

44° ENCONTRO NACIONAL DE ECONOMIA (ANPEC 2016)

Área ANPEC: Área 10 - Economia Regional e Urbana

Vizinhança setorial na indústria de transformação brasileira

Milene Simone Tessarin (FEA/USP)

Paulo César Morceiro (FEA/USP)

André Luis Squarize Chagas (FEA/USP)

Resumo

A inovação ocorre devido ao investimento direto da empresa, motivada por interesses econômicos, como lucro, parcela de mercado, economias de escopo, etc. Essas decisões podem ser influenciadas por choques indiretos que ocorrem nas indústrias mais próximas, como uma inovação na indústria automobilística que pode motivar mudanças em sua cadeia produtiva. O objetivo deste trabalho é avaliar os efeitos de pares sobre os indicadores de inovação na indústria de transformação brasileira. Para isso, foi proposto um novo modo de medir a proximidade das indústrias. Consideramos os produtos típicos produzidos por uma indústria no subsetor de origem e em outros subsetores. Estes subsetores são considerados vizinhos porque eles usam a mesma base produtiva e tecnológica. A proximidade setorial foi construída através de uma “matriz de diversificação setorial” detalhada de empresas produtoras de bens em um ou mais subsetores (de um total de 103 setores) em 2013. Os dados foram fornecidos pelo IBGE por meio de uma solicitação especial, os quais formaram uma matriz-W relacionando os setores um a um. Como uma *proxy* de inovação considerou-se uma razão entre o número de engenheiros empregados e o número total de trabalhadores em cada subsetor. Este indicador é uma *proxy* reconhecida sobre os esforços inovativos das empresas, pois inovação e intensidade tecnológica estão relacionadas com competências do pessoal técnico empregado. Para testar a hipótese de efeitos de pares consideramos o teste I de Moran e usamos os testes LM e LM robusto para escolher o melhor modelo econométrico que confirme a dependência espacial setorial. Os resultados sugerem a existência de fortes efeitos de pares setoriais entre os subsetores com o mesmo nível tecnológico (baixa/média-baixa e alta/média-alta tecnologia). Além disso, existe um efeito de vizinhança fraca entre ramos de diferentes níveis tecnológicos. Portanto, o nível tecnológico de um setor é um bom indicador das suas relações de vizinhança setorial. Estes resultados podem ser úteis para formuladores de políticas públicas focarem em setores que geram maiores efeitos de *spillovers* sobre a estrutura produtiva e inovativa do país. Entendemos que este trabalho faz uma contribuição para o uso da análise de econometria espacial para além da vizinhança geográfica. Ele também lança luz sobre uma agenda de pesquisa fértil atualmente inexistente na literatura de organização industrial.

Palavras-chave: Vizinhança setorial, Inovação, Econometria espacial.

Código JEL: C23, O3, L14

Abstract

The innovation in industry occurs due to direct investment of the firm, motivated by economic interests, as profit, market share, economies of scope, etc. However, same these decisions can be influenced by indirect shocks occurring in closest industry, as an innovation in automobile industry motivating changes in its chain production. The aim of this paper is to evaluate the peer effects on the innovation indicators in Brazilian manufacturing industry. For this, we proposed a new way to measure the proximity of the industries. We consider the typical goods produced by an industry in the main subsector and in other subsectors. These sub-sectors are considered neighbors because they use the same technological and production bases. The sectorial proximity was building through a detailed “sectorial diversification matrix” of firms producing goods in one or more subsector (of a total 103 subsectors of the Brazilian manufacturing) in 2013. Brazilian Institute of Geography and Statistics (IBGE) provided the data from a specific request, special to this work. The data composed a W-matrix relating one by one sector. As an innovation proxy, we considered the number of engineers employed divided by the total number of

employees in each subsector. This indicator is a proxy recognized of the innovative efforts of companies, because innovation and technological intensity are related to technical staff' skills. To test the peer effects hypothesis we considered the Moran's I test, and we used the LM and the LM robust tests to choose a best econometric model that confirms the sectorial spatial dependence. The results suggest the existence of strong sectorial peer effects among subsectors with the same technological level (low and medium-low, and high and medium-high). Moreover, there is a weak neighborhood effect between subsectors of different technological levels. Therefore, the technological level of a sector is a good indicator of their relations of sectorial neighborhood. These results can be useful for policymakers assign public policies focused on subsectors generating greater spillovers effects on the production and innovation structure. We think this work is a contribution to the use of spatial econometric analysis beyond the geographical neighborhood. It also sheds light on a fertile research agenda, which is currently absent in the industrial organization literature.

Key-words: Sectorial neighborhood, Innovation, Spatial econometrics.

1. Introdução

Diversos fatores fazem com que firmas e, conseqüentemente, setores produtivos possuam uma relação de proximidade de modo a se influenciarem mutuamente. Dentre esses fatores estão a tecnologia, o conhecimento, uso de insumos e matérias-primas em comum, compartilhamento de fornecedores e canais de distribuição, etc. Atualmente, com a evolução da informática e telecomunicações, os transbordamentos de conhecimento e tecnológicos não estão mais limitados por fronteiras geográficas e análise de *clusters* setoriais tornou-se incapaz de explicar as relações existentes entre setores distantes geograficamente. Neste contexto, o conhecimento e a distância cognitiva entre setores ganharam evidência, pois deram indicações de porque os *spillovers* são mais intensos entre alguns setores e como se revela a relação de proximidade entre eles.

Algumas classificações setoriais tentam identificar essa vizinhança utilizando como critério a tecnologia e mostram-se úteis para vários propósitos analíticos. Contudo, buscaremos avaliar a existência de vizinhança entre setores por uma ótica mais detalhada, e para tanto será proposto como critério de vizinhança o fato de uma unidade local produtiva fabricar bens típicos do seu subsetor de origem e também de outros subsetores que são considerados vizinhos porque dependem da mesma base produtiva e tecnológica.

O objetivo deste artigo será avaliar a existência de *vizinhança setorial* na indústria de transformação brasileira no período recente. A principal questão a ser investigada é se os subsetores industriais possuem subsetores industriais vizinhos assim como os territórios geográficos? Se sim, como ocorre essa relação de vizinhança?

Verificou-se que a abordagem da proximidade setorial¹ ainda é pouco estudada, embora haja alguns trabalhos antigos (Marshall, 1890; Penrose, 1959; Jacobs, 1969) e outros recentes (Boschma e Iammarino, 2005; Frenken, van Oort e Verburg, 2007; Hidalgo et al, 2007) que abordaram temas conexos, mas com enfoque ligeiramente diferente. Dada a inexistência de trabalhos sobre esse assunto e a limitação com respeito aos dados, a análise realizada aqui será exploratória. Este estudo possui a vantagem de trabalhar com uma matriz de vizinhança setorial industrial bastante desagregada para o Brasil², a partir da qual foi possível explorar as relações de vizinhança setorial através de análises descritivas e métodos de econometria espacial.

2. Conceitos de proximidade setorial

O conceito de *proximidade* é muito utilizado para tratar de distância geográfica entre agentes, principalmente nos estudos que tratam de aglomerações econômicas. Marshall (1890) foi um dos primeiros autores a apontar fatores que estimulam a concentração geográfica das firmas baseados na presença de trabalhadores especializados, na oferta de insumos não facilmente comercializáveis e no

¹ Não encontramos nenhum trabalho com o termo “vizinhança setorial” ao fazer uma busca no Google Acadêmico.

² A matriz de vizinhança foi obtida com apoio dos técnicos do IBGE, aos quais agradecemos o trabalho.

fluxo de informações que circula melhor a pequenas distâncias. Weber (1929) detalhou a existência, por um lado, de fatores aglomerativos como a proximidade com outras empresas, clientes e fornecedores e o acesso a condições de infraestrutura mais favoráveis e, por outro lado, fatores desaglomerativos, como custo da terra, poluição e dificuldade em se locomover devido a congestionamentos. Esses são também alguns dos fatores que os autores da *New Economic Geography* (Krugman, 1991; Fujita, Krugman e Venables, 1999) chamaram de forças centrípetas e centrífugas, ou seja, forças que promovem a concentração geográfica, como as definições clássicas *marshallianas* sobre fontes de economias externas à firma, e as que desestimulam a concentração espacial, essencialmente devido a fatores imóveis e outras deseconomias externas puras. Von Thünen (1826), Christaller (1935), Lösch (1940) e Isard (1956) apresentaram como ponto focal a importância conferida aos custos de transporte para a decisão da firma quanto à sua localização ótima. Essas são algumas importantes referências que analisaram a proximidade de firmas, inclusive no nível setorial, do ponto de vista geográfico ao explicar os motivos que as firmas consideram para decidir sua localização.

Embora seja mais comum, a proximidade não se restringe à relação de vizinhança geográfica. Diversos autores citam outras possíveis formas de proximidade, apesar delas ainda serem difíceis de ser evidenciadas por meio de dados e estudos quantitativos (Frenken, van Oort e Verburg, 2007; Hausmann et al, 2011), já que esta análise exige o levantamento de inúmeras variáveis a respeito de tecnologias, capacidade produtiva, mercados, produtos, setores, entre outros itens. Torre e Gilly (2000) destacaram que a proximidade organizacional tem função importante para mostrar a familiaridade entre os atores, e envolve relações entre pessoas, instituições e na organização da produção (por exemplo, um setor pode demandar as mesmas instituições que outro setor como uma certificadora ou órgão de normalização). Boschma (2005) fez um levantamento teórico de quatro dimensões da proximidade (além da geográfica): a *proximidade cognitiva* está relacionada à base de conhecimento comum entre agentes, formada pelo conhecimento acumulado (Dosi, 1988) e a capacidade de absorção (Cohen e Levinthal, 1990) para entender e assimilar novos conhecimentos; a *proximidade organizacional* abrange interações de várias naturezas com atores conectados por partilhar o mesmo espaço e conhecimento; a *proximidade social* inclui relações entre agentes no nível microeconômico baseadas em confiança, parcerias e experiência; e a *proximidade institucional* está associada a relações de interação no nível macroeconômico conduzidas por normas e regras de conduta.

Neste trabalho enxergamos a *proximidade setorial* quando as firmas usam a mesma base produtiva e tecnológica para produzir produtos de mais de um subsetor industrial. Neste caso, esses subsetores são vizinhos por partilharem a mesma base produtiva e tecnológica.³ Assim, é coerente considerar que a proximidade setorial captura e condensa alguns tipos de proximidade conforme os conceitos de Boschma (2005).

Esse entendimento está intimamente relacionado com a visão de diversificação produtiva das firmas (Penrose, 1959), uma vez que esta promove sua diversificação *intra* e *inter* setorial através da combinação de fatores produtivos e tecnológicos que detêm, dentre eles, conhecimento, capital, insumos, trabalho etc.

O conceito de variedade relacionada, desenvolvido na virada do século, é muito útil e é uma boa aproximação teórica da proximidade setorial. *Variedade relacionada* é definida como “setores que estão relacionados em termos de competências complementares ou partilhadas” (Boschma e Iammarino, 2009, p. 292-293) que para tanto demandam comunicação e aprendizado interativo. Nooteboom (2000); Boschma e Iammarino (2005); Frenken, van Oort e Verburg (2007); Quattraro (2010); Castaldi, Frenken e Los (2015) utilizam dados de firmas industriais para medir o impacto da variedade relacionada sobre a produtividade, o emprego e a inovação dos países, encontrando efeitos positivos para todos eles. A variedade relacionada foi mensurada por meio de uma medida de entropia para os subsetores industriais como forma de captar a variedade dentro de cada subsetor. No entanto, apesar de destacar os efeitos benéficos da variedade e de argumentar que alguns setores são mais relacionados que outros, nenhum desses trabalhos buscou explicar como a proximidade setorial ocorre e quais setores se relacionam com

³ Isso será mais bem exposto no decorrer do texto e na seção metodológica de construção da matriz de vizinhança.

maior intensidade. Esses autores não testaram a dependência setorial, pois a forma de mensuração da variedade relacionada é calculada por uma *proxy* obtida pela medida de entropia.

Dado essa lacuna na teoria, nosso enfoque recairá em evidenciar a vizinhança existente entre setores produtivos e indicar quais são os vizinhos mais próximos ao *formalizar* a ideia de que alguns setores facilitam a existência de outros por compartilharem conhecimentos, bases produtivas, estruturas organizacionais etc. Entende-se que há similaridade lógica entre setores em termos tecnológicos e produtivos resultante de capacitações acumuladas no tempo (Lundvall, 1992; Cohen e Levinthal, 1989; 1990; Breschi e Malerba, 1997), as quais propiciam um ambiente favorável para firmas avançarem para estruturas produtivas mais diversificadas (Penrose, 1959). Considera-se que as firmas são os vetores que promovem a diversificação essencialmente através do uso da tecnologia e do conhecimento. Desta forma traçamos um paralelo entre firmas, tecnologias e diversificação para chegar à vizinhança setorial. Define-se, portanto *vizinhança setorial* como setores que, por possuírem algumas características compartilhadas, estão mais próximos produtiva e tecnologicamente. Essa proximidade ocorre devido à *spillovers* de conhecimento e tecnologia que desencadeiam oportunidades para firmas estabelecerem a produção em outros setores.

É possível identificar um fio condutor que orienta o processo de mudança tecnológica (Malerba e Orsenigo, 1993) e delimita as áreas possíveis de se estabelecer uma relação de proximidade. Este fio é composto por um conjunto de competências comuns a diversos setores que levam em direção a uma pauta de produção diversificada (Boschma e Iammarino, 2009; Frenken, van Oort e Verburg, 2007; Castaldi, Frenken e Los, 2015). Mas para obter sucesso neste processo de mudança, deve-se ir subindo um degrau por vez da escada tecnológica, deixando, como citou Hirschman (1981), que “uma coisa leva a outra”.

Considerando uma analogia feita por Hidalgo et al. (2007, p. 482), para crescer e promover o desenvolvimento, as firmas precisam realocar capital (físico, humano e institucional) para áreas diferentes daquelas que estão movendo-se de áreas menos nobres para outras mais nobres. Da mesma forma, os macacos precisam buscar áreas com maior oferta de alimentos para crescerem, mas para chegar até elas precisam dar saltos entre as árvores. Entretanto, se as árvores são muito esparsas e pequenas, ficam impossibilitados de se locomoverem pela floresta e acabam aprisionados numa área mais deserta. No caso das firmas, se o conjunto de produtos feitos por ela for tão disperso a ponto de não estar ao alcance de se comunicar, dificilmente ocorrerá um salto tecnológico e as áreas continuarão desconectadas. Essas áreas necessariamente precisam ter alguma conexão, mesmo que não seja explícita, por exemplo, entre produtos de setores distintos como produtos de madeira e papel – que utilizam a mesma base produtiva e o mesmo insumo – ou entre produtos distintos do mesmo setor, como produtos de carne e laticínios – que necessitam de equipamentos de transporte refrigerados, utilizam os mesmos canais de distribuição e dependem das mesmas instituições de vigilância sanitária.

Penrose (1959) aponta dois fatores que revelaria a proximidade entre os setores, o primeiro é a base produtiva e tecnológica, que envolve cada tipo de atividade produtiva que utiliza máquinas, processos, qualificações e matérias-primas; a combinação entre eles demanda competências específicas, que mescladas de formas diferentes, podem ser utilizadas na produção de itens com similaridade tecnológica. O segundo fator apontado é a área de mercado que indica cada grupo de clientes que a firma espera poder alcançar utilizando uma mesma estratégia; assim competências pré-estabelecidas num determinado mercado (ou produto) podem ampliar o campo de atuação conforme a firma define sua estratégia e suas limitações. A combinação desses dois fatores permite conhecer diversas possibilidades de atuação de uma firma, podendo-se, então, ter variações dentro da área de especialização ao ampliar a produção de produtos que utilizam a mesma base tecnológica ou então seguir para áreas mais distantes gerando três alternativas (Penrose, 1959, p. 177): utilizar a *mesma base* produtiva para produzir produtos novos e atuar em *novos mercados* (diversificação para longe); mudar a base produtiva para produzir novos bens e expandir a atuação em um mercado já existente (diversificação próxima); ou ainda, utilizar uma nova base produtiva para produzir novos produtos para novos mercados. Os dois primeiros referem-se à vizinhança setorial, se os produtos para onde ocorre a diversificação são de subsetores distintos (intra ou inter setor). O terceiro, por envolver uma nova base produtiva, é a diversificação do ponto de vista da empresa, não relacionado com a vizinhança setorial como estamos abordando. Todas essas possibilidades estão embasadas necessariamente em um firme domínio de seu campo de especialização incluindo certos

tipos de tecnologias, recursos e também de mercados. Caso contrário, uma “miscelânea de recursos em diversos campos” (Penrose, 1959, p. 215) pode prejudicar o desempenho da firma que não terá sucesso ao expandir para outros setores muito avessos aos que têm conhecimento.

Jacobs (1969) também tratou da proximidade entre setores, porém com um enfoque urbano (ou regional) em que defendeu que a proximidade de empresas, mesmo que diversas, importa para os seus ganhos, produzindo externalidades de diversificação.⁴ Na sua visão, o crescimento de firmas e países decorre da adição de novo trabalho a outros já existentes, por meio de um processo em que “coisas levam explicitamente a outras coisas” (Jacobs, 1969, p. 59). Contudo, nossa proposta de proximidade setorial abarca questões mais relacionadas à *spillovers* tecnológicos que não exigem necessariamente proximidade espacial, não trabalharemos com seu conceito.

Schumpeter (1911) foi seminal ao dizer que a partir de *novas combinações* surgem as inovações, que é um caso bem sucedido de diversificação. Nestas definições nota-se que há uma semelhança entre os argumentos de Penrose (1959) para quem também são as combinações (de áreas de mercado e base tecnológica) que levam à crescente diversificação de produtos e ao crescimento da firma.

Hidalgo et al (2007) e Hausmann et al (2011) destacaram como fator que define a proximidade setorial a base de conhecimento para fabricação dos produtos. Para captar a proximidade, os autores sugerem o mote “*espaço do produto*” – similarmente à Penrose (1959) – e sugerem que os produtos estão conectados por uma mesma base de conhecimento, assim, produtos que possuem bases próximas tem maior chance de serem considerados próximos no espaço do produto. Os produtos que possuem uma base de conhecimento mais ampla são considerados centrais na rede produtiva (por exemplo, produtos de metal, eletrônicos, ferramentas e máquinas e químicos, conforme Hausmann et al, 2011, p. 45) e são eles que dispõem de maiores oportunidades de induzir a produção de outros produtos que demandam capacitações similares (Hausmann et al, 2011; Hidalgo et al, 2007; Boschma e Iammarino, 2009). No agregado, há um conjunto de *core competences* através do qual promove-se melhoramentos contínuos e inovações relativas ao produto (Granstrand, Patel e Pavitt; 1997, p. 18). Quanto mais densa uma rede produtiva ou mais diversificada a produção entre os setores, menor é o processo de *descoberta de custos* (Hausmann e Rodrik, 2003), o qual refere-se a todos os custos necessários para se produzir um novo produto. Quanto menor for a descoberta de custo mais se tem uma espécie de economia de escopo entre os setores produtivos, uma vez que um setor aproveita de condições já estabelecidas por outro para o seu funcionamento⁵.

A criação de um setor exige várias capacitações que quando o país já as domina torna-se mais fácil ou menos custoso instalar outros setores que se beneficiarão dos *spillovers* de conhecimento e tecnológicos gerados (Hidalgo et al, 2007). Como as capacitações são complementares, no caso de um país não dispor de algumas delas, já se supõe que outras capacidades relacionadas não serão demandadas pelo fato das primeiras não estarem presentes, portanto, os países tendem a produzir bens que exigem a mesma capacitação ou que requerem o mesmo tipo de conhecimento incorporado (Hidalgo et al, 2007; Hausmann et al, 2011). Nesse sentido, os autores defendem que as firmas produzem muito menos do que sabem produzir, e os países fazem o que podem e não o que querem fazer.

Observa-se que por trás da proximidade ou vizinhança setorial está o tema conhecimento. Assim, complexidade de uma economia está relacionada à multiplicidade de conhecimentos úteis incorporados nela, a qual é expressa na composição da estrutura produtiva do país (Hausmann et al, 2011, p. 18). Quanto maior a variedade, mais benefícios tecnológicos e econômicos o país pode auferir (Koren e Tenreyro, 2013). Contudo, o conhecimento tem duas características que inviabilizam sua livre transferência: o caráter cumulativo e tácito (Dosi, 1988; Teece, 1986; Hausmann et al, 2011), desta forma, seu domínio não pode ser simplesmente adquirido, pois leva tempo até que competências sejam estabelecidas e dominadas (Malerba e Orsenigo, 1993; Frenken, van Oort e Verburg, 2007). A transferência de conhecimento requer que o receptor possua capacidade de absorção para identificar,

⁴ Num outro extremo, Marshall (1890) disse que as empresas de vários subsetores pertencentes a um *mesmo* setor ou setores próximos localizados numa *mesma* área geográfica são beneficiados por “externalidades de aglomeração ou de especialização”.

⁵ Por exemplo, um país que quer produzir mamão vai se beneficiar da infraestrutura logística e da disponibilidade de embalagens para transporte do produto caso ele já produza melão ou outra fruta.

interpretar e explorar o novo conhecimento (Cohen e Levinthal, 1990; Teece, 1986; Malerba e Orsenigo, 1993; Castaldi, Frenken e Los, 2015).

Entretanto, uma proximidade cognitiva em excesso não é desejável uma vez que ela também limita a dimensão do aprendizado e da inovação (Boschma, 2005, p. 64; Boschma e Iammarino, 2009). Para que o aprendizado e o conhecimento circulem entre agentes, é necessário um corpo de conhecimentos que se complementem e caso eles sejam idênticos, não há aprendizado para nenhuma das partes (Nooteboom, 2009). Além disso, uma só base de conhecimento pode provocar um aprisionamento (ou “*lock-in*”) numa espécie de “armadilha da competência” e dificultar a mudança de processos que deram certo um dia, mas que já estão defasados. Distâncias cognitivas muito pequenas podem indicar falta de novidade e de avanço tecnológico, mas distâncias demasiadamente grandes refletem falta de comunicação entre os atores (Boschma e Iammarino, 2009; Nooteboom, 2009), desta forma o espaço ideal entre diferentes áreas de especialização é aquele não tão perto mas nem tão longe, de modo que o corpo de conhecimento demandado e as competências detidas pelos agentes possam ser permutadas e absorvidas, gerando avanços em áreas vizinhas.

Em resumo, a literatura apontada evidencia que a proximidade entre setores produtivos possui seus fios condutores, que setores vizinhos possuem características que os conectam. Contudo, é preciso ainda identificar essa conexão para conhecer como a proximidade se revela entre os setores e buscaremos evidenciar essa relação de proximidade setorial com dados sobre unidades locais de empresas estabelecidas no Brasil.

3. Dados, construção das variáveis e matriz de vizinhança setorial

Existem classificações de setores e/ou produtos industriais⁶ que são úteis para vários propósitos analíticos e podem dar pistas para a identificação das relações de vizinhanças setoriais, por exemplo, a classificação industrial por intensidade tecnológica da OCDE (1994) que agrega os setores em quatro categorias a partir da razão entre gastos em P&D e faturamento, há também a taxonomia de padrões setoriais de mudança técnica de Pavitt (1984) em que os setores industriais são divididos em quatro grupos de acordo com a origem e o uso de inovações⁷. Todas elas indicam que o fator tecnologia é o principal critério para diferenciar os setores produtivos (e utilizam como *proxy* o indicador de esforço em P&D). Contudo, essas classificações ainda são muito agregadas e, por isso, insuficientes para o escopo deste artigo, motivo pelo qual buscamos uma forma alternativa e nova de observar a existência de vizinha setorial.

Para verificar se há dependência espacial na nossa matriz de vizinhança setorial utilizamos dados da Pesquisa Industrial Anual Empresa de 2013 (PIA Empresa) do IBGE. A PIA Empresa possui diversas informações para as empresas com cinco ou mais empregados no nível setorial bastante desagregado. Neste estudo, trabalharemos com os grupos da CNAE 2.0 para a indústria de transformação que é composto por 103 subsetores.⁸

Variável dependente. Utilizaremos como variável explicada (dependente) o número de engenheiros empregados em cada subsetor dividido pelo total de empregados de cada subsetor para o ano de 2013. Foi utilizada essa variável dependente porque a inovação e a intensidade tecnológica parecem ser os impulsionadores da diversificação setorial. Esse indicador é uma *proxy* do esforço inovativo dos 103 subsetores da indústria de transformação brasileira. Foi utilizado essa *proxy* porque a Pesquisa de Inovação Tecnológica (PINTEC) do IBGE que levanta informações sobre a atividade inovativa no país, além de ter uma periodicidade trienal, tem o nível de agregação setorial muito elevado. Tradicionalmente, o número de engenheiros é uma *proxy* muito reconhecida sobre os esforços inovativos das empresas.

⁶ Neste artigo, o termo *indústria* e *manufatura* são utilizados como sinônimos e referem-se à *indústria de transformação* que está classificada nas divisões 10 a 33 da Classificação Nacional de Atividades Econômicas (CNAE) na versão 2.0 do Instituto Brasileiro de Geografia e Estatística (IBGE).

⁷ Outras classificações muito citadas na literatura econômica são as de Lall (2000) e UNCTAD (2002), que utilizam como critérios básicos de diferenciação e agregação dos setores a intensidade em P&D e o uso de mão de obra qualificada pelos setores.

⁸ Neste artigo utiliza-se a palavra setor para referir-se a uma divisão específica da CNAE e subsetor para um grupo específico da CNAE. A divisão corresponde a desagregação a 2 dígitos da CNAE e o grupo a desagregação a 3 dígitos. A indústria de transformação possui 24 setores a dois dígitos e 103 subsetores a três dígitos de desagregação.

Araújo, Cavalcante e Alves (2014) num estudo sobre *proxy* para gastos com inovação no Brasil encontraram uma correlação superior a 90% entre pessoal técnico ocupado (engenheiros, pesquisadores e profissionais científicos) e gastos empresariais em P&D internos e externos. A mesma elevada correlação foi encontrada por Gusso (2006) ao considerar como pessoal técnico ocupado engenheiros, pesquisadores, profissionais científicos, diretores e gerentes de P&D e professores de ensino superior.

Variáveis de controle. No nível de desagregação a três dígitos da CNAE 2.0 foram criadas algumas variáveis de controle para o ano de 2013, quais sejam:

- i) Coeficiente de Importações: razão entre importações e o valor bruto da produção industrial;
- ii) Produtividade do trabalho: razão entre valor adicionado e o pessoal ocupado;
- iii) Indicador de porte das empresas: empregos das 4, 8 e 12 maiores empresas em relação ao total de empregos do subsetor;
- iv) Indicador de intensidade em capital: razão entre o estoque de capital e pessoal ocupado ou número de estabelecimentos;
- v) Coeficiente de Exportações: razão das exportações e valor bruto da produção industrial;
- vi) Outros indicadores.

Matriz de vizinhança setorial (matriz W). Foi utilizada uma *matriz de diversificação setorial* das empresas manufatureiras que atuaram em mais de uma divisão de atividade em 2013. Essa matriz foi obtida do IBGE mediante pedido de extração especial exclusiva para este estudo.

A matriz utilizada contém informações do valor bruto da produção industrial das unidades locais das empresas que atuaram em mais de uma divisão de atividade para todos os 103 subsetores da indústria de transformação. Utilizou-se como critério de vizinhança o fato de uma unidade local industrial produtiva – planta industrial – produzir produtos típicos do subsetor de origem e, também, de outros subsetores estranhos ao subsetor de origem. Assim, se uma planta industrial produz produtos típicos de alguns subsetores industriais esses mesmos subsetores são vizinhos porque dependem da mesma base produtiva e tecnológica. Acreditamos que essa forma de mensurar a proximidade setorial é mais apropriada que os métodos de entropia dos estudos de *variedade relacionada* e do *espaço do produto* porque os produtos de subsetores vizinhos foram produzidos na mesma base produtiva e tecnológica, e por isso capta melhor os esforços no âmbito da firma para atuar em setores distintos.

De posse dessas informações, foi construída uma matriz de zero e um, em que “um” indica o(s) subsetor(es) que cada um dos 103 subsetores atuam – isto é, a maioria dos subsetores industriais produz produtos de mais de um subsetor de atividade, assim são assinados com “um” os subsetores em que há produção industrial – e “zero” indica os subsetores que cada um dos 103 subsetores não atuam.

Após obter a matriz de zero e um foi zerada toda a diagonal principal já que um subsetor não é vizinho dele mesmo. Feito isso, foi verificado que alguns subsetores não possuíam vizinhos. Para contornar esse problema e garantir que a matriz possua determinante diferente de zero, foram realizados dois procedimentos: (i) agregação de sete subsetores considerando como critério a similaridade produtiva e tecnológica, resultando numa matriz final com 96 subsetores; (ii) a matriz de zero e um com zeros na diagonal principal foi somada com a sua transposta para obter uma matriz simétrica. Após esses procedimentos, a matriz de vizinhança de zero e um com 96 subsetores foi normalizada ao ser dividida pelo seu maior autovalor. Para as análises econométricas, essa matriz foi padronizada na linha, obtendo assim a matriz W.

4. Metodologia de avaliação da dependência espacial⁹

Os procedimentos metodológicos apresentados a seguir são amplamente utilizados em economia espacial para atestar a dependência espacial. Estimou-se o Índice I de Moran que é utilizado para detectar autocorrelação espacial e a testamos por modelos SAR e SEM.

⁹ Esta seção apresenta os procedimentos metodológicos utilizados na sequência deste artigo de modo bem sucinto. Uma exposição mais extensa e completa desses procedimentos metodológicos pode ser encontrada em LeSage e Pace (2009), Almeida (2012) e Elhorst (2014).

Índice I de Moran. O teste de autocorrelação espacial de unidades espaciais mais utilizado em estudo de *cross-section* foi desenvolvido por Moran (1948) e é amplamente conhecido por “I de Moran” (ou Estatística I de Moran) que em notação matricial, segundo Almeida (2012, p. 105), consiste:

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \frac{z'Wz}{z'z}$$

Onde:

- n é o número de unidades espaciais, corresponde aos 96 subsetores da indústria de transformação;
- z é a variável de interesse que neste estudo é um vetor dos resíduos de um modelo de Mínimos Quadrados Ordinários;
- Wz representa os valores médios da variável de interesse padronizada nos vizinhos, definidos segundo uma matriz de ponderação espacial W .
- w_{ij} são elementos da matriz de pesos espaciais referente ao subsetor i e ao subsetor j ;
- $\sum_i \sum_j w_{ij}$ indica a soma de todos os elementos da matriz de pesos espaciais W .

Segundo Almeida (2012, p. 105) a estatística de I de Moran é como um coeficiente de autocorrelação obtido pela autocovariância espacial do denominador composta pelos produtos cruzados $z'Wz$ dividida pela variância dos dados ($z'z$). Observe que se a matriz W for normalizada na linha, o duplo somatório no denominador é igual a n , então, podemos reescrever a última equação como:

$$I = \frac{z'Wz}{z'z}$$

A inferência para o Índice I de Moran baseia-se numa distribuição normal aproximada ao usar um valor Z padronizado abaixo:

$$Z(I) = \frac{I - E(I)}{DP(I)}$$

Onde $E(I)$ e $DP(I)$ são a média e o desvio-padrão, respectivamente, do I de Moran. O Índice I de Moran varia entre -1 e 1, sendo que os valores negativos (positivos) indicam autocorrelação espacial negativa (positiva). Um valor zero para o Índice indica inexistência de padrão espacial nos dados e a hipótese nula de aleatoriedade espacial é aceita.

Modelo autorregressivo espacial ou modelo de defasagem espacial (SAR). O modelo SAR¹⁰ incorpora um parâmetro rho (ρ) aos modelos lineares, representado por:

$$Y = \rho WY + \varepsilon \quad \text{ou} \quad Y = \beta X + \rho WY + \varepsilon$$

Onde:

- Y é um vetor $n \times 1$ de valores observados (a variável a ser explicada);
- ρ é coeficiente escalar autorregressivo;
- X é uma matriz $n \times k$ de variáveis explicativas (os controles);
- β é um vetor $p \times 1$ dos parâmetros das variáveis explicativas;
- W é a matriz de pesos espaciais;
- ε é um vetor $n \times 1$ dos erros, onde $\varepsilon \sim N(\phi, I\sigma^2)$.

O modelo SAR incorpora a autocorrelação espacial como componente adicional do modelo. O coeficiente ρ representa o efeito médio dos vizinhos sobre a variável dependente (Y), isto é, se ρ for estatisticamente diferente de zero quer dizer que uma parcela da variação total de Y é explicada pela dependência de cada observação de seus vizinhos. Em outros termos, um ρ positivo (negativo) significa que um alto valor de Y nos subsectores vizinhos aumenta (diminui) o valor de Y no subsetor i (ALMEIDA, 2012, p. 152). Se não houver autocorrelação espacial $\rho = 0$ e, neste caso, podemos estimar o modelo por MQO. Agora, se $\rho \neq 0$, a estimativa por MQO produzirá estimadores viesados e inconsistentes.

Após manipulações algébricas, a última equação pode ser escrita na forma reduzida como:

$$Y = (I_n - \rho W)^{-1} X\beta + (I_n - \rho W)^{-1} \varepsilon$$

A expressão acima requer que a matriz $(I_n - \rho W)^{-1}$ seja não singular para ser invertida (ALMEIDA, 2012, p. 154-155). As características apresentadas da matriz W na seção 3¹¹ garante a não singularidade da matriz $(I_n - \rho W)^{-1}$.

¹⁰ Em inglês, *Spatial Autoregressive Model* ou *Spatial Lag Model*.

Dado que $|\rho| < 1$, observe que podemos reescrever $[I_n - \rho W]^{-1}$ como:
 $[I_n - \rho W]^{-1} = I_n + \rho W + \rho^2 W^2 + \dots$

A passagem acima é semelhante a matriz de coeficiente técnicos diretos e indiretos do modelo de insumo-produto¹², neste caso, é captado um choque de produção setorial (efeito direto) e seu impacto na cadeia de fornecedores e subfornecedores (efeito indireto). Similarmente, no modelo SAR, a variável dependente (Y) pode ser explicada não apenas pelas variáveis X e ε (efeito direto), específicas a uma região (ou subsetor, como neste estudo), mas também por suas influências captadas pelas ligações com subsetor(es) mais ou menos conectado(s) pela matriz W.

Modelo de erro autorregressivo espacial (SEM). O modelo SEM¹³ considera a autocorrelação espacial no termo de erro, veja:

$$Y = \beta X + \varepsilon \quad \varepsilon = \lambda W \varepsilon + \mu \quad \mu \sim N(0, \sigma^2 I_n)$$

As variáveis que se repetem são as mesmas definidas anteriormente. O parâmetro lambda (λ), quando estatisticamente diferente de zero, reflete a autocorrelação espacial nos erros ou nas variáveis omitidas do modelo. “Neste modelo, os erros associados com qualquer observação são uma média dos erros nas regiões vizinhas mais um componente de erro aleatório” (ALMEIDA, 2012, p. 162), como expresso pela última equação acima. Assim, para $\lambda \neq 0$ um choque em alguma unidade espacial se espalha para as unidades espaciais vizinhas, mais o fato de $|\lambda| < 1$ garante que o efeito da propagação desse choque diminua à medida que se afasta de seu epicentro.

Após manipulações algébricas, a última equação pode ser escrita na forma reduzida por:

$$Y = \beta X + (I_n - \lambda W)^{-1} \mu$$

Semelhante ao já dito na subseção anterior, a matriz $(I_n - \lambda W)^{-1}$ necessita ser não singular, o que é garantido pelas características da matriz W.

Como anteriormente, podemos escrever $(I_n - \lambda W)^{-1}$ como:

$$[I_n - \lambda W]^{-1} = I_n + \lambda W + \lambda^2 W^2 + \dots$$

Neste caso,

(...) a parte da dependência não modelada se manifesta no padrão de erro aleatório entre regiões vizinhas, na forma de que os erros sejam autocorrelacionados espacialmente (...). Além de ser o reflexo de efeitos não modelados que exibem padrão espacial, o modelo de erro espacial pode informar que a influência sobre a variável dependente não é resultado apenas do choque, representado pelo termo de erro, específico a uma região [leia-se subsetor, neste estudo], mas também de transbordamentos de choques de regiões mais conectadas ou menos conectadas pela matriz W, como expresso pelo termo da equação $(I_n - \lambda W)^{-1} \mu$ (...). A variável Y é influenciada pelos choques vindo de todas as outras regiões, isto é, do sistema como um todo. (...) Um choque na região i afeta os vizinhos e os vizinhos dos vizinhos por intermédio das potências de W, e eventualmente, volta para afetar a região i de novo, porém, agora, com o efeito amortecido (ALMEIDA, 2012, p. 162).

Na sequência são apresentadas nossa formulação dos três modelos econométricos:

MQO: $Eng_s = \alpha + \beta M_VBPI_s + \delta VA_VP_s + \varepsilon$ onde,

Eng_s é o percentual de engenheiros em relação ao total de empregados de cada subsetor, M_VBPI_s é o coeficiente de importações subsetorial e VA_VP_s é a produtividade do trabalho subsetorial. α, β e δ são os coeficientes a serem estimados e ε é o termo de erro.

SAR: $Eng_s = \sum X_s \beta + \sum \rho W_{st} Eng_s + \varepsilon$ onde,

¹¹ Lembre-se que a matriz de ponderação W for normalizada na linha, assim sendo, a soma de suas linhas é limitada por um número fixo.

¹² O modelo de insumo-produto de Leontief pode ser escrito por $X = (I - A)^{-1}Y$, onde X é a produção, A é uma matriz de coeficientes técnicos diretos e Y é a demanda final. Podemos escrever $(I - A)^{-1} = I + A + A^2 + \dots$, que é a matriz dos coeficientes técnicos diretos e indiretos.

¹³ Em inglês, *Spatial Error Model*.

Eng_s é o percentual de engenheiros em relação ao total de empregados de cada subsetor, X_s são os controles subsetoriais como o coeficiente de importações e a produtividade do trabalho e W_{st} é a matriz de vizinhança subsetorial. Estamos interessados no valor do coeficiente autorregressivo ρ que capta a dependência espacial.

SEM: $Eng_s = \sum X_s \beta + \varepsilon$ e $\varepsilon = \lambda W_{st} \varepsilon + \mu$ onde,
 Eng_s , X_s , ε , W_{st} e β são as mesmas variáveis do modelo anterior. λ capta a autocorrelação espacial nos erros ou das variáveis omitidas do modelo e μ é um termo de erro aleatório.

5. Resultados

Os resultados serão discutidos em duas subseções, primeiro faremos na Subseção 5.1 uma análise descritiva dos dados e posteriormente, na Seção 5.2 apresentaremos os resultados dos modelos de econometria espacial testados.

5.1 Análise descritiva dos dados e da matriz de vizinhança setorial

Em 2013, das 181.767 empresas com mais de cinco empregados da indústria de transformação brasileira, apenas 1,0% delas, precisamente 1.830, atuaram em mais de uma divisão de atividade da CNAE 2.0 dentre as 33 divisões da indústria de transformação (ver Tabela 1). Pela quantidade de unidades produtivas (ou seja, por unidade local), 5,9% das empresas atuaram em mais de uma divisão de atividade. Essas são as empresas que doravante chamaremos de diversificadas. Embora o número relativo de empresas seja pequeno, as empresas diversificadas empregaram 18,5% do pessoal ocupado na manufatura, foram responsáveis por 30,6% do valor bruto da produção industrial e somaram um terço da receita líquida de vendas, evidenciando que são empresas de porte relevante e de alta produtividade.

Tabela 1: Descrição das empresas da indústria de transformação brasileira – 2013

	Número de empresas (NE)	Número de unidades locais (NUL)	Pessoal ocupado em 31.12 (PO)	Salários, retiradas e outras remunerações (milhões de R\$) (S&R&OR)	Valor Bruto da Produção Industrial (milhões de R\$) (VBPI)	Valor da transformação industrial (milhões de R\$) (VTI)	Receita líquida de vendas (milhões de R\$) (RLV)
Empresas que atuam em apenas uma divisão de atividade (A)	179.937	192.267	6.821.338	185.567	1.563.454	627.559	1.627.918
Empresas que atuam em mais de uma divisão de atividade (B)	1.830	12.100	1.547.065	64.419	689.950	342.092	788.793
Total de empresas (A+B)	181.767	204.367	8.368.403	249.986	2.253.404	969.651	2.416.711

Nota: Unidades de investigação: Empresa e unidade local com 5 ou mais pessoas ocupadas.

Fonte: Elaboração própria a partir da Pesquisa Industrial Anual do IBGE.

Os dados das Tabelas 1 e 2 revelam que as empresas diversificadas são significativamente maiores, remuneram melhor seus trabalhadores e são mais produtivas relativamente ao mar de empresas que atuaram em apenas uma divisão de atividade (doravante empresas não diversificadas).

As empresas não diversificadas empregaram, em média, 35,5 trabalhadores por unidade local contra 127,9 trabalhadores das empresas diversificadas. O valor bruto da produção industrial por unidade local das empresas diversificadas foi 7 vezes maior que das empresas não diversificadas. E a agregação de valor por trabalhador (VTI / PO) foi 2,4 vezes superior nas empresas diversificadas frente as não diversificadas. Esses são apenas alguns indicadores presentes na Tabela 2 que mostram que as empresas diversificadas possuem maior escala e eficiência em relação às empresas não diversificadas da manufatura brasileira.

Tabela 2: Indicadores selecionados de empresas da indústria de transformação brasileira – 2013

	PO / NE	PO / NUL	S&R&OR / PO	VBPI / NE (milhões de R\$)	VBPI / NUL (milhões de R\$)	VBPI / PO (mil R\$)	VTI / NE (milhões de R\$)	VTI / NUL (milhões de R\$)	VTI / PO (mil R\$)
Empresas que atuam em apenas uma divisão de atividade (A)	37,9	35,5	27.203,9	8,7	8,1	229,2	3,5	3,3	92,0
Empresas que atuam em mais de uma divisão de atividade (B)	845,4	127,9	41.639,7	377,0	57,0	446,0	186,9	28,3	221,1
Todas as empresas (A+B)	46,0	40,9	29.872,6	12,4	11,0	269,3	5,3	4,7	115,9

Nota: Variáveis nomeadas na Tabela 1. Unidades de investigação: Empresa e unidade local com 5 ou mais pessoas ocupadas.

Fonte: Elaboração própria a partir da Pesquisa Industrial Anual do IBGE.

Considerando apenas as empresas que atuaram em mais de uma divisão de atividade, é possível identificar as unidades locais que se dedicaram à produção industrial, administração ou outra atividade não-industrial (como serviços de comercialização e *marketing*) (Tabelas 3 e 4). Observa-se que das empresas diversificadas, 5.388 unidades locais (ou 44,5% do total) foram constituídas com objetivos não industriais, ou seja, são unidades de distribuição, revenda de mercadorias, prestação de serviços não-industriais e venda no varejo de produtos de fabricação própria (quando realizada em unidade distinta da unidade de fabricação).

Tabela 3: Empresas diversificadas da indústria de transformação, por tipo de função – 2013

	Número de unidades locais	Pessoal ocupado em 31.12	S&R&OR (milhões de R\$)	VBPI (milhões de R\$)	VTI (milhões de R\$)	RLV (milhões de R\$)
Unidades locais industriais produtivas (C)	4.579	1.273.302	45.620	689.948	342.090	739.974
Unidades locais industriais administrativas (D)	2.133	136.002	14.855	-	-	-
Unidades locais produtivas não-industriais (E)	5.388	137.761	3.944	2	1,3	48.819
Total das empresas (C+D+E)	12.100	1.547.065	64.419	689.950	342.092	788.793

Nota: Unidades de investigação: Empresa e unidade local com 5 ou mais pessoas ocupadas.

Fonte: Elaboração própria a partir da Pesquisa Industrial Anual do IBGE.

Apesar do maior número, essas unidades locais são pequenas, empregam apenas 8,9% do total dos trabalhadores das empresas diversificadas, apresentando uma média de 25,6 trabalhadores ocupados por unidade local com salário médio anual de R\$28,6 mil (ver Tabela 4).

Já as unidades locais diversificadas para outras atividades produtivas representam 37,8% do total (4.579 unidades), neste caso são empresas que desenvolveram capacitações e domínio tecnológico em diferentes áreas, que lhe conferiram um VBPI de cerca de R\$ 690 milhões em 2013, equivalente a um VBPI de R\$ 542 por trabalhador. Em relação ao emprego gerado, representam 82,3% dos empregados das empresas diversificadas, os quais possuem salário médio anual de R\$ 35,8 mil (Tabelas 3 e 4).

Nas unidades não-industriais e nas administrativas praticamente não há VTI e VBPI porque elas efetivamente não realizam produção, apenas operam como escritórios de apoio às unidades produtivas.

Tabela 4: Indicadores selecionados das empresas diversificadas da indústria de transformação, por tipo de função – 2013

	Pessoal ocupado / número de unidades locais	Salário médio anual	VBPI / Número de unidades locais (milhões R\$)	VBPI / pessoal ocupado (mil R\$)	VTI / Número de unidades locais (milhões R\$)	VTI / pessoal ocupado (mil R\$)
Unidades locais industriais produtivas (C)	278,1	35.828,4	150,7	541,9	74,7	268,7
Unidades locais industriais administrativas (D)	63,8	109.227,2	-	-	-	-
Unidades locais produtivas não-industriais (E)	25,6	28.628,1	0,0	0,0	0,0	0,0
Total das empresas (C+D+E)	127,9	41.639,7	57,0	446,0	28,3	221,1

Nota: Unidades de investigação: Empresa e unidade local com 5 ou mais pessoas ocupadas.

Fonte: Elaboração própria a partir da Pesquisa Industrial Anual do IBGE.

Ao observar apenas as informações das unidades locais industriais produtivas das empresas diversificadas ficam ainda mais discrepantes seus indicadores frente às unidades locais das empresas não diversificadas (comparação da Tabela 4 e Tabela 2). As unidades locais industriais produtivas das empresas diversificadas obtiveram, em média, VTI de R\$ 74,7 mil em relação a R\$ 3,3 mil das empresas não diversificadas, uma diferença de 22,6 vezes.

Quanto à origem da RLV das empresas diversificadas (exposto na Tabela 5), é possível incluir também a diversificação que ocorre para outros subsetores externos à indústria de transformação. Em 2013, 81,4% da RLV adveio do subsetor de origem da atividade principal. O restante está dividido em: (i) diversificação dentro do setor de origem com 5,0% da RLV; e (ii) diversificação fora do setor de origem com 13,6% da RLV, sendo este distribuído entre diversificação na própria indústria de transformação (4,9%); diversificação para a indústria extrativa (2,5%) e diversificação para os demais subsetores da CNAE 2.0, notadamente para os serviços (6,2%).

Tabela 5: Formas de diversificação das empresas da indústria de transformação: diversificação global – 2013

	RLV (milhões R\$)	RLV (em %)
Apenas no subsetor de origem (atividade principal)	641.851	81,4
Subsetores do mesmo setor de origem (diversificação dentro do setor de origem)	39.300	5,0
Subsetores de um setor diferente da origem (diversificação fora do setor de origem)	39.043	4,9
Subsetores da indústria extrativa (diversificação fora da indústria de transformação)	19.780	2,5
Subsetores das demais seções da CNAE 2.0, essencialmente serviços (diversificação fora da indústria de transformação)	48.819	6,2
Total da RLV das empresas que atuaram em mais de uma divisão de atividade	788.793	100,0

Nota: Unidades de investigação: Empresa e unidade local com 5 ou mais pessoas ocupadas.

Fonte: Elaboração própria a partir da Pesquisa Industrial Anual do IBGE.

Ao considerar a produção apenas da indústria de transformação das empresas diversificadas, verificou-se que 88,7% da produção manufatureira ocorreu apenas no subsetor de origem da atividade principal dessas empresas em 2013 (Tabela 6). O restante da produção adveio: (i) de subsectores pertencentes ao mesmo setor de origem da atividade principal, com 5,2% da produção manufatureira e (ii) de subsectores de um setor diferente da origem, com 6,1% da produção manufatureira.

Sendo assim, a diversificação para setores diferentes foi um pouco superior a diversificação dentro de um mesmo setor quanto ao valor da produção de bens manufatureiros.

Tabela 6: Empresas que diversificaram para a própria indústria de transformação: diversificação dentro versus fora do subsetor de origem – 2013

	VBPI (milhões R\$)	VBPI (em %)
Apenas no subsetor de origem (atividade principal)	561.935	88,7%
Subsetores do mesmo setor de origem (diversificadas para dentro do setor de origem)	32.991	5,2%
Subsetores de um setor diferente do setor de origem (diversificadas para fora do setor de origem)	38.487	6,1%
Total da produção manufatureira das empresas diversificadas	633.414	100,0%

Nota: Unidades de investigação: Empresa e unidade local com 5 ou mais pessoas ocupadas.

Fonte: Elaboração própria a partir da Pesquisa Industrial Anual do IBGE.

Em termos da quantidade de setores em que existem empresas diversificadas, observa-se que este número é elevado. Dos 103 subsectores, 84 apresentaram diversificação para pelo menos um outro

subsetor da indústria de transformação. Embora os 19 subsetores restantes não tenham se diversificado para outras divisões da indústria de transformação, eles também são diversificados porque registraram produção em outras divisões de atividade da indústria extrativa ou outra divisão de atividades não industriais.

Vale ressaltar que o foco deste estudo é a indústria de transformação, especialmente os 84 subsetores que apresentaram diversificação para o interior da manufatura. Para destaca-los, selecionamos os 20 subsetores mais diversificados neste contexto, os quais estão apresentados na Tabela 7 abaixo.

Tabela 7: Subsetores mais diversificados da indústria de transformação brasileira (2013)

Cód.	Grupos (subsetores) da CNAE 2.0	Número de subsetores em que registram produção
294	Fabricação de peças e acessórios para veículos automotores	18
222	Fabricação de produtos de material plástico	15
283	Fabricação de tratores e de máquinas e equipamentos para a agricultura e pecuária	14
101	Abate e fabricação de produtos de carne	12
259	Fabricação de produtos de metal não especificados anteriormente	12
281	Fabricação de motores, bombas, compressores e equipamentos de transmissão	12
254	Fabricação de artigos de cutelaria, de serralheria e ferramentas	10
329	Fabricação de produtos diversos	10
206	Fabricação de sabões, detergentes, produtos de limpeza, cosméticos, produtos de perfumaria e de higiene pessoal	9
233	Fabricação de artefatos de concreto, cimento, fibrocimento, gesso e materiais semelhantes	9
243	Produção de tubos de aço, exceto tubos sem costura	9
271	Fabricação de geradores, transformadores e motores elétricos	9
310	Fabricação de móveis	9
104	Fabricação de óleos e gorduras vegetais e animais	8
106	Moagem, fabricação de produtos amiláceos e de alimentos para animais	8
141	Confecção de artigos do vestuário e acessórios	8
162	Fabricação de produtos de madeira, cortiça e material trançado, exceto móveis	8
174	Fabricação de produtos diversos de papel, cartolina, papel-cartão e papelão ondulado	8
251	Fabricação de estruturas metálicas e obras de caldeiraria pesada	8
286	Fabricação de máquinas e equipamentos de uso industrial específico	8

Fonte: Elaboração própria a partir de extração especial da Pesquisa Industrial Anual do IBGE.

A Tabela 7 elenca os 20 subsetores que registraram produção manufatureira mais diversificada do Brasil em 2013. O líder é o subsetor de “fabricação de peças e acessórios para veículos automotores”, em que sua produção está distribuída em 18 subsetores manufatureiros. Isso indica que as empresas desse subsetor principal têm maior potencial de expandir e/ou aprofundar a diversificação no futuro, haja vista que quanto mais diversificados são esses subsetores maior é o potencial deles se diversificarem ainda mais, conforme Hausmann et al (2011) e Rodrik (2010). Também são subsetores menos suscetíveis a crises e choques externos devido a sua facilidade de transitar em vários segmentos de mercado, os quais podem se balancear diante de crises momentâneas em um mercado específico, conforme mostraram Frenken, van Oort e Verburg (2007) e Boschma e Iammarino (2009) sobre o efeito positivo na absorção de choques em regiões com produção mais diversificada.

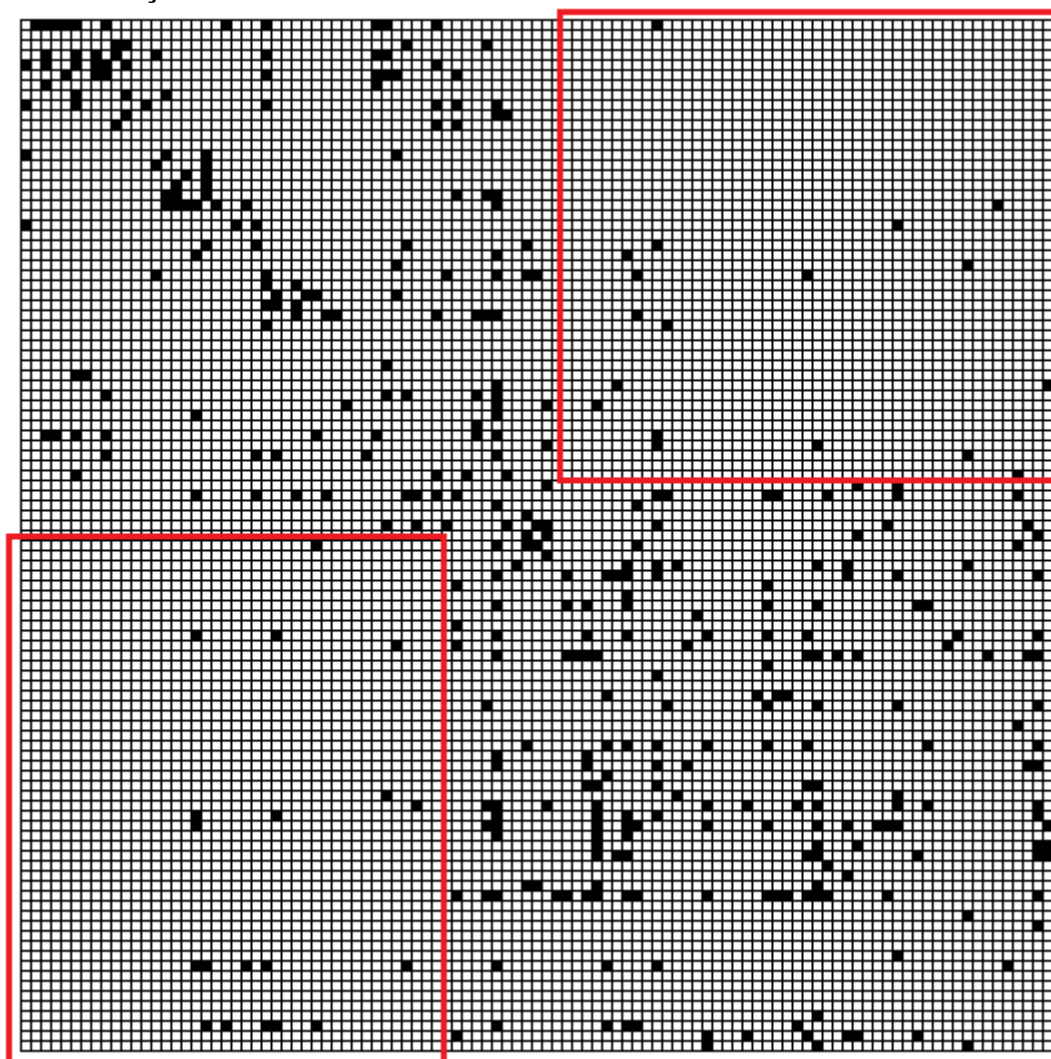
Necessita-se de estudos adicionais para entender porque esses 20 subsetores são os mais diversificados. Será que a característica estrutural de alguns deles contribui para serem muitos diversificados? Por enquanto, buscaremos uma resposta mais simples, que é entender quais setores são vizinhos entre si na estrutura produtiva brasileira.

Elaboramos a Figura 1 com uma matriz quadrada dos 103 subsetores da indústria de transformação, de acordo com a ordem em que eles aparecem na CNAE 2.0. Observe que nos dois quadrantes em destaque – quadrante inferior-esquerdo e quadrante superior-direito – praticamente não há relação de vizinhança bem definida.

Já nos outros dois quadrantes – superior-esquerdo e inferior-direito – a relação de vizinhança é mais explícita. No quadrante superior-esquerdo estão alocados, em sua maioria, os subsetores de baixa e média-baixa tecnologia que possuem relações de vizinhança principalmente com os subsetores de baixa e

média baixa tecnologia. Já no quadrante inferior-direito estão dispostos, em sua maioria, os subsetores de alta e média-alta tecnologia, que também possuem relações de vizinhança bem explícita com os subsetores de mesma intensidade tecnológica.

Figura 1: Matriz de vizinhança/proximidade setorial dos 103 subsetores da indústria de transformação brasileira



Nota: Matriz de 0 e 1, sendo a cor preta = 1. A diagonal principal não está preenchida porque um subsetor não é vizinho dele mesmo. Considerou-se a diversificação ocorrida apenas nos subsetores da indústria de transformação. Unidades de investigação: Empresa e unidade local com 5 ou mais pessoas ocupadas.

Fonte: Elaboração própria a partir de extração especial da Pesquisa Industrial Anual do IBGE.

Nesse sentido, com raras exceções, as relações de vizinhança setorial da Figura 1 parecem indicar um padrão em que os subsetores se relacionam com vizinhos próximos e dentro do mesmo nível tecnológico. As relações de vizinhança dos subsetores de alta e média-alta tecnologia com subsetores de baixa e média-baixa tecnologia, ou vice-versa, são menos intensas (veja os dois quadrantes em destaque da Figura 1). Assim, pode-se considerar que a origem tecnológica do subsetor é um bom sinalizador de suas relações de vizinhança com os demais subsetores.

Apesar de a Figura 1 sugerir algum padrão de dependência espacial, a sua comprovação estatística requer análises de métodos mais apropriados como o Índice I de Moran e o teste LM exibidos na próxima seção.

5.2 Resultados das especificações econométricas

Existem alguns testes para detectar a autocorrelação espacial, sendo o Índice I de Moran e o teste LM (Multiplicador de Lagrange) os mais utilizados pela literatura.

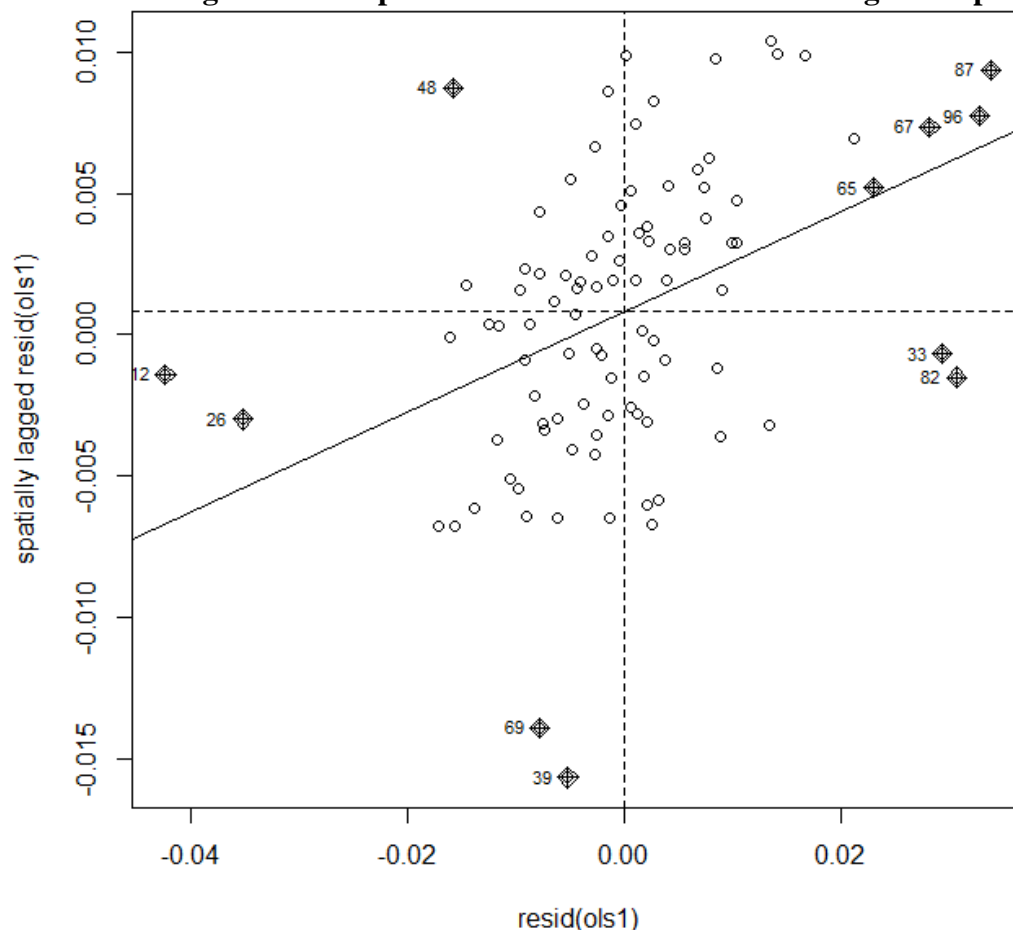
Tabela 8: Coeficiente do Índice I de Moran aplicado aos resíduos da regressão por MQO

Valor observado	Valor esperado	Valor p	Nível de significância
0,17722168	-0,0136963	0,001245	***

Nota: A regressão pelo método de MQO está exposta na Tabela 13 abaixo. *** significante ao nível de 1%.

Fonte: Elaboração própria.

Gráfico 1: Diagrama de dispersão de Moran dos resíduos da regressão por MQO



Nota: A regressão pelo método MQO está exposta na Tabela 13 abaixo.

Fonte: Elaboração própria.

A Tabela 8 mostra que o Índice I de Moran é positivo no valor de 0,177 e que pode-se rejeitar a hipótese nula da aleatoriedade espacial num nível de significância de 1%, pois há uma *similaridade* no modelo de regressão entre o número de engenheiros e a sua distribuição subsetorial. Isto significa que há uma autocorrelação espacial positiva em que altos valores do número de engenheiros (y) tendem a estar circundados por altos valores desta variável em subsetores vizinhos (W_y), ao passo que baixos valores de y tendem a estar rodeados por baixos valores também de y em subsetores vizinhos (W_y) conforme evidenciou Almeida (2012, p. 106). Neste caso, a chance de se obter num subsetor vizinho um valor parecido com o que se encontra em um determinado subsetor é elevada, o que já era sugerido visualmente pela Figura 1 da subseção 5.1.

O diagrama de dispersão de Moran (Gráfico 1) apresenta concentrações intensas nos quadrantes baixo-baixo (BB) e alto-alto (AA), ou seja, vizinhos parecidos são circundados por vizinhos parecidos (conforme já sugerido pela Figura 1). As aglutinações nos quadrantes baixo-alto (BA) e alto-baixo (AB) são menos intensas e indicam o caso inverso, em que vizinhos diferentes são circundados por vizinhos diferentes. De qualquer forma, o sinal da autocorrelação visto pelo Gráfico 1 indica que não há aleatoriedade do comportamento espacial.

O Índice I de Moran constatou que os resíduos da regressão por MQO apresentaram autocorrelação espacial, que inviabilizou o prosseguimento da análise por MQO. Vale ressaltar que quando existe autocorrelação espacial: (i) na variável dependente, as estimativas de MQO são viesadas e inconsistentes ou (ii) no termo de erro, não há viés nem inconsistência, mas o estimador de MQO deixa de ser o mais eficiente.

Sendo assim, foi realizado o teste LM, baseado na estimação por máxima verossimilhança, para a escolha do modelo espacial mais adequado (Tabela 9). O teste LM apontou que os modelos LMlag (SAR), SARMA e LMerr (SEM) são, nesta ordem, os modelos mais apropriados a serem estimados, pois apresentaram os menores valor p . Esses três modelos revelaram resultados diferentes de zero ao nível de 1% de significância, o que indica, mais uma vez, a presença da autocorrelação espacial.

Tabela 9: Teste LM para escolha do modelo espacial

Modelo	Estatísticas	gl	valor p
LMerr (SEM)	7,27662	1	0,0070
LMlag (SAR)	10,56800	1	0,0012
RLMerr	0,08309	1	0,7732
RLMlag	3,37447	1	0,0662
SARMA	10,65109	2	0,0049

Fonte: Elaboração própria.

Na Tabela 10 são apresentados os resultados apenas das regressões para os modelos SAR e SEM, além do modelo MQO que foi utilizado para calcular o Índice I de Moran anteriormente.

Tabela 10: Resultados dos modelos econométricos para o ano de 2013

Mínimos Quadros Ordinários (MQO)					
	Coefficiente estimado	Desvio padrão	Valor t	Pr(> t)	Nível de significância
Intercepto	-0,01020	0,00220	-4,637	0,00001	***
Coefficiente de Importações	0,01937	0,00429	4,513	0,00002	***
Produtividade do Trabalho	0,00017	0,00001	14,968	0,00000	***
R^2 : 0,7169 R^2 ajustado: 0,7108 Estatística F: 117,7					
Modelo Auto regressivo espacial (SAR)					
	Coefficiente estimado	Desvio padrão	Valor z	Pr(> z)	Nível de significância
Intercepto	-0,014299	0,00203	-7,0317	0,0000	***
Coefficiente de Importações	0,017850	0,00386	4,6259	0,0000	***
Produtividade do Trabalho	0,000162	0,00001	15,0759	0,0000	***
Rho (ρ)	0,510610	0,09870	5,1733	0,0000	***
AIC: -580,63 Log likelihood: 295,3125					
Modelo de Erro Espacial (SEM)					
	Coefficiente estimado	Desvio padrão	Valor z	Pr(> z)	Nível de significância
Intercepto	-0,009852	0,00354	-2,7801	0,0054	***
Coefficiente de Importações	0,015237	0,00404	3,7738	0,0002	***
Produtividade do Trabalho	0,000170	0,00001	15,5037	0,0000	***
Lambda (λ)	0,600720	0,12152	4,9434	0,0000	***
AIC: -577,08 Log likelihood: 293,5411					

Nota: Coeficiente de importações = razão de importações e o valor bruto da produção industrial. Produtividade do trabalho = valor adicionado dividido pelo pessoal ocupado. *** significante ao nível de 1%.

Fonte: Elaboração própria.

No modelo de MQO, apenas duas variáveis independentes explicam mais de 71% do percentual de engenheiros em relação ao emprego total no âmbito setorial (Tabela 13).¹⁴ Entretanto, o Índice I de Moran mostrado anteriormente detectou a existência de autocorrelação espacial positiva e, nessa situação, a estimação por MQO não é a mais adequada. Por isso, foram estimados os modelos espaciais SAR e SEM apontados pelo teste LM como apropriados para esse caso.

Os coeficientes de todos os modelos apresentados na Tabela 10 são estatisticamente significantes ao nível de 1% e possuem os sinais esperados. Assim, o percentual de engenheiros em relação aos empregados totais dos subsetores tende a variar positivamente com o aumento da produtividade do trabalho e com a elevação do coeficiente de importações. Ademais, os coeficientes de autocorrelação espacial ρ do modelo SAR e λ do modelo SEM foram estatisticamente diferentes de zero ao nível de 1% de significância e positivos, confirmando a presença de autocorrelação espacial positiva.

Esta autorrelação espacial é forte e não marginal, pois os coeficientes estimados ρ e λ foram superiores a 0,5. Isso coloca em evidência a dependência espacial no âmbito subsetorial da manufatura brasileira. Ademais, o melhor modelo é o SAR por apresentar o menor valor de AIC e maior valor do Log.

Tabela 11: Efeito direto e indireto

Variáveis	Efeito Marginal Médio		
	Direto	Indireto	Total
Coeficiente de Importações	0,0186040	0,0193258	0,0379298
Produtividade do Trabalho	0,0001692	0,0001762	0,0003453

Fonte: Elaboração própria.

A Tabela 11 apresenta os efeitos marginais médios obtidos através do modelo SAR por meio de 10.000 simulações.¹⁵ Os efeitos direto e indireto são bem próximos, sendo o efeito indireto ligeiramente superior tanto para o coeficiente de importações quanto para a produtividade do trabalho. Desse modo, um aumento de R\$ 1.000,00 na produtividade do trabalho em um subsetor eleva a participação de engenheiros no emprego total em apenas 0,035% direta e indiretamente. Já um aumento de um ponto percentual no coeficiente de importações em um subsetor gera um impacto total de 3,78% na participação de engenheiros no emprego total. Logo, o coeficiente de importações produz um impacto maior que a produtividade do trabalho; mas é importante ressaltar que o coeficiente de importações tem um teto máximo e muito antes de se chegar neste teto as importações podem ocasionar uma desindustrialização absoluta e demissão de engenheiros ao invés de contratação.

Tabela 12: Simulações com os efeitos marginais médios

Aumento de 10% no Coeficiente de Importações atual			
	Direto	Indireto	Total
Aumento no número de engenheiros (em quantidade)	3.744	3.890	7.634
Aumento no número de engenheiros em relação ao total de empregados (em porcentagem)	5,06%	5,26%	10,33%
Aumento de 10% no nível da Produtividade do Trabalho atual			
	Direto	Indireto	Total
Aumento no número de engenheiros (em quantidade)	241	251	492
Aumento no número de engenheiros em relação ao total de empregados (em porcentagem)	0,33%	0,34%	0,67%

Fonte: Elaboração própria.

¹⁴ Foram testadas diversas variáveis explicativas ou de controle, mas na Tabela 13 estão expostas apenas as duas com maior poder de explicação.

¹⁵ Elhroost (2014, p. 20-25) apresenta trechos interessantes sobre os efeitos diretos e indiretos.

A Tabela 12 apresenta uma situação com os efeitos diretos e indiretos mostrados na Tabela 14. Um aumento de 10% no coeficiente de importações de cada subsetor contribuiria para aumentar o número de engenheiros em 7.634, sendo 3.744 novos engenheiros diretos e 3.890 indiretos (Tabela 15), representando respectivamente um acréscimo de 10,33%, 5,06% e 5,26% engenheiros na indústria. Contudo, a quantidade adicional de engenheiros para um aumento de 10% no nível atual da produtividade do trabalho seria bem mais modesta.

6. Considerações finais

Este artigo investigou a questão da vizinhança ou proximidade setorial de 103 subsetores da indústria de transformação brasileira no ano de 2013. Utilizou-se como critério de vizinhança o fato de uma unidade local industrial produtiva – planta industrial – produzir produtos típicos do subsetor de origem e de outros subsetores estranhos ao subsetor de origem. Assim, os subsetores industriais são vizinhos porque dependem da mesma base produtiva e tecnológica. Para tanto, neste estudo, foi utilizado uma matriz de diversificação setorial bem detalhada das empresas que atuaram em mais de uma divisão de atividade – empresas diversificadas – elaborada pelo IBGE através de uma tabulação especial da Pesquisa Industrial Anual Empresa de 2013. Nesse sentido, procuramos responder a pergunta se os subsetores industriais possuem subsetores vizinhos assim como os territórios geográficos? Se sim, como ocorre essa relação de vizinhança?

Verificamos através de análises descritivas e de métodos econométricos de economia espacial que os subsetores industriais da manufatura brasileira possuem vizinhos industriais bem definidos, isto é, eles são espacialmente dependentes ou autocorrelacionados. Isto foi confirmado pelos elevados coeficientes rho (ρ) do modelo SAR e lambda (λ) do modelo SEM e também pelo Índice I de Moran e pelo teste LM; todos esses métodos atestaram a dependência espacial e foram estatisticamente significativos ao nível de 1% de significância.

Há um padrão de vizinhança subsetorial mais nítido entre os subsetores da indústria de baixa e média-baixa tecnologia com os subsetores dessa mesma indústria e, também, entre os subsetores da indústria de alta e média-alta tecnologia e os subsetores dessa mesma indústria. A vizinhança subsetorial entre subsetores de níveis tecnológicos muito distintos foram menos intensas. Portanto, com raras exceções, os subsetores se relacionam com vizinhos próximos e, com maior probabilidade, dentro do mesmo nível tecnológico. Então, a origem tecnológica de um subsetor é um bom sinalizador de suas relações de vizinhança com os demais subsetores.

As empresas diversificadas – que atuam em mais de uma divisão de atividade – são significativamente maiores, mais produtivas e eficientes que as empresas não diversificadas, embora elas constituam um conjunto de empresas relativamente pequeno. Ademais, na indústria brasileira, a diversificação ocorre tanto dentro do mesmo setor quanto para setores diferentes, com maior relevância para esta última.

Quanto à agenda de pesquisa, levantamos alguns pontos importantes a partir desse trabalho. Primeiro, é interessante verificar se a vizinhança setorial brasileira aumentou ou diminuiu durante o tempo, no período de 1996 (início da PIA/IBGE) a 2013. Segundo, comparar o grau da diversificação e da vizinhança setorial brasileira com a de outros países. E terceiro observar como opera a influência da vizinhança setorial mediante estímulos de políticas setoriais.

7. Referências

- ALMEIDA, E. (2012). *Econometria especial aplicada*. Campinas, Editora Alínea.
- ARAÚJO, B. C.; CAVALCANTE, L. R.; ALVES, P. (2009). Variáveis proxy para os gastos empresariais em inovação com base no pessoal ocupado técnico-científico disponível na Relação Anual de Informações Sociais. In: *Radar: tecnologia, produção e comércio exterior*, n. 5, p. 16-21, Brasília: Ipea
- BOSCHMA, R. (2005) Proximity and Innovation: A Critical Assessment. *Regional Studies*, 39(1), pp. 61-74.
- BOSCHMA, R. A.; IAMMARINO, S. (2009). Related variety, trade linkages and regional growth in Italy. *Economic Geography*, 85, pp. 289–311.

- BRESCHI, S.; MALERBA, F. (1997). Sectoral innovation systems: technological regimes, Schumpeterian dynamics, and spatial boundaries. In: Edquist, C. (org.), *Systems of innovation: Technologies, institutions and organizations*, p. 130-156. London: Pinter A Cassell.
- CASTALDI, C.; FRENKEN, K.; LOS, B. (2015). Related Variety, Unrelated Variety and Technological Breakthroughs: An analysis of US State-Level Patenting. *Regional Studies*, 49(5), pp. 767-781.
- CHRISTALLER, W. (1935). *Central Places in Southern Germany*. Prentice-Hall.
- COHEN, W. M.; LEVINTHAL, A. D. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 35(1), pp. 128-152, Special Issue: Technology, Organizations, and Innovation.
- COHEN, W. M.; LEVINTHAL, A. D. (1989). Innovation and learning: the two faces of R&D. *The Economic Journal*, 99, pp. 569-596.
- ELHORST, J. P. (2014). *Spatial econometrics: from cross-sectional data to spatial panels*. Springer.
- FRENKEN, K., Van OORT, F. G.; VERBURG, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies*, 41, pp. 685–697.
- GRANSTRAND, O., PATEL, P., PAVITT, K. (1997). Multi-technology corporations: why do they have “distributed” rather than “distinctive core” competencies. *California Management Review*, 39, pp. 8-25.
- GUSSO, D. (2006). Agentes da inovação: quem os forma, quem os emprega? In: De Negri, J.A.; De Negri, F.; Coelho, D. (org.). *Tecnologia, exportação e emprego*. Brasília: Ipea.
- HAUSMANN, R.; HIDALGO, C. A.; BUSTOS, S.; COSCIA, M; CHUNG, S.; JIMENEZ, J; SIMOES, A.; YILDIRIM, M. A. (2011). *The Atlas of Economic Complexity: Mapping paths to prosperity*. Boston, Center for International Development. Harvard University.
- HIDALGO, C. A; KLINGER, B; BARABÁSI, AL; HAUSMANN, R. (2007). The Product Space Conditions the Development of Nations. *Science*. 317, pp. 482-487.
- HIRSCHMAN, A. O. (1958). *The strategy of economic development*. London: Yale University Press.
- HIRSCHMAN, A. (1981). *Policy-making and policy analysis in Latin America: A return journey. Essays in Trespassing: Economics to Politics and Beyond*. Cambridge: Cambridge University Press.
- IMBS, J.; WACZIARG, R. (2003). Stages of Diversification. *The American Economic Review*, 93(1), March, pp. 63-86.
- ISARD, W. (1956). *Location and space economy: a general theory relation to industrial location, market areas, land use trade and urban structure*. Cambridge: MIT Press.
- JACOBS, J. (1969). *The Economy of cities*. New York: Random House.
- KOREN, M.; TENREYRO, S. (2013). Technological diversification. *American Economic Review*, 103(10), pp. 378-414.
- KRUGMAN, P.R. (1991), *Geography and Trade*. Boston: MIT Press.
- LALL, S. (1992). Technological capabilities and industrialization. *World development*. 20(2), pp. 165-86.
- LALL, S. (2000). The technological structure and performance of developing country manufactured exports, 1985-1998. *Oxford Development Studies*, 28(3), pp. 337-369.
- LESAGE, J.; PACE, R. K. (2009). *Introduction to spatial econometrics*. London: CRC Press.
- LÖSCH, A. (1940). *The economics of location*. New Haven: Yale University Press.
- LUNDVALL, B. (1992). *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning*. London: Pinter.
- MALERBA, F.; ORSENIGO, L. (1993). Technological Regimes and Firm Behavior. *Industrial and Corporate Change*, 2(1).
- MARSHALL, A. (1890). *Principles of Economics: An Introductory*. London: Macmillan.
- MORAN, P. A. P. (1948). The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society*, 10(2), pp. 243-251.
- NOOTEBOOM, B. (2000) *Learning and Innovation in Organizations and Economies*. Oxford University Press, Oxford.
- OCDE (1994). *Globalisation and Competitiveness: Relevant Indicators*. OECD Directorate for Science, Technology and Industry, Paris.
- PAVITT, K. (1984). Sectoral patterns of technical change: towards a taxonomy and a theory. *Research Policy*, 13(6), p. 343-373.

- PENROSE, E. (1959). *The theory of the growth of the firm*. John Wiley and Sons, New York.
- QUATRARO F. (2010). Knowledge coherence, variety and productivity growth: manufacturing evidence from Italian regions. *Research Policy*, 39, pp. 1289–1302.
- RODRIK, D. (2010). Políticas de diversificação econômica. *Revista Cepal*, Número Especial em português, p. 27-43.
- SCHUMPETER, J. A. (1911). *The Theory of Economic Development*. Oxford University Press, Nova York.
- SCHUMPETER, J. A. (1939). *Business cycles: a theoretical, historical, and statistical analysis of the capitalist process*. New York: McGraw-Hill.
- SCHUMPETER, J. A. (1942). *Capitalismo, Socialismo e Democracia*. Rio de Janeiro: Editora Fundo de Cultura.
- TEECE, D. J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing, and public policy. *Research Policy*, 15(6), pp. 285-305.
- THÜNEN, J. H. V. (1826). *The isolated state*. New York: Pergamon Press.
- TORRE, A; GILLY, J. P. (2000) On the analytical dimension of proximity dynamics. *Regional Studies*, 34, pp. 169-180.
- UNCTAD (2002). *Trade and Development Report 2002*. United Nations Conference on Trade and Development. United Nations, Geneva.
- VIOTTI, E. B. (2002). National learning systems: a new approach on technological change in late industrializing economies and evidence from the cases of Brazil and South Korea. *Technological Forecasting and Social Change*, New York, 69, pp. 653-680.
- WEBER, A. (1929). *Theory of the location of industries*. Chicago: The University of Chicago Press.