Universidade do Minho
Escola de Economia e Gestão

Ana Filipa Ribeiro Freitas

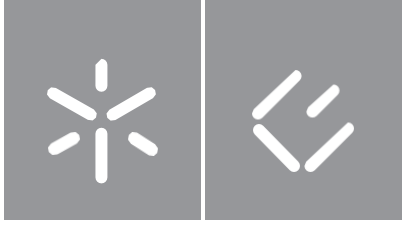
Desigualdades regionais na criação de empresas: o efeito das externalidades de aglomeração

abril de 2026

Desigualdades regionais na criação de empresas: o efeito das externalidades de aglomeração

Ana Filipa Ribeiro Freitas

UMinho | 2026



Universidade do Minho
Escola de Economia e Gestão

Ana Filipa Ribeiro Freitas

Desigualdades regionais na criação de empresas: o efeito das externalidades de aglomeração

Dissertação de Mestrado
Economia Industrial e da Empresa

Trabalho efetuado sob a orientação da
Professora Doutora Isabel Correia

abril de 2026

Direitos de Autor e Condições de Utilização do Trabalho Por Terceiros

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos. Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada. Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.



Atribuição
CC BY

https://creativecommons.org/licenses/by/4.0/

Agradecimentos

A realização desta dissertação representa o fim de um percurso, repleto de desafios, conquistas e alegrias. Para mim, é não só um marco académico, mas também a prova de que conseguimos desafiar-nos a nós próprios e superar as dúvidas sobre o nosso potencial. No entanto, isto não seria possível sem o apoio e incentivo de várias pessoas.

Primeiramente, quero agradecer à minha orientadora Professora Doutora Isabel Correia pelo apoio, acompanhamento e por acreditar em mim quando eu própria duvidava.

Agradeço também aos meus pais, pelo amor, paciência, esforço e compreensão, e por me incentivarem desde cedo a lutar pelos meus sonhos. Ao meu irmão Dinis, por trazer leveza, alegria e um abraço nos momentos mais difíceis.

Às minhas melhores amigas, Gabriela e Marta, obrigada por me acompanharem em todas as etapas e por acreditarem sempre. Agradeço também à minha amiga Sofia, pelo apoio, pelas palavras e por nunca me deixar desistir.

Por fim, quero agradecer à minha avó, que sempre me transmitiu o quão importante é a educação, e me viu começar este processo, mas a vida não permitiu que me visse terminar, espero que estejas orgulhosa onde quer que estejas, "vó".

A todos os meus amigos e família, pela motivação constante que contribuíram para que concluísse esta etapa.

Declaração de Integridade

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducente à sua elaboração. Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho.

Declaração de Uso de Ferramentas de Inteligência Artificial (IA)

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri a qualquer forma de utilização indevida de ferramentas de Inteligência Artificial (IA), nem à falsificação ou omissão de informações, fontes, dados, resultados ou processos em nenhuma das etapas conducentes à sua elaboração. Declaro ainda que qualquer utilização de ferramentas de IA foi feita de forma transparente e responsável, encontrando-se devidamente assinalada/descrita no trabalho, e que procedi à verificação crítica da exatidão do conteúdo produzido com o seu apoio, assumindo integralmente a responsabilidade pelo que submeto. Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho, bem como as normas aplicáveis de integridade académica e de proteção de dados.

Resumo

A criação de empresas tem um papel importante na economia de um país, dado o seu potencial para estimular a inovação, a competitividade e a criação de emprego. Porém os benefícios não são uniformes a nível territorial, podendo acentuar desigualdades regionais. Deste modo, torna-se crucial identificar os fatores que influenciam a criação de empresas.

Esta dissertação centra-se nos efeitos da estrutura sectorial existente, analisando os efeitos diferenciados da urbanização, da aglomeração e da diversificação na taxa de criação de empresas, nos 278 municípios portugueses de Portugal Continental. Calcula-se a taxa média de natalidade das empresas no período 2018-2023, considerando a indústria transformadora e os serviços (incluindo atividades de informação e comunicação, atividades imobiliárias, atividades de consultoria, científicas, técnicas e similares e atividades administrativas e dos serviços de apoio). Analisa-se se a criação de novas empresas é mais favorecida por uma estrutura sectorial diversificada ou mais especializada, considerando quer a diversificação relacionada quer a não relacionada.

Utilizam-se técnicas de análise de dados espaciais, usando dados publicados do Instituto Nacional de Estatística (INE).

Os resultados dos testes confirmam a presença de autocorrelação espacial. Ou seja, que a taxa de natalidade observada num município está correlacionada com as taxas de natalidade observadas nos municípios vizinhos. Este resultado indica que os modelos clássicos (OLS) não são adequados para explicar as diferenças regionais na criação de empresas, sendo necessário recorrer a modelos de econometria espacial que permitem incorporar estes efeitos.

Os resultados econométricos indicam que existe uma associação positiva entre a urbanização, diversificação em geral, diversificação relacionada, diversificação não relacionada, e a criação de empresas. Em contraste, a especialização apresenta uma associação negativa e estatisticamente significativa na formação de empresas.

Palavras-chave: *criação de empresas*, *urbanização*, *aglomeração*, *diversificação*, *análise espacial*

Abstract

Firm formation plays an important role in a country's economy, given its potential to stimulate innovation, competitiveness and job creation. However, the benefits are not evenly distributed across the country and may increase regional inequalities. It is therefore crucial to identify the factors that influence firm creation.

This dissertation focuses on the effects of the existing sectoral structure, analysing the differentiated effects of urbanisation, agglomeration and diversification on the business creation rate across the 278 municipalities of mainland Portugal. The average business birth rate for the period 2018–2023 is calculated, considering manufacturing and services (including information and communication activities, real estate activities, professional, scientific and technical activities, and administrative and support service activities). The study analyses whether the creation of new firms is more favoured by a diversified or a more specialised sectoral structure, considering different types of diversification, namely related and unrelated diversification.

Spatial data analysis techniques are employed, using published data from the National Statistics Institute (INE).

The test results confirm the presence of spatial autocorrelation. In other words, the birth rate observed in a municipality is correlated with the birth rates observed in neighbouring municipalities. This result indicates that classical models (OLS) are not suitable for explaining regional differences in business creation, making it necessary to use spatial econometric models that allow these effects to be incorporated.

The econometric results indicate that there is a positive association between urbanisation, diversification in general, related diversification, unrelated diversification, and firm creation. In contrast, specialisation shows a negative and statistically significant association with firm formation.

Keywords: firm creation, urbanisation, agglomeration, diversification, spatial analysis

Índice

Direitos de Autor e Condições de Utilização do Trabalho Por Terceiros i
Agradecimentos ii
Declaração de Integridade iii
Declaração de Uso de Ferramentas de Inteligência Artificial (IA) iii
Resumo iv
Abstract v
Índice de Tabelas viii
Índice de Figuras ix
1. Introdução 1
2. Revisão de Literatura 3
2.1. Criação de Empresas e Desenvolvimento Regional 3
2.2. Determinantes da Criação de Empresas 3
2.3. Externalidades de Aglomeração 5
2.3.1. Urbanização 6
2.3.2. Especialização 7
2.3.3. Diversificação 8
2.3.3.2. Diversificação Relacionada 9
2.3.3.2. Diversificação Não Relacionada 10
2.4. Especialização e Diversificação: Evidência Empírica 11
3. Metodologia 15
3.1. Objetivos da Investigação e Hipóteses 15
3.2. Dados, Período e Unidade Geográfica De Análise 15
3.3. Variável Dependente 15
3.4. Variáveis Independentes 17
3.4.1. Variáveis de Controlo 20
3.5. Modelos Econométricos 26
3.6. Autocorrelação Espacial 27
3.7. Índice I de Moran 27
3.8. Método de Estimação dos Modelos Econométricos 29
4. Apresentação de Resultados 32

4.1. Análise Descritiva das Taxas de Natalidade 32
4.2. Análise Exploratória de Dados Espaciais 34
4.2.1. Distribuição Espacial: mapas de quantis 34
4.2.2. Dependência Espacial: variável dependente 38
4.2.3. Dependência Espacial: variáveis independentes 39
4.3. Estimação dos Modelos Econométricos 41
5. Análise e Discussão dos Resultados 44
6. Conclusão 48
7. Limitações e Investigação Futura 49
Referências 50
Anexos 56

Índice de Tabelas

Tabela 1- Evidência empírica: determinantes 4
Tabela 2 - Especialização e Diversificação: Evidência Empírica 13
Tabela 3 - Tabela síntese das variáveis 21
Tabela 4 - Estatística Descritiva 22
Tabela 5 - Matriz de correlação de Pearson 25
Tabela 6 - Estatística descritiva da taxa de natalidade para cada ano 32
Tabela 7 - Os 10 municípios com maiores/menores taxas de natalidade 35
Tabela 8 - Índice de Moran Univariado 40
Tabela 9 - Resultados dos Testes de Multiplicador de Lagrange 42
Tabela 10 - Resultados de estimação dos modelos econométricos usando OLS e SEM 43

Índice de Figuras

Figura 1 - Evolução do Coeficiente de Gini 33
Figura 2 - Distribuição da taxa de natalidade em Portugal Continental 34
Figura 3 - Distribuição Espacial do Índice de Especialização 36
Figura 4 - Distribuição Espacial da Diversificação em geral, Diversificação Relacionada e Diversificação Não Relacionada 37
Figura 5 - Índice de Moran Univariado 39
Figura 6 - Mapa de clusters da taxa de natalidade (LISA) 40

1. Introdução

A criação de empresas atua como um indicador de competitividade da economia de um país, (Fritsch & Kublina, 2019; Kato-Vidal, 2022). Favorece o desenvolvimento económico, criando efeitos diretos (criação de emprego) e indiretos (mudanças na dinâmica do mercado) (Bosma et al., 2008), e é considerada crucial para a inovação e para a redução da pobreza (Pietak, 2021).

No entanto, os seus efeitos e a sua intensidade podem variar (Van Stel & Suddle, 2008) e, quando o nascimento não ocorre de forma homogénea no espaço, pode acentuar desigualdades regionais, manifestando-se por exemplo na concentração de recursos humanos e oportunidades. Estas disparidades regionais afetam então o crescimento económico, tornando-se importante compreender o papel das mesmas na economia (Pietak, 2021).

A literatura tem identificado diversos determinantes que explicam as diferenças regionais na formação de novas empresas, como por exemplo, o capital humano, a educação, a criatividade, o rendimento, e a estrutura industrial (Armington & Acs, 2002; Gajewski & Kutan, 2018; Jagódka & Snarska, 2023; O'Leary et al., 2023; Turečková & Nevima, 2020).

No entanto, os efeitos da concentração espacial das atividades económicas, nomeadamente a especialização e a diversificação ainda permanecem em debate na literatura. Por um lado, existem estudos que demonstram que a especialização impulsiona a criação de empresas, dado que a proximidade existente entre empresas do mesmo setor permite que existam mercados de trabalho e *inputs* mais especializados e eficientes (relacionada com as externalidades Marshall-Arrow-Romer). Por outro lado, existem estudos que mostram que é a diversificação industrial que fomenta a troca de ideias e permite que potenciais empreendedores identifiquem oportunidades. No mesmo sentido, a literatura mais recente incorporou os conceitos de diversificação relacionada e diversificação não relacionada, baseando-se na ideia de que a existência de setores relacionados entre si/não relacionados entre si influenciam a criação de empresas.

Deste modo, torna-se importante investigar o efeito da estrutura industrial e da aglomeração das atividades económicas na criação de empresas nos setores da indústria transformadora e dos serviços.

Concretamente, pretende-se responder às seguintes questões de investigação:

- Existe uma associação positiva entre a urbanização e a taxa de criação de empresas?
- A especialização está positivamente associada à taxa de criação de empresas, embora com efeitos heterogéneos entre setores?
- A diversificação está associada à criação de empresas?

- A diversificação relacionada está positivamente associada à taxa de criação de empresas?
- A diversificação não relacionada está associada à criação de empresas?

A seguir a esta introdução, o presente trabalho prossegue com a revisão da literatura, com foco nos fatores determinantes das diferenças regionais na criação de empresas, em particular nos relacionados com a estrutura industrial. No capítulo 3 é apresentada a metodologia do estudo, incluindo as variáveis dependentes e independentes, o método e os modelos empíricos utilizados. O capítulo 4 apresenta os resultados, e no capítulo 5 a sua análise e discussão. O trabalho termina com uma breve conclusão, a identificação das principais limitações do estudo e algumas recomendações para investigação futura.

2. Revisão de Literatura

2.1. Criação de Empresas e Desenvolvimento Regional

De acordo com Audretsch et al. (2021), o empreendedorismo consiste num processo em que as empresas identificam e exploram oportunidades que favorecem o desenvolvimento económico, podendo gerar efeitos diretos e indiretos no mesmo (Van Stel & Suddle, 2008). Os efeitos diretos relacionam-se, por exemplo, com a criação de emprego no início da atividade. Os efeitos indiretos estão associados a mudanças na dinâmica do mercado, como a eliminação de concorrentes menos eficientes, a destruição de postos de trabalho nas empresas menos competitivas (Gajewski & Kutan, 2018), ou a melhoria da competitividade do setor devido ao aumento da concorrência (Fritsch & Mueller, 2004).

Adicionalmente, as novas empresas estimulam a inovação, nomeadamente através da introdução de novos produtos no mercado, além de potencialmente contribuírem para o aumento da quantidade procurada de produtos e serviços, através da diminuição dos preços das empresas estabelecidas para resistirem às ameaças das novas (Van Stel & Suddle, 2008).

Porém, os efeitos da criação de empresas e a sua intensidade podem variar, dependendo do setor em que as novas empresas são criadas e da sua qualidade (Van Stel & Suddle, 2008). Além disso, embora as novas empresas possam desempenhar um papel crucial na promoção da inovação, do crescimento económico sustentável e na redução da pobreza, os seus benefícios não são uniformes a nível territorial, podendo acentuar as desigualdades regionais (Criscuolo et al., 2014; Decker et al., 2014; Gajewski & Kutan, 2018).

Efetivamente, Fritsch & Kublina (2019) salientam que a criação de empresas depende de fatores como investimentos, infraestruturas urbanas, instituições e pessoas qualificadas. Para estes autores, em períodos de crise económica, catástrofes naturais e pandemias, a formação de novas empresas tende a persistir ao longo tempo nas regiões que já eram empreendedoras antes destas crises .

Piętak (2022) constatou que as desigualdades regionais são um reflexo do desenvolvimento de muitos países que se manifesta, por exemplo, na concentração de recursos humanos e oportunidades.

Deste modo, torna-se importante identificar os fatores que explicam a diferenças regionais na criação de novas empresas.

2.2. Determinantes da Criação de Empresas

A literatura tem procurado identificar as razões pelas quais as empresas se formam, e porque é que este nascimento de empresas não acontece de forma igual no espaço, contribuindo para as desigualdades entre regiões (Armington & Acs, 2002).

São apresentadas diversas razões para explicar este fenómeno, nomeadamente as características dos empreendedores, como a idade e a experiência, das empresas, do mercado (intensidade de capital, novas entradas) e das condições económicas e de localização (taxa de desemprego, mão de obra qualificada, disponibilidade de capital humano) (Brixy et al., 2007; Power et al., 2019, 2020; Westlund et al., 2014). Especificamente, a dimensão do mercado (população e poder de compra), a diversidade, a cultura empreendedora, a disponibilidade de recursos, o rendimento e as limitações no financiamento, são frequentemente apontados como fatores que influenciam a dinâmica da criação de empresas (Armington & Acs, 2002; Gajewski & Kutan, 2018; Sutaria & Hicks, 2004).

Mais especificamente, alguns estudos têm investigado a relação entre a criação de empresas e a estrutura industrial existente na região (Basile et al., 2017; Corradini & Vanino, 2022; Fertala, 2008; Müller, 2016; Renski, 2014), existindo alguma evidência de que a urbanização, a especialização e a diversificação são fatores cruciais na estrutura industrial de uma região, que podem explicar a criação de novas empresas (O'Leary et al., 2023).

Desta forma, é possível perceber que existe uma grande variedade de fatores que influenciam a criação de empresas, pelo que a literatura propõe várias tipologias para agrupar/classificar os determinantes da criação de empresas (Tabela 1).

Tabela 1- Evidência empírica: determinantes

	Fatores determinantes	**Variáveis**
Davidsson et al. (1994)	▪ Disponibilidade de indivíduos motivados e capazes	▪ Nível de escolaridade e taxa de desemprego
	▪ Condições de mercado regional	▪ População, densidade populacional e crescimento do rendimento per capita
	▪ Acesso a capital	▪ Disponibilidade de capital privado, despesas governamentais de apoio às empresas
	▪ Clima favorável	▪ Cultura empreendedora, condições de vida
Sutaria & Hicks (2004)	▪ Condições da procura local	▪ Crescimento da população e crescimento do rendimento per capita
	▪ Mercado de trabalho	▪ Taxa de desemprego e variação da taxa de desemprego
	▪ Reestruturação industrial	▪ Dimensão média das empresas e taxa de saída de empresas
	▪ Capital financeiro	▪ Depósitos bancários *per capita*
	▪ Setor público	▪ Despesas governamentais
Bosma et al. (2008)	▪ Fatores de procura e oferta	▪ Crescimento da população, rendimento e desemprego
	▪ Aglomeração	▪ Economias de localização e de urbanização
	▪ Ambiente político	▪ Mudança política

Noutra perspetiva, há autores que salientam que os determinantes regionais da formação de novas empresas são diferentes consoante o setor de atividade. Por exemplo, o elevado desemprego leva ao aumento das taxas das *startups* apenas em empresas de serviços relativamente simples, enquanto os efeitos positivos do apoio às empresas podem ser identificados apenas em relação às *startups* na indústria (Davidsson, Lindmark, & Olofsson, 1993).

Dado que o foco deste trabalho incide sobre os determinantes de natureza territorial, nomeadamente as condições do mercado e a estrutura industrial que caracterizam o contexto regional e local, assume particular relevância a análise dos efeitos da aglomeração, da especialização/diversificação e das externalidades de aglomeração nas desigualdades regionais na criação de empresas.

2.3. Externalidades de Aglomeração

De acordo com Kim et al. (2022), a aglomeração é essencial para compreender a economia regional, e a criação de novas empresas tem um papel preponderante na formação e na manutenção destas aglomerações.

A aglomeração é um processo natural que se manifesta na tendência para existir uma concentração das atividades económicas no espaço e que surge da necessidade de indivíduos e empresas minimizarem custos de transporte e de transação. A proximidade geográfica facilita o acesso a mão de obra qualificada e diversificada, a partilha de fatores de produção de custo elevado e complexos, como máquinas e infraestruturas comuns de comunicação e transporte, bem como a troca de conhecimento tácito (Caragliu et al., 2016). Ou seja, a concentração de atividades económicas surge porque as empresas usufruem de vantagens ao localizarem-se próximas umas das outras e de outos agentes económicos (Frenken et al., 2007; Neffke et al., 2011). Esta dinâmica de aglomeração cria um ciclo de crescimento, que varia consoante o número e setor de atividade das empresas da região. Ou seja, quanto maior a concentração de empresas numa região, mais acentuados são os efeitos sobre a produtividade (Ebert et al., 2019). Este aumento da produtividade torna a região mais atrativa, atraindo emprego e, consequentemente uma maior aglomeração de empresas (Ebert et al., 2019). Neste contexto, a aglomeração depende da dimensão e da concentração espacial da atividade económica (Crowley et al., 2021).

Alguma literatura distingue três tipos de externalidades de aglomeração: a especialização, a diversificação e a urbanização (Basile et al., 2017; Neffke et al., 2011; O'Leary et al., 2023).

2.3.1. Urbanização

As externalidades de urbanização estão associadas à densidade da atividade económica (Basile et al., 2017).

A literatura sugere que áreas mais urbanizadas facilitam a acessibilidade a redes de transporte (sendo importante para as empresas, dado que a atividade económica envolve o transporte de *inputs* e *outputs*), o acesso a mercados locais ou globais, a capital humano com elevadas qualificações, a serviços empresariais especializados e a equipamentos culturais. Áreas mais urbanizadas permitem ainda explorar economias de escala (devido a uma procura maior), além de estarem associadas à existência de centros de investigação e desenvolvimento e universidades (Arauzo-Carod et al., 2010; Neffke et al., 2011; Renski, 2011).

As externalidades de urbanização podem influenciar a criação de empresas, ao impactar a procura local e o acesso a uma maior oferta de produtos e serviços (Basile et al., 2017). As empresas usufruem de vantagens ao localizarem-se em regiões com maior atividade económica (Badr et al., 2019; McCann & Folta, 2008), nomeadamente a disponibilidade de mão de obra, acesso a informação, *spillovers* tecnológicos e proximidade com clientes e fornecedores (Monseny et al., 2011). Além disso, empresas localizadas em áreas com maior densidade populacional tendem a ter custos mais baixos, dado que existe uma concentração da produção numa determinada região (Müller, 2016). Estas áreas com maior densidade populacional podem conduzir a níveis mais elevados de concorrência (Cainelli et al., 2014; Combes et al., 2012), porém a entrada de novas empresas pode também atrair mais clientes para uma dada região, e aumentar as vendas das empresas que já estão estabelecidas (Matsuyama, 1995).

No entanto, a urbanização também acarreta custos, como preços de terrenos mais elevados, mão de obra mais cara, congestionamento, poluição e maiores taxas de criminalidade (Arauzo-Carod et al., 2010; Capozza et al., 2018; Neffke et al., 2011; Renski, 2011). Desta maneira, as empresas só têm vantagem em localizarem-se em áreas com elevada densidade se as externalidades de urbanização superarem os custos de congestionamento (Basile et al., 2017).

Concluindo, as regiões centrais tendem a atrair mais empresas, dada a proximidade a possíveis clientes e fornecedores e a um mercado local mais vasto. No entanto, existem também desvantagens, como o congestionamento e os preços mais elevados dos terrenos, que podem desencadear um "equilíbrio de dispersão" (uma distribuição uniforme da atividade económica pelas regiões), ou um "equilíbrio de aglomeração" (uma concentração desproporcional da atividade) (Van et al., 2006).

Empiricamente, Renski (2014) encontrou uma relação positiva entre a densidade populacional

e a formação de empresas nos EUA. Também Andersson et al. (2019) para três cidades da Suécia, e Motoyama & Malizia (2017) para os EUA, demonstram que os benefícios existentes associados a áreas com maior densidade populacional podem ser um estímulo para impulsionar a criação de novas empresas. Recentemente, O'Leary et al. (2023), com base em 196 regiões da União Europeia de 16 países, no período de 2008 – 2017, concluíram que a urbanização (medida pela densidade populacional) afeta positivamente a criação de empresas. Mostraram também que a densidade populacional aumenta a formação de empresas, em consonância com Jacobs et al. (2014), que demonstram que aglomerações com maior densidade populacional podem ter um impacto positivo na criação de empresas, podendo dever-se ao facto de que áreas com elevada densidade têm maior acesso aos consumidores, bem como a fatores de produção que são necessários às empresas (Guo et al., 2016; Wagner & Sternberg, 2004).

Piacentino et al. (2017) consideram relevante distinguir entre dois tipos de comportamento empreendedor: empreendedorismo replicativo, em que as empresas replicam modelos de negócio já existentes, dependendo fortemente do contexto local, e empreendedorismo inovador, menos dependente das características do território. Usando econometria espacial e dados relativos à Itália, Piacentino et al. (2017) mostram que os efeitos da urbanização na criação de empresas são heterogéneos, dependendo do tipo de comportamento empreendedor: o comportamento replicativo das novas empresas parece depender significativamente da urbanização, enquanto que uma dependência semelhante não foi encontrada para o comportamento inovador.

2.3.2. Especialização

Marshall (1920) foi pioneiro a perceber as vantagens das empresas do mesmo setor se concentrarem numa determinada região. De acordo com o mesmo, a proximidade existente entre empresas do mesmo setor permite que existam mercados de trabalho e *inputs* mais especializados e eficientes, por consequência, os custos de transação são mais baixos e melhora a qualidade da troca de ideias, pessoas e bens.

Segundo Marshall (1923), quando as empresas se localizam perto umas das outras, tendem a surgir atividades complementares a nível de fornecedores, *inputs* e serviços especializados, bem como equipamentos específicos do setor. Esta concentração favorece uma divisão do trabalho mais eficiente e reduz os custos de produção para todas as empresas localizadas nesse território. Além disso, nas regiões onde estão localizadas empresas do mesmo setor desenvolve-se um mercado de trabalho especializado. As empresas beneficiam de oferta de trabalho especializado, podendo ajustar o número de trabalhadores de forma mais flexível consoante o ciclo de vida da empresa, e os trabalhadores não dependem apenas de uma empresa, aumentando a segurança laboral. As empresas beneficiam ainda do acesso a mão de

obra qualificada, contribuindo para a redução dos custos de contratação e para a diminuição do tempo e dinheiro na formação de novos funcionários (Noonan et al., 2021). Por fim, a proximidade geográfica favorece os *spillovers* de conhecimento, isto é, fluxos de informação não codificada, relacionados com as competências e conhecimentos específicos do setor, que se difundem entre as empresas localizadas próximas umas das outras. A partilha deste conhecimento tácito conduz à formação de um *know-how* próprio do setor, sendo este efeito mais forte quando as empresas são tecnologicamente semelhantes (Aharonson et al., 2006; Callejón, 2003; Ellison et al., 2010).

Marshall (1920), Arrow (1962) e Romer (1986), consideram que o conhecimento é, em grande medida, específico de cada setor e que a especialização regional gera externalidades intra-sectoriais pelo que a concentração das atividades económicas no espaço contribui para melhorar o desempenho económico. Neste sentido, regiões com estruturas produtivas especializadas num dado setor tendem a ser mais inovadoras nesse setor, dado que a proximidade entre empresas permite que o conhecimento se espalhe entre empresas semelhantes (Van Der Panne, 2004).

Adicionalmente, níveis mais elevados de especialização industrial, têm sido associados a níveis mais baixos de concorrência, o que pode incentivar a entrada de novas empresas mercado (Johan & Vania, 2022). Contudo, esta relação não é consensual na literatura. Numa perspetiva oposta, Ebert et al. (2019) referem que este tipo de externalidades pode promover o aumento da concorrência local devido à concentração de empresas com atividade económica semelhante. Para Joffe (2024), quando os níveis de especialização são elevados, isto pode indicar que existe um mercado com elevadas barreiras à entrada, limitando o potencial de criação de novas empresas.

Empiricamente, Garofoli (1994), para a Itália, conclui que a estrutura de produção, nomeadamente a especialização produtiva local é das variáveis explicativas mais importantes para a formação de novas empresas, apresentando um efeito positivo.

2.3.3. Diversificação

Jacobs (1969) apresenta a ideia de que a diversificação industrial fomenta a troca de ideias e permite que potenciais empreendedores identifiquem oportunidades. Esta perspetiva alega que a proximidade entre pessoas e atividades empresariais diversificadas pode ser estimulante, enquanto a existência de homogeneidade pode limitar a criatividade e a capacidade de gerar novas ideias. Deste modo, quando existe uma estrutura industrial diversificada, dá-se uma interação, adaptação e formação de ideias entre diferentes setores de atividade (Basile et al., 2017).

Jacobs (1969) argumenta que as cidades se desenvolvem com indústrias diversificadas, uma vez que o contacto entre trabalhadores de diferentes setores estimula a inovação. Esta perspetiva assenta

na ideia de que os *spillovers* de conhecimento podem ocorrer entre indústrias diferentes, visto que ideias desenvolvidas por um dado setor podem ser utilizadas noutros setores. Esta diversidade de setores impulsiona o conhecimento complementar, facilitando a exploração e a experimentação associadas ao processo de inovação. Desta forma, as externalidades de diversificação resultam de uma estrutura setorial diversificada, estando associada a rendimentos crescentes (Van Der Panne, 2004). Assim, este tipo de externalidades facilita a partilha de conhecimento, permitindo a invenção de novos produtos e serviços, e criando novos mercados (Ebert et al., 2019). Seguindo este raciocínio, os setores e as empresas beneficiam por se localizarem geograficamente próximas de empresas diferentes, uma vez que o conhecimento pode ser adquirido de empresas não relacionadas. Audretsch & Belitski (2020) denominam este processo como *spillovers* interindustriais. Estes *spillovers* interindustriais relacionam-se com as externalidades de Jacobs (1969). Ou seja, as principais fontes de conhecimento advêm de fora do setor principal que a empresa atua, e é o conhecimento que existe entre vários setores que leva a maiores retornos (Acs et al., 2006).

Frequentemente, a literatura distingue entre diversificação relacionada e diversificação não relacionada (Caragliu et al., 2016). Boschma (2005) refere que as externalidades podem resultar da variedade intraindustrial (diversificação relacionada) e da variedade interindustrial (diversificação não relacionada). No entanto, os efeitos destas diversidades não são mutuamente exclusivos (Caragliu et al., 2016; Frenken et al., 2007).

2.3.3.2. Diversificação Relacionada

A diversificação relacionada baseia-se na ideia de que a proximidade tecnológica entre setores facilita a inovação e *spillovers* de conhecimento intersectoriais (Frenken et al., 2007). Neste sentido, argumenta-se que o conhecimento tende a disseminar-se sobretudo entre empresas de setores relacionados, uma vez que existe complementaridade ao nível das competências e das bases de conhecimento (Capozza et al., 2018).

Desta forma, a diversificação relacionada traduz-se num equilíbrio entre a diversidade cognitiva e a similaridade entre empresas (Crowley et al., 2021). A informação que flui entre as empresas deve ser suficientemente semelhante para que seja compreendida, mas simultaneamente diferente para que seja nova (Boschma & Iammarino, 2009). Quando os indivíduos conseguem interligar este conhecimento diferente, mas similar, identificam oportunidades empreendedoras, influenciando positivamente a criação de empresas (Content et al., 2019). Assim, quanto maior a diversificação relacionada numa dada região, maior tenderá a ser a circulação do conhecimento, contribuindo de forma positiva para a formação de empresas (Crowley et al., 2021).

No entanto, o equilíbrio referido pode também gerar externalidades negativas para as empresas. Efetivamente, níveis elevados de similaridade podem criar problemas de dependência cognitiva, situação que ocorre quando a informação que flui entre as empresas deixa de ser útil porque já não é nova (Crowley et al., 2021).

Assim, Boschma & Iammarino (2009) consideram que há um nível adequado de proximidade cognitiva. A proximidade cognitiva é importante para que haja *spillovers* de conhecimento, mas em demasia pode levar a problemas de dependência cognitiva. Por outro lado, demasiada diversificação não relacionada impediria a difusão de conhecimento, originando bloqueios cognitivos que dificultariam a inovação e a criação de ideias, levando à redução do número de empresas criadas (Cohen & Levinthal, 1990; Thrane et al., 2010).

Em suma, entre os argumentos que suportam a hipótese de que a diversificação relacionada favorece a criação de empresas, salientam-se:

- O conhecimento dissemina-se apenas entre empresas de setores relacionados, existindo complementaridade no que diz respeito a competências (Capozza et al., 2018);
- A proximidade cognitiva promove os *spillovers* de conhecimento (Cohen & Levinthal, 1990; Frenken et al., 2007; Thrane et al., 2010)
- A combinação de conhecimento diferente, mas relacionado, que facilita a identificação de oportunidades empreendedoras (Content et al., 2019).

2.3.3.2. Diversificação Não Relacionada

A diversificação não relacionada ocorre quando, numa região, há uma concentração de empresas que operam em setores tecnologicamente diferentes (Aarstad et al., 2016; Crowley et al., 2021). A literatura sugere que este tipo de diversificação pode promover a troca de diferentes tipos de conhecimento, pelo que regiões com uma maior diversidade não relacionada poderiam estar relacionadas com maiores taxas de criação de empresas (O'Leary et al., 2023).

Numa perspetiva complementar, a diversificação não relacionada também pode ser entendida como uma estratégia de portefólio, na medida em que a diversificação permite proteger melhor as regiões da exposição a choques assimétricos (choques económicos) na procura. Segundo Frenken et al. (2007), as regiões tornam-se menos vulneráveis a crises que afetam particularmente sectores específicos, o que contribui para mitigar o aumento do desemprego nesses períodos.

Desta forma, presume-se que a diversificação não relacionada esteja associada a um maior dinamismo regional, dada a existência de vários setores numa região que permitem distribuir o risco (Boschma & Iammarino, 2009; Tavassoli & Jienwatcharamongkhol, 2016).

Em suma, há argumentos que suportam a hipótese de que a diversificação não relacionada favorece indiretamente a criação de empresas ao reduzir a volatilidade e o risco:

- Maior proteção contra choques económicos assimétricos (Frenken et al., 2007);
- Um maior dinamismo regional que permite distribuir o risco (Boschma & Iammarino, 2009; Tavassoli & Jienwatcharamongkhol, 2016).

2.4. Especialização e Diversificação: Evidência Empírica

A literatura tem vindo a estudar o papel da especialização industrial na criação de empresas, existindo diversos estudos que encontraram um efeito positivo na criação de novas empresas (Tabela 2).

Alguns estudos têm testado os efeitos da especialização em setores específicos. Por exemplo, Anyadike-Danes & Hart (2006), mostraram que, no Reino Unido, a especialização em serviços empresariais tem efeito positivo na formação de novas empresas, sendo estas regiões também caracterizadas, em geral, por uma maior densidade populacional. Van Soest et al. (2006), com base em dados relativos aos Países baixos, encontraram evidência de que o efeito da especialização não é uniforme, contribuindo para a criação de empresas nos setores dos serviços às empresas e aos consumidores, mas não observaram efeitos estatisticamente significativos na indústria transformadora.

Também Bishop (2012) no Reino Unido, refere que os efeitos positivos da diversificação na criação de empresas dependem do setor.

Barboza & Pede (2024) observaram que os efeitos dos *spillovers* de conhecimento intraindustriais são mais fortes em níveis mais baixos de agregação de dados, mas, a nível regional, os elevados níveis de especialização industrial influenciam negativamente a criação de novas empresas.

Lasch et al. (2013) analisaram o papel da diversificação industrial e da especialização, mostrando que o efeito da diversificação foi estatisticamente significativo, ao passo que a especialização não teve um resultado significativo, levando a que os autores concluíssem que os efeitos da estrutura industrial são parcialmente rejeitados. Estes resultados foram de acordo por exemplo, com o estudo de Bade & Nerlinger (2000). Não obstante, para Frenken et al. (2007) e Ejdemo & Örtqvist (2020) a diversificação só tem impacto positivo se as indústrias ou conhecimentos estiverem tecnologicamente relacionados.

No que diz respeito à diversificação relacionada, existem estudos que evidenciam uma relação positiva com a criação de empresas (Content et al., 2019; Ejdemo & Örtqvist, 2020).

Um estudo realizado por Kato-Vidal (2022), concluiu que regiões onde existe uma maior diversificação industrial relacionada têm uma maior resiliência e contribuem mais para a formação de novas empresas. Além disso, observou que na crise de COVID-19 existiu uma maior resiliência em regiões

com maior diversificação relacionada, dado que os setores que estavam relacionados entre si sofreram menos impactos negativos. Persistiu uma maior abertura de empresas em regiões com maior variedade de setores relacionados e a densidade empresarial manteve-se alta. Em regiões com baixa diversificação relacionada, deu-se uma quebra na densidade empresarial, visto que existiram menos empresas novas a abrir e mais empresas a fechar.

Também autores como Bishop (2012), Bishop & Brand (2014) e Baltzopoulos et al. (2016) e encontraram efeitos positivos da diversificação relacionada na criação de empresas, na indústria transformadora e nos serviços. Em contraste, Kanellopoulos & Fotopoulos (2018) e O'Leary et al. (2023) concluíram que a diversificação relacionada tem efeito negativo e significativo na criação de empresas.

No que diz respeito à diversificação não relacionada, Bishop (2012) e Colombelli (2016) encontraram evidência de que a diversificação não relacionada impulsiona a formação de empresas. No entanto, O'Leary et al. (2023) verificaram que a diversificação não relacionada não tem um efeito estatisticamente significativo na criação de empresas, mas quando se tem em conta efeitos espaciais, verifica-se que níveis mais elevados de diversidade não relacionada estão associados a um aumento na criação de empresas.

Já Bishop (2012) encontrou evidência de que tanto a diversificação relacionada como a não relacionada têm impacto positivo na criação de empresas no Reino Unido, mas que a primeira aparenta ser mais importante.

Por fim, os resultados do estudo de Kato-Vidal (2022) mostraram que a diversificação não relacionada influencia mais a densidade empresarial do que a criação de novas empresas. Bishop & Brand (2014) e Kanellopoulos & Fotopoulos (2018) verificaram que a diversificação não relacionada não teve efeito estatisticamente significativo na formação de novas empresas na indústria transformadora, porém, observou-se um impacto positivo e significativo no setor dos serviços.

A Tabela 2 sintetiza a evidência empírica acerca da especialização, diversificação em geral, diversificação relacionada e diversificação não relacionada.

Tabela 2 - Especialização e Diversificação: Evidência Empírica

Artigo	País(ses)	Setor(es)	Especialização	Diversificação	Diversificação relacionada	Diversificação não relacionada
Anyadike-Danes & Hart (2006)	Reino Unido	Agricultura, Produção, Distribuição e Serviços empresariais	+			
Van Soest et al. (2006)	Países Baixos	Serviços empresariais, Indústria transformadora e Serviços ao consumidor	+ (depende do setor)	+		
Lasch et al. (2013)	França	Empresas dos setores de tecnologias de informação e comunicação		+		
Fotopoulos (2014)	Reino Unido	15 setores (agricultura, energia, indústria transformadora, construção, distribuição, hotéis e restauração, serviços, ...)	+	-		
Guo et al. (2016)	China	Indústria transformadora	+	Efeito ambíguo	+	-
Kanellopoulos & Fotopoulos (2018)	Grécia	Indústria transformadora	+	-	-	Não significativa
Bishop (2019)	Grã Bretanha	Indústria transformadora de alta tecnologia Serviços intensivos em conhecimento			Não significativa	+
Kim et al. (2022)	Austrália	Indústria transformadora de alta e baixa tecnologia, serviços intensivos em conhecimento, outros serviços	+	Não significativa		
O'Leary (2022)	148 regiões da Europa	Todos	-		-	+
O'Leary et al. (2023)	16 países da Europa	Todos	+ (sem pesos) Não significativa (com pesos)		-	- (sem pesos) + (com pesos)
Barboza & Pede (2024)	Itália	19 setores	+ (nível província) - (nível regional)	+		

Conclusão

A literatura tem identificado diferentes fatores que explicam as diferenças entre regiões nas taxas de criação de novas empresas, nomeadamente o emprego, a densidade e crescimento populacional, a estrutura industrial, o nível de rendimento, acesso a capital humano e disponibilidade de financiamento (Armington & Acs, 2002).

No entanto, não existe consenso sobre os efeitos das externalidades de aglomeração, quer da especialização quer da diversificação na criação de novas empresas. Existem estudos, que encontram evidências sobre os efeitos positivos e negativos da diversificação e da especialização, e da coexistência dos dois como Campi et al. (2004). Van Oort & Stam (2005) observaram que a especialização e a diversificação estão associados positivamente à criação de novas empresas e ao crescimento das empresas existentes. Campi et al. (2004) verificaram que em Espanha, os efeitos da especialização e da diversificação coexistem numa mesma região.

3. Metodologia

3.1. Objetivos da Investigação e Hipóteses

Neste capítulo é apresentada a metodologia de análise utilizada para responder ao objetivo principal desta investigação, isto é, analisar os efeitos das externalidade de aglomeração e da estrutura industrial na criação de empresas.

Com base na revisão de literatura, formularam-se as seguintes hipóteses:

H1: Existe uma associação positiva entre a urbanização e a taxa de criação de empresas.

H2: A especialização está positivamente associada à taxa de criação de empresas, embora com efeitos heterogéneos entre setores.

H3: A diversificação está associada à criação de empresas.

H4: A diversificação relacionada está positivamente associada à taxa de criação de empresas.

H5: A diversificação não relacionada está associada à criação de empresas.

3.2. Dados, Período e Unidade Geográfica De Análise

A unidade geográfica de análise corresponde aos 278 municípios que compõem o território de Portugal Continental.

Todos os dados utilizados são provenientes do INE e dos Censos para o período de análise 2018-2023.

Para o cálculo das taxas de criação de empresas foram considerados os setores da indústria transformadora e dos serviços (incluindo neste último apenas as atividades de informação e comunicação, atividades imobiliárias, atividades de consultoria, científicas, técnicas e similares e atividades administrativas e dos serviços de apoio). A escolha destes serviços relaciona-se com o facto de serem serviços mais intensivos em conhecimento, permitindo excluir setores como a restauração que têm uma maior volatilidade. Na indústria transformadora foi excluída a indústria do tabaco que, por ser um setor fortemente regulado, enfrenta barreiras à entrada elevadas.

3.3. Variável Dependente

A literatura propõe três abordagens para medir a taxa de criação de novas empresas: (i) a abordagem do mercado de trabalho, (ii) a abordagem populacional e (iii) a abordagem ecológica, as quais diferem na variável utilizada para fazer a normalização necessária para ser possível realizar comparações. Efetivamente, a análise do número de empresas criadas não permitiria fazer comparações

devido às diferenças de tamanho existentes entre regiões (Jung et al., 2012; Kangasharju, 2000).

A abordagem do mercado de trabalho considera o número de novas empresas em relação ao número de trabalhadores. Esta abordagem parte do pressuposto de que os empreendedores criam empresas no mesmo setor onde estavam a trabalhar ou a procurar emprego, uma vez que adquirem experiência nessa área (Audretsch & Fritsch, 1994; Fotopoulos, 2014; Kangasharju, 2000).

A abordagem populacional relaciona o número de novas empresas com a população. Esta abordagem pressupõe que as empresas são criadas numa região pelos indivíduos que vivem nessa região. A criação de empresas é motivada principalmente pelas oportunidades ou necessidades específicas daquela região (Jung et al., 2012; Kangasharju, 2000).

Por fim, a abordagem ecológica (dos negócios estabelecidos) tem em conta a relação entre as novas empresas e o stock de empresas existentes (Audretsch & Fritsch, 1994). Esta abordagem pressupõe que o número de empresas criadas em cada região não é independente das empresas nela estabelecida. As empresas existentes funcionam como incubadoras para novas empresas e ambientes de aprendizagem para os novos empreendedores. Esta abordagem traz desvantagens porque sobrestima o número de novas empresas em zonas onde predominam grandes empresas, e subestima em áreas onde predominam as pequenas empresas, e também porque as pequenas empresas acabam por incentivar mais pessoas a tornarem-se empreendedores do que as grandes empresas, uma vez que os gestores estão mais próximos a nível operacional nas pequenas empresas (Kangasharju, 2000; Lawton Smith, 1991; Spilling, 1996).

Seguindo Piacentino et al. (2017), a variável dependente é definida pelo rácio entre o número de empresas no município e a população ativa entre os 15 e os 64 anos de idade, enquadrando a criação de empresas como um fenómeno ligado à base populacional potencialmente ativa:

$$TN = \frac{nº\ empresas\ criadas_{\ ano\ t}}{população\ em\ idade\ ativa_{\ ano\ t}} \times 1000 \qquad (1)$$

Esta abordagem tem a vantagem de captar a propensão empreendedora da população em idade ativa, independentemente de estar empregada ou desempregada. Entre as desvantagens, salienta-se o facto de ignorar a estrutura industrial existente.

Primeiramente, foi calculada taxa de natalidade para cada ano e para cada município, através do quociente entre o número de empresas criadas (na indústria transformadora e nos serviços) e a população (1000 habitantes) em idade ativa no mesmo ano.

Seguindo o procedimento usado em outros estudos (por exemplo em Bishop (2012) e Piacentino

et al. (2017)), foi calculada a taxa média de natalidade de empresas para o período 2018–2023 com o objetivo de suavizar eventuais variações estocásticas. No entanto, é de salientar que se observa uma correlação quase perfeita entre anos consecutivos na taxa de empresas criadas em cada município. Evidência semelhante foi reportada Bishop (2012), relativamente ao Reino Unido (acima de 0,95 para todos os pares de anos adjacentes).

3.4. Variáveis Independentes

Urbanização

As externalidades de urbanização são medidas pela variável URBAN. De acordo com a literatura, as externalidades de urbanização estão associadas à densidade da atividade económica e podem impactar a procura local, influenciando a criação de empresas. Tal como em O'Leary et al. (2023) esta variável é definida pelo logaritmo da densidade populacional (habitantes por quilómetro quadrado).

Especialização

A especialização é operacionalizada pela variável SPECIAL. Marshall (1920), Arrow (1962) e Romer (1986) defendem que a concentração no espaço das atividades económicas melhora o desempenho económico. Por um lado, níveis mais elevados de especialização industrial, estão relacionados com níveis mais baixos de concorrência, incentivando empreendedores a entrarem no mercado (Johan & Vania, 2022). Por outro, níveis de especialização elevados, pode indicar que existe um mercado com elevadas barreiras à entrada (Joffe, 2024).

Seguindo a literatura, a especialização é calculada de acordo com Piacentino et al. (2017), Fotopoulos (2014) e Kanellopoulos & Fotopoulos (2018):

$$SPECIAL = \frac{1}{2} \times \Sigma_i \left| \frac{E_{ir}}{E_r} - \frac{E_{in}}{En} \right| \quad (2)$$

onde E_{ir} é o número de estabelecimentos na indústria i na região r, E_r é o número total de estabelecimentos na região, E_{in} é o número de estabelecimentos no setor i a nível nacional e E_n é o número total de estabelecimentos a nível nacional. Este índice de especialização varia entre 0 (ausência de especialização: a estrutura produtiva é idêntica à nacional) e 1 (máxima especialização: estrutura produtiva concentrada num único setor), quando o valor é 1 significa que apenas um setor está presente na região e 0 quando a estrutura setorial da região é semelhante à do país como um todo.

O cálculo deste índice foi considerado um nível de desagregação da Classificação das Atividades Económicas (CAE) a 2 dígitos. Devido a não haver dados disponíveis ao nível do município, com este

grau de desagregação, para o emprego, foi utilizado como *proxy* o número de estabelecimentos por localização geográfica e atividade económica referente ao ano de 2017.

Diversificação

Jacobs (1969) apresenta a ideia de que a diversificação industrial fomenta a troca de ideias e permite que potenciais empreendedores identifiquem oportunidades, explicando que a proximidade entre pessoas e atividades empresariais pode ser estimulante. Consequentemente, dá-se uma interação, adaptação e formação de ideias entre diferentes setores de atividade (Basile et al., 2017).

Seguindo a literatura a diversificação (DIVERSIF) foi calculada da seguinte forma (Fotopoulos, 2014; Frenken et al., 2007):

$$DIVERSIF = \sum \left(\frac{Pr_i}{P_r}\right) \log \left(\frac{P_r}{P_{ri}}\right) \quad (3)$$

onde P_{ri} é a proporção do número de estabelecimentos do setor i na região r, e P_r é a proporção do número de estabelecimentos total da região. Para limitar ao intervalo [0,1], divide-se o valor por log (80), uma vez que são incluídos 80 subsetores (CAE 2 dígitos). Quando o valor é 0 apenas um setor está presente na região, e tem o valor máximo ln (n) quando todos os setores têm o mesmo número de estabelecimentos.

No cálculo deste índice também foi considerado um nível de desagregação da CAE a 2 dígitos, e o número de estabelecimentos por localização geográfica e atividade económica relativo ao ano de 2017.

Diversificação Relacionada

A diversificação relacionada (DIVERS_R) baseia-se na ideia de que setores tecnologicamente interdependentes facilitam a inovação e *spillovers* de conhecimento entre setores (Frenken et al., 2007). Assim, quando indivíduos conseguem interligar este conhecimento, identificam oportunidades empreendedoras, influenciando positivamente a criação de empresas (Content et al., 2019). No entanto, elevados níveis de similaridade podem levar a problemas de dependência cognitiva (Crowley et al., 2021).

Nesta dissertação, a diversificação relacionada é definida com base nos artigos de Crowley et al. (2021) e O'Leary et al. (2023), considerando o nível de desagregação CAE a 2 dígitos:

$$DIVERS_R = \sum_j \; P_{j,i} \left(\sum_{h \in (j,i)} \frac{P_{h,i}}{P_{j,i}} log_2 \left(\frac{1}{\frac{P_{h,i}}{P_{j,i}}} \right) \right) \quad (4)$$

onde, i é o município, j representa os setores a CAE a 1 dígito, h reflete os setores a um nível de desagregação CAE a 2 dígitos, $P_{j,i}$ é a proporção de estabelecimentos do setor j no total do município i, $\frac{P_{h,i}}{P_{j,i}}$ representa a proporção do setor h (CAE 2 dígitos) dentro do setor j (CAE 1 dígito) e $log_2\left(\frac{1}{\frac{P_{h,i}}{P_{j,i}}}\right)$ é o logaritmo de base 2 do inverso da proporção de estabelecimentos, para medir o grau de diversificação interna. Quanto maior os valores do índice da diversificação relacionada, maior é a variedade dentro dos setores de atividade.

Diversificação Não Relacionada

A diversificação não relacionada caracteriza-se pela coexistência, numa região, de empresas de setores distintos e sem relações de complementaridade (Aarstad et al., 2016; Crowley et al., 2021). A ideia é de que a diversificação entre setores pode, em determinados contextos, promover a troca de diferentes tipos de conhecimento, pelo que, regiões com uma maior diversificação não relacionada poderiam estar relacionadas com maiores taxas de criação de empresas (O'Leary et al., 2023). Contudo, elevados níveis de diversificação não relacionada podem também desencadear bloqueio cognitivo e, consequentemente, a redução do número de empresas criadas (Cohen & Levinthal, 1990; Thrane et al., 2010).

Nesta dissertação a diversidade não relacionada ($DIVERS_NR$) é definida com base nos artigos de Crowley et al. (2021) e O'Leary et al. (2023) considerando o nível de desagregação CAE a 2 dígitos:

$$DIVERS_{NR} = \sum_j P_{ji}, log_2\left(\frac{1}{P_{ji}}\right) \quad (5)$$

onde, i é o município, j são os setores agregados a CAE a 1 dígito, e P_{ji} é a proporção do número de estabelecimento do setor j no total de estabelecimentos do município i, e $log_2\left(\frac{1}{P_{ji}}\right)$ representa o logaritmo de base 2, do inverso da proporção do número de estabelecimentos, permitindo medir a forma como os estabelecimentos estão distribuídos pelos diferentes setores. Quanto maior o índice de diversificação não relacionada, maior é a variedade de setores de atividade diferentes.

Apesar de a probabilidade de correlação ser muito elevada, Aarstad et al. (2016) defendem que os conceitos de diversificação industrial relacionada e não relacionada divergem na conotação e, por isso, merecem ser tratados como conceitos distintos.

3.4.1. Variáveis de Controlo

Desemprego

O papel do desemprego na criação de empresas tem sido muito debatido na literatura. Por um lado, pessoas desempregadas podem estar mais propícias a iniciar os seus próprios negócios, o que contribui para a criação de empresas e para redução do desemprego (Armington & Acs, 2002). Este fenómeno está associado ao "*push mechanism*", em que os desempregados são forçados a trabalhar por conta própria por não encontrarem outras alternativas no mercado de trabalho (Calá et al., 2016).

Por outro lado, pode dar-se o "*pull mechanism*", pois um elevado nível de desemprego pode demonstrar problemas na economia regional, como procura fraca e desajustamentos estruturais no mercado de trabalho, diminuindo a atratividade e, consequentemente, fazendo com que exista menos propensão para criar empresas (Calá et al., 2016).

Assim, a criação de empresas pode ser influenciada por estas forças de sinal contrário: o desemprego pode levar alguns indivíduos a empreender por necessidade, mas também pode desencorajar outros pelas expectativas de sucesso diminutas. Gajewski & Kutan (2018) referem que a coexistência destas forças opostas de "*push*" *and* "*pull*" dependem das características específicas de cada país (como, por exemplo, as medidas de apoio ao desemprego) e podem ter efeitos diferentes no tempo. No curto prazo, quando as pessoas perdem o emprego, podem ser forçadas a trabalhar por conta própria devido à falta de oportunidades. No entanto, no longo prazo, se o desemprego persistir elevado na região pode afetar negativamente a criação de empresas.

Esta ideia está de acordo com a evidência observada por Storey (1991) de que diversas análises de séries temporais indicam que o desemprego está relacionado positivamente com a formação de novas empresas, enquanto dados transversais ou em painel mostram o oposto. Já Tervo & Niittykangas (1994) encontram evidência de que a formação de empresas a nível regional está negativamente relacionada com o nível de desemprego, mas positivamente com o crescimento do desemprego. Ou seja, o primeiro está relacionado com a falta de oportunidades de negócio, e o segundo com a diminuição das oportunidades de emprego remunerado, que leva indivíduos desempregados a trabalharem por conta própria.

O desemprego pode impulsionar ou retrair a criação de empresas, não existindo um consenso sobre o efeito do mesmo. A variável desemprego (DESEMP) foi calculado pelo número de desempregados por cada 100 ativos, segundo dados dos Censos de 2011.

Densidade empresarial

A criação de empresas funciona como um indicador de competitividade das economias, sendo

necessário investimentos, infraestruturas urbanas e recursos humanos qualificados. Desta forma, em regiões mais dinâmicas, existe uma maior densidade empresarial, isto é, um maior número de empresas, podendo ser medida por habitante ou por quilómetro quadrado (Kato-Vidal, 2022).

Neste estudo, utiliza-se a variável densidade empresarial (DENSEMP), calculada pelo quociente entre o número de empresas e a área de superfície em cada município.

O uso desta variável, tal como é medida, tem como vantagem ser robusta face a diferenças na dimensão da área das regiões, e captar os benefícios que se dão devido à concentração espacial das atividades económicas (Ciccone & Hall, 1996; Ganau, 2015), permitindo controlar as diferenças na distribuição do tamanho dos estabelecimentos entre regiões (Li et al., 2011).

Nos estudos de Kanellopoulos & Fotopoulos (2018) para a Grécia, Okamuro & Nobuo (2005) para o Japão, e Colombelli & Quatraro (2018) para a Itália, a densidade empresarial foi calculada pelo número de empresas por quilómetro quadrado, apresentando efeitos positivos e estatisticamente significativos na criação de empresas. Já Armington & Acs (2002) nos Estados Unidos, e Barboza & Pede (2024) na Itália, usaram o quociente entre o número de empresas e a população, mostrando também efeitos positivos e significativos, porém, neste último estudo verificou-se um efeito mais forte a nível de províncias do que a nível regional.

Tabela 3 - Tabela síntese das variáveis

Variáveis	**Definição**	**Efeito esperado**	**Fonte**
Variáveis independentes			
Urbanização (*URBAN)*	Logaritmo da densidade populacional por quilómetro quadrado em 2017	Positivo	INE
Especialização (*SPECIAL)*	Índice de especialização calculado pelo número de estabelecimentos em 2017	Ambíguo	INE
Diversificação (*DIVERSIF*)	Índice de diversificação calculado através do número de estabelecimentos, em 2017.	Ambíguo	INE
Diversificação relacionada *(DIVERS_R)*	Diversificação dentro de cada setor a CAE 2 dígitos, calculada com base no número de estabelecimentos, em 2017	Ambíguo	INE
Diversificação não relacionada *(DIVERS_NR)*	Grau de diversificação entre os setores a CAE 2 dígitos, através do número de estabelecimentos em 2017.	Ambíguo	INE
Variáveis de controlo			
Desemprego *(DESEMP)*	Número de desempregados por cada 100 ativos em 2011	Ambíguo	INE
Densidade empresarial *(DENSEMP)*	Número de empresas por quilómetro quadrado em 2017	Positivo	INE

Na Tabela 4, apresenta-se a estatística descritiva das variáveis utilizadas nos modelos econométricos.

Tabela 4 - Estatística Descritiva

Variável	Observações	Média	Desvio Padrão	Mínimo	Máximo
TNAT	278	9.642	3.897	3.408	32.224
URBAN	278	4.307	1.521	1.524	8.888
SPECIAL	278	0.260	0.119	0.080	0.640
DIVERSIF	278	0.652	0.090	0.324	0.747
DIVERS_R	278	0.919	0.217	0.292	1.311
DIVERS_NR	278	3.204	0.373	1.753	3.662
DESEMP	278	12.498	2.833	5.090	22.850
DENSEMP	278	36.818	115.279	0.561	1112.484

Na análise da Tabela 4, são notórias as diferenças nas taxas de natalidade de empresas entre municípios. Em média, foram criadas cerca de 9.6 novas empresas por 1000 habitantes em idade ativa. No período analisado, o município que obteve a maior média de taxa de natalidade foi Lisboa, com uma taxa de natalidade de 32‰, mais de dez vezes superior à menor taxa, que foi registada no município de Mesão Frio (3.4‰ novas empresas por 1000 habitantes em idade ativa).

No que diz respeito à variável urbanização, apresenta uma média de 4.307, oscilando entre o valor mínimo de 1.5 e o valor máximo de 8.89. Os municípios com valores próximos do mínimo correspondem a municípios com uma menor urbanização, como por exemplo, Alcoutim, Mértola e Idanha - a - Nova. Já os municípios com valores próximos do máximo estão associados a uma maior urbanização, por exemplo, Amadora, Odivelas e Porto.

O índice de especialização evidencia alguma heterogeneidade entre municípios no grau de especialização sectorial. Como referido anteriormente, este varia entre 0 e 1, onde 0 significa ausência de especialização e 1 representa a máxima especialização. Entre os municípios com valores mais elevados, encontramos Santa Marta de Penaguião, São João da Pesqueira e Alfândega da Fé, associados a estruturas económicas mais especializadas. Os valores mais baixos do índice de especialização em 2017 foram observados nos municípios de Caldas da Rainha, Mafra e Palmela. Apesar da literatura não estabelecer limiares absolutos para o índice de especialização, alguns estudos empíricos associam valores inferiores a 0.30 – 0.35 a estruturas económicas relativamente diversificadas e próximas da média nacional (Traistaru et al., 2002). Assim, os resultados do cálculo desta variável mostram que os municípios de Santa Marta de Penaguião (0.64), São João da Pesqueira (0.63) e Alfândega da Fé (0.60),

apresentam elevados índices de especialização e acima da média (0.260). Em contrapartida, os municípios de Caldas da Rainha (0.080), Mafra (0.084) e Palmela (0.086), apresentam estruturas económicas menos especializadas com valores inferiores à média nacional.

Relativamente à variável que mede a diversificação em geral, regista uma média de 0.652, tendo um valor mínimo de 0.324 e um valor máximo de 0.747. Salienta-se que esta variável também varia entre 0 e 1, onde 0 significa a ausência de diversificação e 1 representa a máxima diversificação, quando os setores têm o mesmo peso na estrutura económica (após ser normalizado pela divisão ln(n)). Municípios com maior diversificação tendem a ter uma maior variedade de atividades económicas, como Vila Nova de Cerveira, Águeda e Porto de Mós. Por outro lado, os três municípios com menor diversificação de indústrias são Santa Marta de Penaguião, São João da Pesqueira e Vinhais. Deste modo, Vila Nova de Cerveira (0.747), Águeda (0.747) e Porto de Mós (0.745) apresentam níveis de diversificação acima da média (0.652), contrariamente, Santa Marta de Penaguião (0.324), São João da Pesqueira (0.344) e Vinhais (0.357) revelam valores inferiores à média nacional e consistentes com os resultados do índice de especialização, e, consequentemente, uma menor diversificação de atividades económicas.

A variável diversificação relacionada apresenta uma média de 0.92. Os valores mais baixos estão associados a municípios como Santa Marta de Penaguião, São João da Pesqueira, Vinhais, Mesão Frio e Barrancos, onde a estrutura económica é constituída por atividades semelhantes. Em contrapartida, os valores mais elevados são observados em municípios como Trofa, Águeda, Ourém, Guimarães e Oliveira do Hospital, que apresentam um maior grau de diversificação de setores relacionados entre si.

A variável diversificação não relacionada varia entre 1.73 e 3.33, apresentando um valor médio de 3.2. Os municípios de Vinhais e Murça estão entre os que apresentam níveis mais baixos de diversificação não relacionada. Os valores mais elevados deste índice estão associados aos municípios de Constância (3.662) e Arruda dos Vinhos (3.603).

Por não existirem dados mais recentes para os municípios, as taxas de desemprego usadas são as relativas aos Censos de 2011, utilizadas como *proxy* das condições estruturais do mercado de trabalho. Baseado nos Censos de 2011, em média, existe uma taxa de desemprego de 12.5%. Os cinco municípios com uma menor taxa de desemprego são Oleiros (5.1%), Arruda dos Vinhos (7.7%), Batalha e Ferreira do Zêzere (7.9%) e Arouca (8.12%). Por outro lado, os municípios de Mourão (22.9%), Ponte de Sor (20.9%), Serpa (20.5%), Cinfães e Moura (20%) apresentavam as taxas de desemprego mais elevadas.

A densidade empresarial apresenta uma elevada dispersão, evidenciando diferenças significativas entre municípios. Alcoutim é o município com menor concentração de empresas, enquanto Lisboa tem maior concentração empresarial.

Após a realização da estatística descritiva das variáveis, calculou-se, a matriz de correlação entre as variáveis (Tabela 5). A matriz de correlação de *Pearson* é utilizada como um instrumento exploratório para identificar potenciais problemas de colinearidade entre as variáveis explicativas, permitindo analisar a existência de relações lineares entre pares de varáveis.

Tabela 5 - Matriz de correlação de Pearson

Variáveis	*TNAT*	*URBAN*	*SPECIAL*	*DIVERSIF*	*DIVERS_R*	*DIVERS_NR*	*DESEMP*	*DENSEMP*
TNAT	1.000							
URBAN	0.549***	1.000						
SPECIAL	- 0.411***	- 0.563***	1.000					
DIVERSIF	0.380***	0.465***	- 0.929***	1.000				
DIVERS_R	0.360***	0.456***	- 0.809***	0.936***	1.000			
DIVERS_NR	0.369***	0.444***	- 0.944***	0.979***	0.844***	1.000		
DESEMP	0.039	0.174***	- 0.031	- 0.034	- 0.077	- 0.007	1.000	
DENSEMP	0.648***	0.609***	- 0.169***	0.153**	0.151**	0.145**	0.123**	1.000

Nota: A tabela apresenta os coeficientes de correlação de Pearson, efetuados no Stata. Níveis de significância: **** p<0.01, ** p<0.05, * p<0.1*

Observando a matriz de correlação ilustrada na Tabela 5, é possível verificar que quase todas as variáveis estão associadas positivamente com a variável dependente aos níveis de significância de 1%, 5% e 10%. As exceções são as variáveis SPECIAL que apresenta uma associação negativa, e a variável DESEMP que não mostra uma relação estatisticamente significativa.

Estes resultados preliminares sugerem que municípios com maiores níveis de especialização parecem estar associados a menores taxas de natalidade empresarial. Em sentido contrário, a urbanização e todos os tipos de diversificação aparentam estar associadas a maiores taxas de natalidade nos municípios. Importa salientar, no entanto, que estas associações têm um carácter exploratório. Uma interpretação mais aprofundada só é possível através de uma análise econométrica multivariada.

No que diz respeito às variáveis independentes, observa-se uma correlação elevada entre a diversificação e a especialização (- 0.929), bem como entre a diversificação relacionada e a diversificação não relacionada (0.844). Estes valores sugerem a existência de problemas de multicolinearidade que desaconselham a inclusão simultânea destas variáveis no modelo econométrico.

3.5. Modelos Econométricos

Tendo em conta a análise da matriz de correlação, procedeu-se à especificação dos modelos econométricos. Dada a elevada correlação entre algumas variáveis explicativas, nomeadamente a especialização, diversificação, diversificação relacionada e diversificação não relacionada, estas variáveis foram introduzidas separadamente nos modelos, para evitar potenciais problemas de multicolinearidade.

Assim, foram estimados os seguintes modelos:

Modelo1:

$$TNAT = \alpha + \beta_1 URBAN + \beta_2 SPECIAL + \beta_3 DESEMP + \beta_4 DENSEMP + \varepsilon$$

Modelo 2:

$$TNAT = \alpha + \beta_1 URBAN + \beta_2 DIVERSIF + \beta_3 DESEMP + \beta_4 DENSEMP + \varepsilon$$

Modelo 3:

$$TNAT = \alpha + \beta_1 URBAN + \beta_2 DIVERS_R + \beta_3 DESEMP + \beta_4 DENSEMP + \varepsilon$$

Modelo 4:

$$TNAT = \alpha + \beta_1 URBAN + \beta_2 DIVERS_NR + \beta_3 DESEMP + \beta_4 DENSEMP + \varepsilon$$

em que, TNAT é a média da taxa de natalidade empresas no período de análise; URBAN representa a urbanização; SPECIAL mede a especialização sectorial; DIVERSIF refere-se à diversificação em geral, DIVERS_R representa a diversificação relacionada; e DIVERS_NR representa a diversificação não relacionada. As variáveis DESEMP e DENSEMP, são as varáveis de controlo, e referem-se ao desemprego e à densidade empresarial, respetivamente.

3.6. Autocorrelação Espacial

A econometria espacial revela-se mais adequada quando são identificados problemas de autocorrelação espacial, estando esta associada à interdependência económica existente entre regiões (Van Stel & Suddle, 2008). No modelo clássico de regressão linear assume-se que as observações são independentes entre si (não se correlacionam), pelo que os modelos estimados por MQO deixam de ser apropriados quando existe correlação espacial entre regiões, uma vez que este pressuposto é violado (Ponciano & Scalon, 2010; Zhao & Yu, 2021). Deste modo, torna-se necessária a utilização de modelos econométricos que incorporem explicitamente a dependência espacial, de forma a captara heterogeneidade espacial subjacente aos dados.

A análise espacial procura descrever os padrões existentes nos dados espaciais e estabelecer relações entre as diversas variáveis (Ponciano & Scalon, 2010). Quando estes dados espaciais apresentam autocorrelação espacial, as estimativas do modelo devem ter em conta essa estrutura espacial, dado que a dependência existente altera a capacidade explicativa do modelo. Logo, se tiver isto em conta, a robustez dos resultados vai ser maior, e consequentemente a capacidade preditiva do modelo vai aumentar (Lopes et al., n.d.).

De acordo com Druck et al. (2002), a caracterização da dependência espacial é importante na análise exploratória dos dados espaciais, dado que mostra como os valores de uma variável estão correlacionados no espaço. Entre os principais métodos estatísticos utilizados na literatura para esse fim, destacam-se as medidas de autocorrelação espacial, como o Índice I de Moran e o variograma.

3.7. Índice I de Moran

De acordo com a literatura (Bishop & Shilcof, 2017; Cerdán & Hernández, 2013), ao longo do tempo, foram desenvolvidos diversos métodos estatísticos para analisar o grau de dependência espacial, tanto global como local. A dependência espacial global avaliando a existência de autocorrelação espacial na distribuição de uma variável no território, ou seja, se valores semelhantes tendem a localizar-se próximos no espaço. A análise local da dependência espacial permite identificar a existência de padrões espaciais específicos em determinadas áreas geográficas, e baseando - se no cálculo de indicadores locais de autocorrelação espacial (*Local Indicators Spatial Association* - LISA).

O índice I de Moran mede a autocorrelação espacial global, avaliando a associação entre o valor de uma variável numa região (y_i) e a média ponderada dos valores dessa variável nas regiões vizinhas de i, expressa pelo termo Wy_i que é definido a partir de uma matriz de pesos espaciais (Caragliu et al., 2016; Cerdán & Hernández, 2013).

A matriz de pesos espaciais W tem um papel central na análise espacial, na medida em que é esta matriz que define a estrutura de vizinhança entre as unidades geográficas. Ou seja, é a matriz de pesos (ou ponderadores) espaciais que descreve a forma como cada unidade espacial (município) está relacionada com as restantes: cada elemento da matriz w_{ij} indica a intensidade da ligação espacial entre as regiões (municípios) i e j. Esta matriz pode ser construída com base em diferentes critérios, sendo os mais utilizados na literatura os seguintes (Bishop, 2012):

- Contiguidade, baseada na existência de uma fronteira comum entre unidades geográficas;
- Distância, que considera vizinhas as unidades localizadas dentro de um determinado raio;
- K - vizinhos mais próximos, em que cada unidade espacial é associada aos k unidades mais próximas.

Normalmente, a matriz é normalizada por linhas, de modo que a soma dos pesos de cada região seja igual a 1, permitindo interpretar Wy como uma média ponderada dos valores da variável nas regiões vizinhas.

O I de Moran é calculado da seguinte forma (Caragliu et al., 2016; Cerdán & Hernández, 2013):

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (6)$$

onde n é o número de regiões analisadas;

wij é um elemento da matriz de pesos espaciais;

y_i é o valor da variável na região i;

$\bar{y}$ é a média da variável no conjunto das regiões;

e S_0 é a soma de todos os elementos da matriz w_{ij}:

$$S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \quad (7)$$

Quando a matriz de pesos espaciais **W** está normalizada por linhas (*row-standardized*), o quociente $\frac{n}{S_0} = 1$. O índice *I* de Moran é, portanto, uma medida de autocorrelação espacial global, em que o numerador é uma medida da covariância espacial e o denominador é a variância da variável em análise. Sob a hipótese nula de ausência de autocorrelação espacial, o valor esperado do índice é dado por: $E\ (I) = -\frac{1}{n-1}$.

Deste modo, verifica-se a existência de autocorrelação espacial positiva quando valores semelhantes da variável tendem a agrupar-se no espaço, formando clusters espaciais. A autocorrelação espacial negativa ocorre quando valores dissemelhantes se encontram próximos no espaço. Este índice pode, assim, assumir valores positivos ou negativos, consoante o padrão espacial observado. Quanto à magnitude, em geral, com matrizes de pesos espaciais normalizadas, o índice tende a assumir valores entre -1 e 1. Valores próximos de zero indicam autocorrelação espacial fraca ou inexistente e valores mais afastados de zero indicam associações espaciais mais fortes.

Adicionalmente, o diagrama de dispersão de Moran permite identificar quais as observações que contribuem de forma positiva e negativa para a autocorrelação espacial global. Além disso, permite detetar *outliers* espaciais que podem exercer um efeito desproporcionado sobre o I de Moran, afetando a sua interpretação.

No diagrama de dispersão de Moran, o eixo dos x apresenta os valores normalizados da variável em análise, e no eixo dos y representa o respetivo desfasamento espacial. A existência e o tipo de autocorrelação espacial são avaliados a partir da reta de regressão ajustada aos pontos do diagrama (Anselin, 1988). Uma inclinação positiva indica autocorrelação espacial positiva, enquanto uma inclinação negativa reflete autocorrelação espacial negativa. Quando a inclinação é próxima de zero, a autocorrelação espacial é fraca ou inexistente.

No que diz respeito aos quadrantes, a concentração das observações no primeiro (alto-alto) ou no terceiro (baixo-baixo) quadrantes, indica associações espaciais positivas, correspondendo a situações em que valores acima/abaixo da média são rodeados por valores também acima/abaixo da média. O segundo (baixo-alto) e quarto (alto-baixo) quadrantes correspondem a associações espaciais negativas, em que os valores observados numa dada área são diferentes dos observados na sua vizinhança. Ou seja, o segundo quadrante inclui as observações com valores inferiores à média que estão rodeadas por valores superiores à média. Por outro lado, os que estão no quarto quadrante têm valores superiores à média, enquanto as observações mais próximas apresentam valores inferiores à média (Cerdán & Hernández, 2013).

3.8. Método de Estimação dos Modelos Econométricos

O comportamento empreendedor depende das características da área geográfica onde a nova empresa é formada e das características das regiões vizinhas. Existindo evidências de efeitos espaciais, as regiões não podem ser tratadas como unidades estatísticas espacialmente independentes. Deste modo, os modelos de regressão linear clássicos, ao não terem em conta a estrutura espacial dos dados,

podem conduzir a estimativas ineficientes e a inferência estatística incorreta nos casos em que existe autocorrelação espacial (LeSage & Kelley Pace, 2009; Piacentino et al., 2017).

Nestes casos, é necessário recorrer a modelos espaciais para explorar a intensidade e a significância das interações espaciais presentes nos dados. Estas interações podem ser endógenas (quando o comportamento de uma unidade espacial depende do valor da variável dependente noutros locais), exógenas (quando o comportamento é influenciado pelas variáveis explicativas observadas nos locais vizinhos) ou resultar de efeitos correlacionados, quando fatores não observados levam a comportamentos semelhantes entre unidades espaciais (Piacentino et al., 2017).

Uma forma de testar a existência e a natureza destas interações espaciais consiste em partir da estimação de um modelo de regressão linear mais geral, que engloba todos estes efeitos (Manski, 1993; Piacentino et al., 2017):

$$Y = \rho WY + X\beta + WX\theta + u \ ;$$

$$u = \lambda Wu + \varepsilon \ ,$$

em que a interação espacial é especificada pela matriz de pesos espaciais $\mathbf{W}$, que é uma matriz quadrada de dimensão N por N, cujos elementos genéricos w_{ij} medem a intensidade do efeito espacial entre as unidades espaciais i e j. De um modo geral, os elementos da matriz são não negativos e os elementos diagonais da matriz são iguais a 0 (Anselin et al., 1996).

Como foi referido atrás, a matriz de pesos espaciais pode ser obtida considerando como critério de vizinhança, como a contiguidade ou a distância, estando associadas às características físicas das unidades espaciais num mapa (Caragliu et al., 2016; Piacentino et al., 2017).

A especificação da dependência espacial num modelo de regressão linear pode ser abordada de maneiras diferentes. Entre as especificações mais utilizadas destacam-se o *Spatial Lag Model*, no qual é adicionado um termo que corresponde à variável dependente com desfasamento espacial, e *o Spatial Error Model*, em que a dependência espacial está no termo de erro, refletindo a influência de fatores não observados correlacionados no espaço (Piacentino et al., 2017). O *Spatial Durbin Model* (SDM) corresponde a uma abordagem mais geral, que inclui simultaneamente a variável dependente desfasada espacialmente e as variáveis explicativas desfasadas, permitindo modelar tanto efeitos espaciais endógenos como exógenos.

Vários estudos têm utilizado ferramentas de econometria espacial para analisar o fenómeno de criação de novas empresas, evidenciando a importância da dependência espacial. Por exemplo, Bishop (2012), utilizando um *Spatial Lag Model,* analisou o impacto da diversidade de conhecimento entre setores na criação de empresas, tendo encontrado evidência de dependência espacial no Reino Unido. No mesmo sentido, Levratto (2015) estudou a criação de novas empresas em regiões francesas utilizando diferentes modelos espaciais, concluindo que existem efeitos de contágio espacial negativos.

Normalmente, os estudos começam por estimar um modelo de regressão linear sem efeitos espaciais, através do método dos mínimos quadrados ordinários (OLS). Em seguida, depois de definir a matriz de pesos espaciais a utilizar, recorrem aos testes do Multiplicador de Lagrange (LM), clássico e robusto, propostos por Anselin et al. (1996), de modo a detetar a presença de dependência espacial nos dados e a identificar a sua natureza dominante. Quando estes testes identificam autocorrelação espacial nos resíduos ou dependência espacial na variável dependente, o modelo OLS revela-se inadequado, devendo ser usado um modelo de regressão espacial apropriado. Após a estimação do modelo espacial, compara-se os modelos, utilizando nomeadamente o teste *Likelihood Ratio* (Piacentino et al., 2017).

A análise foi realizada com recurso ao software GeoDa.

4. Apresentação de Resultados

Neste capítulo são apresentados os resultados da análise desenvolvida. Primeiramente, é realizada uma análise descritiva das desigualdades regionais na criação de empresas e da persistência dessas desigualdades ao longo do tempo. Posteriormente, apresenta-se a análise exploratória dos dados espaciais, com o objetivo de identificar padrões de associação espacial (aglomerados espaciais ou *clusters*) e eventuais observações atípicas (valores extremos ou *outliers*). Por fim, são apresentados os resultados da estimação dos modelos econométricos.

4.1. Análise Descritiva das Taxas de Natalidade

Esta secção apresenta os resultados da análise descritiva das taxas de natalidade, considerando cada ano individualmente, de modo a evidenciar a evolução das desigualdades regionais no nascimento de empresas ao longo do tempo, recorrendo a medidas de dispersão. Esta análise é complementada pelo cálculo do coeficiente de Gini, a fim de avaliar o nível de desigualdade relativa da distribuição. Esta secção termina, com a identificação dos municípios com maiores e menores taxas de natalidade.

Na Tabela 6 é apresentada a estatística descritiva da taxa de natalidade das empresas no setor da indústria transformadora e dos serviços em Portugal Continental, nos anos de 2018 a 2023, permitindo analisar a evolução temporal da taxa de natalidade.

Tabela 6 - Estatística descritiva da taxa de natalidade para cada ano

TNAT	**2018**	**2019**	**2020**	**2021**	**2022**	**2023**
Observações	278	278	278	278	278	278
Média	9.354	8.892	7.559	9.284	11.047	11.714
Desvio-padrão	3.578	4.003	2.806	3.756	5.483	5.412
Coeficiente de variação	0,38	0.45	0.37	0.40	0.50	0.46
Mínimo	3.599	1.999	1.458	2.270	2.492	1.853
Máximo	30.655	29.876	22.351	31.570	44.518	42.071

Nota: cálculos efetuados com recurso ao package estatístico Stata

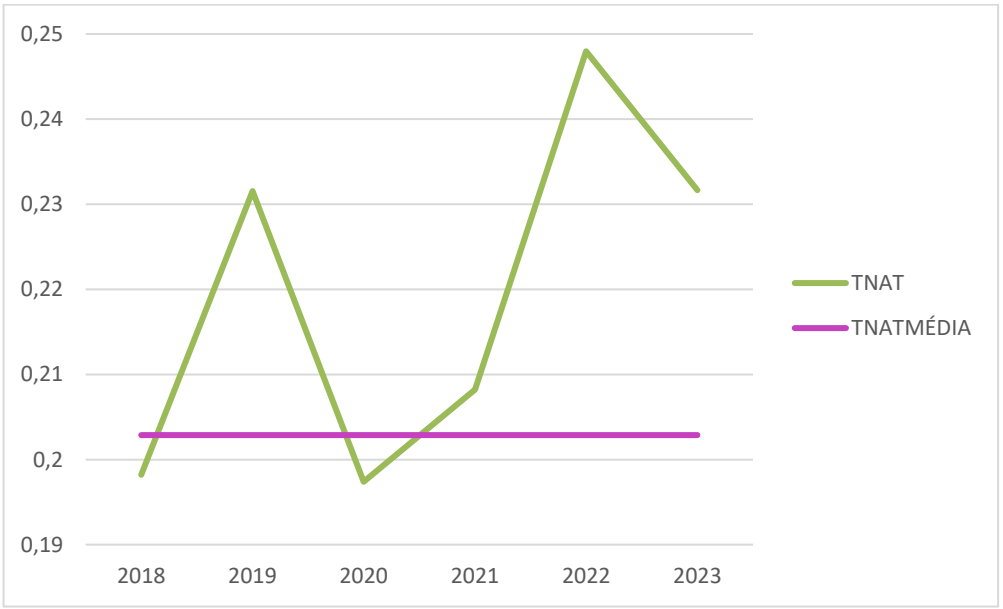
Verifica-se na Tabela 6 uma tendência de crescimento da taxa de criação de novas empresas a partir de 2020. O ano de 2023, registou a taxa de natalidade mais elevada do período, tendo sido criadas, em média, quase 12 empresas por cada 1000 indivíduos em idade ativa. No extremo oposto, o ano de 2020, que corresponde ao ano mais crítico da crise pandémica (COVID19), registou a taxa de natalidade mais baixa do período.

No que concerne às desigualdades regionais, quer o desvio – padrão, quer o coeficiente de variação indicam uma tendência de aumento da dispersão em termos absolutos e relativos. Esta evolução sugere um aumento das diferenças regionais.

De forma a medir o nível de desigualdade relativa na distribuição das taxas de natalidades entre 2018-2023, utilizou-se o coeficiente de Gini (Figura 1). O cálculo deste coeficiente complementa a análise da dispersão, pois permite captar em que medida a distribuição das taxas de natalidade se afastam de uma distribuição perfeita. Por ser normalizado (varia entre 0 e 1, aumentando à medida que a desigualdade aumenta) facilita comparações temporais e espaciais (Bishop & Shilcof, 2017).

Na Figura 1, observa - se que os anos de 2018 e 2020 apresentam os valores mais baixos do coeficiente de Gini, sugerindo uma distribuição mais homogénea das taxas de natalidade entre os municípios. No lado oposto, destaca-se o ano de 2022, com um coeficiente de aproximadamente 0.25. Apesar de os coeficientes refletirem, em geral, uma desigualdade baixa entre os municípios do Continente, não deixa de ser relevante notar uma tendência de agravamento nos anos seguintes ao da crise pandémica, que merece análise adicional.

Figura 1 - Evolução do Coeficiente de Gini



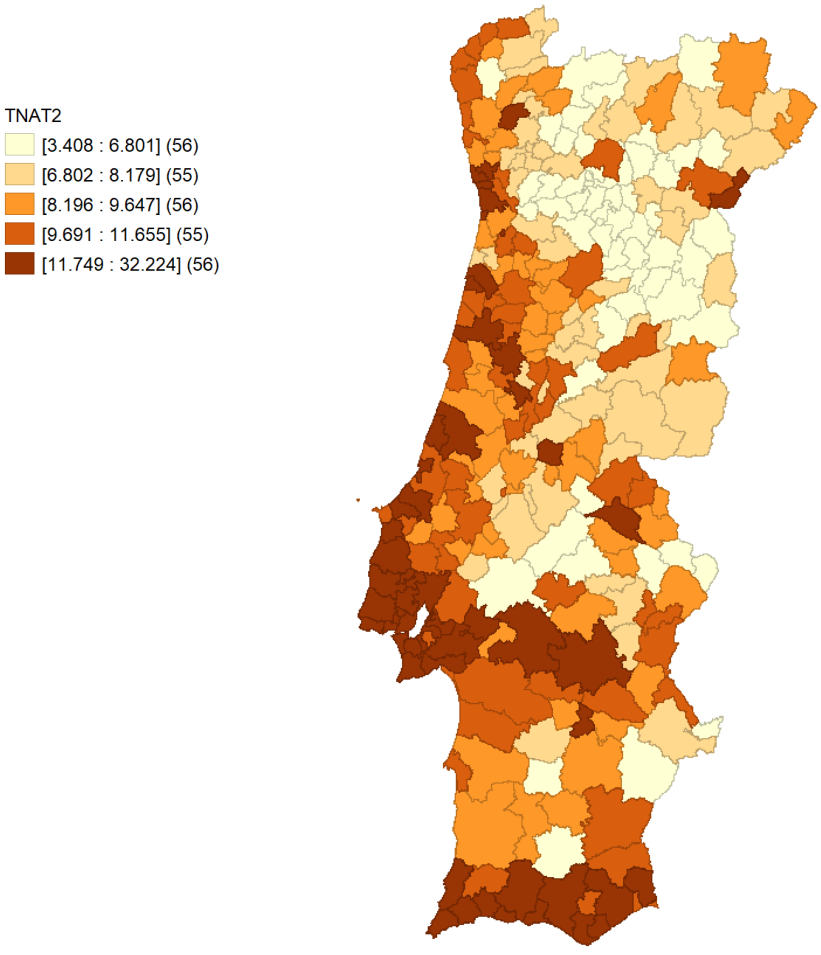
4.2. Análise Exploratória de Dados Espaciais

Nesta secção apresenta-se uma análise exploratória dos dados com recurso ao software GeoDa. Em primeiro lugar, são apresentados os mapas de quantis referentes à variável dependente (taxa de natalidade) e às variáveis independentes relacionadas com a especialização e a diversificação. Seguidamente, investiga-se a existência da dependência espacial das variáveis em estudo, com base em estatísticas de autocorrelação espacial. Por fim, usando mapas LISA, identifica-se padrões de associação local, com o objetivo de detetar *clusters* espaciais.

4.2.1. Distribuição Espacial: mapas de quantis

As desigualdades na distribuição espacial da taxa de natalidade de empresas (média no período 2018 – 2023) estão ilustrada na Figura 2.

Figura 2 - Distribuição da taxa de natalidade em Portugal Continental



A Figura 2 apresenta um mapa de quantis com cinco classes, que permite identificar a posição relativa dos municípios na distribuição da taxa de natalidade. A figura sugere que há diferenças evidentes

entre o litoral e o interior, sobretudo nas regiões do Norte e Centro. Efetivamente, é possível observar que os municípios que estão nos quantis mais baixos se concentram sobretudo em territórios do interior. Os quantis superiores mostram que apenas 10% dos municípios apresentam valores acima de 13,88 ‰, e que somente 5% dos municípios têm uma taxa de natalidade superior a 16.47. Em contrapartida, metade dos municípios portugueses têm taxas de natalidade inferiores a 8,75, e, 20% tem taxas de natalidade inferiores a 7.27‰ (Anexo 2).

Posteriormente, foi elaborada uma tabela onde se identificam os dez municípios com maiores e menores taxas de natalidade em Portugal Continental, de modo a perceber as assimetrias existentes no território nacional (Tabela 7).

Na Tabela 7, apresenta-se a lista dos dez municípios com maiores e menores taxas de natalidade. Como é possível constatar, os municípios com maiores taxas localizam-se no litoral, mais especificamente nas regiões de Lisboa e do Porto e no Algarve. No lado oposto, os municípios com menores taxas de natalidade localizam-se essencialmente em territórios do interior, normalmente associados a menores níveis de atividade económica e dinamismo empresarial.

Tabela 7 - Os 10 municípios com maiores/menores taxas de natalidade

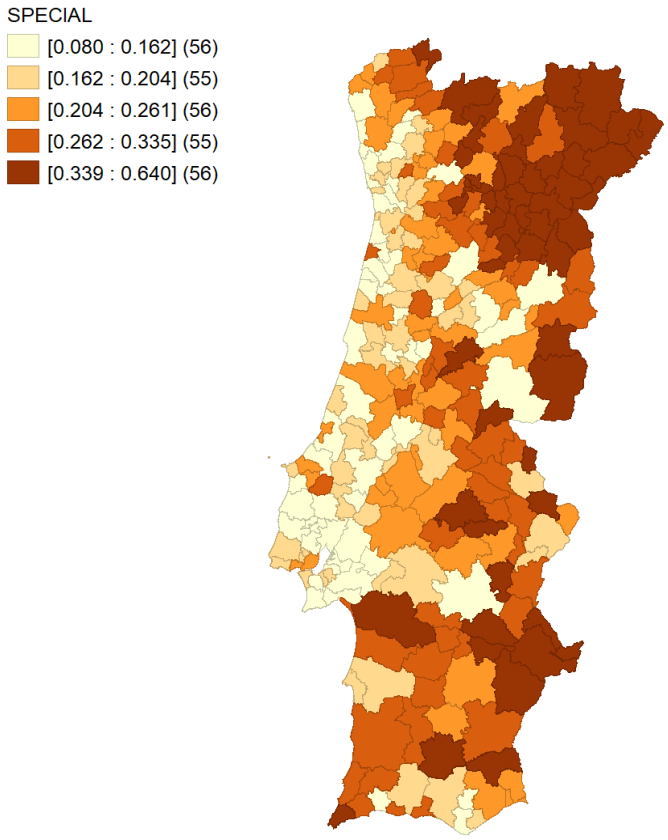
Municípios com maiores taxas de natalidade	**TNAT (‰)**	**Municípios com menores taxas de natalidade**	**TNAT (‰)**
Lisboa	33.224	Mesão Frio	3.408
Vila Nova de Gaia	27.096	Castro Daire	3.933
Cascais	25.217	Armamar	4.407
Porto	24.732	Santa Marta de Penaguião	4.425
Oeiras	22.633	Fornos de Algodres	4.666
Amadora	20.328	Aguiar da Beira	4.749
Odivelas	20.323	Resende	4.800
Loulé	20.144	Cinfães	4.883
Almada	19.479	Baião	4.886
Albufeira	18.976	Cabeceiras de Basto	4.900

Especialização

Depois de analisada a distribuição espacial da taxa de natalidade média do período 2018-2023, passa-se à caracterização espacial das variáveis explicativas relacionadas com as externalidades de aglomeração.

A Figura 3 apresenta a distribuição espacial do índice de especialização nos municípios portugueses, evidenciando a existência de assimetrias entre municípios.

Figura 3 - Distribuição Espacial do Índice de Especialização



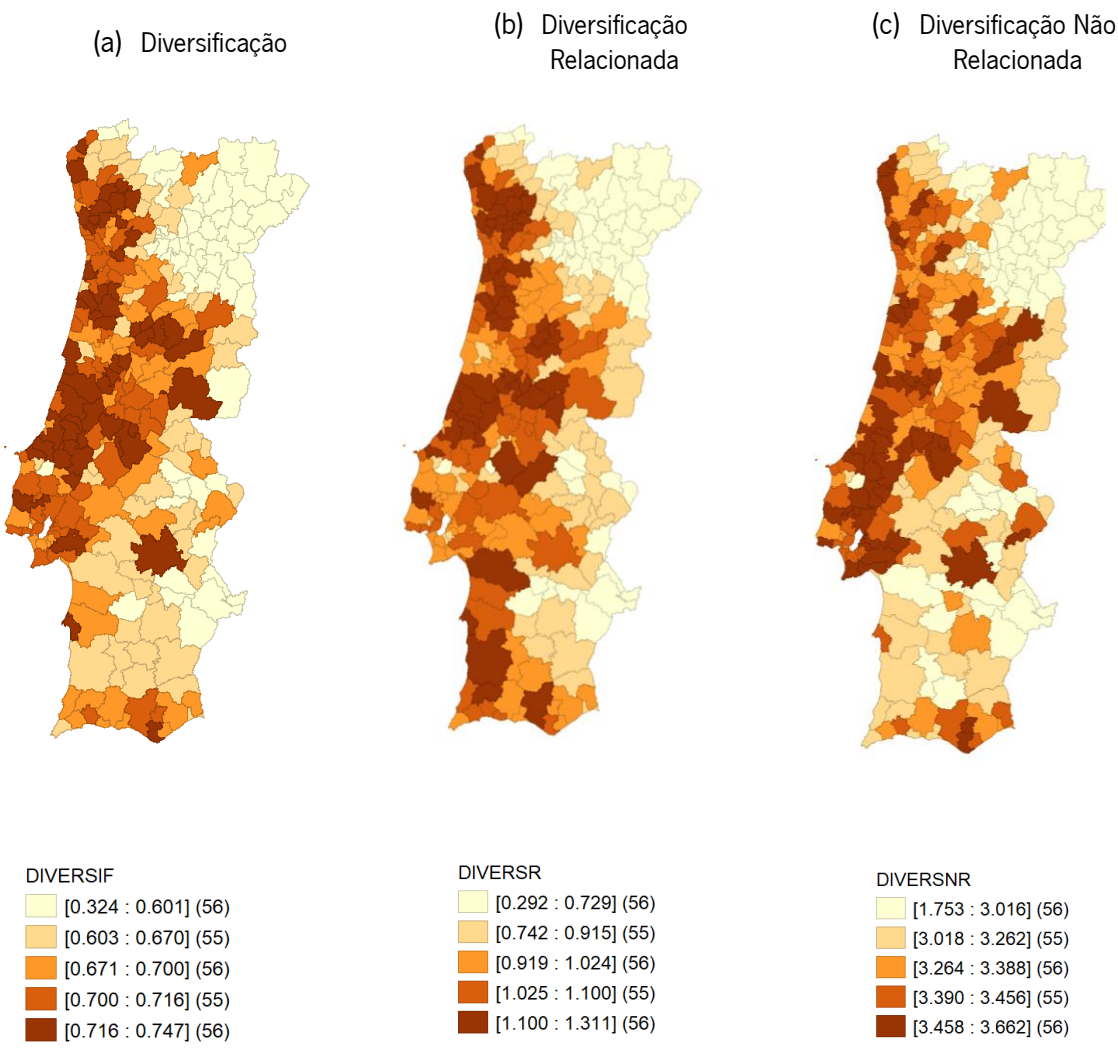
Na Figura 3, observam-se elevados níveis de especialização no Interior Norte e no Alentejo. Em contraste, níveis mais baixos de especialização encontram-se essencialmente no Litoral Norte e Centro, particularmente na Região do Porto e na Região de Lisboa. Adicionalmente, observa-se que metade dos municípios tem um índice de especialização inferior a 0.23, o que sugere uma estrutura produtiva pouco

concentrada. Os índices de especialização mais elevados (superiores a 0.54) estão concentrados em apenas 5% dos municípios (Anexo 3). A diferença entre a média (0,26) e o quartil superior sugere uma distribuição assimétrica à direita, ou seja, poucos municípios com níveis elevados de especialização

Diversificação

Na Figura 4, está representado o padrão de diversificação de Portugal Continental.

Figura 4 - Distribuição Espacial da Diversificação em geral, Diversificação Relacionada e Diversificação Não Relacionada



Como podemos observar, existe alguma heterogeneidade territorial nos níveis de diversificação. Três das cinco classes concentram-se entre 0.60 e 0.71, o que indica uma forte concentração dos municípios em níveis intermédios de diversificação. É no litoral, principalmente no Litoral Norte e Centro,

onde existe uma atividade económica mais diversificada. Níveis mais baixos de diversificação observam-se maioritariamente no interior do país, existindo 97 municípios com valores inferiores à média. Por outro lado, metade dos municípios apresentam índices de diversificação superiores a 0.685, sendo que 25% dos municípios portugueses registam valores acima de 0.713 (Anexo 4).

Relativamente à diversificação relacionada, o intervalo [0.29; 1.31] mostra que os valores se concentram num intervalo relativamente estrito. Os índices mais altos tendem a concentrar-se em municípios do Litoral. Contrariamente, as regiões do Interior Norte, apresentam níveis de diversificação relacionada mais baixos. Importa salientar que, 50% dos municípios registam um índice de diversificação relacionada superior a 0.98 (mediana). Apenas 10% dos municípios apresentam um índice superior a 1.15 (Anexo 5).

No que respeita à diversificação não relacionada, observa-se uma distribuição mais dispersa, com valores mais elevados observados tanto em municípios do litoral como em alguns municípios do interior. Apesar disso, a variabilidade do índice é relativamente estreito. Constata-se que metade dos municípios apresenta um índice de diversificação não relacionada superior a 3.34 (mediana). Apenas 5% têm índices superiores a 3.53 (Anexo 6).

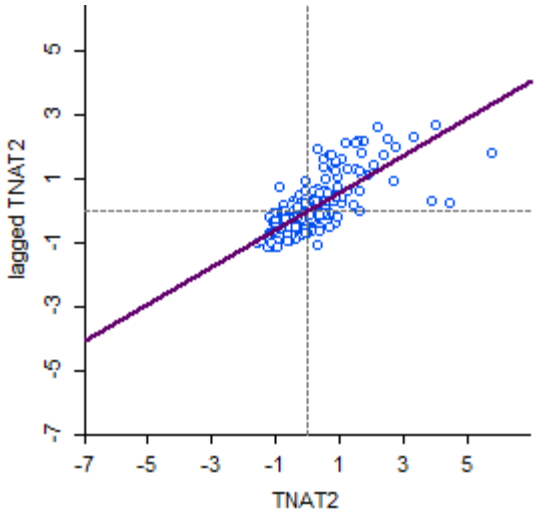
4.2.2. Dependência Espacial: variável dependente

Na análise da dependência espacial foi utilizada uma matriz de pesos espaciais baseada na distância.

O GeoDa permite a construção de diferentes matrizes de pesos espaciais, incluindo matrizes de contiguidade (*rook* e *queen*), matrizes baseadas em distância e matrizes dos k-vizinhos mais próximos. No entanto, a matriz de distâncias é considerada particularmente adequada quando os efeitos espaciais ultrapassam as fronteiras administrativas, pelo que a proximidade física é mais relevante do que a contiguidade formal.

Neste caso, a vizinhança é definida a partir de um raio máximo, determinado automaticamente de modo a garantir que todas as unidades espaciais tenham pelo menos um vizinho e está, por defeito, normalizada por linhas.

Figura 5 - Índice de Moran Univariado



O diagrama de dispersão de *Moran* referente à taxa de natalidade (Figura 5) indica a presença de autocorrelação espacial global positiva. O valor do I de *Moran* (0.580) é estatisticamente significativo (p-value = 0.002), ou seja, que os municípios com taxas de natalidade elevadas tendem a localizar-se próximos de municípios que têm também taxas elevadas, enquanto municípios com taxas mais baixas tendem a localizar-se na vizinhança de outros municípios com valores baixos. Seguindo o procedimento comum na literatura (Anselin et al., 1996), a significância estatística do I *Moran* foi avaliada com base em testes de permutação (499 permutações aleatórias), sob a hipótese nula de ausência de autocorrelação espacial.

É ainda possível perceber que, a maioria dos municípios se concentra no primeiro e terceiro quadrante. Os municípios que se localizam no segundo e quarto quadrante (como os municípios de Almodôvar e Vila Real), correspondem a situações de autocorrelação negativa, frequentemente interpretados como potenciais *outliers* espaciais.

Perante a presença de interdependência espacial entre os municípios, justifica-se o recurso a modelos de econometria espacial.

4.2.3. Dependência Espacial: variáveis independentes

Tendo identificado anteriormente a presença de dependência espacial na taxa de natalidade das empresas, procede-se, nesta secção, ao diagnóstico de dependência espacial das variáveis explicativas (Tabela 8).

Tabela 8 - Índice de Moran Univariado

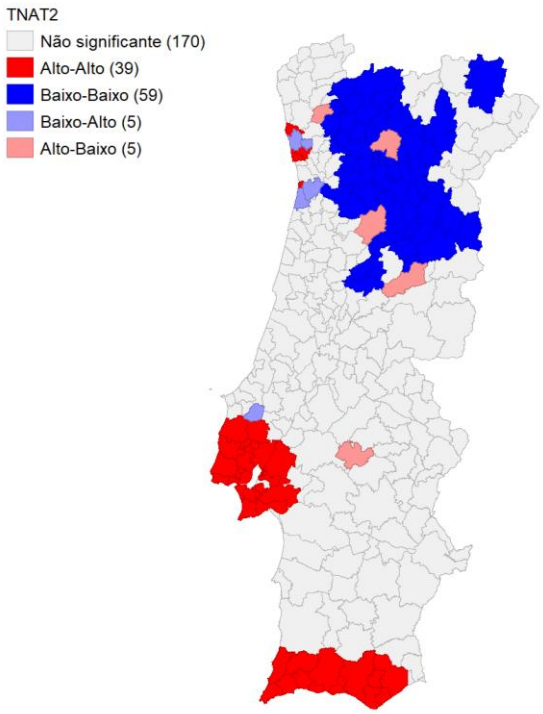
Variáveis	Índice de *Moran* Univariado
TNAT	0.580 (0.002)
URBAN	0.743 (0.002)
SPECIAL	0.622 (0.002)
DIVERSIF	0.696 (0.002)
DIVERS_R	0.717 (0.002)
DIVERS_NR	0.640 (0.002)
DESEMP	0.443 (0.002)
DENSEMP	0.324 (0.002)

Nota: p-value = 0.002, calculado com base em 499 permutações

Na Tabela 8 observa-se a presença de autocorrelação espacial global em todas as variáveis explicativas, estatisticamente significativa a 1%, demonstrando a existência de interdependência espacial.

Tendo sido confirmada a presença de autocorrelação espacial global através do Índice I de *Moran*, procede-se agora à análise local, recorrendo aos mapas LISA, com o objetivo de identificar a localização dos *clusters* estatisticamente significativos.

Figura 6 - Mapa de clusters da taxa de natalidade (LISA)



Nota: Average Spatial Lag. Nível de significância: 5%

Está evidenciado na Figura 6, a presença de *clusters* estatisticamente significativos do tipo Alto-Alto abrangendo 39 municípios, que correspondem a regiões com elevadas taxas de natalidade. Especificamente, estes *clusters* Alto–Alto localizam-se essencialmente na Área Metropolitana de Lisboa e no Algarve. Identificam-se também *clusters* do tipo Baixo–Baixo, abrangendo 59 municípios, associados a regiões com baixas taxas de natalidade. Os *clusters* Baixo–Baixo concentram-se sobretudo no Interior Norte.

Identificam-se ainda associações do tipo Alto–Baixo e Baixo–Alto, cada uma abrangendo cinco municípios, localizadas nas regiões do Litoral Norte e Centro. Nas associações do tipo Alto-Baixo, os municípios de Mora, Covilhã e Braga são estatisticamente significativos a 5%, e Viseu e Vila Real são estatisticamente significativos a 1%. Nas associações do tipo Baixo-Alto, os municípios de Cadaval, Ovar, Santa Maria da Feira, Trofa e Vila do Conde são estatisticamente significativos a 5%.

A identificação destes *clusters* assume particular relevância, uma vez que correspondem a padrões estatisticamente significativos de associação espacial local, não explicáveis por processos aleatórios. Em particular, estas associações do tipo Alto-Baixo e Baixo-Alto representam *outliers* espaciais – são territórios com dinâmicas próprias, que apresentam um comportamento que difere do observado na vizinhança. Isto pode estar relacionado com fatores locais específicos de natureza socioeconómica, demográfica ou institucional.

4.3. Estimação dos Modelos Econométricos

Embora o Índice de *Moran* permita a identificação de autocorrelação espacial, não indica que especificação alternativa ao MQO deve ser utilizada. Deste modo, para ultrapassar esta limitação, é comum na literatura recorrer-se aos testes do *Multiplicador de Lagrange (LM),* para determinar se o modelo mais adequado *é* um modelo de desfasagem espacial (*Spatial Lag Model*) ou um modelo de erro espacial (*Spatial Error Model*) (Caragliu et al., 2016).

Com base na Tabela 9, conclui-se que ambos os testes LM robustos são estatisticamente significativos em todas as especificações consideradas, pelo que não há uma escolha inequívoca entre os modelos SLM e SEM. Neste caso, o modelo recomendado seria o *Spatial Durbin Model* (SDM), por se tratar de uma especificação mais geral, que inclui simultaneamente a variável dependente desfasada espacialmente e as variáveis explicativas desfasadas, permitindo modelar tanto efeitos espaciais endógenos como exógenos.

No entanto, o software utilizado (GeoDa) não permite estimar um SDM. Assim, de acordo com a literatura, procedeu-se à estimação e comparação de ambas as especificações SLM e SEM com base nos critérios de informação (AIC e BIC), no valor do *log-likelihood* e na significância dos coeficientes estimados. Com base nestes critérios, a análise econométrica foi conduzida recorrendo ao *Spatial Error Model*.

Tabela 9 - Resultados dos Testes de Multiplicador de Lagrange

Teste	Modelo 1	Modelo 2	Modelo 3	Modelo 4
LM (Lag)	176.919 (0.00000)	183.197 (0.00000)	189.316 (0.00000)	182.972 (0.00000)
LM Robust	17.984 (0.00002)	17.619 (0.00003)	19.412 (0.00001)	17.072 (0.00004)
LM (Error)	187.032 (0.00000)	190.234 (0.00000)	189.952 (0.00000)	193.171 (0.00000)
LM Robust (Error)	28.097 (0.00000)	24.656 (0.00000)	20.0480 (0.00001)	27.270 (0.00000)

De modo a responder às questões de investigação, foram estimados os seguintes modelos econométricos:

M1:$TNAT = \rho W\ TNAT + \beta_1 + \beta_2 URBAN + \beta_3 SPECIAL + \beta_4 DESEMP + \beta_5 DENSEMP + \varepsilon$

M2:$TNAT = \rho W\ TNAT + \beta_1 + \beta_2 URBAN + \beta_3 DIVERSIF + \beta_4 DESEMP + \beta_5 DENSEMP + \varepsilon$

M3:$TNAT = \rho W\ TNAT + \beta_1 + \beta_2 URBAN + \beta_3 DIVERS_R + \beta_4 DESEMP + \beta_5 DENSEMP + \varepsilon$

M4:$TNAT = \rho W\ TNAT + \beta_1 + \beta_2 URBAN + \beta_3 DIVERS_{NR} + \beta_4 DESEMP + \beta_5 DENSEMP + \varepsilon$

Os resultados da estimação estão apresentados na Tabela 10.

Tabela 10 - Resultados de estimação dos modelos econométricos usando OLS e SEM

	OLS	SEM	OLS	SEM	OLS	SEM	OLS	SEM
	(1)		(2)		(3)		(4)	
Constante	11.812*** (9.9348)	9.9871*** (7.7352)	1.519 (1.0359)	2.719 (1.5737)	4.5471*** (4.1562)	5.0926*** (3.8418)	0.5220 (0.3163)	2.2372 (1.2218)
URBAN	0.1006 (0.5915)	0.3967** (2.0161)	0.2728* (1.7216)	0.5525*** (2.9972)	0.3179** (1.9933)	0.6384*** (3.5286)	0.3044* (1.9497)	0.5433*** (2.9381)
SPECIAL	- 9.5950*** (-5.5278)	- 6.4828*** (-3.6257)	————	————	————	————	————	————
DIVERSIF	————	————	10.633*** (4.9748)	7.4606*** (3.0429)	————	————	————	————
DIVERS_R	————	————	————	————	3.9314*** (4.4009)	2.3253** (2.2187)	————	————
DIVERS_NR	————	————	————	————	————	————	2.4795*** (4.8971)	1.6968*** (3.1099)
DESEMP	- 0.0668 (-1.1324)	- 0.0685 (-1.1995)	- 0.0538 (-0.8983)	- 0.06109 (-1.0586)	- 0.0451 (-0.7412)	- 0.06136 (-1.0522)	- 0.0654 (-1.0942)	- 0.0652 (-1.1325)
DENSEMP	0.01963*** (10.550)	0.01456*** (9.7416)	0.0186*** (10.094)	0.0137*** (9.393)	0.0184*** (9.8858)	0.0133*** (9.1327)	0.0185*** (10.037)	0.01377*** (9.4281)
Lambda (SEM)		0.7338***		0.7326***		0.7320***		0.7348***
N	278	278	278	278	278	278	278	278
R^2	0.5163	0.7267	0.5069	0.7228	0.4978	0.7185	0.5056	0.7235
Log Likelihood	- 671.162	- 607.0656	- 673.847	- 608.939	- 676.382	- 611.0567	- 674.206	-608,73
Teste Breusch-Pagan	60.778*** (0.00)	92.849*** (0.00)	61.562*** (0.00)	91.594*** (0.00)	67.376*** (0.00)	91.843***	57.011*** (0.00)	90.504*** (0.00)
Likelihood Ratio Test		128.193***		129.817***		130.651***		130.952***

Nota: A tabela apresenta os resultados dos modelos OLS e SEM usando o software GeoDa. O valor entre parênteses no modelo OLS correspondem os valores do *t-statistc, enquanto no modelo SEM correspondem ao z – value.. Níveis de significância: ***p<0.01, **p<0.05, *p<0.1.*

5. Análise e Discussão dos Resultados

Na Tabela 10, é possível verificar que o lambda é positivo e estatisticamente significativo em todos os modelos, ao nível mais elevado de significância. Este resultado indica a presença de autocorrelação espacial no termo de erro, confirmando a existência de autocorrelação espacial positiva associada a fatores não observados que afetam os municípios vizinhos. Este resultado indica que os modelos clássicos OLS não são adequados para explicar as diferenças regionais na criação de empresas, justificando o recurso ao modelo *SEM* para estimar os quatro modelos econométricos.

Apesar disso, a literatura recomenda a apresentação e discussão dos resultados obtidos por OLS, com o objetivo de ilustrar o que acontece quando a dependência espacial é ignorada. Assim, os resultados obtidos por OLS são apresentados a título ilustrativo e comparativo, para avaliar de que forma a incorporação da dependência espacial altera a magnitude e a significância das associações estimadas.

(i) Modelos de regressão linear estimados por OLS

No que concerne às variáveis independentes, os resultados obtidos com os modelos OLS indicam que a urbanização apresenta coeficientes positivos, sendo estatisticamente significativa ao nível 5% apenas no Modelo 3. Nos Modelos 2 e 4 os coeficientes estimados são apenas marginalmente significativos, e no Modelo 1 esta variável não é estatisticamente significativa aos níveis convencionais.

Em contraste, a especialização apresenta coeficientes negativos e estatisticamente significativos ao nível de 1%, indicando uma associação negativa entre a especialização produtiva e a criação de empresas.

A variável diversificação, apresenta coeficientes positivos e estatisticamente significativos, sugerindo uma associação positiva com a criação de empresas. Resultados semelhantes são observados para as variáveis que medem a diversificação relacionada e não relacionada, que apresentam coeficientes positivos e estatisticamente significativos em todas as especificações estimadas.

Relativamente às variáveis de controlo, o desemprego evidencia coeficientes negativos, mas não significativos, em todos os modelos. Em contraste, a densidade empresarial apresenta uma associação positiva e estatisticamente significativa à criação de empresas.

Em suma, os resultados dos modelos OLS sugerem que a diversificação – geral, relacionada e não relacionada bem como a densidade empresarial, se encontram positivamente associadas à criação de empresas nos municípios portugueses. Pelo contrário, a especialização apresenta uma associação

negativa. A urbanização também evidencia uma associação positiva, ainda que não seja estatisticamente significativo em todos os modelos.

(ii) *Spatial Error Model*

Tendo em consideração os resultados dos modelos Spatial Error Model, a urbanização passa a apresentar significância estatística no Modelo 1 (ao nível de 5%), e nos restantes três modelos torna-se significativo aos níveis convencionais de significância. Efetivamente, os resultados dos modelos SEM evidenciam maior estabilidade dos coeficientes associados à urbanização (com coeficientes positivos e estatisticamente significativos). Estes resultados sugerem uma associação positiva entre a urbanização e a criação de empresas, que se torna consistente quando se controla a dependência espacial do termo de erro. Esta evidência está de acordo com a apresentada por Renski (2014), Andersson et al. (2019) e O'Leary et al. (2023) de que áreas mais urbanizadas tendem a beneficiar da proximidade a clientes e fornecedores, maior disponibilidade de capital humano com elevadas qualificações e melhor acesso a informação, fatores que são frequentemente associados a taxas de formação de empresas mais elevadas. Sendo assim, a hipótese (H1) não é rejeitada:

- H1: Existe uma associação positiva entre a urbanização e a taxa de criação de empresas – não rejeitada.

A variável especialização mantém - se estatisticamente significativa, mas com sinal negativo, sugerindo uma associação negativa entre a especialização e a criação de empresas. Este resultado não está de acordo com os reportados em Anyadike-Danes & Hart (2006) e Fotopoulos (2014). A associação negativa observada poderá estar relacionada, por exemplo, com o aumento da concorrência devido à concentração de empresas com atividade semelhante. Perante estes resultados, a hipótese (H2) é rejeitada com base na evidência empírica:

- H2: A especialização está positivamente associada à criação de empresas, embora com efeitos heterogéneos entre setores - rejeitada.

A diversificação em geral, mantém-se positivamente associada à criação de empresas, apresentando coeficientes positivos e estatisticamente significativos nos modelos estimados. Estes resultados são consistentes com os apresentados por Barboza & Pede (2024), que associam contextos produtivos mais diversificados a um ambiente favorável à partilha de conhecimento e ideias entre diferentes setores de atividade, estimulando a criação de empresas. Desta forma, a hipótese (H3) não é rejeitada:

- H3: A diversificação está associada à criação de empresas – não rejeitada.

Por outro lado, a diversificação relacionada continua a estar positivamente associada à criação de empresas, embora com um nível de significância inferior (estatisticamente significativa a 5%). Estes resultados sugerem que estruturas produtivas caracterizadas por maior diversidade entre setores tecnologicamente próximos se encontram associados a taxas mais elevadas de formação de empresas. Esta evidência vai ao encontro de Guo et al. (2016) e Content et al. (2019). Em suma, os resultados são consistentes com a hipótese (H4), não permitindo a sua rejeição:

- H4: A diversificação relacionada está positivamente associada à criação de empresas – não rejeitada.

A diversificação não relacionada mantém-se positivamente associada à criação de empresas, apresentando coeficientes positivos e estatisticamente significativos ao nível de 1%. Este resultado sugere que contextos caracterizados por maior diversificação não relacionada podem estar associados a maiores taxas de natalidade de empresas. Esta evidência está de acordo com a reportada em O'Leary (2022) e O'Leary et al. (2023), que associam a diversificação não relacionada a diferentes tipos de conhecimento e ideias, que promovem a criação de empresas.

Posto isto, a hipótese (H5 não é rejeitada):

- H5: A diversificação não relacionada está associada à criação de empresas – não rejeitada.

Relativamente às variáveis de controlo, a variável desemprego continua a apresentar com coeficientes negativos, mas não estatisticamente significativos, assim como em Capozza et al. (2018) e Piacentino et al. (2017). A densidade empresarial mantém coeficientes positivos e estatisticamente significativos em todos os modelos, sugerindo uma associação positiva com a criação de empresas, tal como em Colombelli & Quatraro (2018) e Barboza & Pede (2024).

Por último, o teste *Likelihood Ratio* e o teste *Breusch-Pagan* apresentam em todos os modelos, coeficientes de sinal positivo e estatisticamente significativos, revelando que ainda existe heteroscedasticidade, mesmo introduzindo um modelo econométrico espacial.

Em suma, os resultados sugerem que as externalidades de aglomeração se encontram associadas à criação de empresas em Portugal Continental, com particular destaque para o papel da diversificação. Foi também encontrada evidência de que as taxas de natalidade em cada município são influenciadas por fatores não observados nos municípios vizinhos, demonstrando ser necessário considerar explicitamente a autocorrelação espacial neste contexto.

Após a estimação dos modelos por OLS, foram calculados os fatores de inflação da variância

(usando Stata) como diagnóstico de multicolinearidade (Anexo 1). A análise dos valores obtidos indicam que não se detetam problemas de multicolinearidade.

Para efeitos de teste de robustez, utilizou-se também uma matriz de contiguidade do tipo rainha, verificando-se que não há alterações significativas quer nos sinais dos coeficientes quer na significância estatística, com exceção da variável diversificação relacionada, que se revela apenas marginalmente significativa (Anexo 9).

6. Conclusão

A criação de novas empresas assume um papel fundamental no desenvolvimento regional, pois está associada a diversos efeitos diretos e indiretos, como a criação de emprego e a melhoria da competitividade impulsionada pelo aumento da concorrência. No entanto, a literatura tem mostrado que a dinâmica empreendedora não é homogênea no território, existindo fatores que contribuem para uma maior ou menor atratividade das regiões e tendem a acentuar as desigualdades existentes.

Este estudo analisou a relação entre as externalidades de aglomeração e as taxas de natalidade das empresas nos setores da indústria transformadora e dos serviços, no período de 2018 – 2023, nos 278 municípios de Portugal Continental.

A taxa de natalidade foi calculada pelo rácio entre o número de empresas criadas e a população em idade ativa, que enquadra a criação de empresas como um fenómeno ligado à base populacional potencialmente ativa.

Através da matriz de correlação, verificou-se que poderiam surgir problemas de multicolinearidade entre algumas variáveis, como a especialização, diversificação, diversificação relacionada e diversificação não relacionada, caso fossem utilizadas no mesmo modelo. Por isso, foram estimados quatro modelos para permitir avaliar de forma mais rigorosa a associação de cada uma dessas variáveis com a formação de novas empresas.

Utilizou-se econometria espacial, visto que permite ter em conta a dependência espacial existente entre os municípios, e consequentemente, uma maior robustez nos resultados. O modelo *Spatial Error Model*, mostrou ser o mais adequado, tendo se observado que taxas de natalidade mais elevadas, estão associadas a regiões com um maior nível de urbanização, diversificação em geral, diversificação relacionada e diversificação não relacionada. Por outro lado, regiões onde existe um maior nível de especialização estão associadas a taxas de natalidade mais baixas.

Assim, conclui-se que as externalidades de aglomeração se encontram estreitamente associadas às dinâmicas regionais da criação de empresas, destacando-se o papel da diversificação. Adicionalmente, foi encontrada evidência da relevância de incorporar o espaço na análise, uma vez que as taxas de natalidade observadas num município se relacionam com fatores não observados presentes nos municípios vizinhos. Estes resultados reforçam a importância de considerar explicitamente a autocorrelação espacial na análise das diferenças regionais na criação de empresas.

7. Limitações e Investigação Futura

Este estudo apresenta algumas limitações que devem ser consideradas. Uma das principais limitações relaciona-se com os problemas de endogeneidade entre a variável dependente e as variáveis explicativas. Ou seja, variáveis como a especialização e a diversificação explicam a criação de empresas, mas também são influenciadas pela criação de empresas. Apesar da utilização de variáveis explicativas diferidas contribuir para atenuar a endogeneidade, não a elimina totalmente. Por esta razão, não é possível estabelecer uma relação de causalidade.

Outra das limitações deste trabalho prende-se com o facto de ter sido utilizada como proxy os "Estabelecimentos por Localização Geográfica" para o cálculo da especialização e para os diferentes tipos de diversificação, em vez de "Pessoal ao Serviço", como estava previsto inicialmente. Esta opção deveu-se à existência de "missings" e valores nulos a nível dos municípios, que não permitia calcular quer para CAE 1 dígito quer para CAE 2 dígitos. Contudo, a utilização do número de estabelecimentos, traz implicações nomeadamente a atribuição do mesmo peso a unidades produtivas que podem ter dimensões muito distintas, independentemente do número de trabalhadores. Esta limitação é particularmente relevante em economias locais onde predominam micro empresas (comércio e serviços). Pelo contrário, pode estar a sub-representar setores intensivos em capital – poucas grandes empresas. Assim, a análise reflete sobretudo a estrutura empresarial, em vez da estrutura do emprego.

Paralelamente, os resultados dos testes de dependência espacial foram significativos quer para o *Spatial Lag Model* quer para o *Spatial Error Model*. Deste modo, trabalhos futuros poderiam estimar um modelo que captasse simultaneamente os dois tipos de dependência espacial, como o *Spatial Durbin Model*. Além disso, seria pertinente o uso de metodologias que permitam lidar com os problemas de endogeneidade referidos anteriormente.

Por fim, não só a utilização de dados em painel poderia captar melhor a evolução temporal da criação de empresas, bem como a análise dos setores da industria transformadora e dos serviços separadamente poderia melhorar a precisão dos resultados.

Referências

Aarstad, J., Kvitastein, O. A., & Jakobsen, S. E. (2016). Related and unrelated variety as regional drivers of enterprise productivity and innovation: A multilevel study. *Research Policy*, *45*(4), 844–856. https://doi.org/10.1016/j.respol.2016.01.013

Acs, Z. J., Armington, C., & Zhang, T. (2006). *The Determinants of New-firm Survival across Regional Economies*. https://doi.org/https://doi.org/10.1111/j.1435-5957.2007.00129.x

Aharonson, B. S., Baum, J. A. C., & Feldman, M. P. (2006). *Desperately Seeking Spillovers? Increasing Returns, Industrial Organization and the Location of New Entrants in Geographic and Technological Space*. https://doi.org/http://dx.doi.org/10.2139/ssrn.899598

Andersson, M., Larsson, J. P., & Wernberg, J. (2019). The economic microgeography of diversity and specialization externalities – firm-level evidence from Swedish cities. *Research Policy*, *48*(6), 1385–1398. https://doi.org/10.1016/j.respol.2019.02.003

Anselin, L., Bera B', A. K., Florax, R., & Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. In *Regional Science and Urban Economics* (Vol. 26). https://doi.org/https://doi.org/10.1016/0166-0462(95)02111-6

Anyadike-Danes, M., & Hart, M. (2006). The impact of sector, specialisation, and space on business birth rates in the United Kingdom: A challenge for policy? *Environment and Planning C: Government and Policy*, *24*(6), 815–826. https://doi.org/10.1068/C0549

Arauzo-Carod, J. M., Liviano-Solis, D., & Manjón-Antolín, M. (2010). Empirical studies in industrial location: An assessment of their methods and results. *Journal of Regional Science*, *50*(3), 685–711. https://doi.org/10.1111/j.1467-9787.2009.00625.x

Armington, C., & Acs, Z. J. (2002). The determinants of regional variation in new firm formation. *Regional Studies*, *36*(1), 33–45. https://doi.org/10.1080/00343400120099843

Arrow, K. J. (1962). *The Economic Implications of Learning by Doing* (Vol. 29, Number 3). https://about.jstor.org/terms

Audretsch, D. B., & Belitski, M. (2020). Knowledge complexity and firm performance: evidence from the European SMEs. *Journal of Knowledge Management*, *25*(4), 693–713. https://doi.org/10.1108/JKM-03-2020-0178

Audretsch, D. B., Belitski, M., & Caiazza, R. (2021). Start-ups, Innovation and Knowledge Spillovers. *Journal of Technology Transfer*, *46*(6), 1995–2016. https://doi.org/10.1007/s10961-021-09846-5

Audretsch, D. B., & Fritsch, M. (1994). The Geography of Firm Births in Germany. *Regional Studies*, *28*(4), 359–365. https://doi.org/10.1080/00343409412331348326

Bade, F.-J., & Nerlinger, E. (2000). The spatial distribution of new technology- based firms: Empirical results for West - Germany. *Papers in Regional Science*. https://doi.org/https://doi.org/10.1111/j.1435-5597.2000.tb00766.x

Badr, K., Rizk, R., & Zaki, C. (2019). Firm productivity and agglomeration economies: evidence from Egyptian data. *Applied Economics*, *51*(51), 5528–5544. https://doi.org/10.1080/00036846.2019.1613506

Baltzopoulos, A., Braunerhjelm, P., & Tikoudis, I. (2016). *Spin-offs: Why geography matters*. https://doi.org/https://doi.org/10.1093/jeg/lbv006

Barboza, G., & Pede, V. (2024). Weak vs Strong Knowledge Spillover Effects. Evidence from Geographic Distribution of Innovative Startup in Italy. *The Review of Regional Studies*, *54*, 163–191. https://doi.org/10.52324/001c.123323

Basile, R., Pittiglio, R., & Reganati, F. (2017). Do agglomeration externalities affect firm survival? *Regional Studies*, *51*(4), 548–562. https://doi.org/10.1080/00343404.2015.1114175

Bishop, P. (2012). Knowledge, diversity and entrepreneurship: A spatial analysis of new firm formation in Great Britain. *Entrepreneurship and Regional Development*, *24*(7–8), 641–660.

https://doi.org/10.1080/08985626.2011.617786
Bishop, P. (2019). Knowledge diversity and entrepreneurship following an economic crisis: an empirical study of regional resilience in Great Britain. *Entrepreneurship and Regional Development*, *31*(5–6), 496–515. https://doi.org/10.1080/08985626.2018.1541595
Bishop, P., & Brand, S. (2014). Human capital, diversity, and new firm formation. *Service Industries Journal*, *34*(7), 567–583. https://doi.org/10.1080/02642069.2014.886193
Bishop, P., & Shilcof, D. (2017). The spatial dynamics of new firm births during an economic crisis: the case of Great Britain, 2004–2012. *Entrepreneurship and Regional Development*, *29*(3–4), 215–237. https://doi.org/10.1080/08985626.2016.1257073
Boschma, R. (2005). Proximity and innovation: A critical assessment. *Regional Studies*, *39*(1), 61–74. https://doi.org/10.1080/0034340052000320887
Boschma, R., & Iammarino, S. (2009). *Related Variety,Trade Linkages, and Regional Growth in Italy*. https://doi.org/10.1111/j.1944-8287.2009.01034.x
Bosma, N., Van Stel, A., & Suddle, K. (2008). The geography of new firm formation: Evidence from independent start-ups and new subsidiaries in the Netherlands. *International Entrepreneurship and Management Journal*, *4*(2), 129–146. https://doi.org/10.1007/s11365-007-0058-8
Brixy, U., Grotz, R., Brixyy, U., & Grotzz, R. (2007). Regional patterns and determinants of birth and survival of new firms in Western Germany. *Entrepreneurship and Regional Development*, *19*(4), 293–312. https://doi.org/10.1080/08985620701275510
Cainelli, G., Montresor, S., & Vittucci Marzetti, G. (2014). Spatial agglomeration and firm exit: A spatial dynamic analysis for Italian provinces. *Small Business Economics*, *43*(1), 213–228. https://doi.org/10.1007/s11187-013-9532-6
Calá, C. D., Manjón-Antolín, M., & Arauzo-Carod, J. M. (2016). Regional determinants of firm entry in a developing country. *Papers in Regional Science*, *95*(2), 259–280. https://doi.org/10.1111/PIRS.12128
Callejón, M. (2003). En busca de las economías externas. *Ekonomiaz. Revista Vasca de Economía*, *53*(2), 74–89. https://doi.org/10.69810/ekz.0670
Campi, M. T. C., Blasco, A. S., & Marsal, E. V. (2004). The Location of New Firms and the Life Cycle of Industries. In *Economics* (Vol. 22). Kluwer Academic Publishers. https://doi.org/10.1023/B:SBEJ.0000022215.02146.40
Capozza, C., Salomone, S., & Somma, E. (2018). Local industrial structure, agglomeration economies and the creation of innovative start-ups: evidence from the Italian case. *Entrepreneurship & Regional Development*, *30*, 749–775. https://doi.org/10.1080/08985626.2018.1457087
Caragliu, A., De Dominicis, L., & De Groot, H. L. F. (2016). Both Marshall and Jacobs were right! *Economic Geography*, *92*(1), 87–111. https://doi.org/10.1080/00130095.2015.1094371
Cerdán, F., & Hernández, F. (2013). *Distribución Espacial De La Actividad Económica En La Comarca De Cartagena. Un Análisis Basado En La Demografía Empresarial Usando Arcview Y Geoda*. *4*, 97–118. https://dialnet.unirioja.es/servlet/articulo?codigo=7897429
Ciccone, A., & Hall, R. E. (1996). Productivity and the Density of Economic Activity. *The American Economic Review*, *86*(1), 54–70.
Cohen, W. M., & Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. In *Quarterly* (Vol. 35, Number 1).
Colombelli, A. (2016). The impact of local knowledge bases on the creation of innovative start-ups in Italy. *Small Business Economics*, *47*(2), 383–396. https://doi.org/10.1007/s11187-016-9722-0
Colombelli, A., & Quatraro, F. (2018). New firm formation and regional knowledge production modes: Italian evidence. *Research Policy*, *47*(1), 139–157. https://doi.org/10.1016/j.respol.2017.10.006
Combes, P. P., Duranton, G., Gobillon, L., Puga, D., & Roux, S. (2012). The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection. *Econometrica*, *80*(6), 2543–2594.

https://doi.org/10.3982/ecta8442
Content, J., Frenken, K., & Jordaan, J. A. (2019). Does related variety foster regional entrepreneurship? Evidence from European regions. *Regional Studies*, *53*(11), 1531–1543. https://doi.org/10.1080/00343404.2019.1595565
Corradini, C., & Vanino, E. (2022). Path dependency, regional variety and the dynamics of new firm creation in rooted and pioneering industries. *Journal of Economic Geography*, *22*(3), 631–651. https://doi.org/10.1093/jeg/lbab021
Criscuolo, C., Gal, P. N., & Menon, C. (2014). *The Dynamics of Employment Growth: new evidence from 18 countries* (OECD Science, Technology and Industry Policy Papers, Vol. 14). https://doi.org/10.1787/5jz417hj6hg6-en
Crowley, F., Doran, J., & Mccann, P. (2021). The vulnerability of European regional labour markets to job automation: the role of agglomeration externalities. *Regional Studies*. https://doi.org/10.1080/00343404.2021.1928041
Davidsson, P., Lindmark, L., & Olofsson, C. (1994). New Firm Formation and Regional Development in Sweden. *Regional Studies*, *28*(4), 395–410. https://doi.org/10.1080/00343409412331348356
Decker, R., Haltiwanger, J., Jarmin, R., & Miranda, J. (2014). The role of entrepreneurship in us job creation and economic dynamism. In *Journal of Economic Perspectives* (Vol. 28, Number 3, pp. 3–24). American Economic Association. https://doi.org/10.1257/jep.28.3.3
Druck, S., Carvalho, M., Câmara, G., & Monteiro, A. (2002). *Análise Espacial de Dados Geográficos*. http://www.dpi.inpe.br/gilberto/livro/analise/
Ebert, T., Brenner, T., & Brixy, U. (2019). New firm survival: the interdependence between regional externalities and innovativeness. *Small Business Economics*, *53*(1), 287–309. https://doi.org/10.1007/s11187-018-0026-4
Ejdemo, T., & Örtqvist, D. (2020). Related variety as a driver of regional innovation and entrepreneurship: A moderated and mediated model with non-linear effects. *Research Policy*, *49*(7). https://doi.org/10.1016/j.respol.2020.104073
Ellison, G., Glaeser, E. L., & Kerr, W. R. (2010). What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review*, *100*(3), 1195–1213. https://doi.org/10.1257/aer.100.3.1195
Fertala, N. (2008). The shadow of death: Do regional differences matter for firm survival across native and immigrant entrepreneurs? *Empirica*, *35*(1), 59–80. https://doi.org/10.1007/s10663-007-9051-2
Fotopoulos, G. (2014). On the spatial stickiness of UK new firm formation rates. *Journal of Economic Geography*, *14*(3), 651–679. https://doi.org/10.1093/jeg/lbt011
Frenken, K., Van Oort, F., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies*, *41*(5), 685–697. https://doi.org/10.1080/00343400601120296
Fritsch, M., & Kublina, S. (2019). Persistence and change of regional new business formation in the national league table. *Journal of Evolutionary Economics*, 891–917. https://doi.org/10.1007/s00191-019-00610-5
Fritsch, M., & Mueller, P. (2004). Effects of new business formation on regional development over time. *Regional Studies*, *38*(8), 961–975. https://doi.org/10.1080/0034340042000280965
Gajewski, P., & Kutan, A. M. (2018). Determinants and Economic Effects of New Firm Creation: Evidence from Polish Regions. *Eastern European Economics*, *56*(3), 201–222. https://doi.org/10.1080/00128775.2018.1442226
Ganau, R. (2015). Productivity, Credit Constraints and the Role of Short-Run Localization Economies: Micro-Evidence from Italy. *Regional Studies*. https://doi.org/10.1080/00343404.2015.1064883
Garofoli, G. (1994). New Firm Formation and Regional Development: The Italian Case. *Regional Studies*, *28*(4), 381–393. https://doi.org/10.1080/00343409412331348346

Guo, Q., He, C., & Li, D. (2016). Entrepreneurship in China: The role of localisation and urbanisation economies. *Urban Studies*, *53*(12), 2584–2606. https://doi.org/10.1177/0042098015595598

Jacobs, J. (1969). *The Economy of Cities*. New York: Vintage.

Jacobs, W., Koster, H. R. A., & Van Oort, F. (2014). Co-agglomeration of knowledge-intensive business services and multinational enterprises. *Journal of Economic Geography*, *14*(2), 443–475. https://doi.org/10.1093/jeg/lbs055

Jagódka, M., & Snarska, M. (2023). Regional Disparities as a Result of Differences in Human Capital and Innovativeness on the example of Poland. *Technological and Economic Development of Economy*, *29*(2), 696–716. https://doi.org/10.3846/tede.2023.18536

Joffe, M. (2024). Profit rate dynamics in US manufacturing. *International Review of Applied Economics*, *38*(1–2), 194–223. https://doi.org/10.1080/02692171.2022.2154917

Johan, S., & Vania, I. (2022). *The Application of Herfindahl-Hirschman Index in Measuring the Concentration Level of Financial-Technology Industry*. https://doi.org/10.2991/aebmr.k.220501.002

Jung, A., Camacho, M., Jung, A., & Camacho, M. (2012). Industry and regional factors associated to new firm formation in Uruguay. *Cuaderno de Economía*, 29–44. https://www.researchgate.net/publication/271443652

Kanellopoulos, V., & Fotopoulos, G. (2018). The effect of knowledge spillovers on regional new firm formation: The Greek manufacturing case. *Environment and Planning A*, *51*(4), 1005–1030. https://doi.org/10.1177/0308518X18820078

Kangasharju, A. (2000). Regional variations in firm formation: Panel and cross-section data evidence from Finland. In *Papers Reg. Sci* (Vol. 79). https://doi.org/10.1007/PL00011482

Kato-Vidal, E. L. (2022). Industrial Variety, Firms and Resilience. *Studies in Business and Economics*, *17*(2), 141–159. https://doi.org/10.2478/sbe-2022-0030

Kim, J., Kollmann, T., Palangkaraya, A., & Webster, E. (2022). Does local technological specialisation, diversity and dynamic competition enhance firm creation? *Research Policy*, *51*(7). https://doi.org/10.1016/j.respol.2022.104557

Lasch, F., Robert, F., & Le Roy, F. (2013). Regional determinants of ICT new firm formation. *Small Business Economics*, *40*(3), 671–686. https://doi.org/10.1007/s11187-011-9382-z

Lawton Smith, H. (1991). The role of incubators in local industrial development: the crygenics industry in Oxfordshire. *Entrepreneurship & Regional Development*, *3*(2), 175–194. https://doi.org/10.1080/08985629100000012

LeSage, J., & Kelley Pace, R. (2009). *Introduction to Spatial Econometrics*. https://doi.org/https://doi.org/10.1201/9781420064254

Levratto, N. (2015). Imitation or opposition? A focus on neighbouring effects at the local level in the business creation process. *Journal of Innovation Economics & Management*, *No 17*(2), 165–192. https://doi.org/10.3917/e.jie.017.0165

Li, H., Cheng, S., & Haynes, K. E. (2011). The employment effects of new business formation: A regional perspective. *Economic Development Quarterly*, *25*(3), 282–292. https://doi.org/10.1177/0891242411407310

Lopes, S. B., Cristina, N., & Brondino, M. (n.d.). *Análise do Desempenho de Modelos de Regressão Espacial na Previsão de Demanda por Transportes*. Retrieved May 13, 2025, from http://www.redpgv.coppe.ufrj.br/index.php/es/produccion/articulos-cientificos/2006-1/254-lopes-brondino-silva-panam2006/file

Manski, C. F. (1993). Manski 1993. *The Review of Economic Studies*, *60*, 531–542. https://doi.org/https://doi.org/10.2307/2298123

Marshall, A. (1920). Principles of Economics. *London: Macmillan*. http://oll.libertyfund.org/title/1676

Matsuyama, K. (1995). Complementarities and Cumulative Processes in Models of Monopolistic

Competition. In *Source: Journal of Economic Literature* (Vol. 33, Number 2).

McCann, B. T., & Folta, T. B. (2008). Location matters: Where we have been and where we might go in agglomeration research. In *Journal of Management* (Vol. 34, Number 3, pp. 532–565). https://doi.org/10.1177/0149206308316057

Monseny, J., Marín-López, R., & Viladecans-Marsal, E. (2011). The mechanisms of agglomeration: Evidence from the effect of inter-industry relations on the location of new firms. *Journal of Urban Economics*, *70*(2–3), 61–74. https://doi.org/10.1016/j.jue.2011.05.002

Motoyama, Y., & Malizia, E. (2017). Demand pull or supply push? Metro-level analysis of start-ups in the United States. *Regional Studies, Regional Science*, *4*(1), 232–246. https://doi.org/10.1080/21681376.2017.1379885

Müller, S. (2016). A progress review of entrepreneurship and regional development: What are the remaining gaps? In *European Planning Studies* (Vol. 24, Number 6, pp. 1133–1158). Routledge. https://doi.org/10.1080/09654313.2016.1154506

Neffke, F., Henning, M., Boschma, R., Lundquist, K. J., & Olander, L. O. (2011). The dynamics of agglomeration externalities along the life cycle of industries. *Regional Studies*, *45*(1), 49–65. https://doi.org/10.1080/00343401003596307

Noonan, L., O'Leary, E., & Doran, J. (2021). The impact of institutional proximity, cognitive proximity and agglomeration economies on firm-level productivity. *Journal of Economic Studies*, *48*(2), 257–274. https://doi.org/10.1108/JES-07-2019-0345

Okamuro, H., & Nobuo, K. (2005). *Determinants of Regional Variations in the Start-up Ratio: Evidence from Japan*. https://www.researchgate.net/publication/30788156

O'Leary, D. (2022). Unemployment and entrepreneurship across high-, middle- and low-performing European regions. *Regional Studies, Regional Science*, *9*(1), 571–580. https://doi.org/10.1080/21681376.2022.2118072

O'Leary, D., Doran, J., & Power, B. (2023). Urbanisation, concentration and diversification as determinants of firm births and deaths. *Regional Studies, Regional Science*, *10*(1), 506–528. https://doi.org/10.1080/21681376.2023.2204143

Piacentino, D., Giuliani, D., Bono, F., & Cracolici, M. F. (2017). A spatial analysis of new business formation: Replicative vs innovative behaviour. *Spatial Statistics*, *21*, 390–405. https://doi.org/10.1016/j.spasta.2017.02.004

Piętak, Ł. (2022). Regional disparities, transmission channels and country's economic growth. *Journal of Regional Science*, *62*(1), 270–306. https://doi.org/10.1111/jors.12564

Ponciano, P. F., & Scalon, J. D. (2010). Spatial analysis of the dairy yield using a conditional autoregressive model. *Semina:Ciencias Agrarias*, *31*(2), 487–496. https://doi.org/10.5433/1679-0359.2010v31n2p487

Power, B., Doran, J., & Ryan, G. (2019). The effect of agglomeration economies on firm deaths: A comparison of firm and regional based approaches. *Urban Studies*, *56*(16), 3358–3374. https://doi.org/10.1177/0042098018817428

Power, B., Ryan, G., & Doran, J. (2020). A micro-analysis of Irish firm deaths during the financial crisis (2006–2010). *The Irish Journal of Management*, *39*(1), 1–16. https://doi.org/10.2478/ijm-2010-0004

Qualls, D. (1972). Concentration, Barriers to Entry, and Long Run Economic Profit Margins. In *Source: The Journal of Industrial Economics* (Vol. 20, Number 2).

Renski, H. (2011). External economies of localization, urbanization and industrial diversity and new firm survival. *Papers in Regional Science*, *90*(3), 473–502. https://doi.org/10.1111/j.1435-5957.2010.00325.x

Renski, H. (2014). The influence of industry mix on regional new firm formation in the United States. *Regional Studies*, *48*(8), 1353–1370. https://doi.org/10.1080/00343404.2012.722202

Romer, P. M. (1986). Increasing Returns and Long-Run Growth. *The Journal of Political Economy*, *94*(5), 1002–1037. https://www.jstor.org/stable/1833190

Spilling, O. R. (1996). Regional variation of new firm formation: The norwegian case. *Entrepreneurship and Regional Development*, *8*(3), 217–244. https://doi.org/10.1080/08985629600000014

Storey, D. J. (1991). *The Birth of New Firms-Does Unemployment Matter? A Review of the Evidence*. https://doi.org/doi.org/10.1007/BF00400022

Sutaria, V., & Hicks, D. A. (2004). New firm formation: Dynamics and determinants. *Annals of Regional Science*, *38*(2), 241–262. https://doi.org/10.1007/s00168-004-0194-9

Tavassoli, S., & Jienwatcharamongkhol, V. (2016). Survival of entrepreneurial firms: the role of agglomeration externalities. *Entrepreneurship and Regional Development*, *28*(9–10), 746–767. https://doi.org/10.1080/08985626.2016.1247916

Tervo, H., & Niittykangas, H. (1994). The Impact of Unemployment on New Firm Formation in Finland. *International Small Business Journal*, *13*(1), 38–53. https://doi.org/10.1177/0266242694131003

Thrane, S., Blaabjerg, S., & Møller, R. H. (2010). Innovative path dependence: Making sense of product and service innovation in path dependent innovation processes. *Research Policy*, *39*(7), 932–944. https://doi.org/10.1016/j.respol.2010.04.003

Traistaru, I., Nijkamp, P., & Longhi, S. (2002). *Regional specialization and concentration of industrial activity in accession countries*. https://hdl.handle.net/10419/39633

Turečková, K., & Nevima, J. (2020). The cost benefit analysis for the concept of a smart city: How to measure the efficiency of smart solutions? *Sustainability (Switzerland)*, *12*(7). https://doi.org/10.3390/su12072663

Van, A., Kashifa, S., & Zoetermeer, S. (2006). *The Impact of New Firm Formation on Regional Development in the Netherlands*. https://doi.org/10.1007/s11187-007-9054-1

Van Der Panne, G. (2004). Agglomeration externalities: Marshall versus Jacobs. *Journal of Evolutionary Economics*. https://doi.org/10.1007/s00191-004-0232-x

Van Oort, F. G., & Stam, E. (2005). *Agglomeration economies and entrepreneurship: Testing for spatial externalities in the Dutch ICT industry*. https://www.researchgate.net/publication/5012198_Agglomeration_economies_and_entrepreneurship_testing_for_spatial_externalities_in_the_Dutch_ICT_industry

Van Soest, D. P., Gerking, S., & van Oort, F. G. (2006). Spatial impacts of agglomeration externalities. *Journal of Regional Science*, *46*(5), 881–899. https://doi.org/10.1111/j.1467-9787.2006.00488.x

Van Stel, A., & Suddle, K. (2008). The impact of new firm formation on regional development in the Netherlands. *Small Business Economics*, *30*(1), 31–47. https://doi.org/10.1007/s11187-007-9054-1

Wagner, J., & Sternberg, R. (2004). Start-up activities, individual characteristics, and the regional milieu: Lessons for entrepreneurship support policies from German micro data. *Annals of Regional Science*, *38*(2), 219–240. https://doi.org/10.1007/s00168-004-0193-x

Westlund, H., Larsson, J. P., & Olsson, A. R. (2014). Start-ups and Local Entrepreneurial Social Capital in the Municipalities of Sweden. *Regional Studies*, *48*(6), 974–994. https://doi.org/10.1080/00343404.2013.865836

Zhao, X., & Yu, G. (2021). Data-Driven Spatial Econometric Analysis Model for Regional Tourism Development. *Mathematical Problems in Engineering*, *2021*. https://doi.org/10.1155/2021/6631833

Anexos

Anexo 1- Testes VIFs

	Modelo 1	Modelo 2	Modelo 3	Modelo 4
	VIF	**VIF**	**VIF**	**VIF**
URBAN	2.48	2.12	2.10	2.05
SPECIAL	1.58			
DIVERSIF		1.35		
DIVERS_R			1.35	
DIVERS_NR				1.30
DESEMP	1.04	1.05	1.07	1.04
DENSEMP	1.71	1.65	1.64	1.64
VIF Médio	1.70	1.54	1.54	1.51

Anexo 2 - Taxa de natalidade por percentis

```
. summarize tnat2, detail

                Taxa de Natalidade Média 2018-2023
-------------------------------------------------------------
      Percentiles      Smallest
 1%     4.406953       3.407893
 5%     5.237791       3.933277
10%     5.756631       4.406953       Obs                 278
25%      7.26792        4.42517       Sum of wgt.         278

50%     8.749136                      Mean           9.641695
                       Largest        Std. dev.      3.897401
75%     11.10208       24.73208
90%     13.88313       25.21701       Variance       15.18973
95%     16.47351         27.096       Skewness       1.993094
99%     25.21701       32.22426       Kurtosis       9.407814
```

Anexo 3 - Especialização por percentis

```
.
. summarize special, detail

                      Especialização
-------------------------------------------------------------
      Percentiles      Smallest
 1%     .0859468       .0804193
 5%     .1222839       .0844464
10%     .1396595       .0859468       Obs                 278
25%     .1716786       .0935303       Sum of wgt.         278

50%     .2298491                      Mean           .2596194
                       Largest        Std. dev.      .1189148
75%     .3190592        .601358
90%      .425599       .6041353       Variance       .0141407
95%     .5394024       .6337541       Skewness       1.139291
99%     .6041353       .6398433       Kurtosis       3.940759
```

Anexo 4 - Diversificação em geral por percentis

```
. summarize diversif, detail
```

Diversificação

	Percentiles	Smallest		
1%	.3573201	.323519		
5%	.4234297	.3437171		
10%	.5210032	.3573201	Obs	278
25%	.6225637	.3647465	Sum of wgt.	278
50%	.685409		Mean	.6521862
		Largest	Std. dev.	.0898859
75%	.7127139	.7431141		
90%	.7266741	.7448873	Variance	.0080795
95%	.732041	.7466321	Skewness	-1.671902
99%	.7448873	.7472652	Kurtosis	5.265717

Anexo 5 - Diversificação relacionada por percentis

```
. summarize divers_r, detail
```

Diversificação relacionada

	Percentiles	Smallest		
1%	.3269351	.2919647		
5%	.4671681	.3031895		
10%	.5968466	.3269351	Obs	278
25%	.7972263	.3676668	Sum of wgt.	278
50%	.9813004		Mean	.9190846
		Largest	Std. dev.	.2171151
75%	1.074918	1.229675		
90%	1.14562	1.240657	Variance	.0471389
95%	1.198879	1.293774	Skewness	-.8811396
99%	1.240657	1.311259	Kurtosis	3.153952

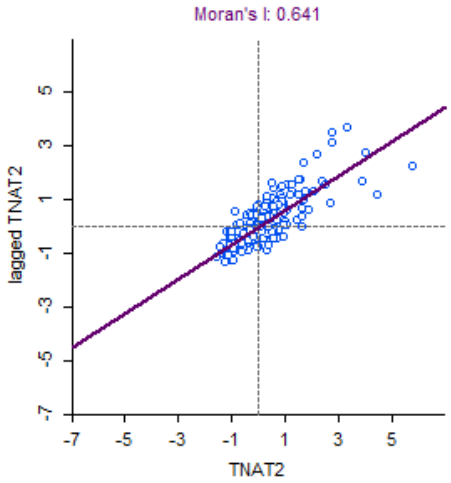
Anexo 6 - Diversificação não relacionada por percentis

```
. summarize divers_nr, detail
```

Diversificação não relacionada

	Percentiles	Smallest		
1%	1.8797	1.753299		
5%	2.24738	1.869765		
10%	2.723554	1.8797	Obs	278
25%	3.115922	1.932017	Sum of wgt.	278
50%	3.335353		Mean	3.20399
		Largest	Std. dev.	.3730166
75%	3.435639	3.591546		
90%	3.501281	3.594678	Variance	.1391413
95%	3.525723	3.602795	Skewness	-1.875593
99%	3.594678	3.661836	Kurtosis	6.133028

Resultados utilizando uma matriz de contiguidade do tipo rainha:



Anexo 7 - I de Moran Univariado, contiguidade tipo rainha

Variáveis	Índice de *Moran* Univariado
TNAT	0.641 (0.002)
URBAN	0.764 (0.002)
SPECIAL	0.626 (0.002)
DIVERSIF	0.704 (0.002)
DIVERS_R	0.726 (0.002)
DIVERS_NR	0.648 (0.002)
DESEMP	0.481 (0.002)
DENSEMP	0.542 (0.002)

Anexo 8- I de Moran Univariado para todas as variáveis, contiguidade tipo rainha

Anexo 9 - Resultados dos modelos OLS e SEM usando matriz de contiguidade do tipo rainha

	OLS	**SEM**	**OLS**	**SEM**	**OLS**	**SEM**	**OLS**	**SEM**
	(1)		(2)		(3)		(4)	
Constante	11.812*** (9.9348)	10.110*** (7.5596)	1.519 (1.0359)	2.2171 (1.3131)	4.5471*** (4.1562)	4.7933*** (3.7269)	0.5220 (0.3163)	1.6338 (0.9037)
URBAN	0.1006 (0.5915)	0.4627** (2.2952)	0.2728* (1.7216)	0.6261*** (3.3601)	0.3179** (1.9933)	0.7296*** (4.0082)	0.3044* (1.9497)	0.6147*** (3.2854)
SPECIAL	- 9.5950*** (-5.5278)	- 6.7355*** (-3.7502)	————	————	————	————	————	————
DIVERSIF	————	————	10.633*** (4.9748)	8.1875*** (3.3449)	————	————	————	————
DIVERS_R	————	————	————	————	3.9314*** (4.4009)	2.5645* (2.4764)	————	————
DIVERS_NR	————	————	————	————	————	————	2.4795*** (4.8971)	1.8708*** (3.4087)
DESEMP	- 0.0668 (-1.1324)	- 0.00946 (-1.5957)	- 0.0538 (-0.8983)	- 0.08625 (-1.4432)	- 0.0451 (-0.7412)	- 0.0886 (-1.4685)	- 0.0654 (-1.0942)	- 0.08936 (-1.4982)
DENSEMP	0.01963*** (10.550)	0.01356*** (7.71148)	0.0186*** (10.094)	0.01262*** (7.3252)	0.0184*** (9.8858)	0.0122*** (7.0599)	0.0185*** (10.037)	0.01271*** (7.3667)
Lambda (SEM)		0.6965***		0.6959***		0.6959***		0.6971***
N	278	278	278	278	278	278	278	278
R^2	0.5163	0.7252	0.5069	0.7224	0.4978	0.7175	0.5056	0.7229
Log Likelihood	- 671.162	- 609.791	- 673.847	- 611.169	- 676.382	- 613.620	- 674.206	-610.957
Teste Breusch-Pagan	60.778*** (0.00)	70.4159*** (0.00)	61.562*** (0.00)	68.556*** (0.00)	67.376*** (0.00)	67.485*** (0.00)	57.011*** (0.00)	68.785*** (0.00)
Likelihood Ratio Test		122.742***		125.357***		125.524***		126.498***