{INST})$	(0.595)	(0.881)	(0.950)	(1.000)	(0.862)
Δ Trade $ECI_{t10,t}$	2.606	3.207	0.980	2.669	4.693**
	(0.573)	(0.411)	(0.837)	(0.432)	(0.011)
Δ WGI $PV_{t10,t}$	6.165***	5.024**	6.061**	4.617***	10.326***
	(0.001)	(0.010)	(0.013)	(0.000)	(0.003)
Δ WGI $RQ_{t10,t}$	12.385***	10.639**	5.410	13.161***	0.030
	(0.004)	(0.019)	(0.107)	(0.000)	(0.981)
Adj. R^2	0.866	0.843	0.840	0.524	0.555
R^2 within	0.621	0.518	0.409	0.346	0.278
N (countries)	84	86	89	84	87

Table 4A.3 reports an alternative second-stage specification estimated across different panel constructions. In this specification, the IO ECI variable is excluded from the productive vector residuals, and the IO ECI change is omitted from the second-stage regressions. The set of countries is no longer restricted by the availability of IO ECI data. Each column reports the results of two-way fixed-effects regressions using a distinct set of base years and growth periods. The first three columns correspond to overlapping ten-year panels starting in 1999, 2000, and 2001, respectively. The last two columns present non-overlapping five-year panels beginning in 1999 and 2000. Significance levels: p*** < 0.01, p** < 0.05, p* < 0.10.

C. Regressions with Aggregated Structural Residual

In this section, we test an aggregated residual vector by consolidating all structural variables — economic complexity, human capital, and institutions — into a single first-stage regression, as reported in Table 4A.4. This approach simplifies the model by generating a unified structural gap that captures the joint explanatory power of these key capabilities.

The first-stage regression shows a stronger fit to current GDP per capita when these variables are combined into an aggregated vector, as reflected by the higher adjusted R-squared in Table 4A.4 compared to Table 4.2.

Table 4A.4 – First-Stage Regressions - Aggregate Vector

		Dependent variable: GDP per capita (PPP)			
Vector	Variable	2000	2003	2006	2009
Aggregate	Trade ECI	0.371*	0.255	0.174	0.116
		(0.083)	(0.156)	(0.458)	(0.333)
	Input-Output ECI	-0.100	-0.127	-0.087	-0.159*
		(0.318)	(0.110)	(0.288)	(0.052)
	Interaction Term	0.009	0.173	0.123	0.371**
		(0.976)	(0.434)	(0.645)	(0.014)
	Control of Corruption (CC)	0.073	0.130	0.387*	0.384
		(0.822)	(0.552)	(0.055)	(0.178)
	Government Effectiveness (GE)	0.266	0.014	-0.159	-0.344
		(0.376)	(0.963)	(0.687)	(0.295)
	Political Stability and Absence of Violence (PV)	0.212*	0.120	0.118	0.061
		(0.076)	(0.200)	(0.190)	(0.412)
	Rule of Law (RL)	-0.122	0.116	-0.007	0.120
		(0.719)	(0.763)	(0.988)	(0.694)
	Regulatory Quality (RQ)	0.223	0.155	0.142	0.150
		(0.524)	(0.656)	(0.631)	(0.632)
	Voice and Accountability (VA)	-0.110	-0.124	-0.129	-0.247
		(0.718)	(0.560)	(0.608)	(0.167)
	log Human Capital Index	0.619	0.776	1.222**	1.424***
		(0.181)	(0.114)	(0.017)	(0.008)
	Adj. R^2	0.828	0.844	0.829	0.858

Table 4A.4 reports the results of cross-sectional OLS regressions for four base years, in which log GDP per capita is regressed jointly on all foundational capability vectors: multidimensional economic complexity, human capital, and institutional quality. Each coefficient is reported alongside its HC3 p-value in parentheses. Asterisks denote statistical significance: p*** < 0.01, p** < 0.05, and p* < 0.10. Adjusted R^2 values are shown at the bottom of each block. The sample includes 63 countries observed consistently across all four years, forming a balanced panel for the second-stage growth analysis.

In Table 4A.5, we present the second-stage results using the aggregated residual vector. The structural residuals show negative coefficients across all panels, as expected, and are statistically significant in every specification. However, the association between this model and economic growth weakens considerably when compared to the baseline results in Table 4.3. For example, in the first panel, the within R-squared decreases from 0.584 to 0.332. This may reflect overfitting in the first stage, where the aggregated residual absorbs much of the variation in current income but fails to capture the distinct facets of a country's capabilities that drive long-term growth.

D. Regressions with Interaction Terms between Residual Vectors

This section includes specifications testing additional pairwise interaction terms between the residuals of economic complexity, human capital, and institutions, exploring the complementarities among these structural drivers of long-run growth.

In Table 4A.6, we replace the interaction term between the residuals of human capital

Table 4A.5 – Panel Comparison - Second-Stage Results with Aggregate Residuals

Dependent variable:	10-year growth			5-year growth	
Base years:	1999, 2002, 2005, 2008	2000, 2003, 2006, 2009	2001, 2004, 2007, 2010	1999, 2004, 2009, 2014	2000, 2005, 2010, 2015
log NRR	-0.117	-0.223	0.004	-0.104	0.025
	(0.508)	(0.420)	(0.980)	(0.685)	(0.919)
ε^{AGG}	-2.452***	-3.144***	-2.196**	-2.606**	-2.945***
	(0.003)	(0.000)	(0.019)	(0.014)	(0.007)
Δ Trade $ECI_{t10,t}$	3.066	3.792	6.090	3.349	5.985
	(0.585)	(0.601)	(0.370)	(0.465)	(0.143)
Δ IO $ECI_{t10,t}$	7.726*	10.205**	3.497	6.346***	8.036**
	(0.097)	(0.010)	(0.474)	(0.006)	(0.038)
Δ WGI $PV_{t10,t}$	7.787***	5.355**	7.418***	3.026**	5.332***
	(0.008)	(0.031)	(0.003)	(0.03)	(0.004)
Δ WGI $RQ_{t10,t}$	13.626***	12.281***	0.476	11.702***	7.413**
	(0.003)	(0.010)	(0.932)	(0.000)	(0.014)
Adj. R^2	0.830	0.833	0.829	0.543	0.603
R^2 within	0.332	0.306	0.182	0.248	0.227
N (countries)	63	63	65	63	63

Table 4A.5 reports an alternative second-stage specification estimated across different panel constructions. In this specification, the residuals are aggregated into one variable, according to fist-stage regressions of Table A.4. Each column reports the results of two-way fixed-effects regressions using a distinct set of base years and growth periods. The first three columns correspond to overlapping ten-year panels starting in 1999, 2000, and 2001, respectively. The last two columns present non-overlapping five-year panels beginning in 1999 and 2000. Significance levels: $p^{***} < 0.01$, $p^{**} < 0.05$, $p^{*} < 0.10$.

and institutions with the interaction between the residuals of economic complexity and human capital. The coefficient for this interaction does not exhibit statistical significance in any of the panels. Meanwhile, the other variables in the model retain coefficient magnitudes and significance levels that are very similar to those observed in the baseline results, indicating that the inclusion of this alternative interaction term does not materially affect the overall model performance.

In Table 4A.7, we include the interaction term between the residuals of economic complexity and institutions. This interaction term is statistically significant in two of the five panels, with negative coefficients in all specifications. Overall, the inclusion of this interaction does not lead to loss of significance or changes in the sign of the coefficients of the other variables in the model, which remain consistent with the baseline results.

In Table 4A.8, we include all pairwise interaction terms between the residuals of economic complexity, human capital, and institutions. The only interaction term that shows statistical significance in any of the panels is that between human capital and institutions, which presents a negative coefficient, consistent with the baseline results. There are no notable changes in the significance levels or the signs of the other regressors. The only

Table 4A.6 – Panel Comparison - Second-Stage Results with Interaction Term between ECI and HCI Residuals

Dependent variable:	10-year growth			5-year growth	
Base years:	1999, 2002, 2005, 2008	2000, 2003, 2006, 2009	2001, 2004, 2007, 2010	1999, 2004, 2009, 2014	2000, 2005, 2010, 2015
log NRR	-0.137	-0.170	-0.036	-0.182	-0.056
	(0.375)	(0.465)	(0.790)	(0.427)	(0.817)
ε^{ECI}	-3.247***	-2.835***	-2.364***	-1.982**	-1.186
	(0.000)	(0.000)	(0.000)	(0.024)	(0.161)
ε^{HCI}	-2.732***	-2.646***	-3.112***	-2.766**	-1.785*
	(0.007)	(0.001)	(0.006)	(0.019)	(0.054)
ε^{INST}	-1.066***	-1.822***	-1.341**	-1.910*	-2.148**
	(0.010)	(0.001)	(0.029)	(0.056)	(0.017)
Interaction	-0.441	-0.573	-0.683	-1.359	-0.339
(ε^{ECI}, ε^{HCI})	(0.765)	(0.679)	(0.683)	(0.244)	(0.793)
Δ Trade $ECI_{t10,t}$	7.483	8.977	10.875	5.669	8.017**
	(0.116)	(0.128)	(0.128)	(0.121)	(0.045)
Δ IO $ECI_{t10,t}$	12.084***	14.182***	8.786*	7.571***	7.402**
	(0.006)	(0.000)	(0.054)	(0.001)	(0.034)
Δ WGI $PV_{t10,t}$	5.617***	4.328***	5.018***	3.007**	5.608***
	(0.003)	(0.009)	(0.004)	(0.033)	(0.002)
Δ WGI $RQ_{t10,t}$	12.646***	13.032***	4.750	12.099***	8.175**
	(0.002)	(0.003)	(0.302)	(0.000)	(0.02)
Adj. R^2	0.886	0.884	0.874	0.593	0.619
R^2 within	0.559	0.525	0.407	0.341	0.271
N (countries)	63	63	65	63	63

Table 4A.6 reports an alternative second-stage specification estimated across different panel constructions. In this specification, the interaction term between human capital (HCI) and institutional residuals is replaced by the interaction term between economic complexity (ECI) and HCI residuals. Each column reports the results of two-way fixed-effects regressions using a distinct set of base years and growth periods. The first three columns correspond to overlapping ten-year panels starting in 1999, 2000, and 2001, respectively. The last two columns present non-overlapping five-year panels beginning in 1999 and 2000. Significance levels: p*** < 0.01, p** < 0.05, p* < 0.10.

exception is in the last panel, where the residuals for complexity and human capital lose statistical significance.

Taken together, the results from Tables 4A.6 to 4A.8 suggest that introducing pairwise interaction terms between the structural residuals does not substantially alter the core findings of the model. The interactions generally do not exhibit consistent statistical significance across panels, with the exception of the human capital–institution interaction, which aligns with the baseline results when significant. Importantly, the inclusion of these interaction terms does not affect the significance or direction of the coefficients associated with the main residual vectors in most specifications, reinforcing the robustness of the baseline structural convergence framework.

Table 4A.7 – Panel Comparison - Second-Stage Results with Interaction Term between ECI and Institutions Residuals

Dependent variable:	10-year growth			5-year growth	
Base years:	1999, 2002, 2005, 2008	2000, 2003, 2006, 2009	2001, 2004, 2007, 2010	1999, 2004, 2009, 2014	2000, 2005, 2010, 2015
log NRR	-0.145	-0.178	-0.045	-0.183	-0.053
	(0.316)	(0.466)	(0.720)	(0.428)	(0.823)
ε^{ECI}	-2.896***	-2.767***	-2.320***	-1.793***	-1.227
	(0.002)	(0.000)	(0.000)	(0.009)	(0.114)
ε^{HCI}	-2.837***	-2.516***	-2.976**	-2.568**	-1.704*
	(0.003)	(0.002)	(0.011)	(0.036)	(0.059)
ε^{INST}	-1.026**	-1.829***	-1.284**	-1.970*	-2.210**
	(0.018)	(0.001)	(0.045)	(0.059)	(0.011)
Interaction	-1.497*	-0.648	-0.532	-1.864*	-0.760
(ε^{ECI}, ε^{INST})	(0.064)	(0.313)	(0.541)	(0.069)	(0.410)
Δ Trade $ECI_{t10,t}$	6.832	9.010	10.590	5.518	8.035**
	(0.154)	(0.131)	(0.141)	(0.144)	(0.045)
Δ IO $ECI_{t10,t}$	11.183***	13.740***	8.647**	7.501***	7.425**
	(0.008)	(0.000)	(0.049)	(0.001)	(0.034)
Δ WGI $PV_{t10,t}$	5.063***	4.256***	4.973***	2.975**	5.588***
	(0.006)	(0.009)	(0.002)	(0.033)	(0.001)
Δ WGI $RQ_{t10,t}$	12.756***	13.315***	5.052	12.307***	8.336**
	(0.001)	(0.002)	(0.261)	(0.000)	(0.016)
Adj. R^2	0.892	0.884	0.874	0.597	0.620
R^2 within	0.581	0.526	0.405	0.349	0.273
N (countries)	63	63	65	63	63

Table 4A.7 reports an alternative second-stage specification estimated across different panel constructions. In this specification, the interaction term between human capital (HCI) and institutional residuals is replaced by the interaction term between economic complexity (ECI) and institutional residuals. Each column reports the results of two-way fixed-effects regressions using a distinct set of base years and growth periods. The first three columns correspond to overlapping ten-year panels starting in 1999, 2000, and 2001, respectively. The last two columns present non-overlapping five-year panels beginning in 1999 and 2000. Significance levels: p*** < 0.01, p** < 0.05, p* < 0.10.

Table 4A.8 – Panel Comparison - Second-Stage Results with Interaction Term between All Residuals

Dependent variable:	10-year growth			5-year growth	
Base years:	1999, 2002, 2005, 2008	2000, 2003, 2006, 2009	2001, 2004, 2007, 2010	1999, 2004, 2009, 2014	2000, 2005, 2010, 2015
log NRR	-0.150	-0.184	-0.037	-0.224	-0.065
	(0.293)	(0.462)	(0.790)	(0.331)	(0.787)
ε^{ECI}	-2.776***	-2.667***	-2.373***	-1.797**	-1.076
	(0.004)	(0.000)	(0.000)	(0.025)	(0.214)
ε^{HCI}	-2.602***	-2.481***	-3.079***	-2.305**	-1.565
	(0.006)	(0.003)	(0.007)	(0.029)	(0.121)
ε^{INST}	-1.014**	-1.811***	-1.315*	-2.036**	-2.100**
	(0.018)	(0.001)	(0.057)	(0.040)	(0.032)
Interaction	0.572	-0.050	-0.527	-0.048	0.540
(ε^{ECI}, ε^{HCI})	(0.549)	(0.958)	(0.706)	(0.971)	(0.704)
Interaction	-0.914	0.261	-0.150	0.250	0.386
(ε^{ECI}, ε^{INST})	(0.307)	(0.685)	(0.81)	(0.868)	(0.828)
Interaction	-1.179*	-1.216	-0.168	-2.989**	-1.987
(ε^{HCI}, ε^{INST})	(0.099)	(0.208)	(0.950)	(0.025)	(0.277)
Δ Trade $ECI_{t10,t}$	7.222	8.203	10.708	5.455	7.737*
	(0.159)	(0.157)	(0.140)	(0.129)	(0.058)
Δ IO $ECI_{t10,t}$	10.753***	13.408***	8.683*	7.205***	6.721*
	(0.008)	(0.000)	(0.064)	(0.001)	(0.052)
Δ WGI $PV_{t10,t}$	4.751**	4.151**	4.926***	2.746*	5.421***
	(0.012)	(0.012)	(0.003)	(0.060)	(0.002)
Δ WGI $RQ_{t10,t}$	12.674***	13.189***	4.833	12.154***	8.500**
	(0.001)	(0.002)	(0.288)	(0.000)	(0.016)
Adj. R^2	0.893	0.885	0.873	0.609	0.623
R^2 within	0.589	0.535	0.407	0.374	0.286
N (countries)	63	63	65	63	63

Table 4A.8 reports an alternative second-stage specification estimated across different panel constructions. In this specification, all pairwise interaction terms among the residual vectors (economic complexity, human capital, and institutions) are included simultaneously. Each column reports the results of two-way fixed-effects regressions using a distinct set of base years and growth periods. The first three columns correspond to overlapping ten-year panels starting in 1999, 2000, and 2001, respectively. The last two columns present non-overlapping five-year panels beginning in 1999 and 2000. Significance levels: p*** < 0.01, p** < 0.05, p* < 0.10.

5 Discussion

Development is not a destination. It is a trajectory traced across a space of possibilities. This dissertation explores how the structure of that space influences the paths countries take and, more importantly, the paths they avoid. If we think of economic development as a process of accumulating and recombining knowledge, then industrial policy becomes the art of helping countries move through that space with direction, precision, and purpose. The three essays presented here offer tools for navigating that journey—not just more efficiently, but more wisely.

Navigating Traps: When Local Maxima Become Global Problems

The first essay highlights a simple but profound insight: not all local maxima in product space are good stepping stones. When a country gets comfortable at the top of a low hill, it may lose the incentive to climb the mountain next door. Local optima—products that deliver high short-term returns but offer low complexity prospects and poor connectivity in the product space can create complexity traps. Their danger lies not in what they instantly offer, but in what they crowd out: the incentives to explore, to take risks, to build capabilities beyond the comfort zone.

This mechanism has deep implications for industrial policy. It suggests that growth strategies based only on comparative advantage may be misleading, not because they are wrong, but because they are incomplete. Policymakers must recognize that some products, while lucrative, are structurally isolating and complexity short-sighted — they offer immediate gains at the expense of long-term capability accumulation, anchoring countries in trajectories that reward exploitation over exploration. Avoiding or escaping these traps requires:

- Sequenced exploration, where initial diversification is deliberately guided toward products that open multiple future paths towards complexity enhancing;
- Strategic patience, resisting the temptation to overexploit high-peak, low PCI products before building the capabilities to move beyond them;
- And a long-term vision, where economic complexity is treated not as a byproduct of growth, but as a policy objective in itself.

In a way, industrial policy must function like a GPS with a zoom function — it cannot just show the next turn; it must help countries see the shape of the terrain and choose the paths that lead beyond local hills, toward systemic transformation.

Filtering Noise: Seeing the Possibilities that Matter

If the first essay deals with traps, the second tackles illusions. The product space, rich in connections, is also rich in noise. Not all proximities are meaningful. Some are statistical artifacts. Others reflect transient coincidences. When policymakers use such noisy maps to guide diversification, they may end up chasing mirages.

After all, less can be more. By applying network filtering techniques, this study shows that we can reveal the real structure hiding behind the noise. The filtered product space becomes a cleaner, more reliable guide for identifying viable opportunities. The gains are not just statistical. They are profoundly practical. Countries with low diversity, in particular, benefit from a map that does not confuse proximity with potential.

The policy implications are immediate:

- Use filtered relatedness to prioritize diversification targets based on empirical strength, not surface similarity;
- Develop opportunity dashboards that rank potential products by filtered density and strategic relevance.

In this filtered landscape, industrial policy becomes an exercise in strategic discernment, eliminating noise to reveal the pathways that truly matter. But more than that, it becomes a tool to help countries navigate complexity not by chasing every option, but by amplifying the signals that resonate with their real, accumulated capabilities.

Aligning Structures: Complexity Is Only Part of the Story

While the first two essays focus on productive structure, the third essay makes a broader claim: economic convergence requires coherence between what countries produce (economic complexity), the rules and norms that govern collective action (institutions), and the cognitive foundation of the population (human capital). Development is not a puzzle of pieces that fit neatly into silos. It is an ecosystem of interdependent systems.

By introducing a multidimensional framework — including multidimensional complexity, institutions, and human capital — this essay shows that growth happens when these vectors align. More than that, it shows that positive capability gaps — instances in which a country's structural endowments exceed its income level — are strong predictors of future convergence. This insight reframes development not as a linear process of catching up, but as a dynamic race between the accumulation of diverse structural capabilities and the resolution of constraints.

For policy, the message is clear:

- Think beyond brick and mortar investments —industrial policy should induce coordinated systems of production, learning, and adaptation. Support economic complexity with institutions that foster coordination, trust, and adaptability, and with human capital that sustains learning, problem-solving, and the continuous recombination of knowledge;

- Identify the binding constraint. Is it productive capabilities? Institutional quality? Human capital? Then prioritize accordingly;
- Use structural diagnostics not only to assess where a country is, but where it could be, and what's holding it back.

In this view, industrial policy becomes less about sectoral targeting and more about orchestrating a coalition of complexity: a policy ensemble in which productive capabilities, institutions, and cognitive skills move together toward convergence.

A Final Note: Paths That Could Have Been

Industrial policy, as understood through the lens of this dissertation, is less about planning the future than choosing among futures that already exist in potential form. The product space, filtered or not, is a landscape of branching possibilities — some inviting, others deceptive, many invisible until the moment has passed. Some countries reach complexity through deliberate coordination. Others fall into traps not because they lacked ambition, but because the structure of their incentives led them astray.

As Jorge Luis Borges once wrote:

"Creía en infinitas series de tiempos, en una red creciente y vertiginosa de tiempos divergentes, convergentes y paralelos... Esa trama de tiempos que se aproximan, se bifurcan, se cortan o que secularmente se ignoran, abarca todas la posibilidades." (El jardín de senderos que se bifurcan, 1941)

This idea mirrors the developmental challenge faced by nations. Each economy stands at a fork in the path, surrounded by routes it could take, some of which converge toward high complexity and shared prosperity, others that spiral into stasis or specialization without depth. The job of the policymaker is not to predict the future, but to intervene in the present so that better futures remain accessible.

In this light, economic complexity is not just a measure of what a country does; it is a map of what it could become. And industrial policy is not merely technical strategy; it is the craft of shaping futures — a deliberate act of choosing which possible stories a nation will strive to make real. This thesis offers tools to clarify that map, to avoid the mirages, and to help countries choose among their many possible selves.

References

ACEMOGLU, D.; JOHNSON, S. Unbundling institutions. **Journal of political Economy**, The University of Chicago Press, v. 113, n. 5, p. 949–995, 2005.

ACEMOGLU, D.; JOHNSON, S.; ROBINSON, J. A. The colonial origins of comparative development: An empirical investigation. **American economic review**, American Economic Association, v. 91, n. 5, p. 1369–1401, 2001.

ACEMOGLU, D.; JOHNSON, S.; ROBINSON, J. A. Institutions as a fundamental cause of long-run growth. **Handbook of economic growth**, Elsevier, v. 1, p. 385–472, 2005.

AL-MARHUBI, F. Export diversification and growth: an empirical investigation. **Applied economics letters**, Taylor & Francis, v. 7, n. 9, p. 559–562, 2000.

ALBEAIK, S.; KALTENBERG, M.; ALSALEH, M.; HIDALGO, C. A. Improving the economic complexity index. **arXiv preprint arXiv:1707.05826**, 2017.

ALBORA, G.; PIETRONERO, L.; TACCHELLA, A.; ZACCARIA, A. Product progression: a machine learning approach to forecasting industrial upgrading. **Scientific Reports**, Nature Publishing Group UK London, v. 13, n. 1, p. 1481, 2023.

ALSHAMSI, A.; PINHEIRO, F. L.; HIDALGO, C. A. Optimal diversification strategies in the networks of related products and of related research areas. **Nature Communications**, v. 9, n. 1, p. 1328, 2018. ISSN 2041-1723. DOI 10.1038/s41467-018-03740-9. Available at: https://doi.org/10.1038/s41467-018-03740-9.

ANGELINI, O.; GABRIELLI, A.; TACCHELLA, A.; ZACCARIA, A.; PIETRONERO, L.; MATTEO, T. D. Forecasting the countries' gross domestic product growth: The case of technological fitness. **Chaos, Solitons & Fractals**, Elsevier, v. 184, p. 115006, 2024.

AUTY, R. **Sustaining development in mineral economies: The resource curse thesis**. [*S.l.*]: Routledge, 2002.

BADEEB, R. A.; LEAN, H. H.; CLARK, J. The evolution of the natural resource curse thesis: A critical literature survey. **Resources Policy**, Elsevier, v. 51, p. 123–134, 2017.

BALASSA, B. Trade liberalisation and "revealed" comparative advantage 1. **The manchester school**, Wiley Online Library, v. 33, n. 2, p. 99–123, 1965.

BALLAND, P.-A.; BOSCHMA, R. **An evolutionary approach to regional development traps in European regions**. [*S.l.*]: Utrecht University, Human Geography and Planning, 2024.

BALLAND, P.-A.; BOSCHMA, R.; CRESPO, J.; RIGBY, D. L. Smart specialization policy in the european union: relatedness, knowledge complexity and regional diversification. **Regional studies**, Taylor & Francis, v. 53, n. 9, p. 1252–1268, 2019.

BALLAND, P.-A.; RIGBY, D. The geography of complex knowledge. **Economic geography**, Taylor & Francis, v. 93, n. 1, p. 1–23, 2017.

BARBIERI, N.; CONSOLI, D.; NAPOLITANO, L.; PERRUCHAS, F.; PUGLIESE, E.; SBARDELLA, A. Regional technological capabilities and green opportunities in europe. **The Journal of Technology Transfer**, Springer, v. 48, n. 2, p. 749–778, 2023.

BARRETT, C. B.; CARTER, M. R. The economics of poverty traps and persistent poverty: empirical and policy implications. **The Journal of Development Studies**, Taylor & Francis, v. 49, n. 7, p. 976–990, 2013.

BAUMOL, W. J. Productivity growth, convergence, and welfare: what the long-run data show. **The american economic review**, JSTOR, p. 1072–1085, 1986.

BECKER, G. S. Book. **Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, Third Edition**. [*S.l.*]: The University of Chicago Press, 1994.

BEKAERT, G.; HARVEY, C. R.; LUNDBLAD, C. Does financial liberalization spur growth? **Journal of Financial economics**, Elsevier, v. 77, n. 1, p. 3–55, 2005.

BEN-DAVID, D. Convergence clubs and subsistence economies. **Journal of Development Economics**, Elsevier, v. 55, n. 1, p. 155–171, 1998.

BENHABIB, J.; SPIEGEL, M. M. The role of human capital in economic development evidence from aggregate cross-country data. **Journal of Monetary economics**, Elsevier, v. 34, n. 2, p. 143–173, 1994.

BERGER-TAL, O.; NATHAN, J.; MERON, E.; SALTZ, D. The exploration-exploitation dilemma: a multidisciplinary framework. **PloS one**, Public Library of Science San Francisco, USA, v. 9, n. 4, p. e95693, 2014.

BOSCHMA, R.; BALLAND, P.-A.; KOGLER, D. F. Relatedness and technological change in cities: the rise and fall of technological knowledge in us metropolitan areas from 1981 to 2010. **Industrial and corporate change**, Oxford University Press, v. 24, n. 1, p. 223–250, 2015.

BOSCHMA, R.; CAPONE, G. Institutions and diversification: Related versus unrelated diversification in a varieties of capitalism framework. **Research Policy**, Elsevier, v. 44, n. 10, p. 1902–1914, 2015.

BUSTOS, S.; GOMEZ, C.; HAUSMANN, R.; HIDALGO, C. A. The dynamics of nestedness predicts the evolution of industrial ecosystems. **PloS one**, Public Library of Science San Francisco, USA, v. 7, n. 11, p. e49393, 2012.

ČADIL, J.; PETKOVOVÁ, L.; BLATNÁ, D. Human capital, economic structure and growth. **Procedia economics and finance**, Elsevier, v. 12, p. 85–92, 2014.

CALDAROLA, B.; MAZZILLI, D.; NAPOLITANO, L.; PATELLI, A.; SBARDELLA, A. Economic complexity and the sustainability transition: A review of data, methods, and literature. **Journal of Physics: Complexity**, 2024.

CAVALCANTI, T. V. d. V.; MOHADDES, K.; RAISSI, M. Growth, development and natural resources: New evidence using a heterogeneous panel analysis. **The Quarterly Review of Economics and Finance**, Elsevier, v. 51, n. 4, p. 305–318, 2011.

COSCIA, M.; NEFFKE, F. M. Network backboning with noisy data. *In*: IEEE. **2017 IEEE 33rd international conference on data engineering (ICDE)**. [*S.l.*], 2017. p. 425–436.

CRISTELLI, M.; TACCHELLA, A.; PIETRONERO, L. The heterogeneous dynamics of economic complexity. **PloS one**, Public Library of Science San Francisco, CA USA, v. 10, n. 2, p. e0117174, 2015.

CUNZO, F. de; PETRI, A.; ZACCARIA, A.; SBARDELLA, A. The trickle down from environmental innovation to productive complexity. **Scientific Reports**, Nature Publishing Group UK London, v. 12, n. 1, p. 22141, 2022.

DAM, A. V.; FRENKEN, K. Variety, complexity and economic development. **Research Policy**, Elsevier, v. 51, n. 8, p. 103949, 2022.

DHILLON, I. S. Co-clustering documents and words using bipartite spectral graph partitioning. *In*: **Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining**. [*S.l.*: *s.n.*], 2001. p. 269–274.

DRAGHI, M. **The Future of European Competitiveness Part B: In-depth analysis and recommendations**. [*S.l.*], 2024.

EVENETT, S.; JAKUBIK, A.; MARTÍN, F.; RUTA, M. The return of industrial policy in data. **The World Economy**, Wiley Online Library, v. 47, n. 7, p. 2762–2788, 2024.

FEENSTRA, R. C.; INKLAAR, R.; TIMMER, M. P. The next generation of the penn world table. **American economic review**, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, v. 105, n. 10, p. 3150–3182, 2015.

FOTI, N. J.; HUGHES, J. M.; ROCKMORE, D. N. Nonparametric sparsification of complex multiscale networks. **PloS one**, Public Library of Science San Francisco, USA, v. 6, n. 2, p. e16431, 2011.

FRENKEN, K.; NEFFKE, F.; DAM, A. V. Capabilities, institutions and regional economic development: a proposed synthesis. **Cambridge Journal of Regions, Economy and Society**, Oxford University Press UK, v. 16, n. 3, p. 405–416, 2023.

GALA, P.; CAMARGO, J.; MAGACHO, G.; ROCHA, I. Sophisticated jobs matter for economic complexity: An empirical analysis based on input-output matrices and employment data. **Structural Change and Economic Dynamics**, Elsevier, v. 45, p. 1–8, 2018.

GALA, P.; ROCHA, I.; MAGACHO, G. The structuralist revenge: economic complexity as an important dimension to evaluate growth and development. **Brazilian journal of political economy**, SciELO Brasil, v. 38, n. 2, p. 219–236, 2018.

GILL, I. S.; KHARAS, H. The middle-income trap turns ten. **World Bank Policy Research Working Paper**, n. 7403, 2015.

GYLFASON, T. Natural resources, education, and economic development. **European economic review**, Elsevier, v. 45, n. 4-6, p. 847–859, 2001.

HANUSHEK, E. A.; WOESSMANN, L. The role of cognitive skills in economic development. **Journal of economic literature**, American Economic Association, v. 46, n. 3, p. 607–668, 2008.

HARTMANN, D.; ZAGATO, L.; GALA, P.; PINHEIRO, F. L. Why did some countries catch-up, while others got stuck in the middle? stages of productive sophistication and smart industrial policies. **Structural Change and Economic Dynamics**, Elsevier, v. 58, p. 1–13, 2021.

HAUSMANN, R.; HIDALGO, C. A. The network structure of economic output. **Journal of economic growth**, Springer, v. 16, p. 309–342, 2011.

HAUSMANN, R.; HIDALGO, C. A.; BUSTOS, S.; COSCIA, M.; SIMOES, A. **The atlas of economic complexity: Mapping paths to prosperity**. [*S.l.*]: Mit Press, 2014.

HAUSMANN, R.; HWANG, J.; RODRIK, D. What you export matters. **Journal of economic growth**, Springer, v. 12, p. 1–25, 2007.

HERZER, D.; D, F. N.-L. What does export diversification do for growth? an econometric analysis. **Applied economics**, Taylor & Francis, v. 38, n. 15, p. 1825–1838, 2006.

HESSE, H. *et al.* Export diversification and economic growth. **Breaking into new markets: emerging lessons for export diversification**, Washington, DC, Banco Mundial, v. 2009, p. 55–80, 2009.

HIDALGO, C. A. Economic complexity theory and applications. **Nature Reviews Physics**, Nature Publishing Group UK London, v. 3, n. 2, p. 92–113, 2021.

HIDALGO, C. A. The policy implications of economic complexity. **Research Policy**, Elsevier, v. 52, n. 9, p. 104863, 2023.

HIDALGO, C. A.; BALLAND, P.-A.; BOSCHMA, R.; DELGADO, M.; FELDMAN, M.; FRENKEN, K.; GLAESER, E.; HE, C.; KOGLER, D. F.; MORRISON, A. *et al.* The principle of relatedness. *In*: SPRINGER. **Unifying Themes in Complex Systems**

IX: Proceedings of the Ninth International Conference on Complex Systems 9. [*S.l.*], 2018. p. 451–457.

HIDALGO, C. A.; HAUSMANN, R. The building blocks of economic complexity. **Proceedings of the national academy of sciences**, National Acad Sciences, v. 106, n. 26, p. 10570–10575, 2009.

HIDALGO, C. A.; KLINGER, B.; BARABÁSI, A.-L.; HAUSMANN, R. The product space conditions the development of nations. **Science**, American Association for the Advancement of Science, v. 317, n. 5837, p. 482–487, 2007.

HIRSCHMAN, A. O. The strategy of economic development. New Haven, Conn.: Yale University,, 1958.

HIRSCHMAN, A. O. A generalized linkage approach to development, with special reference to staples. **Economic development and cultural change**, University of Chicago Press, v. 25, p. 67, 1977.

IM, F. G.; ROSENBLATT, D. Middle-income traps: a conceptual and empirical survey. **Journal of International Commerce, Economics and Policy**, World Scientific, v. 6, n. 03, p. 1550013, 2015.

IMBS, J.; WACZIARG, R. Stages of diversification. **American economic review**, American Economic Association, v. 93, n. 1, p. 63–86, 2003.

JAMES, A. The resource curse: a statistical mirage? **Journal of Development Economics**, Elsevier, v. 114, p. 55–63, 2015.

JR, R. E. L. On the mechanics of economic development. **Journal of monetary economics**, Elsevier, v. 22, n. 1, p. 3–42, 1988.

JUHÁSZ, R.; LANE, N.; RODRIK, D. The new economics of industrial policy. **Annual Review of Economics**, Annual Reviews, v. 16, 2023.

KAUFFMAN, S. A. **Investigations**. [*S.l.*]: Oxford University Press, 2000.

KAUFMANN, D.; KRAAY, A. The worldwide governance indicators: methodology and 2024 update. **Available at SSRN 5154675**, 2024.

KOCH, P.; SCHWARZBAUER, W. Yet another space: Why the industry space adds value to the understanding of structural change and economic development. **Structural Change and Economic Dynamics**, Elsevier, v. 59, p. 198–213, 2021.

KOGLER, D. F.; RIGBY, D. L.; TUCKER, I. Mapping knowledge space and technological relatedness in us cities. *In*: **Global and Regional Dynamics in Knowledge Flows and Innovation**. [*S.l.*]: Routledge, 2015. p. 58–75.

KRAAY, A.; MCKENZIE, D. Do poverty traps exist? assessing the evidence. **Journal of Economic Perspectives**, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203-2418, v. 28, n. 3, p. 127–148, 2014.

KREMER, M.; ONATSKI, A.; STOCK, J. Searching for prosperity. *In*: ELSEVIER. **Carnegie-Rochester Conference Series on Public Policy**. [*S.l.*], 2001. v. 55, n. 1, p. 275–303.

LEDERMAN, D. **Natural resources: neither curse nor destiny**. [*S.l.*]: Stanford University Press, 2007.

LEE, S. H. Network nestedness as generalized core-periphery structures. **Physical Review E**, APS, v. 93, n. 2, p. 022306, 2016.

LIN, J. Y. New structural economics: A framework for rethinking development. **The World Bank Research Observer**, Oxford University Press, v. 26, n. 2, p. 193–221, 2011.

MACKINNON, J. G.; NIELSEN, M. Ø.; WEBB, M. D. Fast and reliable jackknife and bootstrap methods for cluster-robust inference. **Journal of Applied Econometrics**, Wiley Online Library, v. 38, n. 5, p. 671–694, 2023.

MANKIW, N. G.; ROMER, D.; WEIL, D. N. A contribution to the empirics of economic growth. **The quarterly journal of economics**, MIT Press, v. 107, n. 2, p. 407–437, 1992.

MARCH, J. G. Exploration and exploitation in organizational learning. **Organization science**, INFORMS, v. 2, n. 1, p. 71–87, 1991.

MEALY, P.; FARMER, J. D.; TEYTELBOYM, A. Interpreting economic complexity. **Science advances**, American Association for the Advancement of Science, v. 5, n. 1, p. eaau1705, 2019.

MEHLUM, H.; MOENE, K.; TORVIK, R. Institutions and the resource curse. **The economic journal**, Oxford University Press Oxford, UK, v. 116, n. 508, p. 1–20, 2006.

MINCER, J. **Human capital and economic growth**. [*S.l.*]: National Bureau of Economic Research Cambridge, Mass., USA, 1981.

MUELLER, B. The coevolution of everything, everywhere, all at once: Institutions, culture, and the great enrichment. *In*: **Handbook on Institutions and Complexity**. [*S.l.*]: Edward Elgar Publishing, 2025. p. 126–157.

MUNEEPEERAKUL, R.; LOBO, J.; SHUTTERS, S. T.; GOMÉZ-LIÉVANO, A.; QUBBAJ, M. R. Urban economies and occupation space: Can they get “there” from “here”? **PloS one**, Public Library of Science San Francisco, USA, v. 8, n. 9, p. e73676, 2013.

NEFFKE, F.; HENNING, M.; BOSCHMA, R. How do regions diversify over time? industry relatedness and the development of new growth paths in regions. **Economic geography**, Taylor & Francis, v. 87, n. 3, p. 237–265, 2011.

NORTH, D. C. Institutions, institutional change and economic performance. **Cambridge University**, 1990.

OECD. **OECD Inter-Country Input-Output Database**. [*S.l.*]: OECD Paris, France, 2023.

OSIOBE, E. U. *et al.* A literature review of human capital and economic growth. **Business and Economic Research**, Macrothink Institute, v. 9, n. 4, p. 179–196, 2019.

O'CLERY, N.; YILDIRIM, M. A.; HAUSMANN, R. Productive ecosystems and the arrow of development. **Nature communications**, Nature Publishing Group UK London, v. 12, n. 1, p. 1479, 2021.

PANIZZA, U.; PRESBITERO, A. F. Public debt and economic growth: is there a causal effect? **Journal of Macroeconomics**, Elsevier, v. 41, p. 21–41, 2014.

PEREIRA, A. J.; SILVA, G. J. d.; LARRUSCAIM, I. d. M. Economic complexity and the brazilian production structure: Export pattern and national economic development in the 21st century. **Revista de Economia Contemporânea**, SciELO Brasil, v. 27, p. e232717, 2023.

PINHEIRO, F. L.; HARTMANN, D.; BOSCHMA, R.; HIDALGO, C. A. The time and frequency of unrelated diversification. **Research Policy**, Elsevier, v. 51, n. 8, p. 104323, 2022.

PLOEG, F. v. d. Natural resources: curse or blessing? **Journal of Economic literature**, American Economic Association, v. 49, n. 2, p. 366–420, 2011.

PUGLIESE, E.; CIMINI, G.; PATELLI, A.; ZACCARIA, A.; PIETRONERO, L.; GABRIELLI, A. Unfolding the innovation system for the development of countries: coevolution of science, technology and production. **Scientific reports**, Nature Publishing Group UK London, v. 9, n. 1, p. 16440, 2019.

PUGLIESE, E.; TACCHELLA, A. **Economic complexity for competitiveness and innovation: A novel bottom-up strategy linking global and regional capacities**. [*S.l.*], 2020.

QUAH, D. Empirical cross-section dynamics in economic growth. Institute for Empirical Macroeconomics, Federal Reserve Bank of Minneapolis, 1992.

QUEIROZ, A. R.; ROMERO, J. P.; FREITAS, E. Complejidad económica y empleo en los estados del brasil. **Revista CEPAL**, 2023.

RIGBY, D. L. Technological relatedness and knowledge space: entry and exit of us cities from patent classes. **Regional Studies**, Taylor & Francis, v. 49, n. 11, p. 1922–1937, 2015.

RODRIK, D. **The future of economic convergence**. [*S.l.*], 2011.

RODRIK, D.; SANDHU, R. **Servicing development: productive upgrading of labor-absorbing services in developing economies**. [*S.l.*], 2024.

RODRIK, D.; SUBRAMANIAN, A.; TREBBI, F. Institutions rule: the primacy of institutions over geography and integration in economic development. **Journal of economic growth**, Springer, v. 9, p. 131–165, 2004.

ROMER, P. M. Endogenous technological change. **Journal of political Economy**, The University of Chicago Press, v. 98, n. 5, Part 2, p. S71–S102, 1990.

SACHS, J. D.; WARNER, A. **Natural resource abundance and economic growth**. [*S.l.*]: National bureau of economic research Cambridge, Mass., USA, 1995.

SBARDELLA, A.; PUGLIESE, E.; ZACCARIA, A.; SCARAMOZZINO, P. The role of complex analysis in modelling economic growth. **Entropy**, MDPI, v. 20, n. 11, p. 883, 2018.

SCHUMPETER, J. A.; SWEDBERG, R. **The theory of economic development**. [*S.l.*]: Routledge, 2021.

SERRANO, M. Á.; BOGUNÁ, M.; VESPIGNANI, A. Extracting the multiscale backbone of complex weighted networks. **Proceedings of the national academy of sciences**, National Acad Sciences, v. 106, n. 16, p. 6483–6488, 2009.

SOUSA, R.; MUELLER, B. Economic complexity, institutions, and industrial policy. *In*: **Handbook on Institutions and Complexity**. [*S.l.*]: Edward Elgar Publishing, 2025. p. 325–348.

STOJKOSKI, V.; KOCH, P.; HIDALGO, C. A. Multidimensional economic complexity and inclusive green growth. **Communications Earth & Environment**, Nature Publishing Group UK London, v. 4, n. 1, p. 130, 2023.

TACCHELLA, A.; CRISTELLI, M.; CALDARELLI, G.; GABRIELLI, A.; PIETRONERO, L. A new metrics for countries' fitness and products' complexity. **Scientific reports**, Nature Publishing Group UK London, v. 2, n. 1, p. 723, 2012.

TACCHELLA, A.; MAZZILLI, D.; PIETRONERO, L. A dynamical systems approach to gross domestic product forecasting. **Nature Physics**, Nature Publishing Group UK London, v. 14, n. 8, p. 861–865, 2018.

TACCHELLA, A.; ZACCARIA, A.; MICCHELI, M.; PIETRONERO, L. Relatedness in the era of machine learning. **Chaos, Solitons & Fractals**, Elsevier, v. 176, p. 114071, 2023.

TERZA, J. V.; BASU, A.; RATHOUZ, P. J. Two-stage residual inclusion estimation: addressing endogeneity in health econometric modeling. **Journal of health economics**, Elsevier, v. 27, n. 3, p. 531–543, 2008.

TORNELL, A.; WESTERMANN, F.; MARTINEZ, L. **The positive link between financial liberalization, growth and crises**. [*S.l.*]: National Bureau of Economic Research Cambridge, Mass., USA, 2004.

VU, T. V. Does institutional quality foster economic complexity? the fundamental drivers of productive capabilities. **Empirical Economics**, Springer, v. 63, n. 3, p. 1571–1604, 2022.

World Bank. **Global Economic Prospects, June 2019: Heightened Tensions, Subdued Investment**. [*S.l.*]: The World Bank, 2019.

WRIGHT, S. *et al.* The roles of mutation, inbreeding, crossbreeding, and selection in evolution. na, 1932.

YASSIN, A.; HAIDAR, A.; CHERIFI, H.; SEBA, H.; TOGNI, O. An evaluation tool for backbone extraction techniques in weighted complex networks. **Scientific Reports**, Nature Publishing Group UK London, v. 13, n. 1, p. 17000, 2023.

ZACCARIA, A.; CRISTELLI, M.; TACCHELLA, A.; PIETRONERO, L. How the taxonomy of products drives the economic development of countries. **PloS one**, Public Library of Science San Francisco, USA, v. 9, n. 12, p. e113770, 2014.

UnB