



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



RAQUEL COELHO REIS

**REDES DE COPATENTEAMENTO NO BRASIL: DETERMINANTES REGIONAIS
E ESTRUTURAIS DAS LIGAÇÕES NACIONAIS E INTERNACIONAIS**

JUIZ DE FORA
2017

RAQUEL COELHO REIS

**REDES DE COPATENTEAMENTO NO BRASIL: DETERMINANTES REGIONAIS
E ESTRUTURAIS DAS LIGAÇÕES NACIONAIS E INTERNACIONAIS**

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Faculdade de Economia da Universidade Federal de Juiz de Fora como parte das exigências para obtenção do título de Mestre, com área de concentração em Economia.

Orientador: Prof. Dr. Eduardo Gonçalves
Coorientadora: Prof.^a Dra. Juliana G. Taveira

JUIZ DE FORA
2017

RAQUEL COELHO REIS

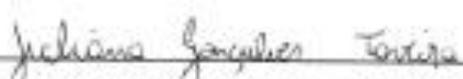
**REDES DE COPATENTEAMENTO NO BRASIL: DETERMINANTES REGIONAIS
E ESTRUTURAIS DAS LIGAÇÕES NACIONAIS E INTERNACIONAIS**

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Faculdade de Economia da Universidade Federal de Juiz de Fora como parte das exigências para obtenção do título de Mestre em Economia Aplicada.
Área de Concentração: Economia.

Aprovada em: 20 de Janeiro de 2017



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



Prof.ª Dra. Juliana Gonçalves Taveira - Coorientadora
Universidade Federal de Juiz de Fora



Prof.º Dr. Veneziano de Castro Araújo - Convidado
Universidade Federal de São Paulo



Prof.º Dr. Eduardo Simões de Almeida - Convidado
Universidade Federal de Juiz de Fora

AGRADECIMENTOS

Primeiramente, agradeço ao apoio incondicional da minha família durante toda a etapa do meu mestrado. Aos meus pais, Mozar e Wanilda, e ao meu irmão Kleber, muito obrigada pelo apoio, confiança e motivação nos momentos difíceis. Sem o apoio de vocês eu jamais teria conseguido chegar ao fim desta etapa. Vocês sempre foram e serão meu porto seguro.

Ao Prof. Dr. Eduardo Gonçalves, que aceitou essa tarefa de ser meu orientador, agradeço imensamente a paciência, os incentivos, correções e a atenção depositada. Principalmente, obrigada aos conselhos e toda a dedicação na orientação não só deste trabalho, mas de toda a minha vida acadêmica. À Prof.^a Dr.^a Juliana G. Taveira, obrigada pelo suporte e atenção para a realização deste trabalho. Por ter aceitado esse desafio de me coorientar e por, desde as primeiras conversas, se mostrar paciente e disposta a colaborar para a melhor conclusão desta pesquisa.

Ao Prof. Dr. Eduardo Almeida, por sua dedicação e disposição em fiscalizar toda a evolução do meu trabalho e pelas contribuições de grande valia durante e fora das reuniões. Ao Prof. Dr. Admir Betarelli, muito obrigada pelas contribuições em prol de melhoras não somente desta pesquisa, mas por possibilitar uma evolução das que ainda estão por vir.

Agradeço aos meus amigos, que me acompanharam de perto ou de longe toda esta jornada. Principalmente, aos companheiros discentes do PPGE/UFJF que estiveram comigo dia a dia e colaboraram de alguma forma para o sucesso e conclusão desta etapa do mestrado. Sem vocês, alguns momentos teriam sido mais difíceis de suportar. Agradeço aos amigos de sempre, que mesmo longe fisicamente nunca deixaram de estar presentes. Por fim, muito obrigada a todo o corpo docente e funcionários do PPGE/UFJF que, de alguma maneira, contribuíram para a conclusão desta etapa da minha vida tão importante.

Um agradecimento especial eu estendo à equipe do INPI, responsável pela BADEPI, sem a qual este trabalho não teria sido possível.

RESUMO

O estudo tem como objetivo investigar os determinantes regionais e estruturais da ocorrência de ligações entre regiões nas redes sociais de colaborações no Brasil. Para isto, utiliza-se de dados de copatenteamento das produções tecnológicas, realizadas tanto por inventores brasileiros como destes em parceria com estrangeiros. Como possíveis determinantes das ligações nas redes de copatenteamento regional consideram-se tanto os aspectos econômicos, tecnológicos e de densidade demográfica das regiões quanto as próprias estruturas de topologia dos nós nas redes. Os dados para a pesquisa são fornecidos pelo Instituto Nacional de Propriedade Industrial (INPI), Relação Anual de Informações Sociais (RAIS), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e Instituto Brasileiro de Geografia e Estatística (IBGE). Sendo assim, foi possível a construção de uma base de dados longitudinal para o período de 2001-2011 com abrangência geográfica mais desagregada de REGIC imediata, e investigação por meio de técnicas econométricas de dados em painel. A fim de atingir o objetivo proposto identificam-se duas redes no Brasil: rede inter-regional e internacional. Portanto, para ambas as redes aplicam-se duas abordagens. Primeiramente, utiliza-se um modelo voltado para dados truncados, Tobit, cuja variável dependente é dada pela razão entre o número de ligações e o total de patentes geradas pela região. E posteriormente, aplica-se um modelo para dados de contagem, Binomial Negativo Inflado em Zeros (ZINB), que considera o número absoluto de ligações de cada região nas redes. Como principais resultados evidencia-se que as ligações dos nós (regiões) são mais impactadas por aspectos da infraestrutura local, como nível de P&D, e por sua estrutura de posição na rede. No entanto, outras características locais e estruturais também merecem ser destacadas como a participação econômica das regiões, assim como os níveis de participação da indústria.

Palavras-chave: Redes de inovação. Copatentes. Tobit. ZINB. REGIC.

ABSTRACT

The study aims to investigate the regional and structural determinants of the links among regions in social networks of collaboration in Brazil. For this, we use patent database that possess information on Brazilian inventors' residence address and on the nationality of all inventors responsible for the patent application in Brazilian Patent Office. We investigate which are the main regional determinants of the links in the Brazilian co-patenting networks taking into account variables that capture economic, technological and demographic density aspects of Brazilian regions, as well as the topology structures of the nodes in the networks. The database is provided by the National Institute of Industrial Property (INPI), the Annual Social Information Relation (RAIS), the Coordination for the Improvement of Higher Education Personnel (CAPES) and the Brazilian Institute of Geography and Statistics (IBGE). We consider a longitudinal database for the period 2001-2011 and apply panel data techniques in spatial unit of analysis corresponding to the Intermediate Areas of Urban Connectivity, according to the concept of Regions of Influence of Brazilian Cities (REGICs). Two types of networks are considered in Brazil: the inter-regional and the international ones. Therefore, for both types of networks two approaches are applied: 1) a truncated data model is used, i.e. Tobit, whose dependent variable is given by the ratio of the number of links to the number of patents generated by the region; and 2) a Zero-inflated negative binomial (ZINB) model is used, which considers the absolute number of inter-regional links among inventors in a certain region. The main results show that node (region) connections are more affected by aspects of local infrastructure, such as R&D level, and their position in the network. However, other local and structural characteristics also deserve to be highlighted as the economic participation of regions, as well as levels of capital goods imports and industry participation.

Keywords: Innovation networks. Co-patenting. Tobit. ZINB. REGIC.

LISTA DE SIGLAS

AEDE - Análise Exploratória de Dados Espaciais
ARS - Análise de Redes Sociais
CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CBO - Classificação Brasileira de Ocupações
FAPEMIG - Fundação de Amparo a Pesquisa de Minas Gerais
IBGE - Instituto Brasileiro de Geografia e Estatística
INPI - Instituto Nacional de Propriedade Intelectual
LISA - *Local Indicator of Spatial Association*
MCMC - *Markovian-Chain Monte Carlo*
MEC - Ministério da Educação
MLE - *Maximum Likelihood Estimation*
MQO - Mínimos Quadrados Ordinários
OECD - *Organization for Economic Cooperation and Development*
P&D - Pesquisa e Desenvolvimento
PIB - Produto Interno Bruto
PINTEC - Pesquisa Industrial de Inovação Tecnológica
POTEC - Pessoal Ocupado Técnico Científico
RAIS - Relação Anual de Informações Sociais
REGIC - Região Geográfica de Influência Econômica
SECEX - Secretaria de Comércio Exterior
ZIP - *Zero-Inflated Poisson*
ZINB - *Zero-Inflated Negative Binomial*

LISTA DE FIGURAS

Figura 1 – Ligações de redes: não orientadas e orientadas.....	29
Figura 2 – Representação transformação rede two-mode em uma rede one-mode.....	30
Figura 3 – Representação do caminho mínimo entre dois nós na rede.....	33
Figura 4 – Construção da rede de copatenteamento regional.....	36
Figura 5 - Gráficos das distribuições das variáveis ligações por patentes (LigaBBpat e LigaBEpat). Período de análise 2001-2011.....	54
Figura 6 - Análise de Cluster LISA – Ligações por patentes (2001-2011).....	66
Figura 7 - Análise de Cluster LISA – Ligações absolutas (2001-2011).....	67
Figura 8 – Mapa das ligações das redes inter-regionais e internacionais (2001-2011).....	72

LISTA DE QUADROS

Quadro 1 - Resumo de trabalhos que utilizam dados de copatentes para investigar redes.....	27
Quadro 2 - Definição das redes de copatenteamento identificadas no Brasil.....	35
Quadro 3 - Descrição das variáveis de interesse do modelo.....	51
Quadro 4 – <i>Ranking</i> 15 regiões com maiores coeficientes de centralidades (2000-2010).....	73

LISTA DE TABELAS

Tabela 1 - Análise descritiva das variáveis dependentes ligações por patentes (LigaBBpat e LigaBEpat). Período de análise 2001-2011.....	54
Tabela 2 – Análises descritivas das variáveis de ligações absolutas (LigaBB e LigaBE). Período de análise 2001-2011.....	55
Tabela 3 – Resultados do Teste <i>Z-Vuong</i> para modelos inflados em zero. Período de análise 2001-2011.....	62
Tabela 4 – Estatísticas <i>I de Moran</i> para variadas matrizes de ponderação (W). Para cada variável dependente das redes e período de análise 2001-2011.....	64
Tabela 5 - Análise descritiva das variáveis independentes (2001-2011).....	68
Tabela 6 - <i>Ranking</i> das 15 regiões brasileiras com mais ligações por patentes (2001-2011).....	70
Tabela 7 - <i>Ranking</i> das 15 regiões brasileiras com maior número de ligações absolutas (2001-2011).....	71
Tabela 8 - Determinantes das ligações inter-regionais por patentes (LigaBBpat) no Brasil. Painel Tobit com Efeitos Aleatórios (2001-2011).....	75
Tabela 9 - Determinantes das ligações absolutas inter-regionais (LigaBB) no Brasil. Painel <i>Pooled ZINB</i> robusto (2001-2011).....	78
Tabela 10 - Determinantes das ligações internacionais por patentes (LigaBEpat) no Brasil. Painel Tobit com Efeitos Aleatórios (2001-2011).....	81
Tabela 11: Determinantes das ligações absolutas internacionais (LigaBE) no Brasil. Painel <i>Pooled ZINB</i> robusto (2001-2011).....	83

SUMÁRIO

1 INTRODUÇÃO	10
2 REVISÃO DE LITERATURA.....	14
2.1 Inovações: definições e formas de mensuração.....	14
2.2 O papel das inovações no crescimento econômico.....	16
2.3 Análise de redes sociais (ARS).....	20
2.3.1 Redes sociais de inovação	21
2.3.1.1 Arquiteturas de redes sociais	28
3 METODOLOGIA.....	34
3.1 Discussão de hipóteses.....	37
3.2 Base de dados.....	39
3.3 Variáveis do modelo	43
3.3.1 Variável dependente: ligações por patentes e total de ligações.....	43
3.3.2 Variáveis de interesse	44
3.3.3 Variáveis de controle.....	50
3.4 Modelos de estimação.....	52
3.4.1 Modelo Tobit.....	55
3.4.2 Modelo binomial negativo inflado em zero (ZINB).....	59
4 RESULTADOS.....	63
4.1 AEDE.....	64
4.2 Estatísticas das variáveis de interesse.....	68
4.3 Resultados das regressões da rede inter-regional.....	74
4.4 Resultados das regressões da rede internacional.....	80
5 CONSIDERAÇÕES FINAIS.....	84
REFERÊNCIAS BIBLIOGRÁFICAS	87
ANEXO.....	100
APÊNDICE	101

1 INTRODUÇÃO

Desde estudos realizados por Schumpeter (1982), as inovações tem se inserido em discussões econômicas. Isto ocorre devido aos possíveis impactos em termos de crescimento e desenvolvimento de instituições e regiões que são causados pelas atividades inovativas. Essas, todavia, tendem a se concentrar geograficamente. Assim, considerar os aspectos e estruturas regionais de onde estas ocorrem, pode auxiliar na compreensão da dinâmica e trajetória dos níveis tecnológicos das economias mundiais.

Em geral, os recentes bens e serviços produzidos avançaram em termos tecnológicos e estão mais intensivos em conhecimentos. No entanto, a geração de inovação não ocorre de forma homogênea no mundo, ainda existindo regiões que não inovam ou possuem níveis tecnológicos precários. Contudo, como os efeitos das inovações não se limitam às regiões onde estas surgem, seus conhecimentos e benefícios podem ser transbordados entre agentes e regiões sob a forma de externalidades.

Deste modo, regiões mais atrasadas e deficientes tecnologicamente buscam aprimorar seus níveis de competição no mercado e produção de inovações por meio da absorção de conhecimentos provenientes de regiões mais avançadas tecnologicamente (VIOTTI, 1997). A capacidade tecnológica de uma região reflete tanto suas habilidades em criar e desenvolver novos conhecimentos, que podem ser incorporados às novas tecnologias, quanto sua eficiência em absorver, compreender e utilizar os já existentes (TAKASHI, 2002).

Sugere-se, no entanto, que os efeitos das externalidades geradas pelas atividades de inovações decaiam conforme maiores distâncias geográficas (JAFFE *et al.*, 1993; ANSELIN *et al.*, 1997). Em análises espaciais evidenciam-se que os conhecimentos incorporados nas inovações enfrentam dificuldades de transmissão a longas distâncias, neste caso, apenas os conhecimentos de caráter codificado seriam transmitidos com eficiência. Deste modo, a proximidade geográfica pode auxiliar na compreensão e absorção de conhecimentos mais complexos. Tais conhecimentos, ditos tácitos, são mais facilmente transmitidos por contatos face a face (GLAESER, 1999; FISCHER; VARGA, 2003; MORENO *et al.*, 2005).

Por outro lado, trabalhos recentes têm questionado a relevância da proximidade física entre agentes para a ocorrência de trocas de conhecimentos, destacando-se que a proximidade relacional, identificada pela interação entre agentes em redes sociais, é capaz de gerar

atividades de inovações tecnológicas, justamente pela facilidade em que seus participantes têm de trocar informações (BOSCHMA, 2005; MAGGIONI; UBERTI, 2007; BRESCHI; CATALINI, 2010; GUAN *et al.*, 2015).

Inovações podem surgir da combinação de conhecimento e experiências de vários agentes (LUNDEVALL, 1992). Sendo assim, redes de inovação são estruturas sociais que dependem da interação entre agentes e do compartilhamento de conhecimentos e recursos entre eles objetivando gerar inovações por meio de colaborações (GRANT, 1996).

Acredita-se que atividades tecnológicas geradas por mais de um inventor possa obter resultados mais satisfatórios, em termos de geração de tecnologia e inovação, do que se produzidas por um único agente. Este fato ocorre devido às copatentes, isto é, patentes geradas por colaborações, incorporarem conhecimentos originários das experiências de cada agente colaborador, de modo que, as distintas habilidades e conhecimentos podem auxiliar na superação de barreiras locais e geração de inovação (BRESCHI; LISONI, 2009; SINGH; FLEMING, 2010).

Redes sociais de invenção, identificadas e medidas com auxílio de dados de copatenteamento das atividades inventivas, são ditas capazes de transbordar tanto o conhecimento codificado quanto o tácito. A explicação disso ocorre devido às redes gerarem confiança mútua entre seus agentes durante o processo de produção tecnológica por compartilharem objetivos e recursos em comum (GRANOVETTER, 1973; 2005). A eficácia na geração de inovações se relaciona com as características individuais dos agentes que executam as atividades inventivas, do local onde se inserem, e principalmente, se estes fazem parte de estruturas de redes. Sendo, tais estruturas de redes meios pelos quais os agentes absorvem conhecimentos externos através de ligações com demais agentes (ROTHWELL *et al.*, 1974).

O presente estudo se encaixa no contexto de análise de redes sociais e tem como objetivo principal estudar quais são os determinantes locais, em termos econômicos, tecnológicos e de densidade demográfica, que influenciam na ocorrência de laços (ligações) entre regiões durante o processo de copatenteamento de atividades tecnológicas no Brasil. Considera-se para isto, tanto as redes formadas apenas por inventores brasileiros quanto aquelas que têm a presença de inventores estrangeiros. Além disso, analisa-se a relevância da própria estrutura de arquitetura do nó na rede para a ocorrência de laços. O trabalho relaciona as ligações presentes nas redes de colaborações como formas de aquisição de conhecimentos externos,

contribuindo para o surgimento de invenções, mais significativas em termos tecnológicos, e que colaboram para a superação de barreiras locais.

Paralelamente, objetiva-se oferecer algumas contribuições no âmbito de pesquisas voltadas para a Economia da Inovação ao inserir a discussão de redes sociais na dinâmica inventiva das regiões brasileiras. Por meio de uma análise dos determinantes das ligações nas redes, busca-se investigar quais fatores locais devem ser desenvolvidos, via maiores investimentos e políticas regionais, destacando-se as regiões que tendem a absorver maiores conhecimentos tecnológicos estrangeiros, em termos de parcerias na geração de inovações. O país, ao se inserir em redes de colaborações, pode se beneficiar em termos de investimentos de terceiros facilitando na geração de suas atividades tecnológicas, além disso, tendem a obter maior controle sobre seus recursos e novos fluxos de conhecimentos (KANG; PARK, 2013).

A fim de atingir os objetivos propostos a análise, primeiramente, utiliza-se do modelo Tobit, que considera a razão entre o número de ligações e o total de patentes geradas na região. Essa primeira abordagem objetiva investigar os determinantes das ligações daquelas regiões cujas atividades inventivas são mais dependentes de colaborações. O modelo Tobit se mostra adequado às estimativas dado que as variáveis dependentes, ligações por patentes, assumem uma distribuição normal truncada em zero, ou seja, assumem apenas valores nulos ou positivos.

Já a segunda abordagem, é realizada por meio do modelo Binomial Negativo Inflado em Zero (ZINB). Esta segunda abordagem busca analisar a variável dependente medida pelo número de ligações absolutas existentes no nó (região). Logo, a justificativa do modelo é que a variável dependente assume a característica de uma variável de contagem, verificando-se a superdispersão dos dados e excesso de zeros causado pela ausência de ligações.

Ambos os modelos utilizam-se da abrangência geográfica de Região de Influência Econômica (REGIC) imediata em sua forma mais desagregada, composta por 482 regiões. A utilização da REGIC imediata tem como propósito lidar de forma mais realista com a existência de regiões polarizadoras e suas regiões de influência na estrutura urbana-regional do Brasil.

Os aspectos locais considerados na análise são: escala de inventores, número de empregados na indústria, pesquisa e desenvolvimento - P&D - industrial e universitário, aglomeração urbana, aglomeração ao quadrado, participação no Produto Interno Bruto (PIB) e as importações de bens de capital. Já os aspectos das características e estruturas de redes

consideradas são: número de ligações internas (ao nó) por número de patentes geradas na região, centralidade *betweenness* e centralidade *closeness*.

Em termos de distribuição espacial das ligações nas redes no Brasil, considerando-se a razão entre ligações e patentes geradas pela região, destacam-se maiores números de ligações nas regiões em desenvolvimento tecnológico. Evidencia-se que tais regiões ainda sejam mais dependentes de colaborações para a geração de invenções. Por outro lado, ao considerar o número absoluto de ligações, destaca-se que haja maior número de ligações nos grandes centros tecnológicos do país. Fato que pode ocorrer por estas serem regiões com maior autonomia de suas atividades tecnológicas. Contudo, em ambos os casos, estas regiões pertencem em grande maioria à região sul e sudeste do país.

Como principais resultados encontrados para ambas as redes identificadas no Brasil, as variáveis que mais influenciam tanto na ocorrência de ligações por patentes quanto de ligações em termos absolutos são as variáveis de P&D, destacando-se o P&D universitário e a centralidade *betweenness*, que mede a influência de um nó intermediar ligações funcionando como um *hub* na rede. No entanto, as demais variáveis que captam aspectos regionais como importação de bens de capital, participação econômica da região e escala de inventores também impactam a ocorrência de ligações por patentes. A centralidade *closeness*, que mede a influência da proximidade de um nó diante demais nós na rede, por sua vez, influencia o número de ligações ao se considerar ligações em termos absolutos e, também assim como a participação da indústria na região, impacta na rede inter-regional.

Além da presente introdução, a dissertação é dividida em mais quatro seções. Primeiramente, apresenta-se uma revisão de literatura que abrange as definições e a importância das inovações para os agentes e regiões, introduzindo-se a temática de redes sociais nas análises da Economia da Inovação por meio de dados de copatenteamento. Em seguida, a segunda seção trata da metodologia a ser utilizada nas análises das redes, assim como a descrição dos modelos, Tobit e ZINB, e variáveis a serem analisadas. A terceira seção traz uma análise dos dados e os principais resultados obtidos pelas estimações. Por último, a quarta seção possui as conclusões do presente estudo.

2 REVISÃO DE LITERATURA

2.1 Inovações: definições e formas de mensuração

Ao considerar e distinguir a relação existente entre invenção, inovação e difusão, Schumpeter (1982) dava um maior destaque ao termo inovação. Fagerberg (2005), no entanto, salienta que tais termos se relacionam entre si, mas possuem peculiaridades individuais. Deste modo, a materialização e concretização da ideia que impulsiona a criação de bem ou serviço denomina-se invenção. Logo, inovação é a implementação desta invenção no mercado consumidor enquanto a difusão é a disseminação da inovação a demais agentes ou regiões (STONEMAN; DIEDEREN, 1994). Segundo Schumpeter (1982), a inovação merece maior destaque por ser responsável pelas mudanças nos longos ciclos de acumulação de capital, além dos limites tecnológicos. De modo que, a inovação é tratada como um processo intermediário à invenção e difusão tecnológica, isto é, na ausência de invenção não há inovação, e sem inovação não há difusão.

Sendo assim, toda inovação é uma invenção, mas nem toda invenção pode ser considerada, de fato, como uma inovação. A inovação cria ou aprimora bens e serviços existentes na economia, possibilitando novos mercados e o surgimento de novos recursos, insumos ou processos produtivos (SCHUMPETER, 1982). Para a *Organization for Economic Cooperation and Development* (OECD), as inovações além de criar novos ou aprimorados bens e serviços, podem criar novos processos, meios de divulgação e *marketing* e também novas formas organizacionais de negócios, além de ambientes corporativos e relacionais (MORTENSEN; BLOCH, 2005). Porém, para serem consideradas “inovações”, as invenções devem possuir como pré-requisito sua ampla aceitação no mercado consumidor, dado que esta permeia a existência da invenção e sua difusão tecnológica (SCHUMPETER, 1982).

Atividades intensivas em conhecimentos tecnológicos são capazes de gerar inovações por meio da troca e desenvolvimento de novas ideias, impactando atividades de outros agentes ou regiões por meio de externalidades que levam ao crescimento econômico e progresso social. Todavia, atividades que almejam gerar inovações são incertas quanto aos seus resultados finais, sendo tais resultados considerados imprevisíveis e não intencionais (ROSENBERG, 1996).

Arrow (1962) considera as inovações como bens públicos, dado que os conhecimentos gerados por estas são não excludentes e não rivais. Polanyi (1967) também salientava sobre a existência da não rivalidade de conhecimentos, porém acreditava haver nas inovações uma exclusão parcial de informações. Tal consideração se explica por certas inovações não poderem ser codificadas e pelos agentes terem dificuldades em explicitar a terceiros todo o conhecimento técnico que possuem. Sendo assim, haveria uma perda parcial de informação durante o processo de transbordamento de conhecimento. Já para Dosi (1988) as inovações deveriam ser classificadas como bens semipúblicos. Sendo o fator público, o conhecimento codificável, enquanto o fator privado seria dado pelo conhecimento tácito, incorporado nos indivíduos, abrangendo suas experiências e, portanto, difíceis de serem transferidos.

Empiricamente na área da Economia da Inovação, os termos “invenção” e “inovação” são tratados usualmente como sinônimos. Isto ocorre pelas dificuldades em dissociar os termos na prática (FAGERBERG, 2005). As inovações são difíceis de serem mensuradas e identificadas devido às suas características peculiares. No entanto, estatísticas de patentes amplamente são utilizadas como *proxies* de atividades de inovação ou de invenção nas análises empíricas.

De acordo com Nagaoka *et al.* (2010), ao se utilizar estatísticas de patente como *proxies* das atividades inovativas, deve-se destacar que nem todas as patentes representam, de fato, inovações. Além disso, algumas invenções deixam de ser patenteadas, podendo os pedidos de patentes serem realizados apenas em prol de estratégias competitivas. Logo, tais *proxies* podem subestimar a verdadeira dinâmica inovativa e a capacidade inventiva regional. Ademais, para que seja dada a concessão de patente a uma invenção, a mesma precisa atender alguns pré-requisitos básicos: novidade, não obviedade e utilidade, além de aplicabilidade industrial. Contudo, tais critérios e requisitos podem diferir-se ao longo do tempo e entre nacionalidades e regiões (GRILICHES, 1990).

Ademais, estatísticas de patentes podem ser aplicadas em estudos a fim de rastrear fluxos de conhecimento, mapeando a trajetória de criação e aplicação das patentes (JAFFE 1989; JAFFE; TRAJTENBERG, 2002; VERSPAGEN, 2005). Pioneiramente, Jaffe *et al.* (1993) utilizaram dados de citações de patentes como *proxy* dos transbordamentos de conhecimento, sendo posteriormente seguidos por outros autores (ALMEIDA; KOGUT, 1997; BALCONI *et al.*, 2004; SINGH, 2005; LEE *et al.*, 2010). Porém, tais estatísticas não captam todas as trocas de conhecimento ocorridas na economia. Interações relacionais entre agentes, como captadas pelas atividades de copatenteamento, se mostram como formas alternativas de se rastrear

fluxos de conhecimento tácitos, incorporados nos agentes, e que corroboram para o surgimento de novas invenções e inovações (WAL *et al.*, 2009).

É possível identificar por meio dos dados presentes nos pedidos de patentes a existência de redes de inventores. Tal fato ocorre, pois a documentação das patentes concentra informações a respeito da localização e identificação de seus inventores e coinventores (CANTNER; GRAF, 2006; EJERMO; KARLSSON, 2006; MAGGIONI; UBERTI, 2007; 2009; OLIVEIRA, 2016). Sugere-se, no entanto, cautela ao se utilizar as estatísticas de copatenteamento. Primeiramente, pelas limitações das estatísticas, pois consideram apenas agentes com nomes inclusos nos registros de patentes, ignorando os demais que possam ter contribuído para a geração da inovação, seja durante a sua criação ou difusão dentro da rede. Outra questão é que as quantidades de registros de patentes não são homogêneas entre os setores econômicos, fato atrelado a formas alternativas de apropriação, isto é, segredos industriais, *trademark* e vantagens competitivas, ou até mesmo pelos altos custos de patentear conforme o tamanho das organizações. E, além disso, há uma tendência a subestimar patentes universitárias ou órgãos menos propensos a patentear (BOSCHMA *et al.*, 2009).

No entanto, mesmo com certas limitações, as estatísticas de patentes são dados fáceis de serem acessados e possibilitam comparar informações provenientes de localidades distintas. Desta forma, dados de patentes podem ser relevantes em medir o potencial inventivo regional, a troca de conhecimentos entre agentes e, mais recentemente, também tem sido utilizados na investigação da formação e existência de redes de inventores (ALMEIDA; KOGUT, 1999; MORENO *et al.*, 2005, 2005b; SINGH, 2005; MAGGIONI; UBERTI, 2007; 2011; MIGUÉLEZ; MORENO, 2013)

2.2 O papel das inovações no crescimento econômico

O crescimento econômico tem como pré-requisito a existência de inovações como meio de aquisição e geração de conhecimento, além da promoção de progresso tecnológico. Caso houvesse ausência de inovações a nível mundial, haveria a existência de um estado “estacionário”. Isto é, o crescimento econômico das regiões se reduziria até atingir um valor de estagnação nulo, e apenas voltaria ao crescimento caso surgisse novas inovações (SHUMPETER, 1982). Assim, é preciso que as regiões se mantenham com níveis de inovações, mesmo que pequenos, no longo prazo, a fim de se manter em crescimento (PHELPS *et al.*, 2012).

Em geral, os bens e serviços recentemente produzidos são mais abundantes em termos de conhecimento, tendência já delineada pela crescente incorporação do conhecimento científico e tecnológico aos produtos com a quinta Revolução Científica e Tecnológica (PEREZ, 2004). Além disso, a globalização dos mercados, a facilidade de comunicação entre os agentes, as exigências do mercado consumidor e as necessidades de adaptar-se continuamente a uma conjuntura econômica de crises pode acentuar tal tendência. Logo, os conhecimentos provenientes de outras economias podem ser adquiridos por outras regiões por intermédio das trocas de informações no instante em que surge a invenção, ou pela aquisição de bens e serviços tratados como insumos e auxiliares do produto final (PICCI, 2010).

Patel e Pavitt (1991) encontram uma relação positiva entre produção tecnológica e a estabilidade do mercado, sendo as inovações formas de aumentar o nível de competição e desempenho dos agentes e regiões. Porém, a geração de inovações ainda não ocorre de forma globalizada, havendo uma tendência de aglomeração das atividades inovativas em regiões mais delimitadas economicamente, a fim de incorporar e absorver os esforços em investimentos e conhecimentos locais, dominando o mercado regional (ALBAGLI, 1997).

Romer (1990) sugere que as inovações incorporam características peculiares das regiões onde são criadas como mão de obra especializada, além das características da infraestrutura e economia da região. No entanto, inovações se caracterizam por não serem de uso exclusivo dos agentes e regiões que as produzem, levando a externalidades sob a forma de transbordamentos (JAFFE *et al.*, 1993).

Partindo das ideias de Marshall (1890; 2009), Griliches (1991) distingue os transbordamentos em pecuniários (*rent spillover*) e não pecuniários (*knowledge spillover*). Os *rent spillovers* se relacionariam à aquisição (compra) de bens e/ou serviços, apesar de os preços não refletirem o real valor e benefícios gerados por seus esforços. Já o *knowledge spillover*, ou transbordamento de conhecimento, ocorreria quando atividades e melhorias produzidas por um agente (ou região), como os conhecimentos provenientes de investimentos em P&D, beneficiam outros sem que haja, necessariamente, trocas monetárias para isto. Grossman e Helpman (1990) salientam que tais transbordamentos permitem, sem custos de transação de mercado e sem proibição legal, a aquisição de conhecimentos gerados por terceiros.

Para Krugman (1991), os fluxos de conhecimento eram invisíveis, não deixando evidências de sua existência, portanto, também eram imensuráveis. Contudo, desde o trabalho de Jaffe *et al.*

(1993), há diversas constatações, por meio de dados de patentes, de que os transbordamentos de conhecimentos possam ser mensurados e disseminam-se para regiões mais próximas geograficamente, gerando externalidades. No entanto, sugere-se que entre maiores distâncias geográficas haja um decaimento dos efeitos dessas externalidades (ANSELIN *et al.*, 1997; AUDRETSCH *et al.*, 1996).

Estudos que incorporam análises a respeito dos transbordamentos de conhecimento e inovação objetivam explicar o comportamento e a dinâmica das atividades e capacidade de inovar no espaço (BRESCHI; LISSONI, 2005). Os transbordamentos tendem a ser mais concentrados onde exista aglomerações de mão de obra de qualidade e disponível, e amplos investimentos em P&D industrial e universitário, sendo tais aspectos também relevantes para a concentração de atividades inovativas (AUDRETSCH *et al.*, 1996). Contudo, acrescenta-se que as aglomerações das atividades inovativas são impulsionadas por demais fatores estruturais, como a facilidade em acessar conhecimentos locais, disponibilidade de recursos e insumos de produção, e pelas próprias características naturais da região como sua localização geográfica (BRESCHI; CATALINI, 2010).

As aglomerações das atividades de inovações são capazes de gerar externalidades como transbordamentos de conhecimentos localizados, permitindo uma maior capacidade de absorção dos agentes já localizados na região além da atração de novos agentes. Tal fato pode além de contribuir para uma melhora do mercado local, fomentar novas inovações (MALMBERG; MASKELL, 2006). Isto ocorre porque normalmente, agentes são atraídos para onde possam usufruir de benefícios e conhecimentos tecnológicos, além de fatores estruturais locais (MAURSETH, VERSPAGEN, 2002).

No Brasil, as atividades inovativas são mais concentradas nas regiões centro-sul e sudeste do país. E há evidências por meio de estatísticas de registros de patentes, artigos científicos e número de pesquisadores, de que haja também um maior volume de transbordamento de conhecimento nesses locais (GONÇALVES, 2007; GONÇALVES; ALMEIDA, 2009; OLIVEIRA, 2016).

A construção de uma relação de interação entre países pode contribuir para o aprimoramento das atividades tecnológicas e de inovação de regiões menos desenvolvidas tecnologicamente. Tal relação pode ocorrer via importação de bens de capital, absorção de conhecimentos mais complexos, e pelas colaborações entre seus agentes durante a produção de inovação,

poupando-se esforços em investimentos em termos monetários e de infraestrutura local. Desta forma, limitações e barreiras tecnológicas locais podem ser superadas, complementando e combinando conhecimentos adquiridos de outras regiões (COOKE, 2002; TÖDTLING *et al.*, 2006).

Partindo-se de ideias e invenções preexistentes, os transbordamentos de conhecimento podem motivar o surgimento de inovações (BEAUNDRY; SCHIFFAUEROVA, 2009). Grossman e Helpman (1993) acreditam que o crescimento econômico ocorra mais rapidamente e de forma eficaz onde os fluxos de conhecimento ultrapassam as fronteiras nacionais. No entanto, regiões periféricas tendem a obter parcelas menores de transbordamentos de conhecimento locais se comparadas a regiões centrais. Logo, seus agentes procuram compensar e complementar a falta desses transbordamentos por meio de atividades de colaborações com regiões mais desenvolvidas (GRILLITSCH; NILSSON, 2015). Em termos mundiais, Maldonado (1999) destaca que além dos Estados Unidos, Japão e a Europa serem os maiores concentradores de atividades inovativas, também são doadores de conhecimento e criadores das tendências tecnológicas.

Os transbordamentos de conhecimento podem ampliar o conhecimento científico e tecnológico de países menos avançados, auxiliando no crescimento de suas economias e capacidade em inovar (COE; HELPMAN, 1995; LEE, 2005; 2006). Este fato pode ser verificado em países emergentes como a China, em que os transbordamentos de conhecimento internacionais levam à expansão econômica, possibilitando o acesso a tecnologias mais avançadas, via transferências tecnológicas e de P&D industrial (GROSSMAN; HELPMAN, 1993; IMBRIANI; REGANATI, 1997). Análises com foco na difusão de conhecimento e tecnologia estudam como ocorre a interação entre agentes a nível mundial, apesar de, frequentemente, descartarem os casos de análises de regiões menos desenvolvidas como a África e a América Latina (ALBAGLI, 1999).

Os transbordamentos de conhecimento podem ser explicados pela proporção de conhecimento codificado, amplitude de conhecimento tácito e a proximidade entre os agentes. Além disso, transbordamentos podem ser medidos por citações de patentes, mobilidade de trabalhadores e inventores e pela existência de redes sociais de colaborações na produção tecnológica (BATHELT *et al.*, 2004; BRESCHI; LISSONI, 2001). Ademais, salienta-se que a globalização permitiu maior conexão e interação entre regiões distantes, facilitando os fluxos de conhecimento por intermédio da *internet* e outros meios de comunicação de baixo custo.

De forma que, a existência de interação e comunicação a longas distâncias permite que haja transbordamentos via colaborações por pesquisadores pertencentes a países distintos (DING *et al.*, 2010).

As redes sociais se inserem nas discussões sobre a relevância da proximidade geográfica para a ocorrência dos transbordamentos e trocas de conhecimento dada a necessidade de se incluir nas análises tanto o conhecimento tácito quanto o codificado (XIANG *et al.*, 2013). As colaborações entre agentes e contatos face a face possibilitam as transferências de conhecimentos tácitos (NONAKA, 1994; UZZI, 1996; HANSEN, 1999). De acordo com Herstad *et al.* (2014), as colaborações, mesmo de longas distâncias, incorporam ambos conhecimentos devido estreitarem as relações e confiança entre os agentes, melhorando a sua capacidade de absorção de conhecimento.

Segundo Granovetter (1973; 2005), tanto o conhecimento tácito, quanto o codificado são trocados e disseminados com maior facilidade entre agentes de uma mesma rede social. Logo, explora-se a ideia de que laços próximos transbordam maiores conhecimentos, inclusive os tácitos, medidos por meio de registros de copatentes, supondo-se haver trocas de experiências durante o processo de colaboração da produção da inovação (HU; JAFFE, 2003; BALCONI *et al.*, 2004; BRESCHI; LISONI, 2009). Além disso, as redes sociais colaboram para maiores níveis de transbordamentos de conhecimento por possuírem menores custos de transação, mesmo a longas distâncias (GERTLER, 2003).

2.3 Análise de redes sociais (ARS)

A Análise de Redes Sociais (ARS), de acordo com Freeman (1996), origina-se de estudos das áreas da sociologia, psicologia e antropologia, e foram introduzidas na literatura por Jacob Moreno (1946). A ARS é definida como estrutura que abrange os indivíduos, instituições, regiões e/ou organizações, e as relações sociais existentes entre eles.

Estudos sobre as estruturas de conexão entre agentes a níveis micro e macro, desenvolveram-se apenas nas décadas de 1950/1960, tomando-se como base os estudos pioneiros de Barnes (1954), a respeito das influências das relações e interações entre indivíduos no sistema social, e de Euler (1739), sobre as teorias dos grafos nos campos de análises da matemática (GRANOVETTER, 1973). Apesar disso, o interesse em redes sociais em análises econômicas surgiu somente após sua incorporação na abordagem de “Sistemas Complexos”, originados da

física (ALBERT; BARABASI, 2002; NEWMAN, 2003) e da computação (BRODER *et al.*, 2000; DALEY; GANI, 2001). Antes disso, economistas desconfiavam da eficácia de uma abordagem com preços e incentivos econômicos não inclusos de maneira clara em sua análise (MAGGIONI; UBERTI, 2011).

As redes sociais têm sido utilizadas na área da economia da inovação, investigando-se principalmente a formação de redes de copatenteamento e de publicações científicas. O Quadro 1, no final da próxima seção (2.3.1), traz o resumo dos principais estudos que tratam as formações de redes como variáveis dependentes.

2.3.1 Redes sociais de inovação

As redes sociais de inovação são consideradas estruturas que podem canalizar e transbordar fluxos de conhecimento tecnológico entre agentes pertencentes a uma mesma estrutura de rede (POWELL; OWEN-SMITH, 2004). Além disso, são ditas capazes de promover tais trocas de conhecimentos e informações por meio da interação restrita ou não a um espaço geográfico.

Análises de redes devem, no entanto, incluir as características relacionais entre os agentes ao invés de apenas considerar seus aspectos individuais. E, em geral, são representadas por diagramas em que os nós (pontos) representam os agentes (indivíduos, instituições ou organizações) ou regiões; e os laços (linhas) representam as ligações, isto é, as relações existentes entre eles.

Considerando-se que certas tecnologias são mais complexas, deve-se haver uma maior capacidade de absorção dos agentes ou regiões, além de uma estrutura mais sólida de relacionamento e interação entre eles. A eficiência na difusão e transbordamento tecnológico por proximidade geográfica e relacional pode ser notada pelas sociedades colaborativas de Boston e do Vale do Silício (FLEMING *et al.*, 2004). Este exemplo se insere na perspectiva de estrutura de redes sociais, refletindo a relação existente entre uma estrutura de interação social e colaborações, e um conseqüente desenvolvimento de atividades inovativas. Casos de aglomeração de atividades de invenção e inovação como o Vale do Silício, nos Estados Unidos, destacam a relevância da estrutura e características locais no desempenho de agentes e ampliação de transbordamentos de conhecimento.

Ademais, destaca-se a relevância se estudar e compreender os impactos das estruturas das redes sociais no desempenho econômico regional (UZZI, 1996). A direção em que ocorrem as ligações em uma estrutura se relaciona à influência dos nós nas redes. Relações simétricas (não direcionais) indicam fluxos recíprocos e equilibrados, enquanto as assimétricas (direcionais) evidenciam existência de nós mais influentes na rede. Normalmente, identifica-se em algumas economias a ocorrência de inventores de zonas periféricas produzindo inovações para regiões centrais, o que Maggioni e Uberti (2011) denominaram como “fuga de cérebros”, fenômeno que pode ser identificado na Índia, por exemplo.

De acordo com Storper e Venables (2004), a colocalização de empresas, organizações, universidades, centros de pesquisas, indivíduos e demais agentes criam externalidades de conhecimento limitadas espacialmente e denominadas como *buzz* local. Por outro lado, a intensificação da globalização e cadeias de valores globais evidenciam redes de inovação de carácter transnacional e inovação aberta, cujos fluxos de conhecimento ultrapassam as fronteiras nacionais e denominam-se *pipelines* globais (CHESBROUGH, 2006).

Agentes localizados em uma mesma região facilitam a existência de intercâmbio de conhecimento entre eles, pois constroem um ambiente de confiança mútua (MARSHAL, 1890; 2009; AGRAWAL *et al.*, 2006). A proximidade geográfica pode facilitar as interações e diminuir os custos de transação e transporte (MASKELL; MALMBERG, 2006; BRESCHI; CATALINI, 2010). Além disso, a proximidade e a colocalização de agentes tendem a produzir uma rede social densa entre eles (AUDRETSCH, 1998).

Além disso, uma vez que a relação entre agentes é construída, a colaboração entre eles prevalece mesmo a longas distâncias geográficas. Apesar disso, a difusão de conhecimento nas redes está sujeita a decaimento espacial, de modo que agentes mais próximos tendem a obter trocas de informação mais eficientes (COWAN; JONARD, 2004). Contudo, se a proximidade relacional for adicionada às afinidades sociais, intelectuais e culturais há ocorrência transbordamentos de conhecimento, até mesmo tácito a longas distâncias.

Sugere-se que conhecimentos absorvidos, internamente ou externamente, complementam as atividades inovativas locais (CASSIMAN; VEUGELERS, 2006). Sendo que, conhecimentos externos podem superar as barreiras tecnológicas, aprimorando e acelerando a própria dinâmica de inovação interna (POOT *et al.*, 2009). Segundo Guan *et al.* (2015), redes de colaboração internacional funcionam como canais de comunicação, transbordando o

conhecimento para regiões mais distantes e nacionalidades distintas. Herstad *et al.* (2014) sugerem que a absorção de conhecimentos estrangeiros colabora para o desenvolvimento de outras regiões, por meio de novas e aprimoradas tecnologias originadas de culturas mais desenvolvidas.

Pesquisas envolvendo análises de redes de inovação investigam se a capacidade em inovar de uma região é mais influenciada por transbordamentos não intencionais, isto é, transbordamentos, ou por trocas intencionais de conhecimento geradas pelas colaborações (BATHELT *et al.*, 2004; MAGGIONI; UBERTI, 2007). Investiga-se também se a geografia apenas facilita as trocas de conhecimento ou se são necessárias para a ocorrência destas, investigando outras dimensões de proximidade, como a relacional (BOSCHMA, 2005).

Analisando-se firmas, Uzzi (1996) combina teorias organizacionais com teorias de redes sociais e argumenta que as atividades econômicas das mesmas são orientadas pelas estruturas e características dos laços sociais existentes entre as organizações. Sendo assim, laços criam novas oportunidades de parcerias e mercados, além de facilitar e estreitar o acesso aos já existentes.

A competição de mercado provém de organizações com redes sociais locais densas, auxiliando a sobrevivência de firmas em relação àquelas que fazem parte apenas de redes comerciais. Firmas de pequeno e médio porte se inserem em tais redes, pois não são capazes de competir com as demais se forem individualistas (CASAROTO; PIRES, 1999). Pereira (2005) acredita que qualquer organização deva se inserir em uma estrutura de redes, a fim de atingir um crescimento econômico mais rápido. Redes interorganizacionais visam apenas transferir conhecimentos e não os produzir. Firmas incorporam padrões de trocas semelhantes a estruturas de mercado, contudo, quando os laços existentes entre firmas são incorporados nesse padrão, produz-se a estrutura de redes (POWELL, 1990).

Grillitsch e Nilsson (2015) evidenciam que firmas de regiões periféricas e com menor infraestrutura tecnológica tendem a colaborar mais externamente do que as firmas de regiões centrais, a fim de compensar a falta de transbordamento local. Alguns autores evidenciam que as colaborações devem complementar e não compensar a falta de transbordamento local (BATHELT *et al.*, 2004, CAMAGNI, 1995; COOKE, 2002).

Ao se tratar das indústrias, as colaborações em redes são tratadas como formas organizacionais de coordenar a aprendizagem e qualificação dos agentes por meio do acesso à

informação (PISANO, 1991; ARORA; GAMBARDELLA, 1994; POWELL *et al.*, 1996). Redes podem levar a vantagens competitivas devido ao compartilhamento de recursos e objetivos em comum entre os agentes. Além disso, o aprendizado mútuo entre organizações pode colaborar para o desenvolvimento de tecnologias via colaborações, proporcionando maiores lucros, novos mercados e acesso a variados fluxos de informações e recursos locais.

Orsenigo *et al.* (2001) destacam as redes de colaboração entre organizações distintas como formas gerar produção de inovação. Evidencia-se que redes organizacionais densas e que incorporam instituições voltadas para pesquisas, tem sido meios de aprimoramento e evolução das inovações nos últimos anos (POWELL *et al.*, 1996). Tal que, relações entre organizações criam novas oportunidades, levando a maiores trocas de conhecimento por meio de colaborações entre seus agentes, e um conseqüente desenvolvimento tecnológico (XIANG *et al.*, 2013).

De acordo com Lerma *et al.* (2009), colaborações entre organizações buscam melhorar os níveis de competitividade dos parceiros por meio do desenvolvimento mútuo de projetos tecnológicos a partir de redes locais, nacionais e internacionais. De modo que, as empresas multinacionais fazem parte de redes globais/internacionais cujos nós representam firmas subsidiárias pertencentes a redes locais, e toda sua estrutura pode ser utilizada na compreensão da dinâmica de exploração e transbordamento de conhecimentos globais (POWELL, 1990). As empresas subsidiárias, neste caso, incorporam redes locais de cooperação com outros agentes locais (inventores, empresas, fornecedores, universidades e etc.) acessando uma gama maior e mais selecionada de recursos e informações que melhoram seu valor de mercado (PANTZALIS, 2001).

Dantas (1997) evidencia que conglomerados¹ globais utilizam redes de conhecimento para controlar seus financiamentos e investimentos. Desta forma, as redes possibilitam aumentar os níveis de produção destes conglomerados e podem disseminar os bens e serviços produzidos a nível mundial, por meio da globalização que promove vantagens competitivas. Existem redes internas aos conglomerados, em que se mantem certa concentração de informações sigilosas e privilegiadas, além de decisões importantes no poder das sedes ou matriz, que na maioria das vezes, localiza-se nas regiões mais desenvolvidas.

¹ Conglomerados se originam do processo de concentração e reconcentração do capital, sendo compostos geralmente por empresas ou instituições multinacionais que dependem de uma mesma matriz.

As redes sociais de colaboração e pesquisas demonstram-se como um meio importante para garantir a competitividade das regiões (COOKE *et al.*, 1997). Segundo Castells (1996), as redes se referem à difusão tecnológica e são formas mais relevantes de promover a inovação e a competitividade. As interações entre agentes econômicos podem ser utilizadas como formas destes se especializarem pela possibilidade de trocas mútuas de conhecimento. Contudo, países em desenvolvimento, inclusive o Brasil, ainda têm participado muito pouco do processo de geração de inovações, via cooperação e colaborações. Ao avaliar os registros de patentes, verifica-se que patentes com único autor tendem a obter resultados economicamente menos relevantes e propensos ao insucesso na geração de avanços tecnológicos e inovação. Singh e Fleming (2010) valida tal resultado analisando inventores que exercem atividades sem parcerias e colaborações (copatenteamento).

Segundo Fleming *et al.* (2007), experiências de copatenteamento são formas de vínculos sociais, facilitando a difusão tecnológica por meio de redes sociais. Redes de colaboração de longo prazo podem gerar relações íntimas, com confiança mútua entre os seus parceiros (ADEGBESAN; HIGGINS, 2011). O conhecimento adquirido por colaborações não se limita aos contatos face a face entre inventores, mas pode ser transportada e novamente transferida para outros locais, conforme o inventor se locomova espacialmente (BALCONI *et al.*, 2004; BRESCHI; LISSONI, 2009).

Maggioni e Uberti (2007) sugerem que as redes de copatenteamento abrangem o conhecimento tácito e o codificado ao mesmo tempo, sendo interessante analisar tanto a questão da distância geográfica, quanto a relacional como estruturas da ocorrência de rede de conhecimento inter-regional e internacional. Análises de redes devem primeiramente delimitar a população de análise e selecionar dados consistentes para a investigação, além das técnicas de amostragem (WASSERMAN; GALASKIEWICZ, 1994; MARSDEN, 2005). Contudo, ainda é precário o conhecimento de formas alternativas de rastrear os transbordamentos e mensurar os impactos das relações sociais entre agentes e regiões sobre a estrutura das redes internacionais de transbordamento.

Maggioni e Uberti (2007) utilizam matrizes de ponderação espacial e relacional a fim de computar a defasagem espacial, considerando-se vizinhos com fronteiras em comum e ligações diretas entre vizinhos dentro da rede regional. Logo, embora haja concentração espacial da inovação, de mão de obra especializada e de investimentos, é possível haver trocas de conhecimento entre agentes, mesmo a longas distâncias por meio de suas relações. Trocas

de conhecimento entre agentes a longas distâncias podem ser evidenciadas pela presença de redes de colaborações, alianças estratégicas, trabalhos temporários e redes sociais globais em geral (KNOBEN; OERLEMANS, 2006; POWELL *et al.*, 1996).

Duas ideias permeiam as discussões a respeito das redes. Primeiramente, sugere-se que as relações e interações entre agentes ocorram por questões geográficas, sendo a proximidade importante e necessária para a formação de redes. E em segundo lugar, considera-se que a forma com que os agentes se inserem e se posicionam na rede é mais relevante para a absorção e compartilhamento de conhecimento do que a distância física entre eles (PONDS *et al.*, 2010).

Ademais, também há duas visões ao se considerar a posição do nó na rede. Uma delas analisa se há uma relação entre a caracterização de nós como centrais ou intermediários e uma consequente eficiência de alcance dos transbordamentos de conhecimento, provenientes ou captados por estes nós (BURT; TALMUD, 1993; GILSING *et al.*, 2008; BURT, 1997). Dessa forma, agentes intermediários, nós entre dois ou mais subgrupos nas redes, captam maiores graus de novidade e conhecimentos distintos (GILSING *et al.*, 2008). Já a outra visão, investiga as propriedades de estruturas dos nós que evidenciam impactar os níveis de transbordamentos, explorando-se a ideia de que redes com laços de menores comprimentos e mais agrupadas podem transbordar conhecimentos com maior rapidez e alcance, sem grandes perdas de informações (GRANOVETTER, 1985; HE; FALLAH, 2009; BARABÁSI; ALBERT, 1999; NEWMAN, 2003; WATTS; STROGATZ, 1998; EBEL *et al.*, 2002; SEN *et al.*, 2003).

Até onde se sabe, há ausência de abordagens no meio acadêmico brasileiro que combine análises de aprendizagem a nível individual e a nível interorganizacional (redes). E, além disso, há escassez de pesquisas sobre os determinantes das redes e de suas ligações sob uma perspectiva de rede internacional. A presente análise tem o propósito de preencher essa lacuna, analisando a existência de ligações inter-regionais e internacionais em redes brasileiras. Utiliza-se, portanto, das estruturas de topologias de redes como fatores explicativos das ligações e da capacidade de transferir conhecimento entre agentes de regiões distintas e auxiliar na capacidade de inovar das regiões. Deste modo, Quadro 1 reporta os principais trabalhos que investigam redes e, assim como a presente análise, utilizam-se de dados de copatenteamento.

Quadro 1 – Resumo de trabalhos que utilizam dados de copatenteamento para investigar redes

Autor (ano)	Objetivo	Abrangência	Variável dependente	Variáveis independentes	Método	Resultado
Broekel; Boschma (2012)	Avaliar a influência das diferentes formas de proximidade nas redes de conhecimentos para setores industriais da aeronáutica e quais destas explicam o desempenho inovador.	Holanda	Copatentes	Proximidades cognitivas, sociais, organizacionais e geográficas.	Logit	A proximidade cognitiva, social, organizacional e geográfica são cruciais para explicar a rede. Há fortes evidências de que a proximidade cognitiva diminua o desempenho inovador.
Cassi; Plunket (2012)	Investigar os determinantes das redes de coinvenções.	França	Copatentes	Proximidade tecnológica e geográfica; se o inventor e aplicante são semelhantes; grau de centralidade.	Logit condicional e multinomial	Uma vez formada a rede, a proximidade social passa ser mais influente do que a proximidade geográfica ou organizacional.
Grillitsch <i>et al.</i> (2015)	Investigar se firmas em áreas periféricas compensam ou complementam a ausência / insuficiência de acesso a conhecimento local colaborando mais.	Suécia	Copatentes	Acessibilidade a conhecimentos locais; densidade populacional;	Logit	Firmas com menos acesso a transbordamento de conhecimento local tendem a colaborar mais; O tamanho da firma e de sua capacidade em inovar mais influi neste resultado. Firmas com pouca capacidade inovativa dependem mais da infraestrutura regional.
Guan <i>et al.</i> (2015)	Analisar o impacto da rede multinível nas inovações de energias alternativas.	Países do G7	Copatentes	Centralidade do país e cidade; buracos estruturais; coeficiente <i>clustering</i> ;	Binomial Negativo com efeitos fixos	Rede interpaís modera relações e inovações de redes intercidades. Centralidade e buracos estruturais a nível país elevam tais estruturas das cidades e atenuam efeitos negativos dos coeficientes <i>clustering</i> das mesmas.
Xiang <i>et al.</i> (2012)	Explorar por meio de dados de citações de patentes e de copatentes as redes de <i>spillover</i> de conhecimento entre empresas chinesas. Investigando mudanças na rede geradas pela colaboração e o destas sobre <i>spillovers</i> estrangeiros entre empresas nacionais.	China e EUA	Copatentes e citações	Densidade da rede, coeficiente <i>clustering</i> , média nodal e grau de centralidade das redes.	Análise de estrutura da rede (ARS)	Dados de citação não indica bem os fluxos de conhecimento, refletindo poucos canais de troca. Interação entre inventores permite rastrear os fluxos de conhecimento tácito, complementando dados de citações. Grupos similares contribuem para o <i>spillover</i> entre países.

Fonte: Elaboração Própria.

É possível identificar, por meio do Quadro 1 que os recentes trabalhos na literatura utilizam de dados de copatentes abrangendo relação entre inventores e abordam apenas as características individuais, locais ou de proximidades como o caso de Broekel e Boschma (2012), Cassi e Plunket (2012) e Grillistch *et al.* (2015), ou apenas as características de estruturas de redes como Xiang (2012) e Guan *et al.* (2015).

Verifica-se, portanto, que há certa escassez de trabalhos que considerem, ao mesmo tempo, a existência da influência do local e/ou das características das estruturas dos agentes nas redes, apesar de trabalhos recentes de redes de copatenteamento no Brasil como Mejdalani *et al.* (2016) e Gonçalves *et al.* (2016) analisarem a estrutura de redes associadas simultaneamente sob a perspectiva dos indivíduos e das regiões do país. No entanto, ainda há poucos trabalhos na literatura brasileira que investiguem os determinantes das ligações em redes de copatenteamento no Brasil. Logo, o presente estudo se diferencia dos demais trabalhos tanto por investigar as redes brasileiras quanto por seguir uma perspectiva de análise de redes de copatenteamento entre regiões e não apenas entre indivíduos.

2.3.1.1 Arquiteturas de redes sociais²

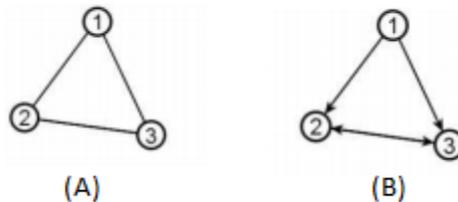
O conceito de topologia foi inserido nas análises de redes sociais, auxiliando a compreensão e estudo das mesmas. A topologia em redes abarca as estruturas de arquiteturas de redes que são definidas como variáveis estáticas. Tais estruturas abrangem o posicionamento dos agentes na rede e destacam-se como sendo referentes às medidas de centralidades, coeficiente *clustering*, conectividade, caminho médio e mínimo, densidade e outras medidas, que além das leis de movimento, que objetivam comparar redes sociais reais com as teóricas.

Primeiramente, em uma análise sobre as estruturas de redes sociais deve-se diferenciar se as ligações (laços) que existem são direcionadas ou não direcionadas. Sendo laços direcionados aqueles em que há um sentido definido em que ocorrem as ligações, isto é, quando é possível identificar quem ou quais são os nós doadores e receptores de conhecimento, bens ou serviços dentro da rede. Na Figura 1, os laços não direcionados são indicados pela ilustração (a); e as ligações direcionadas pela ilustração (b) do lado direito, sendo as direções em que ocorrem os fluxos de conhecimento indicadas pelas setas.

² Os principais conceitos e fórmulas de redes foram retiradas de Jackson (2008).

A maioria das redes de invenção possui o caráter não direcionado, isto é, apenas analisam-se as ligações existentes e não a direção em que estas ocorrem. Muitas vezes, as análises de redes, como no caso do presente estudo, são tratadas como não direcionadas pela falta de informações sobre qual agente é o central, ou seja, por meio do qual as informações se originam e se difundem. No entanto, em casos voltados para firmas ou grupo de pesquisas, o agente “central” pode ser mais facilmente identificado, como a matriz ou o diretor de pesquisa.

Figura 1 – Ligações de redes: não orientadas e orientadas



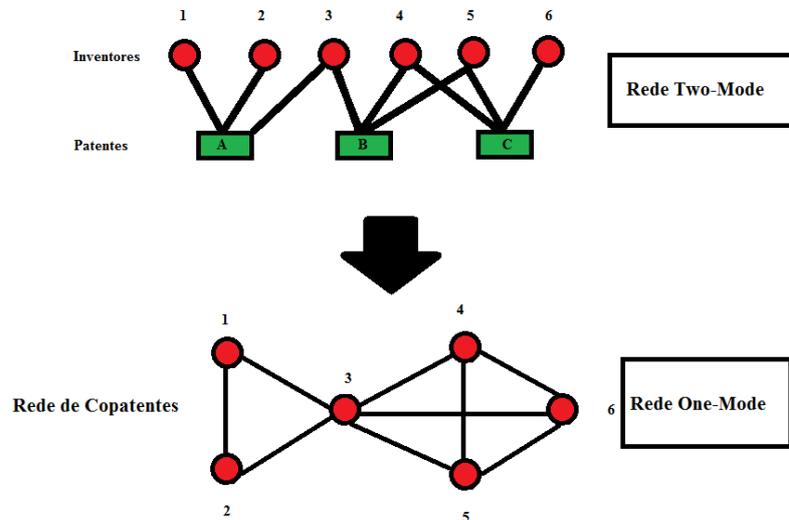
Fonte: Hoppe (2008).

Usualmente, as estruturas de redes são divididas em *one-mode* e *two-mode*. As redes *one-mode* possuem relações internas a um único grupo, sendo representadas por nós genéricos como indivíduos, organizações ou instituições. Já as *two-mode*, redes de afiliação, pertencem a dois grupos sociais contidos em conjuntos distintos e podem ser representadas por organizações empregando trabalhadores e acordos comerciais entre regiões (DOREIAN *et al.* (2004) *apud* DAVIS *et al.*, 1941).

No entanto, redes *two-modes* podem se transformar em redes *one-mode*, como no caso das redes de coinvenção ou copatenteamento. Neste caso, inventores são representados pela rede *one-mode*, e as patentes pela *two-mode*. O copatenteamento cria laços entre inventores ao trabalharem na produção de uma mesma patente, considerando-se que durante este processo estabelecem relações sociais mais próximas e intensas (BRESCHI; LISSONI, 2004).

A representação gráfica dos inventores é dada pelos nós e das patentes pelas ligações (laços) existentes entre eles. Tal representação pode ser vista na Figura 3.

Figura 2 – Representação transformação rede *two-mode* em uma rede *one-mode*



Fonte: Elaboração própria com base em He e Fallah (2014).

O conceito de centralidade visa medir a relevância ou influência de um nó (ou grupo de nós) para a rede. Nessa análise, esta medida tem como objetivo captar a relevância dessa estrutura da região (nó) como determinante para a ocorrência de laços com outras regiões. Há diversos tipos e formas de se mensurar a centralidade das redes, entre elas a centralidade de autovetor, de intermediação de fluxo, de eficiência e de intermediação de percurso. Algumas dessas medidas são difíceis de ser compreendidas e até mesmo mensuradas (FREITAS, 2010), contudo, Freeman *et al.* (1979) resumiram as mais utilizadas e de fácil aplicação em três conceitos básicos: a centralidade de proximidade, de grau e de intermediação.

A centralidade de intermediação (*betweenness*) e a centralidade de proximidade (*closeness*), que são utilizadas nesse estudo, pressupõem que o conhecimento transmitido entre os nós de uma rede só ocorre caso sejam utilizadas distâncias geodésicas, isto é, caminhos mais curtos que ligam dois nós. A centralidade *betweenness* pode ser considerada uma forma de medir os fluxos de informações entre as ligações do nó, enquanto que a centralidade *closeness* se relaciona com o tempo em que demora para o fluxo de conhecimento ser transmitido para outros nós da rede. Já a centralidade de grau mede a influência direta do nó em relação às suas ligações. Formalmente pode ser dada por:

$$d_k = \sum_{j=1}^n a_{kj} \quad (1)$$

considerando-se uma rede conectada ou não, d_k é a centralidade de grau; n é o número de nós da rede e a_{kj} são os elementos pertencentes a matriz adjacente.

A centralidade de intermediação (*betweenness*) pode ser descrita como:

$$C_B(v_k) = \sum_{\substack{1 \leq j \leq n \\ i, j \neq k}} b_{ij}(v_k) \quad (2)$$

de modo que, $C_B(v_k)$ representa a centralidade intermediária; n o número de nós da rede; e v_k representa o nó central intermediador.

$$\text{Assume-se que } b_{ij}(v_k) = \begin{cases} 0, & \text{se não houver caminho entre os nós } v_i \text{ e } v_j \\ \frac{g_{ij}(v_k)}{g_{ij}}, & \text{caso o contrário} \end{cases}$$

sendo g_{ij} a distância entre os dois nós i e j ponderados pelo nó k .

Sendo assim, na investigação dos determinantes dos laços nas redes brasileiras inter-regionais e internacionais, a centralidade *betweenness* tem como objetivo captar o quão importante é a região em termos de se conectar às demais, funcionando como um *hub*, atraindo maior número de ligações, e qual o impacto dessa estrutura na ocorrência de laços do nó com demais nós (regiões) na rede.

Por último, a centralidade de proximidade (*closeness*) é o inverso da soma das distâncias dos nós:

$$C_c(v_k) = \frac{1}{\sum_{j=1}^n \text{dist}(v_j, v_k)} \quad (3)$$

sendo $C_c(v_k)$ a medida de centralidade no nó k ; $\frac{1}{\sum_{j=1}^n \text{dist}(v_j, v_k)}$ é o inverso das distâncias entre os nós v_j e v_k . A centralidade *closeness* mede a facilidade com que uma região se conecta a outra dentro da rede, logo, espera-se que esta seja positivamente relacionada com as ocorrências de laços dentro das redes.

Outros coeficientes de estruturas de redes não serão utilizados nesse estudo, porém merecem destaque. Como é o caso do coeficiente *clustering* dos nós, que indica a presença de aglomerações na rede e possibilita medir a proporção em que os agentes se aglomeram, podendo ser calculado por:

$$C_i = K_i / k_i(k_i - 1) \quad (4)$$

sendo C_i a razão dos nós com ligações com i e os potenciais laços entre os nós da rede; K_i é o número de laços existentes entre nós que se conectam a i , e k_i é o número de laços do nó i . Considerando uma rede não direcional, esta variável permitiria verificar todos os casos em que as regiões A e B se conectam a uma mesma região C, e posteriormente analisaria a probabilidade, em média, de que as regiões A e B também se conectassem na rede, isto é, formando triângulos de conexão entre os agentes na rede. Segundo Easley e Kleinberg (2010), o coeficiente *clustering* representa a probabilidade de um determinado contato do agente A se conectar a um segundo contato deste mesmo agente.

O grau médio do nó representa o número médio de ligações entre os nós existentes nas redes, analisando a proporção de ligações entre os agentes. Quando o valor do grau médio é elevado, há indícios de que os agentes possuem diversos canais para se comunicar entre si, facilitando assim, a capacidade de transmitir e receber o conhecimento. Usualmente utiliza-se a distância geodésica (dimensão) existente entre os nós para medir qual é o caminho mínimo para que haja interação entre eles.

Dentre as medidas de rede mais usuais tem-se a estrutura de densidade, que possibilita mensurar a proximidade existente entre os agentes (nós), por meio da razão entre o número de ligações diretas existentes e o número de ligações possíveis na rede de colaboração. Logo, a densidade pode ser representada como:

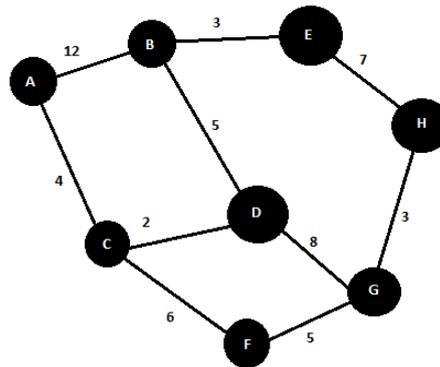
$$D = \frac{L}{[n(n-1)]} \quad (5)$$

sendo D a representação da densidade da rede que tem seu valor variando entre 0 e 1, sendo 1 a representação de uma rede muito densa e com inúmeras ligações. Por seu turno, L representa o número de ligações e n representa o número de nós da rede.

A medida de densidade, no entanto, é ambígua porque ao mesmo tempo em que se justifica que redes densas transmitem maiores fluxos de conhecimento, existe o argumento de que estas podem tender a se fechar em relação ao conhecimento externo a rede e tornam o conhecimento obsoleto (MIGUELLÉZ; MORENO, 2013; LOBO; STRUMSKY, 2008; FLEMING *et al.*, 2007).

Verifica-se na Figura 2 que a menor distância entre os agentes pode não ser a relação direta entre eles. No caso ilustrado, o caminho que liga diretamente A-B não é a distância menor que os conectam e sim a relação dada por A-C-D-B. Sugere-se que o fluxo de informação existente entre dois nós é diretamente proporcional ao número de caminhos que podem conectar os dois agentes e inversamente proporcional à distância das ligações entre eles (GRANOVETTER, 1973).

Figura 3 – Representação do caminho mínimo entre dois nós na rede



Fonte: Elaboração própria com base em Davis Júnior (1997).

Outra questão a respeito das estruturas das redes discutida na literatura são as forças das ligações. De acordo com Cantner e Meder (2008), a força das ligações leva em consideração o tempo e a intensidade da relação entre os nós, considerando-se a existência de confiança mútua e reciprocidade entre os agentes, e podem ser classificadas como: fraca, forte ou inexistente. Assim, quanto mais fortes as ligações, maior tende a ser a intimidade e confiança existente entre os agentes. No entanto, os laços fracos podem criar pontes³ locais e caminhos mais curtos de ligação conectando a nós relevantes e desconsiderá-los pode incorrer em problemas de probabilidade de transmissão (GRANOVETTER, 1973).

O vazio (buraco) estrutural é a inexistência de laços entre dois nós de uma rede, isto é, a ausência de conexão entre agentes. Segundo Burt (2004), essa estrutura controla ligações de lados opostos, além de indicar a existência de informações privilegiadas, já que lados opostos acessam fluxos de informações distintos (BURT; TALMUD, 1993).

³ Segundo Harary *et al.* (1965), uma “ponte” é um caminho único de fluxos de informações entre dois nós que pode interligar indiretamente quaisquer contatos diretos dos agentes centrais, considerados aqueles localizados nas extremidades da ponte.

3 METODOLOGIA

O presente trabalho tem como principal objetivo analisar os determinantes regionais e estruturais da ocorrência de ligações nas redes brasileiras de copatenteamento inter-regional e internacional. Para tal, consideram-se as ligações nas redes para o horizonte de tempo de 11 anos, compreendendo o período de 2001 a 2011, sendo a análise dos dados realizada por meio de técnicas econométricas baseadas em painel de dados.

Este método, associado à inclusão de variáveis de controle das regiões, busca minimizar os possíveis problemas de endogeneidade, que se presentes podem levar à inconsistência das estimativas. Além disso, a utilização de dados em painel objetiva proporcionar uma análise das ocorrências de ligações das regiões ao longo do tempo e controlar os efeitos não observados, além de acomodar a heterogeneidade entre as regiões. Sendo, neste caso, utilizado um painel de dados balanceado.

O painel balanceado tem como vantagem possuir o mesmo número de informações para cada uma das regiões abrangendo todo o período de análise. No entanto, uma de suas desvantagens é a ocorrência de uma possível seleção amostral. Contudo, como há disponibilidade de informações sobre todas as regiões consideradas, e a análise objetiva investigar a dinâmica dessas quanto aos determinantes de suas ligações em redes de colaborações ao longo da trajetória temporal, a utilização de um painel balanceado é adequada.

A identificação das interações entre agentes existentes nas redes de copatentes é realizada pela literatura por meio do valor absoluto ou número de ligações totais relativo à quantidade de inventores existentes na rede (BETTENCOURT *et al.*, 2007; MIGUELLÉZ; MORENO, 2012; 2013; HE; FALLAH, 2014). Sendo assim, nossa análise investiga tanto a razão entre a ocorrência de ligações (laços) e o total de patentes geradas pelo nó (região), quanto o número de ligações absolutas existentes no nó. Considerando-se, para isto, os dados de copatenteamento entre os inventores brasileiros e também as produções tecnológicas que possuem tanto inventores brasileiros quanto estrangeiros.

Como variáveis explicativas das ocorrências de ligações, em redes de copatenteamento regional, recorre-se aos aspectos locais de infraestrutura econômica, tecnológica, densidade demográfica e as características da arquitetura do próprio nó na rede.

Seguindo Grillitsch e Nilsson (2015), que analisam as redes organizacionais da Suécia e as classifica em nacionais, estrangeiras e globais, identifica-se dois tipos de formação de redes no Brasil: (1) rede de copatenteamento nacional, correspondendo às ligações entre inventores brasileiros localizados em regiões distintas do país (rede inter-regional - BB), e (2) rede que abrange as ligações entre inventores brasileiros e estrangeiros (rede internacional - BE).

Ademais, para melhor compreender e detalhar como se estruturam as redes e seus determinantes, serão analisados separadamente os dois tipos de redes formadas, internacional e inter-regional, para cada modelo e período de tempo. As definições sobre a formação de cada uma dessas redes identificadas no Brasil estão detalhadas no Quadro 2.

Quadro 2 – Definição das redes de copatenteamento identificadas no Brasil

Redes	Definição
Rede Inter-regional brasileira (BB)	Considera as ligações inter-regionais formadas apenas por inventores brasileiros;
Rede Internacional (BE)	Considera as ligações que possuam caráter internacional, isto é, ligações entre inventores brasileiros e estrangeiros;

Fonte: Elaboração própria.

A primeira abordagem tem como variável dependente a razão entre o total de ligações do nó (região) e o total de patentes geradas pela região, e utiliza-se do modelo estatístico para dados em painel Tobit com efeitos aleatórios. Esta abordagem, ao considerar a razão entre o número de ligações por patentes geradas, objetiva explicar os determinantes das ligações regionais, investigando as redes cujas regiões possuam inovações, em sua grande maioria, provenientes de colaborações. O modelo Tobit pressupõe a ausência de heterocedasticidade⁴, mas tem a vantagem de poder ser aplicado a dados medidos sob a forma de proporção e captam a presença de truncamento dos dados, de maneira que as observações assumem uma forma de dispersão normal truncada. Na presente análise os dados são truncados em zero, isto é, assumem valores nulos ou proporções positivas.

Já a segunda abordagem tem como variável dependente a contagem do total de ligações que a região possui. Sendo este um valor inteiro, utiliza-se do modelo para dados de contagem Binomial Negativo Inflado em Zero – *Zero-Inflated Negative Binomial* (ZINB)⁵. Este método

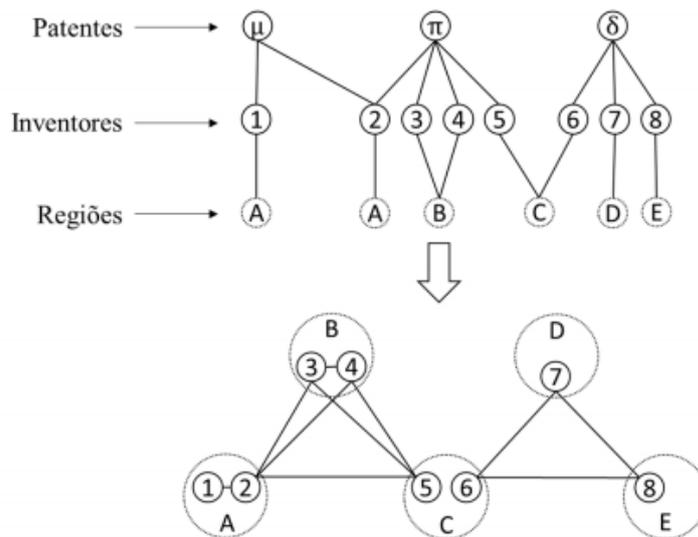
⁴ Neste caso, é necessária a utilização de estimações que considerem erros-padrões robustos ou por *bootstrap*. Como a estimação foi realizada no *software* STATA, utilizando-se o pacote *xttobit* e este permite apenas a estimação por *bootstrap*, esta foi utilizada a fim de gerar robustez às estimações.

⁵ A fim de evitar problemas de heterocedasticidade, foi realizada a estimação considerando-se erros-padrões robustos.

tem como vantagem ser aplicável aos dados de contagem, isto é, que apenas assumem valores inteiros além de tratar problemas de excesso de zeros na base de dados. Assim, visa explicar os determinantes das ligações absolutas, sem considerar a relação da variável dependente, número de ligações, com o total de patentes que a região possui.

Sendo assim, as redes de copatenteamento do Brasil podem ser identificadas como uma estrutura proveniente de uma transformação de uma rede *two-mode* em uma estrutura *one-mode*, como explicado na seção 2.3.1.1 Desse modo, as regiões compõem a rede *one-mode*, enquanto o copatenteamento compõe a rede *two-mode*. A formação da estrutura de redes da presente análise pode ser ilustrada graficamente por meio da Figura 4.

Figura 4 – Construção da rede de copatenteamento regional



Fonte: Gonçalves *et al.* (2016).

Como verificado pela Figura 4, primeiramente identificam-se as patentes e se estas são formadas por mais de um inventor. Posteriormente, localizam-se quais regiões os inventores são pertencentes. Desta forma, os nós representam as regiões cujos inventores se originam, podendo um mesmo nó conter ligações internas, isto é, entre agentes localizados na mesma região. Já as ligações entre os nós, isto é, entre regiões distintas, são identificadas pelas interações entre inventores e coinventores no momento da produção da patente, podendo estas ser ligações inter-regionais ou internacionais, como já explicitado pelo Quadro 2.

3.1 Discussão de hipóteses

Dado que regiões devem se manter com níveis mínimos de inovação a fim de se manterem competitivas, esta investigação parte do argumento de que maiores ligações inter-regionais e internacionais, por meio de redes de copatenteamento, possibilitam a aquisição de conhecimentos externos à região auxiliando na superação de barreiras tecnológicas e em maior produção de atividades inventivas locais. Logo, o presente estudo segue as evidências de Coe e Helpman (1995) e Lee (2005; 2006), e sugere que as interações entre agentes nas redes são formas de transbordamento de conhecimento, tácito e codificado e que, portanto, podem contribuir para maiores fluxos de conhecimento científico e tecnológico nas regiões.

Além disso, supõe-se que haja uma relação entre redes sociais de invenção, que podem ser identificadas por dados de copatentes, e uma elevação na produção de inovação. Sugere-se que invenções, quando realizadas em parcerias, tendem a incorporar maior conhecimento e gerar mais inovações (SINGH; FLEMING, 2010).

Partindo-se do pressuposto de que a ocorrência de ligações e o adensamento das redes, isto é, números elevados de ligações dentro do nó (região), contribuem para a ocorrência de um maior número de ligações inter-regionais, a proximidade entre os agentes, presença de mão de obra qualificada e as atividades de P&D influenciariam a capacidade de absorção e o maior fluxo de conhecimentos entre agentes, a nível local através de interações, compartilhamento de recursos e diminuição de custos de transporte (AUDRETSCH *et al.*, 1996; MASKELL; MALMBERG, 2006; BRESCHI; CATALINI, 2010). Agentes localizados na mesma região tendem, mais facilmente e frequentemente, a trocar conhecimentos, e por meio da geração de confiança mútua pode gerar uma rede social densa entre eles (AUDRETSCH, 1998).

Como este trabalho analisa as ligações entre regiões distintas, parte-se do pressuposto de que redes com ligações internas densas, isto é, ligações entre inventores dentro do mesmo nó (região), tendem tornar o conhecimento local obsoleto. Sendo assim, agentes pertencentes a nós com ligações internas densas procurariam participar de outras redes, ligações externas, a fim de adquirir novos conhecimentos que possam superar barreiras tecnológicas locais. Além disso, dado que regiões com ligações internas densas também estão relacionadas com maiores volumes de produção tecnológica, tais regiões podem se tornar procuradas por demais nós (regiões) a fim de gerar novas parcerias, existindo uma relação bidirecional de ligações.

Segundo Granoveter (1973), um grupo de nós conectados entre si formam um *cluster* cujos nós se conectam por laços fortes. Já a ligação entre *clusters* distintos ocorrem através de laços fracos. Redes com laços fortes tendem a incorporar menor circulação de novos conhecimentos, sendo importante o surgimento de laços fracos a fim de alcançar novos inventores e conseqüentemente novas informações (CAPALDO, 2007). Quase toda relação em uma rede social se inicia com um laço fraco que posteriormente pode evoluir para laços mais fortes. Os laços fracos, ao ligar agentes, podem fornecer diferentes informações colaborando para a inovação na rede (WU, 2012). Do exposto, segue uma primeira hipótese de investigação neste estudo:

H₁: Regiões com muitas ligações internas (laços fortes) tendem a participar mais de colaborações externas (laços fracos).

Ao analisar redes localizadas em zonas urbanas, verifica-se que há uma relação entre a aglomeração populacional e um ambiente propício à concentração de atividades econômicas. Tal relação, conseqüentemente, tende a atrair novos inventores de outras regiões. Essa atração de inventores pode ser identificada pelo aumento da estrutura de centralidade dos nós (regiões), tornando-os centros de atração (*hubs*) para diversos fluxos de conhecimento e elevando a força das redes (MATOS, 2002; MATOS; BRAGA, 2002).

A centralidade dos nós na rede pode influenciar o desempenho das atividades de geração de inovação, uma vez que nós centrais tendem a funcionar como meio de ligação de nós adjacentes através de laços fracos (WASSERMAN; FAUST, 1994). Destaca-se que a centralidade do nó influencie no desenvolvimento de inovações na rede de colaboração, pois o nós centrais tendem incorporar maiores fluxos de conhecimento ao se ligar a demais nós (ROXENHALL, 2013). Tais ideias a respeito de estruturas de centralidade permitem realizar uma segunda hipótese da análise:

H₂: Estruturas de centralidade da rede tendem gerar mais ligações entre seus nós adjacentes tornando-os centro de atração para novas colaborações.

Redes de invenção podem surgir de parcerias em atividades de P&D, permitindo aos agentes participantes da rede os benefícios dos investimentos nas pesquisas, além da possibilidade de aquisição de conhecimentos mais sofisticados (ERNST, 1994). Todavia, tais redes usualmente surgem em países mais desenvolvidos, onde também tendem a se aglomerar, incorrendo-se

em barreiras como alto custo de entrada por parceiros de nações menos desenvolvidas (FREEMAN; HAGEDOORN, 1994).

List (1856) evidencia uma associação entre o capital intelectual, P&D, e a produção tecnológica nas instituições de educação e ciência. Além disso, destaca que é o Estado é o órgão que deve coordenar e executar políticas que impulsionem o desenvolvimento e surgimento de tecnologias em prol do crescimento econômico. Instituições governamentais buscam incentivar a parceria de geração tecnológica entre universidade e demais instituições por meio do registro de parcerias nas patentes e promoção da cooperação entre agentes e instituições (MELO, 2012).

Como redes de copatenteamento representam formas de transferências de conhecimento entre regiões e absorção de conhecimentos estrangeiros e, como o P&D tem sido utilizado como *proxy* de conhecimento desde o trabalho de Griliches (1979), espera-se, que quanto maior os níveis de P&D, maior seja o número de laços das regiões nas redes de copatenteamento, devido à existência de parcerias e colaborações inventivas durante o processo de produção de P&D da região. Com isso, formula-se uma terceira hipótese de investigação:

H₃: Regiões com capacidade para produzir P&D tendem a possuir maior produção tecnológica por meio de colaborações.

3.2 Base de dados

A unidade geográfica de análise utilizada é o agrupamento de municípios brasileiros em uma divisão urbano-regional por intermédio do conceito de Regiões de Influência das Cidades (REGIC). Segundo o Instituto Brasileiro de Geografia e Estatística - IBGE (2008), as REGICs se baseiam em estudos de centralidade e hierarquia urbana, a respeito de localidades centrais como distribuidoras de bens e serviços para a sociedade.

O Projeto de Regiões de Influência das Cidades proposto em 2008 sofreu alterações no ano de 2013 e passou a compreender três tipos de divisões urbanas-regionais: divisão ampliada, intermediária e imediata⁶. A delimitação ampliada compreende 14 regiões de articulação urbana considerando os municípios de maiores influências em termos macrorregionais; a

⁶ A estrutura e composição da divisão urbano-regional estão disponíveis no site do IBGE em ftp://geoftp.ibge.gov.br/divisao_urbano_regional/base_de_dados/divisao_urbano_regional_composicao_das_regioes_xls.zip

subdivisão dessas regiões compreende a delimitação intermediária, composta por 161 regiões com articulação interna voltada para a oferta e demanda de bens e serviços complexos; e a delimitação imediata abrange 482 regiões de articulação urbana, considerada áreas de habitação da população e de deslocamento cotidiano.

Na literatura brasileira, em geral, os trabalhos utilizam unidades geográficas de microrregiões, considerando-se que estas se referem à estrutura de produção econômica, identificadas pela interação de áreas de produção a nível local. Contudo, essa divisão geográfica, assim as macros e mesorregiões, não reflete a real divisão urbana-regional do país e ignora os efeitos de polarização de alguns centros urbanos sobre outros (GUIMARÃES; FARIA, 2006).

Dado que o trabalho envolve as aglomerações urbanas e a influência de uma região sobre a outra, a análise abordará o nível de REGIC mais desagregado, REGIC imediata, a fim de capturar uma divisão urbana-regional mais próxima da realidade de polarização regional econômica brasileira. Contudo, destacam-se que todas as variáveis utilizadas nesta análise, dependentes e independentes, foram capturadas em nível de agregação por município e posteriormente agregadas para o nível geográfico de REGIC imediata através da identificação dos códigos de município e seu código de REGIC correlato, retirados do IBGE.

As variáveis dependentes e também as explicativas que utilizam dados de patentes na sua composição, como a escala de inventores, número de ligações internas e a ligações por patentes geradas no nó, foram elaboradas tendo como origem o Instituto Nacional de Propriedade Intelectual (INPI). A base de dados do INPI contém as informações presentes nos registros de patentes, abrangendo dados sobre a patente e sobre seus inventores e coinventores, assim como informações sobre sua origem.

No entanto, devido à existência de *missings* na base original e de informações com erros/diferenças de digitação, a fim de evitar perdas de informações, fez-se necessário uniformizar os CPF/nomes/siglas através do número da patente, com a conferência de parte destas mudanças através de pesquisas online. Além disso, como o objetivo do estudo é analisar as redes de copatentes, utilizou-se as informações de município contidas na RAIS Identificada, realizando-se a junção dessas bases através do número do CPF do inventor.

Para a identificação das ligações na rede, apenas foram utilizadas aquelas patentes cujas informações finais, após a uniformização e complementação, continham no pedido de patente a informação do inventor e coinventor(es). Além disso, como a base estava inicialmente

identificada apenas pelo número da patente, ela foi manipulada de forma a contabilizar todas as relações possíveis entre seus inventores durante sua produção, sendo assegurado de que cada relação entre inventor e coinventor apenas fosse contabilizada uma única vez por patente. Após estes procedimentos, os dados foram agregados pelo ano de depósito da patente e código de município e, por último, pela sua REGIC correspondente. Assim, identificaram-se as informações para cada ano e cada região e balanceou-se o banco de dados com um total de 482 observações, correspondente a cada uma das REGICs imediatas.

As variáveis dependentes foram construídas considerando a identificação dos inventores e a existência de colaborações durante o processo de produção tecnológica. Logo, cada vez que um inventor de uma região participa da produção de uma patente em parceria com outro agente de outra localidade brasileira e em dado ano, contabiliza-se uma ligação BB para cada região cujos inventores pertençam. As ligações com estrangeiros, por sua vez, foram contabilizadas para as regiões brasileiras como ligações BE, dado que nossa análise possui o objetivo de investigar os determinantes das ligações das regiões do Brasil.

Já para a construção do número de patentes geradas pela região, para o nível geográfico considerado identificou-se o número de processos de patentes distintas, geradas em parcerias ou não, depositadas para cada período considerado na análise. A variável explicativa de ligações internas da região foi construída por meio de uma combinação simples apenas considerando-se o número de ligações entre inventores localizados na mesma região, contabilizando-se uma ligação por cada relação existente entre os inventores no pedido de patente e tendo como denominador o número de patentes geradas pela região. Por exemplo, uma patente produzida em colaboração por quatro inventores, pertencentes a uma determinada região, possui seis ligações internas (inventor 1 e inventor 2; inventor 1 e inventor 3; inventor 1 e inventor 4, inventor 2 e inventor 3; inventor 2 e inventor 4; inventor 3 e inventor 4), não se contabilizando a relação entre dois inventores duas vezes. Caso fosse contabilizada a relação entre inventor 2 e inventor 1, sendo que já se contabilizou a relação existente entre ele através da relação entre inventor 1 e inventor 2, haveria dupla contagem de ligações.

Na construção da variável escala de inventores, o numerador representa a agregação feita pela identificação de inventores distintos localizados na região identificada nos pedidos de patentes. Neste caso, consideraram-se todas as patentes, isto é, invenções em parceria ou não,

durante o período de tempo investigado. Tirou-se então a razão entre o número de inventores da região e o total de inventores no país.

As variáveis referentes ao número de empregados e informações sobre a presença destes nas indústrias, além de dados sobre os trabalhadores que potencialmente desenvolvem atividades inovativas e voltadas para P&D, especificados como Pessoal Ocupado Técnico Científico (POTEC)⁷, têm como fonte a base de dados a Relação Anual de Informações Sociais (RAIS). A RAIS é um relatório imposto pelo Ministério do Trabalho e Emprego do Brasil às pessoas jurídicas que tenham mantido relação de empregadores no ano base. Contudo, tendo em vista que os dados da RAIS são obtidos através de relatórios preenchidos por pessoas jurídicas pode-se incorrer em erros de preenchimento. Além disso, como o número de declarações diferem em relação aos anos é difícil discriminar a real variação do emprego. De acordo com Gusso (2006) e Araújo, Cavalcante e Alves (2009), a POTEC e os gastos em P&D, externos e internos, possuem uma correlação alta, sugerindo que esta seja uma *proxy* adequada para medir o esforço inovador.

Os dados para elaboração da variável de capacidade de P&D universitário provêm da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). 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. No presente estudo, são considerados os dados sobre o número de professores doutores que atuam em centros de pós-graduação (*stricto sensu*) em áreas das Engenharias, Ciências Exatas e da Terra, Ciências Agrárias, Ciências Biológicas e da Saúde, que potencialmente podem criar tecnologias que seriam transferidas para o setor produtivo.

A variável de importação de bens de capital é medida como volume de bens importados em reais, tendo como fonte os dados anuais de indicadores econômicos e comércio dos municípios brasileiros, obtidos da Secretaria de Comércio Exterior (SECEX).

Por último, os dados sobre população, produção econômica e área geográfica das regiões foram extraídos do Instituto Brasileiro de Geografia e Estatística (IBGE), que é o principal provedor de dados estatísticos do país.

⁷ Araújo, Cavalcante e Alves (2009), definiram as ocupações de 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, arquitetos, segundo a Classificação Brasileira de Ocupações (CBO) de 2002.

3.3 Variáveis do modelo

A fim de padronizar a escala em que as variáveis são medidas e evitar uma possível presença de colinearidade entre elas, todas as variáveis, inclusive as variáveis dependentes, foram submetidas à logaritmização. Como este é um método de transformação monotônica sua transformação não altera a ordem original dos dados.

3.3.1 Variáveis dependentes: ligações por patentes e total de ligações

Na formação das redes de copatenteamento do Brasil consideram-se as regiões onde se localizam os inventores como os nós, enquanto as patentes que conectam inventores e coinventores pertencentes a regiões distintas representam os laços existentes na rede. Esses últimos são medidos pela quantidade de vezes em que inventores pertencentes a cada região (nó) aparecem como inventores ou coinventores de uma patente, mecanismo que já foi melhor explicado na seção 3.2, a respeito da manipulação da base do INPI.

Na rede inter-regional (BB), a variável dependente contabiliza para cada região o número de laços existentes entre as regiões, considerando-se, apenas aquelas patentes que possuem exclusivamente inventores e coinventores brasileiros e que são de regiões distintas. Este modelo tem como objetivo analisar a rede nacional, sem influências diretas de conhecimentos estrangeiros.

Na rede internacional (BE), contabiliza-se o número de laços existentes, considerando-se apenas aquelas ligações que possuem pelo menos um inventor ou coinventor brasileiro e outro estrangeiro. Tal rede busca analisar quais os determinantes regionais da ocorrência de ligações com estrangeiros durante o processo de produção tecnológica no Brasil.

Todavia, analisam-se duas estruturas de ligações em cada uma dessas redes, de modo que a variável dependente assume duas formas:

- Razão entre o número de ligações e o número de patentes geradas no nó;
- Número de ligações absolutas do nó na rede;

A primeira estrutura objetiva captar aquelas regiões mais periféricas em termos tecnológicos, mas que tendem a se ligar às demais regiões a fim de compensar a falta de transbordamento e capacidade local em inovar e, portanto, possuem mais produções provenientes de

colaborações. Ao longo do trabalho, ao tratar dessa estrutura de variáveis dependentes, denominar-se-ão as ligações como: $LigaBB_{pat}$ (rede inter-regional) e $LigaBE_{pat}$ (rede internacional).

Já a segunda estrutura não leva em consideração esta característica das patentes geradas no nó serem em sua maioria produzidas por copatenteamentos, mas apenas contabilizam o número de vezes que o nó se liga a outro dentro de uma rede de copatenteamento. Esta estrutura será denominada $LigaBB$ e $LigaBE$ em cada uma das respectivas redes: inter-regional e internacional.

3.3.2 Variáveis de interesse

Os modelos a serem estudados possuem como determinantes das ocorrências de ligações em redes de copatenteamento regional brasileira, variáveis explicativas relacionadas às características e infraestruturas locais, além de variáveis que se referem às próprias estruturas das redes. A fim de analisar o impacto das características tecnológicas e de atividades de inovações na ocorrência de ligações nas redes de copatenteamento regionais brasileiras, resumidamente, utilizaram-se as seguintes variáveis explicativas:

- $Escala_{inv}$, Ped_{priv} , Ped_{univ} e emp_{ind} : representam, respectivamente, a escala de inventores, medida pela proporção de inventores na região; insumos do processo inventivo, medidos pela capacidade de P&D industrial e universitário; e a presença industrial, medida pelo número de trabalhadores na indústria em relação ao total de trabalhadores da região;
- $Aglomeração$ e $Aglomeração\ ao\ quadrado$: denotam variáveis regionais de densidades populacionais (aglomeração urbana). O quadrado desta variável pretende captar deseconomias de escala, que é quando aglomerações muito densas passam a prejudicar a formação de redes;
- Bk_{pib} , Pib_{pibB} : são variáveis que medem o nível de participação econômica da região, sendo medida através da importação de bens de capitais e da participação da região sobre o nível de crescimento econômico do país medido pelo PIB;
- $close_{BB}$, $close_{BE}$, bet_{BB} e bet_{BE} : são características das estruturas de centralidades *closeness* e *betweenness* das redes;

- $LigaINT_{pat}$ é o total de ligações internas à região proporcional ao número de patentes geradas pela região.

Ao fim desta seção, o Quadro 3 resume as variáveis dependentes e as fórmulas das variáveis de interesse utilizadas neste trabalho, assim como suas fontes descritas na seção 3.2 e os sinais esperados pelas estimações. Além das variáveis de interesse, foram incluídas no modelo outras variáveis de controle: d_{metro} e d_{reg} , *dummies* metropolitanas e regionais. Adicionalmente, todas as variáveis explicativas de interesse foram defasadas temporalmente a fim de controlar possíveis problemas de endogeneidade (GONÇALVES; ALMEIDA, 2009; MIGUELÉZ; MORENO, 2013; ARAÚJO, 2013). Assim, para o período de análise de 2001-2011 consideraram-se as variáveis independentes medidas em 2000-2010.

Analisando a relação entre a aglomeração de inventores e a formação de redes, há poucas evidências da relação entre aglomerações de inventores e o efeito disto sobre a formação das redes, acreditando até mesmo que haja efeito negativo (LOBO; STRUMSKY, 2008; BETTENCOURT *et al.*, 2007). A aglomeração de inventores leva a uma maior facilidade de colaboração local entre eles, isto é, colaborações dentro da mesma região. Contudo, ao se considerar a ligação entre inventores em redes sociais, localizados em regiões distintas, esta variável tende a possuir um efeito negativo sobre a formação das redes. Sendo assim, a fim de investigar a relação entre a aglomeração de inventores na região e a ocorrência de ligações em redes inter-regionais e internacionais daquele nó, utiliza-se a variável explicativa escala de inventores ($Escala_{inv}$).

Como já discutido na seção 3.1 na Hipótese 3, sendo redes de copatenteamento meios de transferências de conhecimento, como insumo de conhecimento e a fim de relacionar a capacidade em produzir P&D e a formação de redes sociais de copatenteamento, utiliza-se tanto a *proxy* de P&D industrial (Ped_{ind}) quanto de P&D universitário (Ped_{univ}). E espera-se que estas variáveis tenham coeficientes com sinais positivos. De modo que, regiões com maiores níveis de P&D tenderiam a possuir maiores níveis de ligações com demais regiões por meio da participação em redes de colaboração como formas de captação de conhecimento.

A *proxy* de P&D industrial é construída considerando-se o número de trabalhadores empregados em profissões “técnico-científicas”, definidas por meio da POTECA, seguindo Araújo, Cavalcante e Alves (2009). Em análises para países desenvolvidos, usualmente utilizam-se informações sobre gastos diretos em atividades em P&D, realizados pelas

organizações privadas, como formas de medir esforços e capacidade em inovar, assim como proposto por Jaffe (1989). Contudo, esta é uma medida pouco utilizada em análises empíricas brasileiras, pois tais informações a respeito destes gastos por empresas apenas são divulgadas a níveis regionais mais agregados na Pesquisa Industrial de Inovação Tecnológica do IBGE (PINTEC) (SIMÕES *et al.*, 2005). Portanto, a alternativa de se utilizar dados da POTEC do número de trabalhadores em atividades em áreas tecnocientíficas como *proxy* dos esforços e capacidades em atividades de P&D tem sido uma forma de suprir essa limitação de dados do Brasil. Espera-se, portanto, que esta variável tenha uma relação positiva com a ocorrência de ligações em redes de copatenteamento no Brasil, dado que, atividades de P&D estão associadas a atividades inventivas e neste caso, invenções em colaborações.

Já a *proxy* de P&D universitário é construída tendo como base o trabalho de Gonçalves e Almeida (2009) considerando-se o número de doutores, permanentes ou visitantes, que atuam em centros de pós-graduação e em áreas de pesquisas destacadas como importantes na criação de inovações, como mencionadas na seção 3.2., anteriormente. Considerando que estes agentes possam produzir conhecimentos a serem utilizados na produção tecnológica, espera-se que a *proxy* de P&D universitário tenha uma relação positiva com a ocorrência de ligações em redes no Brasil, sendo este um meio de colaborações.

Outra variável explicativa relacionada às características tecnológicas regionais é representada pela razão entre o número de trabalhadores empregados nas indústrias de extração e transformação em relação ao número total de empregados da região (Emp_{ind}). Esta variável tem como finalidade capturar o grau de industrialização regional, relacionando-a com o processo de produção de inovações, dado que o setor industrial tem maior probabilidade de gerar patentes de caráter inovador do que outros setores econômicos. A relação entre o grau de industrialização e redes, como discutido anteriormente, é positiva em redes locais, dado que o setor industrial tende a se ligar a demais agentes a fim de se manterem competitivos. No entanto, quando se considera redes inter-regionais e internacionais o resultado pode não ser o mesmo, dado que estas indústrias podem estar mais preocupadas em participar de colaborações locais do que capturar conhecimentos de outras regiões. Além disso, evidencia-se que regiões periféricas com menores infraestruturas tecnológicas tendem a colaborar mais em redes externas do que as regiões mais centrais a fim de compensar a falta de um transbordamento local (BATHELT *et al.*, 2004; CAMAGNI, 1995; COOKE, 2002; GRILLITSCH; NILSSON, 2015). Portanto, o sinal esperado por esta variável é indefinido para redes entre regiões.

Uma característica regional que usualmente é considerada em trabalhos empíricos é a representação das economias de urbanização por meio da densidade demográfica ou aglomeração populacional. Neste trabalho, é utilizada a medida de densidade populacional, definida como população por área (km^2) da região. Regiões de menor densidade populacional, composta por vazios populacionais, são consideradas regiões mais periféricas e interiorizadas. Logo, como os grandes centros urbanos, ao contrário das regiões periféricas, tendem a possuir a densidade populacional elevada, relaciona-se a presença de densidade alta como evidência de centros urbanos mais desenvolvidos, em termos econômicos e sociais. As aglomerações populacionais concentram também os maiores níveis de rendas, além de uma demanda mais exigente em produtos e serviços. Assim, a produção de tecnologia e conhecimento científico tende a se aglomerar também nessas regiões (OLIVEIRA, 2008).

O uso de medidas de aglomeração tem como intuito incorporar a ideia de que regiões mais densas em termos populacionais tendem a promover maior interação entre seus agentes, facilitando os contatos face a face e colaborações, impactando positivamente no processo de produção de inovações (STORPER; VENABLES, 2004; MORENO *et al*, 2005; GONÇALVES, 2007). No entanto, aglomerações populacionais podem gerar problemas urbanos denominados como deseconomias de aglomeração como trânsito pesado, excesso de barulho, criminalidade elevada, poluição, serviços públicos insuficientes e custos de bens básicos elevados.

De acordo com Pereira e Lemos (2003), as economias de urbanização crescem de acordo com o aumento populacional e, ao atingir o seu ponto crítico, passam a decrescer devido às deseconomias de escala. Portanto, a fim de capturar a existência de problemas gerados pelo crescimento excessivo da densidade populacional, usa-se o quadrado da variável densidade populacional. Deste modo, espera-se que a densidade populacional tenha uma relação positiva com as ocorrências de ligações em redes, dado que deve gerar ambiente mais favorável a colaborações entre inventores, enquanto as deseconomias de aglomeração tenham uma relação negativa, pois ao invés de atrair inventores, repelem os mesmos.

A fim de analisar se os fatores de participação econômica de cada região estão relacionados com o número de ligações com os demais nós e, portanto, participação em colaborações de atividades inovativas regional, insere-se no modelo a variável de escala econômica (Pib_{pibB}) e a variável de importações de bens de capital (Bk_{pib}).

A variável explicativa de escala econômica é construída a partir da participação do Produto Interno Bruto (PIB) da região em relação ao PIB do país e tem como objetivo verificar se as regiões mais ricas e desenvolvidas economicamente tendem a participar mais de colaborações em atividades inventivas entre elas, ou até mesmo com aquelas menos desenvolvidas economicamente, consideradas periféricas. Estas tenderiam a participar das colaborações a fim de compensar a falta de absorção de conhecimentos locais como discutido anteriormente. Como as regiões que mais se destacam em termos de inovações são regiões mais desenvolvidas economicamente e em termos de infraestrutura, espera-se que a variável tenha uma relação positiva com a ocorrência de laços nas redes.

Já a variável que mede a proporção de gastos em bens de capital importados em relação ao PIB regional tem como objetivo capturar a proporção de transbordamento de conhecimento pecuniário por meio de compras de bens com tecnologias incorporadas. Destacam-se duas formas de absorção de conhecimento de caráter estrangeiro as quais auxiliam na superação de barreiras locais. Uma forma seria por meio de colaborações diretas com estrangeiros durante a produção tecnológica e a outra pode ocorrer por meio de importações de bens de capital, que direta ou indiretamente também podem ser utilizados na produção tecnológica. De acordo com Scmitz (1984), o grau da aprendizagem proveniente de compras tecnológicas é mais eficiente em termos produtivos caso a tecnologia seja importada, pois o uso sobre o bem importado, neste caso, é apropriado pelo país comprador e não supervisionado pelo vendedor. Já List (1856) destaca inclusive que há uma interdependência entre importações de tecnologias estrangeiras e o desenvolvimento técnico local.

Devido aos avanços da globalização, a recente produção de bens e serviços incorporam conhecimentos internacionais. Tais conhecimentos estrangeiros são absorvidos ou por intermédio de informações adquiridas durante a parceria na produção tecnológica, ou por meio da importação direta de bens e serviços, tratados como insumos e que auxiliam na produção do produto final (PICCI, 2010). De acordo com Keller (2010), o conhecimento tecnológico estrangeiro proveniente de importações impacta no aumento do estoque tecnológico nacional e eleva sua produtividade em inovar.

A relação entre importações de bens de capital e a produção de atividades inventivas se dá pelo fator tecnológico incorporado nos bens adquiridos que podem ser utilizados como insumos de produção de atividades inovativas, proporcionando a superação de barreiras locais de conhecimentos tecnológicos. No caso brasileiro, esta variável merece destaque devido à

existência de elevados gastos com importação de maquinaria a ser utilizada na produção, sendo que, por exemplo, esse valor chega a ser superior aos próprios gastos com P&D interno. Logo, espera-se que esta variável tenha uma relação positiva com a ocorrência de ligações nas redes, principalmente sobre a rede estrangeira.

Algumas medidas de redes são mais usualmente utilizadas em trabalhos empíricos sobre redes de colaboração (copatenteamento) de atividades inventivas. Estas são as medidas de densidade, centralidade e conectividade (BETTENCOURT *et al.*, 2007; LOBO; STRUMSKY, 2008; MIGUELLEZ; MORENO, 2013; HE; FALLAH, 2014). As variáveis de estruturas de redes utilizadas neste trabalho foram construídas a partir do *software* de análises de redes *Gephi 0.9.1.*⁸, que são os coeficientes de centralidade *betweenness* e *closeness*.

O coeficiente *betweenness* (betBB e betBE) calcula o quanto um nó se encontra em termos de distâncias geográficas, próximo a outras ligações, isto é, em termos de caminhos geodésicos de modo a transferir mais eficientemente os conhecimentos. O valor deste coeficiente é normalizado e varia de 0 a 1, e quanto mais perto de 1 maior a probabilidade do nó (região) mediar trocas de conhecimento dentro da rede. Um valor elevado deste indicador pode identificar as regiões como pontes. Como esta variável indica que quanto mais um nó exerce a função de *hub*, maior o fluxo de conhecimento que pode passar por ele, no caso de rede de inventores, espera-se que se o nó for um caminho de muitas conexões, ele exerce um poder de aglomeração em torno dele colaborando para maiores ligações no nó. Portanto, o coeficiente de centralidade *betweenness* teria um efeito positivo sobre a ocorrência de laços entre regiões.

O outro coeficiente de centralidade *closeness* (closeBB e closeBE) tem como objetivo analisar as ligações diretas e indiretas dos nós, estudando o comportamento da troca de conhecimentos. É construído através do inverso da distância geodésica, tendo como objeto de análise apenas o nó, excluindo-se os pontos isolados, sendo que seu valor normalizado varia de 0 a 1. O argumento desta variável é que ao medir a facilidade com que um nó se conecta aos demais, seja possível mensurar a distância relacional entre eles, em que nós localizados mais ao centro possuem maior centralidade de proximidade, enquanto que nós mais afastados são menos centrais e, portanto, possuem valor de contribuição menor. Portanto, também se espera que esta seja uma variável que se relaciona positivamente com as ligações nas redes.

⁸ O Gephi 0.9.1. é um *software* colaborativo e livre, bastante utilizado em análises de redes sociais (ARS) que tem como objetivo ilustrar grafos e gerar estatísticas de estruturas de redes.

E por último, foi incluída no modelo uma variável referindo-se a uma característica do nó na rede, identificada pela razão entre o número de ligações internas ao nó por patentes geradas na região ($LigaINT_{pat}$). Esta variável visa capturar se há uma relação entre o fato de a região possuir muitas ligações internas e a existência de conexões inter-regionais e/ou internacionais, isto é, se a característica da região possuir inventores está atrelada ao fato desta produzir mais em parcerias do que individualmente. Espera-se que a variável impacte positivamente tanto na ocorrência de ligações da rede inter-regional quanto da rede estrangeira. A fim de não tornar o conhecimento local obsoleto, tais inventores absorveriam conhecimentos externos, mediante colaborações também com agentes de outras regiões.

3.3.3 Variáveis de controle

Além das variáveis de interesse, faz-se necessário a inclusão de variáveis de controle nos modelos a serem estimados. Como se trata de uma análise regional optou-se por incluir variáveis *dummies* para controlar as características regionais, como a presença de áreas metropolitanas (d_{metro}). O Brasil conta com 26 regiões consideradas metrópoles a nível REGIC imediata, tendo como base⁹ as 26 microrregiões que contém municípios sede de regiões metropolitanas, segundo o IBGE. A variável d_{metro} assume valor 1 caso a região seja considerada uma metrópole e 0 no caso contrário. A inserção da *dummy* de região metropolitana no modelo tem como função controlar o fato de essas regiões concentrarem maiores atividades de patenteamento (CARLINO, CHATTERJEE; HUNT, 2007; ARAUJO, 2016).

Também foram inseridas *dummies* regionais para captar características dos sistemas regionais de inovação. No Brasil, devem-se integrar as políticas regionais aos interesses industriais a fim de amenizar problemas de disparidades regional quanto à presença industrial, urbanização e infraestrutura tecnológica que tendem a se aglomerar apenas no sul e sudeste do país, a importância dessa integração de políticas e interesses de atividades industriais se dá por estas possuírem relação direta com a atividade tecnológica e inventiva local (ALBURQUERQUE *et al.*, 2002, 2005; GONÇALVES; ALMEIDA, 2009, ARAÚJO, 2013; 2016). Essa variável *dummy* regional tem como objetivo controlar aquelas regiões mais inovativas localizadas na região sul e sudeste do país, assumindo o valor 1 caso o nó se encontre nessas regiões e 0, caso contrário.

⁹ A tabela com as divisões das regiões com metrópoles se encontra em Anexo (A1).

Quadro 3 – Descrição das variáveis do modelo

VARIÁVEIS DEPENDENTES	DESCRIÇÃO	FONTE DE DADOS	SINAL
LigaBB	Número de ligações com outras regiões brasileiras	INPI	
LigaBE	Número de ligações com estrangeiros	INPI	
LigaBB_{pat}	$\frac{\text{Número de ligações com outras regiões brasileiras}}{\text{Total de patentes da região}}$	INPI	
LigaBE_{pat}	$\frac{\text{Número de ligações com estrangeiros}}{\text{Total de patentes da região}}$	INPI	
VARIÁVEIS INDEPENDENTES	DESCRIÇÃO	FONTE DE DADOS	SINAL
<i>P&D privado</i> (<i>Ped_{priv}</i>)	$\frac{\text{PÓTEC da região}}{\text{Total de empregados}}$	RAIS	+
<i>P&D universitário</i> (<i>Ped_{univ}</i>)	$\frac{\text{Número de professores doutores}}{\text{População}}$	CAPES/ IBGE	+
<i>Participação de empregados na indústria</i> (<i>Emp_{ind}</i>)	$\frac{\text{Empregados na Indústria}}{\text{Total de empregados}}$	RAIS	?
<i>Escala de inventores</i> (<i>Escala_{inv}</i>)	$\frac{\text{Inventores da região}}{\text{Total de inventores}}$	INPI	-
<i>Densidade populacional</i> (<i>Aglomeración</i>)	$\frac{\text{População}}{\text{Área (km}^2\text{)}}$	IBGE	+
<i>Deseconomia de Escala</i> (<i>Aglomeración</i> ²)	$\left(\frac{\text{População}}{\text{Área (km}^2\text{)}}\right)^2$	IBGE	-
Importações de bens de capital sobre o PIB (Bk_{pib})	$\frac{\text{importação de bens de capital}}{\text{PIB região}}$	SECEX/ IBGE	+
Participação sobre o Pib do Brasil (Pib_{pibB})	$\frac{\text{PIB região}}{\text{PIB Brasil}}$	IBGE	+
Ligações internas por patentes no nó (LigaINT_{pat})	$\frac{\text{núm. de ligações internas ao nó}}{\text{núm. de patentes na região}}$	INPI	+
<i>Centralidade Closeness</i> (<i>CloseBB e CloseBE</i>)	$C_B(v_k) = \sum_{\substack{1 \leq j \leq n \\ i, j \neq k}} b_{ij}(v_k)$	INPI	+
<i>Centralidade Betweenness</i> (<i>betBB e betBE</i>)	$C_C(V_k) = \frac{1}{\sum_{j=1}^n \text{dist}(v_j, v_i)}$	INPI	+

Fonte: Elaboração própria.

3.4 Modelos de estimação

Partindo-se do embasamento teórico de que as transferências de conhecimentos e a interação entre agentes são formas de melhorar a capacidade em inovar de agentes e regiões, segundo Bathelt *et al.* (2004), cabe investigar se transferências intencionais ou transbordamentos de conhecimento seria mais relevante. Tal questão reflete diretamente a importância em se investigar a formação de redes sociais de invenção/inovação, dado que estas transmitem tanto conhecimentos tácitos quanto conhecimento codificado.

Enquanto que para as transferências não intencionais, isto é, transbordamentos, a questão espacial é importante, para as transferências intencionais, as proximidades relacional, tecnológica ou organizacional são o grande enfoque para os fluxos de conhecimentos (GRILLITSCH *et al.*, 2015).

De acordo com Boschma (2005), o grande desafio teórico da economia da inovação tem sido incorporar outras questões de proximidade e não somente a questão geográfica à análise. Logo, visando analisar os determinantes das ligações (ligações absolutas - LigaBB e LigaBE; e a razão de ligações por patentes - LigaBB_{pat} e LigaBE_{pat}) nas redes de invenção no Brasil, as equações a serem estimadas podem ser identificadas de forma generalizada como se segue:

Rede Inter-regional (Modelo BB):

$$\begin{aligned} \text{Log(LigaBB)} = & \beta_0 + \beta_1 \text{Log(LigaINT}_{pat(t-1)}) + \beta_2 \text{Log(aglomeração}_{(t-1)}) + \beta_3 \text{Log(aglomeração} \\ & \text{ao quadrado}_{(t-1)}) + \beta_4 \text{Log(escala}_{inv(t-1)}) + \beta_5 \text{Log(ped}_{univ(t-1)}) + \beta_6 \text{Log(ped}_{ind(t-1)}) + \\ & \beta_7 \text{Log(emp}_{ind(t-1)}) + \beta_8 \text{Log(bk}_{pib(t-1)}) + \beta_9 \text{Log(Pib}_{pibB(t-1)}) + \beta_{10} \text{Log(closeBB}_{(t-1)}) + \\ & \beta_{11} \text{Log(betBB}_{(t-1)}) + d_{reg} + d_{metro} + \varepsilon \end{aligned} \quad (6a)$$

$$\begin{aligned} \text{Log(LigaBB}_{pat}) = & \beta_0 + \beta_1 \text{Log(LigaINT}_{pat(t-1)}) + \beta_2 \text{Log(aglomeração}_{(t-1)}) + \\ & \beta_3 \text{Log(aglomeração ao quadrado}_{(t-1)}) + \beta_4 \text{Log(escala}_{inv(t-1)}) + \beta_5 \text{Log(ped}_{univ(t-1)}) + \\ & \beta_6 \text{Log(ped}_{ind(t-1)}) + \beta_7 \text{Log(emp}_{ind(t-1)}) + \beta_8 \text{Log(bk}_{pib(t-1)}) + \beta_9 \text{Log(Pib}_{pibB(t-1)}) + \\ & \beta_{10} \text{Log(closeBB}_{(t-1)}) + \beta_{11} \text{Log(betBB}_{(t-1)}) + d_{reg} + d_{metro} + \varepsilon \end{aligned} \quad (6b)$$

Rede Internacional (Modelo BE):

$$\begin{aligned} \text{Log(LigaBB)} = & \beta_0 + \beta_1 \text{Log(LigaINT}_{pat(t-1)}) + \beta_2 \text{Log(aglomeração}_{(t-1)}) + \beta_3 \text{Log(aglomeração} \\ & \text{ao quadrado}_{(t-1)}) + \beta_4 \text{Log(escala}_{inv(t-1)}) + \beta_5 \text{Log(ped}_{univ(t-1)}) + \beta_6 \text{Log(ped}_{ind(t-1)}) + \\ & \beta_7 \text{Log(emp}_{ind(t-1)}) + \beta_8 \text{Log(bk}_{pib(t-1)}) + \beta_9 \text{Log(Pib}_{pibB(t-1)}) + \beta_{10} \text{Log(closeBE}_{(t-1)}) + \\ & \beta_{11} \text{Log(betBE}_{(t-1)}) + d_{reg} + d_{metro} + \varepsilon \end{aligned} \quad (7a)$$

$$\begin{aligned} \text{Log(LigaBB}_{pat}) = & \beta_0 + \beta_1 \text{Log(LigaINT}_{pat(t-1)}) + \beta_2 \text{Log(aglomeração}_{(t-1)}) + \\ & \beta_3 \text{Log(aglomeração ao quadrado}_{(t-1)}) + \beta_4 \text{Log(escala}_{inv(t-1)}) + \beta_5 \text{Log(ped}_{univ(t-1)}) + \\ & \beta_6 \text{Log(ped}_{ind(t-1)}) + \beta_7 \text{Log(emp}_{ind(t-1)}) + \beta_8 \text{Log(bk}_{pib(t-1)}) + \beta_9 \text{Log(Pib}_{pibB(t-1)}) + \\ & \beta_{10} \text{Log(closeBE}_{(t-1)}) + \beta_{11} \text{Log(betBE}_{(t-1)}) + d_{reg} + d_{metro} + \varepsilon \end{aligned} \quad (7b)$$

Sendo, LigaBB e LigaBE as variáveis dependentes de número de ligações absolutas das redes inter-regional e internacional, respectivamente; LigaBB_{pat} e LigaBE_{pat} são as variáveis dependentes que contabilizam a razão entre o número de ligações e as patentes produzidas pela região, considerando as respectivas redes; β_0 é o termo constante; ε o termo de erro; e as demais variáveis são as definidas e resumidas anteriormente pelo Quadro 3, na seção 3.3.2.

A fim de justificar o uso dos modelos, as Tabelas 1 e 2 reportam as estatísticas descritivas das variáveis dependentes, número de ligações por patentes e ligações absolutas, respectivamente, considerando-se o período inteiro de análise, 2001 a 2011. O número de observações em ambas tabelas é igual à 5302 (482 REGICs x 11 anos de análise).

Através da Tabela 1, que reporta o número de ligações por patentes existentes nas redes, é possível identificar a proporção de zeros na rede, isto é, regiões que não se ligam a outras por meio de redes de copatenteamento. Verifica-se que na rede inter-regional, mais de 50% das observações possuem valor zero, isto é, são regiões que não possuem ligações com inventores de outras localidades. Na rede internacional, por sua vez, mais de 80% da base é composta por regiões que não possuem ligações com estrangeiros.

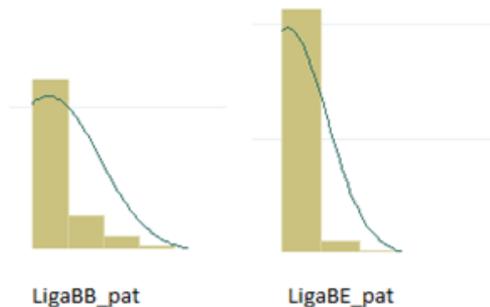
Tabela 1 – Análise descritiva das variáveis dependentes ligações por patentes ($LigaBB_{pat}$ e $LigaBE_{pat}$). Período de análise 2001-2011

REDES	REGIC Imediata					
	Nº de observações	Nº de Zeros	Média	D. Padrão	Min	Max
Rede Inter-regional ($LigaBB_{pat}$)	5302	3871(73%)	0,24	0,56	0	5,42
Rede Internacional ($LigaBE_{pat}$)	5302	5042(95%)	0,01	0,07	0	2,94

Fonte: Elaboração própria através do Stata.

A Figura 5 ilustra os gráficos das distribuições das variáveis dependentes $LigaBB_{pat}$ e $LigaBE_{pat}$. Observa-se, através dos gráficos das distribuições, que a variável dependente, número de ligações por patentes geradas pela região, assume uma distribuição normal¹⁰ truncada à esquerda para os dois tipos de ligações, isto é, $LigaBB_{pat}$ considerando a rede inter-regional e $LigaBE_{pat}$ para a rede internacional.

Figura 5 – Gráficos das distribuições das variáveis ligações por patentes ($LigaBB_{pat}$ e $LigaBE_{pat}$). Período de análise 2001-2011



Fonte: Elaboração própria através de estatísticas do Stata.

A utilização da metodologia de dados truncados Tobit para explicar o número de ligações por patentes geradas pelo nó, isto é, região, para ambas as redes (inter-regional e internacional), pode ser justificada através da verificação de que as observações assumem um comportamento semelhante a variáveis truncadas em zero, possuindo apenas valores nulos ou positivos. Uma estimação por MQO neste caso, por exemplo, incorreria no problema da suposição da existência de uma média linear em x , no entanto, isso só ocorre quando o intervalo de x seja está fortemente limitado, $E(y | x)$.

¹⁰ Além disso, foi realizado testes de normalidade Shapiro-wilk sendo identificada normalidade nos dados.

**Tabela 2 – Análises descritivas das variáveis de ligações absolutas (LigaBB e LigaBE).
Período de análise 2001-2011**

REDES	REGIC Imediata					
	Nº de observações	Nº de Zeros	Média	Variância	Min	Max
Rede Inter-regional (LigaBB)	5302	3871(73%)	0,75	2,34	0	9,49
Rede Internacional (LigaBE)	5302	5042(95%)	0,10	0,29	0	5,76

Fonte: Elaboração Própria.

A Tabela 2 aponta para um excesso de zeros. Como verificado, assim como na Tabela 1, na rede inter-regional, mais de 70% da base de dados é composta por regiões que assumem valores nulos de ligações, enquanto na rede internacional, 95% da base assumem valores iguais à zero.

Portanto, como neste caso se tratam de dados de contagem, é apropriado o uso de um modelo que considere este fenômeno e, ao mesmo tempo trate dos problemas identificados na base, de excessos de zeros. Logo, o modelo proposto é o Binomial Negativo Inflado em Zero (ZINB). Tal modelo além de considerar os excessos de zeros também se mostra adequado para observações com a presença de superdispersão, como também verificado na amostra através dos valores de variâncias superiores às médias, por meio das estatísticas da Tabela 2.

3.4.1 Modelo Tobit¹¹

O modelo estatístico Tobit¹², também conhecido como modelo para dados censurados ou truncados, surgiu em 1958 introduzido por James Tobin. Seu surgimento tinha como objetivo analisar a relação existente entre uma variável dependente não negativa (y_i), com valores censurados, e uma variável ou vetor de variáveis independentes (x_i) não censuradas (TOBIN, 1958).

O Tobit visa analisar variáveis dependentes censuradas ou truncadas, isto é, que são de certa forma, limitadas quanto aos valores que podem assumir, e admitem características tanto de variáveis aleatórias contínuas quanto de variáveis discretas. O caso mais comum de aplicação do modelo é quando a variável resposta, y , é um valor não negativo tendendo à distribuição

¹¹ As principais definições foram retiradas de Wooldridge (2010; 2002).

¹² O nome do modelo se dá devido ao sobrenome de seu criador e por fazer menção ao modelo estatístico *Probit*.

contínua sobre valores positivos e $P(y=0) > 0$. Neste caso, identifica-se a ocorrência do truncamento ou censura em zero.

O termo censura pressupõe que a amostra é extraída da população considerada originalmente e não se pode observar todos os possíveis resultados da variável resposta. Isto é, y é censurado quando observamos x para todas as observações, mas só obtemos o valor real de y para um número restrito (não censurado) destas. Já em uma amostra truncada, restringe o domínio original da variável e a redimensiona sobre o novo domínio. Neste caso, todas as observações, dependentes e independentes, possuem seus valores conhecidos, de modo que, y é truncado quando a amostra só respeita a parte não censurada. Deste modo, não se tem amostra completa de $\{y,x\}$, excluindo-se as observações baseadas nas características de y .

Cabe destacar, contudo, que podem ocorrer variações no modelo Tobit a partir de alterações em quando e onde há a ocorrência da limitação das variáveis. Valores de y de um intervalo são reportados como um valor único ou valores em torno de um número (k). Quando $y = k$ ou $y > k$ para todo y , se considera que y seja censurado a esquerda; e se $y = k$ ou $y < k$ para todo y , então y é censurado à direita. Pode haver casos em que a variável seja censurada também de ambos os lados.

Pressupondo um modelo como o nosso, cuja censura é dada em zero, e dado um grupo de variáveis que compõem o modelo, estas são decompostas em n_1 variáveis censuradas e n_2 variáveis não censuradas. O vetor y_2 corresponde ao grupo de variáveis que não são censuradas, isto é, assumem valores diferentes de 0 (zero). Logo, a técnica baseada no Tobit tem como objetivo gerar uma latente y_1^* para n_1 observações contidas no modelo, sendo y^* $(0, \infty)$, isto é, não há resultados negativos para y . Assim, o modelo assume uma distribuição Normal Condicional Multivariada Truncada à esquerda.

Dada uma variável aleatória contínua, y , a função de distribuição de Y pode ser identificada como $F(Y)$ e em casos de truncamento bilateral à esquerda, c , e à direita, d , a função de distribuição pode ser dada por:

$$Prob(Y \leq y | c < Y < d) = \frac{Prob(c < Y \leq y)}{Prob(c < Y < d)} = \frac{F(y) - F(c)}{F(d) - F(c)} \quad (8)$$

Já em casos de truncamento unilateral à esquerda¹³, como na presente análise, sendo $d = +\infty$ e $F(d) = 1$, a função 8 é modificada e a nova função de distribuição pode ser dada por:

$$F(y | Y > c) = \frac{F(y) - F(c)}{1 - F(c)} \quad (9)$$

e

$$f(y | Y > c) = \frac{f(y)}{1 - f(c)} \quad (10)$$

Em geral, o modelo pressupõe a existência de uma variável latente (y_i^*), que depende linearmente de x_i , sendo a relação entre ambas variáveis dadas por meio de um parâmetro β . Na presente análise, como se trata de um modelo em painel, y_{it} pode ser identificada como a quantidade de ligações da região i no tempo t , dado um conjunto de variáveis explicativas (x_{it}) sobre a região e sobre sua estrutura de centralidade (*betweenness* e *closeness*).

No modelo básico, a variável latente y_{it}^* pode assumir as seguintes formas:

$$y_{it} = y_{it}^* \text{ se } y_{it}^* > 0 \quad (11)$$

ou

$$y_{it} = 0 \text{ se } y_{it}^* \leq 0 \quad (12)$$

Isto é, sendo y_{it} a quantidade de laços em redes da região i no tempo t , $y_{it}^* > 0$ se a região se liga a outra por meio de copatentes e $y_{it}^* = 0$ caso ela não participe de atividades de colaborações em produção de inovações.

Podendo y_{it}^* ser representada por:

$$\begin{aligned} y_{it}^* &= \max(\beta x_{it} + u_{it}); & t &= 1, 2, 3, \dots, T \\ u_{it}/x_{it} &\sim N(0, \sigma^2) \end{aligned} \quad (13)$$

sendo u o termo de erro normalmente distribuído que visa capturar influências aleatórias na relação. Cabe destacar que β estima o efeito de x_{it} sobre a variável y_{it} latente (y_{it}^*) e não sobre a observada (y_i).

¹³ Em casos de truncamento à direita, $c = +\infty$ e $F(c) = 1$ na equação 10.

O modelo tem algumas características notáveis, como não manter a exogeneidade estrita de x_{it} , ou seja, u_{it} é independente de x_{it} . Porém, a relação entre u_{it} e x_{is} , $t \neq s$, não é especificada. Como resultado, x_{it} poderia conter $y_{i,t-1}$ ou demais variáveis afetadas por feedback. Outra questão relevante é que o $\{u_{it}: t=1, \dots, T\}$ possibilita a existência de dependência serial, podendo o y_{it} ser dependente após o condicionamento sobre as variáveis explicativas do modelo.

O estimador *pooled* maximiza a função de log-verossimilhança parcial e pode ser representado como:

$$\sum_{i=1}^N \sum_{t=1}^T l_{it}(\beta, \sigma^2) \quad (14)$$

sendo $l_{it}(\beta, \sigma^2)$, a função log-verossimilhança. No entanto, apenas se aplica, computacionalmente, o modelo aos dados de uma seção transversal longa de tamanho NT. Porém, enquanto a equação da matriz de informações condicionais for válida para todas as suposições (9), é necessário um estimador de matriz de variância robusta para a correlação serial no ponto t .

Assim como no modelo *Probit*, há casos em que se permite a heterogeneidade não observada, sendo possível especificar modelos lineares. De acordo com Wooldridge (2002), os modelos em painel permitem a modelagem dos efeitos não observados por meio de abordagens de efeitos fixos ou aleatórios. No caso de efeitos fixos há pressuposição de que o intercepto de cada região possa se correlacionar com um ou mais regressores. Já em casos de efeitos aleatórios, considera-se que o intercepto, aleatório, das regiões não se correlaciona com as variáveis explicativas.

O modelo Tobit de efeitos não observados pode ser dado por:

$$y_{it} = \max(0, x_{it}\beta + c_i + u_{it}), \quad t = 1, 2, \dots, T \quad (15)$$

$$u_{it}|x_i, c_i \sim Normal(0, \sigma_u^2) \quad (16)$$

onde c_i é o efeito não observado das ligações, e x_i é representa cada x_{it} no tempo t . A equação (12) supõe normalidade, mas também implica que x_i é exogenamente condicionado a c_i

Já considerando uma versão de efeitos aleatórios correlacionados, permitindo que c_i seja correlacionado com x_i , tem-se:

$$y_{it} = \max(0, \Psi + x_{it}\beta + \alpha_i + c_i + u_{it}) \quad (17)$$

$$u_{it}|x_i, \alpha_i \sim \text{Normal}(0, \sigma_u^2) \quad t = 1, 2, \dots, T \quad (18)$$

$$\alpha_i|x_i \sim \text{Normal}(0, \sigma_\alpha^2) \quad (19)$$

sendo, σ_α^2 a variância de α_i , supondo o efeito aleatório dado por $c_i = \Psi + \alpha_i$.

Deste modo, para a seleção do melhor método efeitos fixos ou aleatórios, usualmente utiliza-se o teste *Hausman*¹⁴, que para o presente estudo mostrou como melhor modelo o de efeitos aleatórios¹⁵.

A vantagem em se utilizar o modelo Tobit, ao invés de MQO é que este último ao considerar as variáveis de ligações censuradas em zero, levaria a estimadores viesados e inconsistentes, dado que se pode incorrer no problema de heterogeneidade e variável omitida. Deste modo, utilizando-se uma regressão pelo método de máxima verossimilhança, possibilita-se a existência de um estimador não-tendencioso e assintoticamente eficiente.

3.4.2 Modelo binomial negativo inflado em zero (ZINB)¹⁶

Uma variável de contagem é aquela que assume apenas valores inteiros e não negativos. Usualmente para a análise de dados de contagem utiliza-se do modelo de distribuição *Poisson*, também conhecido como distribuição de eventos raros. O modelo *Poisson* tem como pressuposto que a média seja igual à variância e trata de uma variável Y discreta que admite a seguinte função de probabilidade:

$$\Pr[Y, y] = \frac{e^{-\lambda} \lambda^y}{y!}, y = 0, 1, 2, 3, \dots, n. \quad (20)$$

¹⁴ O resultado do teste *Hausman* se encontra no final da Tabela de resultados estimados por Tobit.

¹⁵ É utilizado o pacote de estimação *xttobit* desenvolvido para o STATA. Tal pacote apenas permite a estimação Tobit via efeitos aleatórios, não tendo sido ainda desenvolvidas estatísticas que permitam incluir uma análise sobre os efeitos fixos. Honore' (1992) sugere um estimador semiparamétrico de efeitos fixos para o Tobit, no entanto, suas estimativas incondicionais são tendenciosas. E por meio do teste *Likelihood-ratio test sigma*, reportados nas tabelas de resultados, verificam-se a eficácia do modelo *xttobit* (efeitos aleatórios) em comparação ao *pooled*.

¹⁶ As fórmulas, definições e descrições foram baseadas em Hall (2000), Wooldridge (2002) e Cheung (2002).

No entanto, existem casos em que as observações não assumem valores de média e variância iguais. Quando a variância é superior à média, o fenômeno é denominado de superdispersão dos dados, que pode ocorrer devido a diversos fatores como: emissão de variáveis não observadas, quantidade elevada de valores zeros na amostra, alta correlação entre os dados ou variabilidade da média. Em geral, em casos em que se identifica o fenômeno da superdispersão dos dados, aplica-se o modelo Binomial Negativo.

O modelo Binomial Negativo corrige o problema da superdispersão presente no modelo *Poisson*, adicionando-se um parâmetro, α , que indica a heterogeneidade não observada. Deste modo, a distribuição das observações continua seguindo uma distribuição *Poisson*, no entanto, acomodando-se à superdispersão. Logo, a função densidade da distribuição do modelo binomial pode ser dada por:

$$f(y|\mu, \varphi) = \frac{\Gamma(y+\varphi)}{\Gamma(y+1)\Gamma(\varphi)} \left(\frac{\varphi}{\varphi+u}\right)^{\alpha-1} \left(\frac{u}{\varphi+u}\right)^y, \varphi > 0, \quad y = 0,1,2, \dots \quad (21)$$

de modo que $\Gamma(\cdot)$ representa a distribuição *Gamma*.

Contudo, em alguns casos, verifica-se também a presença de excesso de zeros na amostra, neste caso, tenta-se contornar o problema através de transformações na variável resposta, a fim de se alcançar a normalidade dos dados. No entanto, a não consideração de modelos voltados para tal inflação de zeros pode levar ao viés dos resultados. Desta forma, sugerem-se o uso de modelos inflados.

O modelo de *Poisson* Inflado em zero – *Zero Inflated Poisson* (ZIP) tem como objeto de análise dados de contagem que possuem um excesso de valores zeros em sua amostra (COHEN, 1963; SINGH, 1963; JOHNSON; KOLTZ, 1969). Este modelo assume duas distribuições, uma se refere à distribuição de *Poisson* ajustada aos valores inteiros e superiores a zero enquanto a outra refere-se à probabilidade de valores nulos. Logo, a distribuição de probabilidade pode ser dada por:

$$P(Y = y) = \begin{cases} p + (1 - p)e^{-\lambda}, & y = 0 \\ (1 - p) \frac{e^{-\lambda} \lambda^y}{y!}, & y > 0 \end{cases} \quad (22)$$

Os parâmetros da distribuição são definidos de duas maneiras: (1) pela distribuição *Poisson* (λ) excluindo-se os valores zeros e ajustando-se aos valores inteiros e positivos; e (2) através da distribuição de proporção de zeros (p), sendo y a representação do número de ligações na região ($y \geq 0$).

De acordo com Lambert (1992), ao se comparar os modelos ZIP e o modelo Binomial Negativo sem considerar os excessos de zeros, as estimações do modelo Binomial Negativo superestimam os resultados enquanto o ZIP obtém resultados mais consistentes. Resultados consistentes por meio do ZIP também foram encontrados por Miaou (1994) e o autor sugere cautela ao utilizar o modelo Binomial Negativo em casos de alta ou moderada superdispersão, pois neste caso estas podem estar sendo causadas pelo excesso de zeros.

Entretanto, Lawless (1987) propôs a distribuição Binomial Negativa Inflada de Zero (ZINB) que tem como característica a distribuição de dois modos, semelhante ao modelo ZIP. A distribuição do modelo ZINB pode ser dada por:

$$P(Y = y) = \begin{cases} p + (1 - p) \left(\frac{1}{1 + \sigma\lambda} \right)^{\frac{1}{\sigma}} & , y = 0 \\ (1 - p) \frac{\Gamma(y + \frac{1}{\sigma})}{\Gamma(\frac{1}{\sigma})\Gamma(1+y)} \left(\frac{\sigma\lambda}{1 + \sigma\lambda} \right)^y \left(\frac{1}{1 + \sigma\lambda} \right)^{\frac{1}{\sigma}} & , y > 0 \end{cases} \quad (23)$$

sendo y o número de ligações ($y \geq 0$); p a proporção de zeros; λ o parâmetro da média e $\Gamma(\cdot)$ a função Gama.

Segundo Zaniboni e Montini (2015), a distribuição ZINB e a ZIP tratam da superdispersão da variância em relação a médias, de formas parecidas. Carvalho e Lavor (2008) e Zaniboni e Montini (2015) encontraram resultados melhores quanto à significância de suas variáveis e estatísticas, para as estimações de ZINB quando comparadas ao modelo ZIP.

Contudo, nesta análise, verifica-se a presença de superdispersão, por meio da Tabela 2 na seção 3.4, e não sabemos se a causa desta é devido ao excesso de zeros ou não. Logo, é necessário estimar ambos os modelos, ZIP e ZINB, e verificar qual melhor se ajusta aos dados. A fim de verificar qual o método de estimação mais adequado, por meio de estimações de Máxima Verossimilhança - *Maximum Likelihood Estimation* (MLE) são estimados e realizados testes de seleção do melhor modelo, este é realizado por meio do teste

desenvolvido por Vuong (1989) que utiliza os resíduos das equações e pode ser representado pela seguinte fórmula:

$$Vuong = \frac{|\log(\sigma_w^2) - \log(\sigma_x^2)|}{\left[n^{0.5} \sum_1^n \left(\frac{e_{w,i}^2}{\sigma_w^2} - \frac{e_{x,i}^2}{\sigma_x^2} \right) \right]^2} \quad (24)$$

de modo que $\sigma_w^2 = (\sigma_y^2)(1 - R_w^2)$ e $\sigma_x^2 = (\sigma_y^2)(1 - R_x^2)$

sendo R^2 o coeficiente de correlação múltipla; n é o número de observações; σ_w^2 representa a variância residual do modelo sem considerar os zeros inflados; σ_x^2 é a variância residual do modelo com zero inflado; e σ_y^2 é a variância da variável dependente.

O teste de z-Vuong tem como hipótese nula que ambos os métodos são semelhantes em explicar a regressão e sua hipótese alternativa é que o método de zero inflado é melhor em explicar o modelo, se $z\text{-Vuong} > 0$ e caso contrário que o modelo sem zero inflado é melhor. Logo, a Tabela 3 reporta o teste *z-Vuong* que compara as distribuições como distribuição normal.

Tabela 3 – Resultados do Teste Z-Vuong para modelos inflados em zero. Período de análise 2001-2011.

Rede	Teste Z-Vuong ZINB	Teste Z-Vuong ZIP
Inter-regional (ligaBB)	62,69***	86,44
Internacional (ligaBE)	5,01***	1,02

Fonte: Elaboração própria a partir do software Stata.

É possível identificar que apenas o modelo ZINB mostrou-se significativo a um nível de significância de 1%, enquanto o ZIP não foi significativo para nenhuma das redes identificadas no Brasil. Ademais, como a presente amostra possui suas médias e variâncias desiguais, indicando a presença de superdispersão e, além disso, o ZINB corrige problemas de superdispersões cujas causas não sejam os excessos de zeros, justifica-se a utilização do modelo ZINB, sendo este adotado, em sua versão *pooled*¹⁷, para explicar os determinantes regionais das ocorrências de ligações absolutas nas redes brasileiras de copatenteamento.

¹⁷ Ao se utilizar dados em painel há a acomodação de heterogeneidade dos dados. Primeiramente foi estimado um painel *xtnbreg* e este se mostrou indiferente a uma estimativa *nb pooled*. Como o teste *Z-vuong* evidenciou que o modelo ZINB é melhor em ajustar os dados do que o modelo em sua forma original *pooled*, sem os zeros inflados, este foi escolhido em sua extensão robusta como o melhor modelo para estimar as regressões, dado que, sua versão robusta visa realizar estimações corrigindo possível viés causado por problemas de heterogeneidade.

4 RESULTADOS

Dado que o principal objetivo deste estudo é analisar quais os determinantes das ligações entre regiões nas redes brasileiras de copatenteamento, inter-regional e internacional, foi definida uma estratégia baseada em investigar tanto a contagem de ligações absolutas do nó na rede (LigaBB e LigaBE) quanto a razão do número de ligações pelo total de patentes geradas na região (LigaBB_{pat} e LigaBE_{pat}).

Como possíveis determinantes do número de ligações entre inventores no Brasil, investigam-se tanto os fatores de infraestrutura e características regionais, como as próprias estruturas de arquitetura dos nós nas redes formadas, sendo os resultados desta análise apresentados ao final desta seção.

Inicialmente, foi realizada uma Análise Explanatória dos Dados Espaciais (AEDE) com o objetivo de identificar e explorar as características espaciais das variáveis dependentes que se referem às ligações na rede e como estas se inserem no espaço geográfico. Para isto, foram utilizadas as estatísticas de *I de Moran* que tem como finalidade identificar a presença de dependência ou autocorrelação espacial dos dados. Isto é, o *I de Moran* é um índice que mensura a influência da média das variáveis de regiões vizinhas na participação de uma dada variável de uma região de análise, tendo como hipótese nula a não dependência espacial (ALMEIDA, 2012). Além deste índice, outras técnicas de análise espacial foram utilizadas como a análise de *clusters*, que tem como objetivo identificar como e se os dados se aglomeram no espaço, e a verificação de *outliers*, por meio da ferramenta *box-plot*, que permite neste caso, identificar aquelas regiões com maiores números de ligações em relação à média de todas as outras regiões consideradas na análise.

Ademais, como as variáveis dependentes já foram analisadas estatisticamente na seção anterior, relacionando o seu comportamento com a escolha dos modelos de regressão, nesta seção, serão reportadas as estatísticas descritivas de todas as variáveis de interesse dos modelos a serem analisados. De forma similar, serão analisadas e identificadas as regiões mais relevantes dentro de cada rede, inter-regional e internacional, por meio de *rankings* de regiões com maior número de ligações. Além de destacar aquelas regiões com maior o coeficiente de centralidade (*closeness* e *betweenness*).

4.1 AEDE

A AEDE é uma abordagem utilizada com o objetivo de encontrar a existência de um padrão na distribuição dos dados no espaço. Neste caso, foram analisadas as distribuições espaciais das variáveis dependentes utilizadas nos modelos, a saber, a contagem de ligações do nó na rede (LigaBB e LigaBE) e o número de ligações por patentes geradas por regiões (LigaBB_{pat} e LigaBE_{pat}).

A Tabela 4 retrata as estatísticas de *I de Moran* para ambas as estruturas de ligações e ambas as redes, considerando-se os dados para o período inteiro de análise 2001 a 2011. As estatísticas foram testadas para diversas matrizes de ponderação¹⁸ (W), como a matriz rainha, torre, k-vizinhos, distância inversa.

Tabela 4 – Estatísticas *I de Moran* para variadas matrizes de ponderação (W) para cada variável dependente das redes e período de análise 2001-2011.

Matriz de ponderação(W)	LigaBB _{pat}	LigaBB	LigaBE _{pat}	LigaBE
<i>Queen</i>	0,05	0,03*	0,03	0,02*
<i>Rook</i>	0,05	0,01	0,03	0,01
Distância inversa	0,04*	0,01	0,04	0,01
k1	0,04	0,02	0,03	0,01
k2	0,03	0,03	0,02	0,02
k3	0,03	0,03	0,02	0,02
k4	0,04	0,04*	0,01	0,03*
k5	0,03*	0,03*	0,02*	0,02
k6	0,04*	0,04	0,03*	0,02
k7	0,05*	0,03*	0,04*	0,02
k8	0,04	0,03	0,03	0,01
k9	0,03	0,02	0,02	0,01
k10	0,03	0,01	0,01	0,01
k15	0,03*	0,01*	0,02	0,01
k20	0,02	0,02	0,01	0,02

Fonte: Elaboração Própria.

Nota: ***: 1% de significância; **: 5% de significância; * 10% de significância.

Verifica-se que tanto a contagem de ligações do nó na rede quanto a razão do número das ligações e o número de patentes geradas na região possuem estatísticas do *I de Moran* não significativas na maioria dos casos e significativas apenas a 10% para algumas matrizes de ponderação, porém, com coeficientes bem pequenos. Este resultado é uma evidência de que pode não haver concentração espacial da distribuição das ligações nas redes brasileiras ou que

¹⁸ As matrizes de ponderação (W) podem ser escolhidas de maneira subjetiva pelo pesquisador, no entanto, usualmente utilizam-se medidas de peso relacionadas às distâncias entre as regiões como o caso da distância inversa; ou do número de vizinhos próximos, *Kn* vizinhos sendo n o grau de vizinhança; ou consideram os vizinhos de primeiro grau, de modo que, *rook* analisa os vizinhos em linhas retas e *queen* considera todos os vizinhos diretos ao redor da região de análise.

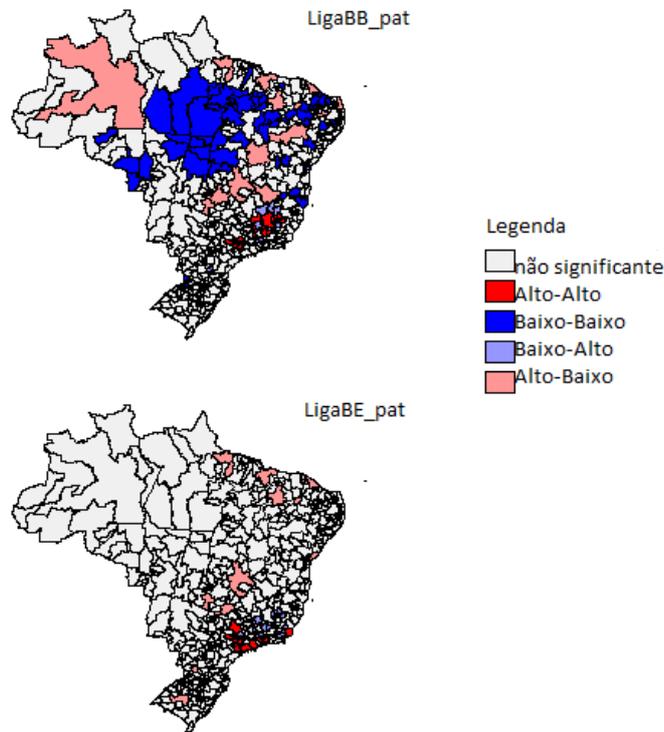
há uma concentração fraca, de modo que, as ligações dos nós vizinhos não influenciam umas às outras. Além disso, tal resultado de inexistência de dependência espacial pode se dar devido às próprias redes já incorporarem em sua estrutura a questão da distância entre os nós. De modo que, se houvesse a autocorrelação espacial dos dados, esta já pode ter sido corrigida ao se considerar as estruturas dos nós nas redes. Sendo assim, a utilização de modelos espaciais na investigação da existência de laços nas redes de copatenteamento regional do Brasil não se mostra necessária.

Por meio da análise de *cluster* LISA (*Local Indicator of Spatial Association*) é possível identificar no mapa os pontos onde regiões com valores elevados de ligações se encontram próximas a regiões também com alto número de ligações (alto-alto) ou próximo a regiões que pouco se ligam às outras (alto-baixo); e regiões com poucas ligações que se encontram próximas a outras com baixo número de ligações (baixo-baixo) ou próximas a regiões com números elevados de ligações (baixo-alto). Tal método permite, portanto, verificar a localização das concentrações geográficas.

As Figuras 6 e 7 trazem, respectivamente, a análise de cluster LISA para as variáveis dependentes da razão entre o número de ligações por patentes geradas na região ($LigaBB_{pat}$ e $LigaBE_{pat}$) e ligações totais do nó na rede ($LigaBB$ e $LigaBE$).

Com base na Figura 6, verifica-se uma concentração de ligações por patentes “alto-baixo”, localizada ao norte do país, isto é, regiões com elevados números de ligações por patentes estão próximas de regiões com números baixos dessas ligações. Evidencia-se que tal concentração “alto-baixo” pertence à região de Manaus. Há também concentrações do tipo “baixo-baixo” na região central do país e notam-se pequenas concentrações de ligações por patentes do tipo “alto-alto” na região sudeste, ao se analisar a rede de ligações inter-regionais. Contudo, ao se analisar a rede de ligações internacionais, é possível identificar poucos pontos significativos, havendo uma concentração de regiões “alto-alto” localizadas em pequenas áreas da região sudeste do país. A concentração inovativa e a importância da região sudeste do país no cenário tecnológico têm sido amplamente evidenciadas por trabalhos que exploram a distribuição espacial dessas atividades no Brasil (ALBURQUERQUE *et al.*, 2002; GONÇALVES, 2007; ARAUJO, 2016).

Figura 6 - Análise de Cluster LISA – Ligações por patentes (2001-2011)



Fonte: Elaboração Própria a partir do *software* econométrico Geoda.

Nota: Para cada rede utilizou-se a matriz W (k7) com o maior valor de I de Moran.

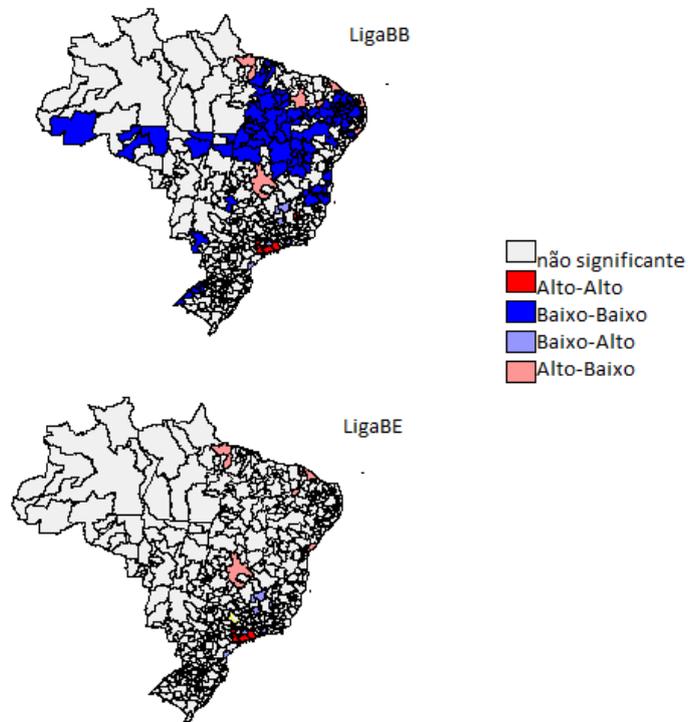
Por meio de uma análise *box-plot*¹⁹, identificam-se *outliers*, isto é, regiões que se destacam com número de ligações por patentes acima da média de suas regiões vizinhas. Na rede inter-regional, identificam-se como *outliers* positivos as regiões de Itabira, Caratinga, Belo Horizonte, João Molenvade, Ipatinga, Rio Claro, Campinas, Piracicaba, São Carlos, São João Del Rei e Congonhas. Já para a rede internacional são identificadas as regiões de Lavras, Guaratinguetá, Campinas, Cruzeiro, São Carlos, Limeira, São José dos Campos, Itajubá, João Molenvade, Campos Goytacazes, Araraquara, Araras, Barretos, Bragança Paulista, Piracicaba, Ribeirão Preto, Rio Claro, São Paulo, Sorocaba. Verifica-se que, ao se comparar as regiões *outliers* quanto ao número de ligações por patentes geradas, estas diferem um pouco, no entanto, todas estão inseridas na região sudeste do país.

Na Figura 7, ao se analisar a variável dependente contagem de ligações do nó na rede verifica-se a concentração “baixo-baixo” do número de ligações na região central do país e uma pequena concentração de “alto-alto” no sudeste do país, possuindo muitas regiões com estatísticas não significativas para a rede internacional.

¹⁹ Os *box-plots* se encontram no Apêndice (B).

Comparando-se as Figuras 6 e 7, é possível notar que em ambos os casos a rede inter-regional possui a distribuição de suas ligações com valores mais significativos estatisticamente do que a rede internacional, que apenas evidencia ligações significativas na metade leste do Brasil. Verifica-se que na rede inter-regional, tanto considerando as ligações por patentes (Figura 6) quanto o número de ligações absolutas (Figura 7) há uma concentração de baixa participação em redes de colaborações na região central do país, enquanto na região sudeste em ambos os casos verifica-se uma maior participação em redes, o que pode ser explicado por esta região concentrar os níveis de atividades inventivas do Brasil.

Figura 7 - Análise de Cluster LISA – Ligações absolutas (2001-2011)



Fonte: Elaboração Própria a partir do *software* econométrico Geoda.

Nota: Para cada rede utilizou-se a matriz W (k_4) com o maior valor de I de Moran.

Os pontos de *outliers*, isto é, as regiões que possuem número de ligações absolutas na rede com valores acima da média das outras regiões foram identificadas na rede inter-regional como São Paulo, Curitiba, São José dos Campos, Itabira, Piracicaba, Sorocaba, Ponta Grossa, São Carlos, Ribeirão Preto, Rio Claro e Jundiaí. Para a rede de ligações internacionais, destacam-se as regiões São Paulo, São José dos Campos, Sorocaba, Curitiba, Limeira, Campinas, Jundiaí, Piracicaba. Como se pode perceber, as mesmas regiões que se destacam com maiores médias de ligações na rede inter-regional também se destacam na rede internacional ao se considerar o número de ligações absolutas, sendo tais regiões pertencentes

à região sudeste do país, onde se localiza a maior concentração de atividades inventivas do Brasil.

4.2 Análise descritiva das variáveis de interesse²⁰

Como forma de verificar a variabilidade das variáveis de interesse dos modelos e compreender como se dá sua distribuição, a Tabela 5 apresenta as estatísticas descritivas de todas as variáveis independentes utilizadas no estudo, considerando o período inteiro de análise, de 2001 a 2011.

Tabela 5 - Análise descritiva das variáveis independentes (2001-2011)

Variável	Média	Desvio Padrão	Máximo	Mínimo
Log(ligaINT _{pat})	0,1759	0,4629	5,0626	0,0000
Log(Ped _{priv})	0,0031	0,0029	0,0619	0,0000
Log(Ped _{univ})	0,0001	0,0002	0,0035	0,0000
Log(Emp _{ind})	0,0102	0,0109	0,2436	0,0000
Log(Escala _{inv})	0,0020	0,0122	0,2577	0,0000
Log(Aglomeracoo)	3,3217	1,1630	7,5670	0,4397
Log(Aglomeracoo _{ao quadrado})	6,5137	2,4609	15,1329	0,2663
Log(Bk _{pib})	0,5248	0,7869	4,1927	0,0000
Log(Pib _{pibB})	0,0020	0,0096	0,1902	0,0001
Log(Close _{BB})	0,0807	0,1542	0,6931	0,0000
Log(Close _{BE})	0,0806	0,1522	0,6931	0,0000
Log(Bet _{BB})	0,0002	0,0013	0,0300	0,0000
Log(Bet _{BE})	0,0002	0,0014	0,0300	0,0000

Fonte: Elaboração própria.

A presença de valores próximos à zero na base de dados é verificada através da coluna de valor mínimo, valor máximo, média e desvio padrão. A título de exemplo, a variável ligaINT_{pat} possui grande parte de suas observações com valores próximos ou iguais a zero, uma vez que o seu valor máximo é zero, seu desvio padrão é maior que um e sua média permanece com valores menores do que um, sendo um indício que a média está sendo puxada para baixo. Isso indica que algumas regiões não apresentam ligações internas, verificando-se ausência de inventores em algumas regiões, assim como a falta de P&D universitário e/ou privado e de importações de bens de capital. De acordo com dados do Ministério da Ciência e

²⁰ No apêndice (A) se encontra a ilustração das correlações entre as variáveis de interesse no modelo. Verifica-se uma correlação alta (acima de 0,7) apenas entre as variáveis de escala de inventores e as variáveis de centralidade *betweenness* e a escala de inventores com a participação da região sobre o PIB nacional; e entre as variáveis de aglomeração e aglomeração ao quadrado. Por causa disso, as regressões foram realizadas com cautela, e estimadas alternando-se as variáveis. No entanto, os efeitos das altas correlações não causaram alterações significativas nas estimativas.

Tecnologia (2008), o Brasil ainda possui investimento privado em atividades de P&D de forma ainda bastante reduzida. O Brasil investe 1,02% do PIB em P&D, sendo que o setor privado investe cerca de 40% deste valor, enquanto que países desenvolvidos investem em média 3% do PIB com participação do setor privado por 75% desse investimento.

Além disso, verifica-se que algumas das regiões não se ligam a nenhuma outra, como apresentado nas análises das variáveis dependentes, o que pode ser verificado também através dos dados das variáveis explicativas de centralidade *closeness* (closeBB e closeBE) e *betweenness* (betBB e betBE) que assumem valores zeros em algumas regiões.

Os resultados das estatísticas analisadas na Tabela 5, com variáveis assumindo valores nulos, são indícios de que existem regiões brasileiras que não participam do processo de geração tecnológica e de inovação. Tal resultado é constantemente encontrado por trabalhos que investigam a concentração espacial de inovações, medidas por dados de patentes, no país evidenciando uma alta concentração dessas atividades na região sul e sudeste do país (ALBUQUERQUE *et al.*, 2002; GONÇALVES, 2007). É possível identificar que apenas as variáveis de participação sobre o PIB (pib_{pibB}), densidade populacional (aglomeração) e escala de aglomeração (aglomeração ao quadrado) possuem valores distintos de zero para todas as regiões.

A fim de analisar quais regiões se destacam ao se considerar cada uma das estruturas da variável dependente, número de ligações por patentes, a Tabela 6 reporta um *ranking* das 15 regiões que mais se ligam dentro de cada uma das redes consideradas, rede inter-regional e rede internacional ($LigaBB_{pat}$ e $LigaBE_{pat}$).

É possível identificar pelo *ranking* que as 15 regiões identificadas como aquelas que mais se ligam dentro da rede a outros nós, considerando o número de ligações por patentes geradas na região, correspondem a 43% do total das ligações da rede inter-regional e 68% das ligações da rede internacional.

Nota-se, no entanto, que não se encontram nesse *ranking* os grandes conglomerados industriais como São Paulo, Curitiba, Rio de Janeiro e Campinas. Esta evidência pode ser explicada por estas regiões serem capazes de gerar patentes sem a necessidade de colaborações com outras regiões, isto é, são mais autossuficientes em produzir invenções, possuindo mais ligações internas à região, em sua maioria. Contudo, verifica-se que as regiões

que se encontram bem posicionadas no *ranking* pertencem, em sua grande maioria à região sudeste.

Tabela 6 - *Ranking* das 15 regiões brasileiras com mais ligações por patentes (2001-2011)

	REGIC	<i>LigaBB_{pat}</i>	REGIC	<i>LigaBE_{pat}</i>
1°	Itabira	545,0	Lavras	42,0
2°	Lavras	347,0	Macaé	8,5
3°	Manaus	177,8	Viçosa	5,7
4°	Viçosa	172,0	Guaratinguetá	4,3
5°	Aracaju	169,8	Nova Prata	4,0
6°	Congonhas	143,0	Alfenas	4,0
7°	Ponte Grossa	127,0	Rio Verde	3,7
8°	Santa Maria	126,0	Joinville	3,5
9°	Ilhéus	124,0	Salvador	3,2
10°	Macaé	102,0	Limeira	2,7
11°	Rio Claro	101,5	Araraquara	2,2
12°	São Luís	93,5	Belo Horizonte	2,2
13°	Caratinga	90,0	Campinas	2,1
14°	Campina Grande	88,0	São Carlos	2,0
15°	Piracicaba	66,0	Cruzeiro	2,0
Percentual 15 primeiros		43%	68%	
Total de ligações na rede		5737	135,2	

Fonte: Elaboração Própria.

Um fato relevante na Tabela 6 é a presença de Manaus na 3ª posição do *ranking* de ligações inter-regionais, o que pode estar atrelado a esta ser uma zona franca o que propicia maior interação entre os agentes, podendo facilitar a ocorrência de colaborações. A região de Lavras também se destaca em ambos os *rankings*, o que pode ser explicado pelas colaborações universitárias. Já Itabira pode estar bem posicionada por meio das colaborações surgidas ou da presença de incubadoras ou de inventores que pertençam ao Vale do Aço. Todas as demais regiões presentes no *ranking* possuem estas mesmas características de facilitação de interação entre agentes ou a presença de instituições que colaboram para a ocorrência de colaborações entre inventores.

Já a Tabela 7, tem a finalidade de analisar o *ranking* de regiões que mais se ligam considerando a contagem absoluta de ligações do nó na rede (LigaBB e LigaBE).

Verifica-se que ao se considerar a contagem de ligações absolutas dos nós na rede há uma presença das regiões industriais brasileiras como São Paulo, Curitiba e Campinas bem como a região de São Carlos que se destaca como centro tecnológico, educacional e científico e onde se localiza o parque tecnológico que originou em 1984 a primeira incubadora da América do Sul, ParqTec (TORKOMIAN, 1994). Nota-se que as 15 regiões que mais se ligam dentro da

rede são responsáveis pela maioria das ligações, sendo este índice de 73% no caso da rede inter-regional e 89% da rede internacional.

Tabela 7 - Ranking das 15 regiões brasileiras com maior número de ligações absolutas (2001-2011)

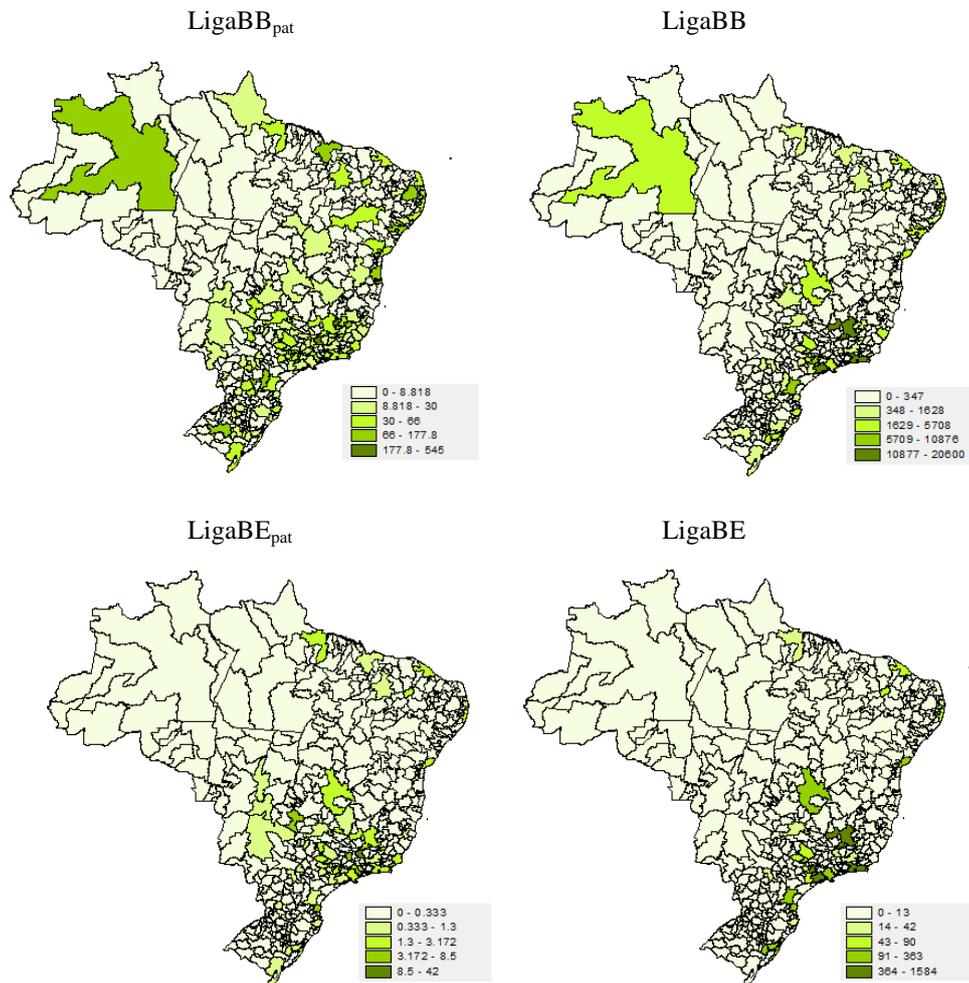
	REGIC	LigaBB	REGIC	LigaBE
1º	São Paulo	20600	São Paulo	1584
2º	Belo Horizonte	20308	Belo Horizonte	670
3º	Rio de Janeiro	17897	Rio de Janeiro	659
4º	Curitiba	10876	Campinas	597
5º	Campinas	10673	Joinville	363
6º	Porto Alegre	5708	Porto Alegre	224
7º	Florianópolis	4434	Salvador	203
8º	Recife	3417	São José dos Campos	179
9º	Salvador	3359	Curitiba	170
10º	São Carlos	2467	Brasília	158
11º	Ribeirão Preto	2450	Recife	90
12º	Aracajú	2377	São Carlos	89
13º	Brasília	2335	Araraquara	75
14º	São José dos Campos	2073	Fortaleza	70
15º	Araraquara	1971	Limeira	66
Percentual - 15 primeiros		73%	89%	
Total de ligações na rede		152618	5835	

Fonte: Elaboração Própria.

Logo, ao se comparar os dois *rankings*, Tabela 6 e Tabela 7, nota-se que ao se considerar a contagem de ligações absolutas do nó na rede, Tabela 7, tem-se redes com ligações mais concentradas em poucas regiões, verificado pelo percentual de ligações das 15 primeiras regiões, e tais regiões são destacadas na literatura brasileira como regiões de centros tecnológicos. Já no *ranking* da Tabela 6 identificam-se regiões em desenvolvimento em termos tecnológicos, isto é, estas regiões possuem potencial para se desenvolver tecnologicamente, mas ainda não são autossuficientes em gerar atividades inovativas sem a colaboração das demais regiões mais inovativas, sendo tais ligações menos concentradas ao se considerar o percentual de ligações no *ranking* em comparação a todas as regiões.

A Figura 8 traz um comparativo sobre a distribuição gradativa das ligações ao se levar em consideração o número de ligações por patentes geradas na região e o número total de ligações existentes no nó, nas redes inter-regionais e internacionais, para o período de análise de 2001-2011.

Figura 8 – Mapa das ligações das redes inter-regionais e internacionais (2001-2011)



Fonte: Elaboração Própria utilizando o *software* Arcview 3.2.

Por meio da Figura 8 é possível identificar na parte superior a comparação entre a distribuição das ligações nas redes internacionais ao se considerar o número de ligações por patentes geradas no nó (esquerda) e o número total de ligações na região (direita). Já na parte inferior compara-se a distribuição das ligações na rede internacional. É possível identificar que as regiões em que ocorrem mais ligações na rede inter-regional, partes mais escuras, ao se considerar o número de ligações por patentes geradas é mais espalhada espacialmente do que ao se considerar o número total de ligações. Nesta última, verifica-se uma maior concentração nas regiões sul e sudeste. Já na rede internacional, ao se considerar o número de ligações por patentes geradas, verifica-se que regiões menores, mais escuras, possuem mais ligações e também se encontram mais dispersas no mapa, enquanto ao se considerar o número total de

ligações, verifica-se uma maior ocorrência de ligações também na região sul e sudeste do país.

Por meio do Quadro 4, é possível identificar que o *ranking* de regiões com maiores valores de centralidades difere-se dentro da mesma rede ao se considerar medidas de centralidades distintas (*closeness* e *betweenness*). Verifica-se que regiões, como Rio de Janeiro e São Paulo, consideradas centros tecnológicos e grandes metrópoles, aparecem no *ranking* de maior centralidade *closeness* e não aparecem no *ranking* de centralidade *betweenness*. Essa evidência pode ser explicada por estas regiões serem consideradas nós que captam e transbordam fluxos de conhecimento com maior rapidez, mas não são consideradas nós intermediários, que servem de caminho para que uma outra região qualquer alcance os demais nós na rede. No entanto, verifica-se que as regiões destacadas são parecidas com as identificadas na Tabela 6 e 7 de *ranking* de regiões com maiores ligações dentro das redes, o que é um indício que as estruturas de centralidades são de fato favoráveis para a formação de laços entre regiões.

Quadro 4 – Ranking 15 regiões com maiores coeficientes de centralidades (2000-2010)

	Rede inter-regional		Rede internacional	
	REGIC <i>closeBB</i>	REGIC <i>betBB</i>	REGIC <i>closeBE</i>	REGIC <i>betBE</i>
1°	São Paulo	Jundiaí	Londrina	Macaé
2°	Rio de Janeiro	Vitória	São Paulo	São Luís
3°	Campinas	Rio Claro	Campinas	São J. dos Campos
4°	Porto Alegre	Catanduva	Rio de Janeiro	Curitiba
5°	Florianópolis	Joinville	São Carlos	Araraquara
6°	Ribeirão Preto	Guaratinguetá	Araraquara	Belém
7°	Fortaleza	Araranguá	Belém	Piracicaba
8°	Brasília	Aracajú	Porto Alegre	Belo Horizonte
9°	Londrina	Divinópolis	Curitiba	Viçosa
10°	Santa Rosa	Itabira	Sorocaba	São Carlos
11°	Recife	Patos de Minas	Viçosa	Sorocaba
12°	Carazinho	Cuiabá	Piracicaba	Salvador
13°	Goiânia	Piracicaba	São J. dos Campos	Jundiaí
14°	Manaus	Manaus	Salvador	Fortaleza
15°	Ilhéus/Itabuna	São J. do Rio Preto	Macaé	Salvador

Fonte: Elaboração Própria

4.3 Resultados das regressões da rede inter-regional

Esta seção visa analisar a ocorrência de ligações na rede inter-regional de copatenteamento brasileira, levando-se em consideração as variáveis dependentes de ligações por patente ($LigaBB_{pat}$), analisando os determinantes das ligações daquelas regiões cujo o

copatenteamento são de grande importância para seus níveis inventivos, e as ligações absolutas na rede (LigaBB).

Logo, as Tabelas 8 e 9 trazem, respectivamente, os resultados obtidos para os determinantes das ligações por patentes (LigaBB_{pat}), utilizando-se regressões em Painel Tobit por Efeitos Aleatórios; e do número absoluto de ligações (LigaBB), utilizando-se de regressões em Painel *Pooled* ZINB.

Visando gerar resultados robustos, as colunas de (1) a (3) representam, respectivamente: (1) as estimações considerando apenas as variáveis explicativas que consideram as estruturas regionais e as variáveis de controle; (2) apenas as variáveis explicativas que abrangem as estruturas de arquitetura de rede e variáveis de controle; e (3) consideram-se todas as variáveis do modelo. Como resultado das estimações sobre os determinantes das ligações das regiões na rede inter-regional, verifica-se através da Tabela 8 e 9, que as três colunas não tem os sinais e significâncias de suas variáveis alteradas. Para as estimações em Tobit também verifica-se a presença da coluna (4). Esta representa as estimações pelo método *bootstrap*.

A ausência de alteração das significâncias e sinais das variáveis pode ser uma evidência de que os efeitos do ambiente e das estruturas de posicionamento dos nós na rede coexistam e um conjunto de variáveis não altera o outro. Como a última coluna de cada tabela é composta por todas as variáveis, refletindo um modelo mais completo e robusto, esta será o foco da análise.

Primeiramente, ao se analisar a Tabela 8 que se refere ao número de ligações por patentes de cada região, objetiva-se captar o que determina a ocorrência das ligações daquelas regiões que inovam mais a partir de colaborações com outras localidades, do que produzindo inovações sozinhas. O número de ligações por patentes geradas internamente à região (LigaINT_{pat}) tem relação positiva com a ocorrência de ligações por patentes entre inventores na rede inter-regional, o que corrobora a Hipótese 1, que regiões cujos inventores se ligam internamente também tendem a se ligar a inventores de outras regiões brasileiras. Do ponto de vista das regiões com mais ligações internas, isso ser uma forma de não deixar que o conhecimento da região se torne obsoleto. Do ponto de vista das regiões mais periféricas, as ligações com as regiões com mais ligações internas podem complementar seu processo inventivo.

Tabela 8 - Determinantes das ligações inter-regionais por patentes (LigaBB_{pat}) no Brasil. Painel Tobit com Efeitos Aleatórios (2001-2011)

Variáveis	(1)	(2)	(3)	(4)
Log(LigaINT_{pat})		0,22*** (0,02)	0,19*** (0,02)	0,20*** (0,05)
Log(closeBB)		0,31*** (0,06)	0,23*** (0,06)	0,33*** (0,08)
Log(betBB)		25,34*** (7,34)	31,02*** (9,52)	21,16* (12,47)
Log(P&Dind)	3,06 (3,30)		1,80 (3,14)	5,09*** (2,09)
Log(P&Duniv)	676,52*** (43,54)		546,48*** (40,96)	519,87*** (56,00)
Log(Emp_{ind})	1,73** (0,84)		1,34** (0,76)	1,88** (0,53)
Log(Escala_{inv})	-11,79*** (3,07)		-14,25*** (2,86)	-15,16*** (1,86)
Log(Aglomeracão)	0,78*** (0,28)		0,68*** (0,25)	0,57*** (0,12)
Log(Aglomeracão ao quadrado)	-0,34** (0,13)		-0,30** (0,12)	-0,25*** (0,05)
Log(Bk_{pib})	0,07*** (0,01)		0,07*** (0,01)	0,06*** (0,01)
Log(Pib_{pibB})	12,64*** (3,95)		12,02*** (3,57)	14,12*** (2,42)
D_{reg}	0,07*** (0,02)	0,13*** (0,02)	0,06*** (0,02)	0,05*** (0,01)
D_{metro}	0,37*** (0,05)	0,41*** (0,05)	0,24** (0,05)	0,22*** (0,07)
_cons	-0,23*** (0,08)	-0,21*** (0,02)	-0,19*** (0,07)	-0,16*** (0,05)
Likelihood-ratio test sigma_u	149,45***	180,26***	52,94***	89,54***
LR chi	529,87***	380,47***	667,23***	
Log Like	-3662,04	-3736,74	-3593,36	-3577,63
Wald				13403,34***
Teste Hausman	chi ² = 88.67			

Fonte: Elaboração própria por meio do *software* Stata 12.

Nota¹: Erro padrão da estatística entre parênteses. Significância: ***0,01; **0,05; * 0,10.

Nota²: O *likelihood-ratio test sigma_u* corrobora a utilização do *xttobit*.

Nota³: Pelo teste Hausman não se rejeita a hipótese de que o modelo eficiente é o de efeitos aleatórios.

A estrutura de rede, centralidade *closeness* do nó (*closeBB*), obteve coeficiente positivo, indicando que quanto mais próximos são os nós na rede no sentido de rapidez de transferência do fluxo de conhecimento, maior é o número de ligações que possuem com outras regiões brasileiras. Esta centralidade é dada pelo caminho mais curto que conecta os nós, sendo que quanto mais central é o nó menor é sua distância para todos os outros nós.

Já a variável de centralidade *betweenness* (*betBB*), além de obter sinal positivo, mostrou-se como a variável de estrutura de rede mais influente para a ocorrência de ligações com o nó, o que vai de acordo com a ideia de que nós, quando considerados *hubs*, isto é, âncoras, tendem a atrair maiores números de ligações com os demais agentes funcionando como um nó

intermediário de novas ligações. Este resultado corrobora a Hipótese 2 de que as estruturas de centralidade dos nós são formas de torná-los centros de atração promovendo novas ligações com nós adjacentes. De acordo com Guan *et al.* (2015), um nó central é menos restrito dentro da rede o que indica que a centralidade é um fator que o torna mais passível de ligações.

Já a variável de P&D industrial foi significativa como determinante do número de ligações por patentes na rede inter-regional. Tal fato sugere que fazer P&D requer laços entre inventores de regiões diferentes. A variável de P&D universitário obteve um coeficiente positivo, além de se destacar com o coeficiente mais elevado em comparação com as demais variáveis, inclusive em comparação com o P&D industrial, o que pode ser explicado pelo grande número de investimentos no Brasil ainda ocorrer via importações de máquinas e equipamentos e o P&D industrial ainda ser baixo em algumas regiões. Estes resultados corroboram a Hipótese 3, de que regiões capazes de produzir P&D fomentam a ocorrência de laços fortes e são tendenciosas a possuir maior produção tecnológica a partir de colaborações em redes. Segundo Lastres e Cassiolato (2009), a inovação ocorre por meio de diversos tipos de colaborações entre agentes, sendo as novas políticas de incentivo a atividades de inovação voltadas para a promoção de atividades de P&D e difusão tecnológica.

A variável que mede o número de empregados na indústria (Emp_{ind}) obteve sinal positivo. Tal resultado evidencia que as indústrias tendem a participar de ligações em redes de colaborações para a geração de inovações. Grillitsch *et al.* (2015) ao analisar firmas de regiões periféricas da Suécia encontrou que as firmas não somente compensam, mas complementam a falta de absorção de conhecimento e capacidade em inovar da região, por meio de maior participação de seus empregados em redes de colaborações.

A escala de inventores ($escala_{inv}$) obteve um resultado que evidência existir uma relação negativa entre concentração de inventores e a ocorrência de ligações em redes de invenção inter-regional, o que corrobora a ideia de que quando agentes se encontram muito próximos tendem a tornar uma rede intrarregional densa e a diminuir as colaborações com inventores de outras localidades.

A aglomeração populacional obteve resultado positivo enquanto a aglomeração ao quadrado, um coeficiente com sinal negativo. Este resultado sugere que regiões com densidade populacional elevada tendem a possuir uma estrutura urbana que atrai inventores e facilita a existência de colaborações entre eles, além de corroborar a ideia de que a escala urbana

facilita e potencializa as ligações entre inventores (HOEKMAN; FRENKEN; VAN OORT, 2009; MAGGIONI; UBERTI, 2009; MIGUELEZ; MORENO, 2013). Porém, as vantagens dessa aglomeração urbana crescem a taxas decrescentes, incorrendo nas chamadas deseconomias de escala, como pode ser constatado pelo sinal negativo da variável aglomeração ao quadrado.

Já a variável que mede a participação da importação de bens de capital sobre o PIB da região (bk_{pib}) foi positiva. Como já mencionado, as importações de máquinas constituem mecanismos de absorção de fluxos de conhecimentos externos incorporados em produtos. Nota-se que há uma associação positiva com o número de colaborações existentes na rede. Este resultado sugere que esta forma de absorção de tecnologia, de conhecimentos externos à região, contribui para a ocorrência de laços por patentes na rede inter-regional.

A participação do nó sobre o PIB do país (Pib_{pibB}) também obteve sinal positivo, o que pode ser uma evidência que as regiões identificadas como aquelas que possuem maior número de ligações por patentes, mesmo sendo mais subdesenvolvidas em termos tecnológicos possuem um nível de participação econômica relevante, relacionando-se com a ideia de que regiões mais ricas tendem a ser mais inventivas, devido sua capacidade em investir e a presença de melhor infraestrutura, atraindo inventores de demais regiões para possíveis colaborações. Regiões mais avançadas economicamente possuem evidências de altas taxas inventivas (Ó HUALLACHÁIN; KANE; KENYON, 2015).

Já com base na Tabela 9, ao se considerar o número de ligações absolutas da rede inter-regional, isto é, sem considerar a relevância das colaborações nos níveis inventivos de cada região, as ligações internas à região por patentes geradas são significativas e obtêm sinal positivo, corroborando também a Hipótese 1 de que regiões com ligações internas (laços fortes) tendem também a se ligar com as demais regiões (laços fracos), o que corrobora a ideia de Granovetter (1973).

Além disso, a centralidade *closeness* também foi significativa, indicando que a posição do nó na rede tem uma influência positiva sobre a ocorrência de ligações. Assim, nós que possuem maior centralidade de proximidade tendem a se ligar mais. Já a variável centralidade *betweenness* obteve o maior coeficiente dentre as variáveis de redes. Este resultado pode ser um indício de que existe maior impacto na ocorrência de ligações quando um nó intermedia a

ligação de outros nós e possui a função de *hub* do que quando se é apenas um nó central de caminho mais curto (*closeness*).

Tabela 9 - Determinantes das ligações absolutas inter-regionais (LigaBB) no Brasil. Painel *Pooled ZINB* robusto (2001-2011)

Variáveis	<i>Pooled Robusto</i>		
	(1)	(2)	(3)
Log(LigaINT_{pat})		0,18*** (0,02)	0,14*** (0,02)
Log(closeBB)		0,74*** (0,09)	0,38*** (0,09)
Log(betBB)		30,82*** (3,32)	13,38*** (3,63)
Log(P&Dind)	18,14*** (4,22)		13,72*** (3,76)
Log(P&Duniv)	299,58*** (26,34)		215,84*** (26,59)
Log(Emp_{ind})	2,58** (1,33)		1,84* (1,01)
Log(Escala_{inv})	-0,73 (1,38)		-0,70 (1,26)
Log(Aglomeração)	-0,14 (0,61)		0,09 (0,63)
Log(Aglomeração ao quadrado)	0,11 (0,29)		-0,01 (0,31)
Log(Bk_{pib})	0,09*** (0,01)		0,08*** (0,02)
Log(Pib_{pibB})	2,81 (1,74)		1,24 (1,64)
D_{reg}	0,14*** (0,04)	0,01 (0,03)	0,12*** (0,03)
D_{metro}	0,37*** (0,03)	0,40** (0,03)	0,28*** (0,03)
_cons	0,24** (0,09)	-0,60*** (0,02)	0,16 (0,09)
Wald	1851,88***	1466***	2300,32***
Log pseudolike	-2366,11	-2396,83	-2339,17
/lnalpha	-204,86***	-18,19***	-24,06***

Fonte: Elaboração própria por meio do *software* Stata 12.

Nota: Erro padrão da estatística entre parênteses. Significância: ***0,01; **0,05; * 0,10.

Verifica-se que a variável de P&D industrial se mostra significativa e positiva. Tal resultado sinaliza que as regiões com capacidade de realizar P&D por parte de instituições privadas tendem a participar mais de colaborações durante o processo de produção da atividade inventiva. Já a variável de P&D universitário também obteve um coeficiente com sinal positivo para explicar a ocorrência de ligações entre inventores de uma região com inventores localizados em outras regiões brasileiras, sendo seu coeficiente superior a qualquer outra variável do modelo. Isso indica que pesquisas acadêmicas são relevantes para a geração de ligações externas à região durante o processo de produção tecnológica.

O número de empregados na indústria (Emp_{ind}) obteve um coeficiente com significância e sinal positivo para explicar a ocorrência de ligações. De acordo com Silva (2007), enquanto as patentes representam “*output*” dos esforços de pesquisas e investimentos de empresas, o número de colaborações das empresas seriam os insumos destes, e estariam atrelados ao desempenho final das mesmas. Tal fato sugere que as indústrias tendem a participar de forma ativa na geração de ligações em redes de colaboração, o que neste caso, pode ocorrer devido à proximidade em que se encontram grandes indústrias inovativas no Brasil, estas podem colaborar entre si na geração tecnológica e participar de redes a fim de complementar seu conhecimento local e se manterem competitivas no mercado.

A variável de escala de inventores não se mostrou significativa em explicar as ligações entre inventores localizados em regiões distintas do Brasil, nem a variável de participação econômica da região em relação ao PIB do Brasil, assim como as variáveis de aglomeração populacional e a escala de aglomeração. Já as importações de bens de capital sobre o PIB da região (BK_{pib}) obtiveram resultados significativos e positivos o que indica que regiões podem estar utilizando da importação de máquinas e conhecimentos externos como insumos de produção tecnológica.

Em ambas as Tabelas, 8 e 9, as variáveis de controle assumiram os sinais positivos esperados de que as regiões metropolitanas tendem a influenciar na ocorrência de ligações nas redes de copatenteamento inter-regional, enquanto a variável regional evidencia que regiões localizadas no sul e sudeste colaboram em atividades inventivas.

4.4 Resultados das regressões da rede internacional

Nesta seção serão reportados os resultados obtidos ao se analisar os determinantes das ligações na rede internacional de copatenteamento, entre inventores brasileiros e estrangeiros. Logo, as Tabelas 10 e 11 reportam os resultados obtidos, respectivamente, em relação aos determinantes de ligações por patentes, por meio do modelo Tobit de efeitos aleatórios, e das ligações absolutas nas redes internacionais através do modelo ZINB robusto. Deste modo, as colunas das Tabelas 10 e 11 possuem a mesma definição das Tabelas 8 e 9, contendo estimações considerando grupo de variáveis de interesse e visando a robustez dos resultados.

Podemos verificar pela Tabela 10, que a única variável de estrutura de rede que se mostrou significativa em explicar as ligações entre inventores brasileiros e estrangeiros foi o coeficiente da variável centralidade *betweenness*, o qual se mostrou com sinal positivo, corroborando a ideia de que nós com as características de *hubs* tendem a atrair e incorporar maiores fluxos de informações. Neste caso, são influentes na atração de colaborações estrangeiras.

Tabela 10 - Determinantes das ligações internacionais por patentes (LigaBE_{pat}) no Brasil. Painel Tobit com Efeitos Aleatórios (2001-2011)

Variáveis	(1)	(2)	(3)	(4)
Log(LigaINT _{pat})		0,01 (0,01)	0,01 (0,01)	0,01 (0,01)
Log(closeBE)		0,02*** (0,01)	0,01 (0,01)	0,01 (0,01)
Log(betBE)		4,65*** (0,98)	4,87*** (1,50)	5,14** (0,02)
Log(P&Dind)	2,01*** (0,47)		1,96*** (0,47)	2,28*** (0,81)
Log(P&Duniv)	30,01*** (5,70)		26,92*** (5,89)	26,87*** (15,88)
Log(Emp _{ind})	0,01 (0,11)		0,01 (0,11)	0,01 (0,03)
Log(Escala _{inv})	-0,97** (0,40)		-1,39*** (0,42)	-1,37*** (0,42)
Log(Aglomeraco)	0,04 (0,03)		0,03 (0,03)	0,02 (0,03)
Log(Aglomeraco ao quadrado)	-0,02 (0,02)		-0,01 (0,02)	-0,01 (0,01)
Log(Bk _{pib})	0,01* (0,01)		0,01* (0,01)	0,01** (0,01)
Pib _{pibB}	1,51*** (0,51)		1,44*** (0,51)	1,36*** (0,52)
D _{reg}	0,01 (0,01)	0,01 (0,01)	0,01 (0,01)	0,001 (0,002)
D _{metro}	0,01 (0,01)	0,02 (0,01)	0,01 (0,01)	0,005 (0,006)
_cons	-0,02** (0,01)	-0,01*** (0,01)	-0,02* (0,01)	-0,017** (0,008)
<i>Likelihood-ratio test sigma_u</i>	29,21***	54,25***	26,18***	0,43***
<i>LR chi</i>	155,89***	93,62***	166,95***	
<i>Log Like</i>	-6359,19	-6328,05	-6364,72	-6351,82
<i>Wald</i>				1982,40***
<i>Teste Hausman</i>	Chi ² = 55.67			

Fonte: Elaboraco prpria por meio do *software* Stata 12.

Nota¹: Erro padro da estatística entre parênteses. Significncia: ***0,01; **0,05; * 0,10.

Nota²: O *likelihood-ratio test sigma_u* corrobora a utilizaco do *xttobit*.

Nota³: Pelo teste Hausman no se rejeita a hiptese de que o modelo eficiente é o de efeitos aleatrios.

Tanto o P&D industrial quanto o P&D universitrio impactam positivamente a ocorrênca de laos entre inventores de regies brasileiras com estrangeiros. Evidenciando-se que regies com capacidade para produzir P&D tendem a produzir mais tecnologia e se ligar a demais

regiões, inclusive regiões estrangeiras, por meio de colaborações, transmitindo e absorvendo maiores conhecimentos.

O número de empregados na indústria não foi significativo em explicar a ocorrência de laços entre regiões brasileiras e estrangeiros, assim como as variáveis de densidade populacional e de escala de aglomeração.

Por outro lado, a escala de inventores demonstra ter uma relação negativa com as ligações estrangeiras, o que pode ser explicado pela existência da facilidade em gerar colaborações locais quando há muita disponibilidade de inventores na região havendo maiores fluxos de conhecimento localizado. Outros autores também identificaram a aglomeração de inventores como tendo um impacto negativo sobre redes como Lobo e Strumsky (2008) e Bettencourt *et al.* (2007).

As importações de máquinas, medidas por meio da variável de importações de bens de capital (bk_{pib}) mostraram-se com sinal positivo, porém com coeficiente bem pequeno para se explicar os laços com estrangeiros.

A participação da região sobre o PIB do país, também sugere que regiões mais desenvolvidas economicamente possam tender a se ligar mais com estrangeiros, por serem regiões que investem mais e possuem melhores infraestruturas se tornam mais atrativas a estes inventores.

As variáveis de controle e as ligações internas à região por número de patentes geradas pela mesma não foram significativas em explicar a ocorrência de ligações entre inventores brasileiros e estrangeiros, assim como a variável de estrutura de redes centralidade *closeness*.

Já ao se analisar a contagem de ligações em termos absolutos na rede internacional, que captam em sua grande maioria ligações geradas pelos grandes centros tecnológicos do Brasil, tem-se que as variáveis de centralidade de proximidade (*closeBE*) e de intermediação (*betBE*) que obtiveram sinais positivos, verificando-se que a centralidade *betweenness* se mostra com um coeficiente mais elevado em explicar as ligações quando comparada à centralidade *closeness*. Este resultado sugere que é mais importante os nós serem intermediários (*betweenness*) do que próximos (*closeness*), pois tendem, assim, a captar maiores ligações funcionando como *hubs*, resultado que também foi encontrado ao se analisar as redes inter-regionais.

A variável de P&D industrial e universitário obtiveram coeficientes positivos e de magnitude semelhante. Esse resultado confirma a hipótese da existência de uma relação entre investimentos em P&D e a ocorrência de maiores fluxos de conhecimento via ligações em redes.

Tabela 11 - Determinantes das ligações absolutas internacionais (LigaBE) no Brasil. Painel *Pooled* ZINB robusto (2001-2011)

Variáveis	<i>Pooled Robusto</i>		
	(1)	(2)	(3)
Log(LigaINT_{pat})		-0,01 (0,08)	0,07 (0,08)
Log(closeBE)		3,36*** (0,70)	2,05*** (0,63)
Log(betBE)		32,90*** (6,26)	11,95** (8,64)
Log(P&Dind)	131,16*** (30,90)		133,56*** (33,78)
Log(P&Duniv)	303,59** (146,65)		187,82* (148,62)
Log(Emp_{ind})	-29,97* (16,25)		-39,65** (17,39)
Log(Escala_{inv})	2,06 (4,05)		3,05 (4,01)
Log(Aglomeração)	-5,56 (3,51)		-5,19 (3,45)
Log(Aglomeração ao quadrado)	2,83 (1,73)		2,63 (1,70)
Log(Bk_{pib})	0,06 (0,09)		0,07 (0,09)
Log(Pib_{pibB})	-1,74 (5,49)		-4,62 (5,74)
D_{reg}	0,04 (0,17)	0,21 (0,15)	0,01 (0,17)
D_{metro}	0,45*** (0,12)	0,41*** (0,12)	0,33*** (0,12)
_cons	-1,26** (0,46)	-0,99*** (0,26)	-1,26*** (0,46)
Wald	220,34***	200,83***	252,56***
Log pseudolike	-823,89	-844,57	-816,33
/lnalpha	-16,64***	-16,61***	-16,64***

Fonte: Elaboração própria por meio do *software* Stata 12.

Nota: Erro padrão da estatística entre parênteses. Significância: ***0,01; **0,05; * 0,10.

A variável de emprego na indústria também foi significativa em explicar as ligações com estrangeiros, porém com sinal negativo, o que pode evidenciar que regiões com maior participação industrial tendem a se ligar menos a inventores estrangeiros em termos de geração de inovação. Talvez isso se dê pelo fato de o Brasil ter empresas que patenteiam inovações incrementais, que dispensam projetos de colaboração com o exterior.

De acordo com Silva (2007), colaborações internacionais remetem a riscos monetários e questões jurídicas, e em países emergentes como o Brasil ainda há falta de incentivos à ambientes de colaborações internacionais que poderiam ser gerados por uma estabilidade econômica, tramites legais mais ágeis e que geram maior credibilidade e incentivos a parcerias. As demais variáveis foram todas não significativas.

Por fim, constata-se que a variável de controle de regiões não foi significativa, enquanto a *dummies* de controle de metrópoles se mostraram como regiões relevantes para a ocorrência de ligações com agentes estrangeiros.

5 CONSIDERAÇÕES FINAIS

A literatura sugere que invenções quando geradas por colaborações tendem a produzir resultados mais eficientes e propensos a gerar inovações. Têm-se evidenciado que redes de copatenteamento são formas diretas de transferência de conhecimentos que incorporam tanto o conhecimento codificado quanto o tácito. Logo, é importante investigar quais são os fatores determinantes das ligações nas redes como forma de compreender e promover maiores fluxos de conhecimento, principalmente, fluxos de conhecimento internacionais.

Nessa perspectiva, o estudo analisou os determinantes e a dinâmica das ligações entre inventores nas redes de copatenteamento identificadas no Brasil, investigando tanto as ligações inter-regionais entre inventores brasileiros pertencentes a regiões distintas quanto a ocorrência de ligações em redes internacionais, entre inventores brasileiros e colaboradores estrangeiros.

A metodologia utilizada foi de inserir a análise da dinâmica inovativa do Brasil, mensurada a partir de dados de patentes, na perspectiva de análise de redes sociais. Para isto, identificam-se os nós das redes como as regiões em que se inserem os inventores que participam de colaborações e, seus laços (ligações) se caracterizam por meio de dados de copatenteamento.

Por meio dessa metodologia, foi possível analisar tanto a influência das estruturas das regiões que participam de redes, quanto das próprias estruturas de arquitetura dos nós nas redes, sob a ocorrência das ligações nas redes identificadas. Ademais, foram utilizadas técnicas econométricas de painel de dados, analisando-se tanto a razão entre a ocorrência de ligações (laços) quanto o total de patentes geradas pelo nó (região). Para tal, utilizaram-se os modelos de estimação Tobit e ZINB, visando explicar os determinantes regionais das ligações entre inventores de diferentes regiões.

Ambas as abordagens se utilizaram dos dados de copatenteamento que abrangem tanto as colaborações entre os inventores brasileiros quanto as produções tecnológicas que possuem tanto inventores brasileiros quanto estrangeiros, para o período de análise de 2001 a 2011 e considerando-se a abrangência geográfica de REGIC imediata.

Em termos de distribuição espacial das ligações no Brasil, constatou-se que a variável que mede a razão entre ligações e patentes geradas pelo nó possui valores maiores em regiões que estão em desenvolvimento tecnológico, se comparadas aos grandes centros tecnológicos

brasileiros. Isso sugere que estas regiões são mais dependentes de colaborações para a geração de atividades tecnológicas. Por outro lado, ao se considerar o número absoluto de ligações, as regiões consideradas grandes centros tecnológicos e educacionais do país são as que mais se destacam, fato que talvez possa estar atrelado aos grandes centros concentrarem também um maior número de inventores.

Além disso, em ambas as redes identificadas no Brasil, as variáveis que mais influenciam tanto na ocorrência de ligações por patentes quanto das regiões que possuem maior número de ligações em termos absolutos, considerando-se as redes inter-regionais são as mesmas: as variáveis de P&D, principalmente o P&D universitário, e a centralidade *betweenness*. Este resultado corrobora a hipótese 2 de que estruturas de centralidade tendem a gerar mais ligações entre seus nós adjacentes tornando-os centros de atração para novas colaborações. Isso demonstra que é importante manter investimentos sobre a capacidade em inovar dessas regiões identificadas como centrais, pois são responsáveis por transferir conhecimentos às demais regiões através de colaborações com inventores de outras localidades. Além disso, confirma-se também a hipótese 3 de que regiões com atividades voltadas para P&D, que é um insumo do conhecimento, tendem a participar mais de ligações em redes de colaboração, que são consideradas formas de transbordamento de conhecimento.

Contudo, outras variáveis impactam a ocorrência de ligações entre as regiões como escala de inventores, número de empregados na indústria, escalas de aglomeração urbana, participação da região sobre o PIB, importações de bens, número de ligações internas (ao nó), centralidade *closeness* e ligações internas a região. Esta última evidencia que a Hipótese 1 também é válida, de que regiões que possuem muitas ligações internas também tendem a possuir mais colaborações com agentes localizados em regiões distintas, por meio de laços fracos.

Ao se analisar a ocorrência de laços com estrangeiros verifica-se que a participação da indústria tem um impacto negativo sobre a ocorrência ligações nas redes internacionais. Este resultado contrastou com aquele da rede inter-regional, pois neste caso foi uma variável com impacto positivo sobre a geração de ligações.

Outra variável de destaque são as importações de bens de capital, que também se mostram como promotoras de ligações entre regiões brasileiras, corroborando a hipótese da existência de uma relação entre existência de laços e a importação de bens de capital. Este é um mecanismo de absorção de conhecimento externo.

Seguindo a ideia de que quanto maior o número de ligações existentes no nó, maior o fluxo de conhecimento dentro da região é importante que haja maior atenção em termos de políticas públicas ao incentivo de colaborações das regiões que concentram os maiores fluxos de conhecimento com as demais regiões brasileiras a fim de promover uma maior produção tecnológica em todo o país e não apenas manter foco nas regiões sul e sudeste como já destacado. Além do mais, devem-se explorar colaborações e analisar mais os determinantes das regiões que concentram as colaborações com estrangeiros, como São Paulo, Belo Horizonte, Rio de Janeiro, Campinas, Joinville e outras já citadas anteriormente, considerando que como estas regiões absorvem conhecimento por meio das redes de colaboração, seus agentes podem, posteriormente, transferir tais conhecimentos adquiridos para demais agentes localizados em outras regiões.

Partindo-se da ideia de que atividades capazes de gerar inovação sejam provenientes de relações entre os agentes e permite a disseminação de conhecimentos, principalmente os tácitos, é importante que novas políticas sejam voltadas para questões de redes. Deste modo, recentes políticas de incentivo a atividades de inovação buscam promover atividades de P&D e incentivo à difusão tecnológica, além da promoção de colaborações e cooperação entre agentes e regiões. No entanto, as políticas ainda são mais tímidas quanto o objetivo explícito em gerar redes de colaborações entre agentes de localidades distintas, havendo em geral cooperações locais.

A limitação deste trabalho ocorre em função da medida dos fluxos de conhecimento por meio de dados de patentes. Haveria outros fluxos não captados por dados de patentes, que estariam fora do alcance desta dissertação. Além disso, como os dados podem incorrer em possíveis *vieses* causados pela endogeneidade, devido à dificuldade em tratar problemas provenientes da própria estrutura do ambiente, procurou-se amenizar esses problemas através de defasagens das variáveis explicativas dos modelos. Reconhece-se que isso não resolve o problema. Entretanto, as correlações estatísticas encontradas continuam válidas.

Em termos de extensão e trabalhos futuros pode-se investigar nas regiões quais nós, agentes e instituições, que funcionam como *gatekeepers*, isto é, nós que seriam responsáveis por criar pontes com demais agentes localizados em redes distintas ou incluir novas variáveis, inclusive novas medidas de arquitetura de redes.

REFERÊNCIAS BIBLIOGRÁFICAS

- ADEGBESAN, J. A.; HIGGINS, M. J. The intra-alliance division of value created through collaboration, **Strategic Management Journal**, v. 32, n.2, p. 187-211, 2011.
- AGRAWAL, A.; COCKBURN, I.; MCHALE, J. Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships. **Journal of Economic Geography**, v. 6, n.5, p. 571–591, 2006.
- ALBAGLI, S. Informação e geopolítica contemporânea: o papel dos sistemas de propriedade intelectual. **Informare**, vol. 3, n. 1/2, Rio de Janeiro: PPCI/IBICT, 1997.
- ALBERT, R.; BARABASI, A. L. Statistical mechanics of complex networks. **Reviews of Modern Physics**, v. 74, n.1, p. 47–99, 2002.
- ALBUQUERQUE E. M.; SIMÕES, R.; BAESSA, A.; CAMPOLINA, B.; SILVA, L. Distribuição espacial da produção científica e tecnológica brasileira: uma descrição de estatísticas de produção local de patentes e artigos científicos. **Revista Brasileira de Inovação**, v. 1, n. 2, p.225-251, 2002.
- ALMEIDA, Eduardo. **Econometria Espacial Aplicada**. Editora: Alinea, 2012.
- ALMEIDA, P.; KOGUT, B. The exploration of technological diversity and the geographic localization of innovation, **Small Business Economics**, v. 9, n. 1, p. 21–31, 1997.
- ALMEIDA, P.; KOGUT, B. Localization of knowledge and the mobility of engineers in regional networks. **Management science**, v. 45, n. 7, p. 905-917, 1999.
- ANSELIN, L.; VARGA, A.; ACS, Z. Local geographic spillovers between university research and high technology innovations. **Journal of urban economics**, v. 42, n. 3, p. 422-448, 1997.
- ARAÚJO, B. C.; CAVALCANTE L. R.; 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**, v. 5, p 16-21, 2009.
- ARAÚJO, Veneziano de Castro. **Dimensão local da inovação no Brasil: determinantes e efeitos de proximidade**. 2014. 188f. Tese (Doutorado em Engenharia de Produção). Universidade de São Paulo (USP), São Paulo, 2014.
- ARAUJO, V. de C.; GARCIA, R. Local determinants of innovation and spatial dependence - A spatial Tobit model applied to brazilian micro-regions. In: Encontro Nacional de Centros de Pós-graduação em Economia, Foz do Iguaçu, 2016. **Anais**, n. 161, 2016.
- ARORA, A.; GAMBARDELLA, A. The changing technology of technological change: general and abstract knowledge and the division of innovative labour. **Research policy**, v. 23, n. 5, p. 523-532, 1994.

ARROW, K. Economic welfare and the all location of resources for inventions. In: NELSON, R.R. (Eds.). **The Rate and Direction of Innovative Activity: Economic and social factors**. Princeton: Princeton University Press, p. 609–625, 1962.

AUDRETSCH, D. B.; FELDMAN, M. P. R&D spillovers and the geography of innovation and production. **The American Economic Review**, v. 86, n.3, p. 630–640, 1996.

AUDRETSCH, B. Agglomeration and the location of innovative activity. **Oxford review of economic policy**, v. 14, n. 2, p. 18-29, 1998.

BARABÁSI, A. L.; ALBERT, R. Emergence of scaling in random networks, **Science**, v. 286, n.5439, p. 509–512, 1999.

BARNES, J. A. Class and committees in a Norwegian island parish. New York. **Human Relations**, v. 7, n. 1, p. 39-58, 1954.

BATHELT H.; MALMBERG A.; MASKELL P. Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. **Progress in Human Geography**, v. 28, n.1, p. 31–56, 2004.

BALCONI, M.; BRESCHI, S.; LISSONI, F. Networks of innovators and the role of academia: an exploration of Italian patent data, **Research Policy**, v. 33, n. 1, p. 127–145, 2004.

BEAUDRY, C.; SCHIFFAUEROVA, A. Who’s right, Marshall or Jacobs? The localization versus urbanization debate. **Research Policy**, v. 38, n. 2, p. 318–337, 2009.

BETTENCOURT, L.; LOBO, J.; HELBING, D.; KÜHNERT, C; WEST, G. B. Growth, innovation, scaling, and the pace of life in cities. **Proceedings of the National Academy of Sciences of the United States of America**, v.104, n. 17, p. 7301–7306, 2007.

BOSCHMA, R. Proximity and innovation: a critical assessment. **Regional Studies**, v. 39, p. 61-74, 2005.

BOSCHMA, R.; ERIKSSON, R.; LINDGREN, U. How does labour mobility affect the performance of plants? The importance of relatedness and geographical proximity. **Journal of Economic Geography**, v. 9, n. 2, p. 169-190, 2009.

BRESCHI S.; LISSONI F. Knowledge spillovers and local innovation systems: a critical survey. **Industrial and Corporate Change**, v. 10, n. 4, p. 975–1005, 2001.

BRESCHI, S.; LISSONI, F. Knowledge networks from patent data. **Handbook of quantitative science and technology research**, Springer Netherlands, p. 613-643, 2004.

BRESCHI, S.; LISSONI, F. “Cross-Firm” Inventors and Social Networks: Localized Knowledge Spillovers Revisited. **Annales d'Economie et de Statistique**, p. 189-209, 2005.

BRESCHI S.; LISSONI F. Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. **Journal of Economic Geography**, v. 9, n.4, p. 439–468, 2009.

- BRESCHI, S; CATALINI, C. Tracing the links between science and technology: An exploratory analysis of scientists' and inventors' networks. **Research Policy**, v. 39, n. 1, p. 14-26, 2010.
- BRODER, A.; KUMAR, R.; MAGHOUL, F.; RAGHAVAN, P.; RAJAGOPALAN, S. STATA, R. TOMKINS, A; WIENER, J. Graph structure in the web. **Computer Network**, V. 33, n.1, p. 309–320, 2000.
- BROEKEL, T.; BOSCHMA, R. Knowledge networks in the Dutch aviation industry: the proximity paradox. **Journal of Economic Geography**, v. 12, n. 2, p. 409-433, 2012.
- BURT, R. S.; TALMUD, L. Market niche, **Social Network**, v. 15, n. 2, p. 133–149, 1993.
- BURT, R. S. The contingent value of social capital. **Administrative Science Quarterly**, v. 42, n. 2, p. 339–365, 1997.
- BURT, R. S. Structural holes and good ideas. **American journal of sociology**, v. 110, n. 2, p. 349-399, 2004.
- CAMAGNI, R. P. The concept of innovative milieu and its relevance for public policies in european lagging regions. **Papers in Regional Science**, v. 74, n.4, p. 317–340, 1995.
- CANTNER, U., GRAF, H. The network of innovators in Jena: an application of social network analysis. **Research Policy**, v. 35, n. 4, p. 463–480, 2006.
- CANTNER, U.; MEDER, A. Regional and technological effects of cooperation behavior. **Jena economic research papers**, n. 14, 2008.
- CAPALDO, A. Network structure and innovation: The leveraging of a dual network as a distinctive relational capability. **Strategic management journal**, v. 28, n. 6, p. 585-608, 2007.
- 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.
- CARVALHO, J. R.; LAVOR, S. C. Repeat property criminal victimization and income inequality in Brazil. **Revista Economia**, 2008.
- CASAROTTO, Nelson Filho; PIRES, Luís Henrique. **Redes de pequenas e médias empresas e desenvolvimento local: estratégias para a conquista da competitividade global com base na experiência italiana**. Editora: Atlas, São Paulo, 1999.
- CASSI, Lorenzo; PLUNKET, Anne. Co-patenting and inventive performance: in search of the proximity paradox. **Conference**, DRUID Society, Conhepagen, Denmark, 2012.
- CASSIMAN, B.; VEUGELERS, R. In search of complementarity in innovation strategy: Internal R&D and external knowledge acquisition. **Management science**, v. 52, n. 1, p. 68-82, 2006.

CASTELLS, M. **The information age: economy, society and culture.** The rise of the network society, v.1, Blackwell Publishers Inc: Oxford, 1996.

CHESBROUGH, H. W. The era of open innovation. **Managing innovation and change**, v. 127, n. 3, p. 34-41, 2006.

COE, D. T.; HELPMAN, E. International R&D spillovers, **European Economic Review**, v. 39, n. 5, p. 859–887, 1995.

COHEN, A. C. Estimation in Mixtures of Discrete Distributions. **Statistical Pub. Society**, Proceedings of the International Symposium on Discrete Distributions, Montreal, Quebec, 1963.

COOKE, P.; URANGA, M. G.; ETXEBARRIA, G. Regional innovation systems: Institutional and organisational dimensions. **Research policy**, v. 26, n. 4-5, p. 475-491, 1997.

COOKE, Philip. **Knowledge economies: Clusters, learning and cooperative advantage.** Routledge, 2002.

COOKE P. Regional innovation systems: general findings and some new evidence from biotechnology clusters. **The Journal of Technology Transfer**, v. 27, n. 1, p. 133–145, 2002.

COWAN, R.; JONARD, N. Network structure and the diffusion of knowledge. **Journal of Economic Dynamics and Control**, v. 28, n. 8, p. 1557–1575, 2004.

DALEY, D. J.; GANI, J. **Epidemic Modelling.** Vol. 25, Cambridge University Press, Cambridge, 2001.

DANTAS, M. Informação-valor e corporações-redes: elementos para discutir um novo padrão de acumulação. **Informare**, vol. 3, n. 1-2, PPCI/IBICT: Rio de Janeiro, 1997.

DAVIS Jr, C. A. Aumentando a eficiência da solução de problemas de caminho mínimo em SIG. **PRODABEL.** Empresa de Informática e Informação do Município de Belo Horizonte, 1997.

DING, W. W.; LEVIN, S. G.; STEPHAN, P. E.; WINKLER, A. E. The impact of information technology on academic scientists productivity and collaboration patterns. **Management Science**, v. 56, n. 9, p. 1439-1461, 2010.

DOREIAN, P.; BATAGELJ, V.; FERLIGOJ, A. Generalized blockmodeling of two-mode network data. **Social networks**, v. 26, n.1, p. 29-53, 2004.

DOSI, G. The nature of the innovative process. **Technical change and economic theory**, v. 2, p. 590-607, London/New York: Pinter Publishers, 1988.

EASLEY, David; KLEINBERG, Jon. **Networks, crowds, and markets: Reasoning about a highly connected world.** Cambridge University Press, 2010.

EBEL, H.; MIELSCH, L; BORNHOLDT, S. Scale-free topology of e-mail networks, **Physical Review**, v 66, n. 3, 2002.

EJERMO, O.; KARLSSON, C. Interregional inventor networks as studied by patent coinventorships. **Research Policy**, v. 35, n. 3, p. 412–430, 2006.

ERNST, D. Network transactions, market structure and technology diffusion: implications for south-south cooperation. In: MYTELKA, L.K. (Ed.), **South-south co-operation in a global perspective**. Paris: OECD, p.89- 124, 1994.

FAGERBERG, J. Innovation: a guide to the literature. In: **The Oxford handbook of innovation**. Oxford University Press: New York, 2005.

FISCHER, M. M.; VARGA, A. Spatial knowledge spillovers and university research. **The Annals of Regional Science**, v. 37 n. 2, p. 303–322, 2003.

FLEMING, L.; COLFE, L.; MARIN, A.; MCPHIE, J. Why the valley went first: Agglomeration and emergence in regional inventor networks. **Academy of Management Conference**, New Orleans, LA. 2004.

FLEMING, L.; KING, C.; JUDA, A. I. Small worlds and regional innovation. **Organization Science**, v. 18, n. 6, p. 938–954, 2007.

FREEMAN, C.; HAGEDOORN, J. Catching up or falling behind: patterns in international interfirm technology partnering. **World Development**, v.22, n.5, p.771-780, 1994.

FREEMAN, L. C.; ROEDER, D.; MULHOLLAND, R. R. Centrality in social networks: II. Experimental results. **Social networks**, v. 2, n. 2, p. 119-141, 1979.

FREEMAN, R. B. Working for nothing: The supply of volunteer labor. **Journal of Labor Economics**, v. 15, n.1, p. 140-166, 1996.

FREITAS, Leandro de Quintanilha. **Medidas de centralidade em grafos**. 2010. 103f. Dissertação (Mestrado em Engenharia de Produção). Universidade Federal do Rio de Janeiro, Rio de Janeiro (UFRJ), Rio de Janeiro, 2010.

GERTLER, M. S. Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there). **Journal of Economic Geography**, v. 3, n. 1, p. 75–99, 2003.

GILSING, V.; NOOTEBOOM, B.; VANHAVERBEKE, W.; DUYSTERS, G; OORD, A. V. D. Network embeddedness and the exploration of novel technologies: technological distance, betweenness centrality and density. **Research Policy**, v. 37, n. 10, p. 1717–1731, 2008.

GLAESER, E. L. Learning in cities. **Journal of urban Economics**, v. 46, n. 2, p. 254-277, 1999.

GONÇALVES, E. O padrão espacial da atividade inovadora brasileira: uma análise exploratória. **Estudos Econômicos**, v. 37, n.2, p. 405-433, 2007.

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

GONÇALVES, E.; FAJARDO, B. A. G. A influência da proximidade Tecnológica e geográfica sobre a Inovação regional no Brasil. **Revista Econômica Contemporânea**, v. 15, n. 1, p. 112-142, 2011.

GONÇALVES, E.; ARAÚJO Jr., I. F.; TAVEIRA, J. G. Identificando local buzz e global pipelines nas regiões brasileiras: Uma análise das redes de coinventores de patentes. In: **Encontro Nacional de Centros de Pós-Graduação em Economia - ANPEC**, Foz do Iguaçu, 2016.

GRANOVETTER M. S. The strength of weak ties. **American journal of sociology**, v.78, n. 6, p.1360–1 380, 1973.

GRANOVETTER, M. Economic action and social structure: the problem of embeddedness. **The American Journal of Sociology**. v. 91, n. 3, p. 481–510, 1985.

GRANOVETTER M. The impact of social structure on economic outcomes. **The Journal of Economic Perspectives**, v. 19, n. 1, p. 33–50, 2005.

GRANT, R. M. Prospering in dynamically-competitive environments: organizational capability as knowledge integration. **Organization science**, v. 7, n. 4, p. 375-387, 1996.

GRILICHES, Z. Issues in assessing the contribution of research and development to productivity growth. **Bell Journal of Economics**, v.10, n. 1, p. 92-116, 1979.

GRILICHES, Z. Patent statistics as economic indicators: a survey. **Journal of Economic Literature**, v. 28, n. 4, p. 1661-1707, 1990.

GRILICHES, Z. The search for R&D spillovers. **The Scandinavian Journal of Economics**, v. 94, 29-47, 1991.

GRILLITSCH, M.; NILSSON, M. Innovation in peripheral regions: do collaborations compensate for a lack of local knowledge spillovers. **The Annals of Regional Science**, v.54, n.1, p. 299-321, 2015.

GROSSMAN, Gene M.; HELPMAN, Elhanan. **Innovation and growth in the global economy**. MIT Press, Cambridge, 1993.

GROSSMAN, G. M.; HELPMAN, E. Trade, knowledge spillovers, and growth. **European Economic Review**, v.35, n. 2-3, p. 517-526, 1990.

GUAN, J.; ZHANG, J.; YAN, Y. The impact of multilevel networks on innovation. **Research Policy**, v. 44, n. 3, p. 545-559, 2015.

GUIMARÃES, E. N; FARIA, G. Integração e desenvolvimento regional: uma proposta de regionalização em Minas Gerais. Belo Horizonte: Cedeplar/UFMG, 2006. Disponível em <http://www.cedeplar.ufmg.br/seminarios/seminario_diamantina/2006/D06A026.pdf>. Acesso em: 10 de Novembro 2016.

GUSSO, D. Agentes da inovação: quem os forma, quem os emprega? In: DE NEGRI, J. A. *et al.* (Orgs.). **Tecnologia, exportação e emprego**. p. 397-444. Brasília: IPEA, 2006

HANSEN, M. T. The search-transfer problem: the role of weak ties in sharing knowledge across organization subunits, **Administrative Science Quarterly**, v. 44, n. 1, p. 82–111, 1999.

HARARY, F.; NORMAN, R. Z.; CARTWRIGHT, D. **Structural models: An introduction to the theory of directed graphs**. New York: Wiley, 1965.

HE, J.; FALLAH, M. H. Dynamics of Inventor Networks and the Evolution of Technology Clusters. **International Journal of Urban and Regional Research**, v.38, n.6, p. 2174-2200, 2014.

HERSTAD, S. J.; ASLESEN, H. W.; EBERSBERGER, B. On Industrial Knowledge Bases, Commercial Opportunities and Global Innovation Network Linkages. **Research Policy**, v. 43, n. 3, p. 495-504, 2014.

HOEKMAN, J.; FRENKEN, K.; OORT, F. V. The geography of collaborative knowledge production in Europe. **The Annals of Regional Science**, v. 43, n.3, p. 721-738, 2009.

HOPPE, B. **Introduction to Network Mathematics**. Boston University, 2008. Disponível em: <<http://webmathematics.net/>> Acesso em: 20 de Março de 2016.

HU, A. G. Z; JAFFE, A. B. Patent citation and international knowledge flows: the case of Korea and Taiwan, **International Journal of Industrial Organization**, v. 21, n.6, p. 849–880, 2003.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **REGIC-Regiões de Influência das Cidades**. Rio de Janeiro: IBGE, 2008. Disponível em: <<http://www.ibge.gov.br/home/geociencias/geografia/regic.shtm?c=7>> Acesso em: 17 de Abril de 2016.

IMBRIANI, C.; REGANATI, F. International efficiency spillovers into the Italian manufacturing sector. **Economia Internazionale**, v. 50, n. 4, p. 583–595, 1997.

JACKSON, Matthew O. **Social and economic networks**, v. 3. Princeton: Princeton university press, 2008.

JAFFE, A. B. Real effects of academic research. **The American Economic Review**, v. 79, n. 5, p. 957-970, 1989.

JAFFE, A. B.; TRAJTENBERG, M.; HENDERSON, R. Geographic localization of knowledge spillovers as evidenced by patent citations. **The Quarterly Journal of Economics**, v. 108, n.3, p. 577- 598, 1993.

JAFFE, Adam B.; TRAJTENBERG, Manuel. **Patents, Citations and Innovation**. Cambridge: The MIT Press, 2002.

JOHNSON, N. L.; KOTZ, S. Developments in Discrete Distributions, 1969-1980, Correspondent Paper. **International Statistical Review**, v. 50, n. 1, p. 71-101, 1982.

KANG, M. J.; PARK, J. Analysis of the partnership network in the clean development mechanism. **Energy policy**, v.52, p. 543-553, 2013.

KELLER, W. International trade, foreign direct investment, and technology spillovers. In: HALL, B. H., ROSENBERG, N. (Eds.), **Handbook of the Economics of Innovation**, v. 2, p. 793-829. Amsterdam: Elsevier, 2010.

KNOBEN J.; OERLEMANS L. A. G. Proximity and inter organizational collaboration: a literature review. **International Journal of Management Reviews**, v. 8, n. 2; p. 71–89, 2006.

KRUGMAN, Paul. **Geography and trade**, v.1. The MIT Press, Cambridge, MA, 1991.

LAWLESS, J. F. Negative binomial and mixed Poisson regression. **Canadian Journal of Statistics**, v. 15, n.3, p. 209-225, 1987.

LASTRES, H. M. M.; CASSIOLATO, J. E. Sistemas de inovação: políticas e perspectivas. **Parcerias Estratégicas**, v. 5, n. 8, p. 237-256, 2009.

LEE, G. Direct versus indirect international R&D spillovers, **Information Economics and Policy**, v. 17, n. 3, p. 334–348, 2005.

LEE, G. The effectiveness of international knowledge spillover channels. **European Economic Review**, v. 50, n. 8, p. 2075–2088, 2006.

LEE, M.; KIM, K.; CHO, Y. A study on the relationship between technology diffusion and new product diffusion. **Technological Forecasting and Social Change**, v. 77, n. 5, p. 796–802, 2010.

LERMA, J. R.; ÁLVAREZ, J. Reconocimiento de caracteres basado en firmas vectoriales. **e-Gnosis**, v. 7, 2009.

LIST, Friedrich. **National System of Political Economy**. JB Lippincott & Company, Philadelphia, 1856.

LOBO, J.; STRUMSKY, D. Metropolitan patenting, inventor agglomeration and social networks: a tale of two effects. **Journal of Urban Economics**, v. 63, n. 3, p. 871–884, 2008.

LUNDEVALL, B. **National innovation system: towards a theory of innovation and interactive learning**. Pinter, London, 1992.

MAGGIONI, M. A.; UBERTI, T. E. Networks and geography in the economics of knowledge flows. **Quality & quantity**, v. 45, n.5, p. 1031-1051, 2011.

MAGGIONI, M. A.; UBERTI, T. E.: Inter-regional knowledge flows in Europe: an econometric analysis. In: FREKEN, K. (ed.) **Applied Evolutionary Economics and Economic Geography**, p. 230–255. Edward Elgar Publishing, Cheltenham, 2007.

MAGGIONI, M. A.; UBERTI, T. E. Knowledge networks across Europe: which distance matters? **The Annals of Regional Science**, v. 43, n. 3, p. 691–720, 2009.

MALDONADO, J. Tecno-globalismo e acesso ao conhecimento. **Informação e globalização na era do conhecimento**. p.105 – 121. Rio de Janeiro: Campus, 1999.

MALMBERG, A.; MASKELL, P. Localized learning revisited. **Growth and Change**, v. 37, n.1, p. 1-18, 2006

MARSDEN, P. V. Recent developments in network measurement. In: CARRINGTON, P. J.; SCOTT, J.; WASSERMAN, S. (ed.) **Models and methods in social network analysis**. Cap 2. P. 8-30, New York: Cambridge University Press, 2005.

MARSHALL, Alfred. **Principles of Economics: Unabridged eight edition**. (1980) Cosimo classics, New York, 2009.

MATOS, R. E. S. A contribuição dos imigrantes em áreas de desconcentração demográfica do Brasil contemporâneo. **Revista Brasileira de Estudos de População**, Campinas, v.19, n.1, p.49-72, 2002.

MATOS, Ralfo; BRAGA, Fernando. Migração e Rede Urbana: procedências e inserção ocupacional. **XIII Encontro Nacional de Estudos Populacionais**, Ouro Preto, 2002.

MAURSETH, P. B.; VERSPAGEN, B. Knowledge spillovers in Europe: a patent citations analysis. **The Scandinavian journal of economics**, v. 104, n. 4, p. 531-545, 2002.

MEJDALANI, A.; GONCALVES, E.; RODRIGUEZ, R. S. Por que cooperar? Os determinantes da formação de conexões inter-regionais na rede brasileira de patentes. In: **XIV Encontro Nacional da Associação Brasileira de Estudos Regionais e Urbanos**, 2016, Aracaju, SE, ENABER, 2016.

MIAOU, S. The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. **Accident Analysis & Prevention**, v. 26, n. 4, p. 471-482, 1994.

MIGUÉLEZ, E.; MORENO, R. Research networks and inventors' mobility as drivers of innovation: evidence from Europe. **Regional Studies**, v.47, n.10, p. 1668-1685, 2013.

MINISTÉRIO DA CIÊNCIA E TECNOLOGIA. **Relatório anual da utilização dos incentivos fiscais: Ano Base 2008**. Brasília, 2010. Disponível em: http://www.mct.gov.br/upd_blob/0214/214324.pdf Acesso em: 18 de Abril de 2016.

MORENO, J. L. Sociogram and sociomatrix: A note to the paper by Foryth and Katz. **Sociometry**, v.1, p. 348–349, 1946.

MORENO, R.; PACI, R.; USAI, S. Geographical and sectoral clusters of innovation in Europe. **The Annals of Regional Science**, v. 39, n. 4, p. 715-739, 2005.

MORENO, R.; PACI, R.; USAI, S. Spatial spillovers and innovation activity in European regions. **Environment and Planning A**, v. 37, n.10, p. 1793-1812, 2005b.

NAGAOKA, S.; MOTOHASHI, K.; GOTO, A. Patent statistics as an innovation indicator. In: **Handbook of the economics of innovation**, v. 2, p. 1083-1127, Elsevier, [s.l.], 2010.

NEWMAN, M. E. J. The structure and function of complex networks. **SIAM Review**, v. 45, n. 2, p. 167–256, 2003.

NONAKA, I. A dynamic theory of organizational knowledge creation. **Organization Science**, v. 5, n. 1, p.14–37, 1994.

Ó HUALLACHÁIN, B.; KANE, K.; KENYON, S. Invention in the United States City System. **Annals of the Association of American Geographers**, v.105, n.6, p.1300-1323, 2015.

OLIVEIRA, Hélio Carlos Miranda de. **Em busca de uma proposição metodológica para os estudos das cidades médias: reflexões a partir de Uberlândia (MG)**. 2008. 187f. Dissertação. (Mestrado em Geografia). Instituto de Geografia. Universidade Federal de Uberlândia (UFU). Uberlândia, 2008.

OLIVEIRA, Priscila Medeiros. **Determinantes territoriais da capacidade de invenção no Brasil: O papel das redes de inventores**. 2016. 111f. Dissertação. (Mestrado em Economia Aplicada). Universidade Federal de Juiz de Fora (UFJF), Faculdade de Economia. Juiz de Fora, 2016.

MORTENSEN, P. S.; BLOCH, C. W. **Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data**, 3rd edition. OECD: Paris, 2005.

ORSENIGO, L.; PAMMOLLI, F.; RICCABONI, M. Technological change and network dynamics: lessons from the pharmaceutical industry. **Research policy**, v.30, n.3, p. 485-508, 2001.

PANTZALIS, C. Does location matter? An empirical analysis of geographic scope and MNC market valuation. **Journal of International Business Studies**, v. 32, n. 1, p. 133-155, 2001.

PATEL, P.; PAVITT, K. Large firms in the production of the world technology: An important case of "non-globalization. **Journal of International Business Studies**, v. 22, n.1, p. 1-21, 1991.

PEREIRA, Breno Augusto Diniz. **Estruturação de Relacionamentos Horizontais em Redes**. 218 f. Tese (Doutorado em Administração) - Programa de Pós-Graduação em Administração, Escola de Administração, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, 2005.

PEREIRA, F. M.; LEMOS, M. B. Cidades médias brasileiras: características e dinâmicas urbano-industriais. Repositório do Conhecimento do IPEA. **Pesquisa e Planejamento Econômico - PPE**, v.33, n.1, 2003. Disponível em: http://repositorio.ipea.gov.br/bitstream/11058/4099/1/PPE_v33_n01_Cidades.pdf. Acesso em: 25 de Maio de 2016.

PEREZ, C. Technological revolutions, paradigm shifts and socio-institutional change. In: REINERT, E. S. (ed.). **Globalization, economic development and inequality: An alternative perspective**, Chapter 7, p. 217-242, New Horizons in institutional and evolutionary economics. Northampton, MA, USA, 2004.

PHELPS, C.; HEIDL, R. WADHWA, A. Knowledge, Networks, and Knowledge Networks: a review and research agenda. **Journal of Management**, v. 38, n. 4, p. 1115-1166, 2012.

PICCI, L. The internationalization of inventive activity: A gravity model using patent data. **Research Policy**, v.39, n.8, p.1070-1081, 2010.

PISANO, G. P. The governance of innovation: vertical integration and collaborative arrangements in the biotechnology industry. **Research Policy**, v. 20, n.3, p. 237-249, 1991.

PONDS, R., OORT, F. V.; FRENKEN, K. Innovation, spillovers and university–industry collaboration: an extended knowledge production function approach. **Journal of Economic Geography**, v. 10, n. 2, p. 231-255, 2010.

POLANYI, Michael. **The tacit dimension**. The University of Chicago Press. Londres: Routledge and Kegan Paul, 1967.

POOT, T.; FAEMS, D.; VANHAVERBEKE, W. I. N. Toward a dynamic perspective on open innovation: A longitudinal assessment of the adoption of internal and external innovation strategies in the Netherlands. **International Journal of Innovation Management**, v. 13, n. 2, p. 177-200, 2009.

POWELL, W. Neither market nor hierarchy: Network forms of organization. In: **Research in Organizational Behaviour**, p. 295-336, Elsevier Science, 1990.

POWELL W. W., KOPUT K W, SMITH D. L. Inter organizational collaboration and the locus of innovation: networks of learning in biotechnology. **Administrative Science Quarterly**, v. 41, n 1, p. 116–145, 1996.

POWELL, W. W; OWEN-SMITH, J. Knowledge networks as channels and conduits: The effects of spillovers in the Boston biotechnology community. **Organization science**, v. 15, n.1, p. 5-21, 2004.

ROMER, P. Endogenous technological change. **Journal of Political Economy**, v. 98, p. 71-102, 1990.

ROSENBERG, N. Uncertainty and technological change. In: LANDAU, R.; TAYLOR, T.; WRIGHT, G. (ed.) **The Mosaic of Economic Growth**. p. 334-356. Stanford University Press: CA, 1996.

ROTHWELL, R.; FREEMAN, C.; HORLSEY, A.; JERVIS, V. T. P.; ROBERTSON, A. B.; TOWNSEND, J. Updated – project SAPPHO phase II. **Research Policy**, v. 3, n.3, p. 258-291, 1974.

ROXENHALL, T. Network structure and innovation in strategic innovation networks. **International Journal of Innovation Management**, v. 17, n. 02, 2013.

SCHUMPETER, J. A. **Teoria do desenvolvimento econômico: uma investigação sobre lucros, capital, crédito, juro e o ciclo econômico**. São Paulo: Abril Cultural, 1982.

SEN, P.; DASGUPTA, S.; CHATTERJEE, A.; SREERAN, P. A.; MUKHERJEE, G.; MANNA, S. S. Small-world network properties of the Indian railway network. **Physical Review**, v 67, n.3, 2003.

SILVA, da D. H. Cooperação internacional em ciência e tecnologia: oportunidades e riscos. **Revista Brasileira de Política Internacional**, v. 50, n.1, p. 5-28, 2007.

SIMÕES, R.; OLIVEIRA, A.; GITIRANA, A.; CUNHA, J.; CAMPOS, M.; CRUZ, W. A Geografia da Inovação: uma Metodologia de Regionalização das Informações de Gastos em P&D no Brasil. **Revista Brasileira de Inovação**, v. 4, n. 1, p. 157-185, 2009.

SINGH, J. Collaborative networks as determinants of knowledge diffusion pattern. **Management Science**, v. 51 n. 5, p. 756–770, 2005.

SINGH, J., FLEMING, L. Lone inventors as sources of breakthroughs: Myth or reality? **Management Science**, v. 56, n. 1, p. 41-56, 2010.

SINGH, S. A note on inflated Poisson distribution. **Annals of mathematical statistics**. vol. 33, n. 3. Institute Mathematical Statistics, Hayward, CA, 1963.

SCHMITZ, H. Industrialisation strategies in less developed countries: some lessons of historical experience. **The Journal of Development Studies**, v. 21, n. 1, p. 1-21, 1984.

STONEMAN, P.; DIEDEREN, P. Technology diffusion and public policy. **The Economic Journal**, v.104, n. 425, p.918-930, 1994.

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

TAKAHASHI, Vania Passarini. **Capacidades tecnológicas e transferência de tecnologia: estudo de múltiplos casos da indústria farmacêutica no Brasil e no Canadá**. 2002. 223f. Tese (Doutorado em Engenharia Mecânica) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2002.

TORKOMIAN, A. L. V. Fundação ParqTec: o órgão gestor do Pólo de Alta Tecnologia de São Carlos. **Ciência da Informação**, v. 23, n. 2, 1994.

TOBIN, J. Estimation of relationships for limited dependent variables. **Econometrica**, v. 26, n. 1, p. 24–36, 1958.

TÖDTLING, F.; LEHNER, P.; TRIPPL, M. Innovation in knowledge intensive industries: The nature and geography of knowledge links. **European Planning Studies**, v. 14, n. 8, p. 1035-1058, 2006.

VERSPAGEN, B. Innovation and economic growth. In: FAGERBERG, J.; MOWERY, D. C.; NELSON, R. R. (ed.) **The Oxford handbook of innovation**, p.487-513, Oxford University Press, 2005.

VIOTTI, Eduardo Baumgratz. **Passive and active national learning systems**. 1997. 310f. tese. (Doutorado em Política e Ciência Social). New School for Social Research, [s.l] 1997.

UZZI, B. The Source and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect. **American Sociological Review**, v.61, n. 4, p. 674-698, 1996.

VUONG, Q. H. Likelihood ratio tests for model selection and non-nested hypotheses. **Econometrica**, v. 57, n. 2, p. 307-333, 1989.

WAL, T.; ANNE, L. J.; BOSCHMA, R. A. Applying social network analysis in economic geography: framing some key analytic issues. **The Annals of Regional Science**, v. 43, n. 3, p. 739-756, 2009.

WATTS, D.; STROGATZ, S. H.: Collective dynamics of “small-world” networks. **Nature**, v. 393, p. 440–442, 1998.

WASSERMAN, Stanley; GALASKIEWICZ, J. **Advances in social network analysis: Research in the social and behavioral sciences**. v. 171. A Sage Publications: London, 1994.

WASSERMAN, Stanley; FAUST, K. **Social network analysis: Methods and applications**. v. 8. Structural Analysis in the social sciences. Cambridge university press, Cambridge: UK. 1994.

WOOLDRIDGE, Jeffrey M. **Econometric analysis of cross section and panel data**. MIT press, Cambridge, Massachusetts, 2010.

WOOLDRIDGE, Jeffrey M. **Introductory Econometrics: A modern approach**, v. 4, South-Western Cengage Learning: USA, 2009.

XIANG, X. Y.; CAI, H.; LAM, S.; PEI, Y. L. International knowledge spillover through co-inventors: An empirical study using Chinese assignees' patent data. **Technological Forecasting and Social Change**, v. 80, n.1, p. 161-174, 2013.

XU, X.; LEE, L. "Maximum likelihood estimation of a spatial autoregressive Tobit model." **Journal of Econometrics**, v. 188, n.1, p. 264-280, 2015.

ZANIBONI, N.; MONTINI, A. Modelos de Poisson inflado de zeros e binomial negativo inflado de zeros na previsão de sinistro de automóveis. **Revista Economia & Gestão**, v. 15, n. 41, p. 159-180, 2015.

ANEXOS

ANEXO A - Tabela Regiões a nível REGIC com cidades sede de metrópoles

Nome da Região	Código REGIC	Código REGIC
	161	482
Manaus	130260	130260
Belém	150140	150140
São Luís	211130	211130
Fortaleza	230440	230440
Natal	240810	240810
Recife	261160	261160
Maceió	270430	270430
Salvador	292740	292740
Belo Horizonte	310620	310620
Ipatinga ¹ , Coronel Fabriciano, Timóteo	313130	313130
Vitória	320530	320530
Rio de Janeiro	330455	330455
Campinas	350950	350950
São Paulo	355030	355030
Santos	354850	354850
Maringá	411520	411520
Londrina	411370	411370
Curitiba	410690	410690
Joinville	420910	420910
Blumenau	420240	420240
Itajaí ²	-	420820
Florianópolis	420540	420540
Criciúma	420460	420460
Porto Alegre	431490	431490
Goiânia	520870	520870
Brasília	530010	530010

Nota¹: No código 313130 (Regic 161 e 482) Ipatinga é a região segundo o IBGE que contém a sede metropole, contudo na agregação Regic a nomenclatura também incorpora outras regiões.

Nota²: A região de Itajaí não foi agrupada em uma regic separada no nível de agregação de Regic 161.

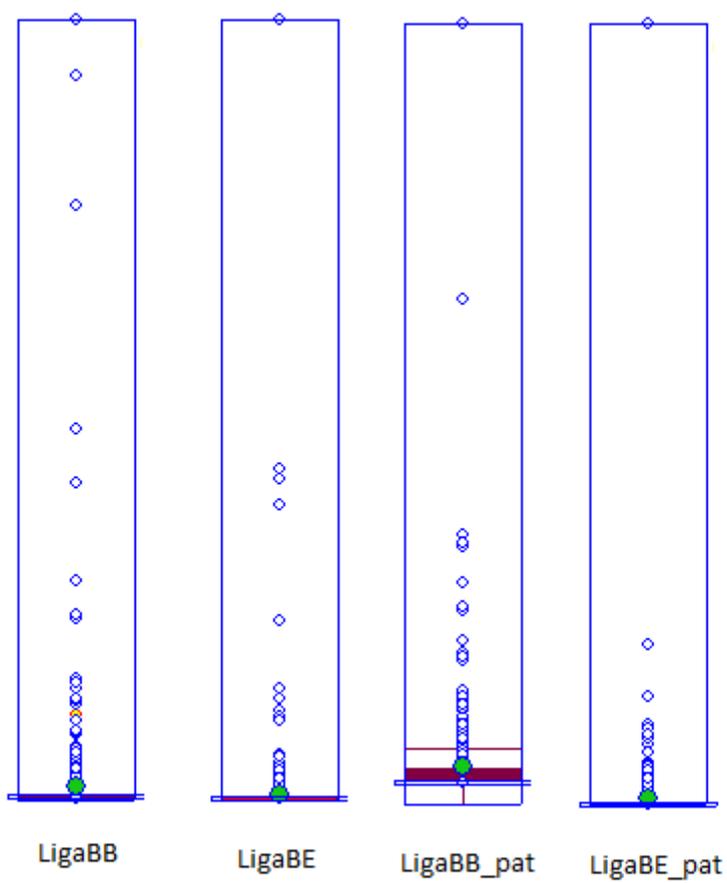
Fonte: IBGE (2008)

APÊNDICE

APÊNDICE A - Matriz de Correlação entre as variáveis de interesse (2000-2011)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1) LigaINT_{pat}	1												
(2) Ped_{univ}	,35	1											
(3) Ped_{priv}	,40	,20	1										
(4) Emp_{ind}	,20	,05	,20	1									
(5) Escala_{inv}	,21	,14	,27	,02	1								
(6) Aglomeração	,42	,17	,22	,26	,30	1							
(7) Aglomeração ao quadrado	,42	,17	,21	,26	,29	,99	1						
(8) Bk_{pib}	,51	,23	,45	,32	,28	,43	,42	1					
(9) Pib_{pibB}	,27	,14	,30	,01	,97	,33	,32	,30	1				
(10) close_{BB}	,61	,24	,35	,31	,22	,53	,52	,49	,27	1			
(11) close_{BE}	,59	,39	,38	,11	,36	,40	,39	,51	,41	-	1		
(12) bet_{BB}	,35	,26	,35	-,01	,85	,36	,35	,35	,85	,31	-	1	
(13) bet_{BE}	,22	,16	,29	-,01	,96	,29	,28	,28	,94	-	,39	-	1

Fonte: Elaboração a partir de saída do Stata 6.

APÊNDICE B – *Box-plot* de verificação de *outliers* das redes (2001-2011)

Fonte: Elaboração a partir de saída do *software* Geoda.