

Internalização de transbordamentos de conhecimento por regiões: Uma medida baseada em autocitações de patentes*

Beatriz Pereira de Almeida¹

Eduardo Gonçalves²

André Suriane da Silva³

Raquel Coelho Reis⁴

Resumo:

Este artigo aplica uma medida de autocitações de patentes como *proxy* para a capacidade de internalizar transbordamentos de conhecimento das regiões. Trabalhos anteriores discutiram a medida ao nível da firma, considerando que firmas com maior capacidade de internalizar conhecimento e, portanto, se apropriar mais dos ganhos de seus investimentos em P&D, são também firmas com maior estoque de conhecimento acumulado. Utilizando informações contidas nas citações de patentes da base de dados de patentes da OECD e demais controles, estimou-se um modelo com variável dependente binária, um modelo logit para dados longitudinais, em um período de dez anos. Verificou-se que regiões com maior grau de generalidade tendem a internalizar mais conhecimento, bem como regiões que realizam maior gasto em P&D e são especializadas em áreas do conhecimento mais tecnológicas. Essas conclusões indicam que as regiões produtoras de ciência básica mais inovadoras, possuem maior habilidade para se apropriar dos ganhos de seu processo inovativo.

Palavras-Chaves: Citações, Internalização de transbordamentos, Patentes.

This paper uses patent self-citation data as a way of measuring the ability to internalize knowledge spillovers in the context of regions. Previous studies have discussed the measure at firm level, considering that firms with a greater ability to internalize knowledge and, therefore, to appropriate more of the gains of their R&D investments, are also firms with a higher stock of accumulated knowledge. Using information contained in patent citations from the OECD patent database and other controls, a binary dependent variable model, a logit model for longitudinal data, was estimated over a ten years period. It was verified that regions with a greater degree of generality tend to internalize more knowledge, as well as regions that carry out greater spending on R&D and are specialized in more technological knowledge fields. These conclusions indicate that the most innovative basic science specialized regions have a greater ability to appropriate the gains of their innovative process.

Keywords: Citations, Internalization of spillovers, Patents.

JEL: O30; O33; R11.

Área 10 - Economia Regional e Urbana

¹ Doutoranda do PPGE/UFJF, beatriz.p.almeida1@gmail.com.

² Professor Associado da Faculdade de Economia, Programa de Pós-Graduação em Economia, PPGE/UFJF, Pesquisador do CNPq, eduardo.goncalves@ufjf.edu.br

³ Economista, Faculdade de Economia, UFJF, andresuriane@gmail.com.

⁴ Doutoranda do PPGE/UFJF, raquelcoelhoreis@gmail.com.

*Os autores agradecem o apoio da FAPEMIG, CNPq e CAPES.

1. Introdução

É amplo o interesse em estudar o conhecimento tecnológico enquanto importante motor do crescimento econômico das nações. Quando o conhecimento pode ser codificado, medir o seu estoque e fluxo entre regiões e firmas é possível por meio da análise das migrações de trabalhadores especializados, das colaborações em artigos científicos e patentes e, também, por meio das citações de patentes.

Quando uma patente cita outras que foram necessárias para o seu desenvolvimento, essas citações são chamadas de citações para trás (*backward citations*), e, quando uma patente em questão é citada por outras desenvolvidas posteriormente, essas citações são chamadas de citações para frente (*forward citations*). As citações para trás são utilizadas para identificar a origem ou base de conhecimento utilizada para desenvolver uma invenção e podem medir os fluxos de conhecimento tecnológico. As citações para frente medem a importância tecnológica da patente (Nagaoka *et al.*, 2010) e, assim, também podem mensurar o valor de mercado da inovação (Breitzman e Thomas, 2002). Esses fluxos de conhecimento tecnológico, quando transpassam os limites das firmas e regiões inventoras, são chamados de transbordamentos de conhecimento (*knowledge spillovers*) (Griliches, 1979; Jaffe *et al.*, 1993).

Se o interesse é medir os fluxos dos transbordamentos de conhecimento, as citações para trás que as próprias firmas e regiões inventoras fazem de si mesmas - as autocitações - não são utilizadas, por não representarem transbordamentos entre agentes e regiões distintas, respectivamente. Entretanto, as autocitações podem ter outras funções. Para Mancusi (2004; 2008), as autocitações de firmas, ou seja, firmas que citam suas próprias patentes indicam que desenvolveram novas tecnologias baseadas em conhecimento prévio gerado por suas pesquisas anteriores no mesmo campo tecnológico. Isso representaria, portanto, o acúmulo de conhecimento interno à firma, de modo que essa seria uma medida de sua capacidade de absorção. No caso de regiões, Mukherji e Silberman (2013) consideram que pode haver um problema na aplicação do conceito de capacidade de absorção baseado em autocitações, já que regiões com alto percentual de autocitações seriam mais limitadas ao conhecimento externo, tendo um comportamento míope (Agrawal *et al.*, 2009), o que não seria compatível com um alto desempenho inovador.

Belezon (2006) também utiliza citações como forma de medir transbordamentos. A internalização de transbordamentos pelas firmas é dada como uma razão entre o volume de citações indiretas para trás das próprias patentes da firma⁵ e o volume total de citações para frente das patentes da firma. Quando ocorre a internalização, ou seja, quando o conhecimento retorna à dinâmica inovativa da firma, essa firma estaria se apropriando de parte de seus transbordamentos, mecanismo pelo qual ela obteria maiores retornos dos investimentos em P&D iniciais. Sua principal hipótese é de que transbordamentos de conhecimento de caráter mais geral, ou de conhecimento básico, são menos prováveis de retroalimentar a dinâmica inovativa da firma.

Esse trabalho busca aplicar a ideia de autocitações indiretas para o contexto de regiões. Nesse sentido, a internalização ocorre quando as citações para trás feitas por uma região se referem a transbordamentos de uma tecnologia desenvolvida pela própria região. A ideia é que regiões especializadas em conhecimento mais básico teriam maior dificuldade em se apropriar

⁵ Belezon (2006) define as autocitações indiretas como as citações que determinada firma inventora faz (*backward citation*) de patentes de outras firmas que, por sua vez, citaram a patente da firma inventora em questão. Essa seria uma forma de se apropriar dos desenvolvimentos tecnológicos de seu próprio conhecimento.

dos retornos do seu investimento em P&D através da internalização. Entretanto, regiões especializadas em conhecimento básico também podem ser dotadas de uma vasta gama de conhecimento tecnológico, que permitem que haja maior apropriação mesmo com os transbordamentos para diferentes campos tecnológicos.

Com uma variável dependente binária, que assume valor unitário no caso em que há internalização de conhecimento, foi utilizado um modelo Logit para dados longitudinais em um período de 10 anos para medir o impacto das variáveis explicativas na probabilidade de ocorrência da internalização do conhecimento regional. A desagregação regional escolhida compreende 516 regiões ano nível NUTS 2 de 44 países. Além das bases de patentes e citações de patentes da OECD, dados referentes à economia e a investimentos em inovação e educação da OECD e da EUROSTAT foram utilizados.

Além da introdução, esse trabalho aborda, na segunda seção, discussões teóricas e da literatura empírica sobre a capacidade de absorção regional. Na terceira seção, a estratégia empírica utilizada e uma explicação sobre as bases de dados e variáveis utilizadas são apresentadas; os principais resultados são discutidos na quarta seção e as conclusões e apontamentos futuros encerram o trabalho.

2. Referencial teórico

2.1 Transbordamentos de conhecimento e capacidade de absorção

Transbordamentos são usualmente definidos como externalidades do processo inovativo. Ou seja, são parte do produto gerado pelo processo inovativo e que não pode ser apropriado pelos agentes inovadores. Também podem ser definidos como transbordamentos de conhecimento de uma firma para outra, que ocorre sem que haja transação de mercado com esse fim (Grossman e Helpman, 1992). Os transbordamentos são difíceis de serem medidos. Geralmente, os fluxos de patentes entre firmas ou indústrias são utilizados com esse fim, já que as patentes são uma forma de conhecimento codificado e, portanto, pode ser mensurado.

Para que o processo de aprendizagem a partir do conhecimento externo ocorra e afete o desempenho inovativo, duas condições são necessárias: que haja transbordamentos e que o agente receptor do conhecimento externo possua capacidade de absorver esses transbordamentos (Filippetti *et al.*, 2017).

A capacidade de absorção foi definida por Cohen e Levinthal (1989) como a habilidade das firmas de aprender a partir de conhecimento externo por meio de processos de identificação de conhecimento relevante disponível, sua assimilação e posterior exploração no processo inovativo. Apesar de o termo ter surgido no contexto microeconômico, pode ser adaptado também para outras instituições, como países e regiões (CARAGLIU e NIJKAMP, 2008). A ideia de capacidade de absorção vai além da discussão ao nível das firmas, sendo também aplicada ao contexto de regiões e países (Cantwell e Iammarino, 2003; Doloreux e Parto, 2005; Mukherji e Silberman, 2013; Roper e Love, 2006; Von Tunzelmann, 2009).

Argumenta-se que a capacidade de absorção de uma região é dada pela agregação da capacidade das firmas e organizações locais, bem como pela sua interação (Abreu, 2011). Outros fatores a serem considerados no contexto regional são as instituições, geografia e a interação que ocorre entre inventores e empresas (Agrawal, Cockburn e Rosell, 2010).

Nesse sentido, a capacidade de inovar é influenciada pela capacidade de absorção, ou a capacidade de aprendizado do conhecimento externo, assim como pela capacidade interna de criação de conhecimento. Essas duas formas de aprendizado, do conhecimento externo e interno, ou as “duas faces” da P&D de Cohen e Levinthal (1989) - inovação e aprendizado - são baseadas em processos, rotinas e estratégias diferentes.

2.2 Literatura empírica

O conceito de capacidade de absorção e transbordamentos do conhecimento tem sido amplamente empregado em o nível de empresas e setores (Cassiman e Veugelers, 2002; Schmidt, 2005; Veugelers, 1997), regiões (Jaffe, Trajtenberg e Henderson, 1993; Maurseth e Verspagen, 2002; Doring e Schnellenbach, 2006) e nações (Mowery e Oxley, 1995; Criscuolo e Narula, 2002, Narula, 2004, Kneller, 2005, Kneller e Stevens, 2006).

O conhecimento externo à região é absorvido por ela de forma mais eficiente quando a região possui um *background* de conhecimento maior. A medida de capacidade de absorção na literatura empírica é dada pela intensidade de P&D de empresas, regiões e países, uma vez que o P&D seria responsável não apenas pela geração de novos conhecimentos, mas pela melhoria das habilidades de empresas e regiões em assimilar e utilizar novas tecnologias (Cohen e Levinthal, 1989).

A capacidade de absorção está relacionada à habilidade das empresas em assimilar e utilizar o conhecimento para fins de melhoria da capacidade inovativa e, por consequência, das vantagens competitivas (Zahra e George, 2002; Doring e Schnellenbach, 2006).

Os processos de aquisição e assimilação do conhecimento tecnológico são considerados como a capacidade de absorção potencial das firmas, enquanto a transformação e exploração do conhecimento, representam a capacidade de absorção efetivamente realizada. A capacidade de absorção potencial depende da disponibilidade de fontes de conhecimento relevantes e do tipo de cooperação aos quais a empresa tem acesso, e a capacidade de absorção realizada depende do grau de adequação da tecnologia (Zahra e George, 2002).

O conceito também foi aplicado ao contexto regional (Cantwell e Iammarino, 2003; Doloreux e Parto, 2005; Mukherji e Silberman, 2013; Roper e Love, 2006; Von Tunzelmann, 2009). Neste contexto, a capacidade de absorção das firmas e organizações determinam a capacidade de absorção agregada das regiões, já que as firmas são agentes basilares do Sistema Regional de Inovação. A dimensão regional é capaz de englobar as conexões entre as firmas, além de ser uma forma de análise mais agregada (Abreu, 2011).

O uso de novas técnicas de gerenciamento, o comportamento colaborativo nas empresas e o número de empregados envolvidos no processo de P&D, por exemplo, afetam a capacidade de inovação regional (Abreu *et al.*, 2008). Discute-se também, no contexto regional, quando o conhecimento desenvolvido por empresas internas influencia ainda mais regiões externas – principalmente em regiões dotadas de grandes empresas. Essas empresas possuem maior capacidade inovativa, medida pelo seu maior grau de patenteamento de tecnologias, o que não afeta de maneira tão efetiva o desenvolvimento local, pelo maior efeito dos transbordamentos (Caragliu e Nijkamp, 2012).

Outros fatores que podem influenciar a capacidade de absorção são os conhecimentos prévios relacionados à tecnologia desenvolvida e as habilidades individuais, já que o conhecimento acumulado e a experiência facilitam a utilização de novo conhecimento (Cohen e Levinthal, 1990). Como consequência, a capacidade de absorção tem natureza cumulativa, o que implica que o grau de escolaridade dos empregados da firma afeta essa capacidade.

Os dados de citações de patentes são uma importante fonte de informações sobre as ligações existentes entre as novas invenções e as passadas. Quando uma patente cita outra, isso indica que o inventor se utilizou de conhecimento disponível e gerado em outra região ou firma para a geração de uma nova patente (Nagaoka *et al.*, 2010). Nesse sentido, as análises das citações das patentes evidenciam a dinâmica do conhecimento no tempo e no espaço (Hu e Jaffe, 2003).

As citações feitas são como o rastro de transbordamentos gerados pela patente originária – medindo como o fluxo de conhecimento ocorre e qual a importância das patentes, de acordo com o seu impacto em termos desse fluxo gerado. As citações podem indicar também quando

uma determinada linha de pesquisa está sendo aprofundada, quando patentes na mesma área tecnológica fazem as citações, ou se uma inovação é capaz de influenciar diversas áreas, quando as citações são feitas por patentes em diferentes campos tecnológicos (Antonelli *et al.*, 2006).

A partir do trabalho de Jaffe *et al.* (1993), convencionou-se excluir as citações de patentes feitas pela mesma organização originária da patente, porque a autocitação não representaria de fato a geração de uma externalidade do conhecimento – ou spillover. Já para Mancusi (2004), as autocitações podem ser utilizadas como medida da capacidade de absorção das firmas. A realização da autocitação significaria que a firma desenvolveu novas tecnologias baseada em conhecimento prévio gerada por suas pesquisas anteriores no mesmo campo tecnológico, o que representaria, portanto, o acúmulo de conhecimento interno à firma. No mesmo sentido, Belenzon (2006) considerou como internalização do conhecimento uma medida de autocitação no contexto de firmas. Entretanto, em sua medida, diferenciou a autocitação direta da indireta. A primeira é a que ocorre quando uma firma cita uma patente desenvolvida internamente para o desenvolvimento de uma tecnologia num período posterior, que indica que a firma desenvolveu o conhecimento inicial de forma direta. A segunda medida é a de autocitação indireta, em que o conhecimento gerado pela firma, subsequentemente desenvolvido por outras firmas, retorna à firma inventora, num processo associado pelo autor à maior capacidade de apropriação do conhecimento.

No contexto de firmas, o seu acúmulo de conhecimento – medido pelo número de autocitações – é uma medida capaz de mensurar a capacidade da firma em desenvolver pesquisas e absorver conhecimento externo. No caso de regiões, Mukherji e Silberman (2013) consideram que pode haver um problema na aplicação do conceito de capacidade de absorção baseado em autocitações, já que regiões com alto percentual de autocitações seriam mais limitadas ao conhecimento externo, tendo um comportamento míope, o que não seria compatível com um alto desempenho inovador. Agrawal *et al.*, (2009) apontam o problema da miopia em relação às regiões geográficas e áreas tecnológicas a que pertencem as patentes citadas. Eles analisam cidades em que a atividade inovativas está concentrada em uma única grande firma e concluem que inventores nesses locais tendem a utilizar mais o seu próprio conhecimento prévio medido pelas autocitações. Seus resultados apontam para um possível menor impacto das inovações geradas nessas regiões que em regiões mais industrialmente diversas.

Belenzon (2006) desenvolve uma metodologia empírica nova baseada em citações de patentes para testar se a apropriabilidade da inovação é menor para o conhecimento básico em um quadro dinâmico de inovação sequencial. No trabalho, os transbordamentos são medidos como desenvolvimentos sequenciais de conhecimento provenientes de fora da empresa inventora e são internalizados quando a empresa inventora absorve transbordamentos advindos de desenvolvimentos da sua inovação. Como a internalização tecnológica é um canal através do qual as rendas privadas são apropriadas por inventores iniciais, o trabalho de Belenzon analisa os incentivos para a inovação em áreas básicas do conhecimento.

Tendo em vista a lacuna da literatura em tratar a questão das autocitações de patentes em nível regional, o amplo interesse na busca por variáveis capazes de medir a capacidade de absorção e as hipóteses testadas em trabalhos anteriores, os seguintes questionamentos são levantados neste trabalho:

H1: A medida de capacidade de absorção baseada em autocitações pode ser estendida para o contexto regional?

H2: Regiões mais diversas tecnologicamente tendem a internalizar mais conhecimento?

H3: Regiões especializadas em conhecimento mais geral, dito conhecimento de base, internalizam mais ou menos conhecimento?

H4: O investimento em P&D feito pela região influencia sua capacidade de internalizar conhecimento?

3. Estratégia Empírica

Os dados foram coletados junto à OECD na base de dados Patent Datasets (February 2016 Edition), que é formada por vários bancos de dados com informações sobre o depósito de patentes na Patent Cooperation Treaty (PCT), na European Patent Office (EPO) e no United States Patent and Trademark Office (USPTO), abrangendo o período entre os anos de 1990 a 2015. Os dados de citação de patentes, obtidos na OECD Citations Database, levam em consideração os depósitos de patentes PCT, que possuíam informações sobre a localização geográfica de origem do inventor, obtidas nas bases OECD REGPAT Database e OECD Triadic Patent Families. As informações contidas nessa base abrangem 44 países de todos os continentes, entre os países membros da OECD e seus principais parceiros econômicos, divididas em 645 regiões ao nível NUTS2 e TL2 de todos os continentes.

Para controlar para as características das regiões, os dados da EUROSTAT Database também foram utilizados. Essa base fornece dados socioeconômicos, geográficos e ligados à ciência e tecnologia ao nível NUTS2 de 28 países da EU e outros.

A variável dependente se baseia na contagem do número de patentes que foram internalizadas por cada região. Os transbordamentos podem ser divididos entre transbordamentos internalizados ($Spill_{Intern_{it}}$) e externalizados ($Spill_{Extern_{it}}$). Eles são considerados como internalizados quando retornam à região de origem após terem sido desenvolvidos por outra região. No diagrama da Figura 1 esse mecanismo é explicitado.

Nesse caso, uma patente A, dita originária, é desenvolvida pela região X representada pela cor cinza claro, e é citada pela patente B da região Y, em azul, o que configura um transbordamento de conhecimento para fora dos limites da região. Nesse caso, a região Y se beneficiou da tecnologia da região X (transbordamento inter-regional). Posteriormente, a mesma região originária X cita a patente B para o desenvolvimento de sua patente C, de modo que ela utiliza de maneira indireta, via patente B, de um conhecimento originalmente desenvolvido por inventores que residem na própria região X. Dessa forma, a região internaliza o transbordamento de seu conhecimento, num processo de autocitação indireta.

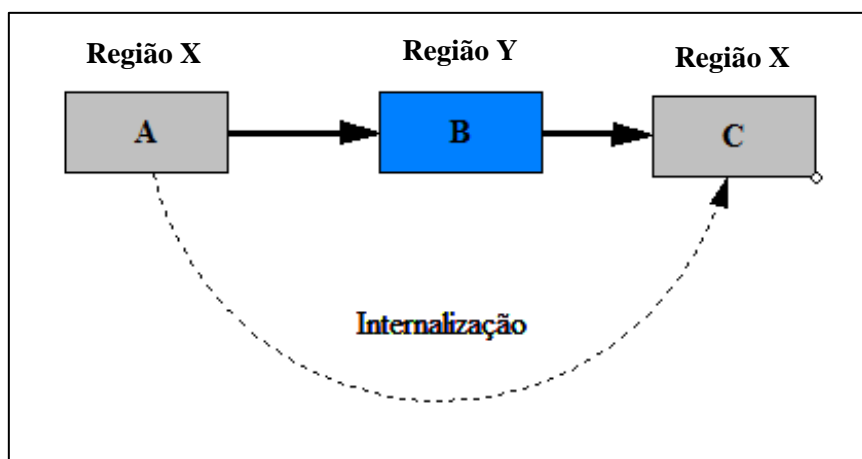


Figura 1: Internalização de transbordamentos de conhecimento

Fonte: Elaboração própria.

Belenzon (2006) utiliza ainda a ideia de linhas de pesquisa em sua análise para o caso de firmas. Na Figura 2, as linhas de pesquisa que surgiram da patente A, da firma originária, são cinco (A-B-D-H; A-B-E-I; A-C-F-I; A-C-F-J; A-C-G-J) e duas delas geram internalização de conhecimento (A-B-E-I; A-C-F-I). As linhas internalizadas são as que terminam com a

patente I, desenvolvida pela mesma região originária de A, que está representada pela cor verde escuro. Essas duas linhas de pesquisa com internalização geram, cada uma, transbordamentos para duas patentes diferentes de firmas externas. Portanto, o transbordamento total contabilizado pelas linhas de pesquisa internalizadas é de 4 (B e E e C e F).

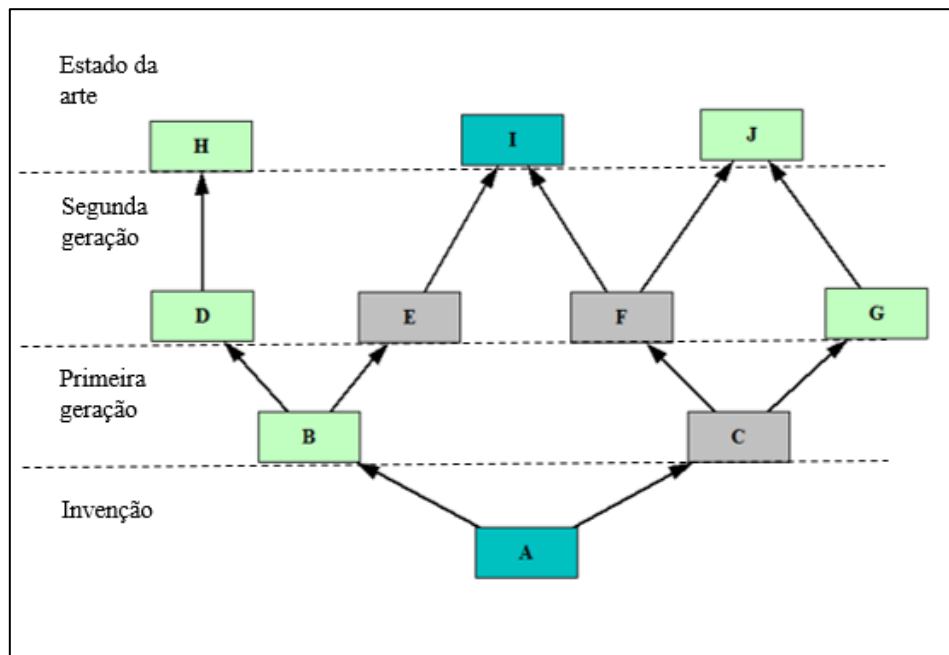


Figura 2: Linhas de pesquisa e internalização de transbordamentos

Fonte: Adaptado de Belenzon (2006)

As variáveis Generalidade e Originalidade foram utilizadas também no trabalho de Belenzon (2006). Elas foram desenvolvidas por Trajtenberg, Henderson e Jasse (1992) em seu trabalho sobre apropriabilidade da pesquisa básica. Uma patente é tanto mais geral quanto mais o conhecimento gerado por ela for disseminado por diferentes campos tecnológicos. O índice computado é dado por um menos um Índice Herfindhal de citações ao nível de 3 dígitos (Nclass). Quanto maior o índice, menos concentradas as citações em determinados campos tecnológicos e mais geral é a patente originária.

Em trabalho subsequente (Trajtenberg, Henderson e Jaff, 1993), os autores testam empiricamente a possível concentração espacial das citações de patentes dado seu grau de generalidade. Eles relacionam patentes com alto percentual de autocitações – que seria uma medida do sucesso do esforço das firmas em se apropriar de seu conhecimento gerado – com o grau de generalidade, e verificam o seu grau de localização.

A medida de originalidade se refere às citações feitas pela patente inicial. Portanto, se uma patente cita patentes anteriores que pertencem a um dado conjunto específico de tecnologias, o índice de originalidade da patente será baixo e, se ela cita um vasto grupo de patentes, seu índice será alto (Trajtenberg, Henderson e Jaff, 1993). Patentes mais originais são também de conhecimento mais básico.

Optou-se por considerar a variável dependente como uma *dummy*, já que apenas apresenta valores positivos para 42 observações. Dessa forma, um modelo de regressão não linear é a forma mais adequada de proceder a estimação. A estimação via Mínimos Quadrados Ordinários, nesse caso, produziria estimadores inconsistentes.

A estimação via Máxima Verossimilhança estima parâmetros que maximizam a probabilidade de uma determinada amostra pertencer à uma população especificada.

Os modelos Logit e Probit são os mais utilizados na literatura empírica, quando se deseja trabalhar com variáveis dependentes de resposta binária (Lima, 1996). Os modelos de escolha binária, ou qualitativos, com dados em painel são formulados da seguinte forma latente:

$$y_{it}^* = x_i' \beta + \alpha_i + \varepsilon_{it} \quad (1)$$

,em que $y_{it} = 1$ se $y_{it}^* > 0$ e $y_{it} = 0$ do contrário. x representa o vetor de variáveis independentes, i denota o indivíduo, t denota o tempo e α_i contempla as variações existentes entre unidades seccionais. Nesse trabalho, y_{it} indica se a região i teve uma internalização positiva de transbordamentos de conhecimento no período t ou não.

Assim, em termos da forma funcional:

$$\begin{aligned} P(y_{it} = 1|x_{it}) &= \Lambda(x_{it}\beta + f_i) \\ P(y_{it} = 0|x_{it}) &= 1 - [\Lambda(x_{it}\beta + f_i)] \end{aligned} \quad (2)$$

Em que f_i é a característica intrínseca a cada unidade seccional – o efeito fixo. O pressuposto é de que a distribuição dos resultados segue uma distribuição logística, aqui denotada pela letra grega Λ .

Nos modelos de variável binária, α_i pode ser considerado como um efeito fixo, ou como um termo de erro aleatório⁶. Em modelos de dados em painel por efeitos fixos (*FE*), o α_i é tratado como uma variável aleatória não observável que pode ser correlacionada com os regressores x_{it} .

Pode-se estimar também modelos de variável binária em painel pelo método de efeitos aleatórios (*RE*), no qual o efeito específico do indivíduo α_i é eliminado ao se integrar à distribuição. Modelos com dados agrupados (*pooled*) definem $\alpha_i = \alpha$.

O modelo a ser estimado, com as respectivas variáveis e seus efeitos, é o seguinte:

$$Y_{it} = \beta_0 + \beta_1 G_{it} + \beta_2 O_{it} + \beta_3 IT_{it} + \beta_4 Q_{it} + \beta_5 gapc_{it} + \beta_6 KH_{it} + \beta_7 P\&D_{it} + \beta_8 PIBpc_{it} + \varepsilon_{it} \quad (3)$$

As unidades seccionais i são as 645 regiões (NUTS2) do mundo e as unidades temporais t compreendem os anos de 2000 a 2010.

1) $Y_{it} = dummy$ que indica se houve internalização de transbordamentos de conhecimento pela região i no tempo t , ou seja, assume valor 1 caso $Spill_{Intern_{it}}$ seja maior que 0.

$$Y_{it} = \begin{cases} Y_i = 0; & \text{se } Spill_{Intern_{it}} = 0 \\ Y_i = 1; & \text{se } Spill_{Intern_{it}} > 0 \end{cases} \quad (4)$$

2) $G_{it} = grau$ de generalidade das patentes da região;

$$G_{it} = 1 - \sum_{n=1}^n \left(\frac{CR_{it,n}}{CR_{it}} \right)^2 \quad (5)$$

⁶ O teste de Hausman é realizado para definir a escolha do melhor modelo. A estatística do teste é apresentada no texto. As hipóteses do teste são definidas como H0: ambos os estimadores são consistentes, mas o estimador de efeitos aleatórios é eficiente; H1: Só o estimador de efeitos fixos é consistente. O teste é distribuído através de uma χ^2 com n graus de liberdade.

em que n se refere às subclasses IPC das patentes, $CR_{it,n}$ é o número de citações recebidas pela região i no tempo t a partir de patentes no campo n e CR_{it} é o número de citações totais recebidas pela patente i no período t. Para corrigir o viés de regiões que foram pouco citadas, a generalidade é estimada da seguinte forma:

$$\hat{G}_{it} = \left(\frac{CR_{it}}{CR_{it-1}} \right) G_{it} \quad (6)$$

3) \hat{O}_{it} = grau de originalidade das patentes geradas pela região i no tempo t:

$$O_{it} = 1 - \sum_n \left(\frac{CF_{in,t}}{CF_{i,t}} \right)^2 \quad (7)$$

$$\hat{O}_{it} = \left(\frac{CF_{it}}{CF_{it-1}} \right) O_{it} \quad (8)$$

$CF_{in,t}$ = número de citações feitas pela região i nos campos tecnológicos n no tempo t;
 CF_{it} = número total de citações feitas pela região i no tempo t.

4) IT_{it} = intensidade tecnológica das patentes da região i no tempo t;

$$IT_{it} = \frac{\frac{\sum_{n=1}^{33} \text{patentes}_{n,it}}{\sum \text{patentes}_{it}}}{\frac{\sum_{n=1}^{33} \text{patentes}_{n,t}}{\sum \text{patentes}_t}} \quad (9)$$

IT_{it} é definido como um índice de especialização, em que $\sum_{n=1}^{33} \text{patentes}_{n,it}$ é soma do número de patentes recebidas pela patente i no tempo t nos campos n (33 subclasses de alta tecnologia) e $\sum \text{patentes}_{it}$ o número total de patentes da região i no período t. $\sum_{n=1}^{33} \text{patentes}_{n,t}$ é a soma das patentes de alta tecnologia de todas as regiões no tempo t e $\sum \text{patentes}_t$ a soma de todas as patentes no período t. As subclasses consideradas de alta tecnologia são as de computadores e equipamentos automatizados, aviação, engenharia genética, lasers, semicondutores e tecnologia de comunicação.

5) Q_{it} = índice de qualidade das patentes da região – razão entre o número de citações das patentes da região i sobre o total de patentes da região i.

$$Q_{it} = \frac{\text{Citações}_{it}}{\text{Patentes}_{it}} \quad (10)$$

6) $gapcit_{it}$ = média de intervalo de tempo entre o depósito e a citação da patente, no período t, da região i;

7) KH_{it} = nível de escolaridade da mão-de-obra da região no período t, definido como a proporção de mão-de-obra com terceiro grau;

8) $P\&D_{it}$ = proporção do PIB dos gastos em Pesquisa e Desenvolvimento da região i no período t;

9) $PIBpc_{it}$ = PIB per capita da região i no período t.

Por causa de limitações da base de dados, os intervalos de geração foram limitados em 10 anos, ou seja, toda patente gerada entre o ano t-1 a t-10 em relação ao ano da patente analisada é considerada uma geração anterior. Para tal, os anos de 1990 a 1999 serviram de base

para a construção da variável dependente, e não foram consideradas no painel, para que todas as regiões tenham o mesmo intervalo de geração para todos os anos.

Para a variável originalidade foram considerados todas as citações que a patente recebeu nos 5 anos seguintes a sua publicação. Esta limitação busca evitar viés nos valores. Assim, os anos de 2011 a 2015 foram excluídos do painel, para que todos os anos tenham o mesmo intervalo de construção da variável.

Uma abordagem bayesiana da análise foi também proposta, com objetivo de, por meio do método de amostragem via cadeias de Markov (MCMC) realizar uma estimação de amostras de uma distribuição por meio de amostragem iterativa das distribuições condicionais. O modelo linear generalizado bayesiano (família Bernoulli) se aproxima do modelo linear (OLS) assim como do GLM. Conforme a amostra aumenta, o resultado converge para os mesmos valores. Um dos benefícios da perspectiva bayesiana para qualquer análise é que permite fazer afirmações críveis sobre o intervalo de confiança. Intervalos críveis são similares a intervalos de confiança, mas na abordagem bayesiana o intervalo contém o verdadeiro parâmetro populacional. Por exemplo, um intervalo de confiança de 95% para um parâmetro é interpretado como há uma probabilidade de 95% que o verdadeiro parâmetro esteja incluso no intervalo. Isso ocorre pois o intervalo é baseado em informações da distribuição posterior.

4. Resultados e discussão

A variável dependente, Internalização, foi criada com o objetivo de aplicar uma medida baseada em autocitações de patentes como uma *proxy* para medir a capacidade de internalização de transbordamentos de conhecimento pelas regiões.

As estatísticas descritivas são apresentadas na Tabela 1. As variáveis PIB per capita, grau de educação e gastos em P&D foram padronizados com o objetivo de resolver possíveis vieses.

Tabela 1: Estatísticas descritivas das variáveis regionais, NUTS 2 (Período: 2000-2010)

Variável	Obs .	Média	Desvio-Padrão	Mínimo	Máximo
Internalização (Y_{it})	4175	0,374	7,692	0,000	341
Grau de educação (KH_{it})	3359	28,083	11,060	4,600	78,800
PIB per capita ($PIBpc_{it}$)	3457	28369	14689	658	168207
Gasto per capita em P&D ($P\&D_{it}$)	2313	1,742	1,269	0,100	10,240
Originalidade (\hat{O}_{it})	4175	0,892	0,369	0,000	2,000
Generalidade (\hat{G}_{it})	4175	0,252	0,382	0,000	2,000
Gap temporal ($gapcit_{it}$)	4175	2,112	1,702	0,000	5,000
Qualidade (Q_{it})	4175	0,212	0,520	0,000	21,500
Intensidade Tecnológica (IT_{it})	4175	0,003	0,009	0,000	0,157

Fonte: Elaboração própria.

A Figura 3 mostra o número de patentes e citações geradas por elas no período da análise. Com o passar dos anos, o impacto gerado em termos de transbordamentos das patentes aumentou muito. O maior fluxo de conhecimento e necessidade das inovações de se basearem em campos tecnológicos cada vez mais vastos, e, portanto, se integrarem mais com outras regiões, torna ainda mais evidente a necessidade de se analisar os diversos aspectos que concernem aos transbordamentos de conhecimento.

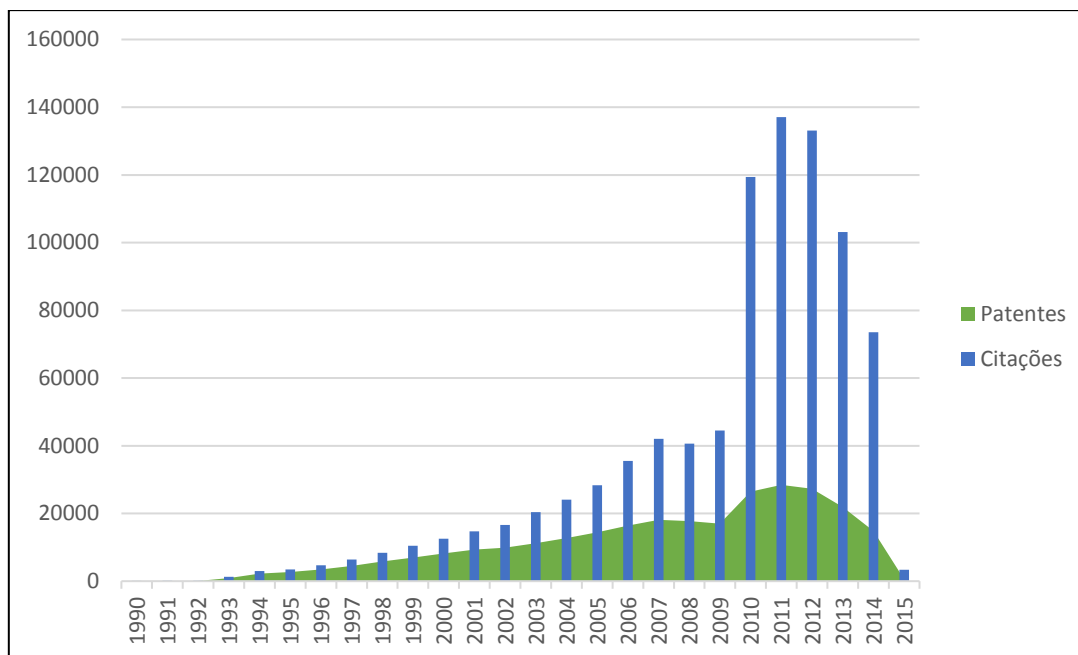


Figura 3: Número de patentes e citações delas originadas das NUTS2 no período de 1990 a 2014

Fonte: Elaboração própria a partir da base de dados da OECD.

Na Figura 4 o número de internalizações de transbordamentos de conhecimento, medido pela contagem de citações para trás de patentes que se basearam no conhecimento da região da patente originária, é apresentado em termos proporcionais. A região que mais realizou internalização, e que, portanto, mais se apropriou posteriormente dos seus transbordamentos de conhecimento e investimentos em P&D foi a região da Califórnia, nos EUA (US06), considerada uma das regiões mais inovadoras – o Vale do Silício. Outras regiões que também internalizaram mais conhecimento estão localizadas também nos EUA, Japão, Dinamarca, Alemanha, Itália, Coreia, Reino Unido, Suécia, China, Holanda e Canadá.

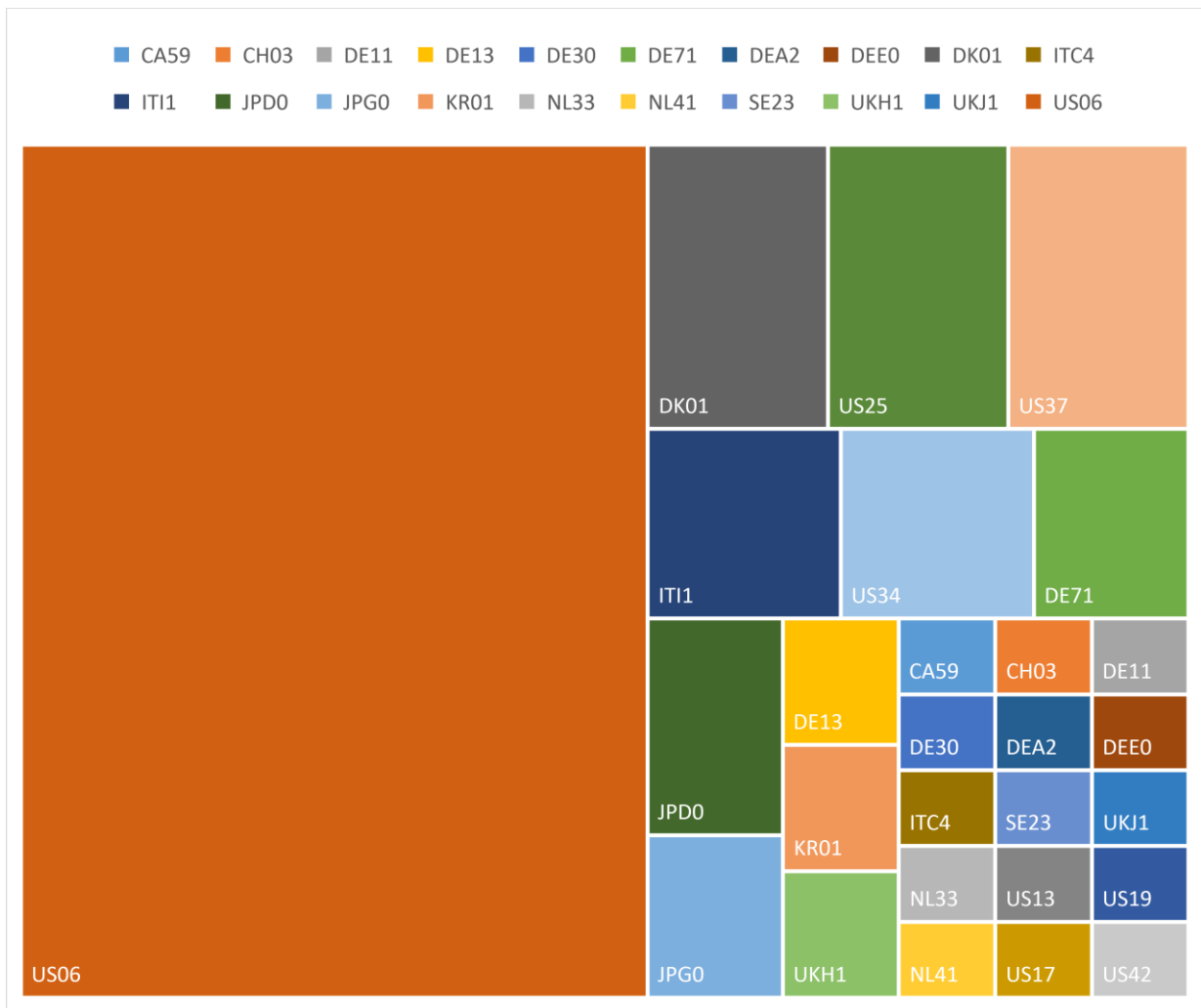


Figura 4: Proporção das autocitações indiretas por regiões que as realizaram
 Fonte: Elaboração própria.

A variável generalidade possui 63,35% de suas observações iguais a zero, o que pode causar um viés na estimação. Por isso, optou-se por transformá-la em um conjunto de *dummies*, de acordo com sua distribuição, mostrada na Figura 4. As variáveis G1, G2 e G3 são, portanto, as *dummies* para cada grau de generalidade da região, sendo G1 igual a 1 se a variável Generalidade assume valor 0 e 0, caso contrário; G2 igual a 1 se a variável Generalidade é maior que 0 e menor que 0,8333 e 0, caso contrário e G3 igual a 1 se a variável Generalidade tem valor maior que 0,8333.

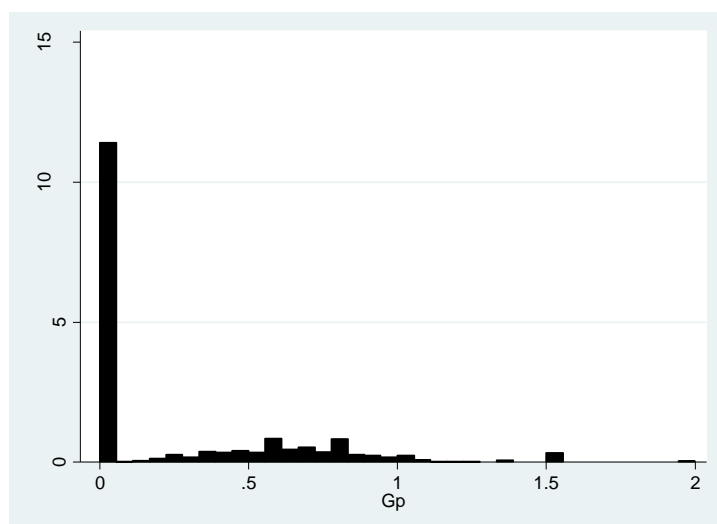


Figura 5: Histograma da variável Generalidade
 Fonte: Elaboração própria.

A variável de gastos em P&D sofre um problema de dados faltantes, o que afeta o número de observações do modelo. Por isso, o modelo foi testado com e sem sua inclusão, para verificar possíveis impactos causados pela perda de observações.

As estimações para os diferentes modelos para dados longitudinais são apresentadas na Tabela 2:

Tabela 2: Resultados do modelo logit de dados em painel para a variável internalização das NUTS2 no período de 2000-2010

Variável	<i>Pooled</i>	EA	EF
Grau de educação (KH_{it})	0,238 (0,292)	0,233 (0,342)	1,464 (1,787)
PIB per capita ($PIBpc_{it}$)	0,133 (0,171)	0,183 (0,225)	3,65* (2,169)
Gasto per capita em P&D ($P\&D_{it}$)	0,351* (0,186)	0,384* (0,223)	-2,162* (1,161)
Originalidade (\hat{O}_{it})	1,05 (1,448)	1,206 (1,616)	11,617 (7,568)
Generalidade (\hat{G}_{it})	1,522** (0,539)	1,596** (0,607)	2,193* (1,286)
Gap temporal ($gapcit_{it}$)	0,472* (0,253)	0,493* (0,278)	0,687 (0,569)
Qualidade (Q_{it})	0,157 (0,442)	0,219 (0,494)	0,386 (0,805)

Intensidade Tecnológica (IT_{it})	29,789*** (5,373)	33,180*** (9,098)	-1,776 (36,991)
Constante	-8,081*** (1,702)	-9,006*** (2,002)	
<hr/>			
Estatísticas			
N	2172	2172	168
Estatística Wald		35.11 (0,000)	23.50 (0,003)

Legenda: * $p < 0.10$; ** $p < 0.05$; *** $p < 0.001$ e erros-padrão entre parênteses.

Fonte: Elaboração própria.

O teste de Hausman (1978) verifica se a diferença entre o efeito fixo e efeito aleatório é estatisticamente significativa. A hipótese nula é de que a diferença entre os coeficientes estimados por efeitos fixos e efeitos aleatórios não é sistemática. O teste assume que o estimador de efeitos aleatórios é eficiente sob a hipótese nula e inconsistente sob a hipótese alternativa. Não se rejeita a hipótese nula, nesse caso, já que a estatística do teste é de 13,20 com p-valor de 0,105. Portanto, a estimação via efeitos aleatórios é mais eficiente e será o principal resultado reportado neste trabalho.

Os efeitos marginais estimados para um modelo Logit com efeitos aleatórios são apresentados na Tabela 3.

A intensidade tecnológica é a variável que mais afeta positivamente a probabilidade de que a região internalize transbordamentos do seu conhecimento tecnológico. A variável de intensidade tecnológica é construída como um índice de especialização setorial utilizando as áreas consideradas de alta intensidade tecnológica das patentes segundo a classificação da Eurostat. Essas áreas são as de computadores e equipamentos automatizados, aviação, engenharia genética, lasers, semicondutores e tecnologia de comunicação. A capacidade de internalizar conhecimento pode ser interpretada como uma vantagem competitiva obtida por meio da realização do esforço inovativo prévio das regiões. Ou seja, regiões que realizam maiores investimentos em P&D e possuem firmas mais inovadoras são mais capazes de internalizar conhecimento e são também as regiões com maior capacidade de absorção dos transbordamentos. As regiões especializadas nas áreas mais tecnológicas, têm, portanto, maiores chances de internalizar e possuem maior capacidade de absorção.

A teoria (Cohen e Levinthal, 1989) aponta para a maior capacidade de absorção de empresas que mais investem em P&D. O investimento em P&D teria não só a função de gerar novos conhecimentos e produtos, mas também a de aumentar a capacidade de assimilação das firmas de conhecimentos externos. Nesse sentido, quanto mais complexo o conhecimento buscado, maiores as exigências em termos de investimentos, e maior a atividade de P&D necessária (Cohen e Levinthal, 1990). Setores mais tecnológicos, em que a aprendizagem é mais complexa, teriam, portanto, maior de capacidade de absorção.

Em segundo lugar, o grau de generalidade da região é o que mais aumenta a probabilidade de que o evento ocorra.

Os investimentos em P&D e o gap temporal médio das citações das patentes também parecem afetar positivamente a probabilidade de ocorrência da internalização.

Tabela 3: Efeitos marginais do modelo logit com efeitos aleatórios para a variável internalização das NUTS2 no período de 2000-2010⁷

Variável	I	II	III	IV
Grau de educação (KH_{it})	0,233	0,263	0,188	0,25
PIB per capita ($PIBpc_{it}$)	0,183	0,301	0,191	0,293
Gasto per capita em P&D ($P\&D_{it}$)	0,384*		0,368*	
Originalidade (\hat{O}_{it})	1,206	0,964	1,065	0,882
Generalidade (\hat{G}_{it})	1,596***	1,439***		
Gap temporal ($gapcit_{it}$)	0,493*	0,346	0,545**	0,425**
Qualidade (Q_{it})	0,219	0,473	0,529	0,709*
Intensidade Tecnológica (IT_{it})	33,180***	36,978***	32,843***	36,925***
Dummy de generalidade (G1)			-0,741	-0,499
Dummy de generalidade (G3)			-0,284	-0,05
Observações	2172	2969	2172	2969

Legenda: *p<0.10; ** p<0.05; *** p<0.001.

Fonte: Elaboração própria.

O PIB per capita foi adicionado como uma forma de controle no modelo para características regionais, apresentando significância estatística em apenas um dos modelos propostos. Apesar de não parecer impactar a variável dependente, seu sinal positivo está de acordo com a intuição econômica em que se baseia esse trabalho, a saber, a de que regiões mais ricas possuem elevada capacidade de absorção tecnológica e maior capacidade de se apropriar de seus ganhos no processo inovativo.

O nível de escolaridade dos trabalhadores da região também não apresentou significância estatística. Como o número de regiões que internalizam conhecimento é pequeno, o nível de escolaridade não parece ser fator diferencial para explicar essa propensão.

Outro controle adicionado foi a variável que mede o *gap* temporal entre o tempo em que a patente surge – quando é publicada - e o momento em que é citada. O coeficiente significativo e positivo indica que a probabilidade de ocorrer internalização dos transbordamentos de conhecimento aumenta com o tempo, conforme esperado.

Um indicador de qualidade das patentes das regiões foi criado com base no impacto em termos de citações que, em média, as patentes das regiões geram. Seu coeficiente positivo indica que regiões mais dinâmicas e inovadoras realizam mais internalização de transbordamentos. Esse resultado corrobora a ideia de que quanto mais citações uma região recebe, mais conhecimento relevante ela gera e maior a sua capacidade inovativa, logo, sua capacidade de absorção de conhecimento.

As medidas de generalidade e originalidade desenvolvidas por Trajtenberg, Henderson e Jaff (1993) são amplamente utilizadas na literatura de inovação e tendem a ser positivamente correlacionadas com o número de citações feitas, no caso da originalidade, e recebidas, no caso da medida de generalidade. Ou seja, quanto mais citações são feitas ou geradas por essas

⁷ A estimativa bayesiana reportada no Anexo corrobora os resultados apresentados por meio da estimativa obtida com modelo logit tradicional. Tanto a significância quanto o impacto dos parâmetros em termos de coeficiente são muito próximos desta análise.

patentes, maior a abrangência de campos tecnológicos, seja das citações para frente ou para trás.

Patentes nos campos tecnológicos de computação e comunicações e biotecnologia possuem maior generalidade. O campo da computação e comunicações se encaixa na noção de tecnologia geral – que atende todas as outras e a biotecnologia desponta como um novo paradigma tecnológico para a indústria farmacêutica em geral. Campos mais tradicionais e menos inovativos como o setor mecânico e de medicamentos tradicionais possuem menores níveis de originalidade e generalidade (Hall *et al.*, 2001).

O grau de generalidade teve impacto positivo na internalização de conhecimento, enquanto o grau de originalidade não apresentou significância estatística. A medida capta quão diversas são as áreas citantes das patentes da região de origem. Regiões que recebem citações de várias áreas tecnológicas diferentes produzem conhecimento que pode ser usado por diversas outras áreas do conhecimento. As patentes de maior impacto – as que são mais citadas - demoram mais para serem concedidas, apresentam maior *gap* temporal médio de citação e são as que apresentam maior nível de generalidade. Em geral, as patentes mais gerais são as de produtos químicos e alguns outros setores de ciência básica (Hall e Trajtenberg, 2004).

Há uma correlação positiva entre originalidade e generalidade. Ou seja, patentes que se baseiam em várias áreas do conhecimento para serem desenvolvidas, também são as que espalham seu conhecimento por mais áreas. Nas estimações de Belenzon (2006), o grau de originalidade das patentes originárias das firmas apresenta coeficiente negativo em relação à capacidade de internalização, enquanto a generalidade é positiva. O autor não vê razões para supor previamente esse resultado e aponta para a existência de diferentes canais através dos quais a generalidade afeta a internalização positivamente e a originalidade afeta negativamente.

5. Conclusões

O objetivo do trabalho foi analisar a aplicação de uma medida baseada em autocitações de patentes ao nível de regiões como forma de medir a capacidade de internalizar transbordamentos de conhecimento dessas regiões. Os transbordamentos de conhecimento podem ser externalizados, na medida em que são direcionados para outras regiões, ou internalizados pela própria região. A hipótese é que regiões que se apropriam mais do próprio conhecimento gerado têm um maior estoque de conhecimento acumulado, o que a torna capaz de se apropriar mais dos ganhos de sua atividade de P&D.

Os resultados indicaram que as regiões com patentes de maior qualidade, mais impactantes em termos de citações, internalizam mais transbordamentos. O nível de renda não parece ter influência sobre a internalização do conhecimento, de um modo geral, o grau de educação da mão-de-obra também não parece afetar a variável dependente num modelo logístico. As regiões com maior *gap* temporal médio de citação, ou seja, com patentes com maior tempo médio entre seu depósito e citação, são também as que tendem a internalizar mais o conhecimento. A originalidade parece não afetar a internalização, já o grau de generalidade afeta positivamente. Além disso, regiões mais especializadas em áreas de alta tecnologia e que realizam mais esforços inovativos – medido pelos seus gastos em P&D – parecem ser as que mais tendem a internalizar conhecimento.

Nesse sentido, regiões produtoras de conhecimento básico tendem a internalizar mais conhecimento próprio, o que vai contra a hipótese e evidência empírica encontradas por Belenzon (2006). Entretanto, como essas regiões produtoras de ciência básica também são mais inovadoras, com sólida base tecnológica, elas parecem possuir maior habilidade para se apropriar dos ganhos do seu processo inovativo.

Alguns outros testes ainda podem ser feitos para explorar melhor a aplicação desta medida em contexto regional. Em trabalhos futuros, a questão da inovação sequencial como em Belenzon (2006) pode ainda ser mais explorada, com utilização de dados ao nível de patentes, para que se possa obter maior número de observações positivas para autocitações indiretas feitas pela região, tendo como base os dados da segunda geração em diante das patentes.

6. Referências bibliográficas

- Abreu, M., (2011). Absorptive Capacity in a Regional Context. In: Cooke, P., Asheim, B., Boschma, R., Martin, R., Schwartz, D., Tödting, F. (Eds.), *Handbook of Regional Innovation and Growth*. Cheltenham: Edward Elgar Publishing, pp. 211–221.
- Abreu, M., Grinevich, V, Kitson, M, Savona M, (2008). Absorptive capacity and regional patterns of innovation DIUS RR-08-11.
- Agrawal, A. K., Kapur, D., McHale, J., (2008). How Do Spatial and Social Proximity Influence Knowledge Flows? Evidence from Patent Data. *Journal of Urban Economics* 64, 258—269.
- Agrawal, A. K., Ajay, Iain Cockburn, e Carlos Rosell. (2010). “Not Invented Here? Innovation in Company Towns,” *Journal of Urban Economics*, 67, 78–89.
- Belenzon, S. (2006). *Basic Research and Sequential Innovation*.
- Cantwell, J., Iammarino, S. (2003). *Multinational corporations and European regional systems of innovation*. London, UK: Routledge.
- Caragliu, A., Nijkamp, P. (2012). The impact of regional absorptive capacity on spatial knowledge spillovers: the Cohen and Levinthal model revisited. *Applied Economics*, 44(11), 1363-1374.
- Caragliu, A., Nijkamp, P. (2008). “The Impact of Regional Absorptive Capacity on Spatial Knowledge Spillovers,” Tinbergen Institute Discussion Paper, Amsterdam, the Netherlands.
- Cassiman, B., R. Veugelers (2002), R&D Cooperation and spillovers: Some empirical evidence from Belgium, *American Economic Review* 44 (3), 1169-1184.
- Cohen W, Levinthal D. (1990). Absorptive capacity: a new perspective on learning and innovation. *Administrative Science Quarterly* 35: 128–152.
- Cohen, W.M., Levinthal, D.A. (1989), “Innovation and Learning: The Two Faces of R&D,” *The Economic Journal*, 99, 569–596.
- Criscuolo, P., R. Narula (2002). ‘A Novel Approach to National Technological Accumulation and Absorptive Capacity: Aggregating Cohen and Levinthal’, Discussion paper, MERIT Infonomics Research Memorandum 2002-016.
- Doloreux, D., Parto, S. (2005). Regional innovation systems: Current discourse and unresolved issues. *Technology in Society*, 27(2), 133–153.
- Doring, T., J. Schnellenbach (2006), ‘What Do We Know about Geographical Knowledge Spillovers and Regional Growth? A Survey of the Literature’, *Regional Studies*, 40.3, 375–395.
- Filippetti, A., Frenz, M., Ietto-Gillies, G. (2017). The impact of internationalization on innovation at countries’ level: the role of absorptive capacity. *Cambridge Journal of Economics*, 41(2), 413-439.
- Greene, W. H. (1997). *Econometric analysis* 3rd edition. New Jersey: Prentice-Hall International.
- Griliches, Z. (1979). Issues in assessing the contribution of research and development to productivity growth. *Bell Journal of Economics* 10:92–116.
- Grossman, G.M., Helpman E. (1992) “Innovation and growth in the global economy”, MIT Press: Cambridge (MA).
- Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). The NBER patent citation data file: Lessons, insights and methodological tools (No. w8498). National Bureau of Economic Research.
- Hall, B. H.; Trajtenberg, M. (2004). *Uncovering GPTs with patent data* (No. w10901). National Bureau of Economic Research.
- Hausman, J. (1978). Specification Tests in Econometrics. *Econometrica*, 46:1251-1272.
- Hu, A. GZ., Jaffe, A. B. (2003). "Patent citations and international knowledge flow: the cases of Korea and Taiwan." *International journal of industrial organization* 21.6, 849-880.

- Jaffe, A.B., Trajtenberg, M., Henderson, R. (1993): Geographic localization of knowledge spillovers as evidenced by patent citations. *Q. J. Econ.* 108(3), 577–98.
- Kneller, R. (2005), ‘Frontier Technology, Absorptive Capacity and Distance’, *Oxford Bulletin of Economics and Statistics*, 67(1), 1–23.
- Kneller, R., P. Stevens (2006), ‘Frontier Technology and Absorptive Capacity: Evidence from OECD Manufacturing Industries’, *Oxford Bulletin of Economics and Statistics*, 68(1), 1–21.
- Mancusi, M. L. (2008). “International Spillovers and Absorptive Capacity: A Cross-Country Cross-Sector Analysis Based on Patents and Citations,” *Journal of International Economics*, 76, 155–165.
- Mancusi, M.L. (2004). International spillovers and absorptive capacity: a cross-country, cross-sector analysis based on European patents and citations [Working paper]. URL: <http://sticerd.lse.ac.uk/dps/ei/ei35.pdf>.
- Maurseth, P.-B., Verspagen, B. (2002). “Knowledge spillovers in Europe: A patent citations analysis”. *Scandinavian Journal of Economics* 104, 531–545.
- Mowery, D.C., J.E. Oxley (1995), Inward technology transfer and competitiveness: The role of national innovation systems., *Cambridge Journal of Economics* 19, 67-93.
- Mukherji, N., Silberman, J., (2013). Absorptive capacity, knowledge flows, and innovation in US metropolitan areas. *J. Reg. Sci.* 53, 392–417, <http://dx.doi.org/10.1111/jors.12022>.
- Nagaoka, S., Motohashi, K., Goto, A. (2010) "Patent statistics as an innovation indicator." *Handbook of the Economics of Innovation* 2, 1083-1127.
- Narula, R. (2004) Understanding absorptive capacities in an ‘innovation systems’ context: consequences for economic and employment growth’. MERIT— infonomics research memorandum series 2004–003.
- Roper, S., Love, J.H., (2006). Innovation and regional absorptive capacity: the labor market dimension. *Ann. Reg. Sci.* 40, 437–447, <http://dx.doi.org/10.1007/s00168-006-0068-4>.
- Schmidt, T., (2005). What determines absorptive capacity. In: DRUID Summer Conference 2005 on Dynamics of Industry and Innovation: Organizations, Networks and Systems. Danish research unit for industrial dynamics, Copenhagen, Denmark.
- Veugelers, R., (1997). Internal R&D expenditures and external technology sourcing. *Research Policy* 26, 303–315.
- Von Tunzelmann, G.N., (2009). Regional Capabilities and Industrial Regeneration. In: Farschi, M., Janne, O., McCann, P. (Eds.), *Technological Change and Mature Regions: Firms, Knowledge and Policy*. Cheltenham, UK, pp. 11–28.
- Zahra, S.A., George, G., (2002). Absorptive capacity: a review, reconceptualization, and extension. *Academy of Management Review* 27 (2), 185–203.

ANEXO

Algoritmo resumido da variável dependente (Internalização)

-
1. Para cada patente i no ano z
 - 1.1 Para patente j citada por i (no ano $z-1$ até $z-10$)
 - 1.1.1 Identificar se região de j é igual a região da patente i (soma 1 se verdadeiro)
 - 1.1.2 para cada patente k citada por j (no ano $z-1$ até $z-10$)
 - 1.1.2.1 Identificar se região de k é igual à região da patente i (soma 1 se verdadeiro)
-

Tabela A: Estimativa de um modelo Bayesiano com aproximação Logit

Variáveis	Média	Desvio Padrão	MCSE	Mediana
Grau de educação (KH_{it})	0,211	0,293	0,031	0,206
PIB per capita ($PIBpc_{it}$)	0,161	0,167	0,022	0,173
Gasto per capita em P&D ($P\&D_{it}$)	0,373	0,205	0,031	0,381
Originalidade (\hat{O}_{it})	0,006	1,111	0,157	-0,033
Generalidade (\hat{G}_{it})	1.319	0,498	0,083	1.323
Gap temporal ($gapcit_{it}$)	0,5845	0,235	0,017	0,579
Qualidade (Q_{it})	0,187	0,410	0,056	0,207
Intensidade Tecnológica (IT_{it})	2.435	4.354	0,879	2,430
Constante	-7.452	1.317	0,151	-7,399
Observações	2.172			

Fonte: Elaboração própria.

Notas: Tamanho da amostra MCMC = 10.000; Número de iterações MCMC = 12.500.