

**UNIVERSIDADE FEDERAL DE MINAS GERAIS**

**Faculdade de Ciências Econômicas**

**Programa de Pós Graduação em Economia**

Calebe Cardia Piacentini

**Laços internacionais complementares e diversificação tecnológica dos  
países: uma análise com dados de patente para os escritórios do USPTO  
e EPO**

Belo Horizonte

2023

Calebe Cardia Piacentini

**Laços internacionais complementares e diversificação tecnológica dos países: uma análise com dados de patente para os escritórios do USPTO e EPO**

Dissertação apresentada ao Programa de Mestrado em Economia do Centro de Desenvolvimento e Planejamento Regional da Universidade Federal de Minas Gerais (CEDEPLAR/UFMG), como requisito parcial para obtenção do título de Mestre em Economia.

Orientador: Gustavo de Britto Rocha

Co-orientador: Leonardo Costa Ribeiro

Belo Horizonte

2023

Ficha catalográfica

P5811 Piacentini, Calebe Cardia.  
2023 Laços internacionais complementares e diversificação  
tecnológica dos países [manuscrito]: uma análise com dados de  
patente para os escritórios do USPTO e EPO / Calebe Cardia  
Piacentini. – 2023.  
134 f.: il., gráfs. e tabs.

Orientador: Gustavo de Britto Rocha  
Coorientador: Leonardo Costa Ribeiro

Dissertação (mestrado) – Universidade Federal de Minas  
Gerais, Centro de Desenvolvimento e Planejamento Regional.  
Inclui bibliografia (f. 68-74).

1. Patentes – Teses. 2. Desenvolvimento organizacional –  
Teses. 3. Economia – Teses. I. Britto, Gustavo. II. Ribeiro,  
Leonardo Costa. III. Universidade Federal de Minas Gerais. Centro  
de Desenvolvimento e Planejamento Regional. IV. Título.

CDD: 330

Elaborada por Rosilene Santos CRB6-2527  
Biblioteca da FACE/UFMG. RSS – 19/2023



UNIVERSIDADE FEDERAL DE MINAS GERAIS  
FACULDADE DE CIÊNCIAS ECONÔMICAS  
CENTRO DE DESENVOLVIMENTO E PLANEJAMENTO REGIONAL  
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

## FOLHA DE APROVAÇÃO

**CALEBE CARDIA PIACENTINI**

LAÇOS INTERNACIONAIS COMPLEMENTARES E DIVERSIFICAÇÃO TECNOLÓGICA DOS PAÍSES: UMA ANÁLISE COM DADOS DE PATENTE PARA OS ESCRITÓRIOS DO USPTO E EPO

Dissertação apresentada ao programa de Pós-graduação em Economia, da Faculdade de Ciências Econômicas da Universidade Federal de Minas Gerais, para obtenção do título de mestre em economia, área de concentração em Economia.

**Aprovado** em Belo Horizonte, 09 de fevereiro de 2023.

### **BANCA EXAMINADORA - participações por videoconferência**

Prof. Gustavo de Britto Rocha (Orientador) (Cedeplar/FACE/UFMG)

Prof. Leonardo Costa Ribeiro (Coorientador) (Cedeplar/FACE/UFMG)

Prof. João Prates Romero (Cedeplar/FACE/UFMG)

Prof. Elton Eduardo Freitas (UFS)

JOÃO BOSCO MOURA TONUCCI FILHO  
Subcoordenador do Programa de Pós-Graduação em Economia



Documento assinado eletronicamente por **Leonardo Costa Ribeiro, Professor do Magistério Superior**, em 09/02/2023, às 14:51, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **João Prates Romero, Professor do Magistério Superior**, em 09/02/2023, às 15:19, conforme horário oficial de Brasília, com fundamento no



art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Elton Eduardo Freitas, Usuário Externo**, em 09/02/2023, às 15:27, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Gustavo de Britto Rocha, Professor do Magistério Superior**, em 10/02/2023, às 09:10, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **João Bosco Moura Tonucci Filho, Professor do Magistério Superior**, em 10/02/2023, às 10:18, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site [https://sei.ufmg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador 2073313 e o código CRC D35901A1.

# Agradecimentos

Este trabalho não seria possível sem o apoio da minha família. Fazer pós-graduação é muitas vezes um desafio e ainda mais quando realizada em um outro Estado. Sempre compreensivos, meu pai, minha mãe e meus avós tornaram as rápidas passagens por Guarulhos muito mais suaves e agradáveis.

Emocionalmente, o auxílio da minha namorada, Rafaela, foi imprescindível. Uma companheira para compartilhar as felicidades e tristezas, fez também o papel de conselheira em inúmeras oportunidades. Ouvido atento e mente afiada, é grande responsável para a efetivação desse trabalho.

Tenho também de agradecer aos meus conselheiros institucionais. Agradeço o Gustavo Britto pelo esforço que sempre desempenhou em abrir espaço na sua conturbada agenda para ouvir meus pensamentos confusos e compartilhar sua bússola comigo. Leonardo Ribeiro, como coorientador, abraçou o projeto e providenciou uma ajuda técnica indispensável, além das ótimas ideias que constituem importantes blocos dessa dissertação. Tantos outros professores que tive a oportunidade de encontrar no mestrado também foram fundamentais nesse processo. Destaco o professor João Romero que contribuiu diretamente em algumas partes dessa dissertação.

A lista continua pelos tantos colegas da pós-graduação do Cedeplar que foram tão receptivos e tornaram diversas etapas mais simples. Agradeço à Samara e à Nara por compartilhar tantos almoços juntos, à colega e vizinha Cinthia, ao Arthur por ter me recebido quando cheguei a Belo Horizonte. Muitas vezes influenciando diretamente nesse trabalho, o Alexandre e o Elton também contribuíram. Indiretamente, muitos outros colegas tornaram esse trabalho possível e agradeço a todos.

Agradeço também ao CNPQ e CAPES por financiarem meus estudos durante o mestrado. Obviamente, qualquer erro que tenha restado nesse trabalho é de minha responsabilidade.

# Resumo

É mais fácil diversificar para setores tecnológicos que são similares àqueles em que já se é especializado. Contudo, países periféricos possuem poucas capacidades e por isso têm maior dificuldade em acessar tecnologias complexas, perpetuando as desigualdades entre os países. Uma possível solução é colaborar com outros países, suprindo a ausência de capacidades locais por meio destes contatos. Assim, não só esse efeito deve ser maior para países menos complexos, como o impacto deve ser superior quando se colabora com países cuja estrutura tecnológica não seja nem tão próxima nem tão distante, mas complementar. Este trabalho testa essas hipóteses usando dados internacionais de patentes para o escritório norte-americano (USPTO) e europeu (EPO) entre 2000 a 2019. Estabelecendo laços através da co-aplicação em patentes e usando do índice de complementariedade proposto por [Balland e Boschma \(2021\)](#), através de modelos de índice Probit e Logit com efeitos-fixos e erros clusterizados, encontra-se evidência de que países que colaboram com países que possuam capacidades similares em uma tecnologia têm maior probabilidade de entrar nela. Este efeito é robusto a diferentes especificações e é maior ainda para países menos complexos. Contudo, encontramos este impacto apenas depois de controlar por efeitos não observáveis dos países, sendo na verdade o oposto quando desconsiderados. Possivelmente isto indica algo que podemos ver na rede de colaborações internacionais de que países desenvolvidos são seus maiores beneficiários, simplesmente porque são os que mais colaboram. Além disso, possuir capacidades similares localmente ainda está associado a maior probabilidade de especialização em novas tecnologias, o que adiciona à evidência do “princípio do relacionamento”. Assim, ao mesmo tempo em que este trabalho fornece evidências inéditas ao nível internacional, coloca novas questões e desafios para trabalhos futuros, principalmente quanto a saber se laços com outros países de fato mitigam ou aprofundam as desigualdades entre os países.

**Palavras-chave:** Complexidade; Relacionamento; Inovação; Colaboração; Patentes; Complementariedade; Desigualdades; Diversificação.

# Abstract

It is easier to diversify into technology sectors that are similar to those in which one is already specialized. However, peripheral countries have few capabilities and therefore have a greater difficulty to access complex technologies, perpetuating inequalities between countries. A possible solution is to collaborate with other countries, supplementing the absence of local capabilities through these contacts. Thus, not only should this effect be greater for less complex countries, but the impact should be greater when collaborating with countries whose technological structure is not as close or as distant, but complementary. This work tests these hypotheses using international patent data for the USPTO and EPO offices between 2000 and 2019. By establishing ties through patent co-application and using the complementary index proposed by Balland and Boschma (2021), through Probit and Logit models with fixed-effects and clustered errors, evidence is found that countries that collaborate with others that have similar capabilities in a technology have a greater probability of entering in that technology. This effect is robust to different specifications and is even greater for less complex countries. However, we find this impact only after controlling for unobserved country effects, the impact being actually the opposite when disregarded. This possibly indicates something that we can see in the network of international collaborations that developed countries are its biggest beneficiaries, simply because they are the ones that collaborate the most. Besides, having similar capabilities locally is still related to a greater probability of specialization in new technologies, which adds to the “principle of relatedness”. Therefore, this work at the same time as providing new evidence at the international level, raises new questions and challenges for future works, especially as to whether links with other countries actually mitigate or deepen inequalities between countries.

**Keywords:** Complexity; Relatedness; Innovation; Collaboration; Patents; Complementarity; Inequalities; Diversification.



# Lista de ilustrações

Figura 1 – Especialização Inteligente . . . . .	29
Figura 2 – Distintas oportunidades . . . . .	31
Figura 3 – Curva S . . . . .	32
Figura 4 – Distância Cognitiva Ótima . . . . .	40
Figura 5 – Rede de Colaborações Internacionais Hipotética . . . . .	56
Figura 6 – Proporção de Aplicantes por Posição . . . . .	60
Figura 7 – Rede de Colaboração Internacional . . . . .	65
Figura 8 – Distribuição da rede de colaborações internacionais . . . . .	68
Figura 9 – Distribuição das colaborações . . . . .	69
Figura 10 – Rede de Colaboração Internacional . . . . .	73
Figura 11 – Rede de Colaboração Internacional . . . . .	74
Figura 12 – Rede de Colaboração Internacional . . . . .	75
Figura 13 – Curva S . . . . .	78
Figura 14 – Correlogramas . . . . .	80
Figura 15 – Correlogramas por Cor . . . . .	81

# Lista de tabelas

Tabela 1 – Correspondência das Variáveis . . . . .	79
Tabela 2 – Estatísticas Básicas . . . . .	82
Tabela 3 – Modelos De Entrada . . . . .	86
Tabela 4 – Efeitos Marginais Médios . . . . .	88
Tabela 5 – Rede Completa de Colaborações (USPTO) . . . . .	107
Tabela 6 – Rede Completa de Colaborações (EPO) . . . . .	112
Tabela 7 – Valores das Redes Coloridas (USPTO) . . . . .	116
Tabela 8 – Valores das Redes Coloridas (EPO) . . . . .	117
Tabela 9 – Valores da Rede Colorida pela Complexidade (USPTO) . . . . .	118
Tabela 10 – Valores da Rede Colorida pela Complexidade (EPO) . . . . .	119
Tabela 11 – Modelos De Entrada . . . . .	121
Tabela 12 – Modelos De Entrada . . . . .	122
Tabela 13 – Modelos De Entrada . . . . .	123
Tabela 14 – Modelos De Entrada . . . . .	124
Tabela 15 – Modelos De Entrada . . . . .	125

# Sumário

<b>1</b>	<b>Introdução</b>	<b>11</b>
<b>2</b>	<b>Diversificação Tecnológica: o privilégio da localidade</b>	<b>17</b>
2.1	Complexidade econômica e a heterogeneidade dos produtos	17
2.2	Diversificação relacionada e o Princípio do Relacionamento	24
2.3	O refúgio no exterior: a importância de colaborações internacionais	33
<b>3</b>	<b>Metodologia</b>	<b>44</b>
3.1	Indicadores	44
3.2	Modelo de regressão estatística	48
3.3	Redes	53
3.4	Dados e Procedimentos Empíricos	56
<b>4</b>	<b>Exercício Empírico</b>	<b>63</b>
4.1	Análise Exploratória	63
4.2	Regressão	85
<b>5</b>	<b>Conclusão</b>	<b>93</b>
	<b>Referências</b>	<b>98</b>
	<b>Anexos</b>	<b>105</b>
	<b>ANEXO A Dados</b>	<b>106</b>
	<b>ANEXO B Regressões</b>	<b>120</b>

# 1 Introdução

A inovação tecnológica tem sido longamente reconhecida como um importante fator na performance econômica. A lógica é simples: avanços tecnológicos permitem firmas aumentarem a produtividade do trabalho. Assim, por um mesmo esforço, mais produtos são gerados, aumentando as possibilidades de consumo da população. Este esforço poupado pode, então, ser redirecionado a outras atividades que, supondo produtivas, aumentariam os bens disponíveis para a sociedade. Normalmente, assumimos que uma quantidade maior de consumo – que derivaria de uma maior abundância de bens – deixa as pessoas mais felizes. Assim, a pergunta fundamental que surge é: o que leva ao avanço tecnológico, isto é, à *inovação*?

Uma inovação tecnológica acontece não por obra do acaso, mas como resultado de um processo de recombinação de conhecimentos e tecnologias já adquiridos (FLEMING; SORENSON, 2001). Seria impossível construir um fogão sem antes descobrir como controlar o calor. Tendo achado uma forma de gerar calor – o fogo! – é que podemos recombinar esse conhecimento com algum outro, digamos o da metalurgia, para poder criar um fogão. Como esse exemplo hipotético ilustra, a inovação seria, então, a efetivação de um processo de busca dentro de um “espaço do conhecimento” (FLEMING; SORENSON, 2001), um ambiente abstrato onde os conhecimentos residem. É como se em cada ponto específico deste espaço de, digamos, 3 dimensões – comprimento, largura e altura – representasse uma tecnologia específica, como a do fogão. O fogão é o resultado da combinação de diferentes conhecimentos: o controle do fogo e da metalurgia (senão, para que seria necessário). É como se cada um desses conhecimentos fosse um ponto num dos eixos desse espaço, por exemplo, fogo sendo  $x = 3$  e a metalurgia  $y = 1$ . Da junção destes dois conhecimentos, temos o fogão,  $(x, y) = (3, 1)$ . A terceira dimensão, a altura, definiria o quão útil essa tecnologia é para a sociedade, tal qual uma função,  $f(x, y) = z$ . No caso, o fogão é muito útil, então  $z$  deve ser um valor relativamente alto. Como adquirir um novo conhecimento é um processo custoso – ou, dito de outra maneira, *mover-se* no “espaço

do conhecimento” é difícil –, a de busca por novas combinações que sejam úteis (gerem um  $z$  elevado) é localizado<sup>1</sup>. A partir do momento que eu já sei gerar o fogo com uma madeira específica, é muito mais fácil eu testar – recombina – com outras lenhas esse mesmo processo do que tentar subitamente congelar as coisas. Quando repassamos essa lógica para o processo produtivo, isso se traduz em países, regiões, firmas e instituições concentrando seus esforços de pesquisa naqueles conhecimentos e áreas mais familiares à sua atividade presente.

Esta dependência de conhecimentos e habilidades previamente adquiridas resulta numa dependência de trajetória (NELSON; WINTER, 1982; BOSCHMA, 2005; MARTIN; SUNLEY, 2006). O que uma firma produz hoje em parte determina aquilo que ela será capaz de produzir amanhã. É mais fácil para uma empresa que produz fogões fazer uma panela – porque ela já domina a metalurgia e assim, por exemplo, tem conhecimento sobre a resistência dos materiais<sup>2</sup> – do que uma da indústria têxtil. Isso quer dizer que o desenvolvimento de uma tecnologia tem um impacto duplo. Por um lado, influencia no aumento de produtividade da economia – uma panela facilita muito no cozimento. Por outro, impacta no desenvolvimento de novas tecnologias. Se hoje a firma inova descobrindo como faz uma panela simples, amanhã ela pode pensar em como fazer uma panela industrial. Essa relação não se dá apenas temporalmente, mas espacialmente também (JAFFE; TRAJTENBERG; HENDERSON, 1993; FELDMAN; FLORIDA, 1994), ou seja, a firma fica constricta às capacidades previamente adquiridas, mas sofre influência de outras firmas ao seu redor, como por exemplo na relação fornecedor-comprador ou por compartilhar do mesmo mercado de trabalho. Portanto, essa dependência de trajetória também é válida para países, regiões e cidades, que se aproveitam dos conhecimentos que já dominam para continuar expandindo suas capacidades a um custo relativamente baixo em direção a setores tecnológicos relacionados.

No entanto, diversificar para tecnologias relacionadas à estrutura tecnológica

---

<sup>1</sup> Este localizado pode ter múltiplas interpretações, uma delas é geograficamente, mas nesta introdução estaremos nos restringindo à ideia do espaço do conhecimento, a não ser quando explicitado o oposto.

<sup>2</sup> Ainda que um conhecimento não formal.

existente nem sempre é favorável, podendo levar uma região a uma situação de *lock-in*. Tomando emprestada a noção de espaço do conhecimento, é um beco sem saída. Para países, podemos dizer que o *lock-in* acontece quando uma região fica presa a uma tecnologia de potencial inferior devido à sua dependência de trajetória (ARTHUR, 1989). No argumento de Arthur (1989), essa situação pode surgir devido a retornos crescentes de escala que tornam eventos históricos aleatórios determinantes na escolha de uma tecnologia em detrimento de outra – potencialmente melhor para o desempenho econômico. Abusando da analogia do fogão, é como se, por conta de um raio, alguém tivesse descoberto o fogo e daí foi capaz de recombina esse conhecimento para inventar o fogão. Mas, não fosse isso<sup>3</sup>, a cidade a que pertencia esse alguém não estaria presa produzindo fogões, mas construindo alguma outra coisa qualquer que trouxesse mais prosperidade àquela região, como navios. Seja qual for o evento “aleatório”, o ponto é que nem sempre é a melhor alternativa para uma região se restringir apenas ao campo de possibilidades que dispõe a partir de capacidades disponíveis localmente. Capacidades são simplesmente aqueles conhecimentos associados a um ente – pessoa, firma, região, país – que são importantes para o processo produtivo (BALLAND et al., 2020). Por um lado, fazer uso dessas capacidades significa aproveitar de oportunidades disponíveis a um baixo custo, *low hanging fruit*<sup>4</sup>. Por outro lado, essas oportunidades podem gerar também menos benefícios do que seria possível caso a região explorasse outros conhecimentos.

Países atrasados estão particularmente suscetíveis a esse lado perverso da dependência de trajetória na medida em que são eles que possuem menos capacidades e, portanto, são menos capazes de inovar. Isto porque, como argumentado, a inovação é um processo localizado de recombinação de conhecimentos já adquiridos. Assim, quão menos tecnologias um país domina, menores são as suas possibilidades de recombinação, reduzindo a velocidade com que ocorrem inovações. Países atrasados possuem um conjunto de capacidades bastante reduzido para buscar combinações que gerem um alto retorno, enquanto países desenvolvidos dispõem de um enorme conjunto. Junto com o decrescente custo de inovar – ou seja, semelhante à ideia de

---

<sup>3</sup> Ou justamente por isso

<sup>4</sup> Expressão em inglês para aquilo que é mais fácil de se obter.

retornos crescentes de escala, mas ao invés de produção, inovação –, este processo acaba resultando em uma lógica de causação cumulativa (MYRDAL, 1957): países mais desenvolvidos inovam mais rapidamente, aumentando ainda mais a distância tecnológica em relação aos demais. Portanto, aprofundando ou, ao menos, perpetuando desigualdades tecnológicas internacionais. Assim, além de também estarem sujeitas a uma situação de *lock-in*, países tecnologicamente atrasados enfrentam um desafio adicional em renovar a seus conhecimentos tecnológicos, no limite os impedindo de alcançar países desenvolvidos.

Como, então, romper essa dependência de trajetória e renovar a estrutura tecnológica? Um possível caminho é através de laços internacionais (BATHELT; MALMBERG; MASKELL, 2004). Descobrir qual alternativa, combinação, funciona melhor para realizar um determinado processo é incerto, trabalhoso e envolve muita tentativa e erro. Contudo, às vezes uma conversa com alguém que já tenha passado pelas mesmas dificuldades serve como um atalho para evitar tentar aquilo que já se sabe não funcionar. Sobre essa intuição, Bathelt, Malmberg e Maskell (2004) argumentam que se conectar com outras regiões pode ser uma forma de encontrar capacidades ausentes localmente. Neste sentido, estabelecer o contato com outros países que dominem técnicas úteis é essencial para o crescimento e a inovação tecnológica. Por meio de canais de comunicação e comércio, é possível transmitir conhecimento produtivo crítico de um país para o outro. Assim, a ausência de capacidades num país pode ser compensada pelo contato com outros países que possuam essas capacidades. Ou seja, um país pode se apoiar em conhecimentos presentes fora dali para se diversificar tecnologicamente, alavancando suas possibilidades de busca no espaço do conhecimento.

A efetividade deste processo está condicionada à capacidade de absorção local, que por sua vez depende das habilidades desenvolvidas localmente. Por um lado, é preciso que esse conhecimento (acessível por meio dos laços estabelecidos por aqueles canais de comunicação) seja suficientemente parecido para que seja compreensível. Alguém nascido na Alemanha muito provavelmente sabe falar alemão. No entanto, seria de pouca ajuda tentar aprender alemão com essa pessoa caso ela não saiba

também português<sup>5</sup>. Por outro lado, se o conhecimento transmitido for muito similar, também o proveito deste contato será pequeno. Para aprender alemão, é preciso achar um professor que saiba português, para que seja possível a comunicação, mas *também* alemão, para poder ensinar. É importante que as conexões estabelecidas deem acesso não aos mesmos conhecimentos já presentes localmente, mas a conhecimentos novos. Ainda mais se levarmos em conta que estabelecer estes canais é um processo custoso (BALLAND; BOSCHMA, 2021), quer dizer, comunicar-se com quem só sabe o que você sabe deve mais atrapalhar do que ajudar. Logo, deve haver uma distância cognitiva ótima (NOOTEBOOM, 2000) no contato entre dois ou mais países que é ideal para induzir o desenvolvimento tecnológico. Dito de outra maneira, os países tem mais a ganhar estabelecendo contato com países que tenham uma estrutura tecnológico-produtiva nem tão próxima e nem tão distante, mas *complementar* à sua.

Entretanto, a literatura sobre diversificação tecnológica tem raramente tratado sobre o impacto desses laços internacionais na trajetória tecnológica local (YEUNG, 2021; BOSCHMA, 2022), menos ainda do ponto de vista da complementariedade. Quando tratados, costumam se concentrar apenas na similaridade entre os conhecimentos que os países dominam. Em especial, Balland e Boschma (2021) desenvolvem um indicador para complementariedade tecnológica entre regiões, mas aplicam esse indicador apenas para regiões europeias. O objetivo deste trabalho é usar deste indicador para avaliar a evidência empírica ao nível internacional, perguntado se o estabelecimento de laços com países cuja estrutura tecnológica é complementar à local auxilia no processo de inovação tecnológica. Para isso, serão utilizados dados de patente dos escritórios dos Estados Unidos (USPTO) e Europa (EPO), abrangendo o período de 2000 a 2019. Assim, apoiando-se em Balland e Boschma (2021), este trabalho contribui para a literatura ao usar de um indicador que trata explicitamente a questão da complementariedade em laços tecnológicos para um amplo conjunto de dados para o qual não há paralelo na literatura de aplicação da metodologia proposta.

---

<sup>5</sup> Supondo que quem está tentando aprender é alguém que fala português.



O trabalho tem a seguinte estrutura. No [Capítulo 2](#), revisa-se a literatura, trazendo as principais contribuições e suas limitações. No [Capítulo 3](#), apresenta-se a metodologia que será utilizada para adereçar as lacunas existentes, primeiro retomando os problemas enunciados, depois, na [seção 3.1](#), passa-se pela construção dos indicadores empíricos utilizados e finalmente, na [seção 3.4](#), apresenta-se a fonte de dados com os quais os indicadores serão calculados bem como algumas estatísticas básicas conforme se apresenta os procedimentos empíricos. No [Capítulo 4](#), primeiro é feita uma análise exploratória dos dados na [seção 4.1](#) onde se usa de diferentes maneiras de apresentar os dados e os indicadores calculados para começar a responder e conduzir as perguntas que serão tratadas mais rigorosamente através de regressões estatísticas na [seção 4.2](#) com modelos PROBIT e LOGIT para a entrada em novos setores tecnológicos. Finalmente, o [Capítulo 5](#) conclui esse trabalho e nos Anexos são apresentadas tabelas com os dados utilizados no [Apêndice A](#) e com análises de robustez das regressões no [Apêndice B](#).

## 2 Diversificação Tecnológica: o privilégio da localidade

O objetivo deste capítulo é apresentar as principais contribuições para a literatura recente de diversificação tecnológica, buscando com isso caracterizar os impasses atuais e no que esse trabalho pode contribuir. Primeiro, parte-se das contribuições que fundamentaram e guiaram a pesquisa atual, que remonta principalmente à abordagem da complexidade econômica e a inovação teórica de levar em conta a heterogeneidade dos produtos, em contraposição à concepção de que todos os produtos influenciam igualmente na produção e crescimento dos países. Depois são resumidos alguns dos principais achados empíricos que a literatura alcançou quanto à diversificação tecnológica, que se centram principalmente na diversificação relacionada – a ideia de que os países diversificam sua estrutura produtiva rumo a setores próximos. Finalmente, concluímos tratando das principais limitações dessa abordagem, que se resumem ao olhar excessivamente auto-centrado na localidade, ignorando outras formas de proximidade que não estejam imbuídas na geografia. Junto a isso, aponta-se algumas contribuições que adereçam parte desses problemas e, ao fim, propondo um objeto de pesquisa que trate da importância de laços internacionais de colaboração para a performance tecnológica local.

### 2.1 Complexidade econômica e a heterogeneidade dos produtos

A literatura de diversificação tecnológica tirou grande proveito das técnicas desenvolvidas através da abordagem da complexidade econômica. [Hidalgo et al. \(2007\)](#) expandiram a ideia de que a produção e, portanto, o crescimento econômico dependem da quantidade de produtos que um país domina tecnologicamente, concepções já presentes nos modelos clássicos de crescimento endógeno, como em [Romer \(1990\)](#).

Em funções de produção do tipo variedade expansiva de produtos<sup>1</sup>, cada produto tem uma contribuição igual à produção, assim, para aumentar a produção basta dominar a produção de um novo tipo de produto. Ao invés de considerar homogeneidade, [Hidalgo et al. \(2007\)](#) consideram que os produtos são diferentes entre si e, logo, possuem uma influência diferenciada na produção de um país, isto é, que eles são *heterogêneos*. Portanto, a motivação das diferenças persistentes entre a renda dos países deixa de ser apenas diferenças tecnológicas, sendo motivada pela estrutura produtiva de cada país.

Este processo de manutenção de diferenciais de renda é em parte motivado pelo “custo de descoberta” ([HAUSMANN; RODRIK, 2003](#)). Empresários de países subdesenvolvidos não conhecem a “função de produção” da cada setor. Isto é, mais do que o simples custo de produzir um item, eles se deparam com uma incerteza em relação ao quanto essa produção irá custar. No entanto, a partir do momento que esse risco é tomado – e se o empreendimento se demonstrar rentável –, o conhecimento é rapidamente difundido no resto da sociedade, gerando uma externalidade positiva não incorporada pelo empreendedor pioneiro. É um problema similar ao de inovação em países desenvolvidos. Neste caso, a solução ao problema da internalização dos lucros adicionais advindos da descoberta são contrabalanceados por uma política ativa de proteção dos direitos de usufruto desses benefícios: patentes. Em contrapartida, muitas vezes não há uma solução centralizadora para o “custo de descoberta”. Assim, não resolvida, esta externalidade atua no sentido da perpetuação das desigualdades entre países desenvolvidos e subdesenvolvidos, na medida em que estes últimos não conseguem acessar produtos que levem a um patamar de renda superior.

Usando de técnicas de análise de rede, [Hidalgo e Hausmann \(2009\)](#) fornecem uma abordagem empírica para trazer à luz algumas dessas ideias e dinâmicas. Os autores constroem o índice de complexidade econômica, mostrando que países mais complexos são também aqueles mais ricos e que crescem mais. Basicamente, o

---

<sup>1</sup> Do inglês, *Expanding Variety of Products*. A comparação não é sem suas limitações, isto porque neste tipo de modelo o que conta são produtos intermediários, não finais. Apesar disso, [Hidalgo et al. \(2007\)](#) estabelece esse paralelo, começando inclusive o artigo justamente tratando destes modelos – e também daqueles de *quality ladder*, como, teoricamente, a inovação é similar, por simplicidade, omite-se aqui.

que torna um país complexo é ser especializado na produção de muitos produtos (diversidade) e em produtos pouco ubíquos<sup>2</sup>, ou seja, raros. Não basta apenas produzir um item raro, como diamante, para ser complexo, é necessário dominar diversos produtos ao mesmo tempo, já que, como se assume, o domínio tecnológico requer um conhecimento tácito e difícil de ser adquirido (HAUSMANN et al., 2014; BALLAND et al., 2022). Ou seja, a produção é um indicativo dos conhecimentos que um país possui. Hidalgo e Hausmann (2009) conseguem juntar essas duas informações – diversidade dos países e ubiquidade dos produtos –, no caso por meio da suas exportações<sup>3</sup>, através do método das reflexões, que interage uma informação com a outra. Assim, países complexos são aqueles capazes de produzir itens complexos, na medida em que esses produtos requerem que múltiplos conhecimentos entrem em cena.

No fundo, a hipótese que se assume é que as atividades que um país ou região realizam estão relacionadas a um conjunto de conhecimentos úteis à produção que chamaremos de *capacidades*<sup>4</sup> (BALLAND et al., 2022). De modo mais formal<sup>5</sup>, podemos pensar numa matriz para países,  $c$ , e produtos,  $p$ , que representa a produção mundial – ou, como *proxy*, suas exportações –,  $M_{cp}$ . Quando assumimos que as atividades desempenhadas por um país no fundo representam, ou são possíveis graças a, as capacidades que estes mesmos países dominam (conhecem), estamos dizendo que existe alguma relação, que daremos o sugestivo nome de *função*, tal que  $f : C_{ca} \rightarrow M_{cp}$ , onde  $C_{ca}$  representa as capacidades dos países, que são basicamente as habilidades,  $a$ , que eles dominam. Mas para transformar a habilidade,  $a$ , em produto,  $p$ <sup>6</sup>, é necessário conhecer ainda uma outra informação, a dizer, quais habilidades são necessárias para produzir um determinado produto,  $P_{pa}$ . Por exemplo, para produzir uma camiseta é necessário usar de diversas habilidades ao mesmo tempo: é preciso saber cultivar o algodão, assim como extraí-lo; transformar essa

<sup>2</sup> Ubíquo é a característica daquilo que está presente em tudo, onipresente.

<sup>3</sup> O que, também, requer a hipótese de que a exportação representa fidedignamente a produção de um país.

<sup>4</sup> Em inglês, *capabilities*.

<sup>5</sup> O formalismo aqui apresentado tem grande inspiração em Balland et al. (2022), que trata justamente sobre o *paradigma da complexidade econômica*.

<sup>6</sup> Como qualquer um familiarizado com álgebra linear já estaria ansioso por fazer.

matéria bruta em um tecido; saber costurar esse tecido num formato específico que os humanos possam vestir sob seu tronco; e, nas civilizações modernas, coordenar todas essas atividades por meio de gerenciamento e providenciar a logística necessária para que a camiseta chegue ao consumidor final. Se apenas uma dessas habilidades não existir, a camiseta simplesmente não pode ser produzida. Podemos resumir essa ideia através da seguinte equação:

$$\mathbf{M} = \mathbf{C}\mathbf{P}^\top, \quad (2.1)$$

onde  $\mathbf{M} = M_{cp}$ ,  $\mathbf{C} = C_{ca}$  e  $\mathbf{P} = P_{pa}$ .

Mais ainda, mesmo que todas essas habilidades coexistam ainda existe uma certa probabilidade de um produto (que necessita delas) não “vir a luz”. Como destaca [Balland et al. \(2022\)](#), a existência de capacidades e a efetivação delas numa atividade não precisa ser determinística – tal como uma Leontief – podendo ser estocástica. Isto é, tendo essas capacidades, existe uma probabilidade de que o país venha a produzir aquele produto – apesar de que na ausência continuaria a ser impossível. De qualquer forma, o ponto fundamental é menos sobre se essa relação é determinística ou estocástica e mais sobre não sabermos o que exatamente está contido em  $\mathbf{C}$  e  $\mathbf{P}$ : podemos ver apenas  $\mathbf{M}$ ! Portanto, mesmo que estejamos interessados em como chegar até as capacidades que nos permitem produzir mais,  $\mathbf{C}$ , ou com as habilidades que precisamos combinar para chegar num determinado produto,  $\mathbf{P}$ , a discussão – principalmente empírica – tem de ficar centrada na produção e assumir uma hipótese, quase por necessidade, de que existe alguma relação tal como em (2.1). Ou seja, o que esperamos é que  $\mathbf{P}$  e  $\mathbf{C}$  sejam de alguma maneira revelados por  $\mathbf{M}$ .

Como medimos  $\mathbf{M}$ ? A estratégia adotada por [Hidalgo et al. \(2007\)](#) e que desde então tem sido extensivamente aplicada foi usar a vantagem comparativa revelada (VCR) ([BALASSA, 1965](#)). A lógica por trás desse “indicador” é que, se um país dedica uma proporção da sua produção superior à que o mundo inteiro dedica, então ele é especializado naquela produção. Por exemplo, se o país A usa 50% de sua capacidade produtiva (do seu PIB) na produção de bananas e o mundo dedica 20%, então A é especializado em bananas. Se, porém, dedica apenas 5% para

a produção de aviões e o resto do mundo usa 10% de sua produção em aviões, então A não é especializado em aviões. Usando uma métrica relativa tão simples quanto essa podemos definir se um país domina ou não a produção de um determinado item. No exemplo citado, dizemos que A *sabe*, ou domina, cultivar (produzir) bananas. O próprio nome já diz muito: comparativa porque é relativa; revelada porque assumimos o domínio tecnológico como algo velado.

Na prática, temos que  $M_{cp} = 1$  se:

$$VCR_{cp} = \frac{X_{cp} / \sum_i X_{cp}}{\sum_r X_{cp} / \sum_r \sum_i X_{cp}} > 1, \quad (2.2)$$

e  $M_{cp} = 0$  caso contrário, onde  $X$  são as exportações de um país  $c$  de um determinado produto  $p$  – novamente, porque assumimos que as exportações são um substituto próximo das informações contidas na produção.

Com isso, temos uma matriz com zeros ou uns da qual podemos tirar aquelas duas informações essenciais a que nos referimos no começo dessa seção: diversidade e ubiquidade. A primeira é apenas quantos produtos um país domina e a segunda quantos o produzem. Em  $\mathbf{M}$  isso significa, respectivamente, a soma na linha (onde está o país) e na coluna (onde está o produto). Hipoteticamente,  $\mathbf{M}$  pode ser representado como:

$$\mathbf{M}_{hip} = \begin{matrix} & \begin{matrix} Soja & Camiseta & Computador \end{matrix} \\ \begin{matrix} Brasil \\ Estados Unidos \\ China \end{matrix} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \end{matrix} \quad (2.3)$$

Formalmente, podemos definir diversidade e ubiquidade como:

$$\begin{aligned} \text{Diversidade} &= k_{c,0} = \sum_p M_{cp} \\ \text{Ubiquidade} &= k_{p,0} = \sum_c M_{cp}, \end{aligned} \quad (2.4)$$

Podemos ponderar uma informação pela outra, através do método das reflexões (HIDALGO; HAUSMANN, 2009), através da seguinte recursão:

$$k_{c,n} = \frac{1}{k_{c,0}} \sum_p M_{cp} k_{p,n-1} \quad (2.5)$$

$$k_{p,n} = \frac{1}{k_{p,0}} \sum_c M_{cp} k_{c,n-1}, \quad (2.6)$$

Ao substituir (2.6) em (2.5):

$$\begin{aligned} k_{c,n} &= \frac{1}{k_{c,0}} \sum_p M_{cp} \left( \frac{1}{k_{p,0}} \sum_{c'} M_{c'p} k_{c',n-2} \right) \\ &= \sum_{c'} \frac{\sum_p M_{c'p} M_{cp}}{k_{c,0} k_{p,0}} k_{c',n-2} \end{aligned} \quad (2.7)$$

De acordo com Hausmann et al. (2014), (2.7) se resolve quando  $k_{c,n} = k_{c,n-2} = 1$ <sup>7</sup>. A ideia é: o método das reflexões é um processo iterativo, assim, o problema está resolvido quando o sistema está parado, “refletir” um dado no outro não altera em mais nada. Aquela condição é satisfeita para a matriz de autovetores da parte na fração,  $\tilde{M} = \frac{\sum_p M_{c'p} M_{cp}}{k_{c,0} k_{p,0}}$ . O primeiro é pouco informativo, são apenas valores iguais, mas o segundo não: este é o nosso índice de complexidade (ECI)<sup>8</sup>,  $\vec{K}$ ! Apenas normalizamos os valores, tirando a média e dividindo pelo desvio-padrão:

$$ECI = \frac{\vec{K} - \langle \vec{K} \rangle}{dp(\vec{K})} \quad (2.8)$$

onde  $\langle \rangle$  representa a média e  $dp$  o desvio-padrão.

<sup>7</sup> Observe que se usarmos  $n$  ímpar teremos uma informação enganadora: não é a diversidade dos países ponderada, é a ubiquidade! O inverso vale caso tivéssemos feito a substituição em equações invertidas.

<sup>8</sup> Do inglês, *Economic Complexity Index*.

[Balland et al. \(2022\)](#) traz uma outra maneira de exposição mais sucinta – menos intuitiva mas que simplifica o procedimento computacional. Primeiro definimos as matrizes diagonais  $\mathbf{D}$  e  $\mathbf{U}$  de dimensões  $c$  e  $p$  respectivamente, cujos valores são o inverso da diversidade e ubiquidade. Isto é:

$$\mathbf{D} = D_{cc'} = \begin{cases} \frac{1}{\sum_p M_{cp}} & , \text{ se } c = c' \\ 0 & , \text{ se } c \neq c' \end{cases}$$

$$\mathbf{U} = U_{cc'} = \begin{cases} \frac{1}{\sum_c M_{cp}} & , \text{ se } p = p' \\ 0 & , \text{ se } p \neq p' \end{cases}$$

Assim, podemos definir as matrizes:

$$\begin{aligned} A_{cp} &= \mathbf{A} = \mathbf{D}\mathbf{M} \\ B_{cp} &= \mathbf{B} = \mathbf{M}\mathbf{U} \end{aligned} \tag{2.9}$$

Com elas, podemos espelhar uma informação na outra a fim de obter os índices de complexidade do país e do produto através do segundo maior autovetor associado a, respectivamente,  $\mathbf{A}\mathbf{B}^\top$  e  $\mathbf{A}^\top\mathbf{B}$ .

Usando dessas técnicas, [Hidalgo e Hausmann \(2009\)](#) observam que o índice de complexidade dos países é um bom preditor do nível de renda per capita dos países. O fundamental sobre essa abordagem metodológica trouxe uma forma empírica de se verificar, ou demonstrar, algumas das ideias que vinham já desde o estruturalismo latino-americano ([GALA; ROCHA; MAGACHO, 2018](#)). Contudo, ao invés de dar um enfoque em diferenças setoriais estilizadas, como a que países subdesenvolvidos centram sua produção em bens primários e os desenvolvidos em bens desenvolvidos, o índice “deixa os dados falarem”, ao não assumir nenhuma classificação *a priori* – a não ser os básicos conceitos de diversidade, ubiquidade e sua interação. Mais do que isso, [Hidalgo et al. \(2007\)](#) assumem que existem um conjunto de fatores produtivos específicos associados a cada produto. Isto é, cada produto está associado a um conhecimento tácito, difícil de ser propagado. Como veremos na próxima seção,



essa concepção abriu um vasto campo de pesquisa e, em seu bojo, novas questões a respeito da convergência de renda entre os países.

## 2.2 Diversificação relacionada e o Princípio do Relacionamento

Uma das conclusões da literatura recente é que, como produzir algo requer um conhecimento específico difícil de ser adquirido, os países costumam explorar suas possibilidades de diversificação em direção a produtos próximos a sua estrutura produtiva. Vale dizer que a própria noção de proximidade entre produtos traz consigo a concepção de que os produtos são diferentes entre si. É impossível estar mais próximo de um produto do que do outro caso os produtos sejam homogêneos: a proximidade seria uma constante apenas! Mas como medir proximidade? Através de uma concepção correlata, [Hidalgo et al. \(2007\)](#) definem que um produto está mais próximo do outro conforme a frequência com que aparecem em uma mesma localidade (no caso, os países).

Outras métricas já foram adotadas, como a de entropia ([FRENKEN; OORT; VERBURG, 2007](#)) que assume que itens dentro de uma mesma classificação (tecnológica, produtiva) são similares. Contudo, essa abordagem metodológica é, primeiramente, estática: mantendo a mesma classificação, dois produtos que eram próximos há 30 anos atrás continuam a ser agora – na verdade, para qualquer horizonte temporal desde a criação da definição. Em segundo, é rígida, ela supõe que classificações são precisas o suficiente para medir a co-presença de um mesmo tipo de conhecimento, excluindo em seu bojo a possibilidade de que conhecimentos de múltiplas denominações interajam entre si. A ideia de co-localização é, ao mesmo tempo que simples e intuitiva, dinâmica e flexível. Por um lado, ela é flexível porque “deixa os dados falarem”. Um exemplo ajuda. Se, quando um país é especializado na produção de uvas, ele costuma ser também na de vinho, então dizemos que uva e vinho são próximos um ao outro. Porém, se observamos que vinho e aviões não são produzidos pelo mesmo país ao mesmo tempo, dizemos que vinho e avião estão distantes. Por outro lado, se uvas e vinho passarem a ser produzidos em lugares

diferentes de uma década para a outra, esses itens provavelmente passaram a envolver capacidades distintas ao longo tempo e a alternativa de observar proximidade por co-localização permite captar essa dinâmica.

Vamos focar aqui na definição de proximidade por co-localização. Seguindo o formalismo desenvolvido na última seção, podemos descrever a proximidade entre dois produtos  $p$  e  $p'$  como:

$$\phi_{pp'} = \frac{\sum_c M_{cp} M_{cp'}}{\max(k_{p,0}, k_{p',0})}, \quad (2.10)$$

onde usamos da abordagem conservadora de dividir pelo valor máximo da ubiquidade entre os dois produtos a fim de pegar a probabilidade condicional mínima de co-existência. No numerador,  $M_{cp} M_{cp'}$  define co-existência: somamos um apenas quando  $p$  existe em  $c$ ,  $M_{cp} = 1$ , e em  $p'$  também,  $M_{cp'} = 1$ .

De (2.10), podemos expandir para o índice de *densidade* que indica quantos conhecimentos similares a um determinado produto,  $p$ , uma localidade tem. Formalmente:

$$\omega_{cp} = \frac{\sum_{p'} M_{cp} * \phi_{pp'}}{\sum_{p'} \phi_{pp'}}. \quad (2.11)$$

Basicamente, (2.11) soma a proximidade de  $p$ , no denominador, com todas as demais tecnologias<sup>9</sup> e, no numerador, com todas os demais produtos que  $c$  é especializado. Assim, mesmo que um país não saiba produzir  $p$ , ele pode ter diversas capacidades (demonstradas) em outros produtos que são relacionados, sendo este o caso, devemos observar empiricamente isto por meio de altos valores de *densidade*.

Empiricamente, a ideia de que os países teriam mais facilidade em se especializar em novos produtos similares àquilo que eles já sabem fazer se mostrou válida sob distintas metodologias (HIDALGO et al., 2018). Porque adquirir um novo

<sup>9</sup> Pela forma que definimos, essa *somanão* resulta em 1 – pois escolhemos a probabilidade condicional mínima.

conhecimento é um processo custoso, quanto mais conhecimentos similares a um produto o país tem, mais provável é que ele diversifique em direção àquele setor. Nos termos da Introdução, não é só que a busca localizada é mais fácil, consciente ou não, é de fato o que acontece com mais frequência. A evidência para esse processo se espalha por diferentes regionalizações e tipos de atividades (HIDALGO et al., 2018). A especificação tradicional encontra que a probabilidade de se especializar, ou entrar, em uma nova atividade – produto, setor, tecnologia ou ocupação – aumenta quando a região é especializada em atividades similares – com alta densidade no produto! A seção 3.1 fornece uma definição formal de entrada numa nova atividade.

Fundamentalmente, o que está em jogo é a geografia. Justamente a literatura da *geografia* econômica se aproveitou daquelas intuições lançadas por Hidalgo et al. (2007) para observar que, quando um conhecimento se encontra disponível num mesmo lugar, ele é muito mais facilmente transmitido (BOSCHMA, 2017). Portanto, o que se faz é privilegiar essa camada de proximidade que está ligada ao espaço, associando em seu bojo conhecimento e geografia.

Inicialmente, o foco foi internacional: será que aquela heterogeneidade entre os produtos se traduz em alguma dinâmica específica no padrão de exportação dos países? Hidalgo et al. (2007) observa para dados internacionais de comércio ao nível SITC-4<sup>10</sup> para 775 classificações de produtos que a probabilidade de um país entrar em um determinado produto aumenta de acordo com a proximidade entre essas classes. Uma das motivações para se usar dados de exportação é porque faltam informações granulares da produção dos países. Parte desse problema tem sido adereçado com estudos que buscam expandir as classificações de tabelas de insumo-produto, que envolvem também uma série de novas potencialidades em termos do fluxo internacional de produtos e da incorporação das técnicas desenvolvidas na literatura da geografia econômica com a de cadeias globais de valor (YEUNG, 2021; BOSCHMA, 2022). No entanto, à época da publicação de Hidalgo et al. (2007), além de inédita, esta abordagem conferia um grau importante de legitimidade da análise.

<sup>10</sup> Do inglês, SITC é “Standardized International Trade Classification” e o 4 se refere a serem 4 dígitos.

Isto porque, como os autores argumentam, dados de comércio são bastante confiáveis uma vez que os países estão muito interessados em controlar o que entra e sai de suas fronteiras, sendo possível inclusive checar as informações de alfândega de um país com a do outro, adicionando ainda outra camada de robustez a eventuais erros de medição.

Posteriormente, essa literatura ganha novos relevos, expandindo o horizonte geográfico de agregação dos países e a lente dos dados de exportação em direção a regionalizações mais desagregadas e fontes distintas de dados como de emprego, pesquisa e patente. [Neffke et al. \(2011\)](#) expandiu essa concepção de países para regiões ao nível subnacional, usando dados de firmas, ao observar que entre 1969 a 2002 uma indústria tinha maior probabilidade de entrar em uma região sueca quão mais similar ela fosse às indústrias pre-existentes naquela região. [Boschma, Balland e Kogler \(2015\)](#) usam dados de patentes do USPTO<sup>11</sup> de áreas metropolitanas dos Estados Unidos entre 1981 e 2010 para observar que a entrada de uma região em uma tecnologia é mais frequente quando a região já dominava outras tecnologias similares. [Balland et al. \(2019\)](#) faz uma análise similar mas incluindo outros países europeus usando dados de patente entre 1990 a 2009 da OCDE-REGPAT derivado da base do PATSTAT. Eles encontram que é mais provável que uma região NUTS-2<sup>12</sup> se especialize em uma nova tecnologia quando esta região já domina outras similares. Mais ainda, eles encontram que, em geral, é mais difícil uma região se especializar em tecnologias complexas<sup>13</sup>, apesar de que esse efeito seja inverso dentre aquelas tecnologias muito próximas à estrutura tecnológica local – indicando que as regiões se aproveitam dessa proximidade avançando rumo a produtos complexos que sejam mais factíveis às suas capacidades. Quanto ao sul global, [Freitas \(2019\)](#) encontra a partir de dados de emprego entre 2006 a 2018 que a probabilidade de uma região brasileira se especializar em um novo setor está positivamente relacionada com a proximidade deste setor à estrutura produtiva local. Portanto, corroborando a ideia de que essas evidências são pervasivas de diferentes regionalizações.

<sup>11</sup> *United States Patent and Trademark Office.*

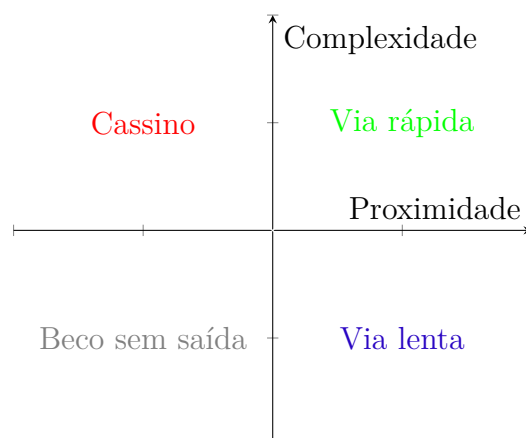
<sup>12</sup> *Nomenclature of territorial units for statistics.*

<sup>13</sup> Considerando apenas a especificação com efeitos-fixos.

Hidalgo et al. (2018) sintetizam esses diferentes achados empíricos sob a cunha de “Princípio do Relacionamento”<sup>14</sup>, que é a ideia de que países e regiões, a diferentes níveis de agregação, são mais propensos a entrar em atividades que sejam mais relacionadas à estrutura produtiva local. Li e Neffke (2022) mostram que este princípio é robusto a diferentes especificações, o que fazem usando dados da Dun and Bradstreet com uma cobertura de 100 milhões de estabelecimentos ao redor do mundo – e quase completa para o Estados Unidos – entre 2011 e 2019. No entanto, sua análise foca principalmente em estabelecimentos baseados nos Estados Unidos, em função da cobertura ser grande o suficiente para representar aquele país, assim limitando também a conclusão para a economia americana. Apesar disto, o que encontram adiciona ainda outra camada ao “princípio do relacionamento”: a ideia de que ele vale para múltiplas especificações.

Se diversificar para setores relacionados é menos custoso, uma estratégia *inteligente* – no sentido de diminuir os custos e aumentar os ganhos – de especialização é mirar em atividades relacionadas e promissoras. Pensando nisso, Balland et al. (2019) argumentam que regiões e países devem focar naquelas atividades mais próximas à estrutura produtiva local e que sejam mais complexas. Isto porque, por um lado, avançar para setores mais próximos envolveria riscos menores, já que a probabilidade de sucesso na entrada nestas atividades é maior, como as evidências do *Princípio do Relacionamento* apontam. Por outro lado, atividades mais complexas costumam estar associadas a economias mais ricas (HAUSMANN et al., 2014), assim, sendo consideradas mais promissoras para o desenvolvimento de uma região. Como vimos, um país “ser complexo” sinaliza que ele domina muitas capacidades e, portanto, seriam aqueles que são capazes de executar muitas atividades distintas, algumas das quais poucos o conseguem. Analogamente, um produto complexo é aquele que incorpora – ou, de modo menos antropomórfico, está associado com, requer – diversos conhecimentos específicos, tácitos, e que, portanto, gerariam efeitos de transbordamento no resto da economia. Isto porque, como assumimos ao não só definir *heterogeneidade* entre os produtos mas similaridade também, algumas atividades devem partilhar capacidades em comum. Usando de um exemplo hipotético,

<sup>14</sup> Ou, em inglês, *Principle of Relatedness*.

**Figura 1** – Especialização Inteligente

Fonte: Elaboração Própria a partir de [Balland et al. \(2019\)](#)

o mesmo cuidado que o cultivador de uvas deve ter ao guardar as uvas que colheu em seu armazém, a vinícola deve ter no zelo do seu estoque de *insumos*. De forma esquemática, a [Figura 1](#) resume essas ideias representando quatro estratégias possíveis que uma região ou país podem seguir<sup>15</sup>. Idealmente, a estratégia a ser seguida é aquela representada pela Via Rápida, envolvendo baixo custo e alto ganho. Os nomes são intuitivos: a estratégia de Cassino, por exemplo, é aquela que envolve um grande prêmio mas a um alto risco<sup>16</sup>.

No entanto, deixados à própria sorte, países e regiões atrasadas enfrentam uma condição estrutural que obsta o seu desenvolvimento. Nos termos do parágrafo anterior, a Via rápida simplesmente não está disponível para países atrasados. É justamente o oposto: a correlação entre setores próximos e complexos, como a evidência empírica tem apontado ([PINHEIRO et al., 2018](#); [HARTMANN et al., 2020](#)), é positiva para países avançados e negativa para países subdesenvolvidos. A [Figura 2](#) ilustra isso, baseada também em discussões feitas por [Boschma \(2021\)](#). Nela são apresentados dois países representativos, indicando o contraste entre um país normalmente definido como subdesenvolvido e outro claramente desenvolvido. Um deles é a Nigéria, cuja correlação entre a complexidade do produto, *pci*, e a proximidade tecnológica, *density*, é negativa, e o outro é os Estados Unidos, cuja

<sup>15</sup> Tradução livre dos termos que os autores usam que, seguindo do primeiro ao quarto quadrante respectivamente, seriam High-Road, Casino, Dead-End, Slow Road.

<sup>16</sup> Talvez, como em um cassino, o jogo seja favorável à casa.

correlação entre proximidade e complexidade tecnológica é positiva. Uma correlação negativa, como a que a Nigéria apresenta, indica que, quão mais complexa a atividade, mais distante dos conhecimentos disponíveis localmente e, portanto, mais difícil seria para um país ou região dominar essa atividade. Inversamente, uma correlação positiva indica que, quão mais complexa, mais próxima da estrutura produtivo-tecnológica local aquela atividade é <sup>17</sup> e, portanto, mais simples para o país se especializar nela. Assim, o que a [Figura 2](#) diz é que a estrutura produtiva-tecnológica de países desenvolvidos<sup>18</sup> gera um ciclo virtuoso em que, por dominar setores complexos, os dá fácil acesso a outros produtos e tecnologias também complexas. Analogamente, países subdesenvolvidos<sup>19</sup> enfrentam um ciclo de pobreza, em que, por não terem capacidades tecnológicas facilmente acessíveis, eles têm dificuldade em ganhar novos conhecimentos.

A evidência empírica corrobora essa ideia de que existem desigualdades persistentes no acesso a oportunidades entre países e regiões ([PINHEIRO et al., 2018](#); [PINHEIRO et al., 2022](#)). A [Figura 3](#) ilustra através de dados internacionais de comércio o que recentemente foi chamado de *Curva S*. Cada ponto do gráfico é um país ou região em um período de tempo, denominado como “país-ano”, definidos pela complexidade da região no eixo das ordenadas <sup>20</sup> e pela correlação entre a complexidade dos produtos potenciais (PCI)<sup>21</sup> com a proximidade tecnológica (Densidade)<sup>22</sup> – os valores são calculados apenas para os produtos em potencial, isto é, aqueles em que não se têm vantagem comparativa revelada (o *option set*). No eixo das ordenadas (y), temos o valor em um dado ano da inclinação da reta tal como na [Figura 2](#). Países subdesenvolvidos estão na base da curva, enquanto países desenvolvidos estão na parte superior dela. A curva S foi encontrada por estudos

<sup>17</sup> Uma limitação do método adotado em pesquisas da área é definir a especialização em termos relativos, através da vantagem comparativa revelada, seguindo [Balassa \(1965\)](#) (para mais detalhes ver [Capítulo 3](#)). Daí a [Figura 2](#) parece indicar uma ideia estranha: países desenvolvidos têm dificuldade de dominar atividades pouco complexas! Essa seria uma conclusão precipitada. O que a figura diz é apenas que, *relativamente*, países desenvolvidos têm mais facilidade em dominar novos setores complexos.

<sup>18</sup> Ou de um pelo menos.

<sup>19</sup> Ou de um pelo menos.

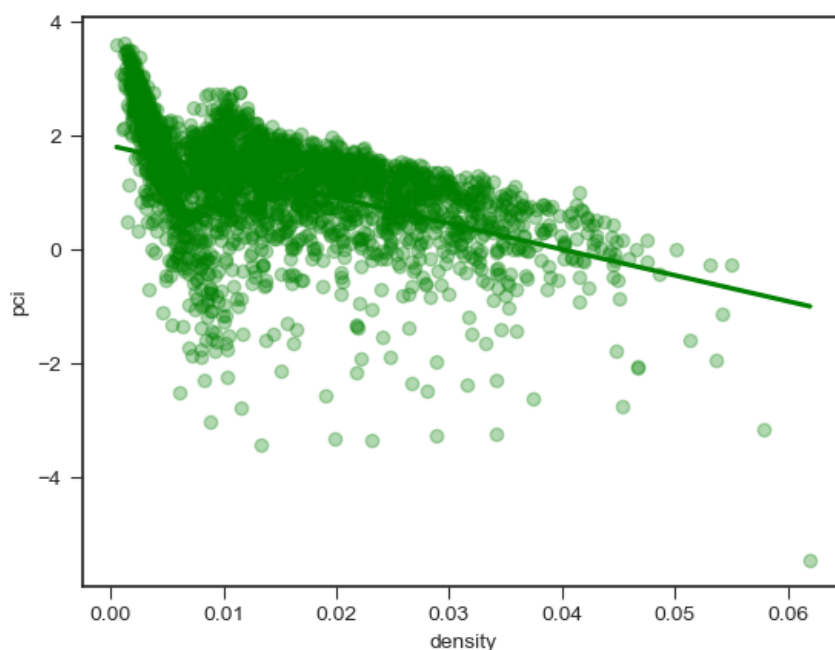
<sup>20</sup> Do inglês, *Economic Complexity Index* ou Índice de Complexidade Econômica.

<sup>21</sup> Do inglês, *Product Complexity Index* ou Índice de Complexidade do Produto.

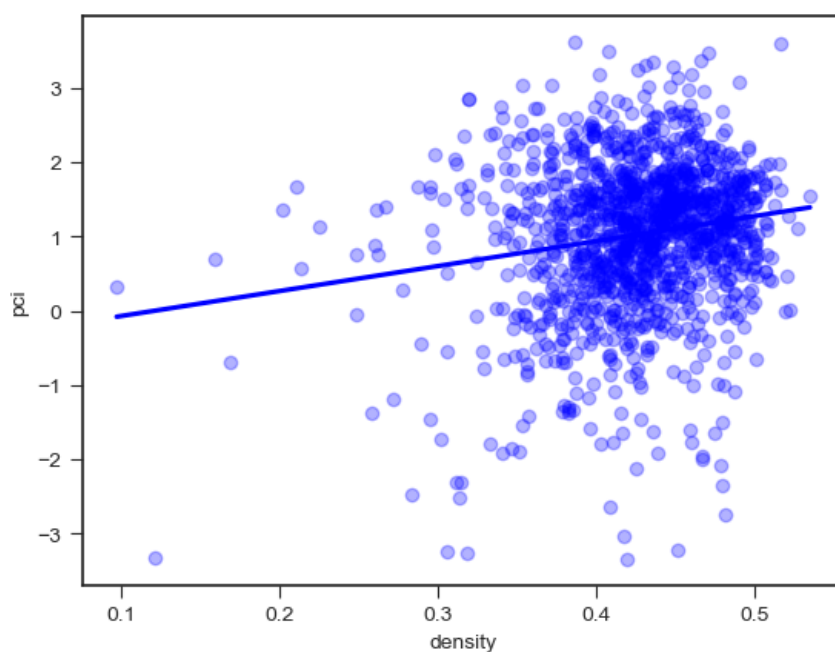
<sup>22</sup> Ver [Capítulo 3](#) para mais detalhes de como definimos Densidade.

**Figura 2** – Distintas oportunidades

(a) Nigéria



(b) Estados Unidos

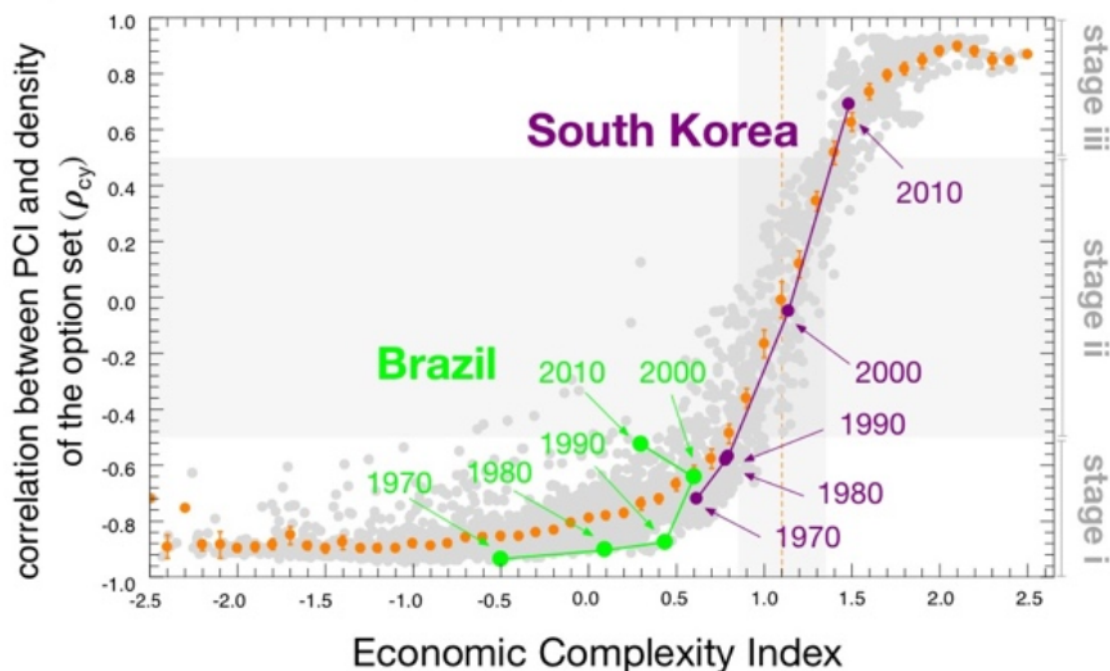


Nota: para mais detalhes de como as variáveis mencionadas são construídas, ver [Capítulo 3](#). Cada um dos pontos representa uma observação ao nível período-país-tecnologia. São apresentados dois países representativos, a Nigéria cuja correlação entre a complexidade do produto, *pci*, e a proximidade tecnológica, *density*, é negativa, enquanto a correlação entre proximidade e complexidade tecnológica é positiva para os Estados Unidos. Esta figura indica uma fonte da persistência das desigualdades entre os países, na medida em que países desenvolvidos têm maior facilidade de acessar tecnologias complexas, que favorecem o crescimento.

Fonte: Elaboração própria baseado em dados do USPTO.



Figura 3 – Curva S



Nota: no eixo das abscissas está a complexidade do país, enquanto no eixo das ordenadas está a correlação entre a complexidade e a densidade dos produtos. A figura destaca o Brasil (verde) e a Coreia do Sul (roxo) como casos representativos.

Fonte: Pinheiro et al. (2018).

usando dados internacionais de comércio (PINHEIRO et al., 2018; HARTMANN et al., 2020) e para dados para a Europa de emprego e patentes (PINHEIRO et al., 2022), reforçando a ideia de que há uma característica estrutural para a persistência das desigualdades entre países e regiões. Pinheiro et al. (2018) encontra que, via de regra, os países diversificam para setores relacionados, mas diversificam para setores não relacionados em um estágio intermediário de desenvolvimento, dando-lhes acesso a setores mais complexos. Ou seja, países neste estágio executam essa transição justamente entre a base e o topo da curva S.

Os icônicos casos limites do Brasil e da Coreia do Sul ilustram esse momento crítico de transição (BRITTO et al., 2019). Até antes de 1990, o PIB per capita brasileiro era consistentemente maior do que o coreano, no entanto, a partir daí a renda per capita coreana passa a divergir da Brasileira chegando a ser mais do que o dobro já no começo da década de 2010. Britto et al. (2019) argumentam que a diferença fundamental foi que, enquanto a Coreia do Sul se concentrou na produção

e exportação de produtos de alta intensidade tecnológica, o Brasil fez um movimento tímido nesse sentido, até vindo a reverter para produtos primários em anos recentes. Esta divergência é refletida na complexidade econômica dos dois países, com o Brasil abandonando uma tendência de complexidade crescente no começo da década de 90 enquanto a Coreia manteve a sua.

A [Figura 3](#) ilustra esse momento crítico de transição com os países acima citados. Como vemos, a correlação no meio da figura entre a proximidade a uma determinada atividade e a complexidade dela é basicamente nula – neste instante, tudo é possível! Tanto o Brasil quanto a Coreia se depararam com esse momento de transição, da base ao topo da curva S por volta dos anos 90 ([PINHEIRO et al., 2018](#)), mas apenas uma das duas economias conseguiu de fato realizar essa transição. Esta “escolha” por arriscar rumo a setores tecnologicamente avançados por parte da Coreia do Sul a alçou a um novo patamar de desenvolvimento, enquanto que o Brasil ficou preso naquilo que [Hartmann et al. \(2020\)](#) consideram a “armadilha da renda média”, um tema que vai além do escopo deste trabalho, ver [Agénor \(2017\)](#) para uma visão mais abrangente. De qualquer forma, o caso Coreia-Brasil apenas reforça a ideia de que países menos complexos ficam sujeitos à uma condição estrutural que lhes mantém em desvantagem em comparação a regiões desenvolvidas quanto às oportunidades de diversificação. Esta conclusão ressoa em diversos estudos que mostram que países e regiões mais complexos são mais ricos ([HIDALGO; HAUSMANN, 2009](#); [BALLAND; RIGBY, 2017](#), por exemplo) e que, quando uma região diversifica para setores relacionados mais complexos, elas crescem mais ([RIGBY et al., 2022](#)). Ou seja, regiões mais complexas se beneficiam de sua posição privilegiada de ter acesso a um rico campo de oportunidades a um baixo custo. Qual alternativa sobra então para países periféricos escaparem dessa condição?

### **2.3 O refúgio no exterior: a importância de colaborações internacionais**

Colaborações internacionais podem ser uma forma de romper o lock-in de países periféricos ao possibilitar o acesso a um estoque de conhecimento e capacidades

ausentes localmente. [Boschma e Iammarino \(2009\)](#) encontram que, entre 1995 a 2003, regiões italianas que comercializaram com outros países cuja produção era similar experimentaram um maior crescimento do emprego, do valor adicionado e da produtividade do trabalho. Usando a mesma regionalização, mas entre 2007 a 2012, [Ascani et al. \(2020\)](#) observam que as regiões são mais inovativas quando multinacionais (MNEs) locais estabelecem subsidiárias em outros lugares com conhecimentos relevantes para a estrutura produtiva local. Desta forma, essas MNEs funcionam como *gatekeepers*<sup>23</sup>, ou sejam, são a porta de entrada de um conhecimento difícil de ser adquirido mas presente no exterior. Nesse sentido, [Ascani et al. \(2020\)](#) contribuem ao destacar a relevância de agentes locais, em especial as MNEs, como mecanismo de transmissão e absorção de conhecimento produtivo de fora da região.

Um *agente* local de particular importância é o trabalhador, através da migração da força de trabalho regiões distintas podem compartilhar conhecimento produtivo. [Eriksson e Lengyel \(2019\)](#) usam de dados de empregado-patrão ao nível da planta para a economia sueca entre 1990 a 2008 e observam que redes sociais de trabalhadores densas são uma importante característica de regiões com alta performance, assim como laços interindustriais e extrarregionais atuam como ferramentas de compensação para a ausência de especialização industrial. Dividindo em dois períodos, os autores analisam cerca de 400 mil indivíduos em 50 mil plantas industriais em 1990 e, respectivamente, 800 mil e 100 mil em 2008. Redes sociais neste caso se refere a relações de trabalho, isto é, duas pessoas fazem parte de uma mesma rede social quando trabalham no mesmo local. Estar em um mesmo ambiente de trabalho facilita a troca de conhecimentos produtivos, assim, quando um trabalhador migra para outra indústria e/ou região, ele deve levar consigo o conhecimento absorvido do seu colega, potencializando efeitos de transbordamento. O tamanho do impacto positivo na produtividade local em decorrência da transmissão dessas capacidades produtivas deve variar também de acordo com a estrutura local. [Eriksson e Lengyel \(2019\)](#) ainda contribuem ao diferenciar entre laços que trazem maior diversidade e aqueles que trazem conhecimento especializado. Seus achados empíricos suportam a ideia de que os primeiros auxiliam mais locais com redes coesas, permitindo uma

---

<sup>23</sup> *i.e.*, porteiras.

diversificam o olhar auto-centrado, e os segundos, com redes abertas, auxiliando a focalizar naquilo que é relevante.

Outros estudos também seguem nesta linha destacando o impacto diferenciado na performance local de laços formados sob tipos distintos de redes. Uma distinção é a entre redes domésticas e internacionais. Scalera, Perri e Hannigan (2018) avaliam como a conexão extrarregional baseada em conhecimento afetou o escopo tecnológico de firmas norte-americanas entre 1990 a 2006. Escopo tecnológico é definido como a quantidade média de classes tecnológicas das patente aplicadas por uma firma, baseando-se nas definições de Hall, Jaffe e Trajtenberg (2001). A motivação dessa variável dependente é que classes tecnológicas são tomadas como blocos de conhecimento tecnológico, assim produzir uma patente com mais classes associadas deve estar relacionado a um processo de busca por muitos campos de conhecimento, sendo assim mais complexa de elaborar. Os autores encontram efeitos diferenciados entre o acesso a conhecimentos através de redes domésticas e internacionais. Os dois tipos de conexão levam a um aumento do escopo tecnológico de firmas americanas, contudo, laços domésticos apresentam um impacto decrescente<sup>24</sup>, enquanto laços internacionais não. A conectividade doméstica de uma firma foi calculada de acordo com a quantidade de endereços únicos dos inventores em uma CBSA<sup>25</sup> nas patentes aplicadas por uma firma. Similarmente, conectividade internacional é definida como a quantidade de endereços únicos que *não* estão naquela localidade.

No entanto, a influência do contato com outra firma depende de mais do que do compartilhamento de fronteiras, depende também da forma como se dá essa relação. Seguindo uma diferente estratégia, Turkina e Assche (2018) focam em conexões internacionais diferenciando entre laços horizontais – parceria de pesquisa, treinamento ou consórcio (*joint venture*) – e verticais – relação de comprador-fornecedor –, encontrando que seu impacto em termos da quantidade patentes aplicadas por clusters regionais varia de acordo com o *tipo* de cluster. Cada tipo

<sup>24</sup> O coeficiente do termo quadrático da variável de conexão doméstica é negativa.

<sup>25</sup> Do inglês, *Core-based Statistical Area*, que é uma definição geográfica norte-americana que une diferentes municípios sob uma mesma circuncrição geográfica, atualmente são 927 CBSA nos Estados Unidos. Para mais detalhes, ver <[https://www.census.gov/geo/reference/gtc/gtc\\_cbsa.html](https://www.census.gov/geo/reference/gtc/gtc_cbsa.html)>.

de conexão, vertical e horizontal, envolve uma rede associada distinta. Assim, para cada uma, é possível definir cluster centrais, aqueles que dentro de uma rede não apenas são altamente conectados mas possuem laços com outros clusters relevantes. Usando desta métrica, [Turkina e Assche \(2018\)](#) observam que clusters intensivos em trabalho se beneficiam mais de assumirem a centralidade em redes de conexões verticais e aqueles intensivos em conhecimento ganham mais com a centralidade na rede de conexões horizontais. Especificamente, um *cluster* em uma indústria específica existe em uma região quando o coeficiente locacional (CL) desta indústria é maior do que um, sendo CL simplesmente uma medida relativa, similar à ideia de vantagem comparativa revelada só que para empregos – na [Equação 2.2](#) trocamos as exportações por empregos (frequentemente, países por regiões e produtos por setores ou ocupações). Para observar isto, eles usam uma base coletada “manualmente”, ou seja, que compila dados de fontes diversas, desde bases oficiais de clusters, passando por relatórios públicos e indo até páginas na internet. No total, a base cobre o período de 2002-2014 para três indústrias intensivas em conhecimento: Aeroespacial, Biofarmacêutica e Tecnologia da Informação e Comunicação (TIC)<sup>26</sup>. A escolha deriva das características dessas indústrias, que vai da combinação de processos intensivos em trabalho e em conhecimento ao alto grau de inovação, colaboração e concentração espacial (permitindo a junção de distintas firmas num só cluster) que exibem. Além disso, os autores controlam seus resultados corrigindo a autocorrelação espacial nos erros e por métricas similares mas intra-firma, sem no entanto diferenciar entre conexões verticais e horizontais já que neste contexto não faria sentido, encontrando coeficientes em geral significativos e positivos na quantidade de patentes aplicadas por um cluster.

Como transparece, existem múltiplas dimensões daquilo que é próximo ([BOSCHMA, 2005](#)) – social, tecnológica, cognitiva, institucional, etc. –, assim, é de se esperar que o contato com outros lugares se manifeste de distintas maneiras, cada qual com diferentes impactos na trajetória local. Neste sentido, diferentes estratégias já foram adotadas para caracterizar empiricamente links internacionais ou extrarregionais – uma análise similar, mas ao nível subnacional –, como através de trocas

<sup>26</sup> ou ICT, do inglês, *Information and Communication Technology*.

comerciais (BOSCHMA; IAMMARINO, 2009), citações em patentes (MIGUELEZ; MORENO, 2018), co-patenteação (NONI; ORSI; BELUSSI, 2018; SANTOALHA, 2019; BARZOTTO et al., 2019; BALLAND; BOSCHMA, 2021), multinacionais (MNEs)<sup>27</sup> (ASCANI et al., 2020) e outras múltiplas fontes (MARROCU; PACI; USAI, 2013; CAPELLO; CARAGLIU, 2018; JUN et al., 2020). Por múltiplas fontes refere-se, por exemplo, a proximidade geográfica, institucional, tecnológica, social e organizacional (MARROCU; PACI; USAI, 2013) (no caso, para regiões europeias). Ao nível nacional, também já foram considerados laços de histórico colonial e idioma em comum (JUN et al., 2020).

Na verdade, essas diferentes formas de se ver a proximidade por meio de relações extra-locais aponta para uma forma alternativa de ver a importância delas: é uma dimensão específica de proximidade. Quando observadas localmente, essas características que auxiliam na performance local, como densidade tecnológica – isto é, o domínio em tecnologias similares facilitando a entrada em um novo setor tecnológico –, apenas ponderam a relevância destes fatores pela distância física, destacando a importância da espacialidade na difusão de conhecimento. Dito de outra forma, quando focamos em transbordamentos locais, estamos simplesmente potencializando alguma das várias dimensões de proximidade (cognitiva, institucional, etc.) pela proximidade espacial. Esta visão “para dentro” dá espaço para uma visão “para fora” a partir do momento em que começamos a tropeçar nas suas limitações, que é justamente aquela tendência à manutenção das desigualdades entre as regiões e países.

Uma limitação da literatura é que a evidência empírica disponível costuma focalizar na Europa, sendo que a evidência internacional tem mais frequentemente adereçado o impacto de conexões com países com conhecimentos similares. Bahar, Hausmann e Hidalgo (2014) desenvolvem uma métrica de similaridade produtiva e observam que a probabilidade um país exportar um produto aumenta em 65% quando seus vizinhos são exportadores deste produto. Jun et al. (2020) adicionam a essa evidência ao encontrar que a similaridade na (i) produção, (ii) exportação e (iii)

---

<sup>27</sup> Do inglês, *MultiNational Enterprises*.

importação está associada a uma maior probabilidade em exportar um produto. Ou seja, é mais provável que um país – o exportador – começar a exportar um produto a um outro país – o importador – quando: (i) já exporta muitos produtos similares outros países; (ii) os vizinhos de seus parceiros comerciais já importavam aquele produto; e (iii) os seus vizinhos já exportavam esse produto para outros lugares. Além disso, [Jun et al. \(2020\)](#) encontram que é mais provável um país exporte um produto que outros países de idioma e histórico colonial em comum já exportam. Esse efeitos aumentam quão mais diversificado, complexo e sofisticado tecnologicamente um país é, indicando que países desenvolvidos se beneficiam ainda mais desse tipo transmissão de conhecimento. Porém, quando esse impacto é reduzido quando se divide fronteiras com um país que já exporta aquele produto, apontando para uma dinâmica de competição pelo mercado internacional. Ao nível regional, há evidência de que, além da Europa, também a China se beneficiou de laços extrarregionais para a diversificação produtiva regional. Especificamente, através de dados regionais de exportação<sup>28</sup> entre 2002 a 2011, [Zhu, He e Zhou \(2017\)](#) encontram que regiões chinesas que importavam mais e recebiam mais investimentos diretos estrangeiros tinham maior probabilidade de se especializar em novos produtos.

Em especial, acessar capacidades não meramente similares mas complementares pode ser importante para a diversificação tecnológica local. [Balland e Boschma \(2021\)](#) desenvolvem uma medida de complementariedade tecnológica a fim de observar se é mais fácil para uma região se especializar em uma nova tecnologia quando há links com regiões que possuam capacidades relacionadas àquela tecnologia. A ideia do indicador é captar capacidades importantes para dominar uma tecnologia que estejam ausentes em uma região mas que sejam acessadas através da conexão com outras regiões – o indicador é descrito detalhadamente no [Capítulo 4](#). Usando dados de patente para 292 regiões NUTS-2<sup>29</sup> da Europa entre 1992 a 2016, [Balland e Boschma \(2021\)](#) mostram que a probabilidade de entrada de uma região numa tecnologia aumenta quando há conexões com regiões que são especializadas em

<sup>28</sup> Quer dizer, não é levada em conta a troca comercial *entre* regiões.

<sup>29</sup> Nomenclatura das Unidades Territoriais para Fins Estatísticos (NUTS). O 2 se refere ao nível da divisão territorial

tecnologias similares. Este efeito aumenta com a quantidade de laços, porém, os laços por si só na verdade diminuem a probabilidade de entrada na tecnologia, indicando que estabelecer laços irrelevantes com outras regiões traz mais malefícios do que benefícios<sup>30</sup>.

Mas por que diferentes tipos de laços têm diferentes impactos? Especificamente, por que laços complementares seriam mais relevantes do que simplesmente estabelecer contato com outros países? [Nooteboom \(2000\)](#) argumenta que existe um custo de oportunidade entre novidade e compreensibilidade. Ter acesso a um conhecimento novo fornece muitas oportunidades, mas ser “novo demais” pode ser simplesmente incompreensível. Assim, o proveito máximo de alguma interação social deve ser no meio do caminho: a uma distância cognitiva ótima. A [Figura 4](#) torna esse conceito didático. Tal como uma parábola invertida, o nível de aprendizado<sup>31</sup> através de um contato com outrem deve ser relativamente pequeno quando a distância cognitiva é muito pequena, é completamente compreensível, mas há muito pouco para se aprender. Assim, conforme a distância aumenta, deve haver um ganho de aprendizado, mas a uma taxa decrescente, pois parte do potencial de absorver aquele conhecimento é gasto compreendendo-o. Finalmente, depois de algum ponto crítico, o conhecimento oferecido por aquele laço passa a ser distante demais daquilo que se sabe, sua utilidade começa a decrescer. O foco em laços complementares é a tentativa de caracterizar justamente aquele ponto ótimo representado na figura.

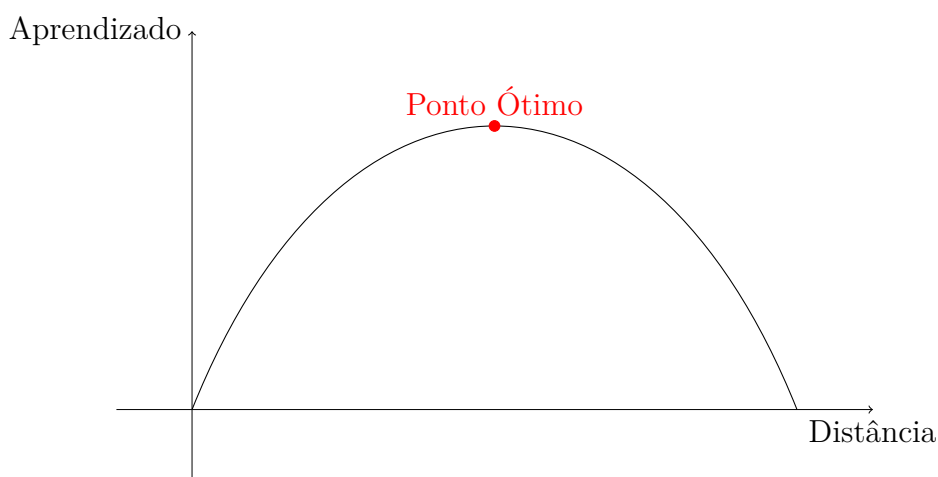
A importância de laços internacionais deve variar não apenas de acordo com a qualidade destes laços, mas também conforme a necessidade local em acessá-los. Assim, estabelecer laços complementares deve ser ainda mais crucial para países e regiões atrasados. De fato, [Balland e Boschma \(2021\)](#) encontram não apenas que laços extrarregionais com regiões cuja estrutura tecnológica é complementar à local são importantes como estas conexões são crescentemente importantes quão menos

---

<sup>30</sup> Econometricamente, sendo *Entrada* em tecnologias a variável dependente e como independente número de laços *NL* e complementariedade produtiva *CL*, o  $\beta$  de *NL* é positivo quando adicionado sozinho, mas se torna negativo depois de adicionar *CL*, enquanto o termo de interação entre ambos é positivo:  $CL * NL > 0$ .

<sup>31</sup> Ou, então, no que será testado adiante, a probabilidade de entrada numa determinada classe tecnológica, ou mesmo a passagem de não ter Vantagem Comparativa Revelada para passar a ter.



**Figura 4** – Distância Cognitiva Ótima

Fonte: Elaboração Própria.

desenvolvida é uma região. Também para a Europa, [Barzotto et al. \(2019\)](#) observam que colaboração regional aumenta a inovação em regiões atrasadas. Conduzindo uma análise para 260 regiões europeias ao nível NUTS-2 espalhadas por 29 países no período de 2000 a 2014, observa-se que tanto a quantidade relativa de patentes à população como o crescimento da atividade de patenteação respondem positivamente à quantidade de conhecimento acumulado localmente e à colaboração entre as regiões, medida por co-aplicação. Contudo, o impacto destes laços extrarregionais decai quão maior é o estoque de conhecimento regional acumulado, indicando que regiões atrasadas são as que mais se beneficiam do contato com outras regiões – são justamente elas que têm um estoque de conhecimento menor!<sup>32</sup> Estes resultados são robustos à diferentes especificações, incluindo uma que leve em conta a possível endogeneidade entre as variáveis, através de um GMM-System.

A evidência para a Europa é recheada e [Noni, Orsi e Belussi \(2018\)](#) adicionam a ela mas usando de estratégias empíricas distintas. Os autores cobrem o período de 2002-2008 e diferenciam o efeito entre regiões atrasadas e desenvolvidas, onde são usadas duas medidas do que é uma região atrasada: i) a partir dos quartis da atividade de patenteação acumulada em 1980-2010; e ii) a partir do índice *Regional*

<sup>32</sup> De fato, é exatamente assim que os autores definem o que é uma região atrasada, a partir dos quantis do estoque de conhecimento. Estoque de conhecimento é medido basicamente pela quantidade de patentes, mas incluindo outras ideias, como a de depreciação.

*Innovation Scorebord* (RIS) da comissão Europeia. Observa-se que colaborar com outras regiões é benéfico e ainda mais quando com uma região altamente inovadora, demonstrando esse efeito de que colaborações atuam como uma ponte de transmissão de conhecimento entre duas regiões. Quanto mais o seu parceiro tem a oferecer, maior o potencial dessa relação.

Santoalha (2019) trabalha também com dados de patente para a Europa no período que vai de 2000 a 2013 e diferencia efeitos de cooperação intra e inter-regionais. Similar à ideia de Scalera, Perri e Hannigan (2018), mas ao nível regional. Analogamente, a estratégia adotada foi a de contar a proporção de patentes que tinham mais de um aplicante local, para o caso da cooperação intra-regional, e não-local, no caso de cooperação inter-regional. O efeito das duas formas de cooperação não é robusto a estimações com efeitos-fixos e GMM-Differences, mas, quando significativos, os coeficientes dessas variáveis eram negativos nos modelos de entrada em novas classes tecnológicas. Em outras palavras, sua evidência aponta que cooperação entre regiões europeias entre 2000 e 2013 na verdade foi, senão inútil, prejudicial à diversificação regional. Contudo, a interação entre essas duas formas de cooperação auxilia na diversificação, indicando que há complementariedade na forma como se chega ao conhecimento externo. Isto é, os agentes locais podem se beneficiar do contato com agentes de outras regiões caso já tenham um alto grau de colaboração interna. Apesar de, relativamente a regiões pobres, regiões com alto PIB per capita se beneficiarem mais da cooperação intra-regional, quando há alto grau de cooperação intra-regional, regiões com baixo PIB per capita ainda ganham em termos de diversificação tecnológica quando cooperam com agentes externos. Em resumo, Santoalha (2019) argumenta que a cooperação intra e inter-regional por si só não bastam para auxiliar na diversificação tecnológica local, é necessário que ambas entrem em ação, para que o conhecimento externamente adquirido seja rapidamente espalhado dentro da fronteira. Não só os impactos de conexões extra e intra fronteiras deve ser distinto, como o custo de estabelecê-los. Como a firma está sujeita à limitada capacidade de decisões da equipe de gerenciamento (PENROSE, 1959), estabelecer conexões desnecessárias, ou em demasia, pode até ser prejudicial

à performance de uma firma ou região (SCALERA; PERRI; HANNIGAN, 2018; BALLAND; BOSCHMA, 2021).

Estes resultados antevêm que, apesar de importantes, estabelecer laços internacionais conectando-se a uma rede global de inovação e colaboração pode ser difícil para países periféricos. Ribeiro et al. (2018) usam de dados coletados do *Web of Science* para até dois milhões de artigos científicos para cada um dos 6 períodos igualmente distribuídos entre 2000 a 2015<sup>33</sup>. A partir dessa enorme base de dados, eles são capazes de ampliar alguns dos achados de Wagner e Leydesdorff (2005) de que há um sistema global de inovação emergente, por conta da rede de colaborações científicas se assemelhar a uma rede sem escala, já que a distribuição da conectividade dos seus nódulos segue uma lei de potência (BARABASI, 2016). Isto indica que o sistema global de inovação é auto-organizado e apresenta alguma hierarquia. Ou seja, existem alguns países que estão muito bem colocados, estabelecem muitas colaborações com outros importantes países, atuando como *hubs* – pontos centrais – dessa rede. Não por acaso, os *hubs* dessa rede são os países desenvolvidos – Estados Unidos, China, Inglaterra, Alemanha e Japão contam com 47,8% das produções na amostra coletada. Contudo, Ribeiro et al. (2018) encontram ainda que a maior parte dos países que tem mais de 60% das suas produções com co-autorias são subdesenvolvidos, servindo a eles como uma forma começar seus sistemas nacionais de inovação. Portanto, por um lado, a importância de acessar capacidades ausentes localmente é tamanha para países periféricos que elas servem como um pontapé para a inovação nesses países. Por outro, países desenvolvidos conseguem explorar muito bem essas redes de colaboração para manter sua posição predominante no sistema global de inovação. Assim, fica ambíguo ainda se laços internacionais de colaboração ajudam ou pioram as desigualdades entre os países.

Este trabalho busca adicionar à evidência empírica existente ao usar de dados internacionais de patentes, agregados ao nível nacional, para avaliar se colaborações internacionais com países com estrutura tecnológica complementar à local podem auxiliar na entrada de novos setores tecnológicos. Especificamente, pretende-se reali-

---

<sup>33</sup> Ou seja, 2000, 2003, 2006 e assim por diante.

zar uma avaliação econométrica desta hipótese para dois dos principais escritórios de economias desenvolvidas – Estados Unidos e Europa – usando dados de patente disponibilizados pelo PATSTAT. Esta discussão é particularmente relevante para aqueles estudos que busquem estabelecer uma estratégia ótima de diversificação produtiva (ALSHAMSI; PINHEIRO; HIDALGO, 2018) ou delimitar políticas de desenvolvimento local (TORRES, 2019), especialmente para regiões periféricas (HARTMANN; BEZERRA; PINHEIRO, 2019; ROMERO et al., 2021; HAUSMANN; CHAUVIN, 2015; HARTMANN; BEZERRA; PINHEIRO, 2019; HAUSMANN; SANTOS; OBACH, 2017) na medida em que, até então, raramente levam em conta o uso de capacidades ausentes localmente mas acessíveis por meio de laços extrarregionais.

## 3 Metodologia

O objetivo deste capítulo é tratar sobre os métodos que serão utilizados no capítulo subsequente. Como argumentado no capítulo anterior, a literatura sobre diversificação tecnológica e produtiva em boa medida ainda dá um peso muito grande às capacidades disponíveis localmente, ignorando o impacto de conexões internacionais e extrarregionais na trajetória local. Assim, esta metodologia mira responder se laços internacionais auxiliam na performance tecnológica dos países, almejando, ao mesmo tempo, diferenciar esse efeito entre países desenvolvidos e subdesenvolvidos. Em especial, o recorte dessas colaborações internacionais é avaliado em termos da complementariedade com a estrutura tecnológica local, já que esse tipo de conexão deve ser mais proveitosa do que aquelas feitas entre países cuja estrutura tecnológica seja muito próxima ou distante.

A estratégia metodológica adotada começa por desenvolver alguns dos principais indicadores utilizados ao longo do resto da dissertação na [seção 3.1](#). Depois disso, a [seção 3.4](#) declara a base de dados utilizada, do PATSTAT, explorando algumas informações fundamentais da amostra coletada. Os métodos aqui propostos serão explorados com maior profundidade no [Capítulo 4](#).

### 3.1 Indicadores

Queremos avaliar se laços internacionais com países cuja estrutura tecnológica seja complementar à local auxiliam na diversificação tecnológica. Em especial, tal como argumentado no [Capítulo 2](#), supomos que esse efeito seja maior para países atrasados. Nesta seção serão expostos os indicadores que utilizaremos para avaliar estas hipóteses.

Seguindo [Balland e Boschma \(2021\)](#), dizemos que um país *entra* em uma

determinada tecnologia quando ele não era e passa a ser especializado nela. Para tanto, primeiro medimos a especialização através da Vantagem Tecnológica Relativa (VTR)<sup>1</sup>, à semelhança de como medimos VCR no [Capítulo 2](#), com a diferença de aqui ao invés de exportações consideraremos patentes.

$$VTR_{cit} = \frac{\text{patentes}_{cit} / \sum_i \text{patentes}_{cit}}{\sum_r \text{patentes}_{cit} / \sum_r \sum_i \text{patentes}_{cit}}, \quad (3.1)$$

onde *patentes* são definidas por titularidade da aplicação de patentes de instituições cuja origem seja no país *c* e classificada na tecnologia *i*, no período *t*. Ver [seção 3.4](#) para mais detalhes.

A fim de não confundir com a matriz de vantagens comparativas que usamos no [Capítulo 2](#), a matriz de especialização tecnológica será denominada de  $V_{cit}$  para um país *c* na tecnologia *i* avaliada no período *t*. Como antes, seguimos [Hidalgo et al. \(2007\)](#) e definimos que um país é especializado quando ultrapassa o limite de 1 para *VTR*. Assim:

$$V_{cit} = \begin{cases} 1, & \text{se } VTR_{cit} > 1 \\ 0, & \text{se } VTR_{cit} \leq 1. \end{cases} \quad (3.2)$$

Isto é, [\(3.2\)](#) mede se um país dedica uma parte da sua estrutura inovativa à uma tecnologia proporcionalmente maior do que a média dos outros países.

Assim, a entrada,  $E_{ci,t+1} = 1$ , de um país numa tecnologia específica é apenas que esse país *c* não tinha vantagem comparativa revelada para um período inicial *t*,  $V_{ci,t} = 0$ , e, em algum período futuro, no caso, definindo pelo subsequente, *t* + 1, passou a ser especializado naquela tecnologia, isto é  $V_{ci,t+1} = 1$ . A manutenção, por sua vez, é quando não há variação, seja continuar especializado,  $V_{ci,t} = V_{ci,t+1} = 1$ , ou continuar não especializado,  $V_{ci,t} = V_{ci,t+1} = 0$ . De modo genérico, definiremos a

<sup>1</sup> Em outros contextos, ao invés de relativa, diz-se revelada. A interpretação é levemente distinta, mas o cálculo igual e ambas as versões são verdadeiras – é um índice relativo e é revelado no sentido de que busca medir algo *velado* mas que deve transparecer nesses dados.

manutenção como a não entrada,  $E_{ci,t+1} = 0$ . Formalmente:

$$E_{ci,t+1} = V_{ci,t+1} - V_{ci,t} = \begin{cases} +1, & \text{entrada;} \\ 0, & \text{manutenção; ou} \end{cases} \quad (3.3)$$

Como já argumentado no [Capítulo 2](#), a evidência empírica aponta que regiões e países tendem a diversificar sua produção em direção a setores relacionados à sua estrutura produtivo-tecnológica. Assim, iremos controlar a entrada em uma tecnologia pela proximidade tecnológica local que a região apresenta. Nesse sentido, primeiro definimos a proximidade entre duas tecnologias  $i$  e  $j$  através da probabilidade,  $\phi_{ij,t}$ , de que ambas estejam presentes simultaneamente em um mesmo país  $c$  no período  $t$ . Segundo, definimos a *densidade* em uma região  $c$  em torno de  $i$  como a proporção de tecnologias que  $c$  domina e são similares a  $i$ . Formalmente:

$$\omega_{cit} = \frac{\sum_j V_{cit} * \phi_{ij,t}}{\sum_j \phi_{ij,t}}. \quad (3.4)$$

Basicamente, [\(3.4\)](#) soma a proximidade de  $i$ , no denominador, com todas as demais tecnologias<sup>2</sup> e, no numerador, com todas as demais tecnologias que  $c$  é especializado. Assim, a densidade indica quanto das capacidades necessárias para dominar  $i$  (que assumimos estar presente naquelas regiões que a produzem e, portanto, produzem outras tecnologias que também usam daquelas capacidades)  $c$  atualmente domina.

Para medir o nível de desenvolvimento de uma região, usamos o Índice de Complexidade Tecnológica (ICT) ([HIDALGO; HAUSMANN, 2009](#); [BALLAND; RIGBY, 2017](#)). O procedimento é idêntico ao descrito para o caso de exportações como desenvolvido por [Hidalgo e Hausmann \(2009\)](#). [Balland e Rigby \(2017\)](#) adaptaram a estratégia usada para dados de exportação construindo o ICT com dados de patente para os Estados Unidos, uma métrica comum na literatura desde então ([BALLAND et al., 2020](#); [PINHEIRO et al., 2022](#), por exemplo).

<sup>2</sup> Pela forma que definimos, essa *somanão* resulta em 1.

Finalmente, para definir laços internacionais complementares, seguimos a métrica desenvolvida por [Balland e Boschma \(2021\)](#). Basicamente, a complementariedade tecnológica em uma tecnologia  $i$  do país  $d$  para o país  $c$  depende de<sup>3</sup>: (I)  $c$  não dominar tecnologias similares a  $i$  e que, portanto, estão associadas a capacidades relevantes para se especializar em  $i$ ; e (II) que  $d$  seja especializado nestas tecnologias. Assim, um laço entre  $c$  e  $d$  deve dar acesso à  $c$  capacidades ausentes localmente mas relevantes para o domínio de  $i$ . Laços internacionais são definidos como em [Balland e Boschma \(2021\)](#) através da co-patenteação. Quer dizer, caso a patente tenha sido aplicada por alguém do país  $c$  e outro alguém do país  $d$  há um laço entre os dois países. Em particular, não restringimos a existência de laços pela titularidade da patente, como fizemos para definir a origem daquela patente. Ver [seção 3.4](#) para mais detalhes.

Especificamente, para construir esse índice de complementariedade, precisamos primeiro definir como um país se liga a outro através da co-aplicação em patentes, medida comum na literatura ([SANTOALHA, 2019](#); [BALLAND; BOSCHMA, 2021](#), por exemplo). Sendo  $L_{cdt}$  os laços extrarregionais entre um país  $c$  e  $d$  em  $t$ :

$$L_{cdt} = \sum_i \sum_p \text{Aplicante}_{cit,p} * \text{Aplicante}_{dit,p}, \quad (3.5)$$

onde  $c \neq d$  e  $\text{Aplicante}_{cit,p} = 1$  se o aplicante da patente  $p$ , que é classificada na tecnologia  $i$ , vem de um país  $c$ , caso contrário,  $\text{Aplicante}_{cit,p} = 0$ . Por exemplo, supondo apenas duas tecnologias,  $i = i_1, i_2$ , se houver 500 patentes classificadas na tecnologia  $i_1$  e 500 patentes em  $i_2$  (ambas em  $t$ ) cujo aplicante tenha origem em  $c$  e houve co-patenteação com  $d$  em 57 patentes em  $i_1$  e 32 em  $i_2$ , então  $L_{cdt} = 57 + 32 = 89$ .

Assim, sendo  $C_{cit}$  um índice de laços internacionais com países tecnologicamente complementares para o país  $c$  na tecnologia  $i$  em  $t$ :

<sup>3</sup> Observe que estamos falando sobre potenciais tecnologias de entrada, logo, estamos apenas considerando produtos que  $c$  não domina em  $t$ , isto é,  $V_{ci,t} < 1$ .



$$C_{cit} = (1 - V_{cit}) \sum_s V_{dit} * \omega_{dit} * L_{cdt}. \quad (3.6)$$

A [Equação 3.6](#) considera apenas tecnologias  $i$  potenciais que  $c$  não é especializado em  $t$ ,  $(1 - V_{cit})$  e  $d$  é,  $V_{dit}$ . Daí, avalia quantas capacidades  $d$  possui em torno da tecnologia  $i$ , ou seja, a densidade,  $\omega_{dit}$ , ponderando o quão próximos são os dois países ao multiplicar pela quantidade de laços entre eles,  $L_{cdt}$ . De maneira sucinta, podemos definir o indicador para apenas aquelas tecnologias  $j$  que  $c$  não domina e apenas para os países parceiros,  $d'$ , que a dominam tal como:

$$C_{cjt} = \sum_{d'} L_{cd't} * \omega_{d'jt}, \quad (3.7)$$

e  $C_{cjt} = 0$  para as demais tecnologias  $j'$  que  $c$  é especializado.

### 3.2 Modelo de regressão estatística

A partir dos indicadores construídos, usaremos de regressões estatísticas para avaliar a hipótese de que laços complementares internacionais são importantes para a performance tecnológica local. Como a variável dependente é binária, entrada em novas classes tecnológicas ([3.3](#)), seguiremos estratégias clássicas de estimação de modelos índice Probit e Logit ([WOOLDRIDGE, 2010](#)). Em especial, queremos usar esses modelos de resposta binária no contexto de dados em painel, considerando também a existência de efeitos-fixos não observáveis. Seguindo [Wooldridge \(2010\)](#), descrevemos o modelo a ser estimado como:

$$P(y_{it} = 1 | \mathbf{x}_{it}) = F(\mathbf{x}_{it}\boldsymbol{\beta}), \quad (3.8)$$

onde  $y_{it}$  é a variável dependente definida para a unidade de *cross-section* (o “indivíduo” que no caso é o país)  $i$ ,  $\mathbf{x}_{it}$  é um vetor de variáveis independentes, cujos coeficientes são definidos por  $\boldsymbol{\beta}$  – incluímos a constante nessa notação. Como é usual,

$P(Z = z|w)$  representa a probabilidade de algum evento  $z$  condicional à ocorrência de  $w$ .  $F(\cdot)$  representa alguma função, que normalmente se assume estar restrito ao intervalo aberto entre 0 a 1,  $0 < F(z) < 1$  para todo  $z \in \mathbb{R}$ . Para modelos Probit,  $F(\cdot)$  é a função de densidade acumulada da normal e, para modelos Logit, a de uma logística – veja que ambas estão definidas no supracitado intervalo.

Podemos chegar a (3.8) também através de um modelo de variável latente:

$$y_{it}^* = \mathbf{x}_{it}\boldsymbol{\beta} + \epsilon_{it}, \quad (3.9)$$

em que, para  $y_{it}^* > 0$ , consideramos que  $y_{it} = 1$ . Assim,

$$P(y_{it} = 1|\mathbf{x}_{it}) = P(y_{it}^* > 0|\mathbf{x}_{it}) = P(\epsilon_{it} > -\mathbf{x}_{it}\boldsymbol{\beta}) = 1 - F(-\mathbf{x}_{it}\boldsymbol{\beta}) = F(\mathbf{x}_{it}\boldsymbol{\beta}),$$

onde na última igualdade usamos do fato que  $F(\cdot)$  é simétrica em torno de zero. Basta então assumir uma determinada distribuição para  $\epsilon_{it}$ , por exemplo, a normal, e temos a função de densidade acumulada. Como (3.8) é não linear, estimamos por máxima verossimilhança, para a qual temos de assumir uma determinada distribuição (comportamento). No caso do modelo de probabilidade linear, teríamos que  $F(z) = z$ , ou seja, a identidade – neste caso, porém, poderíamos ter valores estimados fora do intervalo entre zero e um.

Modelos de índice implicam numa lógica distinta para avaliar efeitos parciais. Ignorando os subscritos e tomando a derivada com relação a  $\mathbf{x}$  de (3.8), temos:

$$\frac{\partial F(y|\mathbf{x})}{\partial \mathbf{x}} = \frac{dF(\mathbf{x}\boldsymbol{\beta})}{d\mathbf{x}\boldsymbol{\beta}} \frac{\partial \mathbf{x}\boldsymbol{\beta}}{\partial \mathbf{x}} = f(\mathbf{x}\boldsymbol{\beta})\boldsymbol{\beta}, \quad (3.10)$$

onde da primeira para a segunda igualdade usamos da regra da cadeia e da segunda para a terceira de que a derivada da função de densidade acumulada é apenas a função de densidade de probabilidade. A (3.10) passa uma informação importante: os efeitos parciais dependem da parte da distribuição em que nos encontramos. Isso

advém da forma não linear desses modelos de índice. Sendo linear, é indiferente em que ponto da curva se está, o efeito marginal é sempre o mesmo. O ganho que temos em controlar o conjunto de resultados possíveis para a nossa variável dependente – no caso, entre 0 e 1 –, perdemos ao aumentar a complexidade na interpretação dos resultados.

Como estamos interessados numa métrica que resuma o impacto de uma determinada variável independente na probabilidade de  $y$ , queremos uma forma de traduzir os efeitos marginais de todas as observações em uma “observação representativa”. Uma maneira conveniente de fazer isso é usar os valores médios de  $\mathbf{x}$ . Contudo, como aponta Wooldridge (2010), o valor médio de algumas variáveis frequentemente não tem sentido econômico<sup>4</sup>, assim, uma abordagem mais comum é usar dos efeitos marginais médios (APE)<sup>5</sup>. Isto é, calculamos (3.10) para cada observação  $n$  da amostra e depois tiramos sua média, tal qual:

$$APE = \frac{1}{N} \sum_n f(\mathbf{x}_n \hat{\boldsymbol{\beta}}) \hat{\boldsymbol{\beta}}, \quad (3.11)$$

onde  $\hat{\boldsymbol{\beta}}$  são os coeficientes *estimados*.

Podemos expandir (3.8) incluindo efeitos-fixos ou não observáveis, do agrupamento em  $i$ <sup>6</sup>,  $c_i$ :

$$P(y_{it} = 1 | \mathbf{x}_{it}, c_i) = F(\mathbf{x}_{it} \boldsymbol{\beta} + c_i), \quad (3.12)$$

ou, em termos do modelo de variável latente:

$$y_{it}^* = \mathbf{x}_{it} \boldsymbol{\beta} + \epsilon_{it} + c_i. \quad (3.13)$$

<sup>4</sup> Isto é mais comum em aplicações com *dummies*, por exemplo, uma de gênero: o que significa ser 0, 6 Homem?

<sup>5</sup> Do inglês, *Average Partial Effect*.

<sup>6</sup> Essa metodologia é expansível para efeitos-fixos em outras agregações, como de tempo e, neste caso, classe tecnológica, por simplicidade, a exposição será feita apenas para um grupo.

Para um modelo de efeitos-fixos, estamos assumindo que a distribuição de  $\epsilon_{it}$  condicional à  $c_i$  é irrestrita. Enquanto restringir  $f(\epsilon_{it}|\mathbf{x}_{it}, c_i)$  à estrita exogeneidade gera um modelo de efeitos-aleatórios (GREENE, 2003), onde  $f(\cdot)$  é a função densidade de probabilidade. Quanto aos efeitos marginais, segue-se de maneira análoga ao caso sem efeitos-efeixos.

Em modelos simples de painel, conseguimos eliminar os efeitos-fixos de grupos de indivíduos que são não observáveis. Esta é inclusive uma grande motivação para usar esse tipo de método, em comparação a uma *cross-section* – ou seja, sem acompanhar indivíduos (ou países) ao longo do tempo. Contudo, por conta da forma não linear de modelos de índice nem sempre isso é possível. Greene (2003) aponta que, em especial, não possuímos uma *estatística suficiente mínima* para modelos probit – em modelos lineares para dados em painel, a média de uma variável,  $\bar{y}_i$ , serve como uma estatística suficiente mínima para  $c_i$ . Esta ausência acarreta no problema de parâmetros incidentais, que é basicamente a ideia de que as estimativas de Probit com efeitos-fixos não serão consistentes, porque, no fundo, teremos que estimar todos os parâmetros associados a  $c_i$ , o que é particularmente danoso por conta de em muitas aplicações  $N \gg T$  – em que  $N$  é a quantidade de indivíduos e  $T$  a quantidade de períodos. Contudo, é possível realizar esse procedimento para uma distribuição logística. Portanto, a análise que segue ainda levará em conta os resultados dos modelos Probit por conta da comparabilidade e ampla utilização na literatura, mas daremos um enfoque maior para os resultados do Logit.

Tendo em vista a construção dos modelos de índice, resta definir qual é a forma funcional, ou o modelo de variável latente, que desejamos estimar. Para responder se laços internacionais podem ajudar um país a renovar sua trajetória local, exploramos a seguinte especificação:

$$E_{ci,t+1} = \beta_0 + \beta_1 C_{cit} + \beta_2 C_{cit} * ICT_{ct} + \beta_3 \omega_{cit} + \boldsymbol{\lambda} \mathbf{X}_{ct} + \gamma_t + \alpha_c + \kappa_i + \epsilon_{cit}, \quad (3.14)$$

onde  $\beta_0$  é uma constante,  $\epsilon$  é um termo de erro,  $\beta_k$  – para  $k = 1, 2, 3$  – e  $\boldsymbol{\lambda}$

são coeficientes a serem estimados, sendo que  $\lambda$  é um vetor de coeficientes para as variáveis de controle  $\mathbf{X}$  e  $\gamma_t$ ,  $\alpha_c$  e  $\kappa_i$  são efeitos fixos para tempo, classes tecnológicas e países, respectivamente. Avaliamos as variáveis para cada país  $c$  nas tecnologias  $i$  em cada período de tempo  $t$ . As demais variáveis seguem a convenção utilizada até aqui, sendo  $E$  a entrada numa tecnologia,  $C$  o índice de complementariedade,  $ICT$  o índice de complexidade tecnológica dos países e  $\omega$  a densidade em torno de uma tecnologia, todas definidas de acordo com seus sub-índices (ver seção 3.1 para mais detalhes). Vale apontar que usar efeitos-fixos para tempo ou *dummies* para cada período leva a resultados muito parecidos.

A hipótese deste trabalho é que  $\beta_1 > 0$ , ou seja, ter contato com outros países que possam oferecer capacidades complementares em uma tecnologia facilita a entrada nela. Além disso, esperamos também que este efeito seja maior para países periféricos, isto é,  $\beta_2 < 0$ , quão menos complexo um país é, mais ele se beneficia de links internacionais que possam complementar a sua estrutura tecnológica. Isto porque esperamos que países “atrasados” possuam maior dificuldade em desenvolver novas capacidades relevantes para dominar uma determinada tecnologia (HIDALGO et al., 2007; BALLAND et al., 2019). Assim, esses laços seriam mais cruciais para a sua diversificação tecnológica do que para países desenvolvidos. Além disso, esse efeito deve ser mais frequente para esses países, já que elas possuem menos capacitações, aumentando as chances de que um outro país possua um desenvolvimento tecnológico complementar ao seu. Adicionalmente, seguindo o “Princípio do Relacionamento”, esperamos que  $\beta_3 > 0$ , já que possuir capacidades similares facilita no domínio de novas tecnologias (HIDALGO et al., 2018).

Para as variáveis de controle, seguimos Balland e Boschma (2021). Incluímos primeiro uma que trata de economias de convergência, no caso, o PIB per capita, tal como os autores – sendo que um coeficiente negativo indicaria que países mais ricos tem maior dificuldade em entrar em novas tecnologias, daí a ideia de convergência: ser mais pobre por si só ajuda na performance tecnológica. Em segundo, adicionamos uma variável para controlar por efeitos de economias de aglomeração, que é a ideia de que quando se aglomera atividades econômicas num mesmo local, há externalidades

positivas. Assim, haveria uma evidência favorável para economias de aglomeração caso encontremos um coeficiente positivo para a *proxy* de densidade populacional – um país densamente povoado deve *aglomerar* também muitas atividades. Aqui, inclusive, divergimos de [Balland e Boschma \(2021\)](#) dado que eles usam apenas da população, o que pareceu menos adequado dado que um país muito populoso pode ser populoso mas com a população amplamente difundida pelo território, o que eliminaria os efeitos de aglomeração. Apesar disto, foram realizados testes incluindo população ao invés de densidade populacional e os resultados ficaram basicamente inalterados.

De fato, essas hipóteses são confirmadas em [Balland e Boschma \(2021\)](#) para regiões europeias ao nível NUTS-2. **O objetivo deste trabalho é oferecer evidência empírica ao nível internacional.** Assim, busca-se responder se laços internacionais são importantes para a diversificação tecnológica local usando dados de patente disponibilizados pelo PATSTAT. Seguindo a lógica expressa em (3.5), laços internacionais serão definidos através de co-patenteação entre aplicantes de diferentes países.

### 3.3 Redes

No [Capítulo 4](#) exploraremos a rede de colaborações internacionais que nossos dados indicam. Isto é, usaremos do indicador descrito em (3.5) para caracterizar como é o padrão de colaboração entre os países. Contudo, o que é uma rede?

Uma rede é uma forma de descrever como componentes interagem entre si ([BARABASI, 2016](#)). Por exemplo, uma rede social descreve como que pessoas se relacionam. Fundamentalmente, sua construção depende de dois elementos básicos: os nódulos (ou componentes) e as conexões (a interação). Juntos, esses elementos constituem uma rede. No caso que temos em mente, estamos preocupados em descrever como é que os países se relacionam quando co-aplicam numa mesma patente, isto porque estamos interessados com a transmissão de capacidades tecnológicas de

um país a outro. Hipoteticamente, poderíamos pensar na rede, ou grafo,  $G$  composto pelo conjunto de nódulos, ou vértices,  $V$ , que são conectados pelos laços, ou arestas,  $L$ .  $G$  é a rede colaborações internacionais, que é composta pelos  $V$  países e que colaboraram em patentes tal como  $L$  descreve. Suponha que existam quatro países: Brasil (BR), Estados Unidos (US), China (CH) e Japão (JP). Então, escrevemos  $V = \{BR; US; CH; JP\}$ . Agora suponha que US e BR estejam relacionados entre si, assim como US e CH e US e JP – ou seja, Estados Unidos se relaciona com todos, mas nenhum dos demais se relaciona entre si. Então teríamos que  $L = \{(US; BR); (US; CH); (US; JP)\}$ . Em especial, (3.5) adiciona uma quantidade que pode ser pensada como um atributo de cada um dos elementos de  $L$ , indo além da descrição binária que assumimos – existe ou não relação (ou co-aplicaram ou não). Contudo, por simplicidade, ignoramos esses atributos.

Podemos descrever esse exemplo hipotético com a [Figura 5](#). Esta representação gráfica deixa explícita o quão bem conectado é os Estados Unidos nesta rede. Uma forma de descrever essa conectividade é através do *grau* dos nódulos. O grau de um nódulo é simplesmente a sua quantidade de conexões. Assim, o nódulo US tem grau 3, enquanto os demais tem grau 1. Podemos descrever uma rede através da sua conectividade, mapeando a distribuição dos graus de seus nódulos, isto é, computamos a frequência, ou probabilidade  $p_k$ , com que nódulos de determinado grau,  $k$ , aparecem em uma rede.

[Barabasi \(2016\)](#) argumenta que redes que encontramos na realidade não possuem uma distribuição aleatória, sua distribuição se assemelha mais a uma lei de potência. Em redes aleatórias encontramos uma quantidade ínfima de nódulos com alto grau – muito bem conectados – e de baixo grau – muito mal conectados. Como seria de se esperar, encontramos mais nódulos cujo grau seja próximo à sua média,  $\langle k \rangle$ . Com isso, a distribuição de uma rede aleatória pode ser simplesmente descrita através de uma distribuição binomial, mas como esperamos que  $N \gg \langle k \rangle$ , onde  $N$  é a quantidade de nódulos na rede, podemos aproximá-la a uma distribuição de Poisson. Contudo, redes encontradas na realidade apresentam um comportamento distinto, em que a probabilidade de encontrar nós com alto e baixo grau é maior

do que esperaríamos de uma rede aleatória – ao que damos o nome de uma “rede sem escala”. Formalmente, podemos representar a probabilidade de que um nóculo possua grau  $k$  da seguinte maneira:

$$p_k \sim k^{-\gamma}, \quad (3.15)$$

em que  $\gamma$  é uma constante positiva. (3.15) nos conta uma ideia simples: existem muitos nóculos com poucas conexões e poucos com muitas conexões, *hubs*. Barabasi (2016) propõe duas explicações do porquê encontramos redes sem escala na realidade, ao invés de redes aleatórias. Primeiro, redes que encontramos na vida real estão em constante crescimento. Sempre há novas instituições surgindo na rede de colaborações internacionais<sup>7</sup>, o que faz com que seu tamanho, em termos de nóculos, esteja em constante crescimento. O que nos leva à segunda motivação: ligação preferencial<sup>8</sup>. Quando um novo nóculo passa a participar de uma rede, ele terá preferência a se conectar com nóculos mais bem conectados. Em conjunto, essas condições explicam porque redes que encontramos na realidade são sem escala e não aleatórias.

Podemos fazer uma transformação logarítmica em (3.15) que facilitará procedimentos empíricos que realizaremos na seção 4.1 para testar se a rede de colaborações internacionais que encontramos através da base de dados descrita na seção 3.4 pode ser caracterizada como uma rede sem escala. Isto é:

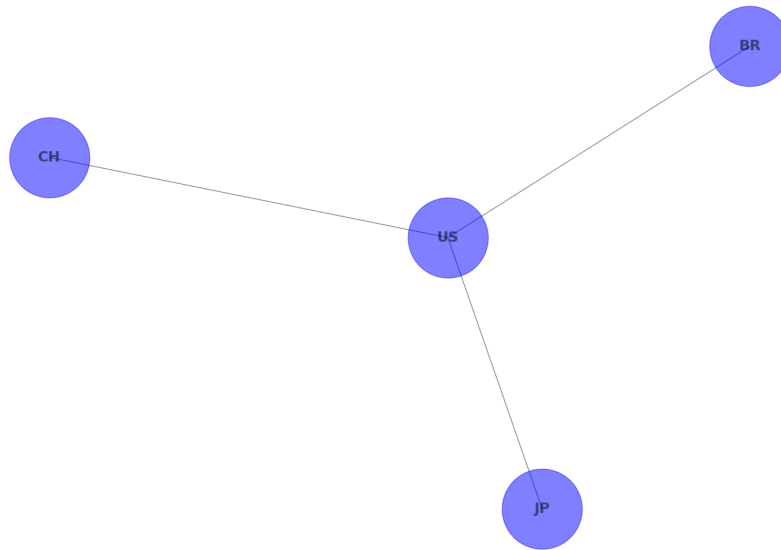
$$\log p_k \sim -\gamma \log k, \quad (3.16)$$

de onde podemos usar de uma regressão estatística para encontrar o coeficiente associado a esta rede,  $\gamma$ . Em especial, estaremos preocupados no quanto que a relação descrita em (3.16) é significativa, para o que observaremos o  $R^2$  da regressão:

<sup>7</sup> Neste trabalho estamos pensando em termos agregados dos “estados nacionais”, contudo, os dados estão sempre refletindo instituições alocadas nesses países, daí ser possível fazer essa comparação.

<sup>8</sup> Tradução livre do inglês, *preferential attachment*.



**Figura 5** – Rede de Colaborações Internacionais Hipotética

Fonte: Elaboração Própria.

$$\log p_k = \alpha_0 + \alpha_1 \log k + u, \quad (3.17)$$

em que  $\alpha_0$  e  $\alpha_1 (= -\gamma)$  são coeficientes a serem estimados e  $u$  é um termo de erro.

### 3.4 Dados e Procedimentos Empíricos

Os dados de patentes a serem utilizados vêm da base do PATSAT. A amostra retirada cobre o período de 2000 a 2019 numa frequência anual. Para as análises dos dados, dividimos a amostra em quatro sub-períodos de igual tamanho e não sobrepostos. Isto significa que o primeiro período vai de 2000 a 2004, o segundo de 2005 a 2009, o terceiro de 2010 a 2014 e o quarto de 2015 a 2019, agregando para cada período os dados. O objetivo desta agregação é reduzir o ruído proveniente de mudanças aleatórias ou sazonalidades nos dados anuais, como é comum na literatura (BALLAND; BOSCHMA, 2021; RIGBY et al., 2022, por exemplo). Com isso podemos observar características estruturais dos países, que é o nível de agregação adotado.

Dois escritórios de aplicação distintos são utilizados, o do Estados Unidos (USPTO)<sup>9</sup> e o da Europa (EPO)<sup>10</sup>, com que será possível observar se há diferenças ou semelhanças regionais – tendo em mente que o foco é internacional, mas países que aplicam em um podem não aplicar no outro, ou simplesmente aplicar com menor frequência. A base de dados necessária para a construção dos indicadores explicados no [Capítulo 3](#) contém cinco informações essenciais: (i) a identificação de uma patente; (ii) o ano de aplicação; (iii) o país de origem do aplicante; (iv) a classe tecnológica a que pertence; e (v) a posição de aplicação do aplicante. A agregação da classe tecnológica é feita em 4 dígitos de acordo com a Classificação Internacional de Patentes (IPC-4)<sup>11</sup> como em outros trabalhos ([RIGBY, 2015](#); [BOSCHMA](#); [BALLAND](#); [KOGLER, 2015](#)). A determinação do local de origem da patente, isto é, a que país ela vai ser atribuída, vai de acordo com a titularidade que, por sua vez, é definida para o primeiro aplicante (de posição igual a 1), uma definição similar ao que [Ribeiro et al. \(2018\)](#) faz ao definir o autor de um artigo científico pelo local de endereço do seu primeiro autor.

Seguindo outros trabalhos ([STOJKOSKI; KOCH; HIDALGO, 2022](#), por exemplo), dois cortes são realizados<sup>12</sup>. São desconsiderados da análise países com (a) uma quantidade total de patentes no inferior a 4 e (b) classes tecnológicas com menos de 4 aplicações, ambas as condições avaliadas para todo o período. Vale ressaltar que os cálculos eliminam uma parcela dos dados, principalmente para calcular a entrada em novos setores tecnológicos, porque, neste caso, a inexistência de patentes num momento futuro coloca em cheque a validade da análise. Oportunamente, esses detalhes serão adereçados. Finalmente, depois de corrigidas inconsistências – por exemplo, de países que não existem<sup>13</sup> –, realizados os cortes e feito os cálculos, a

<sup>9</sup> *United States Patent and Trademark Office.*

<sup>10</sup> *European Patent Office.*

<sup>11</sup> No endereço <<https://ipcpub.wipo.int/?notion=scheme&version=20220101&symbol=A22C&menulang=en&lang=en&viewmode=f&fipcpc=no&showdeleted=yes&indexes=no&headings=yes&notes=yes&direction=o2n&initial=A&cwid=none&tree=no&searchmode=smart>> é possível ver a descrição das classificações adotadas.

<sup>12</sup> É comum ignorar “países” com populações pequenas, contudo, como alguns desses “países”, ou como chamaremos eventualmente, “entidades nacionais”, são bastante representativos, principalmente a depender do escritório de aplicação – como as Ilhas Cayman e as Ilhas Marshall no USPTO –, abdica-se desse corte para permitir a análise deles também.

<sup>13</sup> Isto é, o código, ISO-2, utilizado para aquela observação é um código inválido: não foi definido para país nenhum.

base final utilizada para a regressão estatística na [seção 4.2](#) conta com um total de 7.559.304 patentes para o USPTO distribuídas ao longo de 127 países e 632 classes e 3.416.654 patentes para o EPO distribuídas ao longo de 101 países e 631 classes.

Laços entre países são definidos através da co-patenteação ([SANTOALHA, 2019](#); [BALLAND; BOSCHMA, 2021](#), por exemplo). Isso significa que dois países possuem um link caso tenham colaborado na invenção de uma patente. Assim, se na posição um estiver um aplicante dos Estados Unidos, a patente é considerada norte-americana, pela titularidade, mas tendo como segundo aplicante um brasileiro, esta patente estabelece um laço entre Brasil e Estados Unidos (e vice-versa). No USPTO, aproximadamente 60% das patentes possuem no máximo um aplicante. A proporção vai decaindo até que a partir da oitava posição ela já é menos de 1%. No EPO, essa distribuição é ainda mais concentrada, com mais de 93% das patentes com apenas um aplicante, e cerca de 5% com dois aplicantes, sobrando basicamente 1% das patentes com mais de 2 aplicantes. A [Figura 6](#) demonstra isso, tendo no eixo vertical a proporção de patentes que tem até um número  $x$  de aplicantes que é mostrado no eixo horizontal. Usamos de escala logarítmica para esse último eixo, por fins estéticos<sup>14</sup>, assim, o segundo ponto representado é a segunda posição. Limita-se a exposição até a décima posição, onde, como podemos ver, em ambos os escritórios fica próximo de zero, mas o decaimento é mais veloz no EPO.

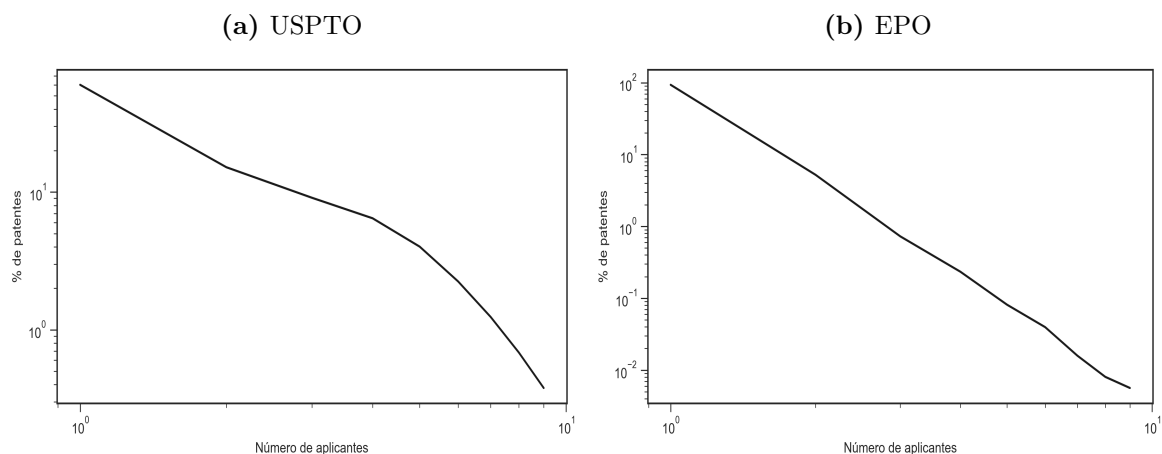
O máximo de inventores que uma única patente aglomera no USPTO é 77. Essa é uma patente com titularidade dos Estados Unidos classificada na área tecnológica G06F (Computação, cálculo e contagem), com data de aplicação de 2002. Apesar de tantos aplicantes, ela só estabelece um link! No caso, entre Estados Unidos e Israel, já que 76 aplicantes são norte-americanos e um é israelense. Depois dela, três outras patentes aglomeram 66 aplicantes e uma 65, juntas representando 0,00006% das patentes totais, ou seja, patentes com muitos aplicantes é um evento bastante raro e não exatamente significativo para os propósitos deste trabalho. Apesar de o EPO ser relativamente menos “colaborador”, a patente com maior número de aplicantes não

<sup>14</sup> E também para incitar a análise de distribuição da quantidade de laços que os países estabelecem na rede de colaborações.

fica tão distante: 62 aplicantes e a segunda 61, sendo ambas representativas, cada, de apenas 0,000037% das patentes. A patente com mais aplicantes tem, contudo, apenas aplicantes de um país, a Alemanha, sendo classificada em duas classes tecnológicas, G07F e G06Q, ou seja, ambos de G, a classe que representa patentes classificadas no campo “Física” – a mesma daquela dos Estados Unidos! Embora uma patente com muitos aplicantes não signifique automaticamente muitas colaborações, é possível que uma só patente possa ser o canal de comunicação entre diversos países. Por exemplo, a patente que mais estabelece links no USPTO possui apenas 12 aplicantes, mas como envolve 10 nacionalidades distintas, essa patente dá origem a 45 interações entre países<sup>15</sup>, juntando as áreas de Física (G) e Química (C), é classificada em três classes tecnológicas – G01N, C07H e C12Q. Contudo, esse tipo de evento também é extremamente raro, sendo que a maioria, mais de 91% dos casos, das patentes que estabelecem algum laço com outro país o fazem com apenas um outro, e os casos em que estabelecem com dois outros países cobrem mais 8% das ocorrências. Portanto, cerca de apenas 1% das patentes que atuam como canais de transmissão de conhecimento o fazem simultaneamente com mais de 3 países no USPTO. Esse é um fato que ressoa no EPO também, tendo patentes com um e dois países distintos cobrindo cerca de 99,31% das patentes! Contudo, curiosamente, existe uma patente que permite laços entre 14 nacionalidades distintas, gerando até 91 colaborações! Talvez esperadamente, essa patente envolve colaborações apenas entre países europeus. Curiosamente, também essa patente envolve a junção das áreas de Química e Física, sendo classificada em G06F e C12Q.

Um último aspecto importante diz respeito à contagem de patentes. Diferente de dados de exportação ou emprego, comuns na literatura, é frequente que uma patente seja classificada em mais de um de setor tecnológico, como já pudemos observar em alguns exemplos acima citados. Assim, uma patente que seja classificada como pertencente às classes (hipotéticas) A e B vai contar como uma patente em A e uma em B, ou seja, duas patentes para o país a que pertence. No entanto, de modo similar ao que vimos quanto a múltiplos aplicantes em uma só patente, não é um evento tão frequente e significativo. Em ambos os escritórios, quase metade

<sup>15</sup> Basta tomar a combinação de 10 em grupos de 2 (porque a ordem não importa):  $\binom{10}{2} = \frac{10!}{2!8!} = 45$ .

**Figura 6** – Proporção de Aplicantes por Posição

Fonte: Elaboração própria.

das patentes tem apenas uma classe tecnológica e quase um terço apenas duas. De fato, mais de 98% das patentes são classificadas em até 5 classes tecnológicas. No USPTO, uma patente de aplicante único de origem japonesa é a que mais concentra classificações e está associada a 38 setores tecnológicos, representando 0,000015% das patentes. No EPO, 5 patentes estão classificadas em 30 classes distintas ao mesmo tempo, cada uma pertencendo a apenas um país – França (2), Grã-Bretanha, Estados Unidos e Liechtenstein –, juntas, elas representam 0.000076% do total das patentes. Portanto, como não é possível determinar qual classe tecnológica mais se adequa a uma patente quando mais de uma classificação lhe ocorre, mantemos uma determinação múltipla, mas que, afinal, não representa uma parcela muito significativa das observações.

Além disso, algumas classes tecnológicas são ignoradas por existirem em apenas parte da amostra. Algumas delas não existiam no começo da amostra e só passam a existir depois. Esse é o caso de algumas classes mais próximas de “outros”, como H99Z, descrita como “assunto não tratado nesta seção”<sup>16</sup>, a seção sendo H, eletricidade. Outras, contudo, apontam para mudanças mais profundas em matéria

<sup>16</sup> Especificamente, em inglês, *subject matter not otherwise provided for in this section*. Inclusive, a descrição já nos dá uma pista de que não a encontraríamos antes de janeiro de 2006: foi quando ela foi criada! De fato, esse é um detalhe relevante, a classificação internacional de patentes (novamente, IPC) foi elaborada de forma a ser flexível à criação de novas classes. Assim, esse é um evento, ao contrário de inesperado, previsto.

de inovação tecnológica. É o caso da G16Y, uma classe voltada à “internet das coisas”<sup>17</sup>, cuja primeira patente em cada escritório foi aplicada apenas em 2017. Existem 12 classes dentro dessa condição (à numeração adotada, de 4 dígitos) no USPTO e 8 classes no EPO, representando, respectivamente, 0,09% e 0,01% das patentes. Também existem algumas classes que, se existem no primeiro período, deixam de existir em algum outro, como a B68F<sup>18</sup> que trata de artigos de couro. Uma escolha possível é considerar que o país ou região simplesmente deixou de se especializar naquela determinada tecnologia. O problema é que não é possível calcular alguns indicadores essenciais para compreender a dinâmica de entrada e saída duma determinada tecnologia. Como vimos no [Capítulo 2](#), a densidade mede a proximidade tecnológica que um país tem numa determinada tecnologia. Não estar especializado numa tecnologia não significa não saber nada sobre ela. Muitas vezes os países dominam tecnologias que são similares àquela e, assim, conseguem dominá-la com maior facilidade. Contudo, se não existe nenhum dado sobre aquela classe, não é possível determinar o que é similar a ela – ela simplesmente não existe naquele período de tempo. Portanto, por simplicidade<sup>19</sup>, iremos considerar apenas aquelas tecnologias que existiram em todo o período.

Como essas muitas escolhas indicam, usar dados de patente para medir performance tecnológica não é sem seus percalços. Por um lado, podemos adotar estratégias para minimizar alguns de seus problemas. É, por exemplo, o que esteve por trás da decisão de juntar em períodos maiores do que um ano. Ao estabelecer sub-períodos de 3, 5 ou 10 anos, é possível suavizar alguns ruídos, como supor que um país deixou de dominar uma certa tecnologia simplesmente porque naquele ano não houve a aplicação de nenhuma patente naquela classe. Também é isso que motiva a agregação da classe tecnológica em menos dígitos do que disponível. Será que (em tradução livre) “seguradores de gás com capacidade variável com peças em forma de anel telescopicamente móveis” (F17B 1/007) é tão diferente de “seguradores de gás

<sup>17</sup> De modo informal, trata-se da comunicação feita via internet entre objetos.

<sup>18</sup> Em inglês, *making articles from leather, canvas, or the like*.

<sup>19</sup> Seria possível, por exemplo, manter aquelas tecnologias que só deixam de existir no último período, já que usamos informações do período base apenas. Porém, isso implica uma complicação adicional que não altera fundamentalmente a análise – daí também o *por simplicidade*.

com capacidade variável com discos móveis” (F17B 1/013)? Talvez agregar ambas sob uma (talvez ainda muito específica) categorização de “seguradores de gás com capacidade variável”<sup>20</sup> (F17B) seja mais realista e capte um bloco de conhecimento que se expressa nas outras duas categorizações. Por outro lado, por mais que algumas dessas limitações possam ser contornadas, não é possível contornar outras. Para começar, patentes medem preferencialmente aquele conhecimento inovador, apesar de que, inovar ao imitar, como é muito frequente na indústria<sup>21</sup>, também é uma forma de adquirir novo conhecimento. Além disso, muitos conhecimentos acabam simplesmente não sendo formalizados na forma de uma patente. Às vezes, por uma questão espacial: assegurar uma patente nos Estados Unidos pode ser mais útil do que assegurá-la em um mercado em desenvolvimento, como o Brasil. Apesar dessas limitações, patentes são uma fonte importante de conhecimento a respeito da capacidade tecnológica que uma instituição, país ou região atualmente domina. Por isso, alavancaremos nossa análise por meio de um exercício empírico usando de dados de patente para os últimos 20 anos para dois dos maiores escritórios do mundo, norte-americano (USPTO) e europeu (EPO), para os primeiros 20 anos do século XXI<sup>22</sup>.

---

<sup>20</sup> Botijão!

<sup>21</sup> Basta olhar os casos de empresas de tecnologia modernas, como Baidu, cujo negócio é um espelho de outros, no caso, da Google.

<sup>22</sup> Paramos em 2019 para evitar o período da pandemia do Covid-19 e começamos em 2000 para evitar ruído quanto ao desaparecimento e surgimento de Estados Nacionais, a exemplo da União Soviética.

## 4 Exercício Empírico

Usando das técnicas apresentadas no capítulo anterior, este capítulo busca aplicar aquelas métricas para a base de dados de patentes do PATSTAT para os escritórios do USPTO e EPO no período 2000-2019. Buscamos com isso responder às lacunas na literatura apontadas no [Capítulo 2](#), conduzindo inicialmente uma análise exploratória dos dados na [seção 4.1](#), em especial com técnicas de análise de rede, e depois testando algumas hipóteses através de modelos índice (Probit e Logit) com efeitos-fixos para entrada em novos setores tecnológicos ao nível do país na [seção 4.2](#).

### 4.1 Análise Exploratória

Este trabalho preenche uma lacuna fundamental em relação a outros trabalhos, como descrito no [Capítulo 2](#) e [Capítulo 3](#): avalia a importância de capacidades disponíveis para os países por meio de laços com outros lugares, em especial ao nível internacional. Assim, uma primeira preocupação é entender o grau e as especificidades das colaborações entre os países.

Uma forma bastante comum de representar interconexões é por meio de uma rede, ou grafo. A [Figura 7](#) cumpre essa função representando a rede de colaboração entre os países a partir de dados de patente dos escritórios dos Estados Unidos (USPTO) e da Europa (EPO) entre 2000 a 2019. Nestas figuras, os dados estão agregados para todo o período. Quanto maior o círculo que representa um país, maior é a quantidade de links que esse país estabeleceu através da co-aplicação em patentes no período. A posição de cada país nessa rede é determinada através do algoritmo de direcionamento à força de Fruchterman-Reingold. A rede foi feita através do pacote *networkx* do *python*<sup>1</sup>. Inicialmente, cada nó recebe uma posição aleatória, daí usa-se a quantidade de conexões (vértices) de cada um para interagir com os demais

<sup>1</sup> Para mais detalhes, ver <https://networkx.org/documentation/stable/index.html>. A função para a posição é a *spring\_layout*.



por um número específico de vezes (no caso, definido em 100) até chegar a uma posição final. Nódulos que estão muito bem conectados, ou seja, que se conectam com uma quantidade maior de países diferentes, possuem uma tendência maior de terminar no centro da rede. Assim, os Estados Unidos (US) por ser o país com a maior quantidade de colaborações terminam no centro das duas redes, enquanto os demais são empurrados para as beiradas. No Apêndice A, as Tabela 5 e Tabela 6 mostram todos os países presentes na Figura 7 para o USPTO e EPO, respectivamente, assim como a quantidade de links e com quantos países distintos cada país se conectou.

Através da Figura 7, uma informação fica clara: países desenvolvidos são centrais nesta rede<sup>2</sup>. Isto talvez signifique que eles são os maiores beneficiários dessa rede de colaboração. Numa lógica inversa, talvez justamente por perceber a relevância desse tipo de colaboração, países desenvolvidos exploram as vantagens de participar dessa rede para manter sua competitividade tecnológica. No entanto, isso aponta para uma fonte de desigualdades persistente entre os países. Enquanto se conectar com outros países e regiões pode ser uma ferramenta relevante para romper o *lock-in*, essa rede sugere que países periféricos não conseguem explorar muito bem essa válvula de escape da falta de capacidades locais. Visto dessa perspectiva, laços internacionais, ao invés de mitigar desigualdades entre os países, podem na verdade *aumentá-las*, independentemente do fato de que, quando capazes de penetrá-la, países em desenvolvimento possam se beneficiar desproporcionalmente mais do que os países desenvolvidos.

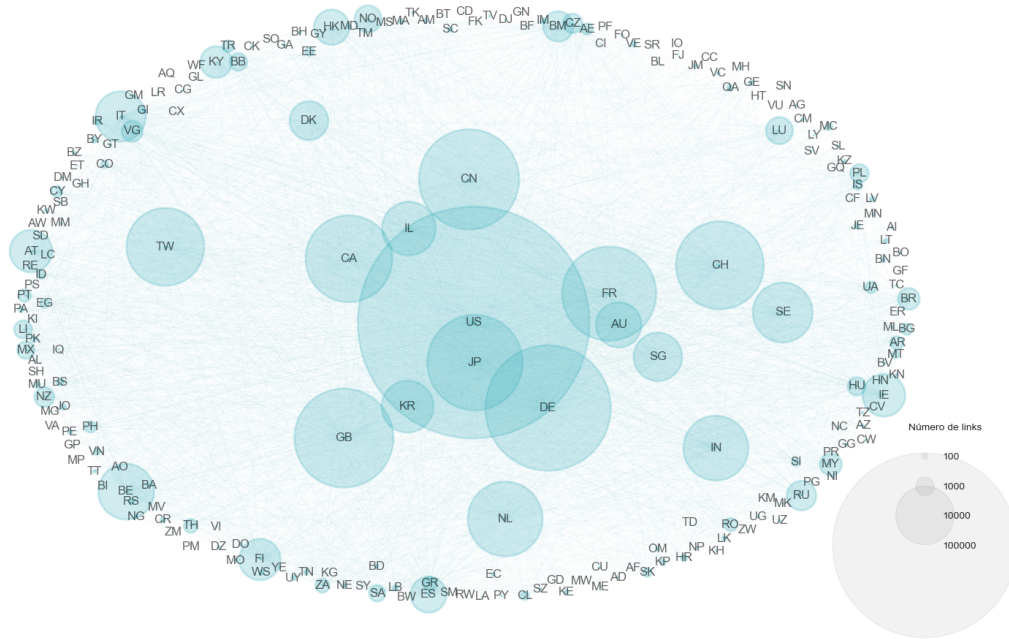
O problema é que o acesso de países periféricos a essa rede de colaborações é simplesmente muito pequeno. De fato, o caso da Suíça (CH) corrobora esse ponto. Apesar de ser um país europeu relativamente pequeno com uma população inferior a 10 milhões de pessoas<sup>3</sup>, ocupa uma posição central nesse rede, colaborando intensamente com outros países também centrais, como Japão (JP), França (FR) e Estados Unidos. No outro extremo, a Índia (IN), apesar de ter uma participação não desprezível nessa rede, a quantidade de colaborações que instituições indianas

<sup>2</sup> O que claramente incita a ideia de Centro-Periferia.

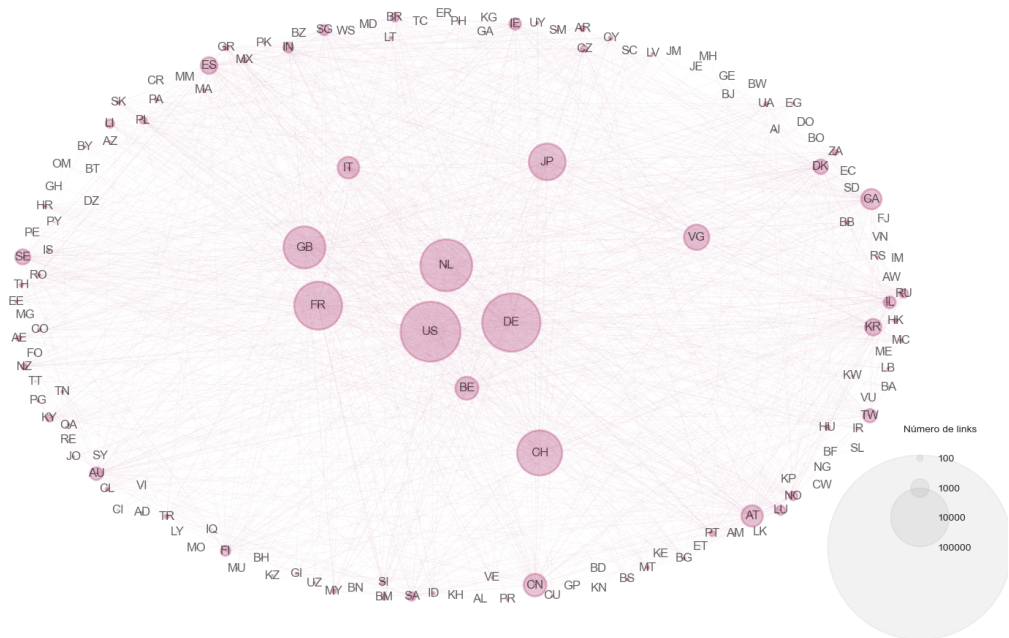
<sup>3</sup> Dados do Banco Mundial para 2021.

**Figura 7 – Rede de Colaboração Internacional**

**(a) USPTO**



**(b) EPO**



Fonte: Elaboração Própria com base nos dados do USPTO e EPO.

punderam estabelecer nesses dois escritórios é menor do que a Suíça. Como seria de se esperar, isso é ainda mais expressivo no escritório europeu, sendo até difícil de ser localizar a Índia no canto superior esquerdo<sup>4</sup>. O contraste entre as duas populações não poderia ser maior: a população indiana é quase 175 vezes a da Suíça! Obviamente, essa comparação é meramente ilustrativa, não querendo significar com isso que países com uma alta taxa de colaboração por habitante sejam aqueles desenvolvidos.

Uma forma de ver como isso acontece é através da distribuição do grau dos nós da rede de colaborações internacionais. Como argumentado na [seção 3.3](#), as redes que encontramos na realidade, ao invés de serem aleatórias, são na verdade sem escala, isto é, com muitos nós mal conectados e alguns poucos muito bem conectados – ambos numa frequência desproporcional com relação a uma rede aleatória. Esse comportamento pode ser descrito aproximadamente por uma lei de potência (BARABASI, 2016). A [Figura 8](#) demonstra esse comportamento para a rede internacional de colaborações em patentes. Através da escala logarítmica, fica clara uma relação linear entre o número de conexões e a frequência com que encontramos países com esse grau de conectividade, tal como esperaríamos de (3.16).

Regredindo uma variável pela outra, tal como em (3.17), encontramos o coeficiente  $-\gamma$  da lei de potência de  $-0.516$  e  $-0.569$  respectivamente para o USPTO e EPO, usando do grau dos nós com uma frequência ao menos maior do que 1%. Realizamos esse corte porque o comportamento de uma rede sem escala deve ocorrer apenas para uma rede finita. Dada a finitude dos dados que dispomos, eliminamos aqueles dados cuja frequência é muito pequena para parecerem significativos. Apesar disso, realizar os cálculos não muda substancialmente as conclusões. O coeficiente do USPTO e EPO passam para  $-0.664$  e  $-0.670$ , respectivamente. Mais relevante ainda, o  $R^2$ , o quanto da variação da dependente que a variável independente é capaz de explicar, é de  $0.820$  e  $0.794$  (USPTO e EPO), quando consideramos apenas parte dos dados (de acordo com aquele corte). Novamente, considerar todos os dados não muda muito a interpretação, conseguimos ainda um grau significativo de *goodness of fit*, com um  $R^2$  de, respectivamente,  $0.660$  e  $0.611$ .

---

<sup>4</sup> Se a rede fosse um relógio, seriam aproximadamente 11 horas.

Em suma, esses resultados indicam que a nossa rede pode ser considerada sem escala. Uma importante consequência desta característica, como argumenta [Ribeiro et al. \(2018\)](#) para a rede citações científicas, é que a rede colaborações internacionais que encontramos é hierárquica. Isto é, alguns nódulos (países) tem uma influência muito grande sobre sua configuração. O desaparecimento dos Estados Unidos nesta rede teria consequências muito mais drásticas do que a da Coreia do Norte. Esta hierarquia pode corroborar a ideia de que, ao invés de mitigar a desigualdade entre os países, esta rede de colaborações pode ser uma maneira dos países desenvolvidos de manter a dianteira na competitividade tecnológica.

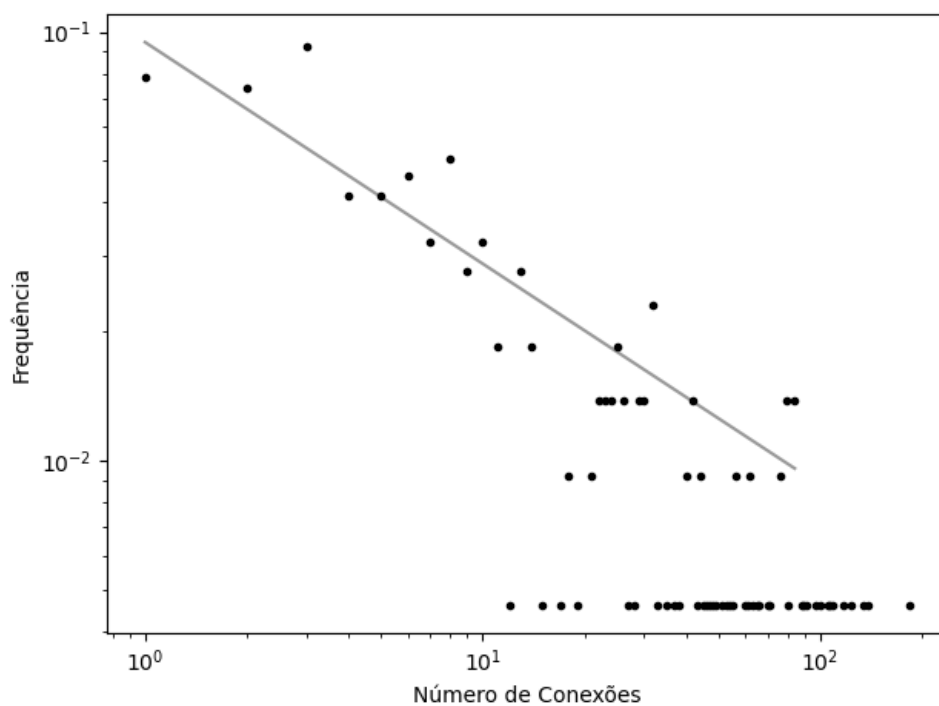
Também a [Figura 7](#), principalmente a parte (a) do USPTO, revela como que não apenas a quantidade mas a qualidade das colaborações importa. Israel (IL) possui praticamente a metade de colaborações em relação a Taiwan (TW) no escritório americano. Contudo, como Israel consegue se juntar a outros países mais relevantes nessa rede, sua posição final acaba sendo mais centralizada. Enquanto Israel possui 11 mil colaborações com os Estados Unidos, Taiwan possui apenas cerca de 7,5 mil colaborações com instituições norte-americanas. Com isso Israel termina mais centralizado na rede, mesmo que tenha menos laços no total do que Taiwan, já que seu principal parceiro é os Estados Unidos, sendo a China (CN) o principal colaborador com Taiwan, com cerca de 20 mil colaborações em todo o período entre os dois países<sup>5</sup>. Por sua vez, instituições chinesas tem como principais colaboradores instituições taiwanesas também, demonstrando como há uma interconexão muito forte entre as duas economias, enquanto que os Estados Unidos é apenas o segundo maior parceiro para os chineses.

A [Figura 7](#) permite também uma comparação interessante em termos de diferenças regionais entre os dois escritórios analisados. Primeiro, a quantidade de colaborações no USPTO é simplesmente muito maior do que no EPO – observe que mantemos a escala a fim de facilitar a comparação. Por um lado, isso reflete o

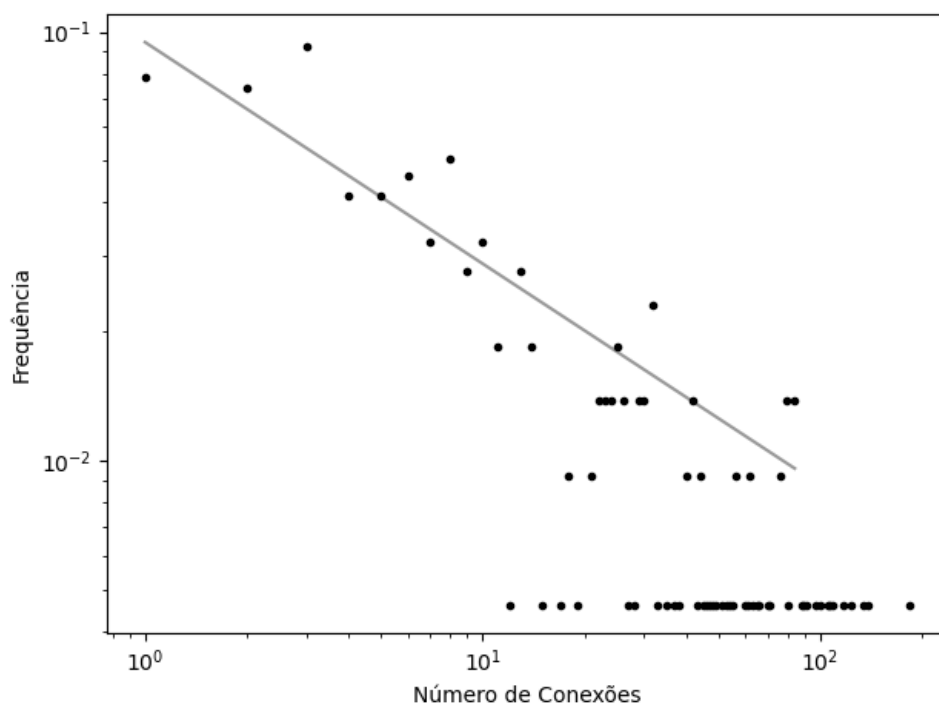
<sup>5</sup> Como os dois *países*, a República Popular da China (CN), também chamada de China Continental, e a República da China (TW) adotam ambas a política de uma só China, essa frase poderia ser absurda, o que se faz aqui é simplesmente adotar a classificação do PATSTAT para a origem das instituições aplicantes, dando o nome de “país” a cada uma dessas definições.

**Figura 8** – Distribuição da rede de colaborações internacionais

(a) USPTO



(b) EPO

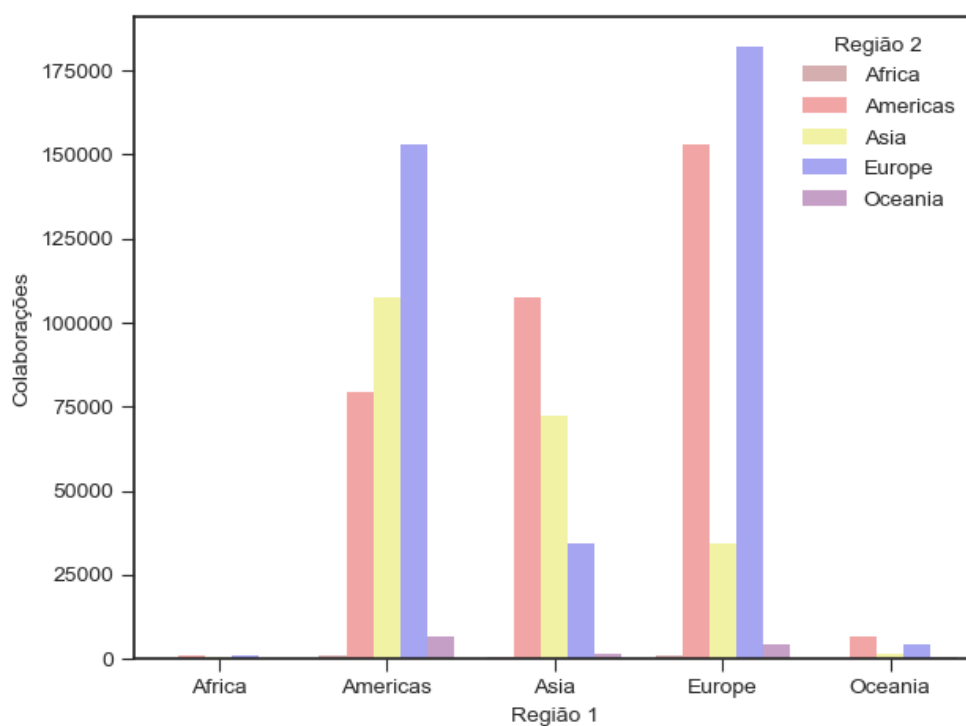


Nota: os coeficientes de inclinação  $|R^2|$  são, respectivamente para os escritórios:  $-0.516|0.820$  e  $-0.569|0.794$ .

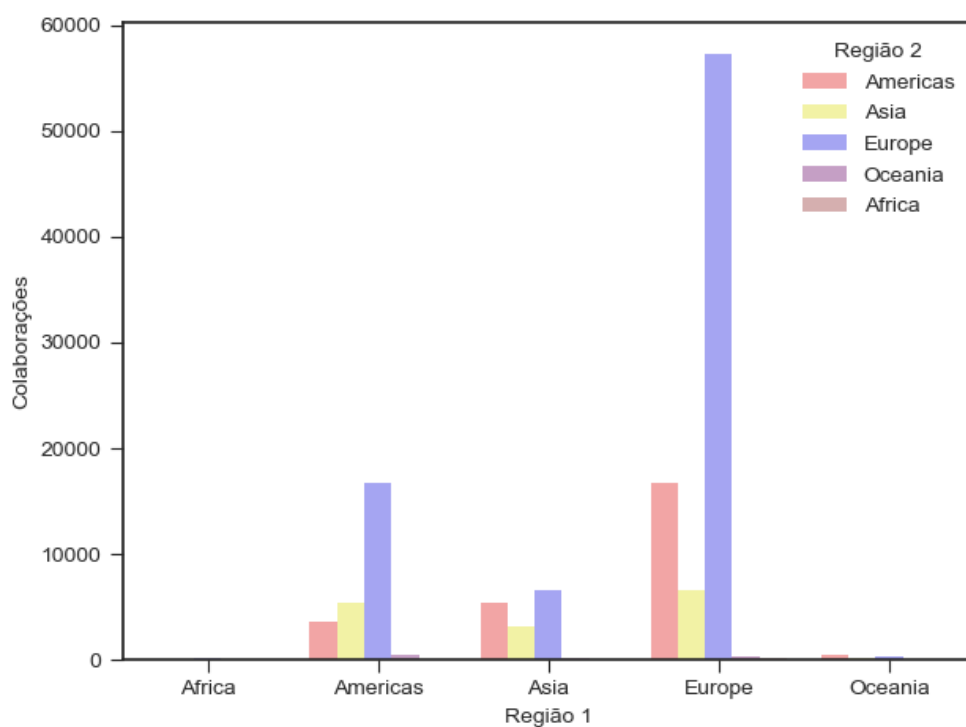
Fonte: Elaboração Própria com base nos dados do USPTO e EPO.

**Figura 9 – Distribuição das colaborações**

(a) USPTO



(b) EPO



Nota: divide-se em região 1 e 2 apenas para diferenciar o laço de “origem” e o de “destino”, conquanto é indiferente considerar um ou outro, já que não são links direcionados. Ou seja, cumpre fins de exposição.

Fonte: Elaboração Própria com base nos dados do USPTO e EPO.

tamanho de cada escritório em termos de aplicações de patentes. No mesmo período, depois de realizados os cortes mencionados na [seção 3.4](#), o USPTO com cerca de 12 milhões de aplicações tem um pouco mais do dobro do tamanho do EPO, este com cerca de 5 milhões de aplicações no mesmo período, 2000 a 2019. Por outro lado, a quantidade total de colaborações é quase 8 vezes maior no USPTO do que no EPO, respectivamente, com cerca de 477 mil e 62 mil colaborações no total do período. Portanto, podemos com isso ver como o perfil de aplicações no escritório dos Estados Unidos é muito mais internacionalizado do que no da Europa. Em especial, isso é curioso dado que seria de se esperar que instituições europeias colaborassem mais entre si, o que deveria ser visível principalmente quando observássemos o escritório europeu, em que essas instituições devem ter uma participação relativamente maior do que nos demais escritórios de patentes ao redor do mundo. E, como existem muito mais estados nacionais para uma mesma população e área geográfica em comparação aos Estados Unidos, poderíamos sem muito furor esperar também que o grau de colaboração internacional fosse maior no escritório Europeu. Apesar de os dados corroborarem tanto que países europeus colaboram mais entre si do que com outros países e que esse efeito é mais expressivo no EPO, como a [Figura 9](#) indica, em geral, o grau de colaboração é muito menor no escritório europeu. Além disso, como a [Figura 9](#) nos permite observar, as colaborações estão muito menos distribuídas no EPO, o que indica o caráter principalmente regional, em comparação à internacionalização do escritório americano.

Essa comparação permite inclusive perceber o quão relevante instituições norte-americanas são no sistema global de inovação. Mesmo no EPO, os Estados Unidos ainda é o país que mais estabeleceu laços com outros países, apesar de nesta rede ser menos central. Contudo, isso é apenas um lembrete da importância de instituições americanas, já que o USPTO é o maior escritório de patentes em termos de aplicação. Apesar de naturalmente o país “natal” do escritório ter uma participação mais expressiva, isso ainda quer dizer que os EUA é o país que mais colabora no maior escritório de patentes do mundo. No USPTO, apenas 16 países tem mais de 1% de aplicações que os Estados Unidos tem, e mesmo no EPO apenas 23 atingem essa

meta. Em ambos os escritórios, o Japão (JP) é o segundo maior aplicante, seguido da Alemanha (DE) em terceiro lugar. Entretanto, enquanto no USPTO o Japão tenha só 37% da quantidade de aplicações dos Estados Unidos (100%) e a Alemanha 11%, no EPO, essas proporções são de cerca de 64% e 56% respectivamente, demonstrando como o escritório é menos dominado por instituições norte-americanas, mas que ainda é um grau de preponderância expressivo. Normalmente, a posição do país em relação à quantidade de aplicantes é similar à sua posição quando à quantidade de colaborações. Os Estados Unidos, por exemplo, é o principal nos dois rankings. O Japão, porém, perde posições nas duas, enquanto a Alemanha acaba tomando o segundo posto. Um caso interessante é o da China, que apesar de em termos de aplicações ser apenas a nona no USPTO, com cerca de 9% da quantidade de aplicações dos Estados Unidos, é o terceiro país cujas instituições mais colaboram internacionalmente. Curiosamente, no EPO, sua posição em termos de aplicações sobe para a sexta colocação, mas em termos de colaboração cede lugar a outros países europeus, como os Países Baixos (NL), terminando na décima colocação<sup>6</sup>.

Em parte, é esse papel central da economia norte-americana que explica como alguns “países” são desproporcionalmente relevantes nessa rede. Basicamente, os Estados Unidos atua como uma porta de entrada para algumas entidades “sub-nacionais” que funcionam na prática como paraísos fiscais. Por um lado, esse comportamento é motivado pelo viés do escritório, que faz com que instituições de determinados locais muito ligados àquela economia que sedia o escritório tenham uma presença desproporcional.

No USPTO, um caso que se destaca é o das Ilhas Cayman (KY). A maior quantidade de patentes das Ilhas Cayman neste escritório se concentra em primeiro lugar na classe H01L, dispositivos semicondutores, em segundo na G06F, processamento de dados digitais, e em terceiro H04L, transmissão de informação digital. Portanto, as três principais classes que instituições registradas como pertencentes às Ilhas Cayman são relacionadas a tecnologias de eletrônicos e de meios digitais de

<sup>6</sup> Nas tabelas disponíveis no [Apêndice A](#) é possível conferir algumas dessas informações. Em termos de ranking, basta se atentar que a contagem começa no 0, não no 1 – isso decorre da estrutura de indexação do *python* e que, por praticidade, é mantida aqui.



comunicação.

No EPO, as Ilhas Virgens Britânicas (VG) ficam mais destacadas, o que já aponta o caráter, digamos, suplementar tanto das Ilhas Virgens Britânicas para o escritório europeu quanto das Ilhas Cayman para o escritório norte-americano. O maior país em termos de aplicações no EPO não são as Ilhas Virgens Britânicas, mas Liechtenstein (LI), que não é tão representativo em termos de colaborações, conquanto aponte no mesmo caráter regional da importância desses paraísos fiscais.

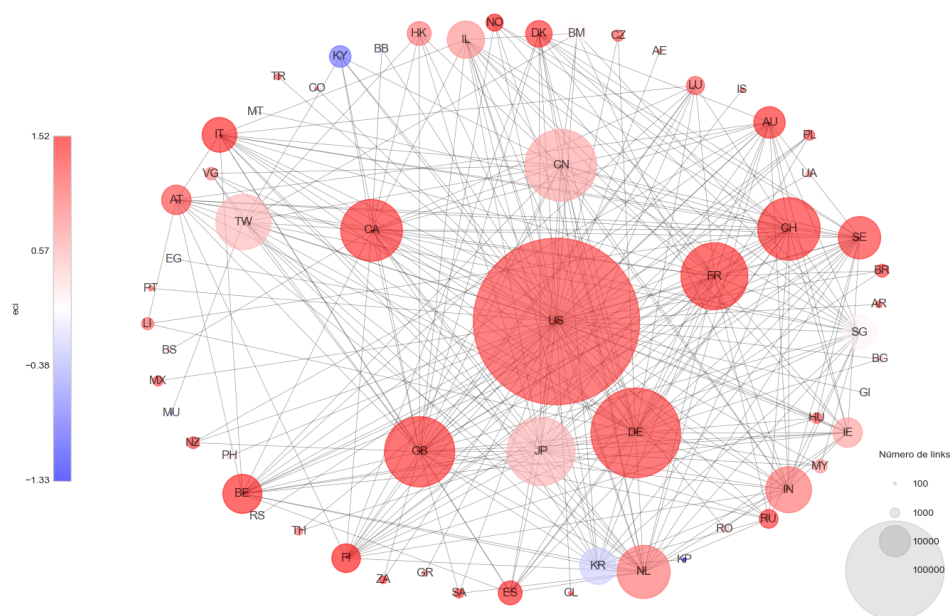
A classe mais comum nas patentes das Ilhas Virgens Britânicas no EPO já apareceu como a terceira maior para as Ilhas Cayman no USPTO: H04L, transmissão de informação digital. Isto indica que o perfil de aplicações para esse tipo de “entidade subnacional” é bastante similar. De fato, em ambos escritórios, a classe mais comum é a G06F, processamento de dados digitais, que também já foi citada, sendo a segunda mais frequente das Ilhas Cayman no USPTO.

Contudo, enquanto esses paraísos fiscais revelam o perfil do escritório e sua presença (desproporcional) na rede de colaborações internacionais está relacionada a ele, por outro lado, essa presença também pode ser explicada pela ausência de capacidades locais, justamente por serem “países” muito pequenos e frequentemente com economias muito dependentes do mercado internacional, ou ainda de países específicos – como as Ilhas Cayman com os Estados Unidos. Portanto, ao invés de serem uma fonte de perturbações na análise que se pretende, observar paraísos fiscais pode ser um objeto bastante útil para avaliar como colaborações internacionais são importantes para a performance local. Em especial, a concentração de patentes em, de maneira genérica, tecnologias digitais que encontramos associadas a esses territórios está muito relacionada à expansão global de companhias de plataformas digitais (NETO et al., 2022), com paraísos fiscais sendo um atrator especial deste tipo de companhia.

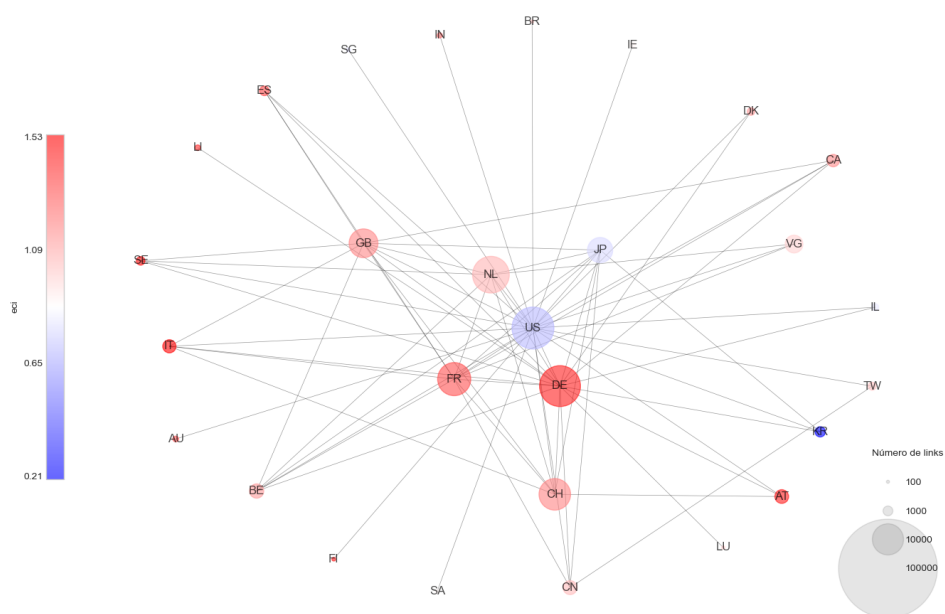
Na [Figura 10](#), expandimos a ideia da figura anterior para avaliar se realmente países desenvolvidos se encontram mais ao centro da rede de colaborações inter-

**Figura 10** – Rede de Colaboração Internacional  
 Colorida pelo Complexidade do País

(a) USPTO



(b) EPO

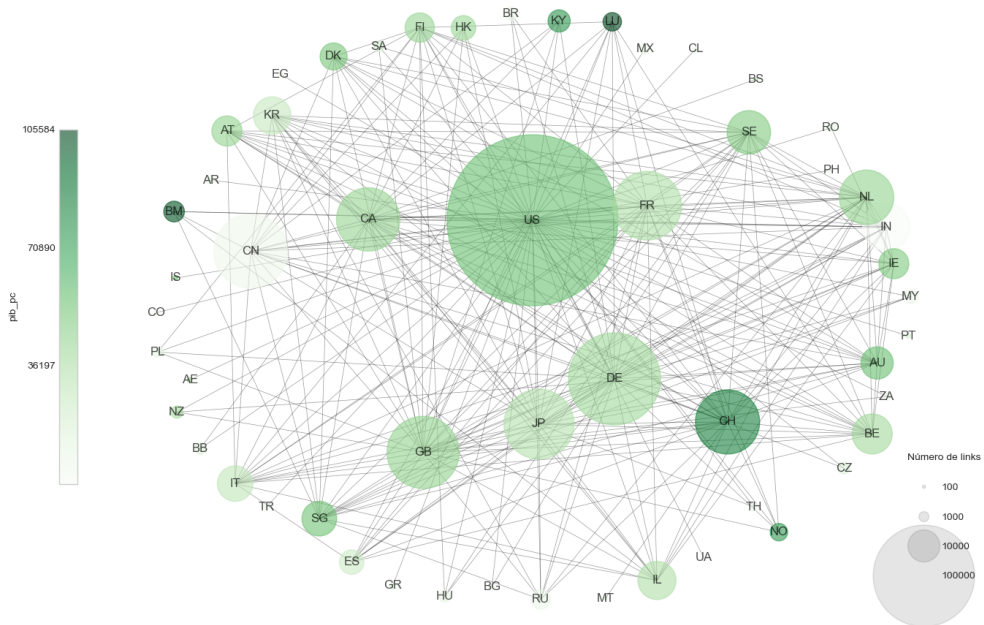


Nota: Nódulos são coloridos pela complexidade das respectivas economias representadas. Quão mais vermelho, mais complexo é o país. Analogamente, quão mais azul, menos complexo é o país.

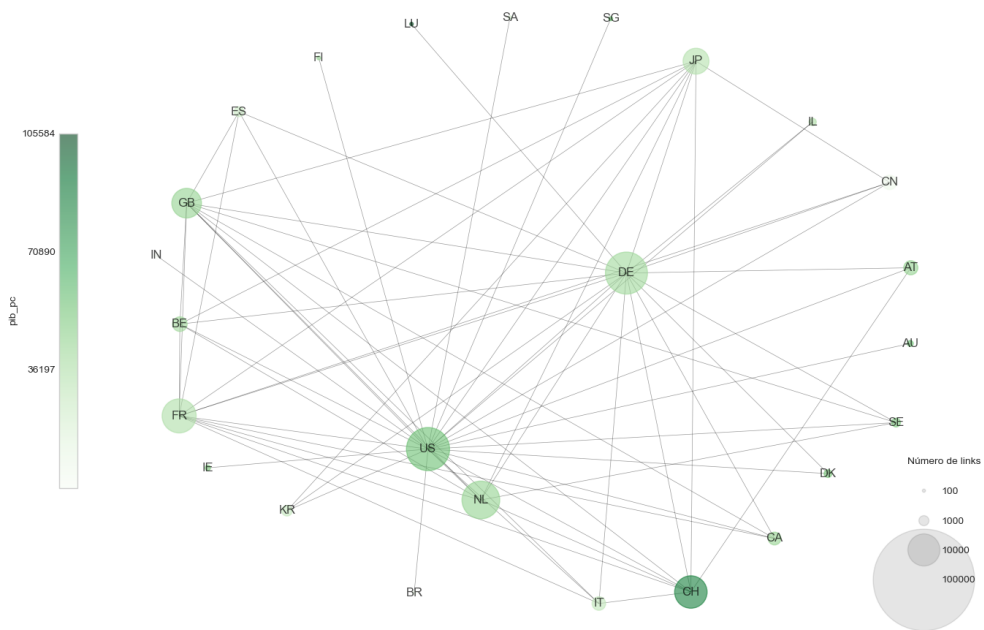
Fonte: Elaboração Própria com base nos dados do USPTO e EPO.

**Figura 11** – Rede de Colaboração Internacional  
Colorida pelo PIB per capita

(a) USPTO



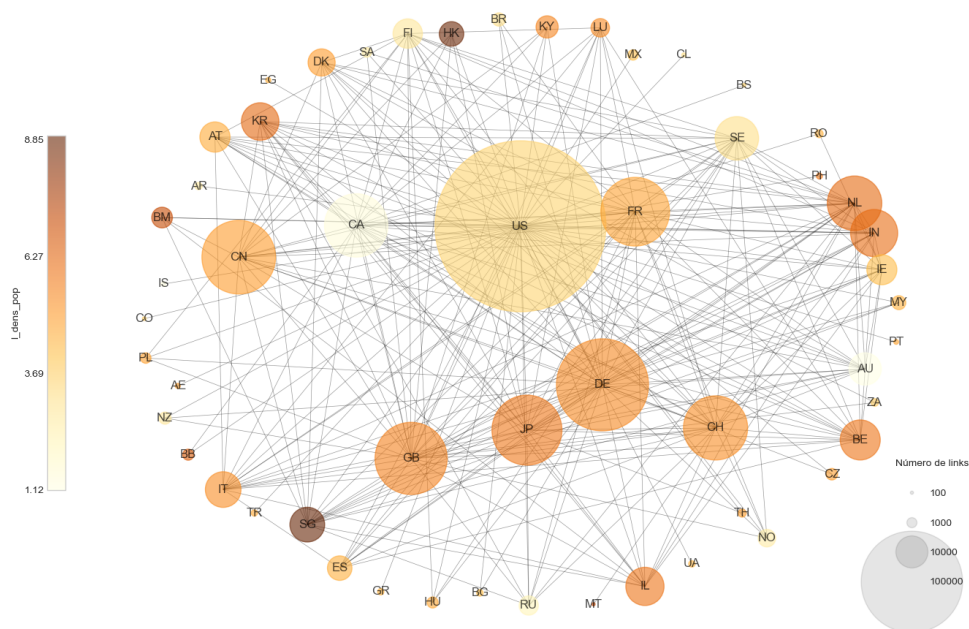
(b) EPO



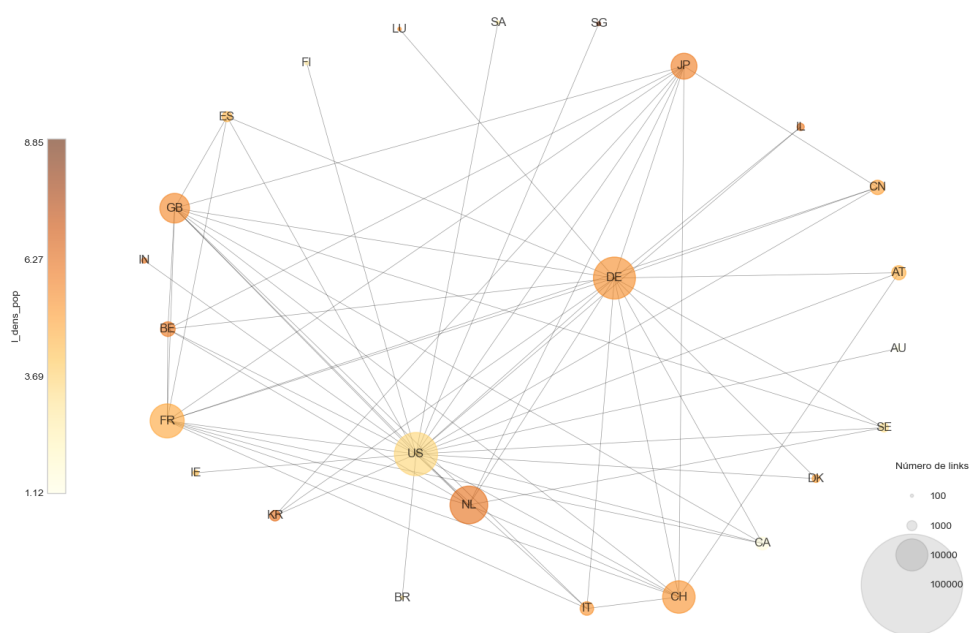
Fonte: Elaboração Própria com base nos dados do USPTO e EPO.

**Figura 12** – Rede de Colaboração Internacional  
 Colorida pela Densidade Popacional

(a) USPTO



(b) EPO



Fonte: Elaboração Própria com base nos dados do USPTO e EPO.

nacionais. Cada nóculo recebe uma cor que varia de acordo com a complexidade tecnológica do país que representa. A fim de destacar a diferença entre os países, a escala de cor adotada apresenta uma ruptura entre azul e vermelho, sendo que, quão mais vermelho, mais complexo um país é – e, analogamente, quão mais azul, menos complexo. Nesta rede, por simplicidade, estão representados apenas os pares de países que tiveram ao menos 100 colaborações<sup>7</sup>. Como transparece pela presença de círculos avermelhados, a maioria dos países que “sobrevivem” nessa rede são relativamente complexos. As Ilhas Cayman (KY) e a República da Coreia (KR), ou Coreia do Sul, se destacam como exceções. Talvez o ponto mais relevante que essa forma de observar os dados permita ver é como os Estados Unidos, apesar de ainda ter uma posição central no escritório europeu, perde relevância ao ponto de se tornar um país “pouco complexo”<sup>8</sup>, similar ao que acontece com o Japão. Comparativamente, isso explica o quase desaparecimento de Israel nesta rede. No USPTO, instituições israelenses alavancam uma quantidade não tão grande de colaborações pela “qualidade” dos laços que fazia com seu principal parceiro, os Estados Unidos. No escritório europeu, porém, essa estratégia já não é mais tão eficaz, fazendo com que Israel perca relevância nela.

A [Figura 11](#) e [Figura 12](#) cumprem a mesma função da [Figura 10](#), mas para PIB per capita e densidade populacional dos países, respectivamente. Ambas passam um padrão claro: a falta de padrão! Ao contrário da [Figura 10](#), temos mais dificuldade em dizer que países ricos, ou mesmo que países densamente povoados – em acordo com a ideia de economias de aglomeração –, estão no centro dessa rede. Por contrário, principalmente para a rede que colore os nóculos de acordo com sua densidade populacional, não são raros os casos “anômalos” de países de baixa densidade populacional povoando o seu centro, como é o caso dos Estados Unidos e do Canadá (CA). Isso não quer dizer, porém, que não há uma influência dessas variáveis em explicar o grau de colaboração de um país com outros, mas ao menos

<sup>7</sup> Em especial, essa escolha será útil para a comparabilidade com as demais redes, que colorimos de acordo com o PIB per capita e a Densidade Populacional, dada a restrição quanto a disponibilidade dos dados. Ver [seção 3.4](#) para como exatamente uma colaboração é definida.

<sup>8</sup> Neste caso, menor do que zero e adotando a escala de cor automaticamente definida pelo pacote do *python*, novamente, *networkx*.

significa que isso não é imediatamente visível. Claro que, em parte, isso é resultado da limitação que existe nesse tipo de representação. Países menos expressivos nessa rede são menores. Assim, é difícil observar qual a cor que tomam, o que dificulta a identificação de padrões. Contudo, ambas as figuras ainda indicam a importância de considerar outras métricas de desenvolvimento tecnológico que levem em conta outras dinâmicas que meramente renda, por exemplo. O indicador de complexidade econômica desenvolvido por Hidalgo e Hausmann (2009) e depois adaptado para o uso com dados de patente em Balland e Rigby (2017), por exemplo, pode ser uma métrica importante para captar o avanço tecnológico dos países, e as redes exploradas corroboram essa ideia.

A Figura 13 revela ainda outra face das desigualdades internacionais entre os países. No eixo horizontal, está a complexidade tecnológica dos países (ECI)<sup>9</sup>, no eixo vertical, está a correlação entre a complexidade da classe tecnológica (PCI)<sup>10</sup> e a densidade em torno dela. Como explicado no Capítulo 2, comumente chamada de “Curva S” (embora aqui lembre o formato da letra apenas vagamente), a Figura 13 ilustra que países mais complexos (mais à direita na abscissa) possuem capacidades que lhes permitem acessar produtos complexos mais facilmente. Cada círculo representa uma observação do tipo país-período. Além disso, o tamanho de cada círculo representa a quantidade de patentes para cada país-período e a cor representa a quantidade de links que o país pôde estabelecer em todo o período da amostra – quanto mais escura a cor, maior a quantidade de links.

Como esperávamos, países menos complexos possuem uma correlação negativa entre densidade, ou proximidade tecnológica, e complexidade das classes que *não* dominam<sup>11</sup>. Assim, as tecnologias que estão facilmente acessíveis para eles são tecnologias pouco complexas, contribuindo para sua permanência num estado de subdesenvolvimento. Conseguimos perceber também que os círculos são menores para países menos complexos, indicando que sua produção tecnológica medida pelo número

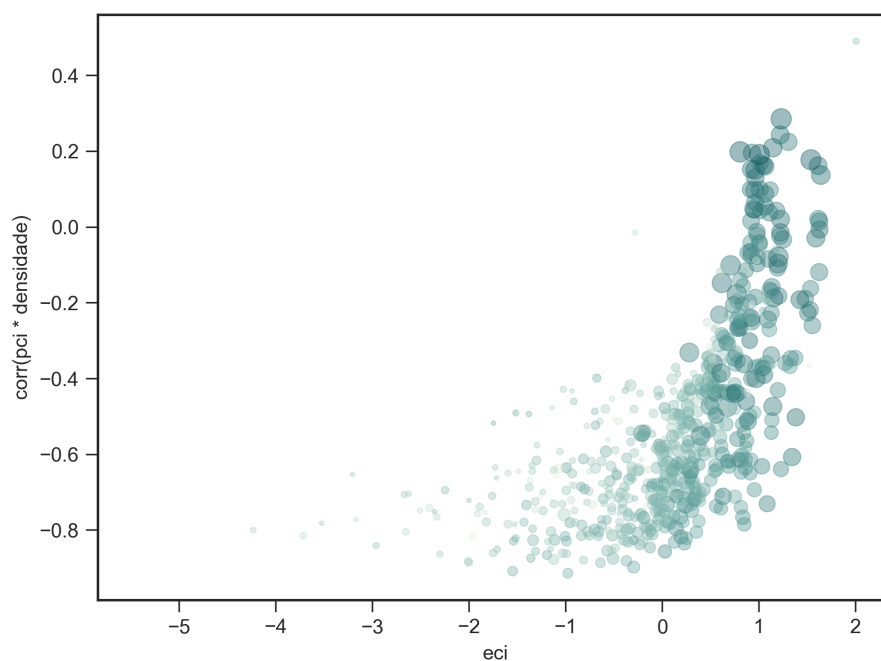
<sup>9</sup> Por simplicidade, adotamos a sigla em para o índice *econômico*, não tecnológico como seria o correto, mas que permite uma compreensão mais imediata do que se trata.

<sup>10</sup> Analogamente à outra sigla, aqui “P” se refere à classe, não ao *produto*.

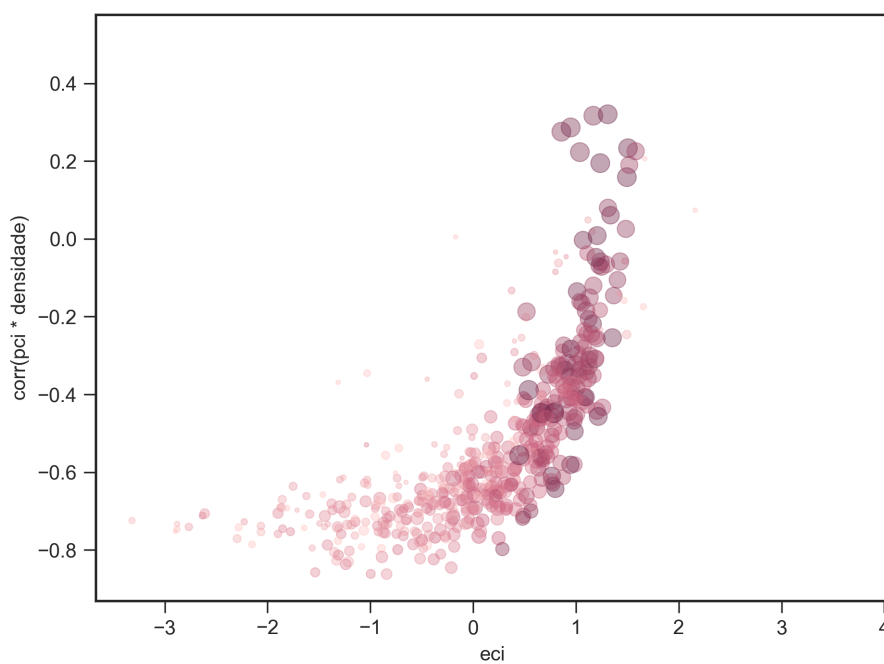
<sup>11</sup> Este é um detalhe importante, porque queremos saber aquelas tecnologias que os países podem passar a se especializar.

**Figura 13 – Curva S**

(a) USPTO



(b) EPO



Nota:  $\text{corr}(\text{pci} * \text{densidade})$  é a correlação entre as variáveis. Tamanho dos círculos está de acordo com o (logaritmo) da quantidade de patentes para cada observação (país-período). A cor dos círculos representa a quantidade (em logaritmo) de colaborações de cada país da observação. É interessante como o formato mais se assemelha a um J do que a um S, cuja relação fica a ser explorada se depende de algum fundamento ou é decorrente da maneira como se escolheu apresentar os dados, com o tamanho e a coloração das bolas.

Fonte: Elaboração Própria com base nos dados do USPTO e EPO.

**Tabela 1** – Correspondência das Variáveis

Código	Nome	Fonte
entry	Entrada	USPTO*
density	Densidade	USPTO*
compl	Complementariedade	USPTO*
pai	Complexidade da Classe Tecnológica	USPTO*
eci	Complexidade Tecnológica do país	USPTO*
pib_pc	PIB per capita	Banco Mundial
dens_pop	Densidade Populacional (hab / km <sup>2</sup> )	Banco Mundial

Nota: \* significa que foi calculada. Variáveis com prefixo “l\_” estão em logaritmo Natural

Fonte: Elaboração Própria

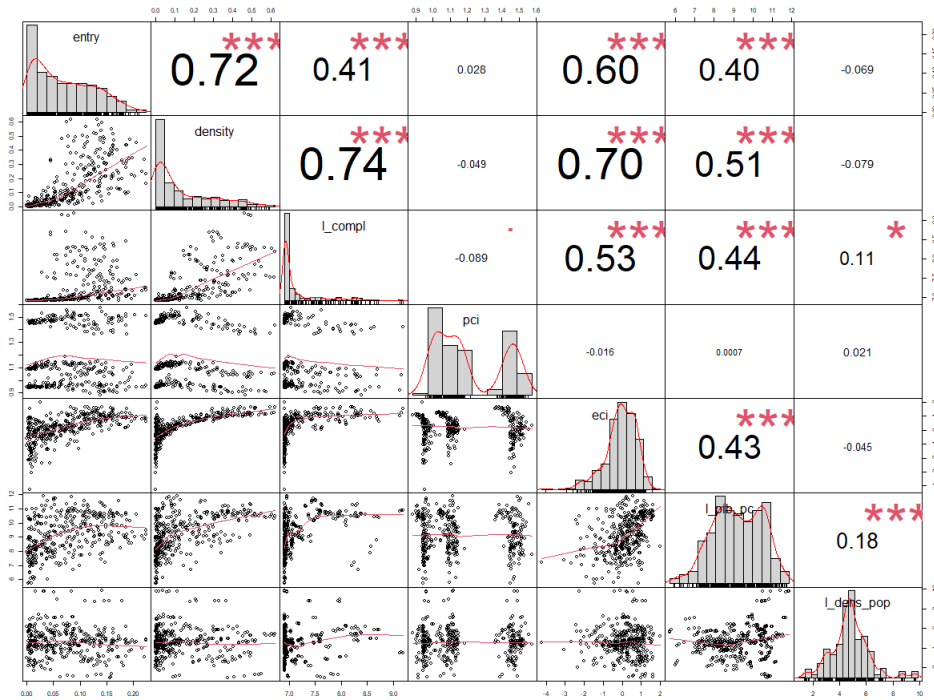
de patentes é relativamente menor do que de países complexos. Mais ainda, adicionar essa informação permite deixar mais claro o padrão, em especial, para o EPO existem algumas observações que destoam principalmente à direita. Sem levar em conta a representatividade dessa observações, poderíamos facilmente suspeitar do padrão ascendente que as demais observações parecem indicar. Além disso, corroborando com aquilo que já tínhamos observado nas redes anteriores, países mais complexos são também aqueles com a maior quantidade de colaborações, o que é representado pela tendência de ter cores mais escuras conforme nos movemos à direita no gráfico. A [Figura 13](#) mascara um pouco essas informações, já que, por fins estéticos, o valor tanto para o tamanho quanto para a cor dos pontos é o logaritmo das respectivas informações. Caso contrário, ou as observações dos países menos complexos sumiriam ou a dos mais complexos dominariam completamente o gráfico. De qualquer forma, como vetor, a figura ainda aponta no mesmo sentido.

A [Figura 13](#) fermenta algumas dúvidas que já vinham num crescente até agora. Será que ter mais laços com outros países auxilia no avanço em novas tecnologias? Será que a evidência internacional mantém os achados do Princípio do Relacionamento, de que é mais fácil diversificar para aquilo que está próximo? A [Figura 14](#) e a [Figura 15](#) começam a responder a essas perguntas por meio de um recurso estatístico básico, apresentando a correlação entre as principais variáveis de interesse, a primeira com a média (país-período) e a segunda sem tirar a média das observações, mas apenas

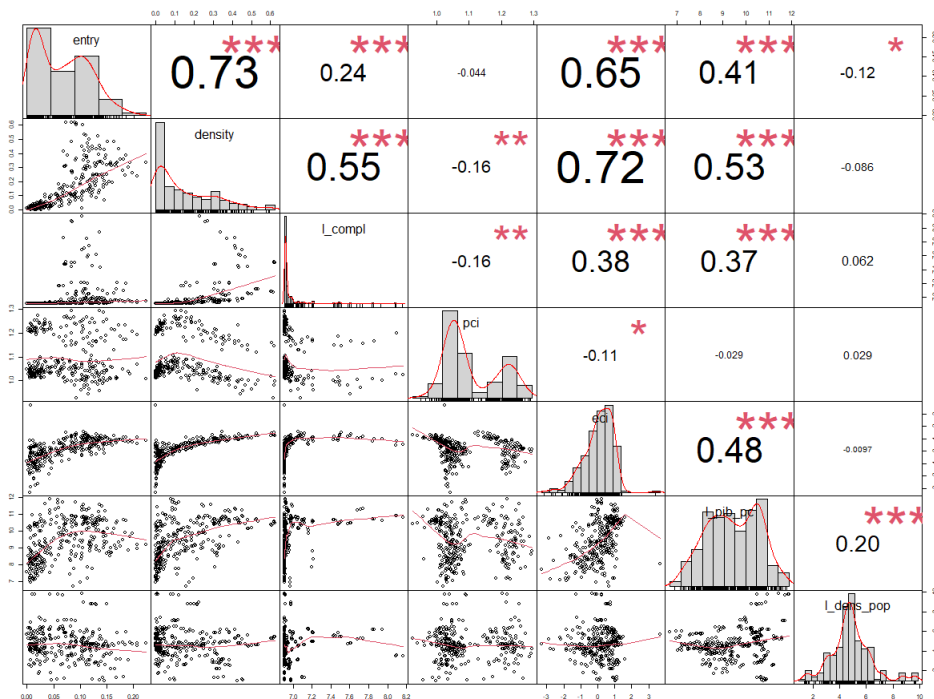


Figura 14 – Correlogramas

(a) USPTO



(b) EPO

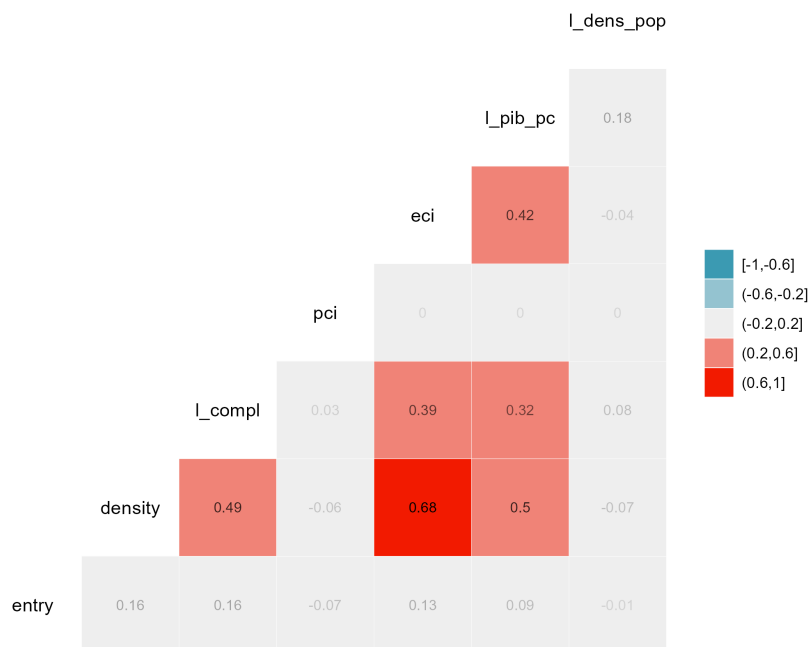


Nota: os dados representados são por médias por período e país (senão os gráficos de dispersão seriam basicamente ilegíveis).

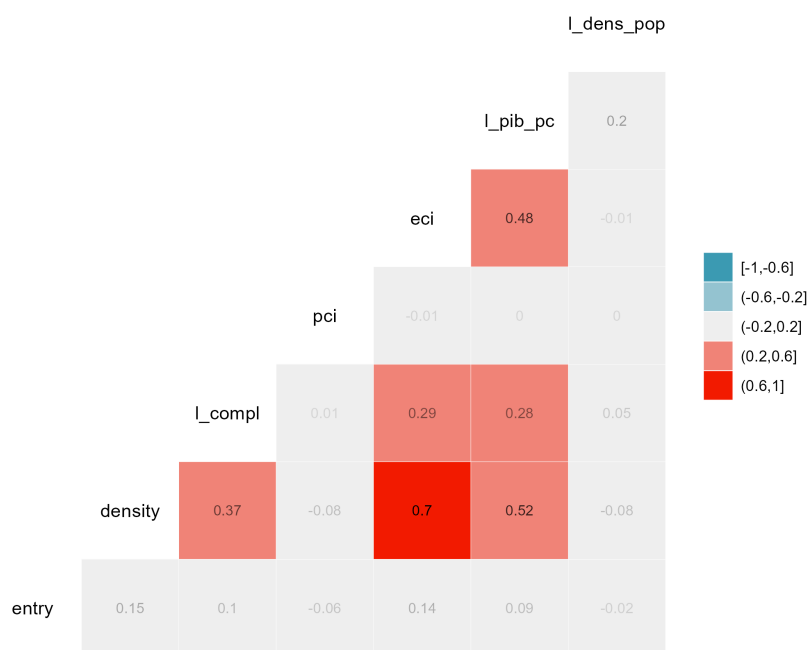
Fonte: Elaboração Própria com base nos dados do USPTO e EPO.

**Figura 15** – Correlogramas por Cor

(a) USPTO



(b) EPO



Fonte: Elaboração Própria com base nos dados do USPTO e EPO.

**Tabela 2** – Estatísticas Básicas

## (a) USPTO

	Min	1Qua	Mediana	Média	3Qua	Max
n_pat	0	0	0	33,170	0	170 333
diversity	1	9	38	80,750	136	369
ubiquity	1	11	16	17,840	23	92
eci	-4,246	-0,329	0,236	0,126	0,810	2,007
pci	-4,252	0,641	1,321	1,174	1,857	3,621
density	0	0,012	0,054	0,129	0,208	1
rca	0	0	0	0,512	0	2 554,344
mcp	0	0	0	0,085	0	1
entry	0	0	0	0,072	0	1
compl	0	2,960	18,850	946,990	132,030	98 297,590
n_link	1	41	174	6 928	1 760	286 501
dens_pop	2,558	46,461	98,990	548,170	206,329	18 788,450
pib_pc	302,800	3 507,700	9 011,200	19 871	32 043,800	144 443,500
pop	28 190	2 313 000	8 002 000	44 950 000	29 530 000	1 354 000 000
l_compl	6,908	6,911	6,926	7,166	7,032	11,506
l_dens_pop	0,939	3,839	4,595	4,609	5,330	9,841
l_pib_pc	5,713	8,163	9,106	9,140	10,375	11,881

## (b) EPO

	Min	1Qua	Mediana	Média	3Qua	Max
n_pat	0	0	0	19,030	0	24 621
diversity	1	15	61	90,570	151	371
ubiquity	1	10	14	15,860	20	68
eci	-3,319	-0,431	0,218	0,070	0,760	3,697
pci	-4,491	0,527	1,240	1,096	1,829	4,417
density	0	0,019	0,084	0,146	0,244	1
rca	0	0	0	0,551	0	4 652,230
mcp	0	0	0	0,094	0	1
entry	0	0	0	0,072	0	1
compl	0	0,462	3,215	146,188	23,876	7 113,223
n_link	1	10	48	1 142	362	19 212
dens_pop	2,558	47,880	100,073	652,472	217,754	18 788,450
pib_pc	815,600	4 414,400	11 631,200	22 454,600	34 807,600	144 443,500
pop	28 190	2 910 000	9 010 000	51 790 000	37 420 000	1 354 000 000
l_compl	6,908	6,908	6,911	6,989	6,931	9,001
l_dens_pop	0,939	3,869	4,606	4,667	5,383	9,841
l_pib_pc	6,704	8,393	9,361	9,391	10,458	11,881

Nota: as estatísticas fornecidas aqui são do tipo período-país-tecnologia e se referem aos dados utilizados nas regressões, portanto, valem apenas para os primeiros três sub-períodos.

Fonte: Elaboração Própria com base nos dados do USPTO e EPO.

com cores<sup>12</sup>. A [Tabela 1](#) traz a correspondência entre os códigos das variáveis e seu nome, bem como a fonte de onde foram extraídas.

Como podemos observar da comparação entre as duas imagens, a correlação só muda quando envolve alguma variável que varia ao nível país-tecnologia-período, como densidade ou entrada. Contudo, para aquelas cujo valor permanece fixo dentro de um mesmo país (*e.g.* PIB pc) ou numa tecnologia (*e.g.* PCI), a correlação é igual. Pela parte superior da [Figura 14](#), vemos que a correlação entre muitas dessas variáveis é expressiva, o que é confirmado pelos fortes tons avermelhados na [Figura 15](#) – inclusive, é curioso como não há nenhuma correlação negativa, o que seria evidenciado por uma tonalidade azul em duas das cinco quebras naturais. Esta impressão da mera correlação, entretanto, é em parte questionada pela parte inferior da [Figura 14](#), já que o gráfico de dispersão não gera uma impressão tão vívida. Na diagonal principal, ainda representamos o histograma dessa variáveis. Muitas delas apresentavam uma dispersão assimétrica à direita, isto é, com muitos valores pequenos – à esquerda – e poucos valores grandes. Portanto, a fim de normalizar a distribuição das variáveis que apresentavam esse comportamento, utiliza-se o logaritmo natural, representado pelo prefixo “l\_”. Em especial, a variável de complementariedade tinha muitos valores próximos a zero, assim, o logaritmo ainda não resolvia a assimetria de sua distribuição. A estratégia adotada foi somar 1000 antes de passar o logaritmo. Apesar desses arranjos, como ainda é possível observar, a assimetria não é inteiramente corrigida, o que pode dentro de uma regressão estatística prejudicar a inferência sobre os modelos. A [Tabela 2](#) traz algumas das principais estatísticas descritas das variáveis utilizadas.

Talvez a característica mais marcante dos dois correlogramas seja a homogeneidade. Se não fosse pelos números, poderíamos até pensar que a parte (a) e (b) da [Figura 15](#) são iguais! Algumas diferenças acabam acontecendo, mas são tão poucas que acabam por reforçar a ideia de que existe um padrão compartilhado entre os dois escritórios. Essa semelhança nos permite fazer uma análise em comum para ambos. Destaca-se, primeiro, que densidade tecnológica e complexidade possuem

---

<sup>12</sup> A razão é que os gráficos de dispersão na parte inferior da [Figura 14](#) ficavam muito poluídos, já que existem alguns milhares de observações.

alta correlação positiva. De fato, as discussões até agora não tornam esse achado surpreendente. Tanto nas redes de colaboração internacional, quanto nas “Curvas S”, já pudemos observar que existe uma tendência para o aprofundamento de desigualdades internacionais. Por um lado, países mais complexos são em parte definidos por serem aqueles que possuem mais capacidades. Assim, é senão natural que a *densidade*, ou a quantidade de conhecimentos similares a uma tecnologia, seja em média mais elevada para esses mesmos países. Por outro lado, nossa análise da rede de colaborações internacionais permitiu observar que países mais complexos colaboram mais e, assim, complementam a abundância que dispõem de capacidades locais com ainda outras, extra-locais. Isto resulta num ciclo de que países desenvolvidos, por serem tecnologicamente complexos, conseguem desenvolver suas capacidades tecnológicas num passo ainda mais rápido. Por isso, países mais desenvolvidos em geral possuem maior densidade nas classes tecnológicas, já que dominam mais capacidades, seja pela sua disponibilidade local ou não local.

Destaca-se também a correlação positiva entre densidade e o índice de complementariedade. Em parte, essa correlação acaba perpassando para outras variáveis, como a de complementariedade e complexidade, uma vez que densidade e complexidade também estão altamente correlacionadas – ou seja, é como se por tabela laços complementares estivessem correlacionados com a complexidade tecnológica dos países. Por sua vez, a correlação entre complexidade e PIB per capita também acaba se refletindo na correlação entre densidade e complementariedade. Uma possível interpretação dessa correlação entre densidade e complementariedade é que os países estabelecem links complementares naquelas classes tecnológicas que já possuem muitas capacidades. Com isso, tais colaborações poderiam servir como um empurrão adicional para entrar numa determinada tecnologia, reforçando a noção de complementariedade. Nesse sentido, a [Figura 14](#) corrobora essa versão, dado que a correlação, na média, entre complementariedade e entrada é positiva e significativa mas menos do que entre densidade e entrada.

Na [seção 4.2](#), usamos de regressões para testar algumas das hipótese sugeridas pela análise conduzida nesta seção. Particularmente, usaremos regressões separadas

para cada um dos escritórios, USPTO e EPO, para modelos de entrada, usando dados do PATSTAT entre 2000 a 2019. Como a variável dependente é binária, especialização ou não, especificações do tipo PROBIT e LOGIT são exploradas numa subperiodização em 4 períodos para toda a amostra.

## 4.2 Regressão

O modelo econométrico representado em (3.14) deve ser estimada por algum método aplicável para variáveis binárias, já que queremos avaliar a probabilidade de se entrar ou não numa determinada tecnologia. Novamente:

$$E_{ci,t+1} = \beta_0 + \beta_1 C_{cit} + \beta_2 C_{cit} * ICT_{ct} + \beta_3 \omega_{cit} + \lambda \mathbf{X}_{ct} + \gamma_t + \alpha_c + \kappa_i + \epsilon_{cit},$$

Usando da função *femlm* do pacote *fixest* do R<sup>13</sup>, estima-se modelo probit e logit de entrada com efeitos-fixos e erros clusterizados para a unidade de tempo e de país, a fim de corrigir por padrões em comum nos erros que poderiam gerar um problema de heterocedasticidade, afetando a inferência dos resultados. A [Tabela 3](#) traz os resultados para os modelos de entrada. Os primeiros quatro modelos são estimados por Probit (*Gaussian*), os demais por Logit, como indicado. Como descrito na [seção 3.4](#), os dados estão divididos em quatro sub-períodos entre 2000 a 2019, contudo, como são modelos de entrada, ignoramos o último período, com o valor das variáveis definidos para o período base. No [Apêndice B](#), várias tabelas indicam variações das escolhas tomadas nessa primeira regressão. Os resultados lá expostos serão eventualmente utilizados para análise de robustez. Além disso, a [Tabela 4](#) resume os efeitos marginais médios para o modelo apresentado na [Tabela 3](#).

Como esperado pelo “Princípio do Relacionamento” ([HIDALGO et al., 2018](#)), as regressões indicam que a densidade em torno de uma classe tecnológica aumenta a probabilidade de que uma região passa a ser especializada numa determinada

<sup>13</sup> A página do [CRAN](#) do *fixest* traz algumas vinhetas (*vignettes*) muito úteis. Além disso, a função tem sua própria [documentação](#).

**Tabela 3** – Modelos De Entrada

(a) USPTO

Dependent Variable:	Entrada							
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Gaussian	Gaussian	Gaussian	Gaussian	Logit	Logit	Logit	Logit
<i>Variables</i>								
density	0.0568 (0.0766)	0.3982*** (0.0553)	0.3915*** (0.0552)	0.3977*** (0.0560)	0.5493 (0.7439)	3.648*** (0.6391)	3.562*** (0.6451)	3.721*** (0.6656)
l_compl		0.0956*** (0.0128)	0.0954*** (0.0128)	0.0962*** (0.0128)		1.259*** (0.1281)	1.258*** (0.1277)	1.284*** (0.1287)
l_pib_pc			0.0330** (0.0164)	0.0363** (0.0165)			0.7638** (0.3128)	0.8215** (0.3228)
l_dens_pop			0.0493*** (0.0163)	0.0526*** (0.0163)			0.9178*** (0.2717)	1.007*** (0.2823)
l_compl × eci				-0.0006** (0.0003)				-0.0211** (0.0099)
<i>Fixed-effects</i>								
period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
country	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>								
Observations	227,925	227,925	227,925	227,925	227,925	227,925	227,925	227,925
Squared Correlation	0.04663	0.06737	0.06759	0.06765	0.04648	0.07226	0.07260	0.07281
Pseudo R <sup>2</sup>	0.37449	0.54691	0.54879	0.54931	0.09632	0.12724	0.12789	0.12813
BIC	19,784.5	14,785.4	14,755.3	14,752.6	107,834.6	104,211.7	104,159.8	104,144.2

Clustered (country & ipc\_class) standard-errors in parentheses  
 Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

(b) EPO

Dependent Variable:	Entrada							
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Gaussian	Gaussian	Gaussian	Gaussian	Logit	Logit	Logit	Logit
<i>Variables</i>								
density	0.0697 (0.0887)	0.3156*** (0.0674)	0.3174*** (0.0676)	0.3187*** (0.0682)	0.6153 (0.8614)	2.949*** (0.7422)	2.974*** (0.7548)	2.991*** (0.7682)
l_compl		0.1709*** (0.0269)	0.1710*** (0.0269)	0.1713*** (0.0270)		2.488*** (0.5112)	2.493*** (0.5120)	2.497*** (0.5139)
l_pib_pc			0.0242* (0.0141)	0.0248* (0.0143)			0.5601* (0.3016)	0.5639* (0.3028)
l_dens_pop			0.0220 (0.0202)	0.0219 (0.0202)			0.6034* (0.3553)	0.5976* (0.3571)
l_compl × eci				-0.0002 (0.0002)				-0.0033 (0.0095)
<i>Fixed-effects</i>								
period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
country	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>								
Observations	179,582	179,582	179,582	179,582	179,582	179,582	179,582	179,582
Squared Correlation	0.04073	0.05210	0.05218	0.05218	0.04055	0.05281	0.05285	0.05285
Pseudo R <sup>2</sup>	0.32292	0.41548	0.41612	0.41616	0.08542	0.10349	0.10376	0.10376
BIC	16,917.2	14,788.7	14,797.9	14,809.2	86,049.2	84,386.4	84,385.4	84,397.1

Clustered (country & ipc\_class) standard-errors in parentheses  
 Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Fonte: Elaboração Própria.

tecnologia de forma significativa em praticamente todos os modelos e para os dois escritórios. Isso continua a acontecer mesmo quando mudamos a estrutura de clusters para corrigir os erros padrões. Contudo, quando incluímos efeitos-fixos para classe tecnológica, independente da estrutura de clusters, a variável da densidade se torna negativa e significativa em alguns em modelos Logit (ver [Tabela 13](#), [Tabela 14](#) e [Tabela 15](#)). Um possível motivo desse resultado “anômalo” é que algumas regiões, por alguma característica não observável, são mais propensas a desenvolver tecnologias em uma determinada área, com que faz que estar próximo em termos de capacidades (*capabilities*) na verdade não influencia tanto no processo de diversificação tecnológica. Por exemplo, por motivos culturais<sup>14</sup>, pode ser que um país centre-se mais na produção de patentes em Física do que em Química, apesar de suas capacidades estarem mais relacionadas a uma do que à outra. Portanto, dependendo da escolha do pesquisador, podemos receber resultados completamente distintos. Apesar disso, sob diferentes escolhas e especificações, quando um país possui maior densidade em trono de uma classe tecnológica, é mais provável que ele entre naquela determinada tecnologia.

Em contraste, a principal variável de interesse, laços complementares, apresenta uma robustez marcante a diferentes especificações, sendo positiva e significativa a 1% de nível de confiança em todas as regressões! Isto indica a importância para os países em ter acesso a capacidades tecnológicas através de laços internacionais de cooperação, em especial quando são complementares. Este achado está de acordo com [Balland e Boschma \(2021\)](#) que propuseram o índice utilizado. Mais importante ainda, está de acordo com a literatura evolucionária que reforça a importância da trajetória local para a performance de uma região e como o acesso a capacidades extrarregionais – no caso a “região” é o país – é importante para fugir do *lock-in* que os países e regiões podem se defrontar ([BATHOLT; MALMBERG; MASKELL, 2004](#)). Finalmente, aponta um vetor importante, no sentido de que a literatura ainda está muito centrada em explorar como as capacidades disponíveis localmente são importantes para um país ou região. Apesar dessa abordagem ter uma clara

---

<sup>14</sup> Um passado rico em pensadores da Física, por conta de algum capricho de um governador no passado, alguma característica geográfica que facilite a observação de eventos físicos específicos, *etc.*



**Tabela 4** – Efeitos Marginais Médios

(a) USPTO

	variável	estimativa	erro-padrão	estatística-z	p-valor	conf.inf	conf.sup
1	density	0.229	0.781	0.294	0.769	-1.301	1.76
2	l_compl	0.078	0.279	0.281	0.779	-0.469	0.626
3	l_pib_pc	0.051	0.196	0.258	0.796	-0.334	0.435
4	l_dens_pop	0.062	0.257	0.242	0.809	-0.441	0.565
5	eci	-0.01	0.034	-0.284	0.776	-0.077	0.057

(b) EPO

	variável	estimativa	erro-padrão	estatística-z	p-valor	conf.inf	conf.sup
1	density	0.188	0.69	0.273	0.785	-1.165	1.541
2	l_compl	0.157	0.589	0.267	0.79	-0.998	1.312
3	l_pib_pc	0.036	0.139	0.256	0.798	-0.236	0.307
4	l_dens_pop	0.038	0.147	0.256	0.798	-0.251	0.326
5	eci	-0.001	0.008	-0.171	0.864	-0.018	0.015

Nota: resultados baseados na especificação da [Tabela 3](#) para o Logit completo.

Fonte: Elaboração Própria.

vantagem em termos de abandonar uma visão homogeneizadora, que desconsidera diferenças regionais no desenvolvimento tecnológico dos países, ela ainda falha ao ser muito auto-centrada no aspecto geográfico da proximidade. Por meio de canais de comunicação, comércio e, em geral, de transmissão de informações, é possível que os países encontrem conhecimentos tácitos não presentes localmente. Em específico, os resultados aqui apresentados mostram que por pelo menos uma via isto de fato acontece, através da co-aplicação em patentes por instituições em diferentes locais do globo. Uma longa tradição na literatura de Cadeias Globais de Valor (GVC)<sup>15</sup> já apontava nesse sentido ([YEUNG, 2021](#)). Os resultados empíricos deste trabalho suportam a ideia de que tanto a literatura de GVC quanto a evolucionária tem a ganhar ao se inspirarem mutuamente ([BOSCHMA, 2022](#)).

Como modelos de índice são não lineares, o impacto *linear* das variáveis no modelo de variável latente, ou seja, os coeficientes, não possuem uma interpretação

<sup>15</sup> *Global Value Chains*.

direta. Apesar de o sinal ser mantido depois da transformação, já que as duas funções de densidade de probabilidade tanto do Probit quanto do Logit são estritamente crescentes, a magnitude tem que ser recalculada de acordo com as observações, já que o efeito parcial sobre a probabilidade de ativação da variável dependente *depende* do ponto em que estamos da distribuição – ver [seção 3.2](#) para mais detalhes. A [Tabela 4](#) traz a média dos efeitos marginais de cada variável em cada observação (veja que não é o efeito da média, o que teríamos utilizando as médias de cada variável e daí calculando o efeito marginal, mas a média do efeito). A análise a seguir se fia do resultado das estimativas, mas com cautela, já que nenhuma das estimativas se mostrou significativa – até um alto grau de confiança.

Em relação à significância econômica desses resultados, a [Tabela 4](#) traz os efeitos marginais médios. Como exposto na [seção 4.2](#), esta métrica é o cálculo do efeito marginal médio para todas as observações, tal como em (3.11). Os efeitos representados consideram o modelo (8) completo do Logit na [Tabela 3](#). Primeiro, porque, novamente como argumentado na [seção 4.2](#), em modelos de índice com efeitos-fixos, a depender da distribuição, ou função, que assumimos, não é possível eliminar esses efeitos-fixos e não observáveis como em modelos de painel simples. Isso acontece para modelos Probit, resultando em um problema de parâmetros incidentais, gerando vieses nas estimativas. Contudo, conseguimos eliminar esses efeitos não observáveis em modelos Logit. Além disso, como o grau de explicação do modelo, indicado pelo Pseudo  $R^2$ , e o critério de informação (BIC) não se altera muito em relação aos demais, por conta da utilidade em poder avaliar todas as variáveis ao mesmo tempo, ficamos com o modelo Logit completo.

Contudo, as estimativas dos efeitos marginais podem passar uma mensagem enganadora. Pelo que diz a tabela, aumentar em 10% o índice de complementariedade leva a um aumento, em média, de 0,78% no USPTO e 1,57% no EPO na probabilidade de entrar em uma determinada classe tecnológica. Por outro lado, um aumento de 0,1 na variável de densidade, ou cerca de 10% da sua variação possível total, está associada a um aumento na probabilidade de entrada numa classe tecnológica de 2,29% no USPTO e 1,88% no EPO. Contudo, como a [Tabela 2](#) mostra, a variação

para o índice de densidade é muito menor, variando de 0 a 1 apenas, enquanto o índice (do log) de complementariedade varia de cerca de 6 a 9. Assim, uma avaliação “global” tem que levar em conta as diferenças quanto à variabilidade possível de cada um. Numa regra simples, poderíamos simplesmente multiplicar por 3 o coeficiente de complementariedade para comparar com o de densidade, já que o primeiro pode variar quase três vezes mais. Como vemos, isto torna o impacto econômico de ambas variáveis muito similares no USPTO e dá uma vantagem que a princípio não daríamos para laços complementares no EPO. É difícil, porém, auferir um significado econômico preciso a qualquer uma dessas variáveis, já que são índices cujos valores não possuem interpretação imediata.

Parte da análise exploratória conduzida na [seção 4.1](#) já indicava os resultados que obtemos para as variáveis de controle. Tanto PIB per capita quanto densidade populacional nem sempre são significativas. Isto é coerente com a falta de padrão que pudemos observar nas redes coloridas de acordo com essas mesmas variáveis. Em geral, PIB per capita parece ter um efeito positivo sobre a entrada de novas classes tecnológicas, quer dizer, países mais ricos conseguem entrar mais facilmente em algumas classes tecnológicas, reforçando a ideia de divergência econômica internacional. Por sua vez, densidade populacional também é robusta a diferentes especificações em termos do sinal do seu coeficiente sendo, em geral, positiva. Em especial, é curioso destacar que a significância dos coeficientes é sensível a diferentes estruturas de clusters. A mais conservadora é a que escolhemos destacar, com erros clusterizados conforme o período e o país. Nas demais, a significância dos coeficientes em geral aumenta, principalmente nos modelos Logit. PIB per capita é significativo com mais frequência do que densidade populacional, sugerindo que padrões de divergência internacional são mais influentes do que economias de aglomeração em explicar o comportamento de inovação tecnológica dos países no período analisado. Finalmente, é mais comum encontrarmos coeficientes significativos para o USPTO do que no EPO, apontando para uma diferença regional no perfil das aplicações em cada escritório.

A magnitude do efeito de ambas as variáveis é similar, sendo um pouco maior nos resultados para o USPTO do que no EPO. Dobrar o PIB per capita

ou a densidade populacional está associado, em média, a um aumento de 5 a 6% no USPTO e cerca de 4% no EPO na probabilidade de entrar numa determinada classe tecnológica, mantendo os demais fatores fixos. Contudo, a não significância dos coeficientes na variável latente e dos efeitos marginais coloca em cheque a validade desses efeitos. De qualquer forma, parece ter uma associação positiva ao aumento de cada uma dessas variáveis, mas cuja significância econômica não é tão clara – não é trivial dobrar o PIB per capita.

O termo de interação entre complementariedade e complexidade tecnológica é negativo e significativo em explicar a entrada em novos setores tecnológicos, confirmando a hipótese inicial deste trabalho. Isto indica que países menos complexos tem um ganho maior ao explorar laços internacionais de colaboração. Contudo, curiosamente, se não incluímos efeitos-fixos de países chegamos a conclusões radicalmente distintas. Além disso, vale destacar que o ajuste da regressão aos dados, o  $R^2$ , fica praticamente inalterado quando incluímos outras variáveis além da de densidade e complementariedade tecnológica.

Este resultado indica que estabelecer laços com outros países cuja estrutura tecnológica é complementar à local pode ser uma alternativa para romper as desigualdades persistentes entre os países. Contudo, a análise das redes de colaboração internacional sugere que devemos tomar essa conclusão com um grão de sal. Se os países desenvolvidos são os que mais exploram essa rede, é insuficiente que haja um impacto desproporcionalmente maior para países menos complexos: eles simplesmente não conseguem estabelecer tantas colaborações. Talvez essa interpretação ofereça uma explicação do porquê quando não levamos em conta efeitos-fixos dos países a conclusão é a oposta. Países complexos exploram tão bem essa rede de colaboração internacional ao ponto que conseguem captar desproporcionalmente mais seus benefícios. Infelizmente, quanto à significância econômica do respectivo coeficiente, ter um termo de interação complica a análise, mas, como seria de esperar, quão mais complexo um país, nesta especificação, menor a probabilidade de entrar numa determinada classe tecnológica.

De maneira geral, a regressão estatística corrobora nossa hipótese inicial: laços complementares auxiliam na entrada de novos setores tecnológicos. Além disso, esse efeito é maior para países menos complexos, atuando, portanto, como uma válvula de escape da ausência de capacidades locais. Isso acontece mesmo controlando por efeitos de convergência (PIB per capita) e economias de aglomeração (densidade populacional). Mais importante ainda, é um efeito mais robusto do que o de proximidade local, densidade, a um determinado produto. Ou seja, ter conhecimentos locais similares a uma determinada classe tecnológica torna o processo de entrada mais fácil, contudo, embora essa robustez não se mantenha no diferencial entre países complexos e não complexos, é um efeito estatístico e economicamente menos significativo do que o de laços complementares.

## 5 Conclusão

O que diferencia países ricos de países pobres? A abordagem de complexidade econômica ofereceu uma explicação simples e potente: países ricos são especializados na produção de produtos complexos. Ao abandonar a suposição de homogeneidade dos produtos [Hidalgo et al. \(2007\)](#) propuseram uma nova abordagem empírica que parte de uma simples pressuposição: a produção de um item está ligada a um conhecimento tácito difícil de ser transmitido específico àquele produto. Assim, países que produzissem muitos produtos e produtos que ninguém mais soubesse fazer, raros ou não ubíquos, seriam países complexos. Usando de técnicas de análise de rede, [Hidalgo et al. \(2007\)](#) propuseram o índice de complexidade econômica e mostraram que países complexos são também aqueles mais ricos.

No fundo, o que eles propuseram foi uma flexibilização daquela antiga hipótese: ao invés de homogeneidade dos produtos, por que não heterogeneidade? Esta flexibilidade deu luz a uma série de estudos posteriores que desaguaram no que se chamou de “princípio do relacionamento” ([HIDALGO et al., 2018](#)). Basicamente, é à ideia de que países, regiões, instituições etc. possuem maior facilidade em entrar (passar a ser especializado) em atividades (produtos, tecnologias, ocupações etc.) que sejam similares à estrutura produtivo-tecnológica local. Isto é, é mais fácil para um país que produz uvas passar a produzir vinho do que vir a produzir aviões. Isto advém daquela concepção de que cada produto/tecnologia está associada a um conjunto específico de conhecimentos úteis à produção, *capacidades*. Assim, a diversificação produtivo-tecnológica está localmente estrangida.

Por um lado, esta visão oferece uma alternativa a soluções *one size fits all*. Isto é, cada região se defronta com capacidades distintas, assim, seu desenvolvimento deve levar em conta essas vantagens. Por outro lado, o que é vantagem em alguns produtos, é desvantagem em outro. Se esta visão “diminui” – relativamente a uma concepção de homogeneidade – o custo para se entrar em atividades próximos, ela

“aumenta” o custo para atividades distantes. Em especial, esta dificuldade é mais manifesta para países periféricos, que possuem menos capacidades e, assim, possuem um conjunto possível de diversificação muito mais limitado. Essa visão dinâmica de dependência de trajetória indica que regiões mais pobres podem ter uma maior tendência ao *lock-in* pela falta de capacidades disponíveis localmente, estabelecendo uma relação persistente de desigualdade entre os países.

Uma possível solução é no exterior: estabelecer laços com outras regiões pode ser uma forma de complementar a ausência de capacidades localmente. Como argumentado no [Capítulo 2](#), a evidência empírica aponta que diversos canais de transmissão de conhecimentos são importantes para a performance tecnológica local. Contudo, o ganho em se colaborar com outros lugares não é linear. Deve ser muito mais proveitoso se conectar com regiões que possuam uma estrutura complementar à local. Assim, o conhecimento que se acessa não é nem parecido demais, adicionando pouco, nem distante demais, sendo incompreensível. O que se almeja é uma distância cognitiva ótima que maximize o aprendizado ([NOOTEBOOM, 2000](#)).

A contribuição deste trabalho está em oferecer uma evidência empírica inédita de que estabelecer laços complementares com outros países auxilia na diversificação tecnológica local. Aproveitando-se do índice proposto por [Balland e Boschma \(2021\)](#), expande-se seus achados ao usar de dados internacionais de patentes para dois dos maiores escritórios em termos de aplicação, o norte-americano (USPTO) e o europeu (EPO). Na posse destes dados, testamos se este índice de laços com países cuja estrutura tecnológica é complementar à local está associada a uma maior probabilidade de entrada num novo setor tecnológico. Confirmamos essa hipótese para o período de 2000-2019 dividindo em 4 sub-períodos não sobrepostos através de modelos de índice Probit e Logit com efeitos-fixos não observáveis adicionando erros clusterizados, a fim de controlar pela heterogeneidade entre grupos (períodos, países e classes tecnológicas). Neste sentido, o trabalho oferece também uma contribuição metodológica, já que [Balland e Boschma \(2021\)](#) não levam em conta efeitos-fixos entre grupos usando de uma especificação com dados empilhados e Probit/Logit simples, apesar de [Freitas \(2019\)](#) por exemplo já ter usado de efeitos-fixos mas

focando no “princípio do relacionamento” para o Brasil.

Além de encontrarmos que o efeito de laços complementares é robusto a diferentes especificações, mais do que proximidade da estrutura local a um setor tecnológico, observamos que países menos complexos tem um maior ganho em estabelecer esse tipo de conexão. Isso está de acordo com a ideia que países menos desenvolvidos possuem menos capacidades localmente, assim, colaborar com outros países é uma forma de preencher esta ausência. Entretanto, apenas encontramos esse efeito após controlar por fatores não observáveis dos países, sendo na verdade o efeito contrário antes disso – países mais complexos se beneficiando mais de laços complementares. Possivelmente a explicação para este achado está naquilo que observamos na análise exploratória da rede de colaborações internacionais. Países mais desenvolvidos – os mais complexos – são centrais nessa rede, estabelecendo uma quantidade elevada de colaborações com outros países e, assim, sendo potencialmente os maiores beneficiários das vantagens de participar desta rede de colaboração. Portanto, enquanto países subdesenvolvidos possam ter um ganho maior em colaborar com outros países, raramente eles o conseguem, o que coloca em cheque o papel desta rede em mitigar desigualdades persistentes entre os países.

De fato, uma das grandes motivações para avaliar o impacto de conexões com outras regiões advém da limitação que o olhar para dentro da localidade oferecia. Países menos complexos por possuírem menos capacidades se encontram numa difícil situação em que, para conseguir novas capacidades, precisam dominar novos setores tecnológicos, mas para isso precisam antes ter as capacidades! A “curva S” resume esta ideia de uma maneira simples. Países mais complexos possuem capacidades que lhes permitem acessar com maior facilidade produtos e tecnologias complexas. Assim, como estas atividades estão associadas a um nível de desenvolvimento maior, países desenvolvidos se veem numa posição privilegiada para manter sua competitividade tecnológica. Este trabalho encontra uma evidência similar, apontando que países mais complexos tem fácil acesso a tecnologias complexas, o que se observa seja para o USPTO quanto para o EPO. Mas, enquanto aquela visão para dentro dá lugar a uma para fora, a fim de romper o ciclo do subdesenvolvimento, o que se vê é que são



também os países mais complexos são também os que mais colaboram com outros países. Assim, o que se monta é um ciclo auto-perpetuador, em que, ao invés de válvula de escape, uma rede internacional de colaboração (em patentes) pode ser ainda outra forma de expandir, ou ao menos manter, as desigualdades entre os países.

Estas contribuições inéditas à literatura lançam desafios e novas questões a serem exploradas em trabalhos futuros. Como seria de se esperar, ainda existe muito espaço para testar distintas especificações, mudando do método às variáveis, principalmente as de controle, que permitiriam não só uma análise de robustez dos resultados mas discussões distintas. De fato, apesar da novidade que esse trabalho traz, ele deixa de responder muitas outras questões que foram trazidas ainda no [Capítulo 2](#). Laços internacionais não são a única maneira de instituições se conectarem. Muitas vezes, elas podem se conectar a instituições nacionais com grande proveito ([SCALERA; PERRI; HANNIGAN, 2018](#); [SANTOALHA, 2019](#), por exemplo). Em especial, conectar-se com instituições numa mesma unidade geográfica pode ser uma maneira de “cortar caminho”. Ao invés de engajar num processo custoso de internacionalização, firmas, por exemplo, podem colaborar com outros parceiros nacionais já internacionalizados e alavancar suas atividades a partir daí. Mais do que simplesmente a origem, o tipo de conexão é um importante aspecto deixado de lado neste trabalho ([TURKINA; ASSCHE, 2018](#), por exemplo). Nem todo laço tem a mesma influência sob a performance tecnológica local. Embora este trabalho analise a influência de uma maneira distinta, complementar, de colaborar com outra região, o impacto de laços extrarregionais deve variar de acordo com outros fatores, como a fonte – *e.g.* comércio ou co-patenteação – e o tipo de instituição – *e.g.* grande ou pequena. Finalmente, ainda fica a responder se as conclusões obtidas aqui e em [Balland e Boschma \(2021\)](#) se mantêm para outros lugares do globo, pois, enquanto essa análise se propôs internacional, primeiro, deixa de avaliar outros escritórios de patente além dos Estados Unidos e Europa e, segundo, avalia isso ao agregado nível nacional. [Balland e Boschma \(2021\)](#) foca apenas na Europa, mas também com isso pôde observar a um nível de regionalização mais desagregado. Trabalhos futuros podem se beneficiar de ambas as perspectivas ao ampliar (e talvez corrigir)

as conclusões retiradas para o USPTO e EPO.

Contudo, seria um erro pensar que os desafios são apenas empíricos, há também ainda importantes contribuições teóricas que este trabalho não se ocupa mas que futuras investigações podem ter como objetivo. O ponto de chegada desta dissertação é um de intersecção entre literaturas. A extensão de métodos desenvolvidos em uma nem sempre é justificável na outra e facilmente pode esbarrar em complicações teóricas em uma das duas. Trabalhos futuros devem explorar os canais pelos quais a teoria de uma linha de pesquisa conversa com a outra.

Em termos de política, esse trabalho mantém conclusões que a literatura já vem ressaltando. É importante que políticas de desenvolvimento econômico e tecnológico local não sigam uma abordagem de *one size fits all*, o que é evidenciado pela confirmação do “Princípio do Relacionamento”. Isto é, que leve em conta as capacidades e limitações daquele país ou região para propor políticas de diversificação tecnológica. Contudo, mais do que isso, este trabalho aponta para um importante vetor que tem sido relativamente negligenciado na literatura que é a importância de se fazer laços com outras regiões. Mais do que apenas se conectar com países desenvolvidos tecnologicamente, os formuladores de política têm de buscar identificar quais são os países e regiões que tenham uma estrutura complementar à local, a partir daí incentivando essas conexões com maior potencial. Em especial, isto vale mais ainda para países periféricos, que frequentemente se veem circundados pela falta de oportunidade local, assim, suprir a ausência de capacidades por meio do contato com outros países pode servir como um atalho ao desenvolvimento.

## Referências

AGÉNOR, P.-R. Caught in the Middle? The Economics of Middle-Income Traps. *Journal of Economic Surveys*, v. 31, n. 3, p. 771–791, 2017. ISSN 1467-6419. \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/joes.12175>. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/joes.12175>. Citado na página 33.

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, dez. 2018. ISSN 2041-1723. Disponível em: <http://www.nature.com/articles/s41467-018-03740-9>. Citado na página 43.

ARTHUR, W. B. Competing Technologies, Increasing Returns, and Lock-In by Historical Events. *The Economic Journal*, v. 99, n. 394, p. 116, mar. 1989. ISSN 00130133. Disponível em: <https://academic.oup.com/ej/article/99/394/116-131/5188212>. Citado na página 13.

ASCANI, A. et al. Global networks, local specialisation and regional patterns of innovation. *Research Policy*, v. 49, n. 8, p. 104031, out. 2020. ISSN 00487333. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0048733320301104>. Citado 2 vezes nas páginas 34 e 37.

BAHAR, D.; HAUSMANN, R.; HIDALGO, C. A. Neighbors and the evolution of the comparative advantage of nations: Evidence of international knowledge diffusion? *Journal of International Economics*, v. 92, n. 1, p. 111–123, jan. 2014. ISSN 00221996. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0022199613001098>. Citado na página 37.

BALASSA, B. Trade Liberalisation and “Revealed” Comparative Advantage<sup>1</sup>. *The Manchester School*, v. 33, n. 2, p. 99–123, 1965. \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9957.1965.tb00050.x>. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9957.1965.tb00050.x>. Citado 2 vezes nas páginas 20 e 30.

BALLAND, P.-A.; BOSCHMA, R. Complementary interregional linkages and Smart Specialisation: an empirical study on European regions. *Regional Studies*, v. 55, n. 6, p. 1059–1070, jun. 2021. ISSN 0034-3404, 1360-0591. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/00343404.2020.1861240>. Citado 15 vezes nas páginas 6, 15, 37, 38, 39, 42, 44, 47, 52, 53, 56, 58, 87, 94 e 96.

BALLAND, P.-A. et al. Smart specialization policy in the European Union: relatedness, knowledge complexity and regional diversification. *Regional Studies*, v. 53, n. 9, p. 1252–1268, set. 2019. ISSN 0034-3404, 1360-0591. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/00343404.2018.1437900>. Citado 4 vezes nas páginas 27, 28, 29 e 52.

- BALLAND, P.-A. et al. The new paradigm of economic complexity. *Research Policy*, v. 51, n. 3, 2022. Citado 3 vezes nas páginas 19, 20 e 23.
- BALLAND, P.-A. et al. Complex Economic Activities Concentrate in Large Cities. p. 10, 2020. Citado 2 vezes nas páginas 13 e 46.
- BALLAND, P.-A.; RIGBY, D. The Geography of Complex Knowledge. *Economic Geography*, v. 93, n. 1, p. 1–23, jan. 2017. ISSN 0013-0095, 1944-8287. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/00130095.2016.1205947>>. Citado 3 vezes nas páginas 33, 46 e 77.
- BARABASI, A.-L. *Network science*. Cambridge: Cambridge University Press, 2016. ISBN 978-1-107-07626-6 1-107-07626-9. Disponível em: <<http://barabasi.com/networksciencebook/>>. Citado 5 vezes nas páginas 42, 53, 54, 55 e 66.
- BARZOTTO, M. et al. Enhancing innovative capabilities in lagging regions: an extra-regional collaborative approach to RIS3. *Cambridge Journal of Regions, Economy and Society*, v. 12, n. 2, p. 213–232, jun. 2019. ISSN 1752-1378, 1752-1386. Disponível em: <<https://academic.oup.com/cjres/article/12/2/213/5479411>>. Citado 2 vezes nas páginas 37 e 40.
- BATHELT, H.; MALMBERG, A.; MASKELL, P. Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, v. 28, n. 1, p. 31–56, fev. 2004. ISSN 0309-1325, 1477-0288. Disponível em: <<http://journals.sagepub.com/doi/10.1191/0309132504ph469oa>>. Citado 2 vezes nas páginas 14 e 87.
- BOSCHMA, R. Proximity and Innovation: A Critical Assessment. *Regional Studies*, v. 39, n. 1, p. 61–74, fev. 2005. ISSN 0034-3404, 1360-0591. Disponível em: <<http://www.tandfonline.com/doi/abs/10.1080/0034340052000320887>>. Citado 2 vezes nas páginas 12 e 36.
- BOSCHMA, R. Relatedness as driver of regional diversification: a research agenda. *Regional Studies*, v. 51, n. 3, p. 351–364, mar. 2017. ISSN 0034-3404, 1360-0591. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/00343404.2016.1254767>>. Citado na página 26.
- BOSCHMA, R. Designing Smart Specialization Policy: relatedness, unrelatedness, or what? p. 32, 2021. Citado na página 29.
- BOSCHMA, R. Global value chains from an evolutionary economic geography perspective: a research agenda. *Area Development and Policy*, v. 7, n. 2, p. 123–146, abr. 2022. ISSN 2379-2949, 2379-2957. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/23792949.2022.2040371>>. Citado 3 vezes nas páginas 15, 26 e 88.
- 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*, v. 24, n. 1, p. 223–250,

fev. 2015. ISSN 0960-6491. Disponível em: <<https://doi.org/10.1093/icc/dtu012>>. Citado 2 vezes nas páginas 27 e 57.

BOSCHMA, R.; IAMMARINO, S. Related Variety, Trade Linkages, and Regional Growth in Italy: ECONOMIC GEOGRAPHY. *Economic Geography*, v. 85, n. 3, p. 289–311, abr. 2009. ISSN 00130095. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1111/j.1944-8287.2009.01034.x>>. Citado 2 vezes nas páginas 34 e 37.

BRITTO, G. et al. The great divide: Economic complexity and development paths in Brazil and the Republic of Korea. *CEPAL Review*, v. 2019, n. 127, p. 191–213, out. 2019. ISSN 16840348. Disponível em: <<https://www.un-ilibrary.org/content/journals/16840348/2019/127/10>>. Citado na página 32.

CAPELLO, R.; CARAGLIU, A. Proximities and the Intensity of Scientific Relations: Synergies and Nonlinearities. *International Regional Science Review*, v. 41, n. 1, p. 7–44, jan. 2018. ISSN 0160-0176, 1552-6925. Disponível em: <<http://journals.sagepub.com/doi/10.1177/0160017615626985>>. Citado na página 37.

ERIKSSON, R. H.; LENGYEL, B. Co-worker Networks and Agglomeration Externalities. *Economic Geography*, v. 95, n. 1, p. 65–89, jan. 2019. ISSN 0013-0095, 1944-8287. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/00130095.2018.1498741>>. Citado na página 34.

FELDMAN, M. P.; FLORIDA, R. The Geographic Sources of Innovation: Technological Infrastructure and Product Innovation in the United States. *Annals of the Association of American Geographers*, v. 84, n. 2, p. 210–229, jun. 1994. ISSN 0004-5608, 1467-8306. Disponível em: <<http://www.tandfonline.com/doi/abs/10.1111/j.1467-8306.1994.tb01735.x>>. Citado na página 12.

FLEMING, L.; SORENSON, O. Technology as a complex adaptive system: evidence from patent data. *Research Policy*, v. 30, n. 7, p. 1019–1039, ago. 2001. ISSN 00487333. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0048733300001359>>. Citado na página 11.

FREITAS, E. E. Indústrias relacionadas, complexidade econômica e diversificação regional: uma aplicação para microrregiões brasileiras. p. 153, 2019. Citado 2 vezes nas páginas 27 e 94.

FRENKEN, K.; OORT, F. V.; VERBURG, T. Related Variety, Unrelated Variety and Regional Economic Growth. *Regional Studies*, v. 41, n. 5, p. 685–697, jul. 2007. ISSN 0034-3404, 1360-0591. Disponível em: <<http://www.tandfonline.com/doi/abs/10.1080/00343400601120296>>. Citado na página 24.

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*, v. 38, n. 2, p. 219–236, jun. 2018. ISSN

1809-4538, 0101-3157. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0101-31572018000200219&lng=en&tlng=en](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-31572018000200219&lng=en&tlng=en)>. Citado na página 23.

GREENE, W. H. *Econometric Analysis*. Fifth. Pearson Education, 2003. ISBN 0-13-066189-9. Disponível em: <<http://pages.stern.nyu.edu/~wgreene/Text/econometricanalysis.htm>>. Citado na página 51.

HALL, B. H.; JAFFE, A. B.; TRAJTENBERG, M. Working Paper, *The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools*. National Bureau of Economic Research, 2001. (Working Paper Series). Disponível em: <<https://www.nber.org/papers/w8498>>. Citado na página 35.

HARTMANN, D.; BEZERRA, M.; PINHEIRO, F. L. Identifying Smart Strategies for Economic Diversification and Inclusive Growth in Developing Economies. The Case of Paraguay. *SSRN Electronic Journal*, 2019. ISSN 1556-5068. Disponível em: <<https://www.ssrn.com/abstract=3346790>>. Citado na página 43.

HARTMANN, D. et al. Why did some countries catch-up, while others got stuck in the middle? Stages of productive sophistication and smart industrial policies. p. 30, 2020. Citado 3 vezes nas páginas 29, 32 e 33.

HAUSMANN, R.; CHAUVIN, J. Discovering Paths for Export Diversification in Rwanda. p. 41, 2015. Citado na página 43.

HAUSMANN, R. et al. *The Atlas of economic complexity: mapping paths to prosperity*. Cambridge, Mass.: Center for International Development, Harvard University, 2014. OCLC: 781855824. Disponível em: <<http://atlas.media.mit.edu/book/>>. Citado 3 vezes nas páginas 19, 22 e 28.

HAUSMANN, R.; RODRIK, D. Economic development as self-discovery. *Journal of Development Economics*, v. 72, n. 2, p. 603–633, dez. 2003. ISSN 03043878. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S030438780300124X>>. Citado na página 18.

HAUSMANN, R.; SANTOS, M. A.; OBACH, J. Appraising the Economic Potential of Panama: Policy Recommendations for Sustainable and Inclusive Growth. *CID Faculty Working Paper*, 2017. Citado na página 43.

HIDALGO, C. A. et al. The Principle of Relatedness. In: MORALES, A. J. et al. (Ed.). *Unifying Themes in Complex Systems IX*. Cham: Springer International Publishing, 2018. p. 451–457. ISBN 978-3-319-96660-1 978-3-319-96661-8. Series Title: Springer Proceedings in Complexity. Disponível em: <[http://link.springer.com/10.1007/978-3-319-96661-8\\_46](http://link.springer.com/10.1007/978-3-319-96661-8_46)>. Citado 6 vezes nas páginas 25, 26, 28, 52, 85 e 93.

HIDALGO, C. A.; HAUSMANN, R. The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, v. 106, n. 26, p. 10570–10575, jun. 2009. ISSN 0027-8424, 1091-6490. Disponível em: <<https://pnas.org/doi/full/10.1073/pnas.0900943106>>. Citado 7 vezes nas páginas 18, 19, 22, 23, 33, 46 e 77.

- HIDALGO, C. A. et al. The Product Space Conditions the Development of Nations. *Science*, v. 317, n. 5837, p. 482–487, jul. 2007. ISSN 0036-8075, 1095-9203. Disponível em: <<https://www.science.org/doi/10.1126/science.1144581>>. Citado 9 vezes nas páginas 17, 18, 20, 23, 24, 26, 45, 52 e 93.
- JAFFE, A. B.; TRAJTENBERG, M.; HENDERSON, R. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. p. 23, 1993. Citado na página 12.
- JUN, B. et al. Bilateral relatedness: knowledge diffusion and the evolution of bilateral trade. *Journal of Evolutionary Economics*, v. 30, n. 2, p. 247–277, abr. 2020. ISSN 0936-9937, 1432-1386. Disponível em: <<http://link.springer.com/10.1007/s00191-019-00638-7>>. Citado 2 vezes nas páginas 37 e 38.
- LI, Y.; NEFFKE, F. *Relatedness in regional development: in search of the right specification*. arXiv, 2022. Number: arXiv:2205.02942 arXiv:2205.02942 [physics]. Disponível em: <<http://arxiv.org/abs/2205.02942>>. Citado na página 28.
- MARROCU, E.; PACI, R.; USAI, S. Proximity, networking and knowledge production in Europe: What lessons for innovation policy? *Technological Forecasting and Social Change*, v. 80, n. 8, p. 1484–1498, out. 2013. ISSN 00401625. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0040162513000528>>. Citado na página 37.
- MARTIN, R.; SUNLEY, P. Path dependence and regional economic evolution. *Journal of Economic Geography*, v. 6, n. 4, p. 395–437, ago. 2006. ISSN 1468-2710, 1468-2702. Disponível em: <<http://academic.oup.com/joeg/article/6/4/395/1014440/Path-dependence-and-regional-economic-evolution>>. Citado na página 12.
- MIGUELEZ, E.; MORENO, R. Relatedness, external linkages and regional innovation in Europe. *Regional Studies*, v. 52, n. 5, p. 688–701, maio 2018. ISSN 0034-3404, 1360-0591. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/00343404.2017.1360478>>. Citado na página 37.
- MYRDAL, G. *Economic Theory and Under-developed Regions*. G. Duckworth, 1957. (Economic Theory and Under-developed Regions). Disponível em: <<https://books.google.com.br/books?id=MnXgAAAAMAAJ>>. Citado na página 14.
- NEFFKE, F. et al. The Dynamics of Agglomeration Externalities along the Life Cycle of Industries. *Regional Studies*, v. 45, n. 1, p. 49–65, jan. 2011. ISSN 0034-3404, 1360-0591. Disponível em: <<http://www.tandfonline.com/doi/full/10.1080/00343401003596307>>. Citado na página 27.
- NELSON, R. R.; WINTER, S. G. *An evolutionary theory of economic change*. digitally reprinted. Cambridge, Mass.: The Belknap Press of Harvard Univ. Press, 1982. ISBN 978-0-674-27228-6. Citado na página 12.

- NETO, V. J. d. S. et al. The global geography of digital platforms: towards platforms international locational determinants. *Textos para Discussão Cedeplar-UFMG*, nov. 2022. Number: 650 Publisher: Cedeplar, Universidade Federal de Minas Gerais. Disponível em: <<https://ideas.repec.org/p/cdp/texdis/td650.html>>. Citado na página 72.
- NONI, I. D.; ORSI, L.; BELUSSI, F. The role of collaborative networks in supporting the innovation performances of lagging-behind European regions. *Research Policy*, v. 47, n. 1, p. 1–13, fev. 2018. ISSN 00487333. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0048733317301567>>. Citado 2 vezes nas páginas 37 e 40.
- NOOTEBOOM, B. Learning by Interaction: Absorptive Capacity, Cognitive Distance and Governance. *Journal of Management and Governance*, v. 4, n. 1, p. 69–92, mar. 2000. ISSN 1572-963X. Disponível em: <<https://doi.org/10.1023/A:1009941416749>>. Citado 3 vezes nas páginas 15, 39 e 94.
- PENROSE, E. *The Theory of the Growth of the Firm*. [S.l.]: OUP Oxford, 1959. Google-Books-ID: 85FLx2NQlaoC. ISBN 978-0-19-162342-4. Citado na página 41.
- PINHEIRO, F. L. et al. Shooting High or Low: Do Countries Benefit from Entering Unrelated Activities? p. 43, 2018. Citado 4 vezes nas páginas 29, 30, 32 e 33.
- PINHEIRO, F. L. et al. The Dark Side of the Geography of Innovation: Relatedness, Complexity, and Regional Inequality in Europe. p. 47, 2022. Citado 3 vezes nas páginas 30, 32 e 46.
- RIBEIRO, L. C. et al. Growth patterns of the network of international collaboration in science. *Scientometrics*, v. 114, n. 1, p. 159–179, jan. 2018. ISSN 1588-2861. Disponível em: <<https://doi.org/10.1007/s11192-017-2573-x>>. Citado 3 vezes nas páginas 42, 57 e 67.
- RIGBY, D. L. Technological Relatedness and Knowledge Space: Entry and Exit of US Cities from Patent Classes. *Regional Studies*, v. 49, n. 11, p. 1922–1937, nov. 2015. ISSN 0034-3404, 1360-0591. Disponível em: <<http://www.tandfonline.com/doi/full/10.1080/00343404.2013.854878>>. Citado na página 57.
- RIGBY, D. L. et al. Do EU regions benefit from Smart Specialisation principles? *Regional Studies*, p. 1–16, mar. 2022. ISSN 0034-3404, 1360-0591. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/00343404.2022.2032628>>. Citado 2 vezes nas páginas 33 e 56.
- ROMER, P. M. Endogenous Technological Change. *JOURNAL OF POLITICAL ECONOMY*, p. 32, 1990. Citado na página 17.
- ROMERO, J. P. et al. Economic complexity and regional economic development: evidence from Brazil. p. 22, 2021. Citado na página 43.



- SANTOALHA, A. Technological diversification and Smart Specialisation: the role of cooperation. *Regional Studies*, v. 53, n. 9, p. 1269–1283, set. 2019. ISSN 0034-3404. Publisher: Routledge \_eprint: <https://doi.org/10.1080/00343404.2018.1530753>. Disponível em: <https://doi.org/10.1080/00343404.2018.1530753>. Citado 5 vezes nas páginas 37, 41, 47, 58 e 96.
- SCALERA, V. G.; PERRI, A.; HANNIGAN, T. J. Knowledge connectedness within and across home country borders: Spatial heterogeneity and the technological scope of firm innovations. *Journal of International Business Studies*, v. 49, n. 8, p. 990–1009, out. 2018. ISSN 1478-6990. Disponível em: <https://doi.org/10.1057/s41267-017-0109-5>. Citado 4 vezes nas páginas 35, 41, 42 e 96.
- STOJKOSKI, V.; KOCH, P.; HIDALGO, C. A. *Multidimensional Economic Complexity: How the Geography of Trade, Technology, and Research Explain Inclusive Green Growth*. arXiv, 2022. ArXiv:2209.08382 [cond-mat, q-fin]. Disponível em: <http://arxiv.org/abs/2209.08382>. Citado na página 57.
- TORRES, G. S. Complexidade econômica: uma proposta metodológica para identificação de produtos estratégicos. p. 143, 2019. Citado na página 43.
- TURKINA, E.; ASSCHE, A. V. Global connectedness and local innovation in industrial clusters. *Journal of International Business Studies*, v. 49, n. 6, p. 706–728, ago. 2018. ISSN 0047-2506, 1478-6990. Disponível em: <http://link.springer.com/10.1057/s41267-018-0153-9>. Citado 3 vezes nas páginas 35, 36 e 96.
- WAGNER, C. S.; LEYDESDORFF, L. Network structure, self-organization, and the growth of international collaboration in science. *Research Policy*, v. 34, n. 10, p. 1608–1618, dez. 2005. ISSN 0048-7333. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0048733305001745>. Citado na página 42.
- WOOLDRIDGE, J. M. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, 2010. ISBN 978-0-262-23258-6. Disponível em: <https://www.jstor.org/stable/j.ctt5hhcfr>. Citado 2 vezes nas páginas 48 e 50.
- YEUNG, H. W.-c. Regional worlds: from related variety in regional diversification to strategic coupling in global production networks. *Regional Studies*, v. 55, n. 6, p. 989–1010, jun. 2021. ISSN 0034-3404, 1360-0591. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/00343404.2020.1857719>. Citado 3 vezes nas páginas 15, 26 e 88.
- ZHU, S.; HE, C.; ZHOU, Y. How to jump further and catch up? Path-breaking in an uneven industry space. *Journal of Economic Geography*, p. lbw047, jan. 2017. ISSN 1468-2702, 1468-2710. Disponível em: <https://academic.oup.com/joeg/article-lookup/doi/10.1093/jeg/lbw047>. Citado na página 38.

# Anexos

## ANEXO A – Dados

Este anexo complementa a [seção 4.1](#) expondo os dados utilizados para a construção das redes de colaborações internacionais.

**Tabela 5** – Rede Completa de Colaborações (USPTO)

	País	Nome	Número de Links	Países Distintos
0	US	United States	286501	184
1	DE	Germany	83981	138
2	CN	China	53556	107
3	GB	United Kingdom	52278	134
4	JP	Japan	48596	105
5	FR	France	47287	123
6	CH	Switzerland	41266	109
7	CA	Canada	40245	117
8	TW	Taiwan, Province of China	32292	80
9	NL	Netherlands	29818	97
10	IN	India	22528	84
11	SE	Sweden	19119	84
12	BE	Belgium	16648	91
13	IL	Israel	15427	88
14	KR	Korea, Republic of	14443	79
15	IT	Italy	13518	100
16	SG	Singapore	12563	79
17	AU	Australia	10998	89
18	IE	Ireland	9707	79
19	AT	Austria	9659	70
20	FI	Finland	9171	71
21	DK	Denmark	8057	76
22	ES	Spain	6914	84
23	HK	Hong Kong	6688	62
24	KY	Cayman Islands	5346	44
25	BM	Bermuda	4916	42
26	RU	Russian Federation	4642	76
27	LU	Luxembourg	3907	61
28	NO	Norway	3883	66
29	MY	Malaysia	2578	63
30	BR	Brazil	2517	65
31	VG	Virgin Islands, British	2319	60
32	NZ	New Zealand	2076	55
33	CZ	Czech Republic	2011	51
34	PL	Poland	1814	48
35	HU	Hungary	1760	54
36	LI	Liechtenstein	1709	32
37	BB	Barbados	1639	30
38	MX	Mexico	1624	56
39	SA	Saudi Arabia	1418	62
40	ZA	South Africa	1048	56
41	TH	Thailand	1005	43
42	GR	Greece	990	46

Continua na próxima página.

	País	Nome	Número de Links	Países Distintos
43	AR	Argentina	931	49
44	RO	Romania	915	42
45	PT	Portugal	838	47
46	TR	Turkey	725	44
47	UA	Ukraine	702	42
48	AE	United Arab Emirates	686	53
49	PH	Philippines	680	32
50	BG	Bulgaria	670	37
51	CY	Cyprus	661	45
52	EG	Egypt	435	35
53	IS	Iceland	426	32
54	SK	Slovakia	410	40
55	EE	Estonia	398	29
56	CL	Chile	371	40
57	SI	Slovenia	322	32
58	ID	Indonesia	304	38
59	MT	Malta	287	24
60	CO	Colombia	274	32
61	BS	Bahamas	271	25
62	MC	Monaco	243	25
63	IR	Iran, Islamic Republic of	214	29
64	HR	Croatia	206	30
65	MU	Mauritius	202	18
66	GI	Gibraltar	200	25
67	VN	Viet Nam	200	24
68	JO	Jordan	183	18
69	BY	Belarus	174	27
70	KP	Korea, Democratic People's Republic of	169	9
71	JE	Jersey	164	14
72	PK	Pakistan	161	33
73	QA	Qatar	155	28
74	LB	Lebanon	154	25
75	RS	Serbia	151	26
76	PA	Panama	149	30
77	IM	Isle of Man	148	29
78	VE	Venezuela, Bolivarian Republic of	143	23
79	CR	Costa Rica	136	21
80	LV	Latvia	136	23
81	UY	Uruguay	125	22
82	SC	Seychelles	121	26
83	MA	Morocco	116	26
84	GE	Georgia	109	22
85	TN	Tunisia	106	21
86	KE	Kenya	102	11
87	LT	Lithuania	100	24

Continua na próxima página.

	País	Nome	Número de Links	Países Distintos
88	LK	Sri Lanka	97	22
89	AM	Armenia	93	10
90	WS	Samoa	89	10
91	PE	Peru	87	23
92	BN	Brunei Darussalam	79	8
93	NE	Niger	78	19
94	JM	Jamaica	78	13
95	OM	Oman	61	14
96	BD	Bangladesh	61	17
97	KW	Kuwait	57	13
98	MD	Moldova, Republic of	55	13
99	BZ	Belize	53	14
100	EC	Ecuador	53	8
101	KZ	Kazakhstan	52	13
102	TT	Trinidad and Tobago	50	14
103	MO	Macao	49	11
104	VC	Saint Vincent and the Grenadines	47	9
105	CM	Cameroon	41	6
106	AZ	Azerbaijan	39	13
107	NG	Nigeria	38	15
108	GA	Gabon	34	12
109	KN	Saint Kitts and Nevis	33	8
110	DO	Dominican Republic	32	6
111	BV	Bouvet Island	28	10
112	SZ	Swaziland	27	10
113	DZ	Algeria	27	11
114	GG	Guernsey	27	10
115	BH	Bahrain	27	13
116	UZ	Uzbekistan	26	11
117	CU	Cuba	24	10
118	HN	Honduras	22	7
119	TK	Tokelau	19	5
120	MN	Mongolia	19	6
121	BA	Bosnia and Herzegovina	19	6
122	NC	New Caledonia	19	10
123	TZ	Tanzania, United Republic of	19	8
124	CK	Cook Islands	19	4
125	AO	Angola	18	9
126	GH	Ghana	18	9
127	CW	Curaçao	18	8
128	TC	Turks and Caicos Islands	17	4
129	ML	Mali	17	8
130	MH	Marshall Islands	17	8
131	SL	Sierra Leone	16	9
132	SM	San Marino	16	3

Continua na próxima página.

	País	Nome	Número de Links	Países Distintos
133	AI	Anguilla	16	5
134	BO	Bolivia, Plurinational State of	16	5
135	DM	Dominica	16	6
136	GM	Gambia	16	8
137	GT	Guatemala	16	5
138	NI	Nicaragua	15	9
139	MK	Macedonia, the Former Yugoslav Republic of	15	7
140	UG	Uganda	14	6
141	AG	Antigua and Barbuda	14	6
142	SD	Sudan	14	8
143	NP	Nepal	14	5
144	SV	El Salvador	13	8
145	TV	Tuvalu	13	3
146	MM	Myanmar	13	5
147	AD	Andorra	13	7
148	ME	Montenegro	13	4
149	SB	Solomon Islands	12	6
150	AL	Albania	12	7
151	PY	Paraguay	12	4
152	FO	Faroe Islands	11	7
153	SN	Senegal	10	6
154	RE	Réunion	10	3
155	GP	Guadeloupe	10	7
156	YE	Yemen	10	8
157	SY	Syrian Arab Republic	10	3
158	LR	Liberia	10	1
159	KH	Cambodia	10	6
160	VI	Virgin Islands, U.S.	9	2
161	MG	Madagascar	8	4
162	IQ	Iraq	8	7
163	DJ	Djibouti	8	3
164	CG	Congo	7	5
165	BW	Botswana	6	3
166	CI	Côte d'Ivoire	6	3
167	PG	Papua New Guinea	6	2
168	LY	Libya	5	5
169	ER	Eritrea	5	5
170	LC	Saint Lucia	4	4
171	PS	Palestine, State of	4	2
172	SR	Suriname	4	3
173	LA	Lao People's Democratic Republic	4	1
174	GN	Guinea	4	2
175	BL	Saint Barthélemy	4	2
176	ET	Ethiopia	4	4
177	KI	Kiribati	4	4

Continua na próxima página.

	País	Nome	Número de Links	Países Distintos
178	PF	French Polynesia	4	3
179	KG	Kyrgyzstan	4	4
180	TM	Turkmenistan	3	3
181	SO	Somalia	3	2
182	ZW	Zimbabwe	3	3
183	SH	Saint Helena, Ascension and Tristan da Cunha	3	3
184	GL	Greenland	3	3
185	AF	Afghanistan	3	1
186	AW	Aruba	3	3
187	BF	Burkina Faso	3	3
188	BT	Bhutan	3	2
189	CD	Congo, the Democratic Republic of the	3	3
190	CF	Central African Republic	3	3
191	CX	Christmas Island	3	3
192	RW	Rwanda	3	3
193	FJ	Fiji	3	3
194	TD	Chad	2	2
195	WF	Wallis and Futuna	2	1
196	HT	Haiti	2	2
197	GF	French Guiana	2	1
198	GD	Grenada	2	2
199	ZM	Zambia	2	1
200	CV	Cape Verde	2	2
201	MS	Montserrat	2	2
202	VA	Holy See (Vatican City State)	2	1
203	GQ	Equatorial Guinea	2	2
204	MV	Maldives	2	2
205	BI	Burundi	2	2
206	MW	Malawi	2	2
207	PR	Puerto Rico	1	1
208	AQ	Antarctica	1	1
209	PM	Saint Pierre and Miquelon	1	1
210	GY	Guyana	1	1
211	VU	Vanuatu	1	1
212	CC	Cocos (Keeling) Islands	1	1
213	MP	Northern Mariana Islands	1	1
214	IO	British Indian Ocean Territory	1	1
215	FK	Falkland Islands (Malvinas)	1	1
216	KM	Comoros	1	1

Fonte: Elaboração Própria.



**Tabela 6** – Rede Completa de Colaborações (EPO)

	País	Nome	Número de Links	Países Distintos
0	US	United States	19212	93
1	DE	Germany	18184	82
2	NL	Netherlands	14395	60
3	FR	France	12280	82
4	CH	Switzerland	10884	60
5	GB	United Kingdom	9432	78
6	JP	Japan	7223	52
7	VG	Virgin Islands, British	3444	30
8	BE	Belgium	2833	60
9	CN	China	2751	47
10	IT	Italy	2510	70
11	AT	Austria	2454	52
12	CA	Canada	2215	59
13	ES	Spain	1501	55
14	KR	Korea, Republic of	1482	39
15	SE	Sweden	1230	44
16	DK	Denmark	1134	41
17	TW	Taiwan, Province of China	969	29
18	AU	Australia	877	45
19	IL	Israel	769	43
20	IE	Ireland	715	42
21	SG	Singapore	559	39
22	IN	India	530	36
23	FI	Finland	506	35
24	LU	Luxembourg	450	32
25	RU	Russian Federation	412	46
26	NO	Norway	392	38
27	LI	Liechtenstein	388	12
28	SA	Saudi Arabia	381	32
29	BR	Brazil	362	31
30	KY	Cayman Islands	250	22
31	PL	Poland	245	29
32	CZ	Czech Republic	233	28
33	NZ	New Zealand	189	26
34	ZA	South Africa	156	21
35	PT	Portugal	155	25
36	GR	Greece	155	32
37	AR	Argentina	152	20
38	AE	United Arab Emirates	142	16
39	BM	Bermuda	141	14
40	MX	Mexico	132	20
41	HU	Hungary	118	22

Continua na próxima página.

	País	Nome	Número de Links	Países Distintos
42	SI	Slovenia	103	22
43	MY	Malaysia	99	19
44	HK	Hong Kong	95	14
45	TR	Turkey	85	20
46	CY	Cyprus	82	24
47	UA	Ukraine	79	24
48	BB	Barbados	76	14
49	CL	Chile	57	15
50	PA	Panama	54	13
51	MC	Monaco	54	13
52	TH	Thailand	54	12
53	RO	Romania	53	22
54	MT	Malta	52	13
55	SK	Slovakia	48	17
56	CO	Colombia	44	10
57	HR	Croatia	41	13
58	LT	Lithuania	31	10
59	IS	Iceland	31	9
60	BG	Bulgaria	31	16
61	EE	Estonia	28	16
62	UY	Uruguay	28	10
63	TN	Tunisia	26	4
64	ID	Indonesia	26	9
65	MA	Morocco	24	6
66	QA	Qatar	24	8
67	BS	Bahamas	24	8
68	PG	Papua New Guinea	22	3
69	EG	Egypt	22	9
70	LV	Latvia	21	13
71	BY	Belarus	20	8
72	LB	Lebanon	20	5
73	PH	Philippines	18	8
74	CU	Cuba	18	11
75	PR	Puerto Rico	17	1
76	RS	Serbia	16	10
77	AD	Andorra	14	3
78	VE	Venezuela, Bolivarian Republic of	12	7
79	SM	San Marino	12	3
80	KW	Kuwait	11	4
81	AI	Anguilla	11	3
82	UZ	Uzbekistan	11	8
83	GI	Gibraltar	10	6
84	SC	Seychelles	10	4
85	IR	Iran, Islamic Republic of	10	8

Continua na próxima página.

	País	Nome	Número de Links	Países Distintos
86	KZ	Kazakhstan	10	5
87	JO	Jordan	10	5
88	MU	Mauritius	9	6
89	AM	Armenia	9	5
90	EC	Ecuador	9	6
91	PK	Pakistan	8	6
92	RE	Réunion	8	4
93	MG	Madagascar	7	2
94	PE	Peru	7	4
95	VN	Viet Nam	7	4
96	BA	Bosnia and Herzegovina	6	5
97	BH	Bahrain	6	5
98	KE	Kenya	6	2
99	AZ	Azerbaijan	5	5
100	FO	Faroe Islands	5	1
101	TC	Turks and Caicos Islands	4	2
102	GE	Georgia	4	3
103	GA	Gabon	4	3
104	DO	Dominican Republic	4	4
105	JM	Jamaica	4	1
106	CR	Costa Rica	4	4
107	SY	Syrian Arab Republic	3	3
108	DZ	Algeria	3	1
109	CI	Côte d'Ivoire	3	2
110	OM	Oman	3	3
111	MD	Moldova, Republic of	3	2
112	JE	Jersey	3	1
113	BZ	Belize	3	2
114	MO	Macao	3	2
115	FJ	Fiji	2	2
116	TT	Trinidad and Tobago	2	2
117	AL	Albania	2	2
118	CW	Curaçao	2	2
119	PY	Paraguay	2	2
120	BJ	Benin	2	2
121	IM	Isle of Man	2	2
122	BF	Burkina Faso	2	2
123	KH	Cambodia	2	1
124	BW	Botswana	2	2
125	GP	Guadeloupe	2	2
126	BO	Bolivia, Plurinational State of	2	1
127	VU	Vanuatu	1	1
128	BN	Brunei Darussalam	1	1
129	VI	Virgin Islands, U.S.	1	1

Continua na próxima página.

---

	País	Nome	Número de Links	Países Distintos
130	MH	Marshall Islands	1	1
131	KG	Kyrgyzstan	1	1
132	WS	Samoa	1	1
133	LK	Sri Lanka	1	1
134	LY	Libya	1	1
135	ME	Montenegro	1	1
136	IQ	Iraq	1	1
137	MM	Myanmar	1	1
138	GH	Ghana	1	1
139	ER	Eritrea	1	1
140	NG	Nigeria	1	1
141	SL	Sierra Leone	1	1
142	ET	Ethiopia	1	1
143	BD	Bangladesh	1	1
144	BT	Bhutan	1	1
145	AW	Aruba	1	1
146	SD	Sudan	1	1
147	KP	Korea, Democratic People's Republic of	1	1
148	KN	Saint Kitts and Nevis	1	1

---

---

Fonte: Elaboração Própria.

**Tabela 7** – Valores das Redes Coloridas (USPTO)

	Código	País	pib_pc	pop	dens_pop	colaborações
0	AR	Argentina	13567.95	42669500	15.59	518
1	AU	Australia	56305.98	23475686	3.06	10129
2	AT	Austria	44245.17	8546356	103.57	8863
3	BS	Bahamas	31689.25	370625	37.03	119
4	BB	Barbados	16187.79	284825	662.38	1199
5	BE	Belgium	40421.42	11209057	370.18	15616
6	BM	Bermuda	101374.47	65138	1206.26	4176
7	BR	Brazil	9214.98	202763744	24.26	1718
8	BG	Bulgaria	6796.69	7223938	66.54	416
9	CA	Canada	43635.10	35437435	3.95	39012
10	KY	Cayman Islands	75246.83	60848	253.53	4795
11	CL	Chile	13367.23	17758969	23.88	171
12	CN	China	7532.77	1371860000	145.56	52706
13	CO	Colombia	6069.19	46967706	42.33	135
14	CZ	Czech Republic	16951.38	10525347	136.30	1302
15	DK	Denmark	52404.76	5643475	141.09	7166
16	EG	Egypt	3489.86	90424668	90.84	265
17	FI	Finland	42710.92	5461512	17.97	8390
18	FR	France	36378.62	66312067	121.11	45898
19	DE	Germany	40851.16	80982500	232.11	82362
20	GR	Greece	18000.42	10892413	84.50	400
21	HK	Hong Kong	41796.59	7229500	6885.24	5849
22	HU	Hungary	12237.31	9866468	108.11	1198
23	IS	Iceland	51233.24	327386	3.27	209
24	IN	India	1503.42	1295600768	435.76	21597
25	IE	Ireland	50010.43	4657740	67.61	8816
26	IL	Israel	35710.89	8215700	379.65	14196
27	IT	Italy	29979.92	60789140	205.56	12351
28	JP	Japan	34386.91	127276000	349.18	47597
29	KR	Korea, Republic of	28094.92	50746659	520.85	13706
30	LU	Luxembourg	105583.94	556319	216.09	3327
31	MY	Malaysia	9601.18	29866606	90.90	1907
32	MT	Malta	23286.12	434558	1357.99	101
33	MX	Mexico	9426.32	120355137	61.91	1058
34	NL	Netherlands	44521.89	16865008	500.59	28894
35	NZ	New Zealand	38002.39	4516500	17.15	1438
36	NO	Norway	73650.31	5137232	14.07	2983
37	PH	Philippines	2866.82	100513137	337.10	349
38	PL	Poland	12059.25	38011735	124.14	1076
39	PT	Portugal	18833.05	10401062	113.54	230
40	RO	Romania	8670.98	19908979	86.53	606
41	RU	Russian Federation	9520.94	143819667	8.78	3608
42	SA	Saudi Arabia	20327.67	30916603	14.38	895
43	SG	Singapore	54682.93	5469724	7714.70	11742
44	ZA	South Africa	6273.57	54544184	44.96	572
45	ES	Spain	24772.34	46480882	92.92	5854
46	SE	Sweden	49855.30	9696110	23.80	18416
47	CH	Switzerland	84348.39	8188649	207.22	40006
48	TH	Thailand	5685.40	68438748	133.96	624
49	TR	Turkey	10549.68	77229262	100.35	395
50	UA	Ukraine	2346.24	45272155	78.15	409
51	AE	United Arab Emirates	36995.69	9214182	129.74	255
52	GB	United Kingdom	44596.19	64602298	267.03	50618
53	US	United States	55675.39	318386329	34.81	283855

Nota: Valores das variáveis de controle utilizadas para o ano de 2014, o último ano dos sub-períodos utilizados nas regressões. Colaborações são calculadas usando o período inteiro.

Fonte: Elaboração Própria.

**Tabela 8** – Valores das Redes Coloridas (EPO)

	Código	País	pib_pc	pop	dens_pop	colaborações
0	AU	Australia	56305.98	23475686	3.06	339
1	AT	Austria	44245.17	8546356	103.57	1952
2	BE	Belgium	40421.42	11209057	370.18	2093
3	BR	Brazil	9214.98	202763744	24.26	103
4	CA	Canada	43635.10	35437435	3.95	1588
5	CN	China	7532.77	1371860000	145.56	2021
6	DK	Denmark	52404.76	5643475	141.09	587
7	FI	Finland	42710.92	5461512	17.97	144
8	FR	France	36378.62	66312067	121.11	11266
9	DE	Germany	40851.16	80982500	232.11	17169
10	IN	India	1503.42	1295600768	435.76	277
11	IE	Ireland	50010.43	4657740	67.61	281
12	IL	Israel	35710.89	8215700	379.65	503
13	IT	Italy	29979.92	60789140	205.56	1790
14	JP	Japan	34386.91	127276000	349.18	6519
15	KR	Korea, Republic of	28094.92	50746659	520.85	1089
16	LU	Luxembourg	105583.94	556319	216.09	116
17	NL	Netherlands	44521.89	16865008	500.59	13837
18	SA	Saudi Arabia	20327.67	30916603	14.38	122
19	SG	Singapore	54682.93	5469724	7714.70	158
20	ES	Spain	24772.34	46480882	92.92	959
21	SE	Sweden	49855.30	9696110	23.80	757
22	CH	Switzerland	84348.39	8188649	207.22	10238
23	GB	United Kingdom	44596.19	64602298	267.03	8529
24	US	United States	55675.39	318386329	34.81	18088

Nota: Valores das variáveis de controle utilizadas para o ano de 2014, o último ano dos sub-períodos utilizados nas regressões. Colaborações são calculadas usando o período inteiro.

Fonte: Elaboração Própria.

**Tabela 9** – Valores da Rede Colorida pela Complexidade (USPTO)

	Código	País	eci	colaborações
0	AR	Argentina	1.06	518
1	AU	Australia	1.41	10129
2	AT	Austria	1.22	8863
3	BS	Bahamas	0.53	119
4	BB	Barbados	-0.03	1199
5	BE	Belgium	1.45	15616
6	BM	Bermuda	0.15	4176
7	BR	Brazil	1.32	1718
8	BG	Bulgaria	0.46	416
9	CA	Canada	1.39	39012
10	KY	Cayman Islands	-0.79	4795
11	CL	Chile	1.25	171
12	CN	China	0.63	52706
13	CO	Colombia	1.00	135
14	CZ	Czech Republic	1.14	1302
15	DK	Denmark	1.38	7166
16	EG	Egypt	-0.01	265
17	FI	Finland	1.53	8390
18	FR	France	1.40	45898
19	DE	Germany	1.32	82362
20	GI	Gibraltar	0.15	101
21	GR	Greece	0.77	400
22	HK	Hong Kong	0.94	5849
23	HU	Hungary	1.10	1198
24	IS	Iceland	1.09	209
25	IN	India	0.98	21597
26	IE	Ireland	0.68	8816
27	IL	Israel	0.77	14196
28	IT	Italy	1.44	12351
29	JP	Japan	0.56	47597
30	KP	Korea, Democratic People's Republic of	-1.34	119
31	KR	Korea, Republic of	-0.20	13706
32	LI	Liechtenstein	1.03	1531
33	LU	Luxembourg	1.12	3327
34	MY	Malaysia	0.74	1907
35	MT	Malta	0.50	101
36	MU	Mauritius	-0.53	105
37	MX	Mexico	1.21	1058
38	NL	Netherlands	0.97	28894
39	NZ	New Zealand	1.28	1438
40	NO	Norway	1.47	2983
41	PH	Philippines	0.37	349
42	PL	Poland	1.20	1076
43	PT	Portugal	1.18	230
44	RO	Romania	0.39	606
45	RU	Russian Federation	1.34	3608
46	SA	Saudi Arabia	1.28	895
47	RS	Serbia	-0.05	102
48	SG	Singapore	0.16	11742
49	ZA	South Africa	1.36	572
50	ES	Spain	1.44	5854
51	SE	Sweden	1.32	18416
52	CH	Switzerland	1.38	40006
53	TW	Taiwan, Province of China	0.53	31376
54	TH	Thailand	0.97	624
55	TR	Turkey	1.17	395
56	UA	Ukraine	0.80	409
57	AE	United Arab Emirates	0.76	255
58	GB	United Kingdom	1.37	50618
59	US	United States	1.27	283855
60	VG	Virgin Islands, British	0.86	1631

Fonte: Elaboração Própria.

**Tabela 10** – Valores da Rede Colorida pela Complexidade (EPO)

	Código	País	eci	colaborações
0	AU	Australia	1.31	339
1	AT	Austria	1.46	1952
2	BE	Belgium	1.13	2093
3	BR	Brazil	1.02	103
4	CA	Canada	1.17	1588
5	CN	China	1.05	2021
6	DK	Denmark	1.12	587
7	FI	Finland	1.46	144
8	FR	France	1.32	11266
9	DE	Germany	1.46	17169
10	IN	India	1.24	277
11	IE	Ireland	0.89	281
12	IL	Israel	0.82	503
13	IT	Italy	1.54	1790
14	JP	Japan	0.77	6519
15	KR	Korea, Republic of	0.20	1089
16	LI	Liechtenstein	1.41	315
17	LU	Luxembourg	0.98	116
18	NL	Netherlands	1.07	13837
19	SA	Saudi Arabia	0.89	122
20	SG	Singapore	0.80	158
21	ES	Spain	1.34	959
22	SE	Sweden	1.50	757
23	CH	Switzerland	1.19	10238
24	TW	Taiwan, Province of China	1.03	645
25	GB	United Kingdom	1.18	8529
26	US	United States	0.69	18088
27	VG	Virgin Islands, British	0.99	3233

Fonte: Elaboração Própria.



## ANEXO B – Regressões

Esta seção apresenta resultados das regressões apresentadas na [seção 4.2](#) com diferentes especificações. Primeiro variando apenas como os erros-padrões serão clusterizados, mas com uma estrutura similar à do texto. Depois, fazemos o mesmo quanto a ter diferentes tabelas com erros clusterizados de distintas maneiras mas dando ênfase na variação de efeitos-fixos. A análise do significado desses resultados é deixado à [seção 4.2](#).

Tabela 11 – Modelos De Entrada

1 Cluster

(a) USPTO

Dependent Variable: Model:	Entrada							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Gaussian	Gaussian	Gaussian	Gaussian	Logit	Logit	Logit	Logit
<i>Variables</i>								
density	0.0568 (0.0744)	0.3982*** (0.0520)	0.3915*** (0.0519)	0.3977*** (0.0527)	0.5493 (0.7203)	3.648*** (0.5980)	3.562*** (0.6029)	3.721*** (0.6233)
l_compl		0.0956*** (0.0127)	0.0954*** (0.0127)	0.0962*** (0.0127)		1.259*** (0.1250)	1.258*** (0.1246)	1.284*** (0.1255)
l_pib_pc			0.0330** (0.0164)	0.0363** (0.0165)			0.7638** (0.3127)	0.8215** (0.3225)
l_dens_pop			0.0493*** (0.0159)	0.0526*** (0.0160)			0.9178*** (0.2670)	1.007*** (0.2779)
l_compl × eci				-0.0006** (0.0003)				-0.0211** (0.0098)
<i>Fixed-effects</i>								
period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
country	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>								
Observations	227,925	227,925	227,925	227,925	227,925	227,925	227,925	227,925
Squared Correlation	0.04663	0.06737	0.06759	0.06765	0.04648	0.07226	0.07260	0.07281
Pseudo R <sup>2</sup>	0.37449	0.54691	0.54879	0.54931	0.09632	0.12724	0.12789	0.12813
BIC	19,784.5	14,785.4	14,755.3	14,752.6	107,834.6	104,211.7	104,159.8	104,144.2

Clustered (country) standard-errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

(b) EPO

Dependent Variable: Model:	Entrada							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Gaussian	Gaussian	Gaussian	Gaussian	Logit	Logit	Logit	Logit
<i>Variables</i>								
density	0.0697 (0.0865)	0.3156*** (0.0649)	0.3174*** (0.0651)	0.3187*** (0.0657)	0.6153 (0.8389)	2.949*** (0.7145)	2.974*** (0.7270)	2.991*** (0.7391)
l_compl		0.1709*** (0.0268)	0.1710*** (0.0268)	0.1713*** (0.0269)		2.488*** (0.5084)	2.493*** (0.5093)	2.497*** (0.5110)
l_pib_pc			0.0242* (0.0140)	0.0248* (0.0142)			0.5601* (0.2998)	0.5639* (0.3010)
l_dens_pop			0.0220 (0.0203)	0.0219 (0.0203)			0.6034* (0.3576)	0.5976* (0.3594)
l_compl × eci				-0.0002 (0.0002)				-0.0033 (0.0088)
<i>Fixed-effects</i>								
period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
country	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>								
Observations	179,582	179,582	179,582	179,582	179,582	179,582	179,582	179,582
Squared Correlation	0.04073	0.05210	0.05218	0.05218	0.04055	0.05281	0.05285	0.05285
Pseudo R <sup>2</sup>	0.32292	0.41548	0.41612	0.41616	0.08542	0.10349	0.10376	0.10376
BIC	16,917.2	14,788.7	14,797.9	14,809.2	86,049.2	84,386.4	84,385.4	84,397.1

Clustered (country) standard-errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Fonte: Elaboração Própria.

**Tabela 12** – Modelos De Entrada  
3 Clusters  
(a) USPTO

Dependent Variable: Model:	Entrada							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Gaussian	Gaussian	Gaussian	Gaussian	Logit	Logit	Logit	Logit
<i>Variables</i>								
density	0.0568 (0.0959)	0.3982*** (0.0914)	0.3915*** (0.0940)	0.3977*** (0.0938)	0.5493 (1.027)	3.648*** (1.001)	3.562*** (1.051)	3.721*** (1.073)
l_compl		0.0956*** (0.0129)	0.0954*** (0.0128)	0.0962*** (0.0129)		1.259*** (0.1283)	1.258*** (0.1282)	1.284*** (0.1274)
l_pib_pc			0.0330** (0.0140)	0.0363*** (0.0138)			0.7638*** (0.2921)	0.8215*** (0.2874)
l_dens_pop			0.0493 (0.0371)	0.0526 (0.0359)			0.9178 (0.7053)	1.007 (0.6569)
l_compl × eci				-0.0006** (0.0003)				-0.0211** (0.0099)
<i>Fixed-effects</i>								
period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
country	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>								
Observations	227,925	227,925	227,925	227,925	227,925	227,925	227,925	227,925
Squared Correlation	0.04663	0.06737	0.06759	0.06765	0.04648	0.07226	0.07260	0.07281
Pseudo R <sup>2</sup>	0.37449	0.54691	0.54879	0.54931	0.09632	0.12724	0.12789	0.12813
BIC	19,784.5	14,785.4	14,755.3	14,752.6	107,834.6	104,211.7	104,159.8	104,144.2

Clustered (country & ipc\_class & period) standard-errors in parentheses  
Signif. Codes: \*\*\*. 0.01, \*\*. 0.05, \*. 0.1

(b) EPO

Dependent Variable: Model:	Entrada							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Gaussian	Gaussian	Gaussian	Gaussian	Logit	Logit	Logit	Logit
<i>Variables</i>								
density	0.0697 (0.0860)	0.3156*** (0.0612)	0.3174*** (0.0612)	0.3187*** (0.0620)	0.6153 (0.8352)	2.949*** (0.6862)	2.974*** (0.6967)	2.991*** (0.7144)
l_compl		0.1709*** (0.0265)	0.1710*** (0.0265)	0.1713*** (0.0266)		2.488*** (0.5108)	2.493*** (0.5122)	2.497*** (0.5131)
l_pib_pc			0.0242** (0.0116)	0.0248** (0.0119)			0.5601** (0.2452)	0.5639** (0.2486)
l_dens_pop			0.0220 (0.0173)	0.0219 (0.0174)			0.6034** (0.2835)	0.5976** (0.2811)
l_compl × eci				-0.0002 (0.0003)				-0.0033 (0.0139)
<i>Fixed-effects</i>								
period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
country	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>								
Observations	179,582	179,582	179,582	179,582	179,582	179,582	179,582	179,582
Squared Correlation	0.04073	0.05210	0.05218	0.05218	0.04055	0.05281	0.05285	0.05285
Pseudo R <sup>2</sup>	0.32292	0.41548	0.41612	0.41616	0.08542	0.10349	0.10376	0.10376
BIC	16,917.2	14,788.7	14,797.9	14,809.2	86,049.2	84,386.4	84,385.4	84,397.1

Clustered (country & ipc\_class & period) standard-errors in parentheses  
Signif. Codes: \*\*\*. 0.01, \*\*. 0.05, \*. 0.1

Fonte: Elaboração Própria.

**Tabela 13** – Modelos De Entrada  
1 Cluster – Diferentes Efeitos-Fixos  
(a) USPTO

Dependent Variable:	Entrada					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
	Gaussian	Gaussian	Gaussian	Logit	Logit	Logit
<i>Variables</i>						
density	0.1426*** (0.0402)	0.3977*** (0.0527)	0.0968* (0.0547)	1.432*** (0.4878)	3.721*** (0.6233)	-2.534*** (0.8239)
l_compl	0.0420*** (0.0106)	0.0962*** (0.0127)	0.0914*** (0.0139)	0.2820*** (0.0825)	1.284*** (0.1255)	1.400*** (0.1474)
l_pib_pc	0.0004 (0.0024)	0.0363** (0.0165)	0.0381** (0.0158)	0.0199 (0.0465)	0.8215** (0.3225)	0.8157** (0.3798)
l_dens_pop	-0.0025 (0.0020)	0.0526*** (0.0160)	0.0634*** (0.0174)	-0.0259 (0.0316)	1.007*** (0.2779)	1.386*** (0.3583)
l_compl × eci	0.0014*** (0.0004)	-0.0006** (0.0003)	-0.0002 (0.0002)	0.0437*** (0.0103)	-0.0211** (0.0098)	-0.0015 (0.0090)
<i>Fixed-effects</i>						
period	Yes	Yes	Yes	Yes	Yes	Yes
country		Yes	Yes		Yes	Yes
ipc_class			Yes			Yes
<i>Fit statistics</i>						
Observations	227,925	227,925	227,925	227,925	227,925	227,925
Squared Correlation	0.03578	0.06765	0.08293	0.02971	0.07281	0.09115
Pseudo R <sup>2</sup>	0.28570	0.54931	0.67887	0.05689	0.12813	0.16640
BIC	20,860.1	14,752.6	18,771.5	110,964.4	104,144.2	107,430.1

Clustered (country) standard-errors in parentheses  
Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

(b) EPO

Dependent Variable:	Entrada					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
	Gaussian	Gaussian	Gaussian	Logit	Logit	Logit
<i>Variables</i>						
density	0.1554*** (0.0525)	0.3187*** (0.0657)	-0.0264 (0.0643)	1.409** (0.6440)	2.991*** (0.7391)	-4.099*** (1.087)
l_compl	0.0494*** (0.0158)	0.1713*** (0.0269)	0.1499*** (0.0247)	0.3604*** (0.1090)	2.497*** (0.5110)	2.162*** (0.4844)
l_pib_pc	0.0005 (0.0029)	0.0248* (0.0142)	0.0205 (0.0156)	0.0290 (0.0553)	0.5639* (0.3010)	0.5303 (0.4614)
l_dens_pop	-0.0031 (0.0022)	0.0219 (0.0203)	0.0111 (0.0267)	-0.0564 (0.0348)	0.5976* (0.3594)	0.2415 (0.6229)
l_compl × eci	0.0021*** (0.0007)	-0.0002 (0.0002)	0.0002 (0.0002)	0.0582*** (0.0171)	-0.0033 (0.0088)	0.0219** (0.0102)
<i>Fixed-effects</i>						
period	Yes	Yes	Yes	Yes	Yes	Yes
country		Yes	Yes		Yes	Yes
ipc_class			Yes			Yes
<i>Fit statistics</i>						
Observations	179,582	179,582	179,582	179,582	179,582	179,582
Squared Correlation	0.02776	0.05218	0.06598	0.02247	0.05285	0.07284
Pseudo R <sup>2</sup>	0.21857	0.41616	0.53000	0.04981	0.10376	0.13673
BIC	18,169.0	14,809.2	19,798.4	88,189.4	84,397.1	88,962.6

Clustered (country) standard-errors in parentheses  
Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1

Fonte: Elaboração Própria.

**Tabela 14** – Modelos De Entrada  
2 Clusters – Diferentes Efeitos-Fixos  
(a) USPTO

Dependent Variable:	Entrada					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
	Gaussian	Gaussian	Gaussian	Logit	Logit	Logit
<i>Variables</i>						
density	0.1426*** (0.0404)	0.3977*** (0.0560)	0.0968* (0.0584)	1.432*** (0.4899)	3.721*** (0.6656)	-2.534*** (0.8490)
l_compl	0.0420*** (0.0107)	0.0962*** (0.0128)	0.0914*** (0.0139)	0.2820*** (0.0828)	1.284*** (0.1287)	1.400*** (0.1514)
l_pib_pc	0.0004 (0.0024)	0.0363** (0.0165)	0.0381** (0.0158)	0.0199 (0.0465)	0.8215** (0.3228)	0.8157** (0.3798)
l_dens_pop	-0.0025 (0.0020)	0.0526*** (0.0163)	0.0634*** (0.0177)	-0.0259 (0.0316)	1.007*** (0.2823)	1.386*** (0.3617)
l_compl × eci	0.0014*** (0.0004)	-0.0006** (0.0003)	-0.0002 (0.0002)	0.0437*** (0.0105)	-0.0211** (0.0099)	-0.0015 (0.0090)
<i>Fixed-effects</i>						
period	Yes	Yes	Yes	Yes	Yes	Yes
country		Yes	Yes		Yes	Yes
ipc_class			Yes			Yes
<i>Fit statistics</i>						
Observations	227,925	227,925	227,925	227,925	227,925	227,925
Squared Correlation	0.03578	0.06765	0.08293	0.02971	0.07281	0.09115
Pseudo R <sup>2</sup>	0.28570	0.54931	0.67887	0.05689	0.12813	0.16640
BIC	20,860.1	14,752.6	18,771.5	110,964.4	104,144.2	107,430.1

Clustered (country & ipc\_class) standard-errors in parentheses  
Signif. Codes: \*\*\*, 0.01, \*\*, 0.05, \*, 0.1

(b) EPO

Dependent Variable:	Entrada					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
	Gaussian	Gaussian	Gaussian	Logit	Logit	Logit
<i>Variables</i>						
density	0.1554*** (0.0527)	0.3187*** (0.0682)	-0.0264 (0.0673)	1.409** (0.6451)	2.991*** (0.7682)	-4.099*** (1.111)
l_compl	0.0494*** (0.0158)	0.1713*** (0.0270)	0.1499*** (0.0248)	0.3604*** (0.1092)	2.497*** (0.5139)	2.162*** (0.4875)
l_pib_pc	0.0005 (0.0029)	0.0248* (0.0143)	0.0205 (0.0156)	0.0290 (0.0553)	0.5639* (0.3028)	0.5303 (0.4618)
l_dens_pop	-0.0031 (0.0022)	0.0219 (0.0202)	0.0111 (0.0266)	-0.0564 (0.0348)	0.5976* (0.3571)	0.2415 (0.6205)
l_compl × eci	0.0021*** (0.0007)	-0.0002 (0.0002)	0.0002 (0.0002)	0.0582*** (0.0172)	-0.0033 (0.0095)	0.0219** (0.0108)
<i>Fixed-effects</i>						
period	Yes	Yes	Yes	Yes	Yes	Yes
country		Yes	Yes		Yes	Yes
ipc_class			Yes			Yes
<i>Fit statistics</i>						
Observations	179,582	179,582	179,582	179,582	179,582	179,582
Squared Correlation	0.02776	0.05218	0.06598	0.02247	0.05285	0.07284
Pseudo R <sup>2</sup>	0.21857	0.41616	0.53000	0.04981	0.10376	0.13673
BIC	18,169.0	14,809.2	19,798.4	88,189.4	84,397.1	88,962.6

Clustered (country & ipc\_class) standard-errors in parentheses  
Signif. Codes: \*\*\*, 0.01, \*\*, 0.05, \*, 0.1

Fonte: Elaboração Própria.

**Tabela 15** – Modelos De Entrada  
3 Clusters – Diferentes Efeitos-Fixos  
(a) USPTO

Dependent Variable: Model:	Entrada					
	(1)	(2)	(3)	(4)	(5)	(6)
	Gaussian	Gaussian	Gaussian	Logit	Logit	Logit
<i>Variables</i>						
density	0.1426*** (0.0406)	0.3977*** (0.0938)	0.0968* (0.0550)	1.432*** (0.5110)	3.721*** (1.073)	-2.534** (1.236)
l_compl	0.0420*** (0.0126)	0.0962*** (0.0129)	0.0914*** (0.0141)	0.2820*** (0.0987)	1.284*** (0.1274)	1.400*** (0.1464)
l_pib_pc	0.0004 (0.0023)	0.0363*** (0.0138)	0.0381*** (0.0131)	0.0199 (0.0454)	0.8215*** (0.2874)	0.8157** (0.3323)
l_dens_pop	-0.0025 (0.0021)	0.0526 (0.0359)	0.0634** (0.0309)	-0.0259 (0.0326)	1.007 (0.6569)	1.386** (0.5563)
l_compl × eci	0.0014*** (0.0004)	-0.0006** (0.0003)	-0.0002 (0.0002)	0.0437*** (0.0076)	-0.0211** (0.0099)	-0.0015 (0.0101)
<i>Fixed-effects</i>						
period	Yes	Yes	Yes	Yes	Yes	Yes
country		Yes	Yes		Yes	Yes
ipc_class			Yes			Yes
<i>Fit statistics</i>						
Observations	227,925	227,925	227,925	227,925	227,925	227,925
Squared Correlation	0.03578	0.06765	0.08293	0.02971	0.07281	0.09115
Pseudo R <sup>2</sup>	0.28570	0.54931	0.67887	0.05689	0.12813	0.16640
BIC	20,860.1	14,752.6	18,771.5	110,964.4	104,144.2	107,430.1

Clustered (country & ipc\_class & period) standard-errors in parentheses  
Signif. Codes: \*\*\*, 0.01, \*\*, 0.05, \*, 0.1

(b) EPO

Dependent Variable: Model:	Entrada					
	(1)	(2)	(3)	(4)	(5)	(6)
	Gaussian	Gaussian	Gaussian	Logit	Logit	Logit
<i>Variables</i>						
density	0.1554*** (0.0519)	0.3187*** (0.0620)	-0.0264 (0.0832)	1.409** (0.6394)	2.991*** (0.7144)	-4.099** (1.693)
l_compl	0.0494*** (0.0163)	0.1713*** (0.0266)	0.1499*** (0.0245)	0.3604*** (0.1224)	2.497*** (0.5131)	2.162*** (0.4880)
l_pib_pc	0.0005 (0.0029)	0.0248** (0.0119)	0.0205 (0.0130)	0.0290 (0.0547)	0.5639** (0.2486)	0.5303 (0.3937)
l_dens_pop	-0.0031 (0.0021)	0.0219 (0.0174)	0.0111 (0.0222)	-0.0564* (0.0328)	0.5976** (0.2811)	0.2415 (0.5221)
l_compl × eci	0.0021** (0.0009)	-0.0002 (0.0003)	0.0002 (0.0003)	0.0582*** (0.0176)	-0.0033 (0.0139)	0.0219 (0.0151)
<i>Fixed-effects</i>						
period	Yes	Yes	Yes	Yes	Yes	Yes
country		Yes	Yes		Yes	Yes
ipc_class			Yes			Yes
<i>Fit statistics</i>						
Observations	179,582	179,582	179,582	179,582	179,582	179,582
Squared Correlation	0.02776	0.05218	0.06598	0.02247	0.05285	0.07284
Pseudo R <sup>2</sup>	0.21857	0.41616	0.53000	0.04981	0.10376	0.13673
BIC	18,169.0	14,809.2	19,798.4	88,189.4	84,397.1	88,962.6

Clustered (country & ipc\_class & period) standard-errors in parentheses  
Signif. Codes: \*\*\*, 0.01, \*\*, 0.05, \*, 0.1

Fonte: Elaboração Própria.