

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA - PPGE

CIRLENE MARIA DE MATOS

**CANAIS DE DIFUSÃO DO CONHECIMENTO: EFEITO DA MOBILIDADE E DA
COLABORAÇÃO INTER-REGIONAL DE INVENTORES SOBRE A INOVAÇÃO
REGIONAL**

JUIZ DE FORA

2018

CIRLENE MARIA DE MATOS

**CANAIS DE DIFUSÃO DO CONHECIMENTO: EFEITO DA MOBILIDADE E DA
COLABORAÇÃO INTER-REGIONAL DE INVENTORES SOBRE A INOVAÇÃO
REGIONAL**

Tese apresentada ao Programa de Pós-Graduação em Economia da Faculdade de Economia da Universidade Federal de Juiz de Fora (PPGE/UFJF), como requisito parcial para a obtenção do título de Doutor em Economia. Área de concentração: Economia.

Orientador: Prof. Dr. Eduardo Gonçalves
Coorientador: Prof. Dr. Ricardo da Silva Freguglia

JUIZ DE FORA

2018

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Matos, Cirlene Maria de.

Canais de difusão do conhecimento : efeito da mobilidade e da colaboração inter-regional de inventores sobre a inovação regional / Cirlene Maria de Matos. -- 2018.

146 p. : il.

Orientador: Eduardo Gonçalves

Coorientador: Ricardo da Silva Freguglia

Tese (doutorado) - Universidade Federal de Juiz de Fora, Faculdade de Economia. Programa de Pós-Graduação em Economia, 2018.

1. mobilidade inter-regional de inventores. 2. coinvenção inter regional. 3. difusão inter-regional do conhecimento. 4. inovação regional. I. Gonçalves, Eduardo, orient. II. Freguglia, Ricardo da Silva, coorient. III. Título.

CIRLENE MARIA DE MATOS

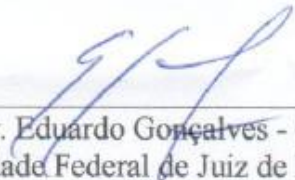
CANAIS DE DIFUSÃO DO CONHECIMENTO: EFEITO DA MOBILIDADE E DA COLABORAÇÃO INTER-REGIONAL DE INVENTORES SOBRE A INOVAÇÃO REGIONAL

Tese apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal de Juiz de Fora, como requisito parcial para a obtenção do grau de Doutor em Economia.


Área de concentração: Economia

Aprovada em 09 de Novembro de 2018

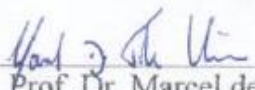
BANCA EXAMINADORA



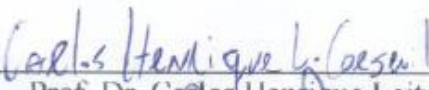
Prof. Dr. Eduardo Gonçalves - Orientador
Universidade Federal de Juiz de Fora (UFJF)



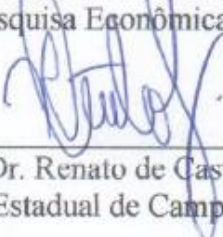
Prof. Dr. Ricardo da Silva Freguglia - Coorientador
Universidade Federal de Juiz de Fora



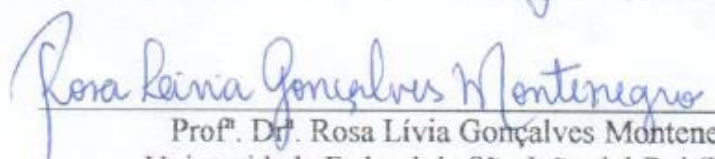
Prof. Dr. Marcel de Toledo Vieira
Universidade Federal de Juiz de Fora (UFJF)



Prof. Dr. Carlos Henrique Leite Corseuil
Instituto de Pesquisa Econômica Aplicada (IPEA)



Prof. Dr. Renato de Castro Garcia
Universidade Estadual de Campinas (UNICAMP)



Prof. Dr. Rosa Livia Gonçalves Montenegro
Universidade Federal de São João del-Rei (UFSJ)

Para
Margarida e Onofre
Bernardo e Rafael

AGRADECIMENTOS

Agradeço primeiramente a Deus por ter me abençoado com a oportunidade de progredir até aqui.

Agradeço aos meus pais, Margarida e Onofre, pelo amor incondicional, pelos ensinamentos valiosos, pelo respeito, pelo exemplo, por terem me apoiado em todos os momentos. A eles eu devo quem eu sou.

Agradeço a minha irmã, Auxiliadora, minha melhor amiga, meu refúgio, meu porto seguro.

Agradeço ao Bernardo, meu companheiro para a vida. Agradeço pelo apoio, por ter suportado a distância, pelas idas e vindas nas frias madrugadas entre Varginha e Juiz de Fora, por trazer alegria aos meus cinzentos finais de semana juiz foranos.

Agradeço ao meu filho, Rafael, por brindar essa árdua jornada com a alegria e o amor que sua chegada trouxe à minha vida.

Agradeço ao meu orientador, Eduardo Gonçalves, pelas valiosas orientações, pela boa vontade, pela paciência e pela compreensão com minhas falhas.

Agradeço ao meu coorientador, Ricardo Freguglia, pelas sugestões e por me despertar o gosto pela econometria.

Agradeço imensamente à Juliana Taveira pelas inúmeras vezes que me auxiliou na manipulação dos dados e no *Stata*. Sua ajuda foi essencial.

Agradeço a todos os meus colegas de doutorado pela convivência e pela solidariedade nos momentos difíceis. Agradeço especialmente ao Inácio Araújo e ao Marcus Vinícius por terem me auxiliado diversas vezes nos percalços dos programas de econometria, sempre de bom grado.

Agradeço ao Instituto Nacional de Propriedade Intelectual (INPI) e ao Ministério do Trabalho e Emprego (MTE) pela disponibilização dos dados, sem o quais essa tese não seria possível.

Agradeço ao PPGE pela oportunidade de cursar o Doutorado.

Agradeço a todos os professores pelos excelentes ensinamentos. Todos contribuíram de forma significativa para meu crescimento intelectual.

RESUMO

Essa tese teve como objetivo principal avaliar o papel da coinvenção e da mobilidade de inventores entre as microrregiões brasileiras na difusão do conhecimento e na inovação regional. Para isso foi utilizada a abordagem da função de produção do conhecimento regional modificada para incorporar as interações regionais promovidas por esses dois fenômenos. Aproveitando-se do instrumental da Econometria Espacial essas interações foram representadas na função de produção por meio de matrizes de pesos construídas para captar as relações bilaterais de coinvenção e mobilidade inter-regional. Assim, os objetivos secundários foram avaliar o efeito das distâncias física, cognitiva e institucional sobre a coinvenção e a mobilidade inter-regional de inventores, e construir matrizes de pesos que refletissem a interação regional decorrente da colaboração para inovar e da mobilidade dos inventores entre as regiões. Essas matrizes foram estimadas ano a ano utilizando-se modelos gravitacionais para a coinvenção e para a mobilidade e o estimador *PPML (Poisson Pseudo Maximum Likelihood)* com efeitos fixos da origem e do destino. A função de produção do conhecimento foi estimada por meio de um painel dinâmico utilizando-se *system gmm*. Os resultados mostraram que o espaço exerce influência relevante na decisão dos inventores de migrar para outras regiões e de colaborar para inovar. Quanto maior a distância geográfica, menor será a mobilidade e a coinvenção inter-regional. Esses achados corroboram a teoria de que os fluxos de conhecimento são localizados no espaço porque a mobilidade e as cooperações entre os agentes da inovação também são. A proximidade tecnológica entre as regiões de origem e de destino contribui para a decisão de migrar do inventor. Por outro lado, essa proximidade não afeta as relações de coinvenção inter-regional. Os resultados da função de produção do conhecimento mostraram que a mobilidade dos inventores entre as regiões brasileiras não promove a difusão do conhecimento e a inovação regional. Isso pode ser decorrente da importância da proximidade tecnológica entre origem e destino na escolha de migração do inventor. Essa proximidade pode gerar sobreposição de conhecimentos, impedindo que o conhecimento externo beneficie a inovação regional. Outra razão possível é a baixa capacidade de absorção regional, revelada pela não significância estatística do gasto industrial com Pesquisa e Desenvolvimento (PeD). Por outro lado, a coinvenção regional influencia significativamente a difusão do conhecimento entre as regiões e contribui para a inovação regional. Isso significa que a colaboração entre inventores de locais diferentes é um mecanismo de transmissão de conhecimento tácito. Os resultados permitiram corroborar o importante papel desse tipo de relação de mercado na transmissão do conhecimento, mostrando que nem toda difusão do conhecimento ocorre via transbordamentos puros. Além disso, a difusão do conhecimento por meio da coinvenção inter-regional também é localizada no espaço, pois os inventores tendem a colaborar com parceiros de regiões próximas geograficamente.

Palavras-chave: Patentes; Mobilidade inter-regional de inventores, Coinvenção inter-regional; Difusão inter-regional do conhecimento; Inovação regional.

ABSTRACT

This thesis had as main objective to evaluate the role of co-invention and of inventors mobility among the Brazilian regions in the diffusion of knowledge and regional innovation. For this purpose, the regional knowledge production function approach was used to incorporate the regional interactions promoted by these two phenomena. Taking advantage of the instruments of Spatial Econometrics, these interactions were represented in the production function by means of matrices of weights constructed to capture the bilateral relations of co-operation and interregional mobility. Secondary objectives were to evaluate the effect of physical, cognitive and institutional distances on co-invention and interregional inventors mobility, and to construct weight matrices that reflect the regional interaction resulting from collaboration to innovate and the mobility of inventors between regions. These matrices were estimated year by year using gravitational models for co-invention and for mobility and the Poisson Pseudo Maximum Likelihood (PPML) estimator with fixed effects of origin and destination. The knowledge production function was estimated by means of a dynamic panel using system gmm. The results showed that space exerts a relevant influence on the decision of the inventors to migrate to other regions and collaborate to innovate. The greater the geographical distance, the less mobility and inter-regional co-operation. These findings corroborate the theory that knowledge flows are located in space because the mobility and cooperation among the agents of innovation are localized in space as well. The technological proximity between the source and destination regions contributes to the inventor's decision to migrate. On the other hand, this proximity does not affect inter-regional co-operation relationships. The results of the knowledge production function showed that the mobility of inventors among the Brazilian regions does not promote the diffusion of knowledge and regional innovation. This may be due to the importance of the technological proximity between origin and destination in the choice of the inventor's migration. This proximity can generate overlapping of knowledge, preventing external knowledge from benefiting regional innovation. Another reason is the low regional absorption capacity, revealed by the non-statistical significance of industrial Research and Development (R&D). On the other hand, regional co-operation significantly influences the diffusion of knowledge between regions and contributes to regional innovation. This means that collaboration between inventors from different locations is a tacit knowledge transmission mechanism. The results allowed to corroborate the important role of this type of market relation in the transmission of knowledge, showing that not all diffusion of knowledge occurs through pure overflows. In addition, the diffusion of knowledge through interregional co-operation is also localized in space as inventors tend to collaborate with partners from geographically close regions.

Keywords: Patents; Inter-regional inventor mobility; Inter-regional co-operation to innovate; Inter-regional knowledge diffusion; Regional innovation.

LISTA DE FIGURAS

| | |
|---|----|
| Figura 01: Principais etapas da construção da base de dados..... | 60 |
| Figura 02: Mapa de <i>clusters</i> para o estoque patentes entre 2000 e 2011, por 100 mil habitantes..... | 76 |

LISTA DE TABELAS

| | |
|---|----|
| Tabela 01: Tipos de patentes dos inventores brasileiros, de 2000 a 2011..... | 60 |
| Tabela 02: Identificação do inventor como Pessoa Física e Pessoa Jurídica, 2000 a 2011 | 61 |
| Tabela 03 Identificação do inventor após recuperação de CPF/CNPJ a partir dos dados dos depositantes – 2000 a 2011..... | 62 |
| Tabela 04: Total de processos com informação de CPF/CNPJ, entre 2000 e 2011..... | 62 |
| Tabela 05: Ampliação do banco de dados dos inventores por meio de busca na internet..... | 63 |
| Tabela 06: Total de patentes com pelo menos um inventor identificado pelo CPF – 2000 a 2011..... | 63 |
| Tabela 07: Distribuição das patentes por ano – 2000 a 2011..... | 64 |
| Tabela 08: Total de patentes por inventor entre 2000 e 2011..... | 64 |
| Tabela 09: Regiões com maior número de patentes entre 2000 e 2011..... | 67 |
| Tabela 10: Taxa de migração e inventores que migraram pelo menos uma vez entre 2000 e 2011..... | 68 |
| Tabela 11: Número de migrações por inventor entre 2000 e 2011..... | 68 |
| Tabela 12: Migrações entre microrregiões e produtividade média dos inventores, entre 2000 e 2011..... | 69 |
| Tabela 13: Entrada e saída total de inventores das 20 maiores microrregiões patenteadoras entre 2000 e 2011..... | 70 |
| Tabela 14: Estados com maior estoque de patentes entre 2000 e 2011: entrada e saída total de inventores..... | 71 |
| Tabela 15: Número de inventores por patente, de 2000 a 2011..... | 72 |
| Tabela 16: Número de patentes resultantes de coinvenção local e inter-regional, de 2000 a 2011..... | 72 |
| Tabela 17: Número de regiões envolvidas na geração de uma patente, 2000 a 2011..... | 73 |
| Tabela 18: de microrregiões com patentes de coinvenção inter-regional no período 2000 a 2011..... | 73 |
| Tabela 19: Regiões com maior estoque de patentes de coinvenção inter-regional - 2000 a 2011..... | 74 |
| Tabela 20: Coeficiente I de Moran para patenteamento per capita no Brasil..... | 75 |
| Tabela 21: Número de microrregiões patenteadoras, quantidade de patentes e de classes | |

| | |
|--|-----|
| tecnológicas: 2000 a 2011..... | 80 |
| Tabela 22: Resultados do modelo gravitacional: mobilidade de inventores entre as microrregiões brasileiras (2001 a 2010) - estimações por PPML sem EF..... | 88 |
| Tabela 23: Resultados do modelo gravitacional: mobilidade de inventores entre as microrregiões brasileiras (2001 a 2010) - estimações por PPML com EF..... | 89 |
| Tabela 24: Resultados do modelo gravitacional: colaborações entre as microrregiões brasileiras (2001 a 2010) estimações por PPML sem EF..... | 91 |
| Tabela 25: Resultados do modelo gravitacional: colaborações entre as microrregiões brasileiras (2001 a 2010) - estimações por PPML com EF..... | 92 |
| Tabela 26: Função de Produção do Conhecimento: estimações por <i>System GMM</i> – 2001 a 2010..... | 94 |
| Tabela 27: Função de produção do conhecimento com defasagem espacial: estimações por <i>System GMM</i> (2001 a 2010)..... | 104 |
| Tabela 28: Função de produção do conhecimento para regiões patenteadoras: estimações por <i>System GMM</i> (2001 a 2010)..... | 112 |
| Tabela A1: Mobilidade de inventores da região i para a região j: 2000 a 2011..... | 128 |
| Tabela A2: Estatísticas descritivas para a mobilidade de inventores: 2000 a 2011..... | 129 |
| Tabela A3: Estatísticas descritivas para a colaboração inter-regional: 2000 a 2011..... | 130 |
| Tabela A4: Coinvenções entre as regiões i e j: 2000 a 2011..... | 131 |
| Tabela A5: Resultados do modelo gravitacional: mobilidade de inventores entre as microrregiões brasileiras (2001 a 2010) estimações por MQO – com EF..... | 132 |
| Tabela A6: Resultados do modelo gravitacional: coinvenção entre as microrregiões brasileiras (2001 a 2010) - estimações por MQO – com EF..... | 134 |
| Tabela A7: Função de Produção do Conhecimento: estimações por <i>System GMM</i> – 2001 a 2010 – lags 2 a 5..... | 136 |
| Tabela A8: Função de produção do conhecimento com defasagem espacial: estimações por <i>System GMM</i> (2001 a 2010) – lags de 2 a 5..... | 139 |
| Tabela A9: Função de produção do conhecimento para regiões patenteadoras: estimações por <i>System GMM</i> (2001 a 2010) – lags (2 a 5)..... | 143 |

SUMÁRIO

| | |
|--|-----|
| 1. INTRODUÇÃO..... | 13 |
| 2. REVISÃO DE LITERATURA..... | 18 |
| 2.1 O uso de patentes como <i>proxy</i> para inovação..... | 18 |
| 2.2 Fluxos localizados do conhecimento..... | 21 |
| 2.3 Canais de transmissão do conhecimento: o papel da mobilidade e da coinvenção inter-regional..... | 27 |
| 3. ESTRATÉGIA METODOLÓGICA..... | 40 |
| 3.1 Função de produção do conhecimento regional..... | 40 |
| 3.2 Interações regionais de conhecimento: uma abordagem gravitacional..... | 43 |
| 3.2.1 Interações de conhecimento por meio da mobilidade inter-regional de inventores..... | 43 |
| 3.2.2 Interações de conhecimento por meio das coinvenções inter-regionais..... | 46 |
| 3.3 Abordagem econométrica..... | 48 |
| 3.3.1 Estimação dos modelos gravitacionais..... | 48 |
| 3.3.2 Estimação da função de produção do conhecimento regional..... | 51 |
| 4. ORGANIZAÇÃO DO BANCO DE DADOS E ANÁLISE DESCRITIVA..... | 59 |
| 4.1 Identificando os inventores..... | 59 |
| 4.2 Identificando a localização do inventor..... | 64 |
| 4.3 Análise descritiva..... | 66 |
| 4.3.1 Analisando a migração dos inventores entre as regiões..... | 67 |
| 4.3.2 Analisando as colaborações inter-regionais..... | 71 |
| 4.4 Análise descritiva espacial..... | 75 |
| 4.5 Descrição das variáveis dos modelos gravitacionais..... | 77 |
| 4.6 Descrição das variáveis utilizadas na Função de Produção do Conhecimento..... | 83 |
| 5. ANÁLISE DOS RESULTADOS..... | 86 |
| 5.1 Resultados dos modelos gravitacionais..... | 86 |
| 5.2 Resultados da Função de Produção do Conhecimento Regional (FPCR)..... | 93 |
| 6. CONSIDERAÇÕES FINAIS..... | 116 |
| REFERÊNCIAS..... | 121 |
| APÊNDICE..... | 128 |

1 INTRODUÇÃO

O progresso técnico é considerado o motor do crescimento econômico de longo prazo de nações e regiões (SOLOW, 1957; ROMER, 1986). Por ser um processo inerentemente cumulativo, o progresso técnico baseia-se na combinação de elementos do conhecimento pré-existente para a geração de um conhecimento novo que seja economicamente útil. O acesso a estes elementos pré-existentes por parte dos agentes envolvidos no processo produtivo é essencial para a determinação da capacidade inovativa regional e do crescimento econômico.

A difusão do conhecimento no espaço implica que o conhecimento gerado em uma região pode ser usado por terceiros em outro local, contribuindo para o desenvolvimento de inovações e o crescimento econômico regional. Assim, o fluxo de conhecimento é um elemento central da teoria do crescimento endógeno, pois sua difusão induz o crescimento econômico (ROMER, 1990).

Entretanto, os fluxos de conhecimento são espacialmente limitados (JAFFE, 1986; JAFFE *et al.*, 1993; BOTAZZI; PERI, 2003; MURATA *et al.*, 2014). A premissa por trás da localização dos fluxos de conhecimento é a existência do conhecimento tácito, cuja transmissão demanda contatos face a face. Essa necessidade de interações presenciais para que o conhecimento tácito seja transmitido confere à proximidade geográfica papel determinante na viabilização dessa difusão. Assim, os fluxos do conhecimento são espacialmente localizados no sentido de que seu transbordamento diminui com o aumento da distância geográfica (PACI; USAI, 2009).

A localização do fluxo de conhecimento pode decorrer de transbordamentos puros localizados do conhecimento ou de mecanismos de difusão baseados em transações de mercado. Diferenciar essas duas formas de transmissão do conhecimento tem sido uma questão central na literatura da geografia da inovação (LISSONI; MIGUÉLEZ, 2014; MIGUÉLEZ; MORENO, 2015).

A abordagem da função de produção do conhecimento regional é a mais utilizada para investigar o processo de produção do conhecimento e sua localização espacial (GREUNZ, 2003; MIGUÉLEZ; MORENO, 2013a; 2013b; MORENO *et al.*, 2005; CHARLOT *et al.*, 2015; LISSONI; MIGUÉLEZ, 2014). A função de produção estabelece uma correlação entre os insumos e o produto do conhecimento, mas sem distinguir os canais pelos quais o conhecimento flui ou transborda. Embora seja possível identificar a existência e a importância das externalidades do conhecimento, não é possível discriminar entre externalidades

tecnológicas puras e externalidades pecuniárias. Sua principal limitação é não distinguir os canais pelos quais o conhecimento é realmente transmitido. Embora tradicionalmente a dependência espacial da inovação identificada por meio dessa abordagem seja interpretada como transbordamentos puros do conhecimento, não há nenhuma evidência na abordagem tradicional que garanta que sejam de fato externalidades puras. A função de produção do conhecimento regional é uma caixa preta que não diz nada sobre a forma como o conhecimento é de fato transmitido entre os agentes (BRESCHI, 2011).

A investigação do transbordamento de conhecimento técnico entre regiões por meio da função de produção do conhecimento regional assume implicitamente que a geografia é uma plataforma suficiente para a difusão espacial do conhecimento. A utilização de matrizes de pesos espaciais baseadas na distância ou na contiguidade para estimar o efeito do conhecimento das regiões vizinhas sobre a inovação local pressupõe que basta haver proximidade espacial para que o conhecimento flua de uma região para outra. Entretanto, estar próximo geograficamente não implica que o conhecimento gerado na região vizinha estará automaticamente disponível para uso na inovação local. O conhecimento, principalmente o tácito, não viaja pelo ar. É necessário que ocorram interações relacionadas ao conhecimento para que seu fluxo entre as regiões se concretize.

Os canais de difusão baseados no mercado são os que mais contribuem para a disseminação do conhecimento imediatamente relevante para o processo de inovação (BRECHI; LISSONI, 2009). Entre esses canais a mobilidade de inventores e a coinvenção inter-regional são considerados os mais importantes na promoção da difusão do conhecimento no espaço (MIGUÉLEZ; MORENO, 2015). Os inventores são os agentes diretamente envolvidos no processo de inovação, sendo os repositórios do conhecimento tácito relevante para inovar. Ao mesmo tempo representam também os veículos de transmissão desse conhecimento no espaço, pois o conhecimento viaja com quem o domina (BRESCHI; LISSONI, 2001). A mobilidade dos inventores e as colaborações para inovar permitem a interação entre os agentes do conhecimento, e constituem canais relevantes de difusão do conhecimento. Esta difusão é essencial para fomentar o desenvolvimento de inovações e, conseqüentemente, o crescimento econômico regional. Nesse sentido, os fluxos de conhecimento são localizados porque a mobilidade dos agentes envolvidos diretamente na geração do conhecimento, bem como as relações interpessoais formais voltadas para inovação, tendem a ser regionais por natureza (MIGUÉLEZ; MORENO, 2013a).

Entender o papel da mobilidade espacial dos inventores e de suas colaborações inter-regionais sobre a difusão do conhecimento técnico contribui para compreender seu efeito sobre a performance inovativa local e sobre o crescimento econômico regional. Existem na literatura evidências de uma forte ligação entre redes e mobilidade inter-regionais de inventores, de um lado, e inovação regional do outro (MIGUÉLEZ; MORENO, 2015).

As interações de conhecimento viabilizadas por estes fenômenos constituem meios específicos pelos quais o conhecimento pode passar de um local para outro. Em contraposição às matrizes de pesos espaciais, a mobilidade de inventores e a colaboração para inovar representam interações que possibilitam a difusão do conhecimento. Estas interações não são limitadas necessariamente pela proximidade espacial, embora sejam influenciadas por ela. Dessa forma, matrizes de fluxos bilaterais que representem essas interações de conhecimento podem ajudar a compreender a importância destas relações para o acesso ao conhecimento externo e para a inovação regional.

O objetivo dessa tese é investigar o papel da mobilidade espacial dos inventores e de suas colaborações inter-regionais para inovar sobre a difusão e a criação de conhecimento técnico entre as regiões brasileiras e, conseqüentemente, sobre a performance da inovação regional. Para isso, propõe-se uma abordagem em duas etapas, similar ao procedimento adotado por Miguélez e Moreno (2013a).

Na primeira etapa pretende-se investigar os determinantes da mobilidade inter-regional dos inventores brasileiros e de suas colaborações inter-regionais para inovar. Nesse processo serão estimadas duas matrizes de pesos regionais, uma que reflita as interações regionais por meio da mobilidade bilateral de inventores e outra que reflita estas interações por meio das colaborações bilaterais para inovar. Essas matrizes serão construídas por meio de modelos gravitacionais estimados por *Poisson Pseudo Maximum Likelihood* (PPML).

O segundo estágio consiste na estimação de uma função de produção do conhecimento regional transformada para incluir as matrizes de interação obtidas na etapa anterior. Será investigado o papel da mobilidade dos inventores e da colaboração para inovar na difusão do conhecimento entre as regiões brasileiras e sua importância relativa diante de outras formas de difusão do conhecimento representadas pela defasagem espacial.

Diferentemente de Miguélez e Moreno (2013a) que estimaram a função de produção do conhecimento por meio do estimador de Efeito Fixo, nessa tese utilizaremos um painel dinâmico estimado por *system gmm*. Esse procedimento possui duas vantagens principais em relação ao adotado pelos referidos autores: contempla o caráter *path dependent* da inovação

regional e permite corrigir a endogeneidade das variáveis explicativas. O conhecimento é cumulativo e sua construção se dá por empilhamento, de forma que o potencial regional para inovar depende diretamente da performance inovativa anterior. A mobilidade dos inventores e as colaborações inter-regionais possuem relações de simultaneidade com a inovação regional, resultando na endogeneidade dessas variáveis. Negligenciar essas duas características da função de produção do conhecimento pode comprometer a confiabilidade dos resultados.

A contribuição principal da tese será compreender a importância da mobilidade de inventores e da coinvenção sobre o processo de inovação regional no Brasil. O entendimento do papel destes mecanismos na difusão do conhecimento técnico permitirá também entender a importância dos transbordamentos puros na localização espacial do conhecimento e no processo de inovação. As evidências encontradas permitirão confirmar, refutar ou qualificar o papel primordial que a literatura tem reservado às externalidades puras na localização espacial dos fluxos de conhecimento e do processo de inovação.

Outra contribuição será a construção das matrizes de pesos regionais que reflitam as interações de conhecimento entre as regiões. A difusão do conhecimento não é automática, não basta estar lá, é necessário haver canais ativos pelos quais ele possa transitar. A utilização das matrizes de interações do conhecimento, estimadas a partir da mobilidade e da coinvenção, no âmbito da função de produção do conhecimento permitirá considerar explicitamente estes canais. Embora existam estudos sobre a mobilidade de capital humano, até onde sabemos não existem para o Brasil estudos sobre a mobilidade de inventores. Além disso, a utilização das matrizes de mobilidade e de coinvenção inter-regional, para ponderar os fluxos de conhecimento externo, também são inéditas para o Brasil.

Serão utilizadas as informações da Base de Dados sobre Propriedade Intelectual para fins Estatísticos (BADEPI v1.0) fornecidas pelo Instituto Nacional de Propriedade Intelectual (INPI) referentes ao período de 2000 a 2011. Essa base identifica as patentes e seus respectivos inventores de forma única, permitindo observar as relações de coinvenção. Devido à grande proporção de *missings* sobre a localização do inventor, essa informação foi combinada com a RAIS-MIGRA (Relação Anual de Informações Sociais do Ministério do Trabalho e Emprego) para identificar a mobilidade dos inventores e as colaborações inter-regionais. Dessa forma, obteve-se um banco de dados inédito que permitiu ampliar as informações sobre a localização dos inventores ao longo do período e, conseqüentemente, melhorar a qualidade dos resultados obtidos.

Essa tese está organizada em 5 (cinco) capítulos além dessa introdução. O capítulo 2 (dois) discute a literatura teórica e empírica sobre os fluxos de conhecimento e seus canais de transmissão. O capítulo 3 (três) apresenta a estratégia metodológica, contemplando os modelos teóricos e os procedimentos econométricos. O capítulo 4 (quatro) explica o processo de organização do banco de dados, apresenta uma análise descritiva dos mesmos e define as variáveis utilizadas nos modelos. O capítulo 5 (cinco) traz a análise dos resultados e o capítulo 6 (seis) faz as considerações finais.

2 REVISÃO DE LITERATURA

2.1 O uso de patentes como *proxy* para inovação

As invenções são novas combinações do conhecimento existente que resultam de uma performance mental acima da média e criam meios potencialmente úteis na produção econômica, que permitem reduzir o custo de produção de bens existentes ou produzir novos bens (KUZNETS, 1962). A invenção possui uma utilidade potencial, caso contrário, o inventor e/ou a firma não dedicariam esforços em seu desenvolvimento. Mas, os inventores não são gênios nem possuem poderes especiais, o que é comum a eles é o grande domínio de tecnologias já existentes sob a forma de operações, componentes, estruturas e métodos. Os inventores estão embebidos na prática e na teoria acerca do princípio ou do fenômeno que irão utilizar (ARTHUR, 2007).

A invenção é diferente da inovação, pois enquanto a primeira se refere a conhecimento novo ou nova descoberta, a segunda é a introdução desta novidade no mercado (SCHUMPETER, 1939; ARTHUR, 2007). Apenas quando a invenção se transforma em uma inovação é que ela gera várias melhorias futuras (KUZNETS, 1962). Seja por falta de utilidade prática ou de novidade, ou por inviabilidade técnica ou econômica, muitas invenções não se concretizam em inovações, e como tal têm pouca ou nenhuma relevância econômica. As inovações, por outro lado, frequentemente geram efeitos diretos sobre o crescimento econômico, de forma que o crescimento regional e nacional depende mais do sucesso na comercialização de novos produtos e processos do que da invenção (CARLINO; KERR, 2015).

A inovação resulta da interação entre o capital humano e dos transbordamentos do conhecimento, o que cria aprendizagem cumulativa. As inovações em uma indústria tendem a se espalhar para as demais, de forma que seu desenvolvimento tem impactos muito mais amplos sobre a economia como um todo do que apenas dentro das próprias firmas inovadoras (FAGGIAN; MCCANN, 2009), gerando progresso técnico e crescimento econômico.

Como a mensuração direta da inovação é difícil, a patente é uma *proxy* para o conhecimento novo economicamente útil (JAFFE, 1989). Por tal motivo, as patentes têm sido cada vez mais utilizadas como uma *proxy* para a inovação. Entretanto, a patente não é a única e nem a principal barreira para evitar o acesso de terceiros ao conhecimento. Existem outros mecanismos de proteção tipicamente mais importantes que a patente, tais como a velocidade

no desenvolvimento do produto, o segredo, as vantagens do aprendizado, esforços de vendas e serviços (LEVIN *et al.*, 1987), design complexo e controle de capacidades complementares (NAGAOKA *et al.*, 2010).

Há várias razões pelas quais as firmas preferem usar estes mecanismos para se apropriar do retorno da inovação. Levin *et al.* (1987) observaram que a possibilidade de *inventing around* por parte dos concorrentes é a principal delas. Outros motivos são a quantidade de informação revelada no requerimento de patente e a dificuldade de demonstrar a novidade (COHEN *et al.*, 2000), bem como o perigo da imitação competitiva e os custos de obtenção e manutenção da patente (KUZNETS, 1962; COHEN *et al.*, 2000).

Para ser patenteada, uma invenção tem que ser relevante, viável tecnicamente e não ser uma extensão óbvia da arte anterior. A patente representa uma quantidade mínima de invenção que passou pelo teste de novidade, realizado pelo exame do escritório de patentes, e implica que a invenção possui importância estratégica e/ou potencial de mercado, uma vez que o inventor e sua firma dedicaram esforço e recursos em seu desenvolvimento (KUZNETS, 1962; GRILICHES, 1990). É a concessão da patente, e não sua solicitação, que garante a novidade e o teste de viabilidade técnica. Portanto, a concessão de patentes é o melhor indicador de invenções com valor econômico potencial, ou seja, é a melhor medida do produto da atividade inventiva (KUZNETS, 1962).

O uso das patentes como indicadores da atividade inovativa também possui limitações: i) A patente reflete apenas o primeiro estágio da inovação, ou seja, a invenção; ii) Não necessariamente a inovação será desenvolvida no mesmo local em que a invenção foi gerada (CARLINO; KERR, 2015); iii) O valor econômico médio da patente difere entre indústrias, ao longo do tempo e no espaço. Existe grande variabilidade tanto em sua importância técnica quanto econômica: muitas patentes refletem apenas pequenas melhorias, sendo, portanto, de baixo valor econômico, enquanto algumas são muito valiosas (KUZNETS, 1962; GRILICHES, 1990; NAGAOKA *et al.*, 2010); iv) Nem todas as invenções factíveis e potencialmente lucrativas são patenteadas (KUZNETS, 1962; GRILICHES, 1979, 1990; NAGAOKA *et al.*, 2010); v) Muitas patentes são usadas de forma estratégica, apenas para impedir a concorrência, viesando as medidas de inovação¹ (NAGAOKA *et al.*, 2010; CARLINO; KERR, 2015)

¹ Muitas firmas recorrem ao patenteamento estratégico, realizando patentes de substitutos próximos de sua patente principal como forma de evitar que a concorrência possa desenvolver e patentear invenções a partir de sua patente principal (CARLINO; KERR, 2015; NAGAOKA *et al.*, 2010).

Além disso, o patenteamento é endógeno às oportunidades de mercado e ao tamanho dos ativos complementares de uma firma, dado que a propensão a patentear de uma firma aumenta com a lucratividade da patente (NAGAOKA *et al.*, 2010). A propensão a patentear também varia entre setores e entre firmas. Por exemplo, a patente é mais eficiente na indústria farmacêutica e na indústria de equipamentos mecânicos simples. De forma geral, a eficácia relativa da patente é maior para produtos de tecnologia discreta ou simples do que para produtos complexos que envolvem vários elementos patenteáveis separadamente (LEVIN *et al.*, 1987). Novos elementos químicos na indústria farmacêutica podem ser protegidos por uma patente e não há como os concorrentes praticarem o *inventing around*. Por outro lado, no caso de equipamentos eletrônicos são várias patentes envolvidas com um único bem e existem muitas tecnologias substitutas para cada patente. Neste tipo de produto, a eficácia de uma patente é limitada quando comparada ao segredo industrial (NAGAOKA *et al.*, 2010). A propensão ao uso de patentes aumenta com o tamanho da firma, pois o custo fixo das patentes é diluído pelo maior número de patentes e, firmas pequenas podem ter dificuldades em fazer cumprir seus direitos de patente, devido aos custos legais (COHEN *et al.*, 2000).

A despeito destas limitações, o fato de as patentes resultarem diretamente do processo inventivo, ao contrário dos gastos com Pesquisa e Desenvolvimento (PeD) (CARLINO; KERR, 2015), e de serem concedidas apenas às invenções tecnicamente viáveis e com potencial econômico as tornam melhores indicadores da atividade inovativa. Seu uso é amplamente difundido na literatura da área, adquirindo especial importância neste trabalho que avalia a mobilidade de inventores e a coinvenção (JAFFE, 1986, 1989; JAFFE *et al.*, 1993; GREUNZ, 2003; MORENO *et al.*, 2005; BRESCHI; LISSONI, 2005; 2009).

Os documentos de patentes contêm várias informações relevantes sobre o processo de inovação. Cada patente é identificada por um código único, permitindo que elas sejam agregadas por inventor e/ou região sem que haja problemas de dupla contagem. A data de registro da patente no escritório de patentes indica o período em que a invenção foi criada, permitindo ao pesquisador acompanhar a evolução do desempenho inovativo do inventor e/ou da região ao longo do tempo.

A patente apresenta a relação com os nomes e endereço de residência de todos os seus inventores. Essas informações permitem que as patentes sejam atribuídas ao local em que as invenções foram desenvolvidas, permitindo que o pesquisador verifique o desempenho inovativo regional por meio da agregação dessas patentes de acordo com o nível regional de interesse da pesquisa. As patentes são consideradas um produto do PeD e um indicador da

inovação, sendo largamente utilizadas na literatura como *proxy* para o potencial inovativo regional (MIGUÉLEZ; MORENO, 2013a; 2013b; 2015; GRINZA; QUATRARO, 2018).

Essas informações permitem a identificação de relações de coinvenção, em nível local e inter-regional, bem como permitem também identificar a mobilidade dos inventores ao longo do tempo. Esses dados possibilitam, por exemplo, realizar estudos sobre a produtividade do inventor, padrões e efeitos regionais da mobilidade do inventor e padrões e efeitos locais das coinvenções sobre a produtividade do inventor e/ou da região (BRESCHI; LISSONI, 2003; 2005; 2009; MORESCALCHI *et al.*, 2015; GORIN, 2016).

As patentes possuem códigos da *International Patent Classification (IPC)* que relacionam as tecnologias produzidas e utilizadas pela patente. Esses códigos permitem analisar as competências tecnológicas em nível regional.

Além de serem facilmente acessíveis, os escritórios de patentes seguem regras nacionais e internacionais para o registro das patentes de forma que as informações são padronizadas, garantindo a confiabilidade dos dados. Esses dados são perenes, pois uma vez registrados no escritório de patentes, a informação permanece a mesma ao longo do tempo, não estando sujeita a modificações.

Portanto, a produção regional de patentes é um indicador confiável da performance inovativa regional. Como tal, o patenteamento per capita tem sido utilizado para construir a variável dependente “produto do conhecimento” na função de produção do conhecimento regional (FPCR). A FPCR é uma das principais abordagens utilizadas na literatura empírica para analisar o alcance espacial da difusão do conhecimento. Essa difusão tende a perder intensidade com o aumento da distância geográfica, resultando na localização espacial dos fluxos de conhecimento, como será discutido na próxima seção.

2.2 Fluxos localizados do conhecimento

As investigações empíricas sobre inovações foram impulsionadas pela adaptação da função de produção neoclássica ao processo de produção de inovações. Originalmente proposta por Griliches (1979) para estudos ao nível da firma, a função de produção do conhecimento relaciona os insumos do conhecimento ao produto do conhecimento:

$$I_i = \alpha (PeD)_i^\beta \cdot (CH)_i^\gamma \cdot \varepsilon_i \quad (1)$$

Onde I_i são as inovações da unidade de observação i , PeD é o gasto com pesquisa e desenvolvimento e CH é o capital humano. A função de produção do conhecimento

estabelece que a produção de inovações (I) depende dos insumos de PeD e de capital humano (CH). Essa abordagem foi subsequentemente expandida por Jaffe (1989) para estudos em nível regional. Enquanto as evidências empíricas sustentaram a validade da relação representada pela função de produção do conhecimento para o nível mais agregado de indústrias e países, o mesmo não ocorreu no nível mais desagregado da firma. Esta constatação sugere a presença de externalidades na geração do produto do conhecimento, levando a literatura a considerar as regiões como importantes unidades de análise (AUDRETSCH; E FELDMAN, 2004).

A importância das externalidades é reconhecida nos modelos de crescimento endógeno, em que o crescimento econômico é função do acesso a novas ideias tecnológicas e da difusão dessas ideias na estrutura produtiva. Os transbordamentos do conhecimento geram retornos crescentes de escala, considerados em nível da unidade espacial, que explicam a divergência crescente entre países e regiões (KRUGMAN, 1991; ROMER, 1986, 1990; LUCAS, 1988; GROSSMAN, HELPMAN, 1991). Embora não enfatize sua dimensão espacial, a difusão do conhecimento na teoria do crescimento endógeno pressupõe que existem externalidades e transbordamentos de capital humano e tecnologia entre setores, indústrias e países (NAGAOKA *et al*, 2010).

Usando dados de patentes e a abordagem da função de produção do conhecimento de Griliches (1979), Jaffe (1986) mostrou que existem transbordamentos do conhecimento entre firmas próximas no espaço tecnológico². O autor mensurou a distância tecnológica entre as firmas e encontrou evidências de que as empresas, que fazem pesquisa em áreas em que já há muita pesquisa sendo feita por outras firmas, possuem, em média, mais patentes por dólar gasto com PeD e maior taxa de retorno do PeD . Dado que parte do conhecimento, especialmente o codificado, tem natureza de bem público, a existência de outras firmas realizando esforços de pesquisa relacionados tecnologicamente à área de uma firma específica, permite que esta firma alcance resultados com menos esforço de pesquisa do que seria necessário na ausência daquelas (JAFJE, 1986).

A literatura empírica sobre a produção de inovações voltou-se para as investigações sobre o papel da geografia na geração e difusão do conhecimento. As evidências mostram que o processo de criação e difusão do conhecimento é espacialmente limitado. Jaffe (1989)

² Na construção da variável ‘*spillovers*’ Jaffe (1986) assume que a atividade total relevante das outras firmas pode ser resumida pelo “estoque de *spillover* potencial” (“*potential spillover pool*”), dado pela média ponderada do PeD das demais firmas, sendo os pesos proporcionais à proximidade das firmas no espaço tecnológico.

modificou a função de produção do conhecimento de Griliches (1979) para incluir a dimensão espacial, dando origem ao que ficou conhecido na literatura como Função de Produção do Conhecimento Regional (FPCR). Jaffe (1989) argumenta que se o mecanismo de transmissão dos transbordamentos do conhecimento for a conversação informal, a proximidade geográfica da fonte do transbordamento pode ajudar ou mesmo ser necessária para capturar os benefícios do transbordamento. Esse autor usou dados dos estados americanos sobre pesquisa universitária e sobre patentes e gastos de *PeD* do setor privado. Os resultados mostraram que existem transbordamentos mediados geograficamente do conhecimento gerado na pesquisa universitária para a inovação do setor privado, mensurada pelo número de patentes, em áreas correlatas³.

Entretanto, os estados americanos são muito grandes para se assumir que seus inventores, empreendedores e gerentes terão mais chances de se encontrar face a face do que com pessoas vivendo em outros lugares, sendo *proxies* muito pobres para as unidades espaciais dentro das quais o conhecimento deve transbordar (BRESCHI; LISSONI, 2001). Os resultados de Jaffe (1989) sobre o papel da proximidade geográfica no transbordamento do conhecimento da pesquisa universitária são nebulosos, devido à ausência de evidências de que a proximidade geográfica dentro do estado também é relevante (ACS *et al.*, 1992). De acordo com o próprio Jaffe (1989), há apenas uma fraca evidência de que os transbordamentos são facilitados pela coincidência geográfica entre universidades e laboratórios de pesquisa dentro do estado.

Acs *et al.* (1992) estimaram o modelo de Jaffe (1989) utilizando como variável dependente a contagem de inovações introduzidas no mercado em vez do número de patentes⁴. Eles encontraram que o impacto do transbordamento da pesquisa universitária é maior sobre as inovações do que sobre as patentes, e que o impacto da coincidência geográfica também é bem maior sobre as inovações do que sobre as patentes, sugerindo que os transbordamentos da proximidade geográfica podem ser mais importantes do que Jaffe (1989) havia concluído.

Audretsch e Feldman (1996) examinam a extensão em que a atividade industrial se aglomera no espaço e relacionam esta concentração geográfica à existência de externalidades

³Jaffe (1989) incluiu também a população dos estados no modelo estimado para controlar pelos diferentes tamanhos entre as unidades espaciais de observação.

⁴O uso de inovações após sua introdução no mercado fornece uma medida mais direta da atividade inovativa do que o uso de patentes. Os dados sobre inovações incluem as invenções que não foram patenteadas, mas que foram introduzidas no mercado, e exclui as invenções que foram patenteadas, mas que não se mostraram economicamente viáveis para serem comercializadas no mercado (ACS *et al.* 1992).

do conhecimento. Os autores argumentam que se a capacidade de receber transbordamentos do conhecimento é influenciada pela distância da fonte do conhecimento, então deve haver uma tendência à concentração geográfica, principalmente nas indústrias onde os transbordamentos do conhecimento são mais importantes. Para explicar porque a concentração geográfica da atividade inovativa varia entre indústrias, os autores controlam pela concentração geográfica da produção.

Os resultados mostraram que, mesmo após controlar pela concentração espacial da produção, a atividade inovativa tende a se aglomerar mais em indústrias onde os transbordamentos do conhecimento têm um papel mais decisivo. Embora estas indústrias também tendam a exibir maior concentração geográfica da produção, os autores concluíram que a tendência da atividade inovativa de se aglomerar é mais atribuível ao papel dos transbordamentos do conhecimento do que à concentração geográfica da produção. São nestas indústrias que o conhecimento tácito é transferido por meio dos transbordamentos do conhecimento, de forma que a atividade inovativa tende a ocorrer dentro de uma proximidade geográfica da fonte deste conhecimento tácito. Esta fonte pode ser um laboratório de pesquisa universitária, o departamento de *PeD* de uma corporação, ou a exposição ao conhecimento incorporado em um trabalhador qualificado (AUDRETSCH; FELDMAN, 1996).

É consenso na literatura que o conhecimento é localizado e que seus transbordamentos diminuem com o aumento da distância geográfica (BOTAZZI; PERI, 2003; GREUNZ, 2003; MURATA *et al.*, 2014). No entanto, os primeiros trabalhos empíricos não indicam qual é o alcance espacial deste transbordamento. Para estudar essa abrangência espacial, a função de produção do conhecimento regional foi refinada usando dados com menor agregação espacial e incluindo variáveis espacialmente defasadas para capturar o efeito da pesquisa (universitária e industrial) realizada em outros lugares dentro de certa distância geográfica. Os transbordamentos inter-regionais de conhecimento são definidos pela soma ponderada dos insumos geradores de conhecimento das regiões vizinhas, onde os pesos espaciais são definidos em função da distância geográfica entre as regiões r e s para refletir a ideia de que a intensidade do transbordamento deve diminuir com a distância espacial entre as áreas geradoras e receptoras (BRESCHI, 2011).

Ao analisar os transbordamentos geográficos e tecnológicos para as regiões europeias, Greunz (2003) observou que os vizinhos geográficos afetam positivamente a atividade de patenteamento de uma região, sendo que esta influência diminui com a ordem de vizinhança

considerada (primeira, segunda e terceira ordens)⁵, o que demonstra o decaimento do transbordamento do conhecimento com o aumento da distância geográfica.

A mesma conclusão foi obtida Bottazi e Peri (2003), ao estimar a elasticidade da inovação regional ao PeD realizado em outras regiões da Europa a várias distâncias espaciais. Eles encontraram uma elasticidade positiva e significativa da inovação regional em relação ao PeD externo, para distâncias de até 300 km. Entretanto, a magnitude da elasticidade é relativamente baixa (dobrar o gasto com PeD em uma região aumenta a inovação em apenas 2 a 3% nas regiões localizadas dentro um raio de 300 km). Os autores argumentam que a limitada abrangência espacial do transbordamento e a baixa elasticidade são consistentes com a ideia de que estes transbordamentos resultam da difusão de conhecimento não codificado, incorporado nos indivíduos, entre pessoas que possuem interações frequentes (mensais ou semanais).

Considerando a proximidade geográfica como determinante do transbordamento do conhecimento, Moreno *et al.* (2005) estimaram a função de produção do conhecimento regional para a Europa. Utilizando uma matriz de pesos espaciais de contiguidade, os autores encontraram uma elasticidade das patentes de uma região em relação às patentes das regiões vizinhas igual a 0,17. Ao utilizar uma matriz de pesos espaciais baseada no inverso do quadrado da distância esta elasticidade foi de 0,25. Isto significa que o transbordamento do conhecimento é importante tanto considerando apenas as regiões vizinhas (contíguas) quanto considerando todas as regiões, atribuindo-se, a cada uma, pesos menores à medida que a distância aumenta. Os resultados mostraram que o conhecimento gerado pela atividade de PeD e pelas inovações transborda para regiões vizinhas, mas apenas até uma distância de 250 km. Os autores observaram também que os efeitos da proximidade espacial são aumentados quando as regiões são tecnologicamente homogêneas.

Jaffe *et al.* (1993) desenvolvem uma metodologia alternativa à abordagem da função de produção do conhecimento regional para investigar a localização dos transbordamentos do conhecimento. Contrariando a afirmação de Krugman (1991, p. 53) de que a mensuração empírica destes transbordamentos seria impossível porque os fluxos do conhecimento são invisíveis e não deixam rastros, Jaffe *et al.* (1993) usam citações de patentes para rastreá-los. As citações de patentes pelos inventores são consideradas uma externalidade pura do conhecimento, na medida em que não há nenhuma compensação monetária envolvida.

⁵ A autora utilizou a matrizes de contiguidade *queen*.

Para utilizar as citações como evidência da localização do transbordamento do conhecimento era necessário controlar a concentração da produção e da inovação. Por isso, Jaffe *et al.* (1993) construíram uma amostra de controle de patentes. Cada patente que cita uma patente anterior foi combinada com uma patente aleatória desta amostra de controle de mesma classe tecnológica e mesma data de requerimento, mas que não citou a mesma patente de origem. Posteriormente, cada patente, a que faz a citação, a citada e a de controle, foi atribuída a uma unidade geográfica de acordo com o endereço do inventor. O experimento consiste em comparar a frequência em que a patente que faz a citação e a citada são da mesma localização geográfica com a frequência em que as patentes de controle e a citada são da mesma localização geográfica. Se a primeira frequência é significativamente maior que a segunda, isto foi interpretado como evidência de efeitos de localização além dos efeitos de aglomeração vindos de outras fontes. Comparando com a amostra de patentes de controle, estes autores mostraram que é três vezes mais provável que as patentes sejam do mesmo estado das patentes citadas, e seis vezes mais provável que sejam da mesma área metropolitana. Cabe observar que, dado que os transbordamentos do conhecimento são altamente localizados, usar estados e áreas metropolitanas tende a subestimar a importância do transbordamento do conhecimento no local da atividade inovativa (CARLINO; KERR, 2015). Além disso, seus resultados indicam que existe localização do conhecimento de forma agregada, mas não analisaram se o grau de localização varia de acordo com a região ou com a tecnologia.

Almeida e Kogut (1999) aplicam a metodologia de Jaffe *et al.* (1993) para o setor de semicondutores dos Estados Unidos e partem da hipótese de que a variação das instituições e das redes de trabalho entre as regiões levam a diferentes graus de localização dos transbordamentos do conhecimento entre estas regiões. Seus resultados mostram que a mobilidade intrarregional dos detentores de patentes na indústria de semicondutores afeta significativamente a transferência do conhecimento, e que estes mecanismos de transmissão estão enraizados nas redes regionais de trabalho. Os autores confirmam também que a localização do conhecimento regional não é um fenômeno universal, havendo padrões diferentes na difusão regional das externalidades do conhecimento devido às diferenças nos mercados regionais de trabalho.

Para lidar com o problema de agregação espacial, Murata *et al.* (2014) utilizam uma abordagem baseada na distância que permite capturar transbordamentos do conhecimento, evidenciados por citações de patentes, também entre fronteiras geográficas. Para evitar

possíveis distorções causadas pela agregação setorial, eles conduzem a análise para diferentes classes tecnológicas. Os resultados mostraram que os transbordamentos do conhecimento são significativamente localizados para a maior parte das classes tecnológicas e são dispersos para um número não desprezível destas. Existe uma heterogeneidade substancial entre as classes tecnológicas no que se refere ao padrão espacial dos fluxos de conhecimento. Isto implica que a agregação das classes tecnológicas em estudos sobre os transbordamentos do conhecimento pode gerar um viés para baixo no grau de localização destes transbordamentos.

Breschi e Lissoni (2009) combinam o uso de citações de patentes *à la Jaffe et al.* (1993) com a análise geográfica da mobilidade dos inventores e redes de coinvenção para setores tecnológicos selecionados nos Estados Unidos. Os resultados mostram que as atividades dos inventores entre as firmas (mobilidade, consultoria, contratos de pesquisa, etc.) explicam uma grande parte dos fluxos localizados do conhecimento. Estas evidências permitem concluir que mecanismos baseados no mercado, tais como a mobilidade no mercado de trabalho, possuem papel relevante na difusão do conhecimento e, conseqüentemente, sobre as inovações e o crescimento regional. Uma vez reconhecida a relevância das externalidades pecuniárias neste processo, e de sua localização espacial, abre-se um questionamento sobre a visão dos transbordamentos puros como sendo os grandes responsáveis pela localização do conhecimento, e indica-se a necessidade de se investigar mais a fundo os mecanismos específicos através dos quais o conhecimento se difunde no espaço.

Como discutido nesse tópico, e evidenciado pela literatura empírica, os fluxos de conhecimento são localizados no espaço. O conhecimento se difunde entre locais diferentes e a intensidade dessa difusão diminui com o aumento da distância geográfica. Essa localização dos fluxos de conhecimento ocorre porque os mecanismos que promovem sua transmissão tendem a ser espacialmente limitados. A próxima seção discute esses mecanismos, enfatizando o papel da mobilidade de capital humano e das relações inter-regionais de coinvenção.

2.3 Canais de transmissão do conhecimento: o papel da mobilidade e da coinvenção inter-regional

Enquanto a informação pode ser formalizada e escrita, o conhecimento tácito é vago e difícil de ser codificado, não podendo ser formalizado e escrito (POLANYI, 1987). A proximidade geográfica é importante na transmissão do conhecimento tácito porque a

interação face a face é o mecanismo essencial pelo qual este conhecimento se difunde. Marshall (1890) já dizia que a proximidade geográfica facilita a transferência deste tipo de conhecimento por meio de interações não planejadas entre indivíduos, promovendo a troca de informações entre trabalhadores e entre firmas. O conhecimento tácito é não rival por natureza, e seu desenvolvimento para qualquer aplicação específica pode transbordar e ter valor econômico em aplicações diferentes (AUDRETSCH; FELDMAN, 2004).

O compartilhamento do conhecimento tácito ocorre por meio de interação social frequente, observação e comunicação (AUDRETSCH; FELDMAN, 2004), o que torna seu custo de transmissão crescente com o aumento da distância geográfica (AUDRETSCH; FELDMAN, 1996). Por não demandar interação, ou interação frequente, a transmissão do conhecimento no formato codificado não é muito sensível à distância (BOTTAZI; PERI, 2003). Apesar disso, a codificação do conhecimento é uma forma menos efetiva que contatos pessoais face a face para difundir o conhecimento tácito (BRECHI; LISSONI, 2009).

A importância da localização geográfica para a geração do conhecimento e para a atividade inovativa parece paradoxal em um mundo de comunicação instantânea. Mas, enquanto o conhecimento codificado pode ser facilmente comunicado e, portanto, tem alcance global, o conhecimento tácito se refere à capacidade específica dos indivíduos, resultante de relações sociais e institucionais dentro de determinado lugar. Este tipo de conhecimento é melhor transferido via contatos face a face e é difícil de ser trocado a longas distâncias (FELDMAN; KOGLER, 2010). Comunicações eletrônicas a distância e contato face a face são complementos. A informação eletrônica (como a internet) abre um grande leque de possibilidades, mas este leque só pode ser experimentado e avaliado quanto à sua utilidade por meio de contatos face a face. Quanto maior a disponibilidade de informação eletrônica mais encontros face a face serão necessários para avaliá-la (HENDERSON, 2007).

A difusão do conhecimento tácito está no cerne das discussões sobre a limitação espacial do conhecimento. Mas ainda existem muitas lacunas na literatura sobre os canais pelos quais esse conhecimento flui. Griliches (1979) diferencia dois tipos de transbordamento do conhecimento: a) os *rent spillovers*, ou externalidades pecuniárias, estão associados à troca de bens e se referem ao conhecimento que é por natureza rival e excludente. Eles ocorrem quando o objeto da transação não é comprado pelo preço que reflete sua qualidade total (*full quality price*); b) por outro lado, os *spillovers* puros, ou externalidades puras, surgem puramente do processo de PeD. Este tipo de transbordamento se refere ao conhecimento que é

não rival e não excludente, podendo ser utilizado por muitos usuários ao mesmo tempo e é livremente acessível.

Na medida em que a troca de conhecimento ocorre em uma transação econômica, as partes envolvidas buscam precificar corretamente o objeto da transação, limitando a extensão de transbordamentos puros (GRILICHES, 1979). Podem surgir externalidades pecuniárias, mas estas não dão origem prontamente a um bem público local. O fluxo de conhecimento decorrente de transações tais como contratos de pesquisa ou recrutamento de cientistas e engenheiros via mercado de trabalho não implica em transbordamentos puros, embora possa incluí-los na medida em que o conhecimento transacionado não seja inteiramente compensado. Ao migrar entre diferentes firmas, os inventores não transferem seu conhecimento de graça (BRESCHI; LISSONI, 2009).

O conceito de transbordamentos localizados do conhecimento é uma caixa preta, cujo conteúdo permanece ambíguo (BRESCHI; LISSONI, 2001). O conhecimento sobre os mecanismos específicos por meio dos quais o conhecimento flui de uma região para outra ainda é limitado (MIGUÉLEZ *et al.*, 2010), a despeito da reconhecida importância destes fluxos para a inovação tecnológica e, conseqüentemente, para o crescimento e o desenvolvimento econômico locais.

As externalidades pecuniárias do conhecimento tácito, decorrentes de relações baseadas em trocas de mercado, também são espacialmente localizadas. O conhecimento tácito, incorporado no indivíduo, pode ser difundido via transações de mercado, tais como contratações no mercado de trabalho, *joint ventures*, redes de coinvenção etc. A transmissão do conhecimento tácito demanda interações pessoais frequentes, independentemente de ser mediada via mercado ou fora dele, o que requer maior proximidade espacial. Além disso, o contato face a face reduz o risco associado ao comportamento oportunista, facilita o monitoramento e contribui para desenvolver confiança entre os parceiros.

Tradicionalmente, os trabalhos empíricos que utilizam a abordagem da função de produção do conhecimento regional tratam toda a externalidade do conhecimento como sendo transbordamento puro (JAFFE, 1986; JAFFE *et al.*, 1993). Esta interpretação negligencia a importância das externalidades pecuniárias no processo de criação e difusão do conhecimento. Além disso, a função de produção do conhecimento ignora completamente os mecanismos pelos quais o conhecimento é de fato transmitido entre os agentes da inovação e entre as regiões.

Para ampliar a compreensão do fenômeno de difusão do conhecimento entre as regiões deve-se avaliar a importância relativa das transferências de conhecimento mediadas pelo mercado e fora dele, e identificar os canais mais importantes pelos quais o conhecimento é transmitido. Isso requer entrar na caixa preta da função de produção do conhecimento regional e distinguir os mecanismos específicos por meio dos quais o conhecimento viaja (BRESCHI, 2011).

Entre as várias formas de contatos pessoais entre cientistas e tecnólogos, destacam-se duas formas de mediação do conhecimento: 1) relações sociais não baseadas no mercado, tais como entre antigos professores e alunos, ou entre colegas afiliados a associações técnicas; e 2) canais baseados em trocas de mercado, tais como mercado de trabalho ou redes de colaboração formal (BRECHI; LISSONI, 2009).

A geografia é relevante para as duas formas de mediação. No primeiro caso porque a proximidade facilita o desenvolvimento das relações sociais não baseadas no mercado, ao expor os agentes a eventos sociais importantes e ao reduzir os custos de se firmar e manter estas relações (AGRAWAL *et al.*, 2006). No segundo caso, porque os trabalhadores intensivos em conhecimento, tais como empregados em atividades de PeD e consultores, tendem a migrar entre firmas, mas dentro do mesmo espaço geográfico (ALMEIDA; KOGUT, 1999; BRESCHI; LISSONI, 2005; 2009; MIGUÉLEZ *et al.*, 2010); e porque as transações econômicas com ativos do conhecimento requerem interações frequentes e construção de confiança, sendo, portanto, mais efetivas se os agentes estão localizados no mesmo espaço geográfico (BRECHI; LISSONI, 2009).

Apesar de as ideias serem simultaneamente os insumos e os produtos em sua própria produção, a geração de conhecimento ocorre dentro de limites espaciais e suas externalidades também são localizadas (ALMEIDA; KOGUT, 1999). Entretanto, não há razão para que as fronteiras geográficas limitem ou impeçam o transbordamento do conhecimento (AUDRETSCH; FELDMAN, 2004). A introdução da distância geográfica no estudo dos transbordamentos inter-regionais do conhecimento reflete a necessidade de se reconhecer a natureza tácita do conhecimento, que é a necessidade de interações frequentes entre os emissores e receptores do transbordamento, na forma de contatos face a face.

Dado que a difusão do conhecimento tácito, incorporado no indivíduo, só pode ocorrer via interações pessoais entre trabalhadores do conhecimento, a mobilidade e a formação de redes de colaboração em pesquisa entre inventores de diferentes regiões são mecanismos essenciais para romper estas barreiras espaciais e ampliar o alcance espacial da difusão do

conhecimento, potencializando a inovação e o crescimento regional. A inovação tecnológica regional pode ser influenciada pela inovação tecnológica das regiões vizinhas por meio desses mecanismos de interações inter-regionais, gerando concentração espacial da atividade inovativa.

A acumulação de capital humano e de indivíduos habilitados é importante para o crescimento e o desenvolvimento regional. Os locais que recebem estes talentos podem se beneficiar do conhecimento incorporado nestes indivíduos e das externalidades positivas de seu capital humano, o que contribui para aumentar a produtividade de outras firmas e indivíduos, incentivando o crescimento regional de longo prazo (LUCAS, 1988). Embora a migração de capital humano seja reconhecida como um importante mecanismo de transferência de conhecimento entre regiões, no contexto das pesquisas sobre inovações ainda se sabe pouco sobre a força deste mecanismo de transferência do conhecimento (FAGGIAN; MCCANN, 2009).

As patentes representam conhecimento codificado, mas o estoque de conhecimento utilizado na elaboração da invenção objeto da patente é, em grande medida, tácito. Para usar este conhecimento de forma produtiva é necessário interagir com os indivíduos que geraram e que dominam este conhecimento, ou seja, os inventores. A geografia e as relações sociais são importantes porque a transferência deste conhecimento tácito requer comunicação direta entre os indivíduos, tais como contatos face a face frequentes, e a eficácia desta comunicação cai significativamente com a distância espacial. A localização dos trabalhadores do conhecimento no mesmo espaço geográfico os expõe a muitas relações sociais potenciais e reduz os custos de formar e manter estas relações (BRESCHI; LISSONI, 2009).

Os trabalhadores altamente qualificados são fonte de inovação e de transbordamento do conhecimento, portanto, sua mobilidade geográfica é importante para o desenvolvimento regional. A mobilidade de inventores é importante porque eles são os veículos do conhecimento tácito. As informações sobre os padrões desta mobilidade, sobre seu efeito sobre a produtividade dos inventores, e sobre suas externalidades, podem fornecer subsídios para o desenho de políticas que visem fomentar a inovação e o desenvolvimento regional (MIGUÉLEZ *et al.*, 2010).

Ao contrário da análise das citações de patentes, que mapeiam o conhecimento codificável, o estudo da difusão do conhecimento por meio da mobilidade das pessoas, baseia-se na ideia de que o conhecimento é enraizado no indivíduo. A dimensão tácita do conhecimento individual não é publicamente disponível, mas é incorporada na pessoa a ser

contratada, e o conhecimento tácito coletivo de uma firma não é facilmente imitado, de forma que o conhecimento tácito é uma fonte de vantagem competitiva (LEONARD; SENSIPER, 1998).

O estudo da mobilidade de indivíduos qualificados para entender transbordamentos do conhecimento tácito se depara com a dificuldade de diferenciar quanto do transbordamento puro realmente é gerado com a mobilidade. O que ocorre quando um indivíduo se move de uma firma para outra é mais uma transferência de conhecimento do que um transbordamento (FELDMAN; KOGLER, 2010), embora esse também possa ocorrer. Diferenciar que parte do fluxo de conhecimento se refere à externalidade pura e que parte foi mediada pelo mercado é uma questão ainda não elucidada pela literatura. A difusão do conhecimento por meio da mobilidade, entre firmas e entre regiões, possibilita o desenvolvimento de inovações e o crescimento econômico, independentemente de serem caracterizadas como externalidades puras ou pecuniárias.

Ao tratar as citações de patentes como transbordamentos do conhecimento (externalidades puras), Jaffe *et al.* (1993, p. 583) assumem que, em geral, o contrato entre duas partes é bastante incompleto, sendo provável que a firma que cita a patente capture alguma renda da invenção original, beneficiando-se de, pelo menos, um transbordamento parcial. De forma similar, é razoável que a contratação de inventores envolva transbordamentos puros parciais, na medida em que o contrato de trabalho pode não compensar adequadamente a troca de conhecimento envolvida na relação. A mobilidade de inventores, tanto regional quanto inter-regional, não apenas promove a difusão do conhecimento, mas também gera externalidades puras positivas. Além disso, o conhecimento possui externalidades naturais, na medida em que a criação de conhecimento novo por uma firma afeta positivamente as possibilidades de produção de outras firmas, pois o conhecimento não pode ser perfeitamente patenteado ou mantido em segredo (ROMER, 1986).

Várias pesquisas sobre mobilidade mostraram que os fluxos de conhecimento são localizados espacialmente porque a mobilidade e a formação de redes entre os agentes do conhecimento tendem a ser limitados geograficamente (BRESCHI; LISSONI, 2009; MIGUÉLEZ; MORENO, 2013). Almeida e Kogut (1999) mostraram que a mobilidade intrarregional e as redes formadas via mercado de trabalho são os determinantes da localização espacial do conhecimento. Concluíram também que a mobilidade inter-regional é

uma fonte de redução desta localização. Esse achado é coerente com o fato de que a mobilidade entre regiões amplia a escala espacial da difusão do conhecimento.

Breschi e Lissoni (2005; 2009) concluem que a mobilidade dos inventores entre as firmas responde pela maior parte da localização geográfica das citações de patentes. Estes inventores móveis funcionam como pontes entre as equipes de inventores, sendo os responsáveis pela criação de redes de coinvenção. A baixa propensão dos inventores de se realocar no espaço resulta em uma rede de coinvenção espacialmente localizada, sendo esta a razão fundamental da localização espacial dos fluxos de conhecimento. Na ausência de movimentos localizados de inventores e a consequente criação de redes de colaboração formal (coinvenções), as relações informais tendem a explicar apenas uma parte secundária do fenômeno observado. A mobilidade e a coinvenção são geralmente baseados em relações de mercado. Os fluxos de conhecimento mediados por estes canais só podem ser classificados como externalidades ou transbordamentos na medida em que forem objetos de um contrato incompleto. Portanto, o papel das relações sociais informais no transbordamento do conhecimento tácito parece ser bem menor do que se pensava.

Miguélez *et al.* (2010) estudaram a mobilidade geográfica dos inventores na Europa e concluíram que a migração de inventores é um fenômeno localizado. Após controlar pelo potencial de inovação e pelo patenteamento das regiões, os autores observaram que a atração de inventores é reservada a apenas algumas regiões e países, sendo os únicos que podem potencialmente se beneficiar do fluxo de conhecimento e das externalidades espaciais do capital humano. Além disto, observaram que há uma autocorrelação espacial positiva da migração de inventores. Concluíram que as características atrativas de uma região, a saber, amenidades, oportunidades de emprego, redes sociais, instalações de pesquisa, firmas multinacionais, contatos com a academia etc., transbordam para as regiões vizinhas, mostrando a importância da geografia e da localização na atração de inventores. Assim, os fluxos de conhecimento são localizados na medida em que a mobilidade dos inventores também é localizada.

Miguélez e Moreno (2013a) investigaram o papel da mobilidade geográfica e das colaborações em pesquisa entre regiões da Europa sobre as inovações regionais, empregando a função de produção do conhecimento regional. As matrizes de pesos espaciais, baseadas na distância ou na contiguidade, tradicionalmente utilizadas nesta abordagem para refletir o decaimento dos transbordamentos do conhecimento ao longo do espaço, foram substituídas por matrizes de pesos baseadas em interações reais entre os agentes do conhecimento das

regiões. Os autores estimaram modelos gravitacionais *cross-section* para cada ano para avaliar o papel de variáveis bilaterais (geográficas, administrativas e cognitivas) em explicar a mobilidade e a colaboração. Os valores ajustados dessas estimações foram usados para construir as matrizes de peso que caracterizam, então, os fluxos de conhecimento entre as regiões. Essas matrizes foram empregadas na função de produção do conhecimento regional para ponderar o estoque externo de conhecimento representado pelo PeD das demais regiões.

Os resultados dos modelos gravitacionais mostraram que as distâncias geográfica e tecnológica afetam negativamente tanto a mobilidade dos trabalhadores do conhecimento entre as regiões quanto as colaborações em pesquisa. Os resultados da função de produção do conhecimento mostraram que a produção do conhecimento regional depende também do conhecimento disponível em outras regiões, acessível através de inventores móveis e de colaborações bilaterais. Além disso, os autores concluíram que as regiões se beneficiam mais dos estoques externos de conhecimento quando a transmissão é por meio das redes de colaboração do que por meio da mobilidade de trabalhadores altamente qualificados. O impacto do último é um terço do impacto do primeiro.

Faggian e McCann (2009) estudam o impacto da migração de capital humano, representado por formandos universitários, sobre a inovação regional na Inglaterra. A relação entre migração de capital humano e inovação regional é um processo endógeno e cumulativo, pois não se pode assumir *a priori* nenhuma causalidade dominante entre estes dois fenômenos, no sentido de um ser a causa do outro. O influxo de capital humano afeta positivamente a performance inovativa regional e, ao mesmo tempo, o dinamismo inovador da região atrai a entrada de capital humano. Usando mínimos quadrados em três estágios e controlando os fatores regionais, os resultados mostraram que o desempenho regional em inovação é um dos principais fatores que atraem formandos para a região e que, ao mesmo tempo, a entrada destes graduandos na região também promove a inovação regional. Os autores interpretaram estas evidências como indicadores de que a performance regional em inovações depende mais da mobilidade inter-regional de capital humano do que da mobilidade intrarregional. A mobilidade de inventores está intrinsecamente relacionada à difusão do conhecimento tácito, permitindo aprendizagem mútua e incentivando o desenvolvimento de inovações. Esta mobilidade pode ser impulsionada pelas redes sociais de pesquisa e vice e versa (MIGUÉLEZ *et al.*, 2010).

As colaborações a longa distância são canais importantes de criação e difusão do conhecimento, mas isso não significa que a distância espacial não importe. A coinvenção

inter-regional é uma forma de rede de colaboração que pode ser governada por um acordo formal ou informal. A rede de coinvenção é afetada pela geografia, pois a proximidade espacial e a colocação facilita a transferência de conhecimento complexo e frequentes interações face a face podem ser necessárias. Embora a facilidade de comunicação e de viagens tenda a reduzir a importância da proximidade espacial, a importância da distância para a coinvenção inter-regional pode depender do grau de complementaridade entre as interações remotas e face a face (MORESCALCHI *et al.*, 2015).

Bercovitz e Feldman (2011) estudaram equipes de inventores para entender se a colaboração a longa distância é influenciada por relações empregatícias e sociais anteriores ou atração de “estrelas”. Os resultados mostram que a maioria dos membros externos teve alguma relação anterior com os membros internos da equipe de inventores, seja como antigos colegas de trabalho ou de faculdade, ou como coautores de longa data. Essas colaborações refletem os rastros de uma geografia anterior, ou seja, a colaboração reflete uma localização anterior no mesmo espaço geográfico. Nos casos em que não houve essa relação anterior o mais importante foi a atração de estrelas.

Dado que o conhecimento tende a viajar com as pessoas que o dominam (BRECHI; LISSONI, 2001), a mobilidade desses indivíduos leva à circulação do conhecimento entre os lugares, e pode incentivar a formação de redes sociais à distância entre o indivíduo móvel e os agentes com os quais ele mantinha relações na localização geográfica anterior. Agrawal *et al.* (2006) exploraram o papel das relações sociais entre inventores por meio de um experimento inspirado em Jaffe, Trajtenberg e Henderson (1993)(JTH) modificado com citações de patentes de inventores que se mudaram pelo menos uma vez entre áreas metropolitanas diferentes. As citações de patentes são consideradas como *proxy* para o fluxo de conhecimento, e os autores partem da hipótese de que esse seria facilitado pelas relações sociais. Eles examinaram qual a fração de citações das patentes, produzidas após a mudança do inventor, vem de patentes geradas em sua localização anterior. Os resultados mostraram que uma parte significativa do conhecimento gerado pelo inventor em seu novo local retorna à região anterior, indicando que a localização no mesmo espaço geográfico facilita a formação de relações sociais que persistem depois que os indivíduos se separam. Isso implica que existe um transbordamento da região que recebeu o inventor para a região que perdeu o inventor, atenuando o *brain drain*.

A difusão inter-regional do conhecimento carrega o potencial de incentivar as inovações regionais. Mas esse potencial só será realizado se a região possuir as condições prévias

necessárias para o desenvolvimento das inovações. Feldman e Florida (1994) e Feldman (1993) testam a hipótese de que a inovação se concentra nos lugares que possuem uma infraestrutura tecnológica bem desenvolvida. A inovação depende da infraestrutura tecnológica regional, capaz de mobilizar recursos técnicos, conhecimento e outros insumos essenciais ao processo inovativo. Essa infraestrutura consiste em fontes do conhecimento: aglomeração de firmas em indústrias relacionadas, que fornecem expertise e conhecimento técnico; concentração de PeD que aumenta as oportunidades de inovação ao fornecer conhecimento de novas descobertas científicas e aplicações; e serviços com *expertise* no posicionamento e comercialização de novos produtos. A concentração geográfica desse tipo de infraestrutura aumenta a capacidade de inovação na medida em que estas regiões desenvolvem e se especializam em dadas tecnologias e setores industriais. Em outras palavras, a geografia serve como o repositório em que os agentes da inovação organizam uma infraestrutura que congrega os recursos cruciais e insumos para o processo de inovação. Os resultados mostram que a inovação é positivamente relacionada com o PeD, universitário e empresarial, com a presença de indústrias relacionadas e com a presença de recursos especializados relacionados às atividades inovativas. Os autores concluem que não apenas as inovações se concentram em locais onde há aglomeração de infraestrutura tecnológica, mas também que estas concentrações espaciais de recursos especializados reforçam mutuamente e positivamente a capacidade regional de inovar.

Além da infraestrutura regional e da proximidade geográfica, a proximidade tecnológica também é importante para que as regiões consigam aproveitar os transbordamentos de conhecimento das regiões vizinhas no desenvolvimento de inovações. A capacidade de uma firma em obter vantagem produtiva do estoque de conhecimento de outras empresas depende do grau de similaridade tecnológica entre elas (JAFFE, 1986). Dado que o sistema produtivo de uma região não é nada mais do que um conjunto de firmas que atuam em diferentes tecnologias, espera-se que os transbordamentos sejam mais fortes entre regiões com perfis tecnológicos semelhantes (GREUNZ, 2003).

De forma geral, o benefício da migração para a região de destino dependerá não só das características locais, mas também do tipo de migrante que a região atrai ou perde (GOLGHER, 2008). As características individuais dos inventores vindos de outras regiões influenciam o efeito da migração sobre a inovação da região receptora. O desempenho do inventor na firma de origem em termos do número de patentes desenvolvidas ou o desempenho inovador da firma de origem influenciam o impacto da migração do inventor

sobre a inovação da firma de destino. Ao analisar o efeito da mobilidade de inventores sobre o patenteamento das firmas, Rahko (2017) observou que a mobilidade de inventores em geral não afeta significativamente o patenteamento da firma contratante, mas a mobilidade de inventores com muitas patentes anteriores ou de inventores vindos de firmas com muitas patentes contribui positivamente para o patenteamento futuro da firma de destino. Para que os benefícios da mobilidade de inventores ocorram, esses inventores têm que ser capazes de transferir conhecimento tecnológico obtido em suas atividades anteriores (KAISER et al., 2015; RAHKO, 2017).

As externalidades do conhecimento não são um subproduto automático da proximidade geográfica e tecnológica. É a interação entre firmas, universidades, cientistas “estrelas”, e engenheiros que condiciona a extensão em que o conhecimento transborda (ALMEIDA; KOGUT, 1999). A proximidade geográfica entre as regiões permite a troca de conhecimento em geral e de conhecimento tácito em particular. Mas os canais pelos quais o conhecimento flui dependem parcialmente da disposição das regiões para incentivá-los. Além disso, o esforço de inovação da região é uma pré-condição para o sucesso da imitação das tecnologias desenvolvidas em outros lugares, pois ele alimenta a base de conhecimento sobre a qual a região constrói sua competência técnica. Em outras palavras, ter uma atividade inovativa confere uma experiência que é necessária para assimilar o conhecimento desenvolvido em outros lugares e para desenvolver o seu próprio (GREUNZ, 2003). Mesmo havendo proximidade espacial, os fluxos de conhecimento não são recebidos automaticamente apenas pelo fato de se estar lá. Ao contrário, os fluxos de conhecimento seguem canais específicos de transmissão, os quais são principalmente baseados nas interações de mercado e externalidades pecuniárias (BRESCHI; LISSONI, 2001). A mobilidade laboral de capital humano e as redes de colaboração em pesquisa são dois canais definidos e estruturados por meio dos quais o conhecimento flui (MIGUÉLEZ; MORENO, 2013b) dentro e entre as regiões.

As externalidades do conhecimento só afetarão a produção de inovações se a busca de fontes externas de conhecimento fizer parte da estratégia de inovações das firmas (CRESCENZI; GAGLIARDI, 2015). A economia evolucionária argumenta que as capacidades da firma são eventualmente padronizadas, criando um *path dependence* interno à firma que é baseado em práticas locais, principalmente se a firma teve algum sucesso inovativo no passado. Ao desenvolver maior *path dependence*, a firma também exibe menor receptividade a fontes externas de conhecimento. Portanto, a efetividade da obtenção de

conhecimento a partir da mobilidade de trabalhadores altamente qualificados depende do grau de *path dependence* (NELSON; WINTER, 1982).

No mesmo sentido, a abordagem da capacidade de absorção se baseia no caráter cumulativo do conhecimento e da inovação. Sob essa visão, o benefício do conhecimento externo acessado por meio da entrada de trabalhadores qualificados e das redes inter-regionais de colaboração para a inovação local dependerá da capacidade regional de reconhecer e aplicar esse conhecimento de forma útil. Desenvolvido para o nível da firma, esse conceito se refere à habilidade de reconhecer, assimilar e aplicar o conhecimento externo em objetivos comerciais (COHEN; LEVINTHAL, 1990). É razoável estender essa noção para o nível regional, assumindo que a capacidade de absorção das firmas locais determina a capacidade de absorção regional, uma vez que as firmas constituem o elemento básico do sistema regional de inovações (MIGUÉLEZ; MORENO, 2015)⁶.

A capacidade de absorver conhecimento externo depende do conhecimento acumulado internamente. Essa acumulação prévia de conhecimento é necessária para que a firma possa compreender e utilizar o conhecimento externo. Nesse sentido, o indicador utilizado para mensurar a capacidade de absorção é o gasto em PeD. A atividade de PeD afeta o desempenho inovador da firma de duas formas: diretamente, como insumo do processo de inovação, e indiretamente por meio de sua importância para a formação da capacidade de absorção (COHEN; LEVINTHAL, 1990). Em nível regional, a capacidade de absorção local do conhecimento externo trazido pela entrada de inventores e pelas redes inter-regionais de colaboração será determinada pelo gasto regional com PeD (MIGUÉLEZ; MORENO, 2015).

A despeito de fazerem parte do mesmo ambiente nacional, as regiões dependem primariamente da capacidade institucional local que reflete diferenças nos serviços de educação e serviços regulatórios. Como resultado de vantagens específicas de capacidades institucionais enraizadas localmente na forma de conhecimento tácito, o sistema regional de inovação é o fator mais importante para o aprendizado local. Interdependências não comercializáveis, tais como conhecimento tácito baseado em contatos face-a-face, rotinas, hábitos e normas, convenções de comunicação e de interação, são ativos importantes que moldam o potencial inovativo de uma região (FELDMAN; KOGLER, 2010).

⁶ Reconhecemos que essa noção é muito simplista, pois a capacidade de absorção regional vai além da simples soma das capacidades individuais das firmas locais, compreendendo as interações e inter-relações entre elas (ABREU, 2011). Entretanto, esse raciocínio tem sido empregado para contextos regionais (MIGUÉLEZ; MORENO, 2015; MUKHERJI; SILBERMAN, 2013).

Além disso, a infraestrutura de transporte e comunicação também é relevante para a troca de conhecimento. Lugares altamente conectados se envolvem em trocas de conhecimento e aprendizagem recíproca de forma muito mais intensa que a distância geográfica entre eles poderia sugerir. Por outro lado, para dois lugares próximos um do outro, mas com infraestrutura precária de transporte e comunicação, o transbordamento do conhecimento será bem menos intenso (FELDMAN; KOGLER, 2010).

Para estudar o fluxo inter-regional de conhecimento, a intensidade inovativa regional tem sido modelada em função da intensidade inovativa de outras regiões. Para isso têm sido utilizadas matrizes espaciais *ad hoc* que descrevem a relação espacial entre as unidades geográficas (baseadas na distância ou na fronteira comum, por exemplo). Os resultados encontrados na literatura geralmente mostram uma associação positiva entre a performance inovativa local e as inovações das regiões vizinhas, o que tem sido interpretado como evidência de transbordamentos localizados do conhecimento (LISSONI; MIGUÉLEZ, 2014). Entretanto, mecanismos de difusão do conhecimento baseados no mercado, tais como a mobilidade de inventores e as redes de coinvenção inter-regional, possuem papel significativo na transmissão do conhecimento e no caráter localizado dessa difusão (BRESCHI; LISSONI, 2009; MIGUÉLEZ; MORENO, 2015). Por isso é relevante identificar empiricamente a contribuição desses dois canais para a difusão do conhecimento entre regiões.

Esse tópico discutiu a importância da mobilidade dos indivíduos que dominam o conhecimento e de suas colaborações em pesquisa como mecanismos para difundir o conhecimento entre as regiões. Tanto a mobilidade quanto as colaborações tendem a ocorrer entre locais mais próximos, o que contribui para a localização espacial dos fluxos de conhecimento. As relações bilaterais de mobilidade e de coinvenção são formas de interações regionais que possibilitam o acesso ao conhecimento externo e, como tal, podem ser utilizadas como pesos de ponderação em substituição aos pesos puramente espaciais para avaliar a difusão inter-regional do conhecimento.

O próximo capítulo explica a abordagem metodológica utilizada para analisar os fluxos bilaterais de mobilidade e de coinvenção inter-regional e o impacto desses fluxos sobre a difusão do conhecimento e a inovação regional.

3 ESTRATÉGIA METODOLÓGICA

A identificação da importância relativa da mobilidade e da rede inter-regional de inventores para a difusão do conhecimento e a inovação regional será baseada no arcabouço da função de produção do conhecimento regional. Mas, diferentemente das abordagens tradicionais que utilizam apenas matrizes de pesos baseadas em critérios espaciais, como contiguidade ou distância (ACS *et al.*, 1994; CHARLOT *et al.*, 2015), serão introduzidas matrizes de pesos que refletem interações reais de conhecimento, construídas a partir da mobilidade e da rede inter-regional de inventores, conforme a estratégia adotada por Miguélez e Moreno (2013a). Dessa forma, será possível distinguir o papel destes mecanismos na difusão espacial do conhecimento, contribuindo para compreender melhor os canais pelos quais o conhecimento viaja entre as regiões.

A utilização de matrizes de pesos baseadas na interação real entre as regiões por meio da mobilidade de inventores e da co-invenção enriquece a abordagem da função de produção regional ao permitir identificar a importância destes canais na difusão do conhecimento. Uma vez identificado o papel destas interações na geração de externalidades pecuniárias, que se traduzem na difusão inter-regional do conhecimento e beneficiam o desempenho inovativo regional, torna-se possível identificar a importância relativa dos transbordamentos puros do conhecimento. A investigação dos canais pelos quais o conhecimento flui entre regiões permite identificar a contribuição das diferentes formas de difusão. Essa compreensão é relevante porque afeta o desenho de políticas de incentivo à inovação.

A estratégia metodológica empregada para investigar o papel da colaboração e da mobilidade bilateral sobre o desempenho das atividades inovadoras das regiões brasileiras será baseada no procedimento utilizado por Miguélez e Moreno (2013a).

3.1 Função de produção do conhecimento regional

A função de produção do conhecimento da firma representativa k da região i ⁷, é dada por:

$$Y_{kit} = A_{it} \cdot f(K_{kit}, L_{kit}) \quad (2)$$

Onde Y_{kit} é o desempenho inovativo da firma k da região i no período t ; A_{it} é um índice de produtividade do PeD, específico da região i no período t ; K_{kit} são os insumos de PeD da

⁷ Serão considerados todos os agentes inovadores das regiões, tais como firmas privadas, centros de pesquisa e universidades.

firma k da região i no período t ; e L_{kit} é a mão de obra qualificada da firma k da região i no período t .

Por simplicidade, assume-se que a função de produção do conhecimento possua a forma de uma *Cobb-Douglas*:

$$Y_{kit} = A_{it} K_{kit}^{\beta} L_{kit}^{\alpha}, \quad 0 < \alpha < 1, \quad 0 < \beta < 1, \quad \alpha + \beta = 1 \quad (3)$$

Agregando-se todas as firmas da região i , tem-se a inovação regional em cada período, Y_{it} :

$$Y_{it} = e^C A_{it} K_{it}^{\beta} L_{it}^{\alpha} \quad (4)$$

Onde K_{it} é o gasto regional com PeD, L_{it} é a mão de obra qualificada da região i , e e^C é o termo constante que capta o impacto de todos os fatores comuns que afetam a inovação de todas as regiões. Dada a diferença populacional entre as regiões, a equação (4) é pré-multiplicada por $1/N$, sendo N a população total de cada região:

$$y_{it} = e^C A_{it} k_{it}^{\beta} l_{it}^{\alpha} \quad (5)$$

Onde $y_{it} = Y/N$, $k_{it} = K/N$, e $l_{it} = L/N$. Dessa forma, a intensidade inovativa das regiões depende do gasto per capita com PeD e da participação da mão de obra qualificada na população regional.

Assume-se que o índice regional de produtividade do PeD (A_{it}) depende das interações inter-regionais, por meio de redes de co-invenção e mobilidade dos inventores entre as regiões, permitindo às firmas da região i acessar o estoque externo de conhecimento (SK_i) bem como de algumas vantagens locacionais gerais da região i , representadas pelo efeito fixo regional, e^{δ_i} ⁸.

Seja SK o estoque de conhecimento acumulado em todas as regiões, exceto a região i , e acessível a todas as firmas da região i no tempo t . Se o conhecimento gerado em uma região fosse perfeitamente difundido às demais, o estoque externo de conhecimento acessível à região i seria a soma do estoque de conhecimento de todas as outras regiões, $SK_{it} = \sum_{j=1, j \neq i}^I SK_{jt}$. Entretanto, como a difusão do conhecimento não é perfeita, o conhecimento

externo acessível à região i é dado por $SK_{it} = \sum_{j=1, j \neq i}^I \phi_{ij} SK_{jt}$ onde $\phi_{ij} \in [0,1]$ é o percentual

do conhecimento gerado na região j que é acessível à região i (PERI, 2005). O conhecimento flui por meio de vários canais, alguns baseados em trocas de mercado e outros fora dele. Nosso interesse é na parte do estoque de conhecimento gerado na região j que é acessível à

⁸ Sabe-se que este índice depende também da mobilidade local entre firmas e das redes de co-invenção internas à região, mas estes fatores não serão abordados nessa tese.

região i através dos canais específicos constituídos pela mobilidade de inventores e pelas relações de coinvenção. A magnitude do estoque externo de conhecimento acessível à região i por meio desses mecanismos dependerá da intensidade das interações entre i e j no tempo t , ou seja:

$$SK_{it} = \left(\prod_{\substack{j=1 \\ j \neq i}}^I SK_{jt}^{\rho^m \omega_{ij,t}^m} \right) \left(\prod_{\substack{j=1 \\ j \neq i}}^I SK_{jt}^{\rho^c \omega_{ij,t}^c} \right) \quad (6)$$

Onde:

SK_{it} é o estoque de conhecimento externo acessível a todas as firmas da região i no tempo t ;

ρ^m é a elasticidade do produto em relação ao potencial da mobilidade em difundir o estoque de conhecimento entre as regiões, logo é um parâmetro de interesse a ser estimado;

$\omega_{ij,t}^m$ é o peso bilateral que mensura a intensidade das interações entre as regiões i e j por meio da mobilidade espacial de inventores;

ρ^c é a elasticidade do produto em relação ao potencial das colaborações em pesquisa em difundir o estoque de conhecimento entre as regiões; logo, é um parâmetro de interesse a ser estimado;

$\omega_{ij,t}^c$ é o peso bilateral que mensura a intensidade das interações entre as regiões i e j por meio das relações geográficas de colaboração para inovar.

Fazendo a transformação logarítmica de (5) e substituindo (A_{it}) pela transformação logarítmica de (6) tem-se a função de produção do conhecimento regional modificada:

$$\ln y_{it} = c + \beta \ln k_{it} + \alpha \ln l_{it} + \rho^m \sum_{j \neq i} \omega_{ij,t}^m \ln SK_{jt} + \rho^c \sum_{j \neq i} \omega_{ij,t}^c \ln SK_{jt} + \delta_i + \varepsilon_{it} \quad (7)$$

Para estimar a equação (7) é necessário antes mensurar a intensidade das interações regionais bilaterais que ocorrem sob a forma de mobilidade de inventores, $\omega_{ij,t}^m$, e de colaborações tecnológicas, $\omega_{ij,t}^c$, para se obter o estoque acessível de conhecimento externo.

Portanto, a investigação empírica será feita em dois estágios: i) no primeiro a magnitude das interações regionais será estimada para cada par de regiões usando dados de mobilidade e de coinvenção entre as regiões como uma função de um conjunto de variáveis bilaterais explicativas; nessa etapa serão construídas as matrizes de pesos $\omega_{ij,t}^m$ e $\omega_{ij,t}^c$ que refletem interações regionais de conhecimento; ii) no segundo estágio as matrizes de pesos construídas a partir dos valores ajustados das estimações da primeira etapa serão usadas para inferir o impacto do conhecimento externo sobre a inovação regional por meio dos parâmetros ρ^m e ρ^c (equação 7).

Na função de produção do conhecimento regional estimada pelas técnicas da econometria espacial, tradicionalmente são estimados modelos de defasagem espacial (*spatial*

autoregressive - SAR). A defasagem espacial é obtida com o uso de matrizes de ponderação baseadas em critérios espaciais, tais como distância ou contiguidade. Nesses modelos a defasagem espacial da inovação das regiões vizinhas é interpretada como evidência de transbordamentos inter-regionais do conhecimento técnico. Nessa tese será estimado um modelo de defasagem nos moldes do *spatial autoregressive (SAR)*, porém utilizando-se como pesos as matrizes de mobilidade e de coinvenção estimadas no primeiro estágio.

3.2 Interações regionais de conhecimento: uma abordagem gravitacional

Os fluxos entre duas regiões ou países podem ser modelados por uma equação gravitacional, inspirada pela lei gravitacional de Newton, em que a magnitude destes fluxos é diretamente proporcional ao tamanho da região, indicado pelo PIB e/ou população, e inversamente proporcional à distância. Inicialmente utilizados na economia para analisar fluxos bilaterais de comércio entre países (TINBERGEN, 1962), os modelos gravitacionais também podem ser utilizados para analisar o fluxo de pessoas e de conhecimento. Estes modelos podem ser usados para avaliar a importância da geografia e de outras características locais para explicar a mobilidade inter-regional de inventores e a colaboração em pesquisa entre inventores de diferentes regiões.

3.2.1 Interações de conhecimento por meio da mobilidade inter-regional de inventores

O modelo gravitacional de migração geralmente se baseia em um modelo de maximização de utilidade (BERTOLI; MORAGA, 2015; RAMOS, 2016; GORIN, 2016). Indivíduos racionais tomam sua decisão de migrar de um lugar para outro com base na diferença entre as utilidades esperadas do destino e da origem. Consideramos um modelo linear onde a utilidade do inventor em trabalhar e morar na região i possui uma parte determinística comum a todos os inventores, u , e um componente estocástico específico a cada indivíduo, ε_i^k . Especificamente, a utilidade do inventor k em trabalhar na região i pode ser expressa como:

$$U_i^k = u(E_i) + \varepsilon_i^k \quad (8)$$

Onde E_i é um vetor de variáveis econômicas e de amenidades da região i e o componente determinístico $u(E_i)$ é uma função deste vetor, afetando as k utilidades desta região, e que é comum a todos os inventores. O componente estocástico ε_i^k inclui quaisquer

fatores que levam o indivíduo a avaliar E_i de forma diferente de um indivíduo selecionado aleatoriamente. Por definição, $E(\varepsilon_i) = 0$, pois qualquer componente sistemático está incluído em u .

O indivíduo k migrará do local i para o j se a utilidade esperada de morar e trabalhar em j for maior que a utilidade esperada de i , menos os custos de realocação, os quais dependem das distâncias geográfica, administrativa e cognitiva, D_{ij} :

$$E(U_j^k) > E(U_i^k) + C(D_{ij}) \quad (9)$$

Se a condição (9) for satisfeita, a variável migração entre i e j (M_{ij}^k) será igual a 1, se ela não for satisfeita a variável migração entre i e j (M_{ij}^k) será igual a 0. Portanto, a migração de i para j implica que não há outra mudança alternativa que resulte em maior aumento da utilidade esperada⁹. Agregando as migrações individuais por região e usando uma forma geral de modelo gravitacional, tem-se:

$$M_{ij} = f(E_i, E_j, D_{ij}) \quad (10)$$

Onde D_{ij} é um conjunto de variáveis bilaterais que podem criar resistência ao fluxo de inventores entre as regiões i e j . Especificamente:

$$D_{ij} = G_{ij}^\varphi e^{\psi_{ij}} T_{ij}^\tau e^{\gamma_{ij}} \quad (11)$$

Onde G_{ij} é a distância euclidiana entre os centroides das regiões i e j . A distância afeta negativamente a mobilidade, pois os riscos e custos de migrar aumentam com a distância entre a origem e o destino (ZIMMERMAN, 2004). Especificamente, a mobilidade bilateral de inventores se relaciona inversamente com a distância geográfica devido, entre outros fatores, a: i) a imperfeição da informação sobre o destino tende a aumentar com o aumento da distância física; ii) os custos de transporte são maiores entre locais espacialmente mais distantes; e iii) encontros com a família, amigos e antigos colegas serão mais esparsos no tempo com distâncias geográficas maiores. Esses efeitos tendem a ser amenizados no caso de regiões contíguas, então, seguindo Miguélez e Moreno (2013a) será adicionada uma *dummy* para regiões com fronteira comum, $e^{\psi_{ij}}$. A proximidade tecnológica, T_{ij} , será incluída como *proxy* do quanto a similaridade cognitiva tende a afetar positivamente a mobilidade entre comunidades epistêmicas fisicamente distantes. Finalmente, para controlar efeitos

⁹ Uma suposição importante do modelo de maximização de utilidade é que a atratividade do destino não é afetada pela migração. Por exemplo, se um destino particular é atrativo devido a seu baixo nível de desemprego em relação à região de origem, grandes influxos de imigrantes poderiam aumentar o desemprego no destino e reduzi-lo na origem. Os modelos gravitacionais não capturam estes efeitos dinâmicos (RAMOS, 2016).

institucionais, históricos e culturais sobre a mobilidade dos inventores será introduzida uma *dummy* indicando se duas regiões pertencem ou não à mesma unidade da federação, $e^{\gamma_{ij}}$.

A abordagem tradicional do modelo gravitacional possui um viés de variável omitida, desconsiderando variáveis relevantes na explicação do fenômeno analisado. Anderson e van Wincoop (2003) argumentam que a equação tradicional é mal especificada, pois ignora os termos de resistência multilateral. No caso da migração, a resistência multilateral associa-se à influência de outras regiões na determinação do fluxo migratório entre duas regiões específicas. A mobilidade entre um par de regiões não depende apenas de seus fatores de atração, mas também dos fatores de atração e repulsão de todas as origens e destinos alternativos (GORIN, 2016). Desconsiderar a influência de destinos alternativos potenciais pode enviesar os resultados da análise (BERTOLI; MORAGA, 2015; RAMOS, 2016). O mesmo raciocínio é válido para a resistência multilateral na colaboração em pesquisa. Ignorar regiões parceiras potenciais pode enviesar os resultados. Os termos de resistência multilateral controlam os custos de transação de uma região em relação a todas as demais regiões com as quais possui algum fluxo. Anderson e van Wincoop (2003) sugerem aumentar a equação gravitacional tradicional com efeitos fixos da origem e do destino para controlar a resistência multilateral. Os termos de resistência multilateral podem incluir variáveis não observáveis, de forma que a inclusão de efeitos fixos mitiga o potencial viés decorrente da omissão de variáveis (ANDERSON; YOTOV, 2012; RAMOS, 2016). Em uma abordagem *cross-section* isto implica incluir *dummies* para a região de origem e para a região de destino (BURGER, 2009).

Seguindo as orientações de Anderson e van Wincoop (2003) que têm sido observadas em várias aplicações empíricas sobre migração (BERTOLI; MORAGA, 2015; RAMOS; SURIÑACH, 2013) o conjunto de variáveis de atração e repulsão específicas a cada região e que influenciam a mobilidade inter-regional, E_i e E_j , serão captadas pelos efeitos fixos da origem e do destino, e^{δ_i} e e^{λ_j} . A inclusão de efeitos fixos da origem e do destino em modelos gravitacionais permite controlar adequadamente os termos de resistência multilateral, sendo consistente com a preocupação teórica em relação à correta especificação destes modelos, gerando estimativas mais consistentes (ANDERSON; VAN WINCOOP, 2003). Introduzindo efeitos fixos da origem e do destino na equação (11) tem-se o modelo gravitacional básico para a mobilidade de inventores:

$$M_{ij} = e^{\beta_0} G_{ij}^{\varphi} e^{\psi_{ij}} T_{ij}^{\tau} e^{\gamma_{ij}} \prod_{i=1}^I e^{\delta_i} \prod_{j=1}^J e^{\lambda_j} \varepsilon_{ij} \quad (12)$$

M_{ij} = mobilidade entre as regiões i e j

e^{β_0} = termo constante que captura o impacto de todos os fatores comuns que afetam a mobilidade entre todas as regiões

G_{ij}^{φ} = distância euclidiana entre os centroides das regiões i e j

$e^{\psi_{ij}}$ = *dummy* igual a 1 (um) se duas regiões possuem fronteira comum

T_{ij}^{τ} = proximidade tecnológica entre as regiões i e j

$e^{\gamma_{ij}}$ = *dummy* igual a 1(um) se duas regiões pertencem o mesmo estado

e^{δ_i} = efeitos fixos da origem

e^{λ_j} = efeitos fixos do destino

ε_{ij} = termo de erro bem comportado

A migração inter-regional será construída pela contagem dos movimentos de inventores entre as fronteiras regionais a cada ano. Serão construídas matrizes assimétricas de mobilidade de 558 linhas e 558 colunas (são 558 microrregiões) para cada ano. Cada elemento na matriz é o número de inventores migrando da região j para a região i . Se um inventor se muda mais de uma vez, ou se retorna para a região anterior, estes movimentos serão considerados separados e independentes. Por definição movimentos da região i para a região i não existem e os elementos da diagonal da matriz são nulos.

3.2.2 Interações de conhecimento por meio das coinvenções inter-regionais

Redes inter-regionais de conhecimento representam troca de conhecimento entre atores localizados em regiões diferentes. Uma troca de conhecimento ocorre sempre que os benefícios obtidos de acessar novo conhecimento forem maiores que os custos associados (MORESCALCHI *et al.*, 2015). Em redes regionais, esses custos se relacionam a diferentes formas de proximidade: física, institucional, cognitiva, social e organizacional (BOSCHMA, 2005). A probabilidade de troca de conhecimento é positivamente afetada por estas diferentes formas de proximidade. Dado o recorte regional de nossa análise, seguiremos a estratégia de Miguélez e Moreno (2013a) e consideraremos apenas as proximidades física, institucional e cognitiva.

Conforme Miguélez e Moreno (2013a), os *payoffs* da coinvenção entre os indivíduos k e h podem ser descritos por:

$$\pi^{kh} = f(X^k, X^h) + \varepsilon^{kh} \quad (13)$$

Onde X^k e X^h são as características observáveis de k e h , e ε^{kh} é o termo de erro estocástico específico ao par k e h . Para cada par de inventores, k e h , um *link* se forma se, e

somente se, os *payoffs* esperados associados à colaboração são, no mínimo, positivos, $\pi^{kh} > 0$. Mais especificamente, a colaboração ocorrerá quando seus *payoffs* forem maiores que os custos, $c(D^{kh})$, ou seja:

$$\pi^{kh} > c(D^{kh}) \quad (14)$$

Se a condição (14) é satisfeita, haverá colaboração e a variável coinvenção (C_{ij}^{kh}) será igual a 1. Se esta condição não for satisfeita, não haverá colaboração e a variável coinvenção (C_{ij}^{kh}) será igual a 0. Agregando todas as decisões individuais por pares de regiões tem-se a especificação do modelo gravitacional de colaboração inter-regional:

$$C_{ij} = f(X_i, X_j, D_{ij}) \quad (15)$$

Onde C_{ij} é a soma das escolhas individuais dos inventores da região i em colaborar com alguém da região j , o que depende das características de um inventor representativo da região i (X_i) e das características de um inventor representativo da região j (X_j). As características dos inventores representativos das regiões i e j são captadas pelos efeitos fixos das regiões de origem, e^{δ_i} , e de destino, e^{λ_j} , de forma que o foco é em D_{ij} , o mesmo conjunto de variáveis bilaterais de distância consideradas no modelo para migração de inventores:

$$D_{ij} = G_{ij}^{\varphi} e^{\psi_{ij}} T_{ij}^{\tau} e^{\gamma_{ij}} \quad (16)$$

Onde G_{ij} é a distância euclidiana entre os centroides das regiões i e j que tende a afetar negativamente as colaborações entre os inventores. A proximidade espacial facilita a formação de redes e colaborações formais, pois facilita a triagem de parceiros potenciais, o gerenciamento de projetos em comum, e o monitoramento do desempenho do parceiro (MIGUÉLEZ; MORENO, 2013a). A contiguidade espacial tende a facilitar a formação de redes, por isso será incluída uma *dummy* para regiões com fronteira comum, $e^{\psi_{ij}}$. A proximidade tecnológica, T_{ij} , será introduzida para representar o quanto a proximidade cognitiva explica o copatenteamento entre comunidades epistêmicas fisicamente distantes. Também será incluída uma *dummy* indicando se duas regiões pertencem à mesma unidade da federação, $e^{\gamma_{ij}}$, para captar o efeito da proximidade institucional.

Introduzindo os efeitos fixos da origem e do destino na equação (16) e combinando com a equação (15) tem-se o modelo gravitacional para as colaborações inter-regionais do conhecimento:

$$C_{ij} = e^{\beta_0} G_{ij}^{\varphi} e^{\psi_{ij}} T_{ij}^{\tau} e^{\gamma_{ij}} \prod_{i=1}^I e^{\delta_i} \prod_{j=1}^J e^{\lambda_j} \varepsilon_{ij} \quad (17)$$

Onde C_{ij} é o número de coinvenções entre as regiões i e j , e^{β_0} é um termo constante que captura o impacto de todos os fatores comuns que afetam a colaboração entre todas as

regiões e ε_{ij} é um termo de erro bem comportado. As demais variáveis são as mesmas utilizadas no modelo gravitacional para mobilidade dos inventores.

As variáveis geográficas utilizadas nos modelos gravitacionais acima também foram empregadas por Miguélez e Moreno (2015) para estimar a mobilidade e a coinvenção inter-regional. Esses autores utilizaram os valores estimados nesses modelos como instrumentos para a mobilidade e a rede de coinvenção em sua função de produção do conhecimento regional, pois são determinados por variáveis exógenas que afetam a inovação por meio da mobilidade e da rede de coinvenção.

As próximas seções descrevem as estratégias econométricas que serão empregadas para estimar os modelos gravitacionais e a função de produção do conhecimento regional apresentados nesse tópico.

3.3 Abordagem econométrica

3.3.1 Estimação dos modelos gravitacionais

Tem sido prática comum nas aplicações empíricas a *log* linearização do modelo gravitacional e sua estimação por Mínimos Quadrados Ordinários (MQO) (SANTOS SILVA; TENREYRO, 2006; RAMOS; SURIÑACH, 2013). Mas esse procedimento possui duas limitações significativas. A primeira refere-se ao valor nulo da variável dependente representando os pares de regiões que não possuem fluxo de migração de inventores ou relações de coinvenção. A estratégia mais usada para lidar com esse problema é omitir estas observações ou adicionar um pequeno valor positivo, usualmente 0,5 ou 1 (EICHENGREEN; IRWIN, 1995; RAMOS; SURIÑACH, 2013; REIS *et al.*, 2014). A primeira alternativa implica em desconsiderar informações relevantes sobre pares de regiões em que não há movimentos migratórios ou coinvenções. A segunda opção também é problemática, pois a escolha da constante a ser adicionada é arbitrária, sem justificativa teórica e/ou empírica, e pequenas variações no número escolhido pelo pesquisador podem provocar grandes variações nos resultados (BURGER, 2009). Esses procedimentos criam problemas de retransformação se o interesse é em $E(y|\mathbf{x})$ e não em $E(\ln y|\mathbf{x})$ (CAMERON E TRIVEDI, 2005), pois os antilogaritmos dessas estimativas tendem a ser enviesados (BURGER, 2009). Esse viés é significativo e não diminui com o aumento do tamanho da amostra, levando a estimadores

inconsistentes dos parâmetros de interesse (SANTOS SILVA; TENREYRO, 2006; 2011; BERTOLI; MORAGA, 2015).

Para que a estimação por MQO da equação gravitacional *log* linearizada seja válida é necessário que o erro ε_{ij} , e, $\ln\varepsilon_{ij}$, sejam estatisticamente independentes dos regressores. Entretanto, o valor esperado do logaritmo de uma variável aleatória depende de sua média e de momentos de ordens maiores da distribuição. Ou seja, $E(\ln \varepsilon_{ij} | X)$ depende da média e da variância. Se a variância do erro ε_{ij} em (12) e (17) depender de alguma variável explicativa, $E(\ln \varepsilon_{ij} | X)$ também dependerá destas variáveis explicativas, violando a condição de consistência do estimador de MQO. Portanto, se ε_{ij} do modelo multiplicativo for heterocedástico a estimação do modelo *log* linear por MQO é inconsistente (SANTOS SILVA; TENREYRO, 2006).

Os dados de contagem são intrinsecamente heterocedásticos, sendo caracterizados por variância crescente com a média (CAMERON; TRIVEDI, 2005). Portanto, a abordagem tradicional de estimação do modelo *log* linearizado por MQO é inviável, uma vez que a condição de consistência deste estimador é naturalmente violada, devido à desigualdade de Jensen¹⁰.

Santos Silva e Tenreyro (2006) propõe a estimação do modelo gravitacional em nível utilizando o estimador *Poisson-Pseudo Maximum Likelihood* (PPML) como uma forma de contornar os dois problemas referidos anteriormente¹¹. O estimador PPML se baseia na distribuição de Poisson, cuja função de probabilidade é dada por:

$$P(Y = y) = \frac{e^{-\mu} \cdot \mu^y}{y!}, \quad y = 0, 1, 2, \dots \quad (18)$$

Onde Y é uma variável de contagem, como o número de migrações ou de colaborações em pesquisa, e μ é a intensidade ou parâmetro de taxa. Os dois primeiros momentos definem a propriedade de equidispersão (igualdade entre a média e a variância condicional) da distribuição de Poisson, frequentemente não observada nas aplicações empíricas (WOOLDRIDGE, 2002):

$$E(Y) = \mu \quad \text{e} \quad V(Y) = \mu \quad (19)$$

¹⁰ A desigualdade de Jensen implica que $E(\ln y) \neq E(y)$.

¹¹ Uma alternativa seria o estimador de mínimos quadrados não lineares da regressão não linear $y = \exp(x'\beta) + u$ (por exemplo, JAFFE; TRAJTENBERG, 2002, capítulo 7). Mas para dados de contagem este estimador é geralmente menos eficiente que o PPML, sendo ótimo apenas em casos de homocedasticidade. O PPML é um modelo melhor para lidar com a heterocedasticidade inerente dos dados de contagem (CAMERON; TRIVEDI, 2005).

O modelo de regressão de Poisson é derivado da distribuição de Poisson por meio da parametrização da relação entre o parâmetro da média, μ , e os regressores, \mathbf{x} . O procedimento padrão é utilizar a parametrização exponencial da média (CAMERON; TRIVEDI, 2005):

$$\mu_i = \exp(\mathbf{x}'_i\beta), \quad i = 1, \dots, N \quad (20)$$

As condições (19) e (20) implicam que a regressão de Poisson é intrinsecamente heterocedástica:

$$V(y_i|\mathbf{x}_i) = \exp(\mathbf{x}'_i\beta) \quad (21)$$

Em contraste com a especificação *log* linear do modelo gravitacional, o modelo de regressão de Poisson gera estimativas de y e não de $\ln y$, evitando o viés da transformação logarítmica; as estimativas são consistentes na presença de heterocedasticidade; e, devido à sua forma multiplicativa, fornece um meio natural de lidar com os zeros da variável dependente (BURGER, 2009; SANTOS SILVA; TENREYRO, 2011; BERTOLI; MORAGA, 2015).

Assumindo que as observações $(y_i|\mathbf{x}_i)$ são independentes e usando (18) e (20) obtém-se a *log* verossimilhança para a observação i :

$$l_i(\beta) = y_i\mathbf{x}'_i\beta - \exp(\mathbf{x}'_i\beta) - \ln y_i! \quad (22)$$

Para a amostra a função de *log* verossimilhança é dada por:

$$L(\beta) = \sum_{i=1}^N l_i(\beta) = \sum_{i=1}^N \{y_i\mathbf{x}'_i\beta - \exp(\mathbf{x}'_i\beta) - \ln y_i!\} \quad (23)$$

O estimador de máxima verossimilhança de Poisson, $\hat{\beta}_P$, é a solução para K equações não lineares, que corresponde à condição de primeira ordem para a máxima verossimilhança:

$$\sum_{i=1}^N [y_i - \exp(\mathbf{x}'_i\beta)]\mathbf{x}_i = 0 \quad (24)$$

O valor esperado do lado esquerdo de (24) será zero se $E(y_i|\mathbf{x}_i) = \exp(\mathbf{x}'_i\beta)$. Portanto, o estimador de máxima verossimilhança de Poisson será consistente sob a hipótese mais fraca de correta especificação da média condicional de y dado \mathbf{x} , ou seja, os dados não precisam ser distribuídos como uma distribuição de Poisson. Neste caso, tem-se o estimador de pseudo-máxima verossimilhança (*Poisson pseudo-maximum likelihood* – PPML) (CAMERON; TRIVEDI, 2005; BERTOLI; MORAGA, 2015)¹².

O modelo de regressão de Poisson é criticado por ser muito restritivo, pois a distribuição é definida em termos de um único parâmetro escalar, a média μ , de forma que todos os momentos da distribuição são uma função de μ . Isto levaria a duas situações: i) o

¹² Também chamado de estimador de quase máxima verossimilhança (*Poisson quase-maximum likelihood estimator* – QMLE) (CAMERON; TRIVEDI, 2005). Sobre a robustez dos estimadores de máxima verossimilhança à má especificação da função de distribuição de probabilidade utilizada na verossimilhança, ver Gouriéroux et al. (1984).

excesso de zeros - a distribuição de Poisson prevê um número de zeros bem menor que o realmente observado nos dados; e ii) a superdispersão: a variância dos dados excede a média, violando a equidispersão assumida pela condição (19). Dado que a inovação no Brasil é altamente concentrada em algumas regiões, a migração de inventores e a colaboração bilateral para inovar também são restritas a algumas regiões, como será visto na análise descritiva dos dados no capítulo 4 (quatro), resultando em elevada proporção de zeros nas variáveis dependentes dos modelos gravitacionais. Assim, é provável que os dados sejam caracterizados pela superdispersão e pelo excesso de zeros.

Mas, desde que a média condicional esteja corretamente especificada, ou seja, que (20) seja válido, o estimador de Poisson por máxima verossimilhança continua consistente na presença de superdispersão¹³. Isto pode ser percebido pela inspeção da condição de primeira ordem dada pela equação (24), pois o lado esquerdo da equação terá valor esperado igual a zero se $E(y_i|\mathbf{x}_i) = \exp(\mathbf{x}_i'\beta)$. O PPML é bem comportado mesmo quando a variância está longe de ser proporcional à média, e sua performance não é afetada mesmo quando proporção de zeros da variável dependente é muito grande (SANTOS SILVA; TENREYRO, 2011; FIGUEIREDO, *et al.*, 2015). No primeiro caso, basta utilizar o erro padrão robusto para que as estimativas sejam consistentes, mesmo se não houver igualdade entre a média e a variância (BERTOLI; MORAGA, 2015).

Estimar o modelo gravitacional com efeitos fixos da origem e do destino utilizando o estimador PPML consiste em uma forma simples de resolver o problema da resistência multilateral abordado por Anderson e Van Wincoop (2003) (FALLY, 2015). Por esses motivos, seguindo as aplicações empíricas mais recentes de modelos gravitacionais (ANDERSON; YOTOV, 2012; FALLY, 2015; SIRRIES, 2015; FIGUEIREDO, *et al.*, 2015) utilizaremos o estimador PPML com efeitos fixos da origem e do destino para estimar os modelos gravitacionais para migração de inventores e para coinvenção inter-regional.

3.3.2 Estimação da função de produção do conhecimento regional

Diferentemente de Miguélez e Moreno (2013a), que estimaram a função de produção do conhecimento regional por meio de um painel de dados usando o estimador de efeito fixo, nessa tese a função de produção do conhecimento regional será estimada por meio de um painel dinâmico espacial usando o estimador *system gmm*.

¹³ A superdispersão possui consequências qualitativamente similares à não observância da hipótese de homocedasticidade no modelo de regressão linear (CAMERON; TRIVEDI, 2005).

A estimação em painel é importante para controlar a heterogeneidade não observada das regiões brasileiras que é invariante no tempo. Esta heterogeneidade se refere às características institucionais que podem afetar a inovação, tais como políticas tecnológicas regionais, habilidades herdadas da comunidade local, perfil das instituições educacionais superiores, cultura inovativa herdada, e todas as características históricas *path dependent* que podem gerar diferenças espaciais nas taxas de inovação (MIGUÉLEZ; MORENO, 2015; GONÇALVES *et al.*, 2018). Estas características são críticas para a distribuição espacial contemporânea da atividade inovativa (e econômica) das regiões brasileiras e têm que ser consideradas no processo de estimação, sob pena de enviesar significativamente as estimativas obtidas.

Mas, além da heterogeneidade não observada das regiões, é importante também considerar explicitamente o caráter *path dependent* da inovação, pois a difusão do conhecimento ao longo do tempo torna a inovação contemporânea dependente das inovações anteriores. Feldman e Florida (1994) mostram que a capacidade regional para inovar depende da infraestrutura tecnológica local, a qual é um legado histórico de concentrações especializadas de PeD, atividade industrial, e serviços de suporte, que se constroem em dado lugar ao longo do tempo. As evidências encontradas por estes autores confirmam que o desempenho contemporâneo das atividades de inovação em uma região é dependente de sua trajetória tecnológica anterior. Além disto, a persistência temporal da inovação também reflete o fato de que os agentes podem não responder imediatamente aos incentivos, demandando tempo para obter informação e para aprendizagem.

Novas tecnologias surgem a partir de um contexto de conhecimentos propiciado pela prática em trabalhar com certas tecnologias existentes (na forma de componentes, operações e métodos) e certos fenômenos recém-descobertos. Esta prática consiste não apenas em conhecer funcionalidades e em como combiná-las, mas também em saber o que provavelmente não funciona, quais métodos utilizar, com quem conversar, a quais teorias recorrer, e, acima de tudo, como manipular o fenômeno recém-descoberto e ainda pouco compreendido. Para isto, o acesso a tecnologias existentes, o investimento em laboratórios e o conhecimento do fenômeno científico são importantes, mas não suficientes. A liderança nacional na criação de tecnologias avançadas decorre dos conhecimentos estabelecidos há longa data de como trabalhar com um novo fenômeno em particular e suas funcionalidades associadas. Essa habilidade precisa ser cultivada ao longo de décadas nos cenários locais com financiamento estável e encorajamento. Ela é alimentada pelas universidades e é localizada,

pois tende a ser compartilhada por um pequeno número de pessoas confinadas em certos laboratórios ou certas regiões (ARTHUR, 2007).

O efeito do conhecimento externo sobre a inovação regional será avaliado no contexto da função de produção do conhecimento regional por meio da defasagem da variável dependente, utilizando como pesos de ponderação: a) a matriz de pesos espaciais tradicional, baseada na geografia, e b) as matrizes de interação do conhecimento estimadas pelos modelos gravitacionais para mobilidade de inventores e para coinvenção inter-regional da primeira etapa. O emprego da matriz de pesos baseada na mobilidade bilateral de inventores e da matriz de pesos baseada na colaboração inter-regional para ponderar o estoque de conhecimento das demais regiões, permitirá avaliar a contribuição destes dois canais de difusão do conhecimento para a inovação regional e sua importância em relação às demais formas de fluxo de conhecimento agrupadas sob a defasagem espacial.

A defasagem temporal da variável dependente é intrinsecamente endógena por ser correlacionada com o efeito não observado. A endogeneidade da defasagem espacial se deve ao efeito espacial multidirecional, que gera simultaneidade com a variável dependente. A presença de ambas endogeneidades inviabiliza a estimação por Mínimos Quadrados Ordinários (MQO) e por Efeito Fixo (EF).

As matrizes de pesos baseadas na geografia são exógenas, sendo adequadas para introduzir variações exógenas nos modelos e permitir a identificação dos efeitos exógenos e endógenos. Por outro lado, as matrizes de interação do conhecimento não são exógenas, pois a migração e a colaboração para inovar afetam e são afetadas pela performance inovadora regional. A imigração de inventores tende a afetar positivamente a inovação regional, uma vez que o inventor imigrante carrega consigo o conhecimento tácito incorporado, extremamente relevante para a inovação. Por outro lado, como mostrado por Faggian e McCann (2009) e Gagliardi (2014) para a Inglaterra, e por Gonçalves *et al.* (2012) para o Brasil, existe simultaneidade entre a migração de mão de obra qualificada e inovação regional. Portanto, as regiões inovadoras também tendem a atrair inventores, havendo uma relação de simultaneidade entre estes fenômenos. Da mesma forma, é provável que haja simultaneidade também entre a colaboração inter-regional para inovar e o desempenho inovador, pois a colaboração estimula a inovação regional, mas a própria atividade inovadora induz ao aumento das colaborações inter-regionais. Esta causalidade reversa na matriz de pesos pode enviesar as estimativas e torná-las inconsistentes, mesmo para amostras infinitas (ALMEIDA, 2012).

Entretanto, as matrizes de interação do conhecimento têm a vantagem de refletir diretamente os canais concretos que permitem a difusão do conhecimento no espaço, não se baseando meramente na proximidade geográfica como mediadora desta difusão. Além disto, a inclusão de variáveis exógenas na estimação dos modelos gravitacionais, tais como a distância euclidiana e *dummies* de contiguidade, permite captar, até certo ponto, o componente exógeno da mobilidade e das colaborações e colocá-lo nas matrizes de peso (MIGUÉLEZ; MORENO, 2013a).

A especificação geral da função de produção do conhecimento regional a ser estimada pode ser representada da seguinte forma:

$$Pat_{i,t} = \beta_0 + \beta_1 Pat_{i,t-1} + \beta_2 W^c Pat_{-i,t} + \beta_3 W^m Pat_{-i,t} + \beta_4 W^e Pat_{-i,t} + \beta_5 PeD_{i,t} + \beta_6 PeD_{uni,t} + \beta_7 CHI_{i,t} + \beta_8 Z_{i,t} + D_t + \alpha_i + \varepsilon_{i,t} \quad (25)$$

Onde:

$Pat_{i,t}$ é o número de patentes da região i a cada cem mil habitantes;

W^c e W^m são as matrizes de pesos estimadas pelos modelos gravitacionais para coinvenção e mobilidade inter-regionais, respectivamente;

$Pat_{-i,t}$ é o número de patentes de todas as regiões, exceto a região i , a cada cem mil habitantes;

W^e é uma matriz de pesos espaciais, baseada na geografia;

$PeD_{i,t}$ é o gasto regional com pesquisa e desenvolvimento das empresas;

$PeD_{uni,t}$ é o gasto regional com pesquisa e desenvolvimento das universidades;

$CHI_{i,t}$ é o capital humano regional;

Z_t representa um conjunto de variáveis locais relevantes para o desempenho regional da inovação;

D_t são variáveis *dummy* anuais, para captar efeitos temporais comuns a todas as regiões;

α_i é o efeito fixo regional não observável e invariante no tempo;

$\varepsilon_{i,t}$ é o erro idiossincrático, independente e identicamente distribuído (iid), com média zero e variância σ^2 .

Considera-se que α_i e $\varepsilon_{i,t}$ sejam independentes entre as regiões a cada período t .

O estimador *system gmm* é adequado para estimar este painel dinâmico espacial, pois permite controlar as fontes de endogeneidade explicadas acima: a heterogeneidade não observada e invariante no tempo das microrregiões brasileiras, endogeneidade intrínseca da

defasagem temporal, da defasagem espacial e das defasagens utilizando-se os pesos das matrizes de interação do conhecimento.

Devido à essas fontes de endogeneidade, a estimação do modelo por mínimos quadrados ordinários gera estimativas enviesadas e inconsistentes. O coeficiente da defasagem temporal da variável dependente é enviesado para cima, pois $Pat_{i,t-1}$ é positivamente correlacionada com o termo de erro, gerando o chamado viés de Nickell (1981). Isso atribui poder preditivo à variável endógena, que na verdade pertence ao efeito fixo que está no termo de erro (ROODMAN, 2006). Uma alternativa para resolver este problema é usar um estimador que expurgue o efeito fixo não observável das microrregiões.

O estimador de efeito fixo elimina o efeito fixo individual, mas a endogeneidade permanece, pois a transformação *within* resulta em $Pat_{i,t-1}^* = Pat_{i,t-1} - \frac{1}{T-1}(Pat_{i,2} + \dots + Pat_{i,T})$ e no erro transformado $\varepsilon_{i,t}^* = \varepsilon_{i,t} - \frac{1}{T-1}(\varepsilon_{i,2} + \dots + \varepsilon_{i,T})$. Observe que o termo $Pat_{i,t-1}$ é negativamente correlacionado com $-\frac{1}{T-1}\varepsilon_{i,t-1}$, e que $-\frac{1}{T-1}Pat_{i,t}$ é negativamente correlacionado com $\varepsilon_{i,t}$, gerando endogeneidade. Essa correlação negativa leva o estimador de efeito fixo a subestimar o coeficiente do *lag* da variável dependente¹⁴. O termo endógeno $Pat_{i,t-1}^*$ não pode ser instrumentalizado com *lags* de $Pat_{i,t-1}$ pois estes *lags* também estão embutidos no erro transformado $\varepsilon_{i,t}^*$.

Outra forma de captar a heterogeneidade individual é a transformação do modelo em primeira diferença. A primeira diferença elimina o efeito fixo, permitindo estimativas consistentes. Arellano-Bond (1991) propôs estimar esta equação pelo método generalizado dos momentos (*gmm*), que minimiza as condições de momento da distribuição. Conhecido como *difference gmm*, este método estima a equação transformada:

$$\Delta Pat_{i,t} = \beta_1 \Delta Pat_{i,t-1} + \beta_2 \Delta W^c Pat_{-i,t} + \beta_3 \Delta W^m Pat_{-i,t} + \beta_4 \Delta W^e Pat_{-i,t} + \beta_5 \Delta PeD_{i,t} + \beta_6 \Delta PeD_{uni,t} + \beta_7 \Delta CHI_{i,t} + \beta_8 \Delta Zi_{i,t} + \Delta Dt + \Delta \varepsilon_{i,t} \quad (26)$$

Embora o efeito fixo tenha sido eliminado, a endogeneidade permanece, pois $\Delta Pat_{i,t-1} = Pat_{i,t-1} - Pat_{i,t-2}$ é correlacionado com $\Delta \varepsilon_{i,t} = \varepsilon_{i,t} - \varepsilon_{i,t-1}$ por meio dos

¹⁴ Um teste informal para avaliar se a estimativa do *lag* da variável dependente é adequada (não enviesada) é observar se o coeficiente está entre os valores estimados pelo efeito fixo e por mínimos quadrados ordinários (BOND, 2002; ROODMAN, 2006).

termos $Pat_{i,t-1}$ e $\varepsilon_{i,t-1}$. Mas, os *lags* anteriores da variável endógena permanecem ortogonais ao erro (ao contrário da transformação *within*) podendo ser usados como instrumentos.

Os instrumentos para $\Delta Pat_{i,t-1}$ podem ser $Pat_{i,t-2}$ e $\Delta Pat_{i,t-2}$, pois estes termos estão relacionados a $\Delta Pat_{i,t-1} = Pat_{i,t-1} - Pat_{i,t-2}$ e são ortogonais ao termo de erro $\Delta \varepsilon_{i,t} = \varepsilon_{i,t} - \varepsilon_{i,t-1}$, desde que $\varepsilon_{i,t}$ seja não correlacionado serialmente. Mas é preferível instrumentalizar com $Pat_{i,t-2}$ em vez de $\Delta Pat_{i,t-2}$ para maximizar o tamanho da amostra (ROODMAN, 2006).

Todos os *lags* da variável em nível podem ser usados como instrumentos para a equação em diferenças (equação 31). Para variáveis endógenas, isto significa *lag* 2 em diante e para variáveis pré-determinadas o *lag* 1 em diante, pois $\Delta \varepsilon_{i,t}$ depende de $\varepsilon_{i,t-1}$ e a variável pré-determinada $Pat_{i,t-1}$ é potencialmente correlacionada apenas com os erros $\varepsilon_{i,t-2}$ em diante. Logo, $\Delta Pat_{i,t-1}$ pode ser instrumentalizada por $Pat_{i,t-2}$. A defasagem espacial bem como as defasagens usando as matrizes de pesos estimadas nos modelos gravitacionais também são endógenas. Portanto, o segundo *lag* e os *lags* mais distantes destas variáveis em nível podem ser utilizados para instrumentalizar as variáveis em diferenças. O mesmo raciocínio é válido para as demais variáveis explicativas endógenas.

Para aumentar a eficiência do estimador, Arellano e Bover (1995) e Blundell e Bond (1998) propõe estimar as equações (25) e (26) simultaneamente, dando origem ao estimador *system gmm*. Nesse sistema, a equação em nível (25) é instrumentalizada pelas primeiras diferenças defasadas e a equação em diferenças (26) é instrumentalizada pelas variáveis em nível defasadas. A primeira diferença dos instrumentos os torna exógenos ao efeito fixo, permitindo o aumento do número de instrumentos e, conseqüentemente, a eficiência do estimador. Isto é válido assumindo-se que as mudanças nos instrumentos não sejam correlacionadas com o efeito fixo. A validade dos instrumentos adicionais no *system gmm* requer que, ao longo do período de estudo, as observações da amostra estejam em um tipo de *steady state*, no sentido de que desvios dos valores de longo prazo, controlado pelos regressores, não são sistematicamente relacionados aos efeitos fixos (ROODMAN, 2006).

Por ser mais eficiente, o estimador *system gmm* será utilizado para a estimação da função de produção do conhecimento regional. O processo gerador dos dados em questão possui as características necessárias para o uso adequado do estimador *system gmm*: 1) o processo é dinâmico, pois a inovação regional em t tende a ser influenciada pelas inovações da região em períodos anteriores; 2) existe efeito fixo individual arbitrariamente distribuído entre as regiões; 3) algumas variáveis explicativas são endógenas; 4) o erro idiossincrático pode ter

padrões individuais específicos de heterocedasticidade e autocorrelação serial; 5) o erro idiossincrático não é correlacionado entre as regiões¹⁵; 6) há um número pequeno de períodos (os dados se referem ao período de 2001 a 2010) e um número grande de observações (558 microrregiões); e 7) os *lags* das variáveis a serem instrumentalizadas estão disponíveis como instrumentos (ROODMAN, 2006).

A consistência do estimador *system gmm* depende da exogeneidade das variáveis explicativas e da ausência de correlação serial dos resíduos em nível. A primeira condição é avaliada pelo teste Hansen (teste J), que testa a hipótese conjunta de correta especificação do modelo e de ortogonalidade. Sua rejeição implica que as condições de ortogonalidade não são satisfeitas¹⁶, o que pode indicar que os instrumentos não são realmente exógenos ou que foram incorretamente excluídos da regressão (BAUM *et. al.*, 2003). A segunda condição para a consistência do estimador *gmm system* é a ausência de correlação serial dos resíduos em nível, o que é verificado pelo teste de autocorrelação de Arellano-Bond (AR). A transformação em primeira diferença introduz a autocorrelação dos resíduos em diferenças. Dado que $\Delta\varepsilon_{i,t} = \varepsilon_{i,t} - \varepsilon_{i,t-1}$ é matematicamente correlacionado a $\Delta\varepsilon_{i,t-1} = \varepsilon_{i,t-1} - \varepsilon_{i,t-2}$, a presença de autocorrelação negativa de primeira ordem em diferenças já é esperada. Portanto, para checar se há autocorrelação serial de primeira ordem em nível, verifica-se a correlação de segunda ordem em diferenças, pois isto irá detectar correlação entre $\varepsilon_{i,t-1}$ e $\varepsilon_{i,t-2}$ presentes nos erros em diferença, $\Delta\varepsilon_{i,t}$ e $\Delta\varepsilon_{i,t-2}$, respectivamente. Para que o modelo seja válido, o teste Arellano-Bond deve indicar presença de autocorrelação de primeira ordem em diferenças e ausência de autocorrelação de segunda ordem em diferenças.

O estimador *system gmm* pode ser calculado em um ou dois passos. A estimação do primeiro passo supõe que os resíduos sejam esféricos, e a do segundo passo utiliza os resíduos da primeira estimação para obter uma estimativa consistente da matriz de variância-covariância. O estimador do segundo passo é mais eficiente e robusto à heterocedasticidade, mas gera erros padrão viesados para baixo quando o número de instrumentos é grande. Windmeijer (2005) propôs uma correção destes erros padrão para pequenas amostras, tornando a estimação do segundo passo, com erros padrão corrigidos, superior ao primeiro passo robusto.

¹⁵ Para atender a esta hipótese, serão incluídas *dummies* de ano para captar choques comuns que afetem todas as regiões.

¹⁶ Se os erros forem esféricos o teste de Hansen coincide com o teste de Sargan para a hipótese nula de exogeneidade dos instrumentos, mas se os erros forem não esféricos, o teste de Sargan é inconsistente e a estatística de Hansen é superior (ROODMAN, 2006).

Tendo apresentado a estratégia metodológica e econométrica para analisar o efeito da mobilidade e da colaboração inter-regional entre inventores sobre a inovação local, o próximo capítulo descreve a organização da base de dados. Além disso, apresenta uma análise descritiva dos mesmos e explica a construção das variáveis que serão utilizadas nos modelos econométricos.

4 ORGANIZAÇÃO DO BANCO DE DADOS E ANÁLISE DESCRITIVA

Os dados sobre as patentes foram obtidos da primeira versão da Base de Dados sobre Propriedade Intelectual para fins Estatísticos (BADEPI v1.0), elaborada pelo Instituto Nacional de Propriedade Intelectual (INPI). Esta base de dados compreende informações sobre patentes depositadas no INPI no período de 2000 a 2011, que receberam o despacho de patente concedida. Esse capítulo descreve o processo de construção e compilação da base de dados, cujas principais etapas são detalhadas nos tópicos 4.1 e 4.2 e são resumidas na figura 01

4.1 Identificando os inventores

As patentes são identificadas pelo número do processo, classificadas em patentes de invenção (PI) e modelos de utilidade (MU)¹⁷. Nessa tese, ambos foram considerados, indistintamente. Os dados da BADEPI estavam originalmente organizados em dois arquivos, um com informações sobre os depositantes das patentes e outro com as informações sobre os inventores. Cada patente possui a identificação do inventor e do depositante por meio do CPF/CNPJ, nome completo e município de domicílio ou sede. A identificação do inventor é essencial para analisar sua mobilidade entre regiões, bem como suas relações de coinvenção com pares de outros locais.

O arquivo inicial de inventores continha observações sobre 87.322 processos diferentes, sendo a maioria de patentes de invenção (55,37%) distribuídos em 1792 municípios diferentes (tabela 01):

¹⁷ Os direitos garantidos pelos modelos de utilidade são muito parecidos com os garantidos pelas patentes de invenção, mas os modelos de utilidade são mais adequados a invenções incrementais. Os requerimentos para obtenção de um modelo de utilidade são menos rígidos que para obtenção de uma patente de invenção. A Lei de Propriedade Industrial (LPI), Lei nº 9279, de 14 de maio de 1996 estabelece que:

Art. 8º É patenteável a invenção que atenda aos requisitos de novidade, atividade inventiva e aplicação industrial.

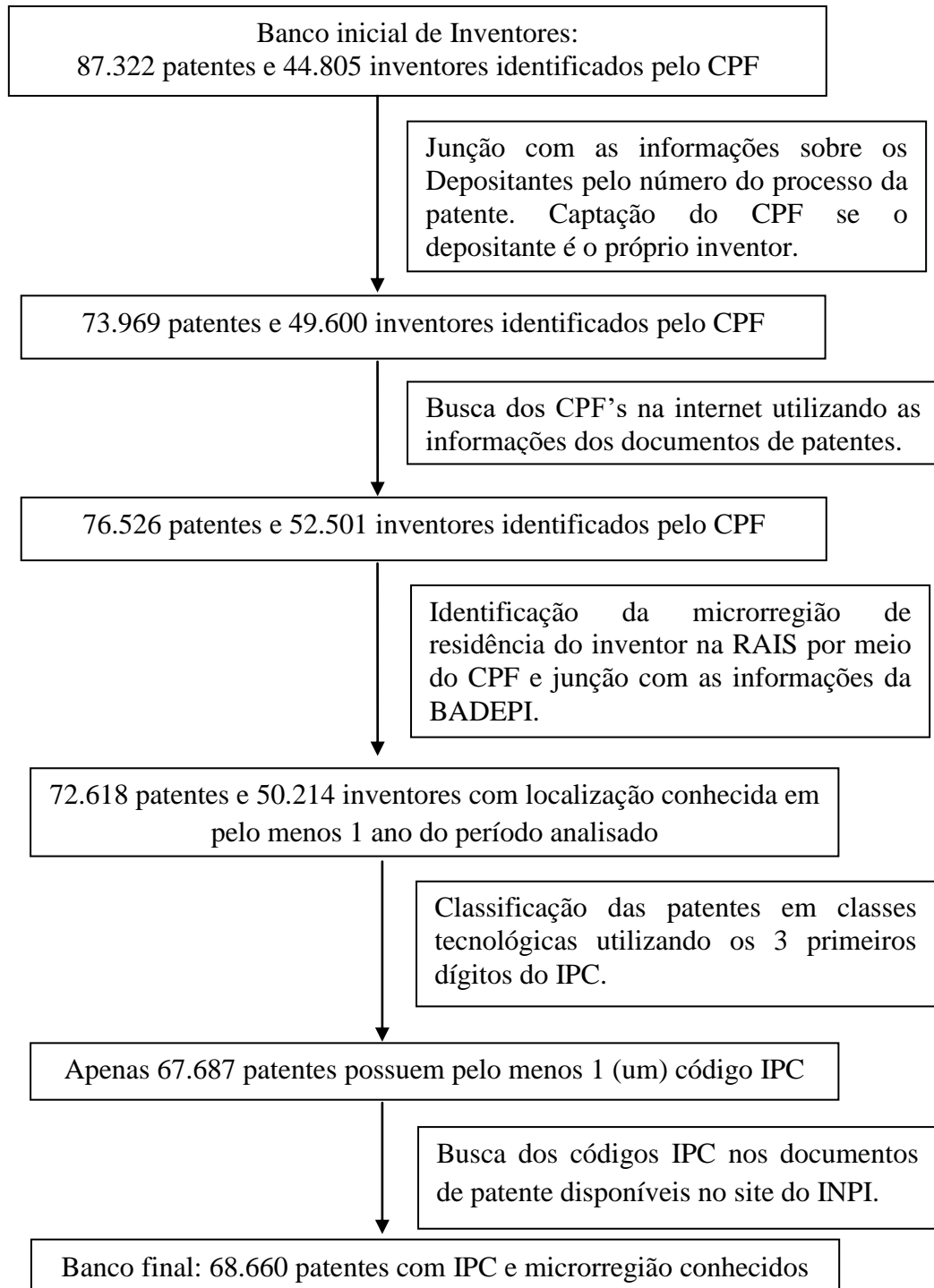
Art. 9º É patenteável como modelo de utilidade o objeto de uso prático, ou parte deste, suscetível de aplicação industrial, que apresente nova forma ou disposição, envolvendo ato inventivo, que resulte em melhoria funcional no seu uso ou em sua fabricação.

Tabela 01: Tipos de patentes dos inventores brasileiros, de 2000 a 2011

| Tipo | Frequência | Percentual (%) |
|---------------------|------------|----------------|
| Modelo de Utilidade | 38.973 | 44,63 |
| Patente de Invenção | 48.349 | 55,37 |
| Total | 87.322 | 100,00 |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI

Figura 01: Principais etapas da construção da base de dados



Fonte: Elaboração própria.

Apesar de apenas pessoas físicas (PF) poderem ser consideradas inventores, existem muitas observações (4,10%) em que o inventor é identificado como pessoa jurídica (PJ) por meio do CNPJ. Além disso, 21% das observações não possui nenhuma informação de CPF ou CNPJ. São 44.805 inventores identificados como pessoa física e 1.562 inventores identificados como pessoa jurídica, responsáveis por 87.322 patentes (tabela 02).

Tabela 02: Identificação do inventor como Pessoa Física e Pessoa Jurídica, 2000 a 2011

| Natureza Jurídica do Inventor | Total de observações ^a | | Total de inventores ^b | |
|-------------------------------|-----------------------------------|----------------|----------------------------------|----------------|
| | Frequência | Percentual (%) | Frequência | Percentual (%) |
| Pessoa Física | 91.574 | 74,80 | 44.805 | 96,63 |
| Pessoa Jurídica | 4.998 | 4,10 | 1.562 | 3,37 |
| Sem informação | 25.819 | 21,10 | - | - |
| Total | 122.391 | 100,00 | 46.367 | 100,00 |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI

^a O total de observações é maior que o total de patentes porque há patentes que possuem mais de um inventor, logo a patente se repete. ^b O total de inventores é menor que o número de patentes porque existem inventores que possuem mais de uma patente.

Dado o grande número de observações sem a informação de CPF, foi feito um trabalho de busca desta informação para ampliar a base de dados, utilizando os nomes dos inventores e as demais informações contidas na patente. Como o interesse é investigar a mobilidade inter-regional e a colaboração inter-regional entre os inventores, as informações sobre CNPJ foram desconsideradas. As informações sobre o CPF são importantes porque identificam cada inventor de forma única, permitindo seu acompanhamento ao longo do tempo para observar se houve mobilidade e possibilita o cruzamento entre os processos para identificar colaborações entre dois ou mais inventores na mesma patente.

O primeiro passo na recuperação de CPF's dos inventores foi cruzar os nomes dos inventores com os nomes dos depositantes. Foi feita a junção das informações sobre os inventores e sobre os depositantes por meio do número do processo da patente. Posteriormente, o nome do inventor foi comparado ao nome do depositante da patente. Se, para o mesmo processo, o nome do depositante é igual ao nome do inventor, é razoável crer que se trata da mesma pessoa. Neste caso, se não havia a informação de CPF para o inventor, e havia a informação de CPF/CNPJ para o depositante, esta informação foi captada para o inventor. Este procedimento permitiu a recuperação de 4.795 CPF's de inventores, totalizando 49.600 inventores diferentes identificados pelo CPF (aumento de 10,7%) e 1634 inventores identificados pelo CNPJ (tabela 03).

Tabela 03: Identificação do inventor após recuperação de CPF/CNPJ a partir dos dados dos depositantes - 2000 a 2011.

| Natureza Jurídica do Inventor | Total de observações | | Total de inventores | | |
|-------------------------------|----------------------|---------------|---------------------|-----------------------|---------------|
| | Frequência | (%) | Frequência | Variação | (%) |
| Pessoa Física | 97.116 | 79,35 | 49.600 | 4.795 (10,70%) | 96,81 |
| Pessoa Jurídica | 5.098 | 4,17 | 1634 | 72 (4,61%) | 3,19 |
| Sem informação | 20.177 | 16,49 | - | - | - |
| Total | 122.391 | 100,00 | 51.234 | 4.867 (10,50%) | 100,00 |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI

Desconsiderando-se as observações em que não há informação de CPF/CNPJ, restaram 102.214 observações. Estas informações referem-se a 76.420 patentes, sendo 55% classificadas como patente de invenção (tabela 04):

Tabela 04: Total de processos com informação de CPF/CNPJ, entre 2000 e 2011

| Tipo | Frequência | (%) |
|--------------|---------------|---------------|
| MU | 34.220 | 44,78 |
| PI | 42.200 | 55,22 |
| Total | 76.420 | 100,00 |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI

Após esta recuperação inicial de CPF's foram identificadas outras duas situações: i) casos em que o inventor possuía mais de uma patente, mas seu nome aparecia de forma diferente em cada registro devido a abreviações e/ou erros na grafia do nome; ii) casos em que a mesma patente possuía mais de um inventor, porém constava o mesmo número de CPF para todos os inventores da equipe. A primeira situação demandou a correção manual dos nomes, caso a caso, para que fosse possível captar todas as informações quando da comparação entre os nomes do inventor e do depositante de uma dada patente. Para resolver o segundo caso foi necessário identificar a qual dos coinventores pertencia o CPF registrado e a buscar o CPF dos demais, para que todos os inventores fossem devidamente considerados na autoria da patente. Essas informações foram pesquisadas na internet utilizando os dados presentes nos documentos de patentes, disponíveis no site do INPI.

Para ampliar a identificação dos inventores foi feita uma pesquisa de CPF's na internet para todos os inventores que não possuíam informação de CPF, apresentavam o CNPJ em vez do CPF, ou que possuíam o mesmo número de CPF que os demais inventores da equipe. Esta busca utilizou o nome completo do inventor e o número do processo da patente, disponíveis nos dados da BADEPI. Com essas informações foi possível acessar a patente no site do INPI

e utilizar as informações existentes nas patentes, tais como o assunto da patente, o nome do depositante, os nomes dos coinventores, o local de residência dos inventores e do depositante. Com esses dados foi possível pesquisar e encontrar o CPF de vários inventores na internet. As informações assim obtidas foram cruzadas criteriosamente com as informações da patente, permitindo identificar o CPF do inventor¹⁸.

Essas recuperações permitiram ampliar o número de inventores identificados pelo CPF. Esse avanço no banco de dados dos inventores é importante não só porque permite acompanhar sua trajetória inventiva, mas também observar sua localização geográfica ao longo do tempo por meio do cruzamento com os dados da RAIS, como será explicado no próximo tópico. O número de inventores identificado pelo CPF aumentou de 49600 para 52501, respondendo por 76526 patentes (tabela 05).

Tabela 05: Ampliação do banco de dados dos inventores por meio de busca na Internet

| | Banco original | Banco ampliado | Variação (%) |
|---|----------------|----------------|--------------|
| Inventores identificados | 49600 | 52501 | 5,85 |
| Patentes com pelo menos um inventor identificado pelo CPF | 73969 | 76526 | 3,46 |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI

Dessa forma, o banco ampliado é composto apenas por inventores denominados como pessoa física e com informações de CPF. Ele possui 101126 observações relativas a 76526 patentes e 52501 inventores no período de 2000 a 2011 (tabela 06):

Tabela 06: Total de patentes com pelo menos um inventor identificado pelo CPF, entre 2000 e 2011:

| Tipo | Patentes | (%) | Observações | (%) |
|--------------|----------|--------|-------------|--------|
| MU | 34260 | 44,77 | 39276 | 38,84 |
| PI | 42266 | 55,23 | 61850 | 61,16 |
| Total | 76526 | 100,00 | 101126 | 100,00 |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI

A quantidade de patentes é bem distribuída ao longo dos anos, embora apresente tendência levemente crescente ao longo do período (tabela 07).

¹⁸ Gorin (2016) e Miguélez e Moreno(2013a) realizaram procedimentos similares para identificar o inventor e observar sua mobilidade ao longo do tempo. Os autores utilizaram as informações constantes nos documentos de patentes para identificar corretamente o inventor, tais como nome do inventor e coinventores, nome do depositante e endereço dos mesmos.

Tabela 07: Distribuição das patentes por ano, 2000 a 2011

| ano | MU | PI | Total |
|-------|-------|-------|-------|
| 2000 | 2914 | 2899 | 5813 |
| 2001 | 3077 | 3144 | 6221 |
| 2002 | 3121 | 3113 | 6234 |
| 2003 | 3096 | 3398 | 6494 |
| 2004 | 2980 | 3471 | 6451 |
| 2005 | 2687 | 3465 | 6152 |
| 2006 | 2637 | 3424 | 6061 |
| 2007 | 2690 | 3792 | 6482 |
| 2008 | 2870 | 3771 | 6641 |
| 2009 | 3053 | 3930 | 6983 |
| 2010 | 2677 | 3877 | 6554 |
| 2011 | 2458 | 3982 | 6440 |
| total | 34260 | 42266 | 76526 |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI

A maior parte dos inventores (67,02%) possui apenas uma patente no período todo e 16,05% depositaram duas patentes entre 2000 e 2011 (tabela 08). Portanto, a grande maioria dos inventores possui baixa produtividade, tendo menos de 3 patentes no intervalo de 12 anos.

Tabela 08: Total de patentes por inventor entre 2000 e 2011

| Nº de Patentes | Inventores | (%) |
|----------------|------------|--------|
| 1 | 35184 | 67,02 |
| 2 | 8429 | 16,05 |
| 3 | 3494 | 6,66 |
| 4 | 1799 | 3,43 |
| 5 | 1080 | 2,06 |
| 6 | 667 | 1,27 |
| 7 ou mais | 1848 | 3,52 |
| total | 52501 | 100,00 |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI

4.2 Identificando a localização do inventor

A princípio, o acompanhamento do inventor ao longo do tempo para identificar sua mobilidade espacial, bem como suas relações de colaboração com inventores de outras regiões seria feito por meio das informações da RAIS-MIGRA/MTE (Relação Anual de Informações Sociais do Ministério do Trabalho e Emprego). Esta possui uma estrutura em painel que permite acompanhar a trajetória geográfica e econômica do trabalhador ao longo do tempo através de seus vínculos de trabalho no mercado formal. A partir do Certificado de

Pessoa Física (CPF) do inventor é possível identificá-lo na RAIS-MIGRA, desde que ele possua vínculo de emprego formal naquele ano, e captar suas informações geográficas e sócio econômicas.

Na identificação dos inventores na RAIS-MIGRA a partir do CPF observou-se que alguns CPF's apareciam nesta base mais de uma vez no mesmo ano. Nestas situações, adotou-se o critério de captar apenas as observações referentes ao vínculo empregatício ativo em dezembro. Como ainda havia duplicações, optou-se por manter apenas as observações sobre o vínculo com maior renda e, para as duplicações remanescentes, mantiveram-se as observações sobre o vínculo mais longo.

Mas este cruzamento de informações permitiu a identificação de apenas 29400 inventores (56% dos 52501 CPF's disponíveis na BADEPI), correspondendo a 1770 municípios diferentes. Além de não obter o município daqueles inventores não localizados na RAIS, para alguns dos inventores encontrados a informação existe apenas para um determinado ano ou para alguns anos do período de interesse (2000 a 2011). Assim, além de não ser possível observar a localização de grande parte dos inventores (44%) em nenhum ano do período, a limitação dos dados encontrados na RAIS leva a uma subestimação da taxa de migração, pois para vários inventores identificados na RAIS é impossível observar se houve mobilidade ao longo do período ou em que ano exato ela ocorreu. Estas restrições limitam a análise de mobilidade e de colaboração inter-regional, pois não há informação sobre a localização geográfica do inventor.

Visando ampliar as informações sobre a localização geográfica do inventor, optou-se por utilizar também o município do inventor disponível na BADEPI. Esta base identifica 1656 municípios diferentes. Para aumentar a qualidade dos dados, optou-se por utilizar o município da BADEPI quando esta informação não estava disponível na RAIS-MIGRA. Com este procedimento obteve-se a informação de 2163 municípios diferentes aos quais foram associadas as microrregiões correspondentes, identificando-se 482 microrregiões. Este processo permitiu identificar a localização geográfica de 50214 inventores (95,64% do total) em pelo menos um ano entre 2000 e 2011, o que representa um ganho substancial de informação em relação ao uso isolado da RAIS-MIGRA ou da BADEPI. Esses inventores respondem por 72.618 patentes.

Desse total de patentes, apenas 67.687 possuíam pelo menos um código IPC. Esse código é necessário para construir as classes tecnológicas das patentes, o que, por sua vez, é requerido para a construção da variável de proximidade tecnológica entre as regiões, como

será descrito no tópico 4.5. Por isso, procedeu-se a uma busca pelos códigos IPC das patentes que não o possuíam. Esses códigos foram obtidos nos próprios documentos de patentes, disponíveis no site do INPI. Dessa forma, o número de patentes com códigos IPC conhecidos aumentou para 68.660, formando a base de dados final que foi utilizada para a estimação dos modelos econométricos.

4.3 Análise descritiva

Para identificar o local em que as invenções patenteadas foram desenvolvidas, observou-se a região de residência do inventor no ano em que a patente foi protocolada junto ao INPI. As patentes com mais de um inventor foram rateadas proporcionalmente entre as regiões dos inventores, supondo que todos os coinventores tiveram o mesmo peso na criação da invenção. Devido a este critério de proporcionalidade na identificação da região de origem da patente e ao fato de que 4919 delas não tem nenhuma informação sobre o local de seu(s) inventor(es), identificou-se a região de origem de 72200 patentes (94,35%). Foram identificadas 482 microrregiões em que as invenções ocorreram, sendo que aquelas com maior número de patentes localizam-se nas regiões Sudeste e Sul do país.

A tabela 09 apresenta as 20 (vinte) microrregiões com maior quantidade de patentes entre 2000 e 2011. Elas representam 4,55% do total de microrregiões que tiveram patentes entre 2000 e 2011 (440 microrregiões) e respondem juntas por 67,8% do estoque nacional de patentes deste período. Destaca-se a importância de São Paulo na inovação nacional, com uma quantidade de patentes quase quatro vezes maiores que a produção do segundo colocado, Rio de Janeiro. Dentre as principais regiões patenteadoras, 10 (dez) encontram-se na região Sudeste do país, sendo 7 (sete) delas no estado de São Paulo, e 6 na região Sul, enquanto apenas 2 (duas) encontram-se na região Nordeste e 2 (duas) no Centro-Oeste do país.

Tabela 09: Regiões com maior número de patentes entre 2000 e 2011

| Microrregião | Patentes | (%) |
|--------------------------|----------|--------|
| São Paulo (SP) | 17157 | 23,76 |
| Rio de Janeiro (RJ) | 4412 | 6,11 |
| Curitiba (PR) | 3830 | 5,30 |
| Belo Horizonte (MG) | 3431 | 4,75 |
| Porto Alegre (RS) | 3162 | 4,38 |
| Campinas (SP) | 2988 | 4,14 |
| Caxias do Sul (RS) | 1782 | 2,47 |
| Joinville (SC) | 1573 | 2,18 |
| Brasília (DF) | 1243 | 1,72 |
| Osasco (SP) | 1224 | 1,70 |
| Florianópolis (SC) | 1096 | 1,52 |
| São José dos Campos (SP) | 976 | 1,35 |
| Blumenau (SC) | 870 | 1,20 |
| Goiânia (GO) | 787 | 1,09 |
| Salvador (BA) | 778 | 1,08 |
| Vitória (ES) | 772 | 1,07 |
| Ribeirão Preto (SP) | 729 | 1,01 |
| Fortaleza (CE) | 724 | 1,00 |
| Guarulhos (SP) | 718 | 0,99 |
| Sorocaba (SP) | 702 | 0,97 |
| Subtotal | 48952 | 67,80 |
| Outras | 23248 | 32,20 |
| Total | 72200 | 100,00 |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI e da RAIS-MIGRA/MTE

4.3.1 Analisando a migração dos inventores entre as regiões

Foram observados 15758 movimentos de migração de inventores entre microrregiões ao longo de todo o período, representando uma taxa de migração de 6,14% em relação ao total de observações disponíveis. Esta taxa é maior que a observada na migração entre Estados e menor que a ocorrida entre Municípios (tabela 10). Aproximadamente 25% dos inventores mudaram de município pelo menos uma vez entre 2000 e 2011 e quase 20% migraram de microrregião pelo menos uma vez neste período. Como já era esperado, o número de inventores que mudou de unidade da federação é bem menor, correspondendo a cerca de 8% do total de inventores.

Tabela 10: Taxa de migração e inventores que migraram pelo menos uma vez entre 2000 e 2011

| Migrações | Movimentos de Migração | | Inventores | |
|---------------------|------------------------|------|------------|-------|
| | Migrações | (%) | Inventores | (%) |
| Entre Municípios | 21318 | 8,31 | 13147 | 25,04 |
| Entre Microrregiões | 15758 | 6,14 | 10283 | 19,59 |
| Entre Estados | 6127 | 2,39 | 4381 | 8,34 |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI e da RAIS-MIGRA/MTE

Algumas observações são necessárias sobre estes dados: 1) foram considerados todos os anos em que a localização do inventor era conhecida e não apenas os anos em que o inventor protocolou patente no INPI. Assim, foram contabilizadas também as migrações ocorridas antes ou depois do pedido da patente. Este procedimento resulta em uma taxa de migração maior do que se fossem observados apenas os anos em que o inventor registrou a patente no INPI.¹⁹; 2) Por outro lado, para muitos inventores sua localização é conhecida apenas em 1 (um) ano ou alguns anos do período, o que leva a crer que as taxas de migração apresentadas estão subestimadas; 3) Por este motivo em vários casos não é possível saber o ano exato em que houve a migração. Por exemplo, se há informação sobre a residência do inventor apenas para 2002 e 2007 e estes locais são diferentes, sabe-se que ele migrou, mas não se conhece o ano exato desta migração.

A tabela 11 mostra que, dos inventores que migraram de município (13147), de microrregião (10283) e de estado (4381) entre 2000 e 2011, mais da metade migrou apenas uma vez no período (7495, 6265 e 2957, respectivamente).

Tabela 11: Número de migrações por inventor entre 2000 e 2011

| Nº de migrações | Entre Municípios | | Entre Microrregiões | | Entre Estados | |
|-----------------|------------------|--------|---------------------|--------|---------------|--------|
| | inventores | (%) | inventores | (%) | inventores | (%) |
| 0 | 39354 | 74,96 | 42218 | 80,41 | 48120 | 91,66 |
| 1 | 7495 | 14,28 | 6265 | 11,93 | 2957 | 5,63 |
| 2 | 3874 | 7,38 | 2953 | 5,62 | 1172 | 2,23 |
| 3 | 1229 | 2,34 | 779 | 1,48 | 197 | 0,38 |
| 4 | 407 | 0,77 | 212 | 0,40 | 42 | 0,08 |
| entre 5 e 10 | 142 | 0,27 | 74 | 0,14 | 13* | 0,02 |
| total | 52501 | 100,00 | 52501 | 100,00 | 52501 | 100,00 |

* 5 ou 6 vezes no caso de migrações entre Estados.

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI e da RAIS-MIGRA/MTE

¹⁹ O banco de dados ainda não está balanceado. As observações referem-se às obtidas da junção da BADEPI com a RAIS. O balanceamento do banco aumentaria o número total de observações e o número de localizações desconhecidas, o que reduziria as taxas de migração apresentadas.

Observa-se que o número de migrações do inventor entre Microrregiões está diretamente relacionado à sua produtividade em termos da quantidade de invenções patenteadas no período (tabela 12). A produtividade média dos inventores que não migrou nenhuma vez no período é 1,51 patentes por inventor entre 2000 e 2011, enquanto a produtividade média dos inventores que migraram 4 vezes ou mais é igual a 4,53 patentes por inventor no período. Note que o aumento da produtividade média é crescente com o número de migrações. Enquanto a produtividade dos inventores que migraram 1 vez ou mais é 16,56% maior que a dos inventores que não migraram nenhuma vez, a produtividade daqueles que migraram 4 vezes ou mais é 55,67% maior do que daqueles que migraram 3 vezes ou mais. Os inventores móveis são, em média, mais produtivos que os não móveis (HOISL, 2007; LENZI, 2009), mas os inventores mais produtivos também são mais propensos a se mudar, havendo uma relação de simultaneidade.

Tabela 12: Migrações entre Microrregiões e Produtividade média dos inventores entre 2000 e 2011

| Nº de migrações | Inventores | Patentes | Produtividade Média (PMe) | Crescimento da PMe (%) |
|-----------------|----------------|----------------|---------------------------|------------------------|
| 0 | 42218 (80,41%) | 63948 (83,56%) | 1,51 | - |
| 1 ou mais | 10283 (19,59%) | 18136 (23,70%) | 1,76 | 16,56 |
| 2 ou mais | 4018 (7,65%) | 8729 (11,41%) | 2,17 | 23,30 |
| 3 ou mais | 1065 (2,03%) | 3097 (4,05%) | 2,91 | 34,10 |
| 4 ou mais | 286 (0,54%) | 1296 (1,69%) | 4,53 | 55,67 |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI e da RAIS-MIGRA/MTE

Essa relação positiva entre patenteamento e migrações também se verifica em nível regional. As 20 (vinte) microrregiões (3,56% do total de regiões) com maior estoque de patentes entre 2000 e 2011 respondem por 67,8% das patentes nacionais do período (tabela 09) e são também as principais protagonistas dos movimentos de migração. Elas receberam em conjunto 57,81% de todas as entradas de inventores no período e registraram 58,6% de todas as saídas (tabela 13).

Tabela 13: Entrada e saída total de inventores das 20 maiores microrregiões patenteadoras entre 2000 e 2011

| Microrregião | Entrada Total | (%) | Saída Total | (%) |
|--------------------------|---------------|--------|-------------|--------|
| Sorocaba (SP) | 238 | 1,51 | 211 | 1,34 |
| Guarulhos (SP) | 265 | 1,68 | 265 | 1,68 |
| Fortaleza (CE) | 95 | 0,60 | 117 | 0,74 |
| Ribeirão Preto (SP) | 148 | 0,94 | 155 | 0,98 |
| Vitória (ES) | 127 | 0,81 | 118 | 0,75 |
| Salvador (BA) | 243 | 1,54 | 239 | 1,52 |
| Goiânia (GO) | 149 | 0,95 | 134 | 0,85 |
| Blumenau (SC) | 101 | 0,64 | 100 | 0,63 |
| São José dos Campos (SP) | 329 | 2,09 | 269 | 1,71 |
| Florianópolis (SC) | 293 | 1,86 | 269 | 1,71 |
| Osasco (SP) | 669 | 4,25 | 645 | 4,09 |
| Brasília (DF) | 330 | 2,09 | 293 | 1,86 |
| Joinville (SC) | 198 | 1,26 | 193 | 1,22 |
| Caxias do Sul (RS) | 150 | 0,95 | 119 | 0,76 |
| Campinas (SP) | 794 | 5,04 | 791 | 5,02 |
| Porto Alegre (RS) | 407 | 2,58 | 466 | 2,96 |
| Belo Horizonte (MG) | 545 | 3,46 | 696 | 4,42 |
| Curitiba (PR) | 575 | 3,65 | 572 | 3,63 |
| Rio de Janeiro (RJ) | 793 | 5,03 | 707 | 4,49 |
| São Paulo (SP) | 2661 | 16,89 | 3002 | 19,05 |
| Subtotal | 9110 | 57,81 | 9361 | 59,40 |
| Outras | 6648 | 42,19 | 6397 | 40,60 |
| Total | 15758 | 100,00 | 15758 | 100,00 |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI e da RAIS-MIGRA/MTE

Estes achados se mantêm quando a análise é feita em nível estadual. As 6 (seis) unidades da federação com maior estoque de patentes entre 2000 e 2011 representam apenas 22,2% dos estados, mas respondem por 88,28% do estoque nacional de patentes, 70,44% das entradas e 73,04% das saídas totais de inventores neste período (tabela 14).

Tabela 14: Estados com maior estoque de patentes entre 2000 e 2011:
entrada e saída total de inventores

| Estado | Patentes | (%) | Entrada | (%) | Saída | (%) |
|-------------------|----------|--------|---------|--------|-------|--------|
| Rio de Janeiro | 5149 | 7,13 | 724 | 11,82 | 652 | 10,64 |
| Santa Catarina | 5774 | 8,00 | 504 | 8,23 | 451 | 7,36 |
| Minas Gerais | 6095 | 8,44 | 639 | 10,43 | 664 | 10,84 |
| Paraná | 6602 | 9,14 | 627 | 10,23 | 634 | 10,35 |
| Rio Grande do Sul | 6964 | 9,64 | 383 | 6,25 | 451 | 7,36 |
| São Paulo | 33158 | 45,92 | 1439 | 23,49 | 1623 | 26,49 |
| Subtotal | 63741 | 88,28 | 4316 | 70,44 | 4475 | 73,04 |
| Outros | 8460 | 11,72 | 1811 | 29,56 | 1652 | 26,96 |
| Total | 72200 | 100,00 | 6127 | 100,00 | 6127 | 100,00 |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI e da RAIS-MIGRA/MTE

A análise descritiva da mobilidade inter-regional de inventores mostrou que, assim como a inovação e a produção industrial, a mobilidade do pessoal diretamente envolvido na criação de inovações é altamente concentrada no espaço, notadamente nas regiões Sudeste e Sul do país. Essa concentração sugere que o acesso ao conhecimento externo incorporado nesses indivíduos e seu efeito potencial sobre a inovação regional beneficia apenas um grupo restrito de regiões. Por outro lado, esses dados indicam que existe alta rotatividade dessa mão de obra.

O fato de que as regiões que apresentaram os maiores fluxos de saída de inventores são também as que mais atraíram inventores de outras localidades indica que poucas regiões concentram a maior parte do movimento inter-regional de inventores e sugere que existe alta rotatividade dessa mão de obra nessas regiões.

Essa rotatividade pode afetar negativamente o desempenho inovativo regional, pois a substituição de trabalhadores prejudica a dinâmica de inovação da firma por causa da importância da aprendizagem individual. O processo de aprendizagem contribui para a acumulação de conhecimento tácito e para o desenvolvimento de rotinas organizacionais, principais fatores da inovação da firma (GRINZA; QUATRARO, 2018).

4.3.2 Analisando as colaborações inter-regionais

Foram consideradas apenas as patentes para as quais a microrregião é conhecida para pelo menos um de seus inventores, o que corresponde a um total de 72.618 patentes. A maioria delas é fruto do esforço inventivo individual (82,01%), 10,92% (7927 patentes) são

produto da colaboração entre 2 (dois) inventores e 7,07% resultam da colaboração entre três ou mais inventores (tabela 15).

Tabela 15: Número de inventores por patente, de 2000 a 2011

| Número de inventores | Patentes | (%) |
|----------------------|----------|--------|
| 1 | 59.553 | 82,01 |
| 2 | 7.927 | 10,92 |
| 3 | 2.714 | 3,74 |
| 4 | 1.212 | 1,67 |
| 5 | 579 | 0,80 |
| entre 6 e 10 | 591 | 0,81 |
| entre 11 e 33 | 42 | 0,06 |
| total | 72.618 | 100,00 |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI e da RAIS-MIGRA/MTE

Considerando-se apenas as patentes para as quais é conhecida a microrregião de pelo menos um de seus inventores, existem 3724 (5,13%) patentes com colaboração entre inventores de regiões diferentes e 11235 patentes com coinvenção local (15,47%) (tabela 16). Observe que a existência de colaboração inter-regional não exclui a colaboração local para a mesma patente, ao passo que as patentes classificadas como resultante de coinvenção local excluem colaborações inter-regionais. A grande maioria das patentes de coinvenção inter-regional (3190 patentes) envolve colaboração entre apenas duas regiões²⁰ (tabela 17):

Tabela 16: Número de patentes resultantes de coinvenção local e inter-regional, de 2000 a 2011

| Coinvenção: | Número de patentes | (%) | Número de regiões |
|----------------|---------------------|---------------------|-------------------|
| Local | 11.235 | 15,47 | 227 |
| Inter-regional | 3.724 | 5,13 | 306 |
| Sem coinvenção | 59.553 | 82,01 | 422 |
| Total | 72.618 ¹ | 100,00 ² | 440 ³ |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI e da RAIS-MIGRA/MTE

^{1, 2, 3} A soma das linhas anteriores é maior que o total apresentado porque existem patentes com coinvenção local e coinvenção inter-regional, sendo contadas duas vezes.

²⁰ A coinvenção inter-regional pode estar subestimada, pois há inventores para os quais não há informação sobre a localização regional.

Tabela 17: Número de regiões envolvidas na geração de uma patente, 2000 a 2011.

| | Nº de patentes | Nº de microrregiões |
|---|----------------|---------------------|
| Patentes sem coinvenção inter-regional | 68894 | 1 |
| | 3190 | 2 |
| | 422 | 3 |
| | 87 | 4 |
| Patentes que envolvem 2 ou mais regiões | 19 | 5 |
| | 3 | 6 |
| | 1 | 7 |
| | 1 | 16 |
| | 1 | 18 |
| Subtotal | 3724 | - |
| Total | 72618 | - |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI e da RAIS-MIGRA/MTE

Existem 306 microrregiões que possuem pelo menos uma patente cujo desenvolvimento ocorreu em colaboração com inventores de outra microrregião, independente do número de inventores das respectivas regiões (tabela 18). A maioria dessas regiões (173) desenvolveu apenas de 1 a 5 patentes em colaboração com outras regiões no período de 2000 a 2011. Observa-se que são poucas as regiões com um número expressivo de patentes inter-regionais. A tabela 19 mostra as regiões com mais patentes resultantes de colaboração inter-regional. São Paulo, Campinas e Rio de Janeiro são os locais com maior número de patentes que foram desenvolvidas em conjunto com inventores de outras regiões, produzindo 1417, 708 e 488 patentes inter-regionais, respectivamente. A maioria das regiões que mais possuem patentes de coinvenção inter-regional situa-se no Sudeste do país (14 microrregiões), concentrando-se principalmente no estado de São Paulo (12 microrregiões). Exceto por uma região do Nordeste e uma do Centro-Oeste, as demais se encontram no Sul do país.

Tabela 18: Número de microrregiões com patentes de coinvenção inter-regional no período 2000 a 2011

| Nº de Patentes | Nº de Microrregiões |
|----------------|---------------------|
| 1 a 5 | 173 |
| 6 a 10 | 35 |
| 11 a 20 | 32 |
| 21 a 40 | 29 |
| 41 a 70 | 16 |
| 71 a 160 | 11 |
| 161 a 710 | 9 |
| 1417 | 1 |
| Total | 306 |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI e da RAIS-MIGRA/MTE

Uma patente de coinvenção inter-regional pode compreender mais de duas regiões, de forma que cada região envolvida na patente possua mais de uma relação inter-regional de colaboração. Por exemplo, se uma patente Z foi desenvolvida por inventores de três regiões distintas, A, B e C, então cada região possui duas relações inter-regionais de colaboração referentes à patente Z, independente do número de inventores das respectivas regiões. Os laços de colaboração inter-regional referem-se às ligações entre regiões para a produção de uma patente, e não entre inventores.

A tabela 19 também mostra que o número de relações inter-regionais de coinvenção é maior que o número de patentes de coinvenção. Isto ocorre porque uma mesma patente pode ser desenvolvida em colaboração com mais de uma região, gerando mais de uma relação de coinvenção. Além disto, a relação de coinvenção se refere a cada patente individual, de forma que se duas regiões desenvolveram duas patentes conjuntamente, então estas regiões possuem duas relações de coinvenção.

Tabela 19: Regiões com maior estoque de patentes de coinvenção inter-regional - 2000 a 2011

| Microrregião | Patentes de coinvenção ^a | Relações de coinvenção ^b |
|--------------------------|-------------------------------------|-------------------------------------|
| Guarulhos (SP) | 88 | 98 |
| Jundiaí (SP) | 91 | 107 |
| Joinville (SC) | 92 | 104 |
| Mogi das Cruzes (SP) | 97 | 127 |
| Sorocaba (SP) | 100 | 146 |
| Piracicaba (SP) | 105 | 153 |
| Araraquara (SP) | 117 | 204 |
| Salvador (BA) | 124 | 152 |
| Ribeirão Preto (SP) | 159 | 244 |
| Brasília (DF) | 160 | 230 |
| São José dos Campos (SP) | 167 | 196 |
| Florianópolis (SC) | 194 | 283 |
| Porto Alegre (RS) | 209 | 282 |
| Osasco (SP) | 213 | 269 |
| Curitiba (PR) | 233 | 304 |
| São Carlos (SP) | 243 | 346 |
| Belo Horizonte (MG) | 248 | 319 |
| Rio de Janeiro (RJ) | 488 | 621 |
| Campinas (SP) | 708 | 927 |
| São Paulo (SP) | 1417 | 1766 |

Fonte: Elaboração própria com base nos dados da BADEPI v1.0/INPI e da RAIS-MIGRA/MTE

^a Refere-se ao número de patentes em colaboração com inventores de outras regiões, independente do número de inventores em cada região e do número de regiões envolvidas na patente. ^b Refere-se à soma do número de regiões com as quais uma região manteve colaboração para o desenvolvimento de uma dada patente. Considera-se cada relação de forma individual, ou seja, se houve 10 patentes em colaboração entre as regiões A e B, esta relação foi contada 10 vezes.

Da mesma forma que a mobilidade inter-regional de inventores se concentra em poucas regiões, a colaboração entre inventores de locais distintos também abrange um pequeno número de regiões. Assim como no caso da mobilidade, as regiões que mais possuem coinvenções com outros locais encontram-se no Sudeste e Sul do país. De forma geral, as regiões que possuem maior mobilidade de inventores são também aquelas que possuem maior número de patentes de coinvenção inter-regional. Essa constatação sugere que são poucas as regiões que têm acesso potencial ao estoque de conhecimento externo por meio dos canais de difusão que estes fenômenos representam.

4.4 Análise descritiva espacial

Os resultados da seção anterior mostram que a inovação brasileira é concentrada em algumas regiões, notadamente nas regiões Sudeste e Sul do país, assim como a migração e a colaboração inter-regional de inventores. Isso sugere que existe uma dependência espacial da inovação regional, ou seja, a inovação de uma região é influenciada pela intensidade inovativa das demais regiões. O padrão espacial da inovação no Brasil pode ser analisado por meio de uma análise exploratória dos dados espaciais sobre as patentes per capita.

A estatística I de Moran (Moran, 1948) indica que existe autocorrelação espacial positiva do patenteamento per capita (tabela 20). Isto significa que regiões com maior desempenho inovador, mensurado pelo número de patentes per capita, tendem a ser vizinhas de regiões que também possuem maior desempenho da atividade inovativa, sugerindo a existência de *clusters* espaciais de inovação.

Tabela 20 : Coeficiente I de Moran
para patenteamento per capita no Brasil

| I de Moran ^a | p-valor ^b |
|-------------------------|----------------------|
| 0,4761 | 0,0010 |

Fonte: Elaboração própria utilizando o *software* GeoDa 1.6

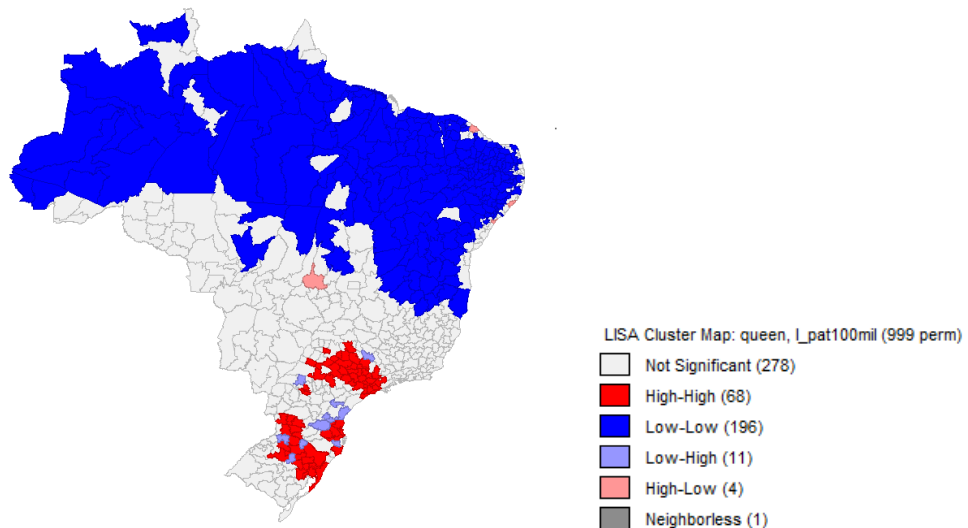
^a Baseado na matriz de contiguidade *Queen*.

^b Pseudo-significância empírica baseada em 999 permutações aleatórias.

O mapa de *clusters* (figura 02) mostra que existe um padrão espacial bem definido, com um grande *cluster* de regiões com baixo nível de patenteamento nas regiões Norte e Nordeste no país, e basicamente dois *clusters* de inovação, com alto nível de patenteamento, localizados no Sudeste e Sul do país. Os agrupamentos do tipo “Baixo-Baixo” compreendem 196 microrregiões caracterizadas por serem regiões com baixo nível de patenteamento per

capta, cujas regiões vizinhas também possuem baixo nível de patenteamento per capita. Os agrupamentos do tipo “Alto-Alto” abrangem 68 microrregiões com alta taxa de patenteamento per capita, cercadas por regiões que também possuem alto nível de patenteamento. A maioria dessas regiões encontra-se no estado de São Paulo, seguido do Rio Grande do Sul, Santa Catarina e Paraná. Destacam-se São Carlos, Guarulhos, Osasco, Campinas, Sorocaba, Jundiaí, Joinville, Blumenau, Itajaí, Caxias do Sul e Porto Alegre. Observa-se também que existem quatro regiões com alto nível de patentes per capita, cercadas por regiões com baixo desempenho inovador, três delas no Nordeste do país (Fortaleza, Maceió e Aracajú) e uma na região Centro-Oeste (Rio Vermelho).

Figura 02: Mapa de *clusters* para o estoque patentes entre 2000 e 2011, por 100 mil habitantes



Fonte: Elaboração própria utilizando a Base de Dados Estatísticas de Propriedade Intelectual – BADEPI/Inpi, v. 1.0; e o *software* GeoDa 1.6

As aglomerações observadas no mapa de *clusters* indicam que existe um padrão espacial na distribuição da atividade inovativa pelo território nacional, com predominância de uma associação espacial positiva. Este fato sugere que existem transbordamentos tecnológicos entre regiões geograficamente próximas, que podem contribuir para a formação destes *clusters*. Ao investigar os determinantes da inovação nas microrregiões brasileiras Gonçalves e Almeida (2009) encontraram evidências que confirmam a existência de transbordamentos inter-regionais do conhecimento, de forma que a atividade tecnológica das regiões vizinhas afeta positivamente a atividade tecnológica local. O mesmo resultado foi verificado por Gonçalves (2007) em uma análise exploratória do padrão espacial da atividade inovadora dos municípios brasileiros entre 1999 e 2001.

A dependência espacial implica que existe uma correlação entre a inovação local e a inovação das demais regiões, mas não permite inferir nada sobre causalidade nem sobre os mecanismos que provocam essa dependência. A proximidade espacial por si só não garante que a difusão do conhecimento entre as regiões vizinhas ocorra, é necessário que haja algum tipo de interação entre os agentes das regiões para que o conhecimento seja efetivamente difundido entre elas. A migração de inventores e a colaboração inter-regional para patenteamento são duas formas de interação que viabilizam o fluxo do conhecimento entre as regiões.

4.5 Descrição das variáveis dos modelos gravitacionais

A análise empírica será feita em duas partes. A primeira consiste na estimação dos modelos gravitacionais para a mobilidade e a coinvenção entre as microrregiões brasileiras. Na segunda etapa será estimada uma função de produção do conhecimento modificada para refletir os efeitos das interações inter-regionais de conhecimento sobre o patenteamento local. Os valores estimados na primeira etapa para a mobilidade e o copatenteamento entre pares de regiões representam a intensidade das interações de conhecimento entre estas regiões. Estes valores estimados serão empregados para construir as matrizes de peso que serão utilizadas na segunda parte da análise empírica para ponderar o estoque externo de conhecimento acessível a cada região. Desta forma, será obtida a elasticidade do patenteamento local ao estoque externo de conhecimento acessível por meio da mobilidade de inventores e a elasticidade do patenteamento local ao estoque externo de conhecimento acessível por meio das relações inter-regionais de coinvenção.

Na primeira etapa, os modelos de mobilidade e de coinvenção foram estimados em função do mesmo conjunto de variáveis explicativas: a) distância geográfica, representada pela distância euclidiana entre os centroides das regiões i e j ; b) proximidade tecnológica, mensurada pela correlação não centrada entre vetores regionais de classes tecnológicas de patentes (JAFFE, 1986); c) *dummy* de contiguidade, assumindo valor igual a 1 se as regiões i e j são contíguas e 0, caso contrário; e d) *dummy* de estado, assumindo valor igual a 1 se as regiões i e j pertencem à mesma unidade federativa e 0, caso contrário.

Espera-se que a distância geográfica afete negativamente tanto a mobilidade bilateral de inventores entre as regiões i e j , quanto a coinvenção entre i e j . No primeiro caso porque a informação sobre o destino torna-se mais imperfeita à medida que a distância aumenta

(MIGUÉLEZ; MORENO, 2013^a; 2015), o que eleva o custo com a coleta de informações e a incerteza da tomada de decisão. Além disto, o custo de transporte tende a ser maior e os encontros com familiares e antigos colegas tendem a ser mais esparsos com o aumento da distância. No segundo caso, porque a proximidade espacial facilita a escolha de parceiros potenciais (STORPER; VENABLES, 2004; MIGUÉLEZ e MORENO, 2013^a; 2015), facilita o gerenciamento conjunto de um projeto comum e o monitoramento do desempenho da outra parte (MIGUÉLEZ; MORENO, 2013a; 2015). Estes fatores estão intimamente associados ao custo de transporte e ao tempo de deslocamento. Quanto maior a distância geográfica, mais imperfeita é a informação sobre parceiros potenciais, dificultando a concretização de relações de coinvenção, e mais difíceis e custosos os eventuais contatos face a face para tratar do projeto comum.

A proximidade tecnológica é uma *proxy* para similaridade cognitiva. Uma maior proximidade tecnológica implica maior similaridade entre os conhecimentos das regiões *i* e *j*. Espera-se que a proximidade tecnológica afete positivamente tanto a mobilidade quanto a coinvenção entre as regiões *i* e *j*. Possuir um conhecimento relacionado à base de conhecimento da região de destino permite ao inventor móvel maiores possibilidades de atuação no novo local. Da mesma forma, possuir conhecimento relacionado e/ou complementar à base de conhecimento dos inventores da região *j*, permite ao inventor da região *i* maiores possibilidades de parceria para realização de coinvenções com os inventores da região *j*.

Para o cálculo da proximidade tecnológica entre as regiões *i* e *j* foram utilizados os 3 (três) primeiros dígitos dos códigos da *International Patent Classification (IPC)*²¹. Foram utilizados todos os *IPC's* disponíveis para cada patente, não apenas o primeiro. Assim, a patente com mais de uma classe tecnológica, representada pelos 3 (três) primeiros dígitos do *IPC*, foi contabilizada uma vez em cada classe. Isto porque cada código *IPC* reflete uma competência tecnológica distinta para a região. Procedimento similar foi empregado por

²¹ O *IPC* tem como objetivo principal funcionar como uma ferramenta de busca para a obtenção de documentos de patentes por escritórios de propriedade intelectual e outros usuários, para permitir o estabelecimento de novidade e a avaliação do passo inventivo ou não-obviedade de divulgações técnicas em pedidos de patentes. O *IPC* classifica todo o corpo de conhecimento técnico usando os níveis hierárquicos: seção, classe, subclasse, grupo e subgrupo, em ordem decrescente de hierarquia. São 8 (oito) seções, indicadas por letras maiúsculas de A a H (A: necessidades humanas; B: *performing operations* e transportes; C: química e metalurgia; D: têxteis e papel; E: construções fixas; F: engenharia mecânica, iluminação, aquecimento, armas, explosivos (*blasting*); G: física; e H: eletricidade). As classes de patentes são formadas pelos 3 (três) primeiros dígitos do *IPC*: uma letra maiúscula indicando a seção, seguida por dois dígitos numéricos (WIPO, 2018).

Gorin (2016) ao classificar as patentes em 40 subclasses e contar a patente associada a mais de uma classe uma vez em cada categoria Apesar de ser o mais importante, considerar apenas o primeiro código *IPC* implicaria em negligenciar as outras habilidades tecnológicas regionais.

Para cada região foram observadas as classes tecnológicas existentes em cada ano, havendo um total de 116 classes no início do período e 119 no final do período (tabela 21). Posteriormente, foram contabilizados quantos pedidos regionais de patentes se encaixam em cada classe²². Estas informações foram utilizadas para construir uma matriz de proximidade tecnológica. Dos 72.618 pedidos de patente com microrregião conhecida para pelo menos um inventor, apenas 68.660 possuem informação de pelo menos um código *IPC*. Dado que o conhecimento do *IPC* é necessário para calcular a variável de proximidade tecnológica, os modelos gravitacionais para mobilidade e para colaboração inter-regional foram estimados utilizando as informações sobre estas 68.660 patentes.

Seguindo a metodologia de Jaffe (1986), a proximidade tecnológica entre as regiões i e j , P_{ij} , é dada pela expressão:

$$P_{ij} = \frac{\sum_{c=1}^c T_{ic}T_{jc}}{[\sum_{c=1}^c T_{ic}^2 \sum_{c=1}^c T_{jc}^2]^{1/2}} \quad (27)$$

Onde:

P_{ij} = proximidade tecnológica entre as regiões i e j ;

T_{ic} : vetor tecnológico da região i , em que c representa o número de classes tecnológicas;

T_{jc} : vetor tecnológico da região j , em que c representa o número de classes tecnológicas.

O valor de P_{ij} varia entre 0 e 1, significando regiões cujos vetores tecnológicos são ortogonais e regiões cujos vetores tecnológicos são idênticos, respectivamente. Quanto maior a proximidade tecnológica entre as regiões i e j , mais próximo da unidade será o valor de P_{ij} . Por outro lado, quanto mais distintos forem os perfis tecnológicos das regiões i e j , mais próximo de 0 será o valor de P_{ij} .

²² Desta forma, as patentes com mais de um código *IPC* foram contabilizadas mais de uma vez. Esta múltipla contagem não prejudica o cálculo da proximidade tecnológica, pois para este cálculo o que interessa são as habilidades tecnológicas da região e não o número de patentes por si só.

Tabela 21: Número de microrregiões patenteadoras, quantidade de patentes e de classes tecnológicas: 2000 a 2011

| Ano | Microrregiões patenteadoras | Classes tecnológicas | Total de patentes*** |
|-------|-----------------------------|----------------------|----------------------|
| 2000 | 252 | 116 | 5363 |
| 2001 | 264 | 116 | 5700 |
| 2002 | 251 | 116 | 5610 |
| 2003 | 251 | 118 | 5975 |
| 2004 | 263 | 118 | 6145 |
| 2005 | 260 | 118 | 5844 |
| 2006 | 259 | 119 | 5749 |
| 2007 | 248 | 119 | 5916 |
| 2008 | 254 | 117 | 5974 |
| 2009 | 272 | 119 | 6099 |
| 2010 | 271 | 119 | 5278 |
| 2011 | 243 | 119 | 5007 |
| Total | 432* | 128** | 68660 |

Fonte: Elaboração própria com base nos dados da BADEPI 1.0/INPI e da RAIS/MTE

* Não é o somatório das linhas anteriores. Refere-se ao total de microrregiões patenteadoras no período de 2000 a 2011. ** Não é o somatório das linhas anteriores. Refere-se ao total de classes tecnológicas do estoque de patentes do período 2000 a 2011.

*** É o total de patentes com microrregião conhecida para pelo menos um inventor e pelo menos um código IPC disponível.

A contiguidade entre regiões tende a facilitar o deslocamento entre elas, reduzindo o efeito da distância sobre a mobilidade e sobre a coinvenção inter-regional. Por outro lado, as regiões que fazem parte da mesma unidade federativa tendem a possuir maior similaridade institucional e cultural, o que tende a facilitar a mobilidade e a colaboração tecnológica entre regiões do mesmo estado. Para controlar estes efeitos foi introduzida uma *dummy* para regiões com fronteira comum, e uma *dummy* para regiões que fazem parte do mesmo estado. Espera-se que ambas tenham sinal positivo tanto no modelo de mobilidade quanto no modelo de coinvenção.

Os modelos gravitacionais de mobilidade e de coinvenção foram estimados ano a ano, de 2001 a 2011. Isto permite acompanhar o papel das variáveis bilaterais de distância geográfica e distância tecnológica na determinação da mobilidade e do copatenteamento inter-regional. Em ambos os modelos, a proximidade tecnológica foi defasada em 1 (um) período para minimizar problemas de endogeneidade devido à possível causalidade reversa entre esta variável e a mobilidade dos inventores, bem como entre ela e a coinvenção inter-regional.

Os modelos foram estimados incluindo efeitos fixos da origem e do destino. A inclusão dos efeitos fixos regionais da origem e do destino mitiga o potencial viés decorrente da omissão de variáveis (ANDERSON; VAN WINCOOP, 2003; RAMOS, 2016). Em uma abordagem *cross-section* isto implica incluir *dummies* para a região de origem e para a região

de destino (BURGER, 2009). No caso da mobilidade, a inclusão dos efeitos fixos permite captar a influência de destinos alternativos na determinação do fluxo migratório entre duas regiões específicas, os chamados termos de resistência multilateral. De forma similar, a introdução dos efeitos fixos no modelo de coinvenção permite considerar a influência de regiões parceiras em potencial. De forma geral, os termos de resistência multilateral (efeitos fixos da origem e do destino) controlam os custos de transação de uma região em relação a todas outras regiões com as quais possui uma relação de migração ou de coinvenção.

No modelo de mobilidade, a variável dependente representa a contagem dos movimentos de migração dos inventores entre as microrregiões em cada ano. Este movimento foi contabilizado no ano em que se identificou a mudança da microrregião de residência do inventor, embora nem sempre este ano corresponda efetivamente ao ano em que ocorreu a migração. Devido à ausência de informações para todos os anos para todos os inventores, nem sempre é possível saber exatamente o ano em que a mudança ocorreu. Por exemplo, se existem informações sobre o inventor apenas para dois anos, digamos 2002 e 2006, e identifica-se que a microrregião de 2006 é diferente da microrregião de 2002, sabe-se que o inventor mudou entre 2002 e 2006, mas não é possível saber o ano exato da mudança. Neste caso, como a mudança foi identificada em 2006, o movimento de migração seria atribuído a 2006.

Para cada ano, foram construídas matrizes assimétricas de mobilidade de 558 linhas e 558 colunas, em que cada elemento representa o número de inventores que migrou da região i para a região j . Por definição, a diagonal principal destas matrizes é nula. Se um inventor se muda mais de uma vez ou se retorna para a região de origem, estes movimentos são contabilizados como migrações separadas e independentes.

A utilização das informações da RAIS permitiu identificar a mobilidade do inventor mesmo para aqueles indivíduos que tiveram apenas uma patente ao longo do período. Esta base de dados confere uma vantagem em relação a trabalhos similares feitos para a Europa e Estados Unidos, onde o local de residência do inventor só é identificado por meio dos documentos de patente (MIGUÉLEZ e MORENO, 2013a;). Nestes casos, só é possível identificar a mobilidade daqueles inventores que tiveram pelo menos duas patentes ao longo do período, uma antes e outra depois da mudança, o que subestima a mobilidade real. A utilização da informação da RAIS sobre a residência do inventor permite uma observação mais precisa acerca desta mobilidade, pois é possível identificar a mudança de residência mesmo daqueles inventores que tiveram apenas uma patente ao longo do período. Dado que a

maioria dos inventores protocolou apenas um pedido de patente entre 2000 e 2011, esta possibilidade é muito relevante, pois, caso contrário, o número de migrações poderia ser significativamente subestimado.

O modelo para coinvenção inter-regional também foi estimado ano a ano e possui as mesmas variáveis explicativas que o modelo de mobilidade. A variável dependente é a contagem do número de patentes da microrregião i realizadas em colaboração com inventores da região j , independente do número de inventores de cada local. Com base no local de residência do(s) inventor(es) no momento do protocolo da patente no INPI foram identificadas e contabilizadas as patentes resultantes da colaboração entre inventores de cada par de regiões. Para cada ano do período de 2000 a 2011 foram construídas matrizes simétricas de colaboração inter-regional com 558 linhas e 558 colunas. Cada elemento da matriz representa o número de patentes feitas em colaboração entre inventores da região i e da região j . Por definição, estas matrizes são simétricas, resultando em $[N(N - 1)/2]$ observações, e possuem diagonal principal nula.

A forma tradicional de estimar modelos gravitacionais aplicando MQO ao modelo log-linearizado possui dois problemas principais: a heterocedasticidade do termo de erro leva a estimativas inconsistentes (SANTOS SILVA; TENREYRO, 2006; 2010) e o grande número de zeros na variável dependente impossibilita a transformação logarítmica. Em relação ao segundo problema, tem sido frequente a estratégia de somar uma constante (geralmente, o número 1) à variável dependente antes de fazer a transformação logarítmica. Obviamente, este procedimento gera resultados duvidosos. Quanto ao primeiro problema apontado, Santos Silva e Tenreyro (2006, 2010) sugerem estimar a forma multiplicativa do modelo por pseudo-máxima verossimilhança de Poisson (PPML).

A mobilidade e a coinvenção inter-regional são variáveis com dados de contagem que assumem valores inteiros não negativos, concentrando-se em alguns pequenos valores discretos. Estes dados possuem distribuição intrinsecamente heterocedástica, com variância crescente com a média²³. Desta forma, o estimador PPML proposto pelos referidos autores é adequado aos dois modelos em questão.

²³ As tabelas com estatísticas descritivas da mobilidade e da coinvenção encontram-se no Apêndice (tabelas A1 a A4).

4.6 Descrição das variáveis utilizadas na Função de Produção do Conhecimento

A variável dependente da função de produção do conhecimento é o logaritmo natural da intensidade inovativa regional, mensurada pelo número de patentes produzidas na região para cada cem mil habitantes ($\ln Patentes$). Os dados sobre as patentes foram obtidos na BADEPI 1.0 do INPI para o período de 2000 a 2011. Cada patente possuía a informação do município dos inventores, de forma que as patentes foram somadas e agregadas por microrregião, unidade espacial de análise escolhida nessa tese.

O efeito da mobilidade inter-regional dos inventores e da coinvenção inter-regional sobre a inovação local foi captado por meio da defasagem da variável dependente utilizando-se os valores estimados nos modelos gravitacionais. No primeiro caso, essa defasagem consiste na média ponderada da intensidade inovativa das demais regiões utilizando-se como pesos os valores estimados para a mobilidade inter-regional ($W^m \ln Patentes$). No segundo caso, os pesos utilizados são os valores estimados no modelo de coinvenção inter-regional ($W^c \ln Patentes$).

Para estimar a dependência espacial foram utilizadas matrizes de pesos espaciais puramente baseados na geografia, como é tradicional na literatura da geografia da inovação ($W^e \ln Patentes$). Dado que a escolha da matriz de pesos espaciais é *ad hoc* e que diferentes matrizes podem gerar resultados substancialmente diferentes (MIGUÉLEZ; MORENO, 2015), o modelo foi estimado utilizando-se várias matrizes espaciais diferentes para verificar a robustez dos resultados.

Devido a indisponibilidade de informações sobre o investimento regional em PeD, foi utilizada como *proxy* da capacidade de PeD industrial a proporção de profissionais em ocupações técnico científicas, pois são as atividades mais propensas a contribuir para o desenvolvimento de inovações. Essa *proxy* tem sido largamente utilizada em trabalhos sobre inovação no Brasil (OLIVEIRA *et al.*, 2016; KANNEBLEY JÚNIOR, *et al.*, 2016). Denominada de PoTec (pessoal ocupado técnico-científico) essa variável foi desenvolvida por Araújo *et al.* (2009). Esses autores observaram uma correlação de 90% entre os gastos em PeD e a PoTec, indicando que essa variável é uma *proxy* adequada para representar a capacidade de PeD privado ($\ln PoTec$). Foram considerados os trabalhadores das seguintes ocupações, de acordo com a Classificação Brasileira de Ocupações (CBO) de 2002: profissionais da biotecnologia e biomédicos, engenheiros, pesquisadores, profissionais da matemática e de estatística, analistas de sistemas computacionais, físicos, químicos,

profissionais do espaço e da atmosfera e arquitetos. Os dados foram obtidos na Relação Anual de Informações Sociais do Ministério do Trabalho e Emprego (RAIS/MTE).

Além do PeD industrial, o PeD universitário (*lnPeDun*) também é um importante insumo de inovação, pois o conhecimento gerado nas universidades pode ser difundido para as empresas e utilizado para a produção de inovações. Como *proxy* para a capacidade regional de PeD universitário foi utilizado o número de professores doutores de centros de pós graduação *stricto sensu* das áreas de Engenharias, Ciências Exatas e da Terra, Ciências Agrárias, Ciências Biológicas e da Saúde, pois essas são as áreas mais propensas a contribuir para o desenvolvimento de inovações. Os dados foram obtidos na Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)²⁴.

Outro importante insumo do conhecimento é o capital humano (*lnCapHumano*). A dotação regional de capital humano afeta positivamente a inovação local. O capital humano regional foi representado pela parcela da população local com curso superior completo. Os dados foram obtidos na RAIS/MTE.

Além das variáveis anteriores, foram incluídas também quatro variáveis de controle. Para captar o efeito das economias de aglomeração foi utilizada a escala urbana (*lnEscUrbana*), representada pela participação da população regional em relação à população total do Brasil. As economias urbanas tendem a beneficiar a atividade inovativa de três formas: compartilhamento de insumos (de fatores de produção indivisíveis e/ou maior variedade de insumos diferenciados), maior eficiência do mercado de trabalho (em locais mais densos os trabalhadores são mais seletivos na escolha de emprego e encontram emprego mais rápido) e transbordamentos do conhecimento (CARLINO *et al*, 2007). Também foi inserido o quadrado da escala urbana (*lnEscUrbanaQd*) para verificar se existem efeitos de congestionamento que prejudiquem a inovação regional. Uma escala urbana muito elevada pode gerar externalidades negativas que aumentem os custos de aluguel e transporte, encarecendo os insumos necessários à atividade inovativa de forma geral. Os dados foram obtidos no Instituto Brasileiro de Geografia e Estatística (IBGE).

Os dois controles remanescentes se referem a características do setor produtivo. Para captar diferenças espaciais na estrutura econômica das regiões foi introduzida a participação do emprego na indústria de transformação no emprego total da região (*lnTransformação*) (MIGUÉLEZ; MORENO, 2015). Essa variável visa controlar o fato de que a indústria de transformação tende a ser mais inovadora que as demais atividades econômicas. Também foi

²⁴ A CAPES é uma fundação do Ministério da Educação (MEC) que tem como objetivo expandir e consolidar os cursos de pós-graduação de mestrado e doutorado em todo o país.

incluído o grau de especialização produtiva da indústria (*lnEspProdutiva*), representado pelo índice Herfindahl. Esse índice foi obtido pelo somatório dos quadrados das participações do emprego em cada setor produtivo no emprego industrial da região. O índice varia entre 0 e 1, sendo que quanto mais próximo da unidade, maior a especialização produtiva regional, e quanto mais próximo de 0, maior a diversidade industrial. Foram utilizados dados do emprego de 23 divisões (da divisão 15 a 37) da Classificação Nacional de Atividades Econômicas (CNAE) do Instituto Brasileiro de Geografia e Estatística (IBGE). As informações foram obtidas na RAIS/MTE.

Esse capítulo explicou o processo de obtenção e organização da base de dados e a construção das variáveis utilizadas na estimação dos modelos gravitacionais de mobilidade de inventores e coinvenção e na estimação da função de produção do conhecimento regional. O próximo capítulo apresenta e discute os resultados de ambos os modelos.

5 ANÁLISE DOS RESULTADOS

5.1 Resultados dos modelos gravitacionais

Primeiramente os modelos gravitacionais foram estimados sem efeito fixo da origem e do destino. As tabelas 22 e 23 apresentam os resultados para o modelo de mobilidade sem e com efeitos fixos da origem e do destino, respectivamente. Os resultados da tabela 22 devem ser vistos com muita cautela, pois a ausência dos efeitos fixos implica que não foram controladas as características específicas de todas as origens e destinos alternativos. Ou seja, os efeitos de atração e de repulsão de todos os destinos e de todas as origens foram negligenciados, respectivamente.

Sem controlar os efeitos fixos da origem e do destino (tabela 22), os resultados mostram que a mobilidade é afetada positiva e significativamente pela proximidade tecnológica e institucional. Incluindo-se os efeitos fixos (tabela 23) a proximidade tecnológica deixa de ser significativa em 4 anos do período analisado e a magnitude dos coeficientes diminuem drasticamente. Mas ainda assim, podemos concluir que a proximidade cognitiva é um fator relevante para o inventor ao decidir a região de destino. Esse resultado sugere que possuir conhecimento relacionado ao perfil tecnológico da região de destino aumenta as possibilidades de inserção no mercado de trabalho no destino. O mesmo resultado foi encontrado por Gorin (2016) para a Europa. O autor observou que a mobilidade de inventores ocorre principalmente entre cidades europeias com a mesma especialização tecnológica, em parte devido à disponibilidade de empregos especializados, em parte porque as relações sociais tendem a se desenvolver dentro da mesma comunidade epistêmica.

A distância geográfica afeta negativamente a mobilidade nas duas estimações, mas apresenta uma ligeira queda na magnitude dos coeficientes quando os efeitos fixos são incluídos nas estimações. Quanto maior a distância geográfica, menor é a mobilidade dos inventores entre as regiões i e j . Essa variável é negativa e altamente significativa em todos os anos do período analisado. O fato de que a mobilidade dos inventores tende a ocorrer entre regiões próximas geograficamente sugere que esta pode ser uma das causas da localização espacial dos fluxos de conhecimento. A intensidade das externalidades inter-regionais do conhecimento diminui com o aumento da distância geográfica precisamente porque a mobilidade dos agentes que possuem o conhecimento tácito incorporado também diminui com o aumento da distância física. Em sua análise da mobilidade de inventores na Europa,

Gorin (2016) observou que a distância geográfica possui papel limitado em inibir a mobilidade e que isso pode decorrer da disponibilidade de transporte rápido entre as cidades. No Brasil, as distâncias são maiores e o transporte não é tão rápido e barato como na Europa, o que torna a distância geográfica um obstáculo mais significativo à mobilidade dos inventores do que na Europa.

A *dummy* de contiguidade não era significativa em nenhum ano do período na tabela 22, passando a ser significativa e positiva em todos os anos quando os efeitos fixos são considerados, indicando que possuir uma fronteira comum facilita a mobilidade inter-regional. A *dummy* de UF também é positiva e altamente significativa nas duas tabelas, mas chama atenção o grande aumento na magnitude dos coeficientes. Isso significa que após controlar pelos efeitos fixos, a proximidade institucional se torna mais relevante na determinação da mobilidade dos inventores. Esse resultado confirma a expectativa de que fatores culturais e institucionais comuns são relevantes para a decisão de migração do inventor. De forma geral, esses resultados são coerentes com o encontrado na literatura (MIGUÉLEZ; MORENO, 2012 e 2013a).

Tabela 22: Resultados do modelo gravitacional: mobilidade de inventores entre as microrregiões brasileiras (2001 a 2010) - estimações por PPML sem EF

| Variáveis | Anos | | | | |
|------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | 2001 | 2002 | 2003 | 2004 | 2005 |
| <i>Ln(proximidade Tecnológica)</i> | 1,757*** (0,000) | 1,753*** (0,000) | 1,678*** (0,000) | 1,425*** (0,000) | 1,601*** (0,000) |
| <i>Ln(distância geográfica)</i> | -0,963*** (0,000) | -1,134*** (0,000) | -1,132*** (0,000) | -1,100*** (0,000) | -1,071*** (0,000) |
| <i>Dummy de contiguidade</i> | 0,055 (0,884) | 0,085 (0,786) | 0,035 (0,907) | 0,251 (0,382) | 0,195 (0,506) |
| <i>Dummy de UF</i> | 1,119*** (0,000) | 0,609*** (0,001) | 0,626*** (0,000) | 0,585*** (0,005) | 0,686*** (0,000) |
| <i>Efeito Fixo da Origem</i> | <i>não</i> | <i>não</i> | <i>não</i> | <i>não</i> | <i>não</i> |
| <i>Efeito Fixo do Destino</i> | <i>não</i> | <i>não</i> | <i>não</i> | <i>não</i> | <i>não</i> |
| <i>Número de observações</i> | 29876 | 30110 | 30284 | 30754 | 35574 |
| <i>Log-likelihood</i> | -5121,63 | -6437,49 | -8323,60 | -7835,07 | -8461,45 |
| <i>r2</i> | 0,13 | 0,19 | 0,19 | 0,18 | 0,15 |

| Variáveis | Anos | | | | |
|------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | 2006 | 2007 | 2008 | 2009 | 2010 |
| <i>Ln(proximidade Tecnológica)</i> | 1,609*** (0,000) | 1,917*** (0,000) | 2,058*** (0,000) | 1,904*** (0,000) | 1,889*** (0,000) |
| <i>Ln(distância geográfica)</i> | -1,050*** (0,000) | -1,107*** (0,000) | -0,999*** (0,000) | -1,151*** (0,000) | -1,046*** (0,000) |
| <i>Dummy de contiguidade</i> | 0,169 (0,580) | 0,076 (0,796) | 0,165 (0,556) | -0,002 (0,994) | 0,041 (0,886) |
| <i>Dummy de UF</i> | 0,696*** (0,000) | 0,683*** (0,000) | 0,697*** (0,000) | 0,388*** (0,012) | 0,504*** (0,002) |
| <i>Efeito Fixo da Origem</i> | <i>não</i> | <i>não</i> | <i>não</i> | <i>não</i> | <i>não</i> |
| <i>Efeito Fixo do Destino</i> | <i>não</i> | <i>não</i> | <i>não</i> | <i>não</i> | <i>não</i> |
| <i>Número de observações</i> | 34694 | 35022 | 34220 | 36188 | 38156 |
| <i>Log-likelihood</i> | -8877,35 | -8772,25 | -8664,83 | -9098,09 | -9528,60 |
| <i>r2</i> | 0,13 | 0,19 | 0,19 | 0,18 | 0,15 |

Fonte: Elaboração própria com o uso do software Stata12.0
p-valor entre parênteses ****p* < 0,01 ***p* < 0,05 **p* < 0,1

Tabela 23: Resultados do modelo gravitacional: mobilidade de inventores entre as microrregiões brasileiras (2001 a 2010) - estimações por PPML com EF

| Variáveis | Anos | | | | |
|------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | 2001 | 2002 | 2003 | 2004 | 2005 |
| <i>Ln(proximidade Tecnológica)</i> | 0,017 (0,929) | 0,212** (0,015) | 0,035 (0,700) | -0,144 (0,263) | 0,214*** (0,003) |
| <i>Ln(distância geográfica)</i> | -0,679*** (0,000) | -0,885*** (0,000) | -0,910*** (0,000) | -0,824*** (0,000) | -0,851*** (0,000) |
| <i>Dummy de contiguidade</i> | 0,470*** (0,009) | 0,231* (0,068) | 0,236** (0,022) | 0,423*** (0,000) | 0,358*** (0,000) |
| <i>Dummy de UF</i> | 2,216*** (0,000) | 1,571*** (0,000) | 1,432*** (0,000) | 1,513*** (0,000) | 1,361*** (0,000) |
| <i>Efeito Fixo da Origem</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> |
| <i>Efeito Fixo do Destino</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> |
| <i>Número de observações</i> | 14089 | 19352 | 21006 | 20552 | 24569 |
| <i>Log-likelihood</i> | -1817,58 | -2651,27 | -3299,73 | -3278,67 | -3659,58 |
| <i>r2</i> | 0,88 | 0,85 | 0,88 | 0,89 | 0,92 |

| Variáveis | Anos | | | | |
|------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | 2006 | 2007 | 2008 | 2009 | 2010 |
| <i>Ln(proximidade Tecnológica)</i> | 0,196*** (0,009) | 0,186*** (0,009) | 0,242*** (0,001) | 0,229*** (0,000) | 0,297*** (0,000) |
| <i>Ln(distância geográfica)</i> | -0,888*** (0,000) | -0,916*** (0,000) | -0,934*** (0,000) | -0,945*** (0,000) | -0,883*** (0,000) |
| <i>Dummy de contiguidade</i> | 0,334*** (0,000) | 0,294*** (0,001) | 0,212** (0,016) | 0,213*** (0,009) | 0,253*** (0,001) |
| <i>Dummy de UF</i> | 1,345*** (0,000) | 1,325*** (0,000) | 1,357*** (0,000) | 1,175*** (0,000) | 1,202*** (0,000) |
| <i>Efeito Fixo da Origem</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> |
| <i>Efeito Fixo do Destino</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> |
| <i>Número de observações</i> | 26304 | 25313 | 27202 | 27225 | 27292 |
| <i>Log-likelihood</i> | -3717,45 | -3773,26 | -4006,58 | -4187,74 | -4173,78 |
| <i>r2</i> | 0,93 | 0,92 | 0,91 | 0,91 | 0,92 |

Fonte: Elaboração própria com o uso do software Stata12.0
p-valor entre parênteses ****p* < 0,01 ***p* < 0,05 **p* < 0,1

As tabelas 24 e 25 apresentam os resultados para as coinvenções inter-regionais sem e com efeitos fixos da origem e do destino, respectivamente. No primeiro caso, a proximidade tecnológica é positiva e altamente significativa em todos os anos. Mas quando os efeitos fixos são introduzidos no modelo essa proximidade só é significativa em 4 ano do período estudado. Portanto, não há evidências robustas de que compartilhar uma base de conhecimento comum facilite as colaborações inter-regionais para inovar. Pode-se concluir que a proximidade tecnológica não é fator relevante para que os inventores façam

colaborações com pares de outras regiões. Esse resultado pode sugerir que os inventores buscam conhecimento externo nas colaborações com colegas de outras regiões para solucionar problemas em áreas tecnológicas distintas ou para desenvolver projetos que utilizem tecnologias distintas, mas complementares.

A distância geográfica afeta negativamente a coinvenção inter-regional nas abordagens sem e com efeitos fixos, mas os coeficientes apresentam leve queda quando esses efeitos são introduzidos. A *dummy* de contiguidade não é significativa quando os efeitos fixos são ignorados, mas se torna significativa quando esses são incluídos no modelo. Portanto, os resultados mostram que as variáveis espaciais são importantes na determinação das relações de copatenteamento inter-regional. A distância geográfica é negativa e altamente significativa em todos os anos do período analisado, confirmando os resultados observados na literatura (MORESCALCHI *et al.*, 2015; MIGUÉLEZ; MORENO, 2013a). Esse achado reflete o efeito dos custos informacionais e de deslocamento sobre as parcerias entre inventores de regiões distintas. Os encontros face a face são necessários para o desenvolvimento de um projeto em comum, ainda que eventuais. A *dummy* de contiguidade é positiva e significativa na maioria dos anos, indicando que compartilhar uma fronteira em comum facilita as relações de copatenteamento inter-regional. A *dummy* de UF é positiva e altamente significativa em todas as estimações, mas os coeficientes são muito maiores quando os efeitos fixos são controlados, mostrando que fatores culturais e institucionais em comum contribuem para estimular as colaborações entre as microrregiões para inovar.

Tabela 24: Resultados do modelo gravitacional: colaborações entre as microrregiões brasileiras (2001 a 2010) estimações por PPML sem EF

| Variáveis | Anos | | | | |
|------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | 2001 | 2002 | 2003 | 2004 | 2005 |
| <i>Ln(proximidade Tecnológica)</i> | 1,860*** (0,000) | 1,597*** (0,000) | 1,461*** (0,000) | 1,405*** (0,000) | 1,558*** (0,000) |
| <i>Ln(distância geográfica)</i> | -0,735*** (0,000) | -0,825*** (0,000) | -0,875*** (0,000) | -0,803*** (0,000) | -0,906*** (0,000) |
| <i>Dummy de contiguidade</i> | 0,198 (0,598) | 0,290 (0,397) | 0,091 (0,790) | 0,335 (0,352) | 0,164 (0,656) |
| <i>Dummy de UF</i> | 0,650*** (0,001) | 0,509*** (0,004) | 0,656*** (0,000) | 0,495*** (0,005) | 0,806*** (0,000) |
| <i>Efeito Fixo da Origem</i> | <i>não</i> | <i>não</i> | <i>não</i> | <i>não</i> | <i>não</i> |
| <i>Efeito Fixo do Destino</i> | <i>não</i> | <i>não</i> | <i>não</i> | <i>não</i> | <i>não</i> |
| <i>Número de observações</i> | 29876 | 30110 | 30284 | 30754 | 35574 |
| <i>Log-likelihood</i> | -4261,65 | -4913,89 | -5717,45 | -6348,34 | -5960,90 |
| <i>r2</i> | 0,08 | 0,08 | 0,09 | 0,09 | 0,09 |

| Variáveis | Anos | | | | |
|------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | 2006 | 2007 | 2008 | 2009 | 2010 |
| <i>Ln(proximidade Tecnológica)</i> | 1,693*** (0,000) | 2,020*** (0,000) | 2,166*** (0,000) | 2,007*** (0,000) | 1,907*** (0,000) |
| <i>Ln(distância geográfica)</i> | -0,921*** (0,000) | -0,814*** (0,000) | -0,730*** (0,000) | -1,044*** (0,000) | -0,828*** (0,000) |
| <i>Dummy de contiguidade</i> | -0,126 (0,733) | -0,003 (0,992) | 0,261 (0,408) | -0,166 (0,563) | 0,075 (0,813) |
| <i>Dummy de UF</i> | 0,889*** (0,000) | 0,971*** (0,000) | 0,896*** (0,000) | 0,458*** (0,013) | 0,533*** (0,001) |
| <i>Efeito Fixo da Origem</i> | <i>não</i> | <i>não</i> | <i>não</i> | <i>não</i> | <i>não</i> |
| <i>Efeito Fixo do Destino</i> | <i>não</i> | <i>não</i> | <i>não</i> | <i>não</i> | <i>não</i> |
| <i>Número de observações</i> | 34694 | 35022 | 34220 | 36188 | 38156 |
| <i>Log-likelihood</i> | -6505,14 | -6683,37 | -6973,04 | -7270,42 | -8310,56 |
| <i>r2</i> | 0,08 | 0,12 | 0,12 | 0,13 | 0,09 |

Fonte: Elaboração própria com o uso do software Stata12.0
p-valor entre parênteses ****p* < 0,01 ***p* < 0,05 **p* < 0,1

Tabela 25: Resultados do modelo gravitacional: colaborações entre as microrregiões brasileiras (2001 a 2010) - estimações por PPML com EF

| Variáveis | Anos | | | | |
|------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | 2001 | 2002 | 2003 | 2004 | 2005 |
| <i>Ln(proximidade Tecnológica)</i> | 0,119 (0,506) | 0,292** (0,024) | 0,163 (0,340) | -0,013 (0,920) | -0,005 (0,974) |
| <i>Ln(distância geográfica)</i> | -0,654*** (0,000) | -0,717*** (0,000) | -0,638*** (0,000) | -0,581*** (0,000) | -0,577*** (0,000) |
| <i>Dummy de contiguidade</i> | 0,345* (0,077) | 0,402* (0,070) | 0,366 (0,100) | 0,498** (0,014) | 0,626*** (0,000) |
| <i>Dummy de UF</i> | 1,129*** (0,000) | 0,963*** (0,000) | 1,233*** (0,000) | 1,089*** (0,000) | 1,196*** (0,000) |
| <i>Efeito Fixo da Origem</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> |
| <i>Efeito Fixo do Destino</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> |
| <i>Número de observações</i> | 9162 | 11124 | 11466 | 12006 | 13918 |
| <i>Log-likelihood</i> | -1615,31 | -2039,22 | -2240,39 | -2523,03 | -2432,47 |
| <i>r2</i> | 0,75 | 0,73 | 0,74 | 0,76 | 0,82 |

| Variáveis | Anos | | | | |
|------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | 2006 | 2007 | 2008 | 2009 | 2010 |
| <i>Ln(proximidade Tecnológica)</i> | 0,237 (0,122) | 0,093 (0,529) | 0,366*** (0,008) | 0,299*** (0,003) | 0,254** (0,027) |
| <i>Ln(distância geográfica)</i> | -0,657*** (0,000) | -0,599*** (0,000) | -0,624*** (0,000) | -0,867*** (0,000) | -0,790*** (0,000) |
| <i>Dummy de contiguidade</i> | 0,362** (0,012) | 0,362** (0,013) | 0,542*** (0,001) | 0,466*** (0,001) | 0,588*** (0,000) |
| <i>Dummy de UF</i> | 1,247*** (0,000) | 1,486*** (0,000) | 1,424*** (0,000) | 1,067*** (0,000) | 1,030*** (0,000) |
| <i>Efeito Fixo da Origem</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> |
| <i>Efeito Fixo do Destino</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>Sim</i> |
| <i>Número de observações</i> | 14472 | 14790 | 18978 | 19752 | 17624 |
| <i>Log-likelihood</i> | -2567,01 | -2829,83 | -3055,72 | -3166,89 | -3494,05 |
| <i>r2</i> | 0,87 | 0,86 | 0,82 | 0,83 | 0,83 |

Fonte: Elaboração própria com o uso do software Stata12.0
p-valor entre parênteses ****p* < 0,01 ***p* < 0,05 **p* < 0,1

5.2 Resultados da Função de Produção do Conhecimento Regional (FPCR)²⁵

Usando o arcabouço da função de produção do conhecimento regional, foi estimado um painel dinâmico para analisar a influência do conhecimento externo sobre a inovação regional. O papel do conhecimento externo foi avaliado por meio da ponderação das patentes das demais regiões usando como pesos os valores estimados da coinvenção inter-regional e da mobilidade de inventores entre regiões.

A defasagem espacial tem sido utilizada para mensurar o transbordamento do conhecimento mediado pelo espaço, incluindo sob sua égide todas as formas pelas quais o conhecimento se difunde, sem diferenciar quais são os canais específicos que viabilizam esta transferência. A inclusão da mobilidade de inventores e das colaborações inter-regionais de pesquisa visa explicitar o papel de dois elementos geralmente embutidos na defasagem espacial, verificando a contribuição relativa destes dois canais e das demais formas de difusão agrupadas pelo termo de defasagem espacial.

Inicialmente, foram estimadas três especificações da FPC regional (tabela 26). A primeira (1) inclui apenas a defasagem temporal da variável dependente e as variáveis de interesse: patentes das demais regiões ponderadas pelo fluxo de coinvenções inter-regionais ($W^c \ln Patentes$) e pela mobilidade inter-regional de inventores ($W^m \ln Patentes$)²⁶. A segunda especificação (2) acrescenta os insumos do conhecimento: *proxies* para o *PeD* privado ($\ln PoTec$), para o *PeD* universitário ($\ln PeDun$) e para o capital humano ($\ln CapHumano$). Além das variáveis anteriores, a especificação (3) inclui as variáveis de controle: escala urbana ($\ln EscUrbana$), escala urbana ao quadrado ($\ln EscUrbanaQd$), participação da indústria de transformação no emprego local ($\ln Transformação$) e especialização produtiva regional ($\ln EspProdutiva$). Foram incluídas *dummies* de ano para remover possíveis choques temporais que afetem todas as regiões, adequando-se à hipótese de ausência de autocorrelação nos erros entre as regiões.

²⁵ Todos os modelos desta seção foram estimados via *system gmm* utilizando os *lags* 2 e 3 das variáveis endógenas como instrumentos. A restrição do número de *lags* é importante para evitar o efeito negativo da proliferação de instrumentos sobre a confiabilidade do teste Hansen. Todas as especificações foram reestimadas utilizando-se os *lags* de 2 a 5 para verificar a robustez dos resultados. De forma geral, os resultados não se alteraram. As tabelas encontram-se no Apêndice (Tabelas A7, A8 e A9).

²⁶ Essas matrizes foram normalizadas na linha, como é usual na literatura. A normalização na linha consiste na divisão de cada elemento da linha, pela soma de todos os elementos desta linha de forma que $\sum_{j=1}^n w_{ij} = 1, i = 1, \dots, n$. Assim, cada elemento w_{ij} da matriz normalizada na linha pode ser interpretado como uma fração da influência espacial total da unidade j sobre a unidade i (ELHORST, 2014; ALMEIDA, 2012).

Tabela 26: Função de Produção do Conhecimento: estimações por *System GMM* – 2001 a 2010¹

| Var. dependente: \ln (Patentes por 100 mil habitantes) | (1) | (2) | (3) |
|--|---------------------|---------------------|---------------------|
| $\ln Patentes_{t-1}$ | 0,134*** (0,002) | 0,149*** (0,000) | 0,147*** (0,001) |
| $\ln Patentes_{t-2}$ | 0,051 (0,172) | 0,064* (0,095) | 0,061 (0,105) |
| $W^c \ln Patentes$ | 0,348*** (0,000) | 0,334*** (0,000) | 0,342*** (0,000) |
| $W^m \ln Patentes$ | -0,033 (0,741) | -0,068 (0,560) | -0,046 (0,657) |
| $\ln PoTec$ | | 0,012 (0,344) | 0,017 (0,145) |
| $\ln PeDun$ | | 0,029* (0,084) | 0,031* (0,069) |
| $\ln CapHumano$ | | 0,250*** (0,001) | 0,192*** (0,001) |
| $\ln EscUrbana$ | | | 0,466* (0,081) |
| $\ln EscUrbanaQd$ | | | 0,026* (0,075) |
| $\ln Transformação$ | | | 0,018 (0,712) |
| $\ln EspProdutiva$ | | | -0,083 (0,412) |
| Constante | 0,130** (0,028) | 0,577*** (0,003) | 2,527** (0,047) |
| Dummy de ano | Sim | Sim | Sim |

(Continua)

| Var. dependente: \ln (Patentes por 100 mil habitantes) | (1) | (2) | (3) |
|---|------------------|------------------|------------------|
| Microrregiões/Observações | 558/4464 | 558/4464 | 558/4464 |
| Lags utilizados como instrumentos ² | 2 e 3 | 2 e 3 | 2 e 3 |
| Nº de instrumentos | 17 | 24 | 32 |
| Teste F | (0,000) | (0,000) | (0,000) |
| AR (1) | (0,000) | (0,000) | (0,000) |
| AR (2) | (0,532) | (0,286) | (0,369) |
| Teste Hansen | (0,299) | (0,331) | (0,641) |
| <i>Difference-in-Hansen: Ho: exogeneidade de subgrupos de instrumentos</i> | | | |
| Instrumentos para a equação em diferenças ³ | (0,731) | (0,468) | (0,612) |
| Instrumentos para a equação em nível ⁴ | (0,142) | (0,246) | (0,521) |
| $\Delta \ln \text{Patentes}_{t-1}$ para a equação em nível | (0,331) | (0,465) | (0,521) |
| $\Delta W^c \ln \text{Patentes}_{t-1}$ e $\Delta W^m \ln \text{Patentes}_{t-1}$ para a equação em nível | | (0,348) | (0,447) |
| $\Delta \ln \text{Controles}_{t-1}$ para a equação em nível | | (0,167) | (0,792) |
| Estimativa de $\ln \text{Patentes}_{t-1}$ por MQO | 0,415 (0,000) | 0,413 (0,000) | 0,404 (0,000) |
| | | | (Continua) |

| Var. dependente: \ln (Patentes por 100 mil habitantes) | (1) | (2) | (3) |
|--|-------------------|-------------------|-------------------|
| Estimativa de $\ln Patentes_{t-2}$ por MQO | 0,337 (0,000) | 0,333 (0,000) | 0,327 (0,000) |
| Estimativa de $\ln Patentes_{t-1}$ por EF | -0,024 (0,367) | -0,025 (0,344) | -0,025 (0,348) |
| Estimativa de $\ln Patentes_{t-2}$ por EF | -0,077 (0,001) | -0,079 (0,001) | -0,079 (0,001) |

Fonte: Elaboração própria com o uso do *software* Stata 12.0.

p -valor entre parênteses *** $p < 0,01$ ** $p < 0,05$ * $p < 0,1$

¹ Foram consideradas como variáveis endógenas: $\ln Patentes_{t-1}$, $W^c \ln Patentes$, $W^m \ln Patentes$, $\ln PoTec$, $\ln PeDun$, $\ln Transformação$ e $\ln EspProdutiva$. Foram consideradas como variáveis exógenas: $\ln CapHumano$, $\ln EscUrbana$ e $\ln EscUrbanaQd$.

² Foram utilizados o segundo e o terceiro *lags* das variáveis endógenas como instrumentos para a equação em diferenças e o primeiro *lag* da diferença para a equação em nível.

³ Instrumentos para a equação em diferenças: $\ln Patentes_{t-2}$, $\ln Patentes_{t-3}$, $W^c \ln Patentes_{t-2}$, $W^c \ln Patentes_{t-3}$, $W^m \ln Patentes_{t-2}$, $W^m \ln Patentes_{t-3}$, $\ln PoTec_{t-2}$, $\ln PoTec_{t-3}$, $\ln PeDun_{t-2}$, $\ln PeDun_{t-3}$, $\ln Transformação_{t-2}$, $\ln Transformação_{t-3}$, $\ln EspProdutiva_{t-2}$, $\ln EspProdutiva_{t-3}$

⁴ Instrumentos para a equação em nível: $\Delta \ln Patentes_{t-1}$, $\Delta W^c \ln Patentes_{t-1}$, $\Delta W^m \ln Patentes_{t-1}$, $\Delta \ln PoTec_{t-1}$, $\Delta \ln PeDun_{t-1}$, $\Delta \ln Transformação_{t-1}$, e $\Delta \ln EspProdutiva_{t-1}$

A validade dos resultados obtidos via *system gmm* depende do atendimento às várias hipóteses do estimador, o que pode ser verificado a partir de testes estatísticos específicos. O estimador requer que exista correlação serial de primeira ordem nos resíduos e que não exista correlação serial de segunda ordem (ARELLANO; BOND, 1991). A autocorrelação de primeira ordem decorre da transformação em primeira diferença. A ausência de autocorrelação de segunda ordem nos resíduos em diferenças implica ausência de autocorrelação de primeira ordem dos resíduos em nível, condição necessária para a validade dos instrumentos e para a consistência do estimador. Como se pode ver na parte de baixo da tabela 28 os testes Arellano Bond indicam que as 3 (três) estimações são consistentes, pois rejeitam a hipótese de ausência de autocorrelação serial de primeira ordem nos resíduos (AR(1)), e não rejeita esta hipótese para a autocorrelação de segunda ordem (AR(2)).

Dado que após a primeira diferença os erros deixam de ser esféricos, as estimações foram feitas com o *system gmm* de dois passos e com a correção dos erros padrão de Windmeijer (2005)²⁷. As estimativas de dois passos são robustas à heterocedasticidade e permitem a obtenção do teste Hansen (*J*), mais adequado para verificar a validade dos instrumentos no contexto da heterocedasticidade. Essa estatística testa a hipótese nula de correta especificação do modelo e de ortogonalidade dos instrumentos (BAUM, 2006).

Os testes Hansen (*J*) para as três especificações não rejeitam a hipótese conjunta de exogeneidade dos instrumentos, indicando que os instrumentos são válidos e que a endogeneidade foi tratada adequadamente. Devido à grande sensibilidade à proliferação do número de instrumentos, o que enfraquece o poder do teste, o número de *lags* dos instrumentos foi restringido ao segundo e terceiro *lags* das variáveis endógenas em todas as estimações. Apesar de não haver uma regra clara sobre o número ideal de instrumentos para evitar o enfraquecimento do teste, recomenda-se que este número seja menor que o número de grupos (KUKENOVA; MONTEIRO, 2008; ROODMAN, 2006 e 2009), preferencialmente bem abaixo desse valor. Em todas as estimações o número de instrumentos é parcimonioso e é bem menor que o número microrregiões (grupos). Os *p-valores* dos testes Hansen reportados na tabela não permitem rejeitar a hipótese nula de exogeneidade dos instrumentos e de correta especificação dos modelos. Além

²⁷ Windmeijer (2005) desenvolveu uma correção para pequenas amostras que melhora as estimativas dos coeficientes do *gmm* de dois passos em relação ao *gmm* de um passo, reduzindo o viés. Os erros padrão se tornam mais precisos, de forma que a estimativa de dois passos com os erros corrigidos é superior ao *gmm* de um passo robusto.

disso, a magnitude do *p*-valor em todas as estimações (0,299; 0,331 e 0,641) confere maior segurança sobre o diagnóstico de validade dos instrumentos, pois atende à recomendação de Roodman (2009, p.142) de utilizar um *p*-valor acima de 0,25 para a não rejeição da hipótese nula.

O teste Hansen (teste *J*) avalia a hipótese conjunta de que todos os instrumentos utilizados no modelo são válidos. Mas é recomendável avaliar também a validade de subgrupos de instrumentos, o que pode ser feito por meio do teste *Difference in Hansen* (teste *C*). O teste para o subgrupo de instrumentos da equação em diferenças mostra que esses instrumentos são válidos e que os modelos estão corretamente especificados. Isso indica que o modelo é dinamicamente completo, o que é necessário para a validade dos instrumentos.

Foram testados também subgrupos de instrumentos para a equação em nível considerados mais importantes para a análise. A utilização do *system gmm* requer que as distâncias iniciais das regiões em relação ao estado estacionário não sejam correlacionadas com os efeitos fixos (BLUNDELL-BOND, 1998; ROODMAN, 2009). Dado que as variáveis explicativas endógenas, *X*, podem conter informações do efeito fixo, a hipótese de que os instrumentos adicionados pelo *system gmm*, ΔX_{t-1} , não são correlacionados com o termo de erro não é trivial. Por isso, é recomendável realizar um teste *Difference in Hansen* para todos os instrumentos para a equação em nível e não apenas para os instrumentos baseados na variável dependente (ROODMAN, 2009).

O teste para o subgrupo de todos os instrumentos usados na equação em nível mostra que as condições de momento adicionadas pelo *system gmm* são ortogonais ao termo de erro nas três especificações apresentadas. Isso indica que o requerimento sobre as condições iniciais de Blundell-Bond (1998), necessário para a validade dos instrumentos, foi atendido. Além disso, o coeficiente da variável dependente defasada é menor que a unidade, o que indica um processo de convergência da variável dependente.

Foram testados também outros três subgrupos de instrumentos para a equação em nível: instrumentos baseados na variável dependente ($\Delta \ln Patentes_{t-1}$), instrumentos baseados nas variáveis endógenas de interesse ($\Delta W^c \ln Patentes_{t-1}$ e $\Delta W^m \ln Patentes_{t-1}$) e instrumentos baseados nas variáveis endógenas de controle ($\Delta \ln Controles_{t-1}$). O teste *Difference in Hansen* indica que os instrumentos de todos subgrupos são válidos nas três estimações.

A defasagem temporal da variável dependente do lado direito da equação capta os efeitos inerciais da inovação, uma vez que a inovação tende a ser *path dependent*. Observa-se que as estimativas dos coeficientes da defasagem temporal geradas pelo *system gmm* encontram-se entre os valores obtidos pelo estimador de mínimos quadrados ordinários (MQO) e o estimador de efeito fixo (EF) para as três estimações. Isto sugere que o *system gmm* estimou adequadamente estes coeficientes, eliminando o viés do painel dinâmico (BOND, 2002; ROODMAN, 2006).

O coeficiente da defasagem temporal para $t-1$ é altamente significativo e mantém-se estável nas três estimações, confirmando que o progresso tecnológico é *path dependent*. Esse resultado mostra que as inovações anteriores afetam as inovações atuais, e também que os determinantes das inovações passadas influenciam as inovações atuais. Isso porque nem sempre os agentes respondem prontamente aos incentivos e porque existe um processo de aprendizado que afeta o progresso tecnológico futuro. Essa persistência temporal implica que a hegemonia regional em inovação tende a permanecer ao longo do tempo, enquanto regiões sem tradição em inovar tendem a permanecer pouco inovadoras. Ao contrário do esperado, a inovação de $t-2$ foi significativa apenas na especificação 2.

A inovação das demais regiões ponderada pelas coinvenções inter-regionais é positiva, altamente significativa e seu coeficiente é estável nas três especificações. Esse resultado mostra que o acesso ao conhecimento externo à região por meio de relações de coinvenção com inventores de outros locais contribui de forma importante para a produção local de inovações. Isso significa que essas relações de coinvenção possibilitam a difusão do conhecimento entre regiões, permitindo uma ampliação do conhecimento local maior do que a obtida apenas com os insumos locais de conhecimento. Ao beneficiar o desempenho inovador local, essas colaborações inter-regionais contribuem também para o desenvolvimento econômico regional e nacional. Portanto, o incentivo às parcerias inter-regionais para inovar pode estimular o progresso tecnológico local e nacional, contribuindo positivamente para a produtividade, a competitividade e o desenvolvimento econômico.

Ao contrário do esperado, a mobilidade inter-regional de inventores não se revelou um canal significativo de difusão do conhecimento e de estímulo à inovação regional. As patentes das demais regiões ponderadas pelos fluxos de mobilidade de inventores não possuem efeito estatisticamente diferente de zero. Apesar de o inventor ser o repositório e o veículo de transmissão do conhecimento tácito, a mobilidade inter-

regional de inventores não se revelou um canal significativo de difusão do conhecimento entre as regiões brasileiras no período em estudo.

Esse resultado pode ser causado por várias razões. Uma delas pode ser a baixa capacidade de absorção regional. Regiões com elevada capacidade de absorção são mais capazes de transformar o conhecimento externo, trazido pela entrada de inventores, em inovações regionais. Essa capacidade de absorção está diretamente ligada a um estoque de conhecimento relativamente grande (MIGUÉLEZ; MORENO, 2015).

A capacidade de internalizar o conhecimento externo é determinada principalmente pelo esforço tecnológico prévio, representado pelo gasto passado com *PeD*, e pelo tempo de emprego dos trabalhadores na firma (COHEN; LEVINTHAL, 1990; De NEGRI, 2006). Quanto maior o esforço de *PeD* maior o conhecimento prévio necessário para aproveitar o conhecimento gerado externamente.

Como pode ser visto na tabela 26, a *proxy* para *PeD* privado (*lnPoTec*) não é significativamente diferente de zero. Isso sugere que o investimento do setor privado em *PeD* é baixo, não contribuindo significativamente para o desempenho da inovação regional, resultado também encontrado por Gonçalves e Fajardo (2011). Uma vez que o aproveitamento de transbordamentos tecnológicos por parte das firmas é, em grande medida, associado à sua capacidade de absorção (De NEGRI, 2006), e essa é determinada pela realização de *PeD*, esse achado sugere que a capacidade de absorção das firmas e, conseqüentemente, da região, pode ser limitada. Para que ocorra a transferência de conhecimento não basta apenas que existam canais que possibilitem essa difusão, é preciso que haja também capacidade de absorção e replicação da tecnologia, sendo que firmas e trabalhadores se completam nesse processo (MENDES *et al.*, 2012).

Além disso, como observado na análise descritiva, as regiões que apresentaram as maiores entradas de inventores no período foram também as que tiveram as maiores saídas, sugerindo que pode haver uma alta rotatividade dos inventores, o que seria prejudicial ao processo de inovação. O tempo de emprego do inventor/trabalhador na empresa possibilita um processo de aprendizagem de conhecimentos específicos à firma que permite o aproveitamento do conhecimento externo. A substituição de trabalhadores é prejudicial para a performance inovativa das firmas, pois provoca a perda de importantes repositórios de conhecimento tácito, obtido com a aprendizagem e a acumulação de conhecimento (GRINZA; QUATRARO, 2018).

Quando os trabalhadores saem da firma, eles levam consigo conhecimento específico sobre o potencial da firma para a combinação de recursos para a criação de inovações. A entrada de novos trabalhadores, com seu próprio conhecimento tácito que poderia ser valioso para a firma, parece não compensar esse efeito negativo (GRINZA; QUATRARO, 2018).

A capacidade de absorção se torna mais eficiente quando as regiões receptoras são capazes de reter os inventores e fornecer oportunidades para interações locais e difusão de ideias (MIGUÉLEZ; MORENO, 2015). O tempo de permanência do trabalhador na firma corresponde a um indicador de aprendizado tecnológico (DE NEGRI, SALERMO; DE CASTRO, 2005).

Miguélez e Moreno (2015) observaram que o desempenho inovativo das regiões mais desenvolvidas da Europa é significativamente beneficiado pela entrada de inventores, enquanto esse efeito nas regiões mais pobres é inexistente ou desprezível. O contrário foi observado em relação ao conhecimento externo obtido via coinvenção inter-regional. As regiões mais pobres são as que mais se beneficiam dessa forma de acessar o conhecimento externo.

Os resultados do modelo gravitacional para mobilidade mostraram que a proximidade tecnológica entre as regiões contribui para aumentar a migração inter-regional de inventores. O fato de a mobilidade ocorrer principalmente entre regiões tecnologicamente relacionadas pode fazer com que o conhecimento transferido seja redundante e tenha impacto econômico limitado (GORIN, 2016). Inventores vindos de firmas com alto desempenho inovativo e com diferentes habilidades tecnológicas trazem habilidades e conhecimento complementares que beneficiam a performance inovativa futura da firma de destino (RAHKO, 2016). O impacto da mobilidade sobre a inovação regional pressupõe algum grau de complementaridade de conhecimento entre o inventor móvel e seus novos colegas de trabalho, de forma que as características individuais do inventor são menos relevantes para a inovação do que a composição da equipe de inventores (HOISL, 2007).

A entrada de inventores na região por si só não afeta a performance inovativa das firmas da região. O impacto da entrada desses indivíduos é altamente dependente das formas pelas quais as firmas buscam seus insumos de conhecimento. O efeito positivo só ocorrerá se existirem interações da firma com o ambiente externo, complementando as fontes internas de conhecimento com as externas, e se houver capacidade de absorção. Dessa forma as firmas podem se beneficiar das externalidades

espaciais positivas geradas pela entrada de indivíduos altamente inovadores em sua vizinhança funcional (CRESCENZI; GAGLIARDI, 2015).

Em relação aos insumos locais do conhecimento, ao contrário do *PeD* industrial (*lnPoTec*), o *PeD* universitário (*lnPeDun*) é positivo e significativo (a 10%) indicando a importância da pesquisa universitária para a produção regional de inovações, resultado verificado também por Gonçalves e Almeida (2009), Gonçalves e Fajardo (2011) e Araujo e Garcia (2016). Um aumento de 1% no *PeD* universitário local gera um aumento de 2,9% na produção regional de patentes (3,1% na especificação 3). A variável capital humano (*lnCapHumano*) é positiva e altamente significativa, mostrando a importância do capital humano como insumo da geração de conhecimento. A elasticidade do capital humano cai de 0,25 na especificação 2 para 0,19 quando são acrescentadas as variáveis de controle (especificação 3). Isso significa que um aumento de 1% no capital humano local induz um crescimento de 0,19% na inovação regional. Como o capital humano é representado pela *proxy* parcela da população com curso superior completo, esse resultado mostra a importância da educação superior para a performance inovadora regional e, conseqüentemente, para o desenvolvimento econômico local.

A escala urbana e a escala urbana ao quadrado são positivas e significativas (a 10%) indicando as vantagens da aglomeração para a inovação. As economias urbanas beneficiam a atividade inovativa de três formas: compartilhamento de insumos (de fatores de produção indivisíveis e/ou maior variedade de insumos diferenciados), maior eficiência do mercado de trabalho (em locais mais densos os trabalhadores são mais seletivos na escolha de emprego e encontram emprego mais rápido) e transbordamentos do conhecimento (CARLINO *et al.*, 2007). Os resultados são coerentes com os achados da literatura, que mostram que os atributos urbanos, como escala industrial e serviços modernos, são requisitos importantes para que uma região possa inovar (SIMÕES *et al.*, 2014), e que a inovação é positivamente afetada pela infraestrutura tecnológica regional (*PeD* industrial, *PeD* universitário, aglomerações de indústrias relacionadas e serviços especializados para os negócios) (FELDMAN; FLORIDA, 1994).

O valor positivo e significativo do quadrado da escala urbana indica a ausência de um efeito de congestionamento sobre a inovação regional. Uma escala urbana muito elevada poderia gerar externalidades negativas que poderiam se traduzir em maiores custos de aluguel e transporte, por exemplo, encarecendo os insumos necessários à atividade inovativa de forma geral. O resultado mostra que essa externalidade negativa

(se existir) é mais do que compensada pelos benefícios da aglomeração. A participação da indústria de transformação no emprego local e a especialização produtiva não são significativas, não afetando a inovação regional.

Para verificar a robustez dos resultados a especificação 3 foi estimada com a inclusão da defasagem espacial da variável dependente. A tabela 27 apresenta os resultados para quatro tipos de matrizes de pesos espaciais: distância inversa ao quadrado (1), distância inversa ao quadrado normalizada pelo método de Ord²⁸ (1975) (2), 15 (quinze) vizinhos mais próximos (3) e contiguidade *queen* de primeira ordem (4). Todas as matrizes, exceto a segunda, foram normalizadas na linha.

²⁸ A normalização na linha das matrizes de distância inversa pode provocar uma má especificação do modelo, pois a interpretação do efeito do decaimento da distância deixa de ser válida (ANSELIN, 1988). A normalização na linha torna a matriz assimétrica e retira o significado de redução da influência de *i* sobre *j* com o aumento da distância entre *i* e *j*. Por isso, Ord (1975) propôs uma transformação que preserva as proporções mútuas entre os elementos da matriz W_0 . Para uma matriz de distância inversa, essa transformação evita que a matriz perca sua interpretação econômica de decaimento da distância: $W = D^{-0,5}W_0D^{-0,5}$. Onde *D* é uma matriz diagonal composta pela soma das linhas da matriz de distância inversa, W_0 .

Tabela 27: Função de produção do conhecimento com defasagem espacial: estimações por *System GMM* (2001 a 2010)¹

| | (1) | (2) | (3) | (4) |
|--|--------------------------------|-------------------------------------|---------------------------|-------------------------|
| Var. dependente: \ln (Patentes por 100 mil habitantes) | W^e : Dist. Inv. Quadr | W^e : Dist. Inv. Quadr. Ord | W^e : 15 vizinhos | W^e : <i>Queen</i> |
| $\ln Patentes_{t-1}$ | 0,139*** (0,001) | 0,139*** (0,001) | 0,132*** (0,001) | 0,127*** (0,003) |
| $\ln Patentes_{t-2}$ | 0,043 (0,256) | 0,042 (0,275) | 0,036 (0,349) | 0,034 (0,373) |
| $W^c \ln Patentes$ | 0,275*** (0,000) | 0,270*** (0,000) | 0,246*** (0,001) | 0,249*** (0,000) |
| $W^m \ln Patentes$ | -0,089 (0,425) | -0,101 (0,368) | -0,121 (0,291) | -0,116 (0,319) |
| $W^e \ln Patentes$ | 0,462** (0,032) | 0,487** (0,024) | 0,704*** (0,000) | 0,469*** (0,006) |
| $\ln PoTec$ | 0,013 (0,205) | 0,013 (0,234) | 0,006 (0,619) | 0,012 (0,279) |
| $\ln PeDun$ | 0,034** (0,017) | 0,033** (0,019) | 0,035*** (0,008) | 0,032** (0,019) |
| $\ln CapHumano$ | 0,123*** (0,010) | 0,125*** (0,010) | 0,063 (0,127) | 0,120** (0,016) |
| $\ln EscUrbana$ | 0,693*** (0,005) | 0,668*** (0,005) | 0,792*** (0,001) | 0,678*** (0,010) |
| $\ln EscUrbanaQd$ | 0,035*** (0,006) | 0,034*** (0,007) | 0,039*** (0,002) | 0,034** (0,019) |
| $\ln Transformação$ | 0,013 (0,722) | 0,012 (0,743) | -0,003 (0,922) | 0,026 (0,462) |

(Continua)

| Var. dependente: ln (Patentes por 100 mil habitantes) | (1) | (2) | (3) | (4) |
|---|---|--|------------------------------------|----------------------------------|
| | W ^e : Dist. Inv. Quadr | W ^e : Dist. Inv. Quadr. Ord | W ^e : 15 vizinhos | W ^e : <i>Queen</i> |
| <i>lnEspProdutiva</i> | -0,083 (0,377) | -0,082 (0,377) | -0,069 (0,455) | -0,090 (0,323) |
| Constante | 3,587*** (0,003) | 3,500*** (0,003) | 4,004*** (0,000) | 3,624*** (0,003) |
| <i>Dummy</i> de ano | Sim | Sim | Sim | Sim |
| Microrregiões/Observações | 558/4464 | 558/4464 | 558/4464 | 558/4464 |
| <i>Lags</i> utilizados como instrumentos ² | 2 e 3 | 2 e 3 | 2 e 3 | 2 e 3 |
| Nº de instrumentos | 35 | 35 | 35 | 35 |
| Teste F | (0,000) | (0,000) | (0,000) | (0,000) |
| AR (1) | (0,000) | (0,000) | (0,000) | (0,000) |
| AR (2) | (0,467) | (0,459) | (0,463) | (0,225) |
| Teste Hansen | (0,680) | (0,699) | (0,624) | (0,617) |
| <i>Difference-in-Hansen</i> : Ho: exogeneidade de subgrupos de instrumentos | | | | |
| Instrumentos para a equação em diferenças ³ | (0,448) | (0,452) | (0,491) | (0,329) |
| | | | | <i>(Continua)</i> |

| Var. dependente: ln (Patentes por 100 mil habitantes) | (1) | (2) | (3) | (4) |
|---|----------------------------------|-------------------------------------|----------------------------------|----------------------------------|
| | W^e : Dist. Inv. Quadr. | W^e : Dist. Inv. Quadr. Ord | W^e : 15 vizinhos | W^e : <i>Queen</i> |
| Instrumentos para a equação em nível ⁴ | (0,740) | (0,762) | (0,615) | (0,782) |
| $\Delta \ln Patentes_{t-1}$ para a equação em nível | (0,989) | (0,987) | (0,845) | (0,985) |
| $\Delta W^c \ln Patentes_{t-1}$, $\Delta W^m \ln Patentes_{t-1}$ e $\Delta W^e \ln Patentes_{t-1}$ para a equação em nível | (0,467) | (0,512) | (0,405) | (0,367) |
| $\Delta \ln Controles_{t-1}$ para a equação em nível | (0,923) | (0,931) | (0,790) | (0,905) |
| Estimativa de $\ln Patentes_{t-1}$ por MQO | 0,396 ^{***} (0,000) | 0,393 ^{***} (0,000) | 0,391 ^{***} (0,000) | 0,391 ^{***} (0,000) |
| Estimativa de $\ln Patentes_{t-2}$ por MQO | 0,316 ^{***} (0,000) | 0,312 ^{***} (0,000) | 0,311 ^{***} (0,000) | 0,311 ^{***} (0,000) |
| Estimativa de $\ln Patentes_{t-1}$ por EF | -0,025 (0,347) | -0,025 (0,348) | -0,026 (0,328) | -0,026 (0,331) |
| Estimativa de $\ln Patentes_{t-2}$ por EF | -0,081 ^{***} (0,001) | -0,081 ^{***} (0,001) | -0,081 ^{***} (0,001) | -0,081 ^{***} (0,001) |

Fonte: Elaboração própria com o uso do *software* Stata 12.0.

¹ Foram consideradas como variáveis endógenas: $\ln Patentes_{t-1}$, $W^c \ln Patentes$, $W^m \ln Patentes$, $W^e \ln Patentes$, $\ln PoTec$, $\ln PeDun$, $\ln Transformação$ e $\ln EspProdutiva$. Foram consideradas como variáveis exógenas: $\ln CapHumano$, $\ln EscUrbana$ e $\ln EscUrbanaQd$.

² Foram utilizados o segundo e o terceiro *lags* das variáveis endógenas como instrumentos para a equação em diferenças e o primeiro *lag* da diferença para a equação em nível.

³ Instrumentos para a equação em diferenças: $\ln Patentes_{t-2}$, $\ln Patentes_{t-3}$, $W^c \ln Patentes_{t-2}$, $W^c \ln Patentes_{t-3}$, $W^m \ln Patentes_{t-2}$, $W^m \ln Patentes_{t-3}$, $W^e \ln Patentes_{t-2}$, $W^e \ln Patentes_{t-3}$, $\ln PoTec_{t-2}$, $\ln PoTec_{t-3}$, $\ln PeDun_{t-2}$, $\ln PeDun_{t-3}$, $\ln Transformação_{t-2}$, $\ln Transformação_{t-3}$, $\ln EspProdutiva_{t-2}$ e $\ln EspProdutiva_{t-3}$.

⁴ Instrumentos para a equação em nível: $\Delta \ln Patentes_{t-1}$, $\Delta W^c \ln Patentes_{t-1}$, $\Delta W^m \ln Patentes_{t-1}$, $\Delta W^e \ln Patentes_{t-1}$, $\Delta \ln PoTec_{t-1}$, $\Delta \ln PeDun_{t-1}$, $\Delta \ln Transformação_{t-1}$ e $\Delta \ln EspProdutiva_{t-1}$.

Em todas as especificações os instrumentos foram restringidos ao segundo e terceiro *lags* das variáveis endógenas para evitar a proliferação de instrumentos e o consequente enfraquecimento do teste Hansen²⁹. Os testes de especificação indicam a validade dos modelos. Os testes AR(1) e AR(2) indicam a presença de autocorrelação serial de primeira ordem e ausência de autocorrelação de segunda ordem, como esperado. O teste Hansen não rejeita a hipótese nula de exogeneidade do grupo total de instrumentos e correta especificação do modelo com *p-valores* acima dos níveis convencionais de significância, como recomendado por Roodman (2009) (0,680; 0,699; 0,624 e 0,617, respectivamente). Os testes *Difference in Hansen* mostram que os subgrupos de instrumentos para a equação em diferenças e para a equação em nível são ortogonais ao termo de erro, expurgando adequadamente a endogeneidade das variáveis. Os testes *Difference in Hansen* para os instrumentos gerados pelas variáveis de interesse do modelo mostram que esses instrumentos são válidos. Portanto, os testes de especificação indicam a confiabilidade dos resultados obtidos.

De forma geral, a magnitude e a significância das variáveis nas estimações espaciais são muito parecidas com os resultados obtidos na estimação não espacial. Os coeficientes estimados para a defasagem temporal da variável dependente encontram-se entre as estimativas obtidas por EF e por MQO, indicando uma especificação correta do modelo. Esses coeficientes são altamente significativos e estáveis entre as especificações, indicando que a inovação regional do período corrente é significativamente influenciada pelas inovações do período anterior.

A defasagem espacial da variável dependente é positiva e significativa para os quatro tipos de pesos espaciais considerados (a 5% utilizando matrizes de distância inversa ao quadrado, e a 1% utilizando a matriz de 15 vizinhos mais próximos e a matriz *queen*). Isso indica que existe transbordamento espacial do conhecimento entre as regiões brasileiras, como verificado por Gonçalves e Almeida (2009) e Araujo e Garcia (2016). Dessa forma, a performance das atividades de inovação em uma região não depende apenas de fatores internos, mas também do estoque externo de conhecimento. Exceto para a matriz de 15 vizinhos mais próximos, os demais coeficientes são muito parecidos. A despeito dessa evidência de que o conhecimento flui entre as regiões brasileiras mais próximas, a defasagem espacial não possibilita

²⁹ Foi feito um teste de robustez, estimando-se os modelos usando os *lags* de 2 a 5 como instrumentos. De forma geral, os resultados são muito parecidos, não havendo mudanças importantes na significância ou na magnitude dos coeficientes. Os resultados encontram-se no Apêndice.

diferenciar quais são os canais pelos quais esse fluxo ocorre. A proximidade espacial não é relevante para a difusão do conhecimento por si só, mas sim porque essa proximidade contribui significativamente para o estabelecimento de interações concretas entre as regiões, por meio das quais o conhecimento pode ser difundido. Acredita-se que o coeficiente da defasagem espacial reflete a soma dos efeitos dos vários canais de difusão do conhecimento para os quais a proximidade espacial é relevante, tais como parcerias em projetos de *PeD*, mobilidade de inventores e de mão de obra qualificada, relações comerciais e de prestação de serviços especializados.

A defasagem das patentes das demais regiões utilizando a matriz de pesos estimada pelas relações de coinvenção inter-regional é altamente significativa e positiva nas quatro estimações da tabela 27, e os coeficientes são muito parecidos entre elas. Isso significa que o acesso ao conhecimento externo por meio das parcerias entre inventores de regiões diferentes para desenvolver uma inovação contribui significativamente para aumentar a base de conhecimento local e a performance inovativa regional. Observa-se uma queda na magnitude do coeficiente em relação ao modelo não espacial da tabela 26, sugerindo que há alguma sobreposição entre os efeitos captados pela defasagem espacial e pela defasagem das coinvenções. A defasagem espacial reflete a importância da proximidade geográfica para a difusão do conhecimento, independente da forma de proximidade espacial considerada (pesos maiores para regiões mais próximas fisicamente, quinze vizinhos mais próximos ou contiguidade). Acredita-se que o efeito positivo da proximidade geográfica para o fluxo do conhecimento se deve à importância da proximidade geográfica para as relações inter-regionais que difundem conhecimento de uma região para a outra. Especificamente, a proximidade espacial contribui para as relações de coinvenção entre regiões. Apesar do uso crescente de tecnologias de comunicação que permitem trocas de informações e de conhecimento através do espaço, reduzindo o efeito da distância, os custos de procurar parceiros potenciais, de negociar entre as partes, formar contratos, acordos sobre o montante de informação e conhecimento que devem ser trocados, gerenciamento do projeto comum, bem como monitorar o desempenho do parceiro, provavelmente serão menores se os agentes puderem interagir e se encontrar frequentemente (MIGUÉLEZ; MORENO, 2013a). E esses encontros tendem a ocorrer mais facilmente se as regiões forem mais próximas fisicamente.

A grande diferença entre os coeficientes da defasagem espacial e os coeficientes da defasagem de coinvenções indica que existem outros canais relevantes de difusão do

conhecimento inter-regional, cujos efeitos estão abrigados sob o coeficiente da defasagem espacial.

A defasagem das patentes utilizando a mobilidade estimada entre pares regiões como pesos de ponderação permanece não significativa, contrariando as expectativas. Esse resultado evidencia que a mobilidade dos inventores brasileiros entre as regiões do país não é uma forma relevante de difusão do conhecimento e de estímulo à inovação regional. Como já discutido, esse resultado pode estar relacionado à baixa capacidade de absorção regional, a uma sobreposição entre o conhecimento dos inventores móveis e o conhecimento da região de destino, como indicado pelo papel significativo da proximidade tecnológica na determinação da mobilidade inter-regional, e à baixa interação das firmas com o ambiente externo.

Dos insumos regionais do conhecimento, apenas o *PeD* universitário (*lnPeDun*) e o capital humano (*lnCapHumano*) são significativos para promover a inovação regional. O *PeD* industrial (*lnPoTec*) permanece não significativo em todas as estimações, indicando o baixo nível de investimento em *PeD* pela indústria brasileira. O *PeD* universitário, ao contrário, permanece positivo e significativo em todas as estimações (a 5% nas estimações 1, 2 e 4, e a 1% na especificação 3) e com coeficiente estável em torno de 0,03 tanto nas estimações espaciais quanto na especificação não espacial (tabela 26). O capital humano permanece positivo e significativo em todas as estimações, exceto a 3, com coeficiente em torno de 0,12. A escala urbana e a escala urbana ao quadrado são positivas e altamente significativas em todas as especificações mostrando que as vantagens da aglomeração contribuem para a performance inovativa regional e que as eventuais externalidades negativas de uma grande escala urbana são compensadas pelos benefícios da aglomeração. A participação da indústria de transformação e a especialização produtiva não são significativas para a inovação regional.

A atribuição das patentes às regiões de origem foi feita utilizando-se a informação sobre a residência do inventor constante nos documentos de patente e fornecida pela BADEPI 1.0 e pela RAIS/MTE. As patentes para as quais não foi possível identificar o local de residência de nenhum de seus inventores não puderam ser consideradas na análise. Utilizando-se as patentes com identificação do local de residência de pelo menos um de seus inventores, observou-se que das 558 microrregiões brasileiras, 130 não produziram nenhuma patente no período analisado. Utilizando as mesmas especificações apresentadas nas tabelas 26 e 27, a função de produção do

conhecimento foi reestimada para o subgrupo de regiões que produziram pelo menos uma patente no período analisado (428 microrregiões) para verificar a robustez dos resultados (tabela 28). A coluna 1 apresenta os resultados para a FPC não espacial, e as colunas 2 a 5 apresentam os resultados considerando-se a defasagem espacial da variável dependente obtida com a matriz de distância inversa ao quadrado, a matriz de distância inversa ao quadrado normalizada pelo método de Ord (1975), a matriz com os 15 vizinhos mais próximos e a matriz de contiguidade *queen* de primeira ordem, respectivamente.

Os testes de especificação indicam que os instrumentos utilizados são ortogonais ao termo de erro, tanto para o grupo total de instrumentos, quanto para os subgrupos mais relevantes: instrumentos para a equação em diferenças, para a equação em nível e os instrumentos baseados nas variáveis de interesse (defasagem temporal e espacial da variável dependente e as defasagens baseadas na coinvenção e na mobilidade). Os coeficientes da defasagem temporal encontram-se entre os valores estimados por EF e por MQO, sugerindo que o modelo está corretamente especificado.

Os resultados obtidos utilizando-se essa subamostra confirmam aqueles encontrados para o total das microrregiões brasileiras em termos de significância e de magnitude para a defasagem temporal e para a defasagem pelas coinvenções da variável dependente. O desempenho inovador da região em $t-1$ influencia positivamente a performance corrente das inovações locais, enquanto o comportamento em $t-2$ permanece não significativo. O acesso ao conhecimento das demais regiões por meio da parceria entre inventores de locais diferentes é positivo e altamente significativo em todas as especificações e com coeficientes muito parecidos aos obtidos para o total das microrregiões brasileiras. Esse resultado confirma a importância da coinvenção inter-regional para difundir o conhecimento através do espaço e para estimular as inovações regionais, podendo, conseqüentemente, contribuir para o desenvolvimento regional. Por outro lado, a mobilidade dos inventores permanece não significativa, corroborando o achado anterior de que, apesar da importância do conhecimento tácito para a inovação, a mobilidade inter-regional dos inventores brasileiros ainda é insuficiente para impactar a inovação regional.

A defasagem espacial permanece significativa em todas as especificações e chama a atenção o aumento substancial dos coeficientes, observado em todas as estimações. Esse aumento reflete o fato natural de que o transbordamento do conhecimento é maior entre as regiões patenteadoras do que entre as regiões que não

realizaram inovações pelo simples fato de que estas últimas não possuem conhecimento (na forma de patentes) para ser difundido.

A *proxy* para o *PeD* industrial (*lnPoTec*) continua não significativa, corroborando o achado anterior de que o investimento em *PeD* industrial é baixo e não contribui para aumentar as inovações regionais. Por outro lado, o *PeD* universitário (*lnPeDun*) permanece significativo em todas as especificações e com coeficiente estável, confirmando a importância da pesquisa universitária local para a performance inovativa regional.

O capital humano é altamente significativo em todas as especificações, exceto na estimação 4 em que a significância é de 5% e cujo coeficiente não foi significativo na mesma especificação para o total das regiões brasileiras. Destaca-se o grande aumento na magnitude dos coeficientes, observado para todas as especificações. Na estimação sem a defasagem espacial (coluna 1) o coeficiente do capital humano aumentou de 0,19 para 0,35 ao considerar-se apenas as regiões patenteadoras. Nas demais especificações (colunas 2 a 5) observa-se que a magnitude do coeficiente do capital humano é duas vezes maior para a subamostra de regiões patenteadoras do que para o total das microrregiões brasileiras. Esse resultado reforça a grande importância do capital humano para geração de inovações. Dado que a *proxy* utilizada para esta variável foi a proporção da população local com curso superior completo, evidencia-se a relevância do investimento na educação superior para o desempenho inovativo regional.

A escala urbana e a escala urbana ao quadrado deixam de ser significativas na estimação sem defasagem espacial e a última deixa de ser significativa nas estimações espaciais das colunas 3 e 5. As colunas de 2 a 5 mostram que houve uma leve redução na magnitude do coeficiente da escala urbana, implicando um impacto menor das vantagens da aglomeração quando se consideram apenas as regiões patenteadoras. Essa mudança na magnitude do coeficiente pode estar relacionada ao aumento na magnitude do efeito do capital humano. O percentual da população com curso superior tende a ser maior em centros urbanos maiores, de forma que parte do efeito que seria atribuído à escala urbana foi captado pelo capital humano. A participação da indústria de transformação no emprego regional e a especialização produtiva permanecem não significativas, não havendo evidências de que afetem a inovação regional.

Tabela 28: Função de produção do conhecimento para regiões patenteadoras: estimações por *System GMM* (2001 a 2010)¹

| | (1) | (2) | (3) | (4) | (5) |
|---|---------------------|---------------------------------------|---|---------------------------------|----------------------------------|
| Var. dependente: ln (Patentes por 100 mil habitantes) | Sem lag espacial | W ^e : Dist. Inv. Quadr. | W ^e : Dist. Inv. Quadr. Ord | W ^e : 15 vizinhos | W ^e : <i>Queen</i> |
| <i>lnPatentes</i> _{t-1} | 0,141*** (0,001) | 0,137*** (0,001) | 0,137*** (0,001) | 0,129*** (0,002) | 0,120*** (0,006) |
| <i>lnPatentes</i> _{t-2} | 0,052 (0,178) | 0,039 (0,306) | 0,037 (0,337) | 0,032 (0,405) | 0,023 (0,562) |
| W ^c <i>lnPatentes</i> | 0,323*** (0,000) | 0,264*** (0,000) | 0,257*** (0,000) | 0,243*** (0,001) | 0,228*** (0,002) |
| W ^m <i>lnPatentes</i> | -0,103 (0,354) | -0,118 (0,320) | -0,139 (0,252) | -0,154 (0,197) | -0,172 (0,173) |
| W ^e <i>lnPatentes</i> | - | 0,522* (0,097) | 0,574* (0,061) | 0,840*** (0,001) | 0,604** (0,016) |
| <i>lnPoTec</i> | 0,027 (0,558) | 0,011 (0,812) | 0,006 (0,892) | -0,013 (0,790) | -0,008 (0,893) |
| <i>lnPeDun</i> | 0,032* (0,066) | 0,035** (0,017) | 0,034** (0,017) | 0,033** (0,013) | 0,031** (0,026) |
| <i>lnCapHumano</i> | 0,354*** (0,001) | 0,242*** (0,004) | 0,250*** (0,003) | 0,157** (0,038) | 0,252*** (0,009) |
| <i>lnEscUrbana</i> | 0,531 (0,142) | 0,622* (0,060) | 0,557* (0,095) | 0,712** (0,029) | 0,604* (0,076) |
| <i>lnEscUrbanaQd</i> | 0,031 (0,123) | 0,032* (0,095) | 0,028 (0,152) | 0,035* (0,066) | 0,030 (0,132) |

(Continua)

| Var. dependente: (ln Patentes por 100 mil habitantes) | (1) | (2) | (3) | (4) | (5) |
|---|-------------------|---------------------------------------|---|---------------------------------|----------------------------------|
| | Sem lag espacial | W ^e : Dist. Inv. Quadr. | W ^e : Dist. Inv. Quadr. Ord | W ^e : 15 vizinhos | W ^e : <i>Queen</i> |
| <i>lnTransformação</i> | 0,022 (0,843) | -0,013 (0,912) | -0,020 (0,856) | -0,066 (0,525) | -0,028 (0,808) |
| <i>lnEspProdutiva</i> | -0,103 (0,491) | -0,113 (0,432) | -0,103 (0,469) | -0,085 (0,562) | -0,108 (0,460) |
| Constante | 2,861* (0,065) | 3,216** (0,023) | 2,986** (0,034) | 3,510*** (0,010) | 3,256** (0,021) |
| <i>Dummy</i> de ano | Sim | Sim | Sim | Sim | Sim |
| Microrregiões/Observações | 428/3424 | 428/3424 | 428/3424 | 428/3424 | 428/3424 |
| <i>Lags</i> utilizados como instrumentos ² | 2 e 3 | 2 e 3 | 2 e 3 | 2 e 3 | 2 e 3 |
| Nº de Instrumentos | 32 | 35 | 35 | 35 | 35 |
| Teste F | (0,000) | (0,000) | (0,000) | (0,000) | (0,000) |
| AR (1) | (0,000) | (0,000) | (0,000) | (0,000) | (0,000) |
| AR (2) | (0,379) | (0,437) | (0,406) | (0,409) | (0,162) |
| Teste <i>Hansen</i> | (0,675) | (0,610) | (0,652) | (0,699) | (0,611) |

(Continua)

| Var. dependente: ln (Patentes por 100 mil habitantes) | (1) | (2) | (3) | (4) | (5) |
|--|----------------------|--|--|---------------------------------|---------------------------|
| | Sem lag espacial | W ^e : Dist. Inv. Quadr. | W ^e : Dist. Inv. Quadr. Ord | W ^e : 15 vizinhos | W ^e : Queen |
| <i>Difference-in-Hansen: (Ho: exogeneidade de subgrupos de instrumentos)</i> | | | | | |
| Instrumentos para a equação em diferenças ³ | (0,443) | (0,323) | (0,339) | (0,426) | (0,279) |
| Instrumentos para a equação em nível ⁴ | (0,732) | (0,779) | (0,817) | (0,789) | (0,834) |
| $\Delta \ln Patentes_{t-1}$ para a equação em nível | (0,439) | (0,967) | (0,964) | (0,910) | (1,000) |
| $\Delta W^c \ln Patentes_{t-1}$, $\Delta W^m \ln Patentes_{t-1}$ e $\Delta W^e \ln Patentes_{t-1}$ para a equação em nível | (0,604) ⁵ | (0,396) | (0,443) | (0,496) | (0,352) |
| $\Delta \ln Controles_{t-1}$ para a equação em nível | (0,868) | (0,911) | (0,943) | (0,866) | (0,905) |
| Estimativa de $\ln Patentes_{t-1}$ por MQO | 0,394*** (0,000) | 0,388*** (0,000) | 0,385*** (0,000) | 0,383*** (0,000) | 0,382*** (0,000) |
| Estimativa de $\ln Patentes_{t-2}$ por MQO | 0,317*** (0,000) | 0,308*** (0,000) | 0,305*** (0,000) | 0,304*** (0,000) | 0,303*** (0,000) |
| Estimativa de $\ln Patentes_{t-1}$ por EF | -0,025 (0,347) | -0,025 (0,347) | -0,025 (0,348) | -0,026 (0,327) | -0,026 (0,328) |
| Estimativa de $\ln Patentes_{t-2}$ por EF | -0,079*** (0,001) | -0,081*** (0,001) | -0,081*** (0,001) | -0,081*** (0,001) | -0,081*** (0,001) |

Fonte: Elaboração própria com o uso do *software* Stata 12.0.

¹ Foram consideradas como variáveis endógenas: $\ln Patentes_{t-1}$, $W^c \ln Patentes$, $W^m \ln Patentes$, $W^e \ln Patentes$, $\ln PoTec$, $\ln PeDun$, $\ln Transformação$ e $\ln EspProdutiva$.

Foram consideradas como variáveis exógenas: $\ln CapHumana$, $\ln EscUrbana$ e $\ln EscUrbanaQd$.

² Foram utilizados o segundo e o terceiro lags das variáveis endógenas como instrumentos para a equação em diferenças e o primeiro lag da diferença para a equação em nível.

³ Instrumentos para a equação em diferenças: $\ln Patentes_{t-2}$, $\ln Patentes_{t-3}$, $W^c \ln Patentes_{t-2}$, $W^c \ln Patentes_{t-3}$, $W^m \ln Patentes_{t-2}$, $W^m \ln Patentes_{t-3}$, $W^e \ln Patentes_{t-2}$, $W^e \ln Patentes_{t-3}$, $\ln PoTec_{t-2}$, $\ln PoTec_{t-3}$, $\ln PeDun_{t-2}$, $\ln PeDun_{t-3}$, $\ln Transformação_{t-2}$, $\ln Transformação_{t-3}$, $\ln EspProdutiva_{t-2}$, e $\ln EspProdutiva_{t-3}$

⁴ Instrumentos para a equação em nível: $\Delta \ln Patentes_{t-1}$, $\Delta W^c \ln Patentes_{t-1}$, $\Delta W^m \ln Patentes_{t-1}$, $\Delta W^e \ln Patentes_{t-1}$, $\Delta \ln PoTec_{t-1}$, $\Delta \ln PeDun_{t-1}$, $\Delta \ln Transformação_{t-1}$ e $\Delta \ln EspProdutiva_{t-1}$.

⁵ Refere-se ao *p*-valor do teste *Difference in Hansen* para o subgrupo de instrumentos composto por $\Delta W^c \ln Patentes_{t-1}$ e $\Delta W^m \ln Patentes_{t-1}$, pois esta especificação não inclui defasagem espacial.

6 CONSIDERAÇÕES FINAIS

Essa tese teve como objetivo principal avaliar o papel da coinvenção e da mobilidade de inventores entre as microrregiões brasileiras na difusão do conhecimento e na inovação regional. Para isso foi utilizada a abordagem da função de produção do conhecimento regional modificada para incorporar as interações regionais promovidas por esses dois fenômenos. Aproveitando-se do instrumental da Econometria Espacial essas interações foram representadas na função de produção por meio de matrizes de pesos construídas para captar as relações bilaterais de coinvenção e mobilidade inter-regional. Assim, os objetivos secundários foram avaliar o efeito das distâncias física, cognitiva e institucional sobre a coinvenção e a mobilidade inter-regional de inventores, e construir matrizes de pesos que refletissem a interação regional decorrente da colaboração para inovar e da mobilidade dos inventores entre as regiões. Essas matrizes foram estimadas ano a ano utilizando-se modelos gravitacionais para a coinvenção e para a mobilidade.

Os resultados do modelo de mobilidade inter-regional mostraram que as proximidades espacial, institucional e cognitiva são relevantes para a decisão de migração dos inventores brasileiros. Eles preferem se mudar para regiões mais próximas geograficamente, que sejam contíguas e que pertençam à mesma unidade da federação. Esses achados confirmam a importância da distância física na escolha dos inventores e implicam que a difusão inter-regional do conhecimento que poderia ocorrer por meio dessa mobilidade seria espacialmente limitada. Além disso, observa-se que a proximidade institucional também é importante, significando que os inventores preferem se mudar para locais com costumes e aspectos culturais parecidos com os de seu local de origem.

A proximidade cognitiva também é um fator relevante na escolha da região de destino dos inventores. Eles preferem se mudar para locais com a mesma base de conhecimento tecnológico. Esse resultado pode ser explicado pelo maior número de oportunidades de trabalho em sua área de *expertise* ou porque as relações sociais e de trabalho tendem a ser desenvolver dentro da mesma comunidade epistêmica.

Apesar da reconhecida importância do conhecimento tácito para a inovação e da relevância das interações face a face para sua difusão, os resultados da função de produção do conhecimento regional mostraram que a mobilidade dos inventores não promove a difusão do conhecimento entre as microrregiões brasileiras. Com base nas

evidências obtidas nas estimações, podemos elencar 3 (três) razões possíveis para que essa mobilidade não contribua para difundir conhecimento. Uma explicação plausível para esse resultado é a alta rotatividade de inventores observada nas principais regiões patenteadoras do país. A aquisição do conhecimento tácito, relevante para que o inventor desenvolva inovações, é cumulativa e depende do processo de aprendizagem vivenciado no âmbito da firma. Quanto maior a rotatividade dos inventores na região, menor é a possibilidade de acumulação de conhecimento específico à firma que possa contribuir para o desenvolvimento de inovações.

Esse conhecimento perdido com a saída de inventores da região pode não ser compensado pela entrada de novos inventores. Dado que a proximidade tecnológica é um fator relevante na escolha da região de destino do inventor, é possível que ocorra uma sobreposição de elementos do conhecimento que acaba por prejudicar uma combinação geradora de novo conhecimento. Esses elementos devem ser relacionados, porém diferentes o suficiente para que se complementem e sua combinação seja capaz de promover a geração de um novo conhecimento ou de contribuir para o desenvolvimento de uma inovação. Nesse sentido, uma grande proximidade cognitiva pode acabar impedindo que a região se beneficie do estoque externo de conhecimento.

Uma terceira razão possível para que a mobilidade dos inventores não promova a difusão do conhecimento entre as regiões brasileiras é a baixa capacidade de absorção dessas regiões. A capacidade de absorção se refere ao conhecimento acumulado necessário para que a firma identifique, assimile e utilize o conhecimento externo para o desenvolvimento de inovações. Essa capacidade é construída por meio das atividades de pesquisa e desenvolvimento (PeD). Como observamos que esse insumo regional do conhecimento não contribui significativamente para o desempenho inovativo regional, podemos concluir que a capacidade de absorção regional é baixa, impedindo que as regiões assimilem e utilizem de forma útil o conhecimento externo disponibilizado pelos inventores móveis.

Da mesma forma que a mobilidade, a coinvenção inter-regional também depende significativamente das variáveis espaciais. Os inventores preferem colaborar com colegas de regiões mais próximas fisicamente, com fronteira comum e que pertençam à mesma unidade da federação. Esse resultado implica que a distância geográfica aumenta os custos de deslocamento para encontros face a face ocasionais, bem como eleva os custos de identificação e seleção de parceiros potenciais e de gerenciamento de um projeto em comum. A proximidade institucional também afeta

positivamente a coinvenção, revelando que os inventores preferem desenvolver projetos com parceiros que possuem as mesmas características culturais.

A proximidade cognitiva não é importante para que os inventores escolham seus parceiros para inovar. Esse resultado sugere que a coinvenção pode ser usada como alternativa para solução de problemas em áreas do conhecimento que o inventor não domina ou para o desenvolvimento de projetos que envolvem tecnologias diferentes, mas complementares.

Dessa forma, a coinvenção inter-regional contribui para a difusão do conhecimento entre as regiões e para o desempenho inovativo local. Os resultados da função de produção do conhecimento mostram que a colaboração entre inventores de locais diferentes é um importante canal de difusão inter-regional do conhecimento, contribuindo significativamente para a performance inovativa das regiões.

Dado que os inventores preferem colaborar com pares de regiões mais próximas geograficamente, a difusão do conhecimento viabilizada por essas colaborações será espacialmente localizada. Esses resultados permitem concluir que a difusão do conhecimento mediada pelos mecanismos de mercado é espacialmente localizada porque os agentes que incorporam e transmitem o conhecimento tendem a se deslocar de forma limitada no espaço. A localização espacial dos fluxos de conhecimento independe de serem fluxos mediados pelo mercado ou de serem transbordamentos puros do conhecimento.

Ao contrário do PeD industrial, o PeD universitário contribui de forma significativa para a inovação regional. Isso indica que o conhecimento gerado nas universidades é difundido, intencionalmente ou não, para o setor industrial e é utilizado no desenvolvimento de inovações. O capital humano também é um fator importante para o desempenho inovativo regional. Considerando que o capital humano é representado pela proporção da população local com curso superior completo, esse resultado reforça a importância das universidades para a promoção da performance inovativa local.

Os resultados obtidos nessa tese permitem inferir que o aumento da intensidade inovativa regional demanda ações que promovam o desenvolvimento das redes de colaboração em pesquisa entre regiões; o investimento em pesquisa e desenvolvimento do setor industrial, o que contribui para o desenvolvimento local de inovações de forma direta enquanto insumo do conhecimento, e de forma indireta por meio da construção de capacidade de absorção do conhecimento externo; o investimento nas universidades em

nível regional de forma a promover a pesquisa universitária e a formação de capital humano.

Além de confirmar o papel da proximidade espacial e tecnológica sobre a coinvenção e a mobilidade inter-regional dos inventores, os resultados dessa tese contribuem para o avanço na literatura sobre a geografia da inovação no Brasil ao permitir a compreensão de uma pequena parte da “caixa preta” da função de produção. Mostrou-se que uma parte da dependência espacial das inovações regionais e dos transbordamentos localizados do conhecimento, verificadas em várias aplicações empíricas para o Brasil, é um reflexo da fricção que o espaço representa nas colaborações inter-regionais para inovar. Uma parte significativa do que é tratado como transbordamento do conhecimento pela literatura da geografia da inovação, na verdade é uma difusão do conhecimento mediada por relações formais de mercado. Essa compreensão dos mecanismos concretos pelos quais o conhecimento se difunde é relevante para que possam ser incentivadas ações que promovam esses mecanismos, aumentando a difusão do conhecimento no espaço e, conseqüentemente, a performance inovativa regional, o crescimento e o desenvolvimento econômico.

Os resultados dessa tese contribuem para o avanço da compreensão sobre o papel da coinvenção e da mobilidade de inventores na inovação das regiões brasileiras. A colaboração entre inventores de regiões distintas é um canal de transmissão de conhecimento entre essas regiões, contribuindo para aumentar a produção local de inovações e, conseqüentemente, o desenvolvimento regional. Dessa forma, políticas públicas que objetivem promover a inovação tecnológica devem contemplar, entre outras medidas, mecanismos que facilitem e incentivem parcerias entre inventores de diferentes regiões. A realização de eventos que disseminem informações sobre as atividades inovadoras das regiões e que coloquem os agentes do conhecimento de uma região em contato com seus pares de outros locais são importantes para viabilizar a troca de ideias e estimular a cooperação para inovar. São exemplos dessas ações as feiras de negócios, *workshops*, congressos, cursos e visitas técnicas a empresas e universidades com foco nos desenvolvimentos de novos produtos e processos bem como no desenvolvimento da ciência básica.

Além de viabilizar a difusão do conhecimento por meio do estímulo à cooperação inter-regional, as políticas de incentivo à inovação devem investir nos insumos locais do conhecimento: capital humano e PeD. Os resultados dessa tese corroboram a teoria e as evidências empíricas acerca da relevância desses insumos para

a promoção da inovação. A ampliação das dotações regionais desses insumos do conhecimento demanda o investimento na criação e expansão das universidades, principalmente em suas atividades de PeD. O conhecimento gerado pelo PeD universitário melhora a performance inovativa regional, implicando que existe uma transmissão desse conhecimento para o setor industrial, seja ela intencional ou não. Dessa forma, além de estimular o PeD universitário, o desenho de políticas públicas para promover a inovação deve também contemplar mecanismos de interação universidade-empresa que intensifiquem essa transmissão do conhecimento.

Ao mesmo tempo em que os resultados dessa tese contribuem para o avanço do conhecimento sobre a inovação regional no Brasil, eles também apontam direções para novas pesquisas. Uma questão instigante é a forma de se mensurar a inovação regional. Nessa tese a inovação regional foi avaliada por meio do patenteamento local, de forma que todos os resultados são analisados à luz dessa variável. O desenvolvimento e a utilização de outro indicador da inovação regional poderão contribuir ainda mais para a compreensão da inovação local e fazem parte da nossa agenda de pesquisa.

Enquanto a cooperação inter-regional de inventores é um meio efetivo de aproveitamento do conhecimento externo, contribuindo para a inovação regional, observamos que a mobilidade dos inventores entre locais distintos não afeta a performance regional das inovações. Portanto, uma avenida de pesquisa aberta por esses resultados é compreender as razões pelas quais a mobilidade de inventores não contribui para a inovação regional. Nesse sentido, pelo menos 3 (três) fenômenos podem ser investigados: os efeitos da rotatividade regional de inventores, a importância da proximidade cognitiva entre as regiões de origem e de destino, e o papel da capacidade de absorção regional no aproveitamento do estoque externo de conhecimento. Dando continuidade à investigação iniciada nessa tese, todas essas questões fazem parte de nossa agenda futura de pesquisa.

REFERÊNCIAS

ABREU, M. Absorptive capacity in a regional context. In: COOKE, P., ASHEIM, B., BOSCHMA, R., MARTIN, R., SCHWARTZ, D., TÖDTLING, F. (Eds.), **Handbook of regional innovation and growth**. Cheltenham: Edward Elgar Publishing, pp. 211–221, 2011.

ACS, Z. J., AUDRETSCH, D. B.; FELDMAN, M.P. Real effects of academic research: comment. **American Economic Review**, 82 (1): 363-367, 1992.

AGRAWAL, A., COCKBURN, I.; McHALE, J. Gone But Not Forgotten: Labor Flows, Knowledge Spillovers, and Enduring Social Capital. **Journal of Economic Geography**, 6(5), pp. 571-591, 2006.

ALMEIDA, E. **Econometria espacial aplicada**. Campinas, Alínea, 2012.

ALMEIDA, P.; KOGUT, B. Localization of Knowledge and the Mobility of Engineers in Regional Redes. **Management Science**, 45 (7), pp. 905-17, 1999.

ANDERSON, J., AND E. VAN WINCOOP Gravity with Gravitas: A Solution to the Border Puzzle, **American Economic Review** 93, pp. 170–192, 2003.

ANDERSON, J. E.; YOTOV, Y. V. Gold standard gravity. NBER **Working Papers** No 17835, 2012.

ARAÚJO, B. C., CAVALCANTE, L. R. E ALVES, P. Variáveis *proxy* para os gastos empresariais em inovação com base no pessoal ocupado técnico-científico disponível na Relação Anual de Informações Sociais (RAIS). **Radar: Tecnologia, Produção e Comércio Exterior**, n. 5, pp. 16-21, 2009.

ARAUJO, V. C.; GARCIA, R. Local determinants of innovation and spatial dependence – A spatial Tobit model applied to Brazilian micro-regions. In: **XLL ENCONTRO NACIONAL DE CENTROS DE PÓS-GRADUAÇÃO EM ECONOMIA**. Anais... Foz do Iguaçu: Anpec, 2016.

ARELLANO, M.; BOND, S. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. **The review of economic studies**, v. 58, n. 2, pp. 277-297, 1991.

ARELLANO, M.; BOVER, O. Another look at the instrumental variable estimation of error-components models. **Journal of econometrics**, v. 68, n. 1, pp. 29-51, 1995.

AUDRETSCH, D. B. e FELDMAN, M. P. R&D spillovers and the geography of innovation and production. **The American Economic Review** 86(3): pp. 630-640, 1996.

AUDRETSCH e FELDMAN. R&D spillovers and the geography of innovation and production. In: Henderson, J.V., Thisse, J.-F. (Eds.), **Handbook of Urban and Regional Economics**, vol. 4. Elsevier, Amsterdam, pp. 2713-2739, 2004.

BAUM, F. C. **An Introduction to Modern Econometrics Using Stata**, Texas: Stata Press, 2006.

BAUM, C. F., SCHAFFER, M. E; STILLMAN, S. Instrumental variables and GMM: estimation and testing. **Stata Journal**, 3, n. 1, p. 1-31, 2003.

BERCOVITZ, J. B. L., FELDMAN, M.P. The mechanisms of collaboration in inventive teams: Composition, social redes, and geography, **Research Policy** 40, pp. 81–93, 2011.

BERTOLI, S., & MORAGA, J. F.-H. The size of the cliff at the border. **Regional Science and Urban Economics**, 51, p. 1–6, 2015.

BLUNDELL, R.; BOND, S. Initial conditions and moment restrictions in dynamic panel data models. **Journal of econometrics**, v. 87, n. 1, p. 115-143, 1998.

BOND, S. R. Dynamic panel models: a guide to micro data methods and practice. Institute for Fiscal Studies/Department of Economics, UCL, CEMMAP (Centre for Microdata Methods and Practice) **Working Paper CWPO 9/02**, 2002.

BOSCHMA, R. Proximity and innovation: A critical assessment. **Regional Studies**, 39(1), 61–74, 2005.

BOTTAZZI e PERI (2003). Innovation and spillovers in regions: Evidence from European patent data. **European Economic Review**, 47(4): pp. 687-710.

BRESCHI, S. The geography of knowledge flows. In: COOKE, P., ASHEIM, B., BOSCHMA, R., MARTIN, R., SCHWARTZ, D., TÖDTLING, F. (Eds.), **Handbook of regional innovation and growth**. Cheltenham/Northampton: Edward Elgar, pp. 132-142, 2011.

BRESCHI, S.; LISSONI, F. Mobility and social networks: localised knowledge spillovers revisited. CESPRI **Working Paper** 142, 2003.

BRESCHI, S., LISSONI, F. Mobility of skilled workers and co-invention redes: an anatomy of localized knowledge flows. **Journal of Economic Geography**, 9 (4): pp. 439–468, 2009.

BRESCHI, S. AND LISSONI, F. Cross-firm inventors and social redes: localised knowledge spillovers revisited. **Annales d’Economie et de Statistique**, n. 79/80, pp. 189-209, 2005.

BRESCHI, S. e LISSONI, F. Knowledge spillovers and local innovation systems: A critical survey. **Industrial and Corporate Change**, 10(4), pp. 975-1005, 2001.

CARLINO, G. A.; CHATTERJEE, S.; HUNT, R. M. Urban density and the rate of invention. **Journal of Urban Economics**, v. 61, n. 3, p. 389-419, 2007.

CARLINO, G.; KERR, W. Agglomeration and Innovation. In: DURANTON, HENDERSON e STRANGE (ed). **Handbook of regional and urban economics**, vol 5. 1653 p. 349-404, 2015.

CAMERON, A. C.; TRIVEDI, P. K. **Regression Analysis of Count Data**. Cambridge: Cambridge University Press, 1998.

CHARLOT S, CRESCENZI R, MUSOLESI A. Econometric modelling of the regional knowledge production function in europe. **Journal of Economic Geography** 15: 1227–1259, 2015.

COHEN, W. M., LEVINTHAL, D. .Absorptive capacity: a new perspective on learning and innovation. **Administrative Science Quarterly**, v. 35, n. 1, p. 128, 1990.

COHEN, W., NELSON, R., WALSH, J. Protecting Their Intellectual Assets: Appropriability Conditions and Why US Manufacturing Firms Patent (or Not). **NBER Working Paper** 7552, 2000.

CRESCENZI, R.; GAGLIARDI, L. Moving people with ideas: innovation, inter-regional mobility and firm heterogeneity. **Working Papers of the Taubman Centre, Harvard Kennedy School of Government**, 1, 2015.

DE NEGRI, F. Determinantes da capacidade de absorção nas firmas brasileiras: qual a influencia do perfil da mão-de-obra? In: DE NEGRI, F.; DE NEGRI, J. A.; COELHO, D. (Org.). **Tecnologia, exportação e emprego**. Brasília: Ipea, p. 523, 2006.

EICHENGREEN, B.; IRWIN, D. A. Trade Blocs, Currency Blocs and the Reorientation of World Trade in the 1930s, **Journal of International Economics** 38, pp. 1-24, 1995.

ELHORST, J. P. **Spatial econometrics: From cross-sectional data to spatial panels**. London: Springer, 2014.

FAGGIAN, A. e McCANN, P. Human capital, graduate migration and innovation in British regions. **Cambridge Journal of Economics**, v. 33, p. 317–333, 2009.

FELDMAN, M.P. An Examination of the Geography of Innovation. **Industrial and Corporate Change** ,vol 2, n. 3, pp.417-437, 1993.

FELDMAN, M. P. e FLORIDA, R. The geographic sources of innovation: technological infrastructure and product innovation in the United States. **Annals of the Association of American Geographers**, Vol. 84, n.2, pp. 210-229, 1994.

FELDMAN, M. P.; KOGLER, D. F. Stylized facts in the geography of innovation. In: Hall; B. H.; Rosenberg, N. (eds.) **Handbook of Economics of Innovation**. Vol 1. Elsevier, cap. 8, 2010.

FIGUEIREDO, O.; GUIMARAES P.; WOODWARD, D. Industry localization, distance decay, and knowledge spillovers: following the patent paper trail. **Journal of Urban Economics**. 89:21–31, 2015.

GAGLIARDI, L. Does skilled migration foster innovative performance? Evidence from British local areas. **Papers in Regional Science**, 2014.

GOLGHER, A. B. As cidades e a classe criativa no Brasil: diferenças espaciais na distribuição de indivíduos qualificados nos municípios brasileiros. **Revista Brasileira de Estudos da População**, v. 25, p. 109-129, 2008.

GONÇALVES, E.; MATOS, C. M.; ARAÚJO, I. F. Path-Dependent Dynamics and Technological Spillovers in the Brazilian Regions. **Applied Spatial Analysis and Policy**, p. 1-25, 2018.

GONÇALVES, E., RIBEIRO, D. R. S., FREGUGLIA, R. S. Migração de mão de obra qualificada e inovação: um estudo para as microrregiões brasileiras. In: **XL Encontro Nacional de Economia**, 2012, Porto de Galinhas (PE) Anais... ANPEC, 2012.

GONÇALVES, E.; ALMEIDA, E. S. Innovation and Spatial Knowledge Spillovers: evidence from Brazilian Patent Data. **Regional Studies**, v. 43. n.4, p.513–528, 2009.

GORIN, C. Patterns and determinants of inventors' mobility across European urban areas, **Working Papers** 1615, Groupe d'Analyse et de Théorie Economique Lyon St-Étienne (GATE Lyon St-Étienne), Université de Lyon. 44 p, 2016.

GREUNZ, L. Geographically and technologically mediated knowledge spillovers between European regions. **The annals of regional Science**, Volume 37, Issue 4, pp 657–680, 2003.

GRILICHES, Z. Issues in assessing the contribution of research and development to productivity growth. **Bell Journal of Economics** 10:92–116, 1979.

GRINZA, E.; QUATRARO, F. Workers' replacements and firms' Innovation dynamics: new evidence from italian matched longitudinal data. **Carlo Alberto Notebooks**, n. 550, 2018.

GROSSMAN, G. M.; HELPMAN, E. Trade, Knowledge Spillovers and Growth, **European Economic Review**, 35:3, 517–26, 1991.

HENDERSON, J. V. Understanding knowledge spillovers. **Regional Science and Urban Economics**, 37, pp. 497-508, 2007.

HOISL, K. Tracing mobile inventors: The causality between inventor mobility and inventor productivity. **Research. Policy**. 36:5, 619-636, 2007.

JAFFE, Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value. **American Economic Review**, Vol. 76, Issue 5, p. 984-100, 1986.

JAFFE, A. B; TRAJTENBERG, M. International Knowledge Flows: evidence from Patent Citations. In: JAFFE, A. B; TRAJTENBERG, M. **Patents, Citations and Innovations: a window on the knowledge economy**. Cambridge, MA MIT Press, 2002.

JAFFE, A. B; TRAJTENBERG, M.; HENDERSON, R. Geographic localization of knowledge spillovers as evidenced by patent citations. **Quarterly Journal of economics**, 108(3), 577-98, 1993.

KANNEBLEY JÚNIOR, S.; SHIMADA, E.; DE NEGRI, F. Efetividade da lei do bem no estímulo aos dispêndios em PED: uma análise com dados em painel. **Pesquisa e Planejamento Econômico**, v. 46, n. 3, 2016.

KUKENOVA, M; MONTEIRO, J. A. **Spatial dynamic panel model and system GMM: a Monte Carlo investigation**. University Library of Munich, Germany, 2008.

KUZNETS, S. Inventive activity: problems of definition and measurement. In: Nelson, R.R. (Ed.), **The Rate and Direction of Inventive Activity: Economic and Social Factors**. Princeton University Press, Princeton, NJ, 1962.

LUCAS. On the mechanics of economic development. **Journal of Monetary Economics**, vol. 22, issue 1, pp. 3-42, 1988.

MARSHALL, A. **Principles of economics**. Macmillan, London, 1890.

MENDES, P.S; GONÇALVES; E.; FREGUGLIA, R.S.; Mobilidade interfirmas de trabalhadores no Brasil formal: composição e determinantes. **Pequisa e Planejamento Econômico** 42 (August (2)), 211–238, 2012.

MIGUÉLEZ, E., MORENO, R., SURIÑACH, J. Inventors on the move: tracing inventors' mobility and its spatial distribution. **Papers in Regional Science**, vol 89, number 2, 2010.

MIGUELEZ, E. e MORENO, R. Do labour mobility and technological collaborations foster geographical knowledge diffusion? The case of European regions, **Growth and Change**, 44(2), pp. 321-354, 2013a.

MIGUÉLEZ, E.; MORENO, R. Skilled labour mobility, networks and knowledge creation in regions: a panel data approach. **The Annals of Regional Science**, 51(1), pp 191–212, 2013b.

MIGUÉLEZ, E.; MORENO, R. Research Networks and Inventors' Mobility as Drivers of Innovation: Evidence from Europe, **Regional Studies**, 47:10, 1668-1685, 2013c.

MORAN, P. A. P. The interpretation of statistical maps. **Journal of the Royal Statistical Society**. v. 10, pp. 243-251, 1948.

MORENO, R., PACI e USAI. Spatial Spillovers and Innovation Activity in European Regions, **Environment and planning A**, 37(10), pp.1793-1812, 2005.

MORESCALCHI, A., PAMMOLLI, F., PENNER, O., PETERSEN, A. M., RICCABONI, M. The evolution of networks of innovators within and across borders: Evidence from patent data. **Research Policy**, 44(3), 651–668, 2015.

MUKHERJI, N., SILBERMAN, J. Absorptive capacity, knowledge flows, and innovation in US metropolitan areas. **Journal of Regional Science**. 53 (3), 392–417, 2013.

MURATA, Y., NAKAJIMA, R., OKAMOTO, R., TAMURA, R. Localized knowledge spillovers and patent citations: a distance-based approach. **Review of Economics and Statistics**, Vol. 96, No. 5, pp. 967-985, 2014.

NICKELL, S. Biases in dynamic models with fixed effects, **Econometrica**, v.49, pp. 1417–1426, 1981.

NAGAOKA, S.; MOTOHASHI, K.; GOTO, A. Patent Statistics as an innovation indicator. In: Hall; B. H.; Rosenberg, N. (eds.) **Handbook of Economics of Innovation**. Elsevier, cap. 25, 2010.

NELSON, R.R., WINTER, S.G. **An Evolutionary Theory of Economic Change**. Belknap Press of Harvard University Press, Cambridge, MA, 1982.

ORD, K. Estimation methods for models of spatial interaction, **Journal of the American Statistical Association**, 70:349, 120-126, 1975.

PACI, R.; USAI, S. Knowledge flows across european regions. **The Annals of Regional Science**. 43(3): 669-690, 2009.

POLANYI, M. The Tacit Dimension. In: PRUSAK, L. **Knowledge in Organizations**. Nova Iorque: Harper & Row, 1987.

RAHKO, J. Knowledge spillovers through inventor mobility: the effect on firm-level patenting. **The Journal of Technology Transfer**, 42(3):585–614, 2017.

RAMOS, R. Gravity models: a tool for migration analysis. **IZA World of Labor**, 239, 2016.

RAMOS, R., SURIÑACH, J. A gravity model of migration between ENC and EU. **IREA Working Paper**, 26 pag, 2013.

REIS, M.; AZEVEDO, A. F. Z.; LÉLIS, M. T. C. Os efeitos do novo regionalismo sobre o comércio. **Estudos Econômicos**, v.44, pp. 351-381, 2014.

ROMER, P. Increasing returns and long-run growth. **The Journal of Political Economy** 94, n. 5, pp.1002–1037, 1986.

ROMER, P. Endogenous Technological Change. **Journal of Political Economy**, Vol. 98, No. 5, Part 2, pp. S71-S102, 1990.

ROODMAN, D. How to do xtabond2: an introduction to “Difference” and “System” GMM in Stata. **Working Paper Number 103**, Central for Global Development, 2006.

ROODMAN, D. Practitioners’s corner: a note on the theme of too many instruments. **Oxford Bulletin of Economics and Statistics**, 71, 1, p.136-158, 2009.

SANTOS SILVA, J.; TENREYRO, S. The log of gravity. **The Review of Economics and Statistics**, 88: 641-58, 2006.

SANTOS SILVA, J.; TENREYRO, S. On the existence of the maximum likelihood estimates in Poisson regression. **Economics Letters**. vol 107, N 2, p. 310–12, 2010.

SANTOS SILVA, J.; TENREYRO, S. Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator. **Economics Letters** vol. 112, n. 2, p. 220–2, 2011.

SANTOS SILVA, J., TENREYRO, S.; WINDMEIJER, F. Testing competing models for nonnegative data with many zeros. **Journal of Econometric Methods** 1-18, 2014.

SIMÕES, R.; MARTINS, Agda ; MORO, Sueli . Innovation, urban attributes and scientific structure: a Zero - Inflated - Poisson Model for biotechnology in Brazil. **Journal of Economics and Development Studies**, v. 2, p. 523-553, 2014.

STORPER, M.; VENABLES, A. Buzz: face-to-face contact and the urban economy. **Journal of Economic Geography** 4(4): 351-370, 2004.

TINBERGEN, J. **Shaping the World Economy; Suggestions for an International Economic Policy**. New York: Twentieth Century Fund, 1962.

WIPO (World Intellectual Property Organization). **Guide to the international patent classification**, 2018. Disponível em:
<http://www.wipo.int/export/sites/www/classifications/ipc/en/guide/guide_ipc.pdf>

APÊNDICE

Tabela A1: Mobilidade de inventores da região i para a região j: 2000 a 2011

| Ano | Migrações de i para j | Frequência | (%) |
|------|-----------------------|----------------|---------------|
| 2000 | 0 | 310,781 | 99,99 |
| | entre 1 e 2 | 25 | 0,01 |
| | <i>Total</i> | <i>310,806</i> | <i>100,00</i> |
| 2001 | 0 | 310347 | 99,85 |
| | entre 1 e 53 | 459 | 0,15 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2002 | 0 | 310524 | 99,91 |
| | entre 1 e 21 | 282 | 0,09 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2003 | 0 | 309961 | 99,73 |
| | entre 1 e 61 | 845 | 0,27 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2004 | 0 | 310083 | 99,77 |
| | entre 1 e 38 | 723 | 0,23 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2005 | 0 | 310008 | 99,74 |
| | entre 1 e 51 | 798 | 0,26 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2006 | 0 | 309943 | 99,72 |
| | entre 1 e 67 | 863 | 0,28 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2007 | 0 | 309911 | 99,71 |
| | entre 1 e 45 | 895 | 0,29 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2008 | 0 | 309918 | 99,71 |
| | entre 1 e 47 | 888 | 0,29 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2009 | 0 | 309829 | 99,69 |
| | entre 1 e 48 | 977 | 0,31 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2010 | 0 | 309746 | 99,66 |
| | entre 1 e 43 | 1060 | 0,34 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2011 | 0 | 309814 | 99,68 |
| | entre 1 e 37 | 992 | 0,32 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |

Fonte: Elaboração própria com base nos dados da BADEPI 1.0/INPI e da RAIS/MTE

Tabela A2: Estatísticas descritivas para a mobilidade de inventores: 2000 a 2011

| Ano | Observações | Média | Desvio Padrão | Mínimo | Máximo |
|------|-----------------------------------|---------|---------------|--------|--------|
| 2000 | 310806 | 0,00008 | 0,009 | 0 | 2 |
| 2001 | 310806 | 0,00338 | 0,213 | 0 | 53 |
| 2002 | 310806 | 0,00151 | 0,081 | 0 | 21 |
| 2003 | 310806 | 0,00543 | 0,231 | 0 | 61 |
| 2004 | 310806 | 0,00416 | 0,171 | 0 | 38 |
| 2005 | 310806 | 0,00457 | 0,186 | 0 | 51 |
| 2006 | 310806 | 0,00476 | 0,198 | 0 | 67 |
| 2007 | 310806 | 0,00533 | 0,195 | 0 | 45 |
| 2008 | 310806 | 0,00520 | 0,199 | 0 | 47 |
| 2009 | 310806 | 0,00541 | 0,183 | 0 | 48 |
| 2010 | 310806 | 0,00582 | 0,195 | 0 | 43 |
| 2011 | 310806 | 0,00554 | 0,177 | 0 | 37 |
| Ano | Observações com migração positiva | Média | Desvio Padrão | Mínimo | Máximo |
| 2000 | 25 | 1,04 | 0,20 | 1 | 2 |
| 2001 | 459 | 2,29 | 5,06 | 1 | 53 |
| 2002 | 282 | 1,67 | 2,11 | 1 | 21 |
| 2003 | 845 | 2,00 | 3,97 | 1 | 61 |
| 2004 | 723 | 1,79 | 3,08 | 1 | 38 |
| 2005 | 798 | 1,78 | 3,22 | 1 | 51 |
| 2006 | 863 | 1,71 | 3,35 | 1 | 67 |
| 2007 | 895 | 1,85 | 3,13 | 1 | 45 |
| 2008 | 888 | 1,82 | 3,25 | 1 | 47 |
| 2009 | 977 | 1,72 | 2,78 | 1 | 48 |
| 2010 | 1060 | 1,71 | 2,88 | 1 | 43 |
| 2011 | 992 | 1,73 | 2,61 | 1 | 37 |

Fonte: Elaboração própria com base nos dados da BADEPI 1.0/INPI e da RAIS/MTE

Tabela A3: Estatísticas descritivas para a colaboração inter-regional: 2000 a 2011

| Ano | Observações | Média | Desvio Padrão | Mínimo | Máximo |
|------|--------------------------------------|--------|---------------|--------|--------|
| 2000 | 310806 | 0,0014 | 0,073 | 0 | 16 |
| 2001 | 310806 | 0,0014 | 0,070 | 0 | 11 |
| 2002 | 310806 | 0,0024 | 0,080 | 0 | 16 |
| 2003 | 310806 | 0,0021 | 0,084 | 0 | 12 |
| 2004 | 310806 | 0,0033 | 0,111 | 0 | 21 |
| 2005 | 310806 | 0,0025 | 0,098 | 0 | 17 |
| 2006 | 310806 | 0,0029 | 0,126 | 0 | 27 |
| 2007 | 310806 | 0,0034 | 0,128 | 0 | 30 |
| 2008 | 310806 | 0,0035 | 0,122 | 0 | 27 |
| 2009 | 310806 | 0,0039 | 0,134 | 0 | 28 |
| 2010 | 310806 | 0,0040 | 0,124 | 0 | 24 |
| 2011 | 310806 | 0,0046 | 0,149 | 0 | 29 |
| Ano | Observações com colaboração positiva | Média | Desvio Padrão | Mínimo | Máximo |
| 2000 | 246 | 1,76 | 1,931 | 1 | 16 |
| 2001 | 262 | 1,69 | 1,733 | 1 | 11 |
| 2002 | 530 | 1,40 | 1,335 | 1 | 16 |
| 2003 | 372 | 1,78 | 1,634 | 1 | 12 |
| 2004 | 580 | 1,76 | 1,882 | 1 | 21 |
| 2005 | 434 | 1,80 | 1,908 | 1 | 17 |
| 2006 | 472 | 1,88 | 2,632 | 1 | 27 |
| 2007 | 562 | 1,86 | 2,373 | 1 | 30 |
| 2008 | 606 | 1,78 | 2,120 | 1 | 27 |
| 2009 | 662 | 1,84 | 2,248 | 1 | 28 |
| 2010 | 768 | 1,61 | 1,912 | 1 | 24 |
| 2011 | 762 | 1,87 | 2,359 | 1 | 29 |

Fonte: Elaboração própria com base nos dados da BADEPI 1.0/INPI e da RAIS/MTE

Tabela A4: Coinvenções entre as regiões i e j: 2000 a 2011

| Ano | Copatentes entre i e j | Frequência | (%) |
|------|------------------------|---------------|---------------|
| 2000 | 0 | 310560 | 99,92 |
| | entre 1 e 16 | 246 | 0,08 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2001 | 0 | 310544 | 99,92 |
| | entre 1 e 11 | 262 | 0,08 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2002 | 0 | 310276 | 99,83 |
| | entre 1 e 16 | 530 | 0,17 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2003 | 0 | 310434 | 99,88 |
| | entre 1 e 12 | 372 | 0,12 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2004 | 0 | 310226 | 99,81 |
| | entre 1 e 21 | 580 | 0,19 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2005 | 0 | 310372 | 99,86 |
| | entre 1 e 17 | 434 | 0,14 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2006 | 0 | 310334 | 99,85 |
| | entre 1 e 27 | 472 | 0,15 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2007 | 0 | 310244 | 99,82 |
| | entre 1 e 30 | 562 | 0,18 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2008 | 0 | 310200 | 99,81 |
| | entre 1 e 27 | 606 | 0,19 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2009 | 0 | 310144 | 99,79 |
| | entre 1 e 28 | 662 | 0,21 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2010 | 0 | 310038 | 99,75 |
| | entre 1 e 24 | 768 | 0,25 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |
| 2011 | 0 | 310044 | 99,75 |
| | entre 1 e 29 | 762 | 0,25 |
| | <i>total</i> | <i>310806</i> | <i>100,00</i> |

Fonte: Elaboração própria com base nos dados da BADEPI 1.0/INPI e da RAIS/MTE

Tabela A5: Resultados do modelo gravitacional: mobilidade de inventores entre as microrregiões brasileiras (2001 a 2010)
estimações por MQO – com EF

| Variável dependente: <i>ln(mobilidade + 1)</i> | Anos | | | | |
|---|----------------------|----------------------|----------------------|----------------------|----------------------|
| | 2001 | 2002 | 2003 | 2004 | 2005 |
| <i>Ln(proximidade Tecnológica)</i> | 0,016*** (0,000) | 0,028*** (0,000) | 0,030*** (0,000) | 0,028*** (0,000) | 0,029*** (0,000) |
| <i>Ln(distância geográfica)</i> | -0,021*** (0,000) | -0,035*** (0,000) | -0,037*** (0,000) | -0,036*** (0,000) | -0,034*** (0,000) |
| <i>Dummy de contiguidade</i> | 0,133*** (0,000) | 0,179*** (0,000) | 0,221*** (0,000) | 0,207*** (0,000) | 0,189*** (0,000) |
| <i>Dummy de UF</i> | 0,033*** (0,000) | 0,044*** (0,000) | 0,052*** (0,000) | 0,049*** (0,000) | 0,049*** (0,000) |
| <i>Constante</i> | 0,101*** (0,000) | 0,179*** (0,000) | 0,194*** (0,000) | 0,172*** (0,000) | 0,148*** (0,000) |
| <i>Efeito Fixo da Origem</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> |
| <i>Efeito Fixo do Destino</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> |
| <i>Número de observações</i> | 34694 | 35022 | 34220 | 36188 | 38156 |
| <i>Log-likelihood</i> | 11181,98 | 10713,71 | 9840,85 | 10761,52 | 11858,35 |
| aic | -32729,38 | -23513,79 | -16486,09 | -18868,19 | -24730,93 |
| bic | -28518,84 | -19099,80 | -12285,31 | -14659,64 | -20245,34 |
| r2 | 0,16 | 0,22 | 0,25 | 0,23 | 0,23 |
| r2_a | 0,15 | 0,21 | 0,23 | 0,22 | 0,22 |

(Continua)

| Variável dependente: <i>ln(mobilidade + 1)</i> | Anos | | | | |
|---|----------------------|----------------------|----------------------|----------------------|----------------------|
| | 2006 | 2007 | 2008 | 2009 | 2010 |
| <i>Ln(proximidade Tecnológica)</i> | 0,029*** (0,000) | 0,032*** (0,000) | 0,027*** (0,000) | 0,031*** (0,000) | 0,031*** (0,000) |
| <i>Ln(distância geográfica)</i> | -0,038*** (0,000) | -0,039*** (0,000) | -0,044*** (0,000) | -0,044*** (0,000) | -0,038*** (0,000) |
| <i>Dummy de contiguidade</i> | 0,205*** (0,000) | 0,224*** (0,000) | 0,232*** (0,000) | 0,204*** (0,000) | 0,218*** (0,000) |
| <i>Dummy de UF</i> | 0,055*** (0,000) | 0,053*** (0,000) | 0,055*** (0,000) | 0,050*** (0,000) | 0,056*** (0,000) |
| <i>Constante</i> | 0,165*** (0,000) | 0,185*** (0,000) | 0,181*** (0,000) | 0,167*** (0,000) | 0,190*** (0,000) |
| <i>Efeito Fixo da Origem</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> |
| <i>Efeito Fixo do Destino</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> |
| <i>Número de observações</i> | 34694 | 35022 | 34220 | 36188 | 38156 |
| <i>Log-likelihood</i> | 11181,98 | 10713,71 | 9840,85 | 10761,52 | 11858,35 |
| <i>aic</i> | -21317,96 | -20385,42 | -18687,70 | -20501,04 | -22622,70 |
| <i>bic</i> | -16896,35 | -15975,82 | -14492,74 | -16159,34 | -17946,16 |
| <i>r2</i> | 0,24 | 0,24 | 0,25 | 0,25 | 0,25 |
| <i>r2_a</i> | 0,23 | 0,23 | 0,24 | 0,24 | 0,24 |

Fonte: Elaboração própria com o uso do *software* Stata 12.0
p-valor entre parênteses ****p* < 0,01 ***p* < 0,05 **p* < 0,1

Tabela A6: Resultados do modelo gravitacional: coinvenção entre as microrregiões brasileiras (2001 a 2010) - estimações por MQO – com EF

| Variável dependente: | Anos | | | | |
|------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| <i>ln(mobilidade + 1)</i> | 2001 | 2002 | 2003 | 2004 | 2005 |
| <i>Ln(proximidade Tecnológica)</i> | 0,020*** (0,000) | 0,025*** (0,000) | 0,024*** (0,000) | 0,023*** (0,000) | 0,020*** (0,000) |
| <i>Ln(distância geográfica)</i> | -0,017*** (0,000) | -0,019*** (0,000) | -0,016*** (0,000) | -0,017*** (0,000) | -0,014*** (0,000) |
| <i>Dummy de contiguidade</i> | 0,111*** (0,000) | 0,137*** (0,000) | 0,129*** (0,000) | 0,113*** (0,000) | 0,102*** (0,000) |
| <i>Dummy de UF</i> | 0,018*** (0,000) | 0,020*** (0,000) | 0,031*** (0,000) | 0,028*** (0,000) | 0,027*** (0,000) |
| <i>Constante</i> | 0,072*** (0,000) | 0,084*** (0,000) | 0,070*** (0,000) | 0,107*** (0,000) | 0,095*** (0,000) |
| <i>Efeito Fixo da Origem</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> |
| <i>Efeito Fixo do Destino</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> |
| <i>Número de observações</i> | 29876 | 30110 | 30284 | 30754 | 35574 |
| <i>Log-likelihood</i> | 19011,56 | 16557,03 | 14083,61 | 13401,06 | 19286,11 |
| <i>aic</i> | -37511,12 | -32578,07 | -27657,22 | -26292,13 | -38038,22 |
| <i>bic</i> | -35385,09 | -30350,29 | -25536,04 | -24167,01 | -35774,23 |
| <i>r2</i> | 0,19 | 0,18 | 0,19 | 0,19 | 0,16 |
| <i>r2_a</i> | 0,17 | 0,17 | 0,17 | 0,18 | 0,15 |

(Continua)

| Variável dependente: <i>ln(mobilidade + 1)</i> | Anos | | | | |
|---|----------------------|----------------------|----------------------|----------------------|----------------------|
| | 2006 | 2007 | 2008 | 2009 | 2010 |
| <i>Ln(proximidade Tecnológica)</i> | 0,022*** (0,000) | 0,027*** (0,000) | 0,024*** (0,000) | 0,026*** (0,000) | 0,026*** (0,000) |
| <i>Ln(distância geográfica)</i> | -0,018*** (0,000) | -0,021*** (0,000) | -0,028*** (0,000) | -0,034*** (0,000) | -0,034*** (0,000) |
| <i>Dummy de contiguidade</i> | 0,107*** (0,000) | 0,128*** (0,000) | 0,162*** (0,000) | 0,135*** (0,000) | 0,126*** (0,000) |
| <i>Dummy de UF</i> | 0,037*** (0,000) | 0,045*** (0,000) | 0,042*** (0,000) | 0,033*** (0,000) | 0,028*** (0,000) |
| <i>Constante</i> | 0,068*** (0,000) | 0,094*** (0,000) | 0,106*** (0,000) | 0,124*** (0,000) | 0,163*** (0,000) |
| <i>Efeito Fixo da Origem</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> |
| <i>Efeito Fixo do Destino</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> | <i>sim</i> |
| <i>Número de observações</i> | 34694 | 35022 | 34220 | 36188 | 38156 |
| <i>Log-likelihood</i> | 16489,67 | 15134,90 | 13425,39 | 14415,97 | 14453,79 |
| <i>aic</i> | -32451,34 | -29743,81 | -26348,78 | -28315,94 | -28355,59 |
| <i>bic</i> | -30219,40 | -27517,85 | -24230,20 | -26123,85 | -25995,94 |
| <i>r2</i> | 0,18 | 0,19 | 0,21 | 0,21 | 0,20 |
| <i>r2_a</i> | 0,17 | 0,18 | 0,20 | 0,20 | 0,19 |

Fonte: Elaboração própria com o uso do *software* Stata 12.0
p-valor entre parênteses ****p* < 0,01 ***p* < 0,05 **p* < 0,1

Tabela A7: Função de Produção do Conhecimento: estimações por *System GMM* – 2001 a 2010¹ – lags 2 a 5

| Variável dependente: ln (Patentes por 100 mil habitantes) | (1) | (2) | (3) |
|---|---------------------|---------------------|---------------------|
| $\ln Patentes_{t-1}$ | 0,127*** (0,001) | 0,131*** (0,001) | 0,137*** (0,001) |
| $\ln Patentes_{t-2}$ | 0,060* (0,075) | 0,062* (0,084) | 0,063* (0,072) |
| $W^c \ln Patentes$ | 0,355*** (0,000) | 0,320*** (0,000) | 0,310*** (0,000) |
| $W^m \ln Patentes$ | 0,079 (0,399) | 0,036 (0,748) | 0,046 (0,633) |
| $\ln PoTec$ | | -0,002 (0,866) | 0,007 (0,444) |
| $\ln PeDun$ | | 0,032* (0,062) | 0,031* (0,089) |
| $\ln CapHumano$ | | 0,205*** (0,004) | 0,145*** (0,009) |
| $\ln EscUrbana$ | | | 0,581** (0,030) |
| $\ln EscUrbanaQd$ | | | 0,032** (0,027) |
| $\ln Transformação$ | | | 0,033 (0,467) |
| $\ln EspProdutiva$ | | | -0,070 (0,415) |
| Constante | 0,055 (0,329) | 0,386** (0,028) | 2,915*** (0,023) |

(Continua)

| Variável dependente: ln (Patentes por 100 mil habitantes) | (1) | (2) | (3) |
|---|----------|----------|---------------------|
| <i>Dummy</i> de ano | Sim | Sim | Sim |
| Microrregiões/Observações | 558/4464 | 558/4464 | 558/4464 |
| <i>Lags</i> utilizados como instrumentos ² | 2 a 5 | 2 a 5 | 2 a 5 |
| Nº de instrumentos | 23 | 34 | 46 |
| Teste F | (0,000) | (0,000) | (0,000) |
| AR (1) | (0,000) | (0,000) | (0,000) |
| AR (2) | (0,480) | (0,375) | (0,465) |
| Teste <i>Hansen</i> | (0,094) | (0,149) | (0,136) |
| <i>Difference-in-Hansen</i> : Ho: exogeneidade de subgrupos de instrumentos | | | |
| Instrumentos para a equação em diferenças ³ | (0,149) | (0,099) | (0,057) |
| Instrumentos para a equação em nível ⁴ | (0,142) | (0,511) | (0,753) |
| $\Delta \ln \text{Patentes}_{t-1}$ para a equação em nível | (0,760) | (0,533) | (0,455) |
| $\Delta W^c \ln \text{Patentes}_{t-1}$ e $\Delta W^m \ln \text{Patentes}_{t-1}$ para a equação em nível | (0,065) | (0,065) | (0,038) |
| | | | (<i>Continua</i>) |

| Variável dependente: ln (Patentes por 100 mil habitantes) | (1) | (2) | (3) |
|---|-------------------|-------------------|-------------------|
| $\Delta \ln \text{Controles}_{t-1}$ para a equação em nível | | (0,281) | (0,205) |
| Estimativa de $\ln \text{Patentes}_{t-1}$ por MQO | 0,415 (0,000) | 0,413 (0,000) | 0,404 (0,000) |
| Estimativa de $\ln \text{Patentes}_{t-2}$ por MQO | 0,337 (0,000) | 0,333 (0,000) | 0,327 (0,000) |
| Estimativa de $\ln \text{Patentes}_{t-1}$ por EF | -0,024 (0,367) | -0,025 (0,344) | -0,025 (0,348) |
| Estimativa de $\ln \text{Patentes}_{t-2}$ por EF | -0,077 (0,001) | -0,079 (0,001) | -0,079 (0,001) |

Fonte: Elaboração própria com o uso do software Stata 12.0.

p -valor entre parênteses *** $p < 0,01$ ** $p < 0,05$ * $p < 0,1$

¹ Foram consideradas como variáveis endógenas: $\ln \text{Patentes}_{t-1}$, $W^c \ln \text{Patentes}$, $W^m \ln \text{Patentes}$, $\ln \text{PoTec}$, $\ln \text{PeDun}$, $\ln \text{Transformação}$ e $\ln \text{EspProdutiva}$. Foram consideradas como variáveis exógenas: $\ln \text{CapHumano}$, $\ln \text{EscUrbana}$ e $\ln \text{EscUrbanaQd}$.

² Foram utilizados do segundo ao quinto lags das variáveis endógenas como instrumentos para a equação em diferenças e o primeiro lag da diferença para a equação em nível.

³ Instrumentos para a equação em diferenças: $\ln \text{Patentes}_{t-2}$, $\ln \text{Patentes}_{t-3}$, $\ln \text{Patentes}_{t-4}$, $\ln \text{Patentes}_{t-5}$, $W^c \ln \text{Patentes}_{t-2}$, $W^c \ln \text{Patentes}_{t-3}$, $W^c \ln \text{Patentes}_{t-4}$, $W^c \ln \text{Patentes}_{t-5}$, $W^m \ln \text{Patentes}_{t-2}$, $W^m \ln \text{Patentes}_{t-3}$, $W^m \ln \text{Patentes}_{t-4}$, $W^m \ln \text{Patentes}_{t-5}$, $\ln \text{PoTec}_{t-2}$, $\ln \text{PoTec}_{t-3}$, $\ln \text{PoTec}_{t-4}$, $\ln \text{PoTec}_{t-5}$, $\ln \text{PeDun}_{t-2}$, $\ln \text{PeDun}_{t-3}$, $\ln \text{PeDun}_{t-4}$, $\ln \text{PeDun}_{t-5}$, $\ln \text{Transformação}_{t-2}$, $\ln \text{Transformação}_{t-3}$, $\ln \text{Transformação}_{t-4}$, $\ln \text{Transformação}_{t-5}$, $\ln \text{EspProdutiva}_{t-2}$, $\ln \text{EspProdutiva}_{t-3}$, $\ln \text{EspProdutiva}_{t-4}$, $\ln \text{EspProdutiva}_{t-5}$.

⁴ Instrumentos para a equação em nível: $\Delta \ln \text{Patentes}_{t-1}$, $\Delta W^c \ln \text{Patentes}_{t-1}$, $\Delta W^m \ln \text{Patentes}_{t-1}$, $\Delta \ln \text{PoTec}_{t-1}$, $\Delta \ln \text{PeDun}_{t-1}$, $\Delta \ln \text{Transformação}_{t-1}$, e $\Delta \ln \text{EspProdutiva}_{t-1}$.

Tabela A8: Função de produção do conhecimento com defasagem espacial: estimações por *System GMM* (2001 a 2010)¹ – lags de 2 a 5

| Var. dependente: ln (Patentes por 100 mil habitantes) | (1) | (2) | (3) | (4) |
|--|---------------------------------|--|---------------------------|-------------------------|
| | W^e : Dist, Inv, Quadr, | W^e : Dist, Inv, Quadr, Ord | W^e : 15 vizinhos | W^e : <i>Queen</i> |
| $lnPatentes_{t-1}$ | 0,121*** (0,002) | 0,116*** (0,002) | 0,121*** (0,001) | 0,113*** (0,003) |
| $lnPatentes_{t-2}$ | 0,043 (0,212) | 0,040 (0,248) | 0,028 (0,428) | 0,036 (0,299) |
| $W^c lnPatentes$ | 0,256*** (0,000) | 0,251*** (0,000) | 0,216*** (0,003) | 0,244*** (0,001) |
| $W^m lnPatentes$ | -0,023 (0,814) | -0,038 (0,702) | -0,062 (0,542) | -0,062 (0,528) |
| $W^e lnPatentes$ | 0,549*** (0,008) | 0,605*** (0,003) | 0,765*** (0,000) | 0,533*** (0,000) |
| $lnPoTec$ | 0,003 (0,750) | 0,001 (0,905) | 0,000 (0,999) | 0,004 (0,683) |
| $lnPeDun$ | 0,035** (0,021) | 0,034** (0,024) | 0,035** (0,011) | 0,035** (0,012) |
| $lnCapHumano$ | 0,083* (0,075) | 0,082* (0,085) | 0,036 (0,358) | 0,093** (0,050) |
| $lnEscUrbana$ | 0,821*** (0,001) | 0,799*** (0,001) | 0,877*** (0,000) | 0,753*** (0,006) |
| $lnEscUrbanaQd$ | 0,041*** (0,002) | 0,040*** (0,002) | 0,044*** (0,001) | 0,038** (0,013) |

(Continua)

| Var. dependente: ln (Patentes por 100 mil habitantes) | (1) | (2) | (3) | (4) |
|--|--|---|--|--|
| | <i>W^e</i> : Dist, Inv, Quadr, | <i>W^e</i> : Dist, Inv, Quadr, Ord | <i>W^e</i> : 15 vizinhos | <i>W^e</i> : <i>Queen</i> |
| <i>lnTransformação</i> | 0,008 (0,811) | 0,004 (0,901) | -0,008 (0,810) | 0,016 (0,625) |
| <i>lnEspProdutiva</i> | -0,037 (0,641) | -0,036 (0,652) | -0,055 (0,490) | -0,067 (0,415) |
| Constante | 4,138*** (0,001) | 4,057*** (0,000) | 4,333*** (0,000) | 3,886*** (0,002) |
| <i>Dummy</i> de ano | Sim | Sim | Sim | Sim |
| Microrregiões/Observações | 558/4464 | 558/4464 | 558/4464 | 558/4464 |
| <i>Lags</i> utilizados como instrumentos ² | 2 a 5 | 2 a 5 | 2 a 5 | 2 a 5 |
| Nº de instrumentos | 51 | 51 | 51 | 51 |
| Teste F | (0,000) | (0,000) | (0,000) | (0,000) |
| AR (1) | (0,000) | (0,000) | (0,000) | (0,000) |
| AR (2) | (0,439) | (0,417) | (0,674) | (0,227) |
| Teste Hansen | (0,247) | (0,235) | (0,429) | (0,260) |

(Continua)

| Var. dependente: ln (Patentes por 100 mil habitantes) | (1) | (2) | (3) | (4) |
|---|-------------------------------------|--|---------------------------|-------------------------|
| | W^e : Dist. Inv. Quadr . | W^e : Dist. Inv. Quadr. Ord | W^e : 15 vizinhos | W^e : <i>Queen</i> |
| <i>Difference-in-Hansen</i> : Ho: exogeneidade de subgrupos de instrumentos | | | | |
| Instrumentos para a equação em diferenças ³ | (0,146) | (0,138) | (0,274) | (0,162) |
| Instrumentos para a equação em nível ⁴ | (0,665) | (0,660) | (0,741) | (0,645) |
| $\Delta \ln Patentes_{t-1}$ para a equação em nível | (0,885) | (0,900) | (0,987) | (0,754) |
| $\Delta W^c \ln Patentes_{t-1}$, $\Delta W^m \ln Patentes_{t-1}$ e $\Delta W^e \ln Patentes_{t-1}$ para a equação em nível | (0,040) | (0,038) | (0,089) | (0,033) |
| $\Delta \ln Controles_{t-1}$ para a equação em nível | (0,327) | (0,312) | (0,692) | (0,346) |
| Estimativa de $\ln Patentes_{t-1}$ por MQO | 0,396*** (0,000) | 0,393*** (0,000) | 0,391*** (0,000) | 0,391*** (0,000) |
| Estimativa de $\ln Patentes_{t-2}$ por MQO | 0,316*** (0,000) | 0,312*** (0,000) | 0,311*** (0,000) | 0,311*** (0,000) |
| Estimativa de $\ln Patentes_{t-1}$ por EF | -0,025 (0,347) | -0,025 (0,348) | -0,026 (0,328) | -0,026 (0,331) |
| Estimativa de $\ln Patentes_{t-2}$ por EF | -0,081*** (0,001) | -0,081*** (0,001) | -0,081*** (0,001) | -0,081*** (0,001) |

Fonte: Elaboração própria com o uso do *software* Stata 12.0.

¹ Foram consideradas como variáveis endógenas: $\ln Patentes_{t-1}$, $W^c \ln Patentes$, $W^m \ln Patentes$, $W^e \ln Patentes$, $\ln PoTec$, $\ln PeDun$, $\ln Transformação$ e $\ln EspProdutiva$.

Foram consideradas como variáveis exógenas: $\ln CapHumano$, $\ln EscUrbana$ e $\ln EscUrbanaQd$.

² Foram utilizados do segundo ao quinto *lags* das variáveis endógenas como instrumentos para a equação em diferenças e o primeiro *lag* da diferença para a equação em nível.

³ Instrumentos para a equação em diferenças: $\ln Patentes_{t-2}$, $\ln Patentes_{t-3}$, $\ln Patentes_{t-4}$, $\ln Patentes_{t-5}$, $W^c \ln Patentes_{t-2}$, $W^c \ln Patentes_{t-3}$, $W^c \ln Patentes_{t-4}$, $W^c \ln Patentes_{t-5}$, $W^m \ln Patentes_{t-2}$, $W^m \ln Patentes_{t-3}$, $W^m \ln Patentes_{t-4}$, $W^m \ln Patentes_{t-5}$, $\ln PoTec_{t-2}$, $\ln PoTec_{t-3}$, $\ln PoTec_{t-4}$,

$\ln PoTec_{t-5}$, $\ln PeDun_{t-2}$, $\ln PeDun_{t-3}$, $\ln PeDun_{t-4}$, $\ln PeDun_{t-5}$, $\ln Transformação_{t-2}$, $\ln Transformação_{t-3}$, $\ln Transformação_{t-4}$, $\ln Transformação_{t-5}$, $\ln EspProdutiva_{t-2}$, $\ln EspProdutiva_{t-3}$, $\ln EspProdutiva_{t-4}$, $\ln EspProdutiva_{t-5}$.

⁴ Instrumentos para a equação em nível: $\Delta \ln Patentes_{t-1}$, $\Delta W^c \ln Patentes_{t-1}$, $\Delta W^m \ln Patentes_{t-1}$, $\Delta W^e \ln Patentes_{t-1}$, $\Delta \ln PoTec_{t-1}$, $\Delta \ln PeDun_{t-1}$, $\Delta \ln Transformação_{t-1}$ e $\Delta \ln EspProdutiva_{t-1}$.

Tabela A9: Função de produção do conhecimento para regiões patenteadoras: estimações por *System GMM* (2001 a 2010)¹ – lags (2 a 5)

| Var. dependente: ln (Patentes por 100 mil habitantes) | (1) | (2) | (3) | (4) | (5) |
|--|--------------------------------|---------------------------------|------------------------------------|---------------------------|-------------------------|
| | W^e : Sem lag espacial | W^e : Dist. Inv. Quadr. | W^e : Dist. Inv. Quadr.Ord | W^e : 15 vizinhos | W^e : <i>Queen</i> |
| $lnPatentes_{t-1}$ | 0,134*** (0,001) | 0,125*** (0,001) | 0,121*** (0,001) | 0,122*** (0,001) | 0,112*** (0,004) |
| $lnPatentes_{t-2}$ | 0,052 (0,147) | 0,041 (0,239) | 0,039 (0,273) | 0,026 (0,473) | 0,028 (0,435) |
| $W^c lnPatentes$ | 0,303*** (0,000) | 0,250*** (0,001) | 0,243*** (0,001) | 0,211*** (0,004) | 0,225*** (0,002) |
| $W^m lnPatentes$ | -0,038 (0,720) | -0,075 (0,468) | -0,093 (0,369) | -0,098 (0,342) | -0,126 (0,221) |
| $W^e lnPatentes$ | - | 0,596** (0,019) | 0,652*** (0,008) | 0,834*** (0,000) | 0,612*** (0,001) |
| $lnPoTec$ | -0,016 (0,607) | -0,027 (0,331) | -0,031 (0,273) | -0,023 (0,398) | -0,028 (0,328) |
| $lnPeDun$ | 0,031* (0,088) | 0,035** (0,021) | 0,035** (0,023) | 0,033** (0,014) | 0,034** (0,014) |
| $lnCapHumano$ | 0,309*** (0,003) | 0,208** (0,017) | 0,213** (0,015) | 0,132* (0,084) | 0,230** (0,012) |
| $lnEscUrbana$ | 0,851** (0,016) | 0,878*** (0,006) | 0,807** (0,012) | 0,862*** (0,006) | 0,768** (0,016) |
| $lnEscUrbanaQd$ | 0,048** (0,014) | 0,045** (0,015) | 0,040** (0,030) | 0,043** (0,019) | 0,038** (0,036) |

(Continua)

| Var. dependente: ln (Patentes por 100 mil habitantes) | (1) | (2) | (3) | (4) | (5) |
|--|---------------------|--|---|------------------------------------|---------------------------|
| | Sem lag espacial | W ^e : Dist. Inv. Quadr. | W ^e : Dist. Inv. Quadr.Ord | W ^e : 15 vizinhos | W ^e : Queen |
| <i>lnTransformação</i> | 0,097 (0,445) | 0,023 (0,841) | 0,014 (0,897) | -0,012 (0,904) | 0,003 (0,973) |
| <i>lnEspProdutiva</i> | -0,058 (0,663) | -0,038 (0,749) | -0,032 (0,785) | -0,067 (0,574) | -0,079 (0,518) |
| Constante | 4,098*** (0,008) | 4,313*** (0,002) | 4,047*** (0,004) | 4,176*** (0,002) | 3,887*** (0,005) |
| <i>Dummy</i> de ano | Sim | Sim | Sim | Sim | Sim |
| Microrregiões/Observações | 428/3424 | 428/3424 | 428/3424 | 428/3424 | 428/3424 |
| <i>Lags</i> utilizados como instrumentos ² | 2 a 5 | 2 a 5 | 2 a 5 | 2 a 5 | 2 a 5 |
| Nº de Instrumentos | 46 | 51 | 51 | 51 | 51 |
| Teste F | (0,000) | (0,000) | (0,000) | (0,000) | (0,000) |
| AR (1) | (0,000) | (0,000) | (0,000) | (0,000) | (0,000) |
| AR (2) | (0,477) | (0,372) | (0,337) | (0,641) | (0,179) |

(Continua)

| Var. dependente: ln (Patentes por 100 mil habitantes) | (1) | (2) | (3) | (4) | (5) |
|--|----------------------|--|---|------------------------------------|---------------------------|
| | Sem lag espacial | W ^e : Dist. Inv. Quadr. | W ^e : Dist. Inv. Quadr.Ord | W ^e : 15 vizinhos | W ^e : Queen |
| Teste <i>Hansen</i> | (0,260) | (0,321) | (0,318) | (0,525) | (0,362) |
| <i>Difference-in-Hansen</i> : (Ho: exogeneidade de subgrupos de instrumentos) | | | | | |
| Instrumentos para a equação em diferenças ³ | (0,143) | (0,250) | (0,256) | (0,415) | (0,300) |
| Instrumentos para a equação em nível ⁴ | (0,715) | (0,541) | (0,521) | (0,642) | (0,521) |
| $\Delta \ln \text{Patentes}_{t-1}$ para a equação em nível | (0,558) | (0,893) | (0,903) | (0,999) | (0,889) |
| $\Delta W^c \ln \text{Patentes}_{t-1}$, $\Delta W^m \ln \text{Patentes}_{t-1}$ e $\Delta W^e \ln \text{Patentes}_{t-1}$ para a equação em nível | (0,045) ⁵ | (0,060) | (0,059) | (0,131) | (0,057) |
| $\Delta \ln \text{Controles}_{t-1}$ para a equação em nível | (0,517) | (0,666) | (0,642) | (0,761) | (0,625) |
| Estimativa de $\ln \text{Patentes}_{t-1}$ por MQO | 0,394*** (0,000) | 0,388*** (0,000) | 0,385*** (0,000) | 0,383*** (0,000) | 0,382*** (0,000) |
| Estimativa de $\ln \text{Patentes}_{t-2}$ por MQO | 0,317*** (0,000) | 0,308*** (0,000) | 0,305*** (0,000) | 0,304*** (0,000) | 0,303*** (0,000) |
| Estimativa de $\ln \text{Patentes}_{t-1}$ por EF | -0,025 (0,347) | -0,025 (0,347) | -0,025 (0,348) | -0,026 (0,327) | -0,026 (0,328) |
| Estimativa de $\ln \text{Patentes}_{t-2}$ por EF | -0,079*** (0,001) | -0,081*** (0,001) | -0,081*** (0,001) | -0,081*** (0,001) | -0,081*** (0,001) |

Fonte: Elaboração própria com o uso do software Stata 12.0.

¹ Foram consideradas como variáveis endógenas: $\ln \text{Pat}_{t-1}$, $W^c \ln \text{Pat}$, $W^m \ln \text{Pat}$, $W^e \ln \text{Pat}$, $\ln \text{PoTec}$, $\ln \text{PeDun}$, $\ln \text{Transf}$ e $\ln \text{EspProd}$. Foram consideradas como variáveis exógenas: $\ln \text{CH}$, $\ln \text{EscUrb}$ e $\ln \text{EscUrbQd}$.

² Foram utilizados do segundo ao quinto lags das variáveis endógenas como instrumentos para a equação em diferenças e o primeiro lag da diferença para a equação em nível.

³ Instrumentos para a equação em diferenças: $\ln Pat_{t-2}$, $\ln Pat_{t-3}$, $\ln Pat_{t-4}$, $\ln Pat_{t-5}$, $W^c \ln Pat_{t-2}$, $W^c \ln Pat_{t-3}$, $W^c \ln Pat_{t-4}$, $W^c \ln Pat_{t-5}$, $W^m \ln Pat_{t-2}$, $W^m \ln Pat_{t-3}$, $W^m \ln Pat_{t-4}$, $W^m \ln Pat_{t-5}$, $\ln PoTec_{t-2}$, $\ln PoTec_{t-3}$, $\ln PoTec_{t-4}$, $\ln PoTec_{t-5}$, $\ln PeDun_{t-2}$, $\ln PeDun_{t-3}$, $\ln PeDun_{t-4}$, $\ln PeDun_{t-5}$, $\ln Transf_{t-2}$, $\ln Transf_{t-3}$, $\ln Transf_{t-4}$, $\ln Transf_{t-5}$, $\ln EspProd_{t-2}$, $\ln EspProd_{t-3}$, $\ln EspProd_{t-4}$, $\ln EspProd_{t-5}$

⁴ Instrumentos para a equação em nível: $\Delta \ln Pat_{t-1}$, $\Delta W^c \ln Pat_{t-1}$, $\Delta W^m \ln Pat_{t-1}$, $\Delta W^e \ln Pat_{t-1}$, $\Delta \ln PoTec_{t-1}$, $\Delta \ln PeDun_{t-1}$, $\Delta \ln Transf_{t-1}$ e $\Delta \ln EspProd_{t-1}$.

⁵ Refere-se ao *p*-valor do teste *Difference in Hansen* para o subgrupo de instrumentos composto por $\Delta W^c \ln Pat_{t-1}$ e $\Delta W^m \ln Pat_{t-1}$, pois esta especificação não inclui defasagem espacial.