



**UNIVERSIDADE ESTADUAL PAULISTA  
FACULDADE DE CIÊNCIAS E LETRAS  
DEPARTAMENTO DE ECONOMIA**

**TAIS DA SILVA PERUSSI**

**UMA ANÁLISE ESPACIAL DO CRESCIMENTO ECONÔMICO DOS  
MUNICÍPIOS PAULISTAS NO PERÍODO DE 2000-2010**

**ARARAQUARA  
JUNHO DE 2013**

**TAIS DA SILVA PERUSSI**

**UMA ANÁLISE ESPACIAL DO CRESCIMENTO ECONÔMICO DOS  
MUNICÍPIOS PAULISTAS NO PERÍODO DE 2000-2010**

**Orientador: Prof. Dr. Alexandre Sartoris Neto**

**Banca Examinadora: Prof. Dr. Elton Eustáquio Casagrande**

Monografia apresentada ao Conselho de Curso  
de Ciências Econômicas da Faculdade de  
Ciências e Letras – UNESP de Araraquara,  
como requisito para obtenção do título de  
Bacharel em Ciências Econômicas

**ARARAQUARA  
JUNHO DE 2013**

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus por iluminar meu caminho e minhas escolhas.

Agradeço a meus pais José e Zilda por todo o amor e apoio, por me guiarem para o caminho certo e por nunca desistirem do meu futuro, apesar de todas as dificuldades que enfrentaram para que eu pudesse chegar até aqui.

Agradeço ao meu namorado Jair por todo o amor e carinho, por sua paciência, por estar sempre ao meu lado e por seu precioso incentivo.

Agradeço a minha irmã Paula pelos momentos divertidos, por sua inteligência e por sua incondicional ajuda.

Agradeço a toda a minha família pelo constante apoio.

Agradeço a meus amigos que estando perto ou distante sempre estarão no meu coração.

Agradeço ao Professor Alexandre Sartoris pelas orientações, sugestões, enfim, por todo o auxílio na realização deste trabalho.

Agradeço a todos os meus professores pelo conhecimento compartilhado.

## **RESUMO**

O presente trabalho tem como objetivo fazer uma análise do crescimento econômico dos municípios paulistas entre os anos de 2000 e 2010 baseando-se, principalmente, no modelo de Glaeser et al. (1995). Essa análise será feita relacionando o crescimento com variáveis socioeconômicas como a renda, infra-estrutura, emprego e educação. Além disso, serão acrescentadas técnicas da econometria espacial para verificar os efeitos da aglomeração e possíveis externalidades provocadas pelo crescimento. Por fim, será feita uma comparação com o trabalho de Vieira (2008) que fez uma análise semelhante dos municípios paulistas para o período de 1980 a 2000.

**Palavras chave:** crescimento econômico, econometria espacial e externalidades.

## SUMÁRIO

Lista de Figuras	VI
Lista de Tabelas	VII
1. Introdução	8
2. Revisão de Literatura	12
2.1 Teorias do Crescimento Econômico	12
2.2 Crescimento Econômico Regional	14
2.3 Crescimento Econômico dos Municípios	16
3. Econometria Espacial e Descrição dos Dados	19
3.1 Econometria Espacial	19
3.2 Matriz de Pesos Espaciais	21
3.3 Modelos Autoregressivos Espaciais	23
3.4 Testes de Especificação	25
3.5 Dados utilizados	26
4. Resultados	34
4.1 Modelo Empírico	34
4.2 Análise dos Resultados Econométricos	34
4.3 Comparação dos Resultados com Vieira (2008)	41
5. Conclusão	44
Referências Bibliográficas	47

## LISTA DE FIGURAS

Figura 1. Diagrama de Moran	20
Figura 2. Dados em treliças	21
Figura 3. Índice de Moran (I)	30
Figura 4. Distribuição espacial da população dos municípios paulistas no ano 2000	31
Figura 5. Distribuição espacial da população dos municípios paulistas no ano de 2010	31
Figura 6. Distribuição espacial da taxa de crescimento populacional médio dos municípios paulistas (período de 2000 a 2010)	32
Figura 7. Mapeamento dos resultados da metodologia LISA	33

## LISTA DE TABELAS

Tabela 1. Variáveis municipais no ano de 2000 e variação da população entre 2000 e 2010	28
Tabela 2. Matriz de correlação das variáveis municipais	29
Tabela 3. Resultado da Estimação por MQO do modelo utilizando a matriz Rainha	35
Tabela 4. Resultado da Estimação para um modelo SAR	36
Tabela 5. Resultado da Estimação para um modelo SAR	37
Tabela 6. Resultado da Estimação para um modelo SAR	38
Tabela 7. Resultado da Estimação para um modelo SAR	39

## 1. INTRODUÇÃO

O estado de São Paulo é formado por 645 municípios distribuídos em uma área de 248.222,801 km<sup>2</sup> e possui cerca de 41 milhões de habitantes, o que o torna o estado mais populoso do Brasil. A população urbana chega ao percentual de 95% e pode ser dividida em 15 regiões administrativas. De acordo com os dados da SAEDE, São Paulo é o estado com maior PIB, com um valor de aproximadamente 1 bilhão de reais, sendo que as principais contribuições para esse valor vêm dos setores secundário e terciário, mas principalmente deste último. As exportações foram de cerca de 65 trilhões e as importações chegaram ao patamar de 77 trilhões de dólares em 2012.

O setor primário representa 1,87% do valor adicionado. A agricultura está baseada principalmente no plantio da cana-de-açúcar, do café e da laranja. Já a criação de animais baseia-se principalmente em bovinos e aves. O setor secundário corresponde a 29% do valor adicionado e tem como principais representantes as indústrias metalúrgica, automobilística, de alta tecnologia, têxtil e de biotecnologia. O setor terciário tem participação de 69% no valor adicionado, o que demonstra a importância que tem os serviços para o PIB do estado.

A infraestrutura dos municípios paulistas abrange quase toda a população do estado, chegando a 98% as casas com acesso à água encanada e 93% dos domicílios com acesso à rede geral de esgoto. A coleta de lixo chega a 99% das casas do estado e o acesso à energia elétrica é quase universal com percentual de 98% de domicílios atendidos.

A taxa de natalidade chega a 14,68 por mil habitantes, enquanto a mortalidade infantil possui um índice muito baixo de 11,55 para cada mil nascidos vivos. A expectativa de vida no estado é de 74,8 anos, o que deixa o São Paulo com a quinta melhor expectativa de vida no Brasil. Já a taxa de homicídios é de 10,47 mortes para cada 100 mil habitantes, um índice bem menor que a média nacional que é de 24,5 mortes para cada 100 mil habitantes.

Com relação à educação os indicadores também apresentam bons resultados, como por exemplo, a população com idades entre 15 e 64 anos tem em média 9,22 anos de estudo e a taxa de analfabetismo da população de 15 anos ou mais é de 4,3%, enquanto que essa taxa nacional atinge 9,6%. O estado é conhecido pelo alto nível do ensino superior, abrigando várias das melhores universidades públicas do país como a USP, UNESP, UNICAMP, ITA, UNIFESP e UFSCar, além de importantes centros de pesquisa como, por exemplo, o Instituto Butantan, o Instituto de Pesquisas Tecnológicas (IPT), o Instituto

Nacional de Pesquisas Espaciais (INPE) e o Instituto de Pesquisas Energéticas e Nucleares (IPEN).

São Paulo possuía uma taxa de desemprego de 9,8% em 2009 e o grau de informalidade na economia chegou a 37%. Com relação ao emprego formal, a participação do emprego no setor primário no total de empregos no estado é de 2,7%, já o emprego na indústria corresponde a 20,9%. A construção contribui com 5,5% do emprego formal. O emprego no comércio atacadista e varejista corresponde a 19,3% e ao setor de serviços corresponde 51,6% dos empregos formais no estado.

Como foi dito anteriormente o estado está dividido em 15 regiões administrativas que são as regiões de Araçatuba, Barretos, Bauru, Campinas, Central, Franca, Marília, Presidente Prudente, Registro Ribeirão Preto, São José dos Campos, São José do Rio Preto, Sorocaba, Baixada Santista e São Paulo. Para cada uma dessas regiões será feita uma breve análise quanto as suas singularidades com base em um trabalho realizado pela Secretária de Economia e Planejamento do Governo do Estado de São Paulo em 2011.

A região metropolitana de São Paulo é a mais populosa do estado com 47,8% da população total e a que tem o maior percentual do PIB, com 56,4% do total do estado. O setor de serviços é o que mais concentra empresas somando 44% do total da região, isso se deve principalmente ao fato de a capital do estado estar em uma região em que há grande concentração de pessoas. Além disso, a região concentra grande parte das sedes de empresas localizadas no estado.

A região de São José dos Campos também possui um setor de serviços bem desenvolvido devido, principalmente, à alta concentração populacional. Possui o terceiro maior PIB do estado correspondendo a 5,2% do total. A região concentra várias empresas de alta tecnologia, como por exemplo, a EMBRAER e institutos de pesquisa como o IMPE.

A região de Campinas igualmente concentra boa parte da população e do PIB do estado, o que faz com que a região tenha o setor de serviços e comércio bem evoluído, além de ser uma importante região agrícola do estado, destacando-se pela produção de flores, café e pelo circuito das frutas. A indústria também possui um espaço de destaque na economia da região, que concilia indústrias menos complexas, como as têxteis em Americana, e indústrias de maior valor agregado, como as de equipamentos eletrônicos.

A região da Baixada Santista tem o setor de serviços e de comércio como seu ponto forte, consequência do turismo na região. A indústria também tem importante contribuição para o PIB local, principalmente em Cubatão, onde está localizado um polo petroquímico.

A região Central possui uma forte influência do setor primário na sua economia. O principal destaque é o cultivo da cana-de-açúcar e a agroindústria ligada a esse cultivo. A região possui tanto indústrias mais simples, como por exemplo, as de bens de consumo não durável, como indústrias de alto valor agregado.

Outras regiões do estado têm uma estrutura muito parecida com a região Central, como as de Bauru e Ribeirão Preto, que têm na produção cana de açúcar e no setor da agroindústria grandes contribuidores do PIB dessas regiões. Além disso, a região de Bauru possui importantes indústrias de calçados, localizadas principalmente na cidade de Jaú, e igualmente influentes na composição da economia do estado estão as indústrias farmacêuticas, farmoquímicas e de produtos eletrônicos, localizadas na região de Ribeirão Preto.

A região de Franca também é um polo da indústria de calçados, além de ser uma região que se destaca na produção de gêneros alimentícios, biocombustíveis e metalurgia, entre outros. A região de Barretos, assim como quase todo o interior do estado, tem como principal produto agrícola a cana de açúcar e seu principal setor industrial está ligado a essa cultura, ou seja, a indústria de biocombustíveis, além de abrigar uma importante indústria alimentícia para a região.

As regiões de São José do Rio Preto, Araçatuba e Presidente Prudente possuem uma dinâmica econômica muito semelhante, nas quais predomina a agroindústria ligada à plantação de cana-de-açúcar e à indústria de bens de consumo não durável. O maior destaque da indústria vem da produção de alimentos, confecções, de móveis e de bebidas. A região de Marília segue a mesma linha das regiões citadas anteriormente, diferenciando-se apenas no fato de possuir uma indústria de maior valor agregado, que produz máquinas, equipamentos e produtos elétricos.

A região de Sorocaba também é marcada pela forte dependência da agricultura, principalmente do cultivo da cana-de-açúcar. A produção florestal também tem importante contribuição para a economia local. A indústria da região destaca-se produzindo tanto bens de consumo não durável como produtos mais complexos. Por fim, a região de Registro é marcada pela produção florestal, e tem nas indústrias alimentícias e de produtos não metálicos, importantes contribuições para a economia da região.

É possível notar que as regiões mais próximas da capital têm o setor de serviços mais evoluído, enquanto que regiões mais distantes demonstram grande dependência do cultivo da cana-de-açúcar e da agroindústria ligada a esse plantio. As indústrias estão relativamente dispersas pelo estado, mas as indústrias de maior valor agregado

concentram-se mais desde a região de São José dos Campos, passando pela região metropolitana de São Paulo, na região de Campinas e chegando até a região Central. O estado possui uma região mais pobre quando em comparação a outras: a região de Registro e uma parte da região de Sorocaba, uma vez que ainda dependem muito do setor primário e não possuem um setor industrial e de serviços expressivos.

Este trabalho procurará analisar o crescimento econômico dos municípios paulistas, na tentativa de identificar que fatores contribuem para o elevado crescimento de algumas cidades, enquanto outras permaneceram, até a presente pesquisa, praticamente estagnadas ou tiveram crescimento negativo.

Para entender a diferença de crescimento entre os municípios paulistas, esse trabalho utilizará as principais contribuições do *mainstream* para a análise do crescimento econômico. Será analisada a renda, o nível educacional e a infraestrutura municipal, variáveis escolhidas com base nos trabalhos de Solow e autores cujas pesquisas discorrem sobre o crescimento endógeno. Já os indicadores sociais, participação da indústria no PIB, urbanização, emprego e desigualdade foram inspirados nos trabalhos da Nova Geografia Econômica (NGE). O modelo a ser utilizado será baseado no trabalho de Glaeser et al. (1995), que analisou o crescimento econômico de cidades americanas.

Serão verificadas também possíveis externalidades causados pelo crescimento de um município sobre os municípios próximos. Para isso, serão utilizadas ferramentas da econometria espacial baseando-se nos trabalhos de Anselin (1988) e Lesage (1999). Por fim, será feita uma comparação com o trabalho de Vieira (2008), que fez uma análise semelhante do crescimento econômico dos municípios paulistas para o período de 1980 a 2000. Por meio dessa comparação será possível saber se ocorreram grandes mudanças no padrão de crescimento dos municípios analisados no período de 2000 a 2010 ou se tais mudanças não se mostraram tão notórias.

## 2. REVISÃO DE LITERATURA

### 2.1 Teorias do Crescimento Econômico

O crescimento econômico há muito tempo vem sendo assunto de grande interesse por parte dos economistas. De acordo com Barro & Sala-I-Martin (1995) foram Smith (1776), Ricardo (1817), Malthus (1798) e tempos depois Ramsey (1928), Young (1928), Knight (1944) e Schumpeter (1934) que forneceram, por meio de seus estudos, as ideias básicas a serem utilizadas pelas novas teorias do crescimento econômico.

Solow foi o um dos pioneiros das teorias do crescimento neoclássicas em 1956 ao publicar o trabalho “A contribution to the theory of economic growth”. Nesse estudo o autor apresenta um modelo teórico para o crescimento dos países com base nos pressupostos neoclássicos de concorrência perfeita, retornos constantes de escala, retornos decrescentes de cada fator de produção e elasticidade de substituição entre fatores de produção positiva. O trabalho apresenta o crescimento do produto dependente da poupança, da taxa de crescimento da população e do progresso tecnológico, sendo que os dois últimos parâmetros não são explicados pelo modelo, ou seja, são determinados exogenamente.

Outro ponto de grande importância no modelo de Solow é a convergência dos níveis de rendas dos países, independente das condições iniciais, devido à suposição de retornos decrescentes do capital. Logo, um país que possui baixa renda *per capita* tende a ter maiores taxas de crescimento do que os países com alto nível de renda *per capita* até chegarem ao mesmo estado estacionário, no qual tal renda crescerá a uma taxa constante.

O modelo de Solow promoveu um grande avanço nos estudos de crescimento econômico, mas, apesar disso, sua teoria continha pontos questionáveis, como a taxa de progresso tecnológico que não era explicada endogenamente e a suposição de que todos os países cresceriam a uma mesma taxa constante no *steady-state*.

Seguindo a linha neoclássica, Cass (1965) e Koopmans (1965) criaram modelos com a taxa de poupança endógena. Segundo Barro & Sala-I-Martin (1995) esse estudo ainda preserva a hipótese de convergência condicional, além de não eliminar a dependência do crescimento da renda *per capita* no longo prazo e a taxa de crescimento tecnológico continuar exógena. Já Arrow (1962) e Sheshinski (1967) se apoiam na ideia de que o conhecimento é um subproduto inesperado da produção.

Anos mais tarde com base nestes trabalhos, surgiu um novo grupo de estudos sobre crescimento econômico que tinha como pretensão aprofundar as análises e encontrar soluções para os problemas enfrentados nesses primeiros modelos. Exemplos disso são Lucas (1988) e Romer (1986) que procuraram fazer com que o crescimento econômico fosse explicado por variáveis internas ao modelo e, por isso, seus estudos ficaram conhecidos como Teoria do crescimento endógeno.

No modelo de Romer (1986) a ênfase é no crescimento em longo prazo, considerado essencial por tais modelos. Segue ainda, a mesma linha de concorrência perfeita, mas com retornos crescentes de escala. Nessa teoria o crescimento econômico é impulsionado pela acumulação de conhecimento, tratado como uma forma de capital, o que gera externalidades positivas uma vez que o conhecimento não fica restrito a um agente, podendo ser aproveitado por todo o restante da economia.

Com essas novas premissas, o modelo de Romer (1986) não supõe a obrigatória convergência do crescimento da renda dos países, mas sim que as rendas podem crescer com taxas diferentes e para patamares distintos, ou muitas vezes, como no caso dos países menos desenvolvidos, podem até diminuir. A velocidade de crescimento também é considerada desigual entre os países. Já em 1990, Romer adota em seu trabalho a concorrência monopolística em seu modelo de crescimento.

Lucas (1988) criou dois modelos para fazer sua análise do crescimento. Em ambos são utilizados alguns dos pressupostos da Teoria exógena, como a concorrência perfeita e uma taxa constante de crescimento da população. O primeiro modelo possui apenas um setor e o que impulsiona o crescimento econômico é a acumulação de capital humano. Já o segundo modelo possui dois setores e o crescimento da renda *per capita* deriva do conhecimento adquirido *learning by doing*, ou seja, a aquisição do aprendizado se dá através do trabalho. Tanto seus estudos como os de Romer (1967, 1990) demonstram que um país pode crescer sem interrupção, pois não pressupõem que os rendimentos do investimento em bens de capital diminuam à medida que a economia se desenvolve (Barro & Sala-I-Martin, 1995).

Os estudos posteriores mantiveram o centro de seus trabalhos nos investimentos em P&D como forma de crescimento econômico dos países. Além disso, para essas teorias, *spillovers* de conhecimento também são necessários para o desenvolvimento. Outro ponto

que ganhou grande destaque nessas novas teorias foram os testes empíricos, considerados agora fundamentais para a real análise desses modelos.

## **2.2 Crescimento econômico regional**

Enquanto os autores citados acima procuravam entender o crescimento de países, outros se dedicavam ao estudo do crescimento de regiões. A principal questão analisada é o que leva a atividade econômica a concentrar-se em uma localidade e, neste sentido, também é questionado porque algumas aglomerações tendem a crescer mais do que outras e porque algumas atividades têm sucesso em certas regiões e em outras acabam fracassando.

A chamada Teoria da localização teve início com Thünen (1826), que criou um modelo para analisar a alocação de terras agrícolas em torno de centros urbanos. O estudo propõe um mercado de terras competitivo, em que a ocupação depende do custo de transporte, ou seja, nas terras próximas ao centro urbano seriam plantados produtos com maior custo de transporte e, por consequência, os que possuíssem menor custo ficariam nas terras periféricas. Essa distribuição ocorre devido à disposição dos agentes a pagar pelas terras de acordo com o produto que pretendem produzir e com a renda que conseguirão alcançar com sua utilização. Esse modelo de ocupação recebeu o nome de Anéis concêntricos. Alonso (1964) passou a aplicar esse modelo de Thünen (1826) aos centros urbanos, utilizando o mesmo conceito de ocupação de terras agora para fins industriais, conforme a distância do centro. Esse modelo ficou conhecido como Cidade monocêntrica e o estudo da distribuição alocativa das atividades nos espaços das cidades passou a ser chamado de Economia urbana.

Marshall (1920) estudou os benefícios da aglomeração industrial para a economia. De acordo com o autor, a proximidade de localização entre as empresas proporciona externalidades de conhecimento e tecnologia que serão aproveitadas pelas indústrias da região, assim os produtores de insumos podem oferecer produtos mais especializados e em maior escala. Além disso, haverá uma grande quantidade de mão de obra disponível e especializada. Todos esses benefícios geram ganhos de escala externos às firmas, porém internos ao setor.

Ao contrário de Marshall (1920) que prevê a especialização como forma de ganhos para indústrias vizinhas, Jacobs (1969) acredita que a concentração de variados tipos de atividades econômicas em localidades próximas gera mais benefícios para o crescimento

das cidades. Outro ponto importante em sua teoria é que essa diversidade promove o aumento de ideias inovadoras.

Christaller (1933) também procurou entender como são alocados os espaços dentro de uma cidade, por meio da Teoria das áreas centrais, na qual as cidades localizam-se próximas para formar um centro urbano maior. Em seu modelo, o custo de transporte tem papel fundamental, já que empresas que produzem bens inferiores devem situar-se perto dos consumidores, ao contrário das produtoras de bens superiores que podem estar mais afastadas, pois seus produtos não são comprados com tanta frequência. O autor propôs um modelo de cidade que diminuísse o número de centros, além de reduzir a concentração de indústrias produtoras do mesmo bem a pouca distância do mesmo mercado consumidor. Lösch (1940) seguiu a mesma linha de Christaller e determinou que o custo de transporte e a elasticidade-preço dos produtos determinam a localização das indústrias nos centros urbanos.

Para aprimorar os conceitos introduzidos pela Economia regional era necessário ajustar as variáveis e os modelos, que eram os mesmos dos estudos sobre nações, para diferentes regiões de um único país. Isso foi possível quando a análise regional passou a incorporar a concorrência monopolística de Dixit-Stiglitz (1977), que permitiu aos novos estudos captarem as desigualdades de crescimento entre as regiões dentro de uma nação. A partir disso surgiu uma nova corrente de pensamento conhecida como A Nova Geografia Econômica (NGE).

Krugman (1991) foi o precursor dessa corrente e de acordo com seu modelo, conhecido como centro-periferia, um fator que induz a concentração de determinada atividade econômica é seguido por mais situações que aumentam essa concentração e acentuam a desigualdade entre as regiões. Como exemplo, esse modelo demonstra que a indústria tende a se concentrar em uma determinada região com fatores econômicos favoráveis, proporcionando economias de escala e diminuição do custo de transporte. Essa concentração da indústria em uma determinada região seria impulsionada por um deslocamento da mão de obra para essas localidades.

A Nova Geografia Econômica propõe que a concentração em torno de uma região promove ganhos tanto para as indústrias quanto para os que ali residem. Para a indústria essa aglomeração gera um aumento de algumas unidades, o que causa o aparecimento da concorrência monopolística e de retornos crescentes. O custo de transporte diminui e pode ocorrer transbordamento de conhecimento devido à proximidade desses estabelecimentos.

Já para os indivíduos essa concentração promove ganhos nos salários além de maior disponibilidade de produtos industriais.

### 2.3 Crescimento econômico dos municípios

Para fazer o estudo do crescimento econômico dos municípios do estado de São Paulo, este trabalho será baseado no de Glaeser et al. (1995) que examina o crescimento econômico de 203 municípios americanos no período de 1960 a 1990. Para isto ele utiliza um modelo em que a principal variável para explicar o crescimento econômico é o crescimento populacional. Além dessa variável, o modelo procura explicar esse crescimento através da renda inicial, do emprego inicial no setor industrial, do desemprego inicial, do nível inicial da educação nessas cidades, da composição racial, da segregação, do crescimento passado e da composição do governo.

Para Glaeser et al. (1995) há maior probabilidade de ocorrer convergência de renda entre cidades do que entre países, pois há maior mobilidade de pessoas, capital e conhecimento e, além disso, as cidades são mais especializadas. Ele também acredita que as externalidades de conhecimento são importantes para o crescimento urbano.

No modelo, o autor assume livre mobilidade de capital e trabalho e, sendo assim, as cidades diferem no nível de produtividade e na qualidade de vida. A função de produção é dada por:

$$A_{i,t} f(L_{i,t}) = A_{i,t} L_{i,t}^{\sigma} \quad (2.3.1)$$

Onde  $A$  é o nível de produtividade,  $L$  a população,  $t$  o tempo,  $i$  a cidade analisada e  $\sigma$  um parâmetro de produção nacional. A renda do trabalho de um imigrante em potencial será dada pela produção marginal do trabalho:

$$W_{i,t} = \sigma A_{i,t} L_{i,t}^{\sigma-1} \quad (2.3.2)$$

A utilidade total é igual ao salário multiplicado pelo índice de qualidade de vida. Esse índice diminui com o aumento do tamanho da cidade.

$$Qualidade\ de\ Vida = Q_{i,t} L_{i,t}^{-\delta} \quad (2.3.3)$$

Sendo que  $\delta$  é maior que zero. No índice de qualidade de vida estão incluídos vários fatores como o crime, o preço da habitação e o congestionamento do trânsito.

A utilidade de um potencial imigrante é dada por:

$$Utilidade = \sigma A_{i,t} Q_{i,t} L_{i,t}^{\sigma-\delta-1} \quad (2.3.4)$$

A partir da equação (2.3.4) tem para cada cidade:

$$\log\left(\frac{U_{t+1}}{U_t}\right) = \log\left(\frac{A_{i,t+1}}{A_{i,t}}\right) + \log\left(\frac{Q_{i,t+1}}{Q_{i,t}}\right) + (\sigma - \delta - 1)\log\left(\frac{L_{i,t+1}}{L_{i,t}}\right) \quad (2.3.5)$$

Também se assume que:

$$\log\left(\frac{A_{i,t+1}}{A_{i,t}}\right) = X'_{i,t}\beta + \varepsilon_{i,t+1} \quad (2.3.6)$$

$$\log\left(\frac{Q_{i,t+1}}{Q_{i,t}}\right) = X'_{i,t}\theta + \xi_{i,t+1} \quad (2.3.7)$$

Onde  $X_{i,t}$  é um vetor de características da cidade no tempo  $t$ , que determina tanto o crescimento da qualidade de vida na cidade quanto o crescimento do nível de produtividade. Combinando (2.3.5), (2.3.6) e (2.3.7) temos:

$$\log\left(\frac{L_{i,t+1}}{L_{i,t}}\right) = \frac{1}{1+\delta-\sigma} X'_{i,t}(\beta + \theta) + \chi_{i,t+1} \quad (2.3.8)$$

$$\log\left(\frac{W_{i,t+1}}{W_{i,t}}\right) = \frac{1}{1+\delta-\sigma} X'_{i,t}(\delta\beta + \sigma\theta - \theta) + \bar{\omega}_{i,t+1} \quad (2.3.9)$$

Onde  $\chi_{i,t}$  e  $\bar{\omega}_{i,t}$  são os termos de erro não correlacionados com características urbanas. Glaeser et al. (1995) afirma que a regressão do crescimento do emprego mostra que as variáveis ao nível da cidade determinam o crescimento da soma da qualidade de vida e da produtividade. A regressão de crescimento do salário pode ser definida como uma média ponderada do crescimento da produtividade e  $(\sigma-1)$  vezes o crescimento da qualidade de vida.

Com esse modelo os autores demonstram que o crescimento das cidades depende praticamente das mesmas características que fazem os países crescerem economicamente. O nível de escolaridade inicial tem papel fundamental nesse crescimento, o desemprego menor e a menor participação da indústria no emprego total das cidades também se mostraram significativas no modelo. Com relação às variáveis sociais, o crescimento da renda urbana está correlacionado positivamente com maior nível de escolaridade dos trabalhadores. A segregação não teve efeitos sobre o crescimento econômico, bem como os gastos governamentais.

Assim como o trabalho de Vieira (2008) e Oliveira (2005) este trabalho também acrescenta ao modelo de Glaeser et al. (1995) uma análise espacial, a fim de captar

possíveis transbordamentos de externalidades entre os municípios. Utilizando técnicas da Econometria espacial e baseado no trabalho de Anselin (1988) será possível especificar o modelo correto e incluir operadores de defasagem espacial.

### 3. ECONOMETRIA ESPACIAL E DADOS UTILIZADOS

#### 3.1 Econometria espacial

O aprimoramento dos estudos sobre Crescimento econômico e Economia regional levou a um aumento do interesse pela questão espacial dos dados analisados. A Econometria tradicional não possuía todos os instrumentos necessários para fazer a análise correta, o que levou ao surgimento de uma derivação, a chamada Econometria espacial. Segundo Anselin (1988) os dados espaciais agregados são caracterizados por dependência (autocorrelação espacial) e heterogeneidade (estrutura espacial), o que pode tornar os resultados obtidos não confiáveis, caso não sejam estimados com os métodos corretos.

A dependência espacial é a relação entre os dados de localidades próximas. Não apenas a proximidade com relação a distância física, mas também a proximidade econômica e política, por exemplo, podem influenciar variáveis. Para Anselin (1988) essa dependência pode ser causada pela arbitrariedade do próprio autor ao delinear unidades de estudo, além da presença de externalidades espaciais e efeitos de *spillover*, ou seja, um evento ocorrido em uma variável não ficava restrito à sua localidade, podendo ultrapassar esse limite e chegar a influenciar variáveis de lugares próximos.

A heterogeneidade espacial é caracterizada pela falta de padrão das variáveis no espaço. Sendo assim, ela pode variar ou formar padrões específicos conforme a localidade ou não formar um conjunto homogêneo de dados, o que poderia gerar heterocedasticidade. Como analisou Anselin (1988) a heterogeneidade espacial pode ser na maioria das vezes, corrigida utilizando a economia padrão, o que não pode ser feito com a autocorrelação espacial. De acordo com Resende (2005) o problema torna-se mais complexo em situações onde a dependência e a heterogeneidade espacial estão presentes ao mesmo tempo, pois nessa situação a Econometria padrão não possui as ferramentas adequadas, o que torna necessário a utilização da Econometria espacial.

Para verificar a dependência espacial existem vários tipos de testes estatísticos. O mais utilizado é o teste de Moran (I), que analisa a existência de algum padrão de valores de uma variável ao longo do espaço. A existência desse padrão indica uma autocorrelação espacial que pode ser positiva ou negativa. Os valores próximos de zero do teste indicam a quase inexistência, enquanto que valores próximos de -1 e 1 indicam uma autocorrelação desses valores. Segundo Vieira (2008) este indicador é dado por:

$$I = \left( \frac{n}{s_0} \right) \left( \frac{Z'WZ}{Z'Z} \right) \quad (3.1.1)$$

Onde  $Z$  é o vetor,  $n$  o número de observações para o desvio em relação à média, e  $S_0$  é um escalar igual à soma de todos os elementos de  $W$ . Sendo o valor esperado:

$$E(I) = -\frac{1}{n-1} \quad (3.1.2)$$

Quando a soma dos elementos de cada linha for igual a 1, podemos reescrevê-la como:

$$I = \frac{Z'WZ}{Z'Z} \quad (3.1.3)$$

O gráfico é dado por:

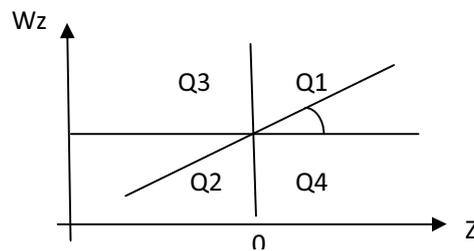


Figura 1: Diagrama de Moran

No primeiro quadrante (Q1) localizam-se altos valores da variável a altos valores da mesma variável em localidades vizinhas, já no terceiro quadrante (Q2) localizam-se baixos valores com baixos valores dos vizinhos. No segundo quadrante (Q3) estão as variáveis com valores baixos e os vizinhos com valores altos, e por fim, no último quadrante (Q4) estão as variáveis com valores altos com variáveis de vizinhos com valores baixos. Os valores localizados em Q1 e Q2 possuem autocorrelação positiva, enquanto que em Q3 e Q4 a autocorrelação é negativa.

Como afirma Vieira (2008) a autocorrelação positiva, apesar de possuir a mesma importância que a negativa, é a mais encontrada nos fenômenos econômicos e é mais intuitiva, enquanto que a negativa possui uma difícil interpretação. Ao contrário da autocorrelação temporal em que valores de uma variável possuem apenas dependência unidirecional, a autocorrelação espacial possui dependência multidirecional, o que pode tornar a utilização dos modelos de econometria tradicional ineficaz para dados com dependência espacial.

### 3.2 Matriz de pesos espaciais

Existem duas maneiras de quantificar a localização dos dados analisados de uma amostra. A primeira consiste, segundo Lesage (1999), na localização em um espaço cartesiano representado pela latitude e pela longitude, que permite encontrar a distância entre pontos no espaço. A segunda maneira seria por meio da contiguidade, que exige o conhecimento do tamanho e da forma das localidades analisadas. Por esse método, dados de localidades são comparados a dados de regiões fronteiriças ou próximas para encontrar possíveis relações entre eles e, para aplicá-lo, é necessário criar uma matriz de pesos espaciais.

A segunda forma é mais aplicada a dados econômicos, mas como afirma Lesage (1999), isso não impede que as duas possam ser usadas juntas, já que a latitude e a longitude podem ser usadas para formular uma matriz de pesos espaciais. A figura 2 mostra um exemplo de região por meio da qual é possível criar relações de contiguidade entre os dados:

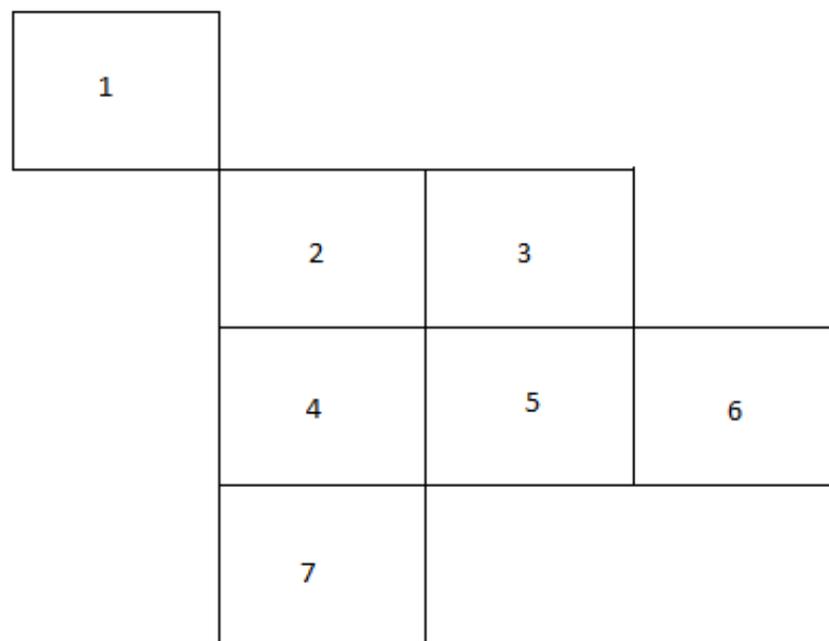


Figura 2: Dados em treliças

São várias as medidas de contiguidade, como por exemplo o critério Rainha (Queen), que aplicado a figura 2 indica que 1 é vizinho de 2 que é vizinho de 1, 3, 4 e 5. Já 3 é vizinho de 2, 4, 5 e 6 e assim por diante. Com essas relações é possível criar uma matriz  $W$ , como a demonstrada a seguir:

$$W = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{pmatrix}$$

Onde a diagonal principal é zero e 1 indica que as localidades são vizinhas de acordo com o critério utilizado. Também é possível criar a matriz  $w$  normalizada, onde o valor de cada elemento é dividido pela soma dos valores da linha. Esse procedimento permite que os valores dessa matriz fiquem em um intervalo entre -1 e 1.

$$w = \begin{pmatrix} 0 & 1/1 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 0 & 1/4 & 1/4 & 1/4 & 0 & 0 \\ 0 & 1/4 & 0 & 1/4 & 1/4 & 1/4 & 0 \\ 0 & 1/4 & 1/4 & 0 & 1/4 & 0 & 1/4 \\ 0 & 1/5 & 1/5 & 1/5 & 0 & 1/5 & 1/5 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 & 0 \end{pmatrix}$$

Lesage (1999) explica que a criação dessa matriz normalizada pode ser explicada multiplicando-a por um vetor de observações associado às regiões como mostrado a seguir:

$$\begin{pmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \\ y_5^* \\ y_6^* \\ y_7^* \end{pmatrix} = \begin{pmatrix} 0 & 1/1 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 0 & 1/4 & 1/4 & 1/4 & 0 & 0 \\ 0 & 1/4 & 0 & 1/4 & 1/4 & 1/4 & 0 \\ 0 & 1/4 & 1/4 & 0 & 1/4 & 0 & 1/4 \\ 0 & 1/5 & 1/5 & 1/5 & 0 & 1/5 & 1/5 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{pmatrix}$$

Onde  $y^*$  representa uma nova variável que indica a média das observações de regiões contíguas:

$$\begin{pmatrix} y_1^* \\ y_2^* \\ y_3^* \\ y_4^* \\ y_5^* \\ y_6^* \\ y_7^* \end{pmatrix} = \begin{pmatrix} y_2 \\ 1/4y_1 & & & & & & \\ & 1/4y_2 & & & & & \\ & & 1/4y_3 & & & & \\ & & & 1/4y_4 & & & \\ & & & & 1/4y_5 & & \\ & & & & & 1/4y_6 & \\ & & & & & & 1/4y_7 \\ & & & & & & & 1/4y_7 \\ & & & & & & & & 1/4y_7 \\ & & & & & & & & & 1/5y_6 & 1/5y_7 \\ & & & & & & & & & & & 1/5y_7 \\ & & & & & & & & & & & & 1/2y_5 \\ & & & & & & & & & & & & & 1/2y_5 \end{pmatrix}$$

Essa é uma maneira de quantificar a expressão  $y_1 = f(y_j), j \neq i$ . Existem outras formas de construir matrizes  $W$  de pesos espaciais levando em consideração, por exemplo, distâncias geográficas ou econômicas. Nesse trabalho, diferente do que foi feito por Vieira (2008), será utilizada apenas uma matriz de pesos espaciais que, no caso, será a matriz binária descrita acima.

### 3.3 Modelos autoregressivos espaciais

De acordo com Lesage (1999), um modelo autoregressivo geral (SAC) pode ser representado da seguinte forma:

$$y = \rho W_1 y + X\beta + u \quad (3.3.1)$$

$$u = \lambda W_2 + \varepsilon$$

$$\varepsilon \sim N(0, \sigma^2 I_n)$$

Onde  $y$  contém  $n \times 1$  vetores de variáveis dependentes,  $X$  representa  $n \times k$  matrizes de variáveis explicativas e  $\varepsilon$  representa o erro aleatório normalmente distribuído. Já  $W_1$  e  $W_2$  representam as matrizes  $n \times n$  de pesos espaciais.  $W_1$  e  $W_2$  normalmente contém relações de continuidade de primeira ordem ou funções de distância, enquanto que  $\rho$ ,  $\lambda$  e  $\varepsilon$  são os parâmetros.

O modelo SAC possui a dependência espacial tanto nas variáveis endógenas do modelo quanto no erro aleatório. Neste caso, os estimadores de MQO são ineficientes e inconsistentes. Para sua correta estimação é necessário otimizar o logaritmo da máxima verossimilhança por meio de uma versão concentrada da função de verossimilhança.

A função de verossimilhança, de acordo com Lesage (1999), é a seguinte:

$$L = C - (n/2) \ln(\sigma^2) + \ln(|A|) + \ln(|B|) - (1/2\sigma^2)(e' B' B e) \quad (3.3.2)$$

$$e = (Ay - X\beta)$$

$$A = (I_n - \rho W_1)$$

$$B = (I_n - \lambda W_2)$$

Enquanto que a função concentrada para  $\beta$  e  $\sigma^2$  é:

$$\beta = (X' A' A X)^{-1} (X' A' A B y) \quad (3.3.3)$$

$$e = B y - X \beta$$

$$\sigma^2 = \frac{e' e}{n}$$

Com essa equação podemos, utilizando os valores de  $\rho$  e  $\lambda$ , encontrar o logaritmo da verossimilhança. Os valores de  $\beta$  e  $\sigma^2$  podem ser encontrados em função de  $\rho$  e  $\lambda$  e os dados amostrais de  $y$  e  $X$ .

A partir desse modelo geral, vários outros podem ser derivados, como por exemplo, quando  $W_2$  e  $X$  são iguais a zero, ou seja, o modelo não possui variáveis explicativas e o erro não apresenta dependência espacial. Pode ser representado da seguinte maneira:

$$y = \rho W y + \varepsilon \quad (3.3.4)$$

$$\varepsilon \sim N(0, \sigma^2 I_n)$$

Nesse modelo a variável  $y$  é explicada somente por uma combinação linear dessa mesma variável nas unidades vizinhas. Outro modelo derivado da representação geral é conhecido como SEM, que é o modelo de erro espacial:

$$\begin{aligned} y &= X\beta + u \\ u &= \lambda Wu + \varepsilon \\ \varepsilon &\sim N(0, \sigma^2 I_n) \end{aligned} \tag{3.3.5}$$

A dependência espacial, como mostra (3.3.5), está no erro. Neste caso,  $y$  contém  $n \times 1$  vetores de variáveis dependentes,  $X$  representa a matriz que possui as variáveis explicativas,  $W$  é a matriz de peso espacial,  $\lambda$  é o coeficiente de correlação espacial do erro e  $\beta$  é um parâmetro. Quando a dependência espacial está localizada no erro significa que essa dependência está em dados que foram omitidos do modelo. Para sua correta estimação é necessário utilizar a função de máxima verossimilhança.

Outro modelo muito comum é o de defasagem espacial (autoregressivo) o SAR:

$$\begin{aligned} y &= \rho Wy + X\beta + \varepsilon \\ \varepsilon &\sim N(0, \sigma^2 I_n) \end{aligned} \tag{3.3.6}$$

Neste caso, a variável de dependência espacial é o próprio  $y$  defasado, além disso, o modelo possui variáveis explicativas. O que se pode concluir é que a variável dependente  $y$ , depende dos valores de  $y$  de seus vizinhos, além de depender de variáveis exógenas ao modelo. Como  $Wy$  é endógena ao modelo, os estimadores de MQO são viesados e inconsistentes, então, para uma correta estimação é necessária a função de máxima verossimilhança, ou a utilização de variáveis instrumentais.

E como analisado por Vieira (2008), existe também a possibilidade da dependência espacial estar correlacionada com a variável exógena, surgindo assim o modelo espacial de Durbin:

$$\begin{aligned} y &= \rho W_1 y + X\beta - \theta W_3 X\beta + \mu \\ \mu &\sim N(0, \sigma^2 I_n) \end{aligned} \tag{3.3.7}$$

Nesse modelo a variável exógena  $X$  depende dos valores das variáveis exógenas das unidades vizinhas, além disso, a variável endógena  $y$  também tem dependência espacial. O problema econométrico desse modelo é que os estimadores de MQO não são eficientes, sendo assim, para sua correta estimação também é utilizada a função de máxima verossimilhança.

### 3.4 Testes de especificação

Como discutido no capítulo anterior, existem vários modelos econométricos com variáveis espacialmente dependentes, como o SAR, o SEM e o SAC entre outros. Nesses modelos podemos encontrar a variável endógena, a variável explicativa ou o erro como sendo dependentes espacialmente e, além disso, essas dependências podem aparecer isoladas, algumas ou todas elas em um mesmo modelo. Durante as últimas décadas, vários autores procuraram encontrar testes para especificar qual o modelo mais adequado para cada base de dados.

Um desses testes é próprio Moran (I) que, como foi detalhado anteriormente, é um teste bidimensional que verifica se existe autocorrelação espacial. O teste varia entre 1 e -1 e caso o valor for próximo de 1, existe uma grande autocorrelação positiva, caso o teste acuse um número negativo, então a autocorrelação também será negativa.

Em sintonia com o trabalho de Vieira (2008) este presente estudo utilizará, além do teste de Moran, o teste de LM (Multiplicador de Lagrange), que inclui o LM-ERR e o LM-LAG. Esse teste utiliza MQO com um  $H_0$  indicando a inexistência de correlação espacial. O modelo LM-ERR pode ser representado da seguinte forma:

$$LM - ERR = \frac{(e'W_1e/s^2)^2}{T_1} \quad (3.4.1)$$

Em que  $s^2 = \frac{e'e}{n}$  e  $T_1 = tr(W_1'W_1 + W_1^2)$ , com  $tr$  como traço operador da matriz. Além disso, segue uma distribuição  $\chi^2$  com 1 grau de liberdade e a hipótese nula é de haver dependência espacial no erro.

Já o modelo LM-LAG é escrito como se segue:

$$LM - LAG = \left( \frac{e'W_1y}{s^2} \right)^2 \frac{1}{(nJ_{\rho-\beta})} \quad (3.4.2)$$

Com  $J_{\rho-\beta} = [T_1 + (W_1X\beta)'M(W_1X\beta)/s^2]$  e  $M = I - X(X'X)^{-1}X'$  que é a matriz de projeção anual. Essa estatística segue uma distribuição  $\chi^2$  com 1 grau de liberdade. A hipótese nula, neste caso, é de haver dependência na variável dependente.

Existem também os testes robustos para essa estatística, que captam possíveis efeitos que os testes simples não apontam. De acordo com Bera & Yoon (1993), o teste LM-EL identifica a dependência espacial no erro testando uma possível dependência espacial da variável dependente, que é representado da seguinte forma:

$$LM - EL = \frac{[e'W_1e/s^2 - T_1(nJ_{\rho-\beta})^{-1}(e'W_1y/s^2)]^2}{[T_1 - T_1^2(nJ_{\rho-\beta})^{-1}]} \quad (3.4.3)$$

Que, assim como os testes simples, também segue uma distribuição  $\chi^2$  com 1 grau de liberdade.

O modelo robusto LM-LE busca captar a dependência espacial na variável endógena testando também uma possível dependência espacial no erro, utilizando o seguinte modelo:

$$LM - LE = \frac{(erW_1y / s^2 - erW_1e / s^2)^2}{nJ_{\rho-\beta-T_1}} \quad (3.4.4)$$

Essa estatística, assim como as anteriores, segue uma distribuição  $\chi^2$  com 1 grau de liberdade.

Com indica Vieira (2008), para seguir um roteiro adequado que identifique a dependência espacial, primeiro é necessário aplicar MOQ ao modelo inicial  $y = X\beta + \varepsilon$ , depois se utilizam os testes de LM-LAG e LM-LRR para verificar a dependência espacial. Com ambos sendo não significativos, então estima-se o modelo apenas por MOQ. Caso o LM-LAG for significativo, mas LM-LRR não, então se utiliza o LM-LAG. Caso ocorra o inverso, ou seja, LM-LRR significativo e LM-LAG não, então o LM-LRR deverá ser utilizado. Se os dois forem significativos, será necessário estimar o mais significativo dos dois por meio do modelo robusto. Contudo, se o modelo robusto mais significativo for LM-EL, se estima utilizando o LM-LRR e se o mais significativo for LM-LE, se estima por meio do LM-LAG.

### 3.5 Dados Utilizados

Os dados utilizados no presente estudo foram obtidos por meio do Instituto Brasileiro de Geografia e Estatística (IBGE). Para que uma comparação com o trabalho de Vieira (2008) seja feita de forma correta, serão utilizadas as mesmas variáveis de seu estudo sobre o crescimento econômico dos municípios do estado de São Paulo entre os anos de 1980 e 2000. Assim como foi feito por Glaeser et al. (1995) e o próprio Vieira (2008), a taxa de variação do crescimento populacional é utilizada como uma *proxy* do crescimento econômico.

Algumas variáveis explicativas como a expectativa de vida, a mortalidade infantil, a taxa de homicídios e a renda *per capita*, foram escolhidas com base nos trabalhos da NGE. Para verificar os efeitos que a educação tem sobre o crescimento foram incorporadas as variáveis anos de estudo e taxa de analfabetismo. Com relação a infraestrutura, o percentual de casas com energia elétrica e com água encanada foi selecionado para captar seus efeitos.

O índice de Theil foi adicionado para captar os efeitos da desigualdade de renda. Este índice varia entre -1 e 1 e quanto mais perto de 1 estiver, pior é a desigualdade de renda do município. Também foi escolhido o percentual do PIB devido à indústria para verificar o efeito desse setor sobre o crescimento. Já a variável percentual do emprego no setor urbano pretende captar os efeitos da aglomeração.

Como já foi analisado por Vieira (2008), para captar os efeitos da concentração de pessoas será utilizado o logaritmo da população. Também será usado o logaritmo da população ao quadrado para identificar os efeitos negativos da concentração de pessoas. E por fim será utilizada a variável distância dos municípios à capital para eliminar essa possível influência.

Na tabela 1 é feita a média, o desvio padrão e os valores máximo e mínimo para cada variável. Alguns municípios tiveram variação da população negativa de até 1,95% no período de 2000 a 2010, enquanto alguns municípios tiveram uma variação positiva de 18,19%. Como a média ficou em torno de 1,16%, o que se pode concluir é que poucos municípios tiveram uma taxa de crescimento muito alta como essa apresentada pelo valor máximo. Uma variável que chama atenção é a taxa de homicídios, que em alguns municípios é praticamente nula e em outros, principalmente na região metropolitana de São Paulo, podem chegar a quase 123 homicídios para cada 100.000 habitantes. A renda *per capita* também possui grandes disparidades entre os municípios, com valores mínimos chegando a 107,89 e valores máximos próximos de 954,65 medidos em salários mínimos.

Tabela 1  
Variáveis municipais no ano de 2000 e variação da população entre 2000 e 2010

Variáveis	Média	Desvio Padrão	Mínimo	Máximo
Varição da população 2000-2010 <sup>1</sup> (em %)	1.16	1.55	-1.95	18.19
Esperança de vida ao nascer	71.66	2.58	61.00	78.18
Mortalidade infantil (por mil nascidos vivos)	15.45	5.06	5.38	42.35
Taxa de homicídios (a cada 100.000 habitantes)	12.91	18.90	0.00	122.95
Escolaridade da população com mais de 25 anos (em anos de estudo)	5.29	0.88	3.30	8.76
Taxa de analfabetismo da população com mais de 15 anos de idade (em %)	11.18	3.55	2.94	21.77
% de casas com energia elétrica	98.95	2.40	71.97	100.00
% de casas com água encanada	96.37	4.42	66.24	100.00
Renda per capita (em salários mínimos)	276.73	90.87	107.89	954.65
Índice de Theil	0.47	0.09	0.29	1.06
% do PIB devido à indústria	21.91	12.25	5.23	91.74
% do emprego no setor urbano	80.36	15.89	20.45	100.00

Fonte: IPEA

Na tabela 2 estão demonstradas as correlações entre as variáveis. Pode-se notar que a correlação positiva entre renda *per capita* e escolaridade é bem alta chegando a 0,85. A renda *per capita* também tem uma relação positiva com a porcentagem de emprego urbano de 0,91 e esta última tem uma correlação positiva com a escolaridade de 0,81. As variáveis domicílios com energia elétrica e casas com água encanada também possuem uma forte correlação positiva de 0,75, o que já era esperado uma vez que no estado de São Paulo são poucos os municípios que não tem acesso a ambos.

Por outro lado, a renda per capita tem relação negativa de -0,74 com a taxa de analfabetismo, o que indica que quanto menor o número de analfabetos em um município, maior é sua renda per capita. Ao contrário dos dados encontrados em Vieira (2008), mas em conformidade com os de Glaeser et al. (1995), a variação populacional e a renda per capita possuem correlação negativa com o percentual da indústria no PIB municipal, com valores respectivamente de -0,10 e -0,75. A escolaridade possui uma forte relação negativa com a porcentagem do setor industrial no PIB total de -0,94, enquanto que uma alta correlação negativa já esperada entre esperança de vida e mortalidade infantil foi corroborada com um valor de -0,99.

<sup>1</sup> Variação média da população no período de 2000 a 2010

Tabela 2: Matriz de correlação das variáveis municipais

	CR00-10	ESPVID	MORTINF	HOMIC	ESC	ANALF	ILUM	ÁGUA	RENPC	THEIL	%IND	%URB	DISTCAP
CR00-10	1,00	0,08	-0,09	0,17	0,12	0,23	0,13	0,10	0,19	-0,09	-0,10	0,16	-0,27
ESPVID	0,08	1,00	-0,99	0,13	0,29	0,33	0,45	0,52	0,35	-0,25	-0,17	0,27	0,06
MORTINF	-0,09	-0,99	1,00	0,12	0,30	0,34	-0,52	0,57	-0,34	0,26	0,17	0,26	-0,06
HOMIC	0,17	-0,13	0,12	1,00	0,30	0,27	-0,07	0,07	0,21	0,14	-0,34	0,22	-0,42
ESC	0,12	0,29	-0,30	0,30	1,00	0,85	0,32	0,38	0,85	0,26	-0,94	0,81	-0,35
ANALF	-0,23	-0,33	0,34	0,27	0,85	1,00	-0,37	0,39	-0,74	-0,09	0,81	0,72	0,54
ILUM	0,13	0,45	-0,52	0,07	0,32	0,37	1,00	0,75	0,32	-0,29	-0,13	0,22	0,06
AGUA	0,10	0,52	-0,57	0,07	0,38	0,39	0,75	1,00	0,37	-0,31	-0,11	0,27	0,08
RENPC	0,19	0,35	-0,34	0,21	0,85	0,74	0,32	0,37	1,00	0,35	-0,75	0,91	-0,37
THEIL	-0,09	-0,25	0,26	0,14	0,26	0,09	-0,29	0,31	0,35	1,00	-0,36	0,24	-0,07
%IND	-0,10	-0,17	0,17	0,34	0,94	0,81	-0,13	0,11	-0,75	-0,36	1,00	0,70	0,39
%URB	0,16	0,27	-0,26	0,22	0,81	0,72	0,22	0,27	0,91	0,24	-0,70	1,00	-0,40
DISTCAP	-0,27	0,06	-0,06	0,42	0,35	0,54	0,06	0,08	-0,37	-0,07	0,39	0,40	1,00

Fonte: Elaborada pela autora

A análise do índice de Moran (I) indica autocorrelação entre o crescimento populacional dos municípios do estado de São Paulo de 0,1721 para uma matriz de pesos espaciais do tipo Rainha. Os dados estão distribuídos principalmente entre o primeiro e o segundo quadrante, o que demonstra que os dados das cidades possuem, em sua grande maioria, autocorrelação positiva.

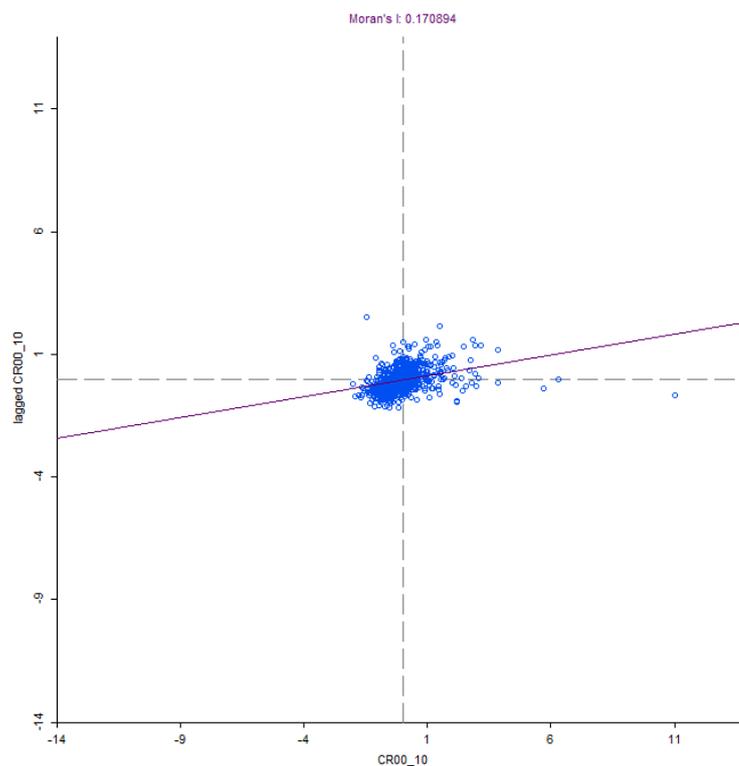


Figura 3: Índice de Moran (I)

Nos mapas das figuras 4 e 5 que mostram a distribuição da população dos municípios paulistas é possível notar uma maior concentração de pessoas na região leste do estado. Como o período analisado é relativamente curto, não é possível notar grandes mudanças na concentração populacional entre os anos em questão.

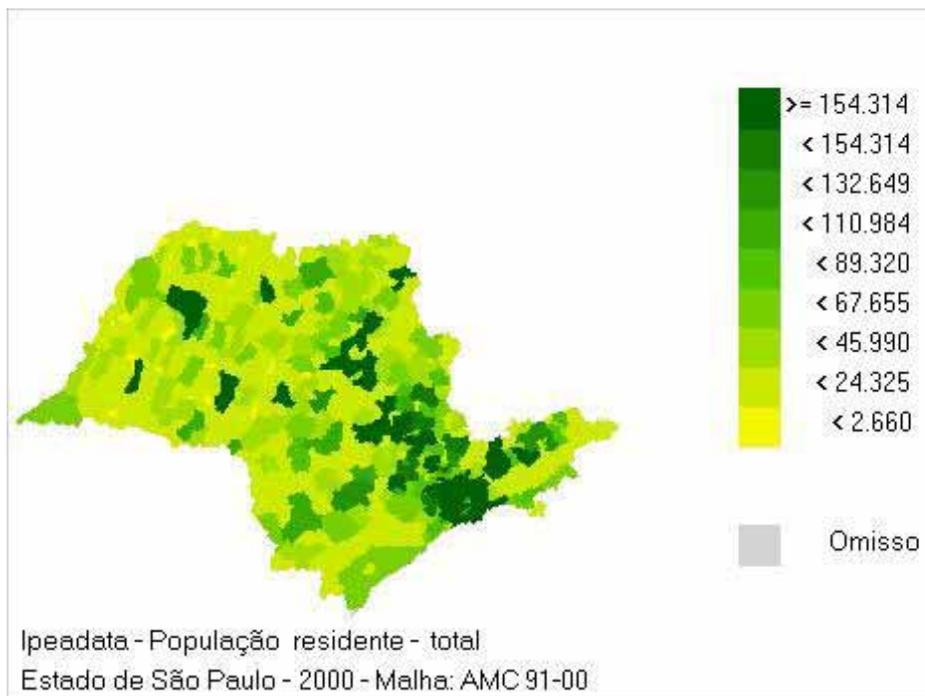


Figura 4: Distribuição espacial da população dos municípios paulistas nos anos 2000

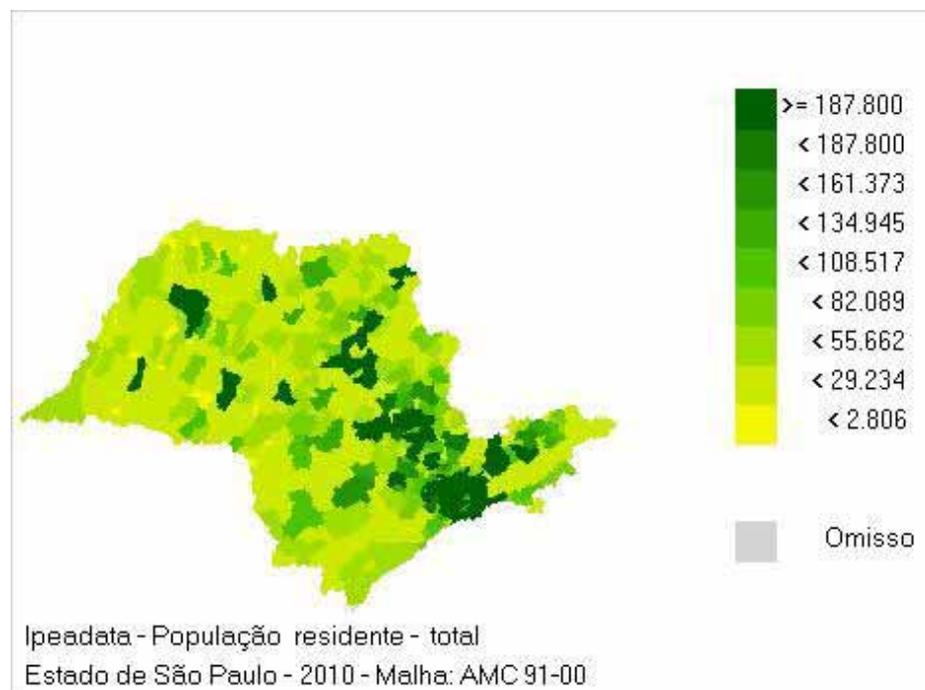


Figura 5: Distribuição espacial da população dos municípios paulistas no ano de 2010

Para verificar de forma adequada a possibilidade de existência de dependência espacial do crescimento dos municípios é necessário a Análise Exploratória de Dados Espaciais (AEDE). A figura 6 mostra a distribuição do crescimento populacional no estado paulista no período de 2000 a 2010. Neste caso também é possível notar que as cidades que mais cresceram estão localizada na área centro-norte do estado, enquanto que o sul concentra a maioria das cidades que cresceram pouco.

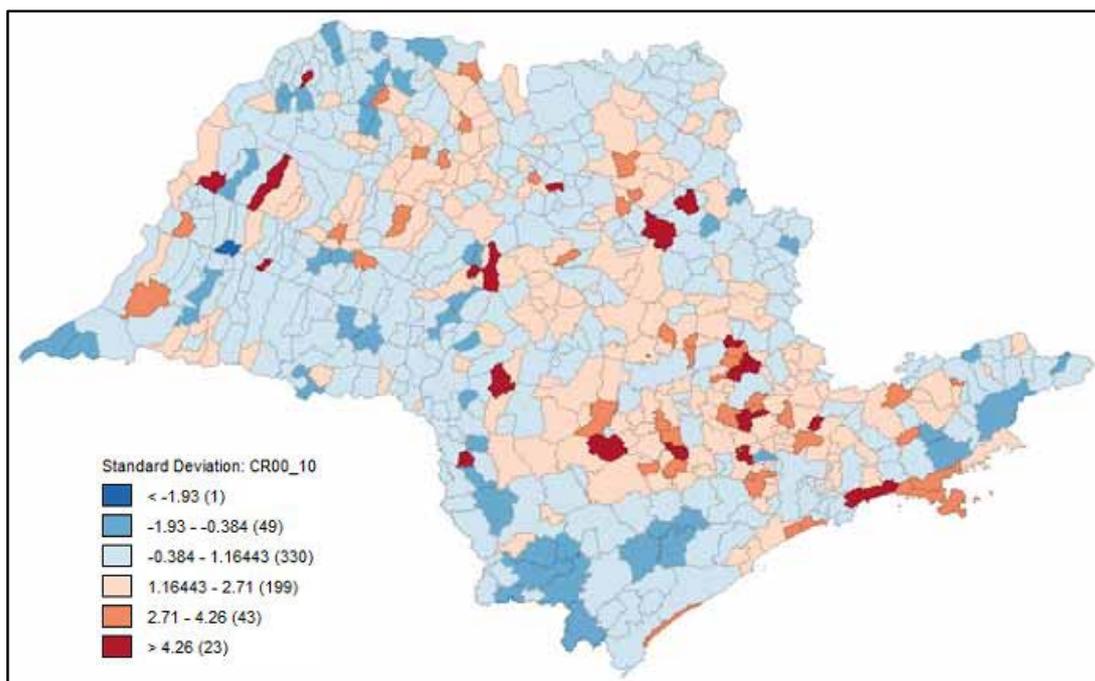


Figura 6: Distribuição espacial da taxa de crescimento populacional média dos municípios paulistas (período de 2000 a 2010)

A metodologia LISA presente no mapa da figura 7 permite a identificação de possíveis *clusters*, ou seja, se há a concentração de cidades com o mesmo padrão de crescimento ou se cidades com alto padrão estão próximas de cidades com baixo padrão de crescimento e vice-versa. No caso dos dados analisados neste estudo, a concentração de cidades com alta taxa de crescimento ocorre na região de Campinas, enquanto que *clusters* de taxas de crescimento baixas ou negativas ocorrem no Sul e à Noroeste do estado.

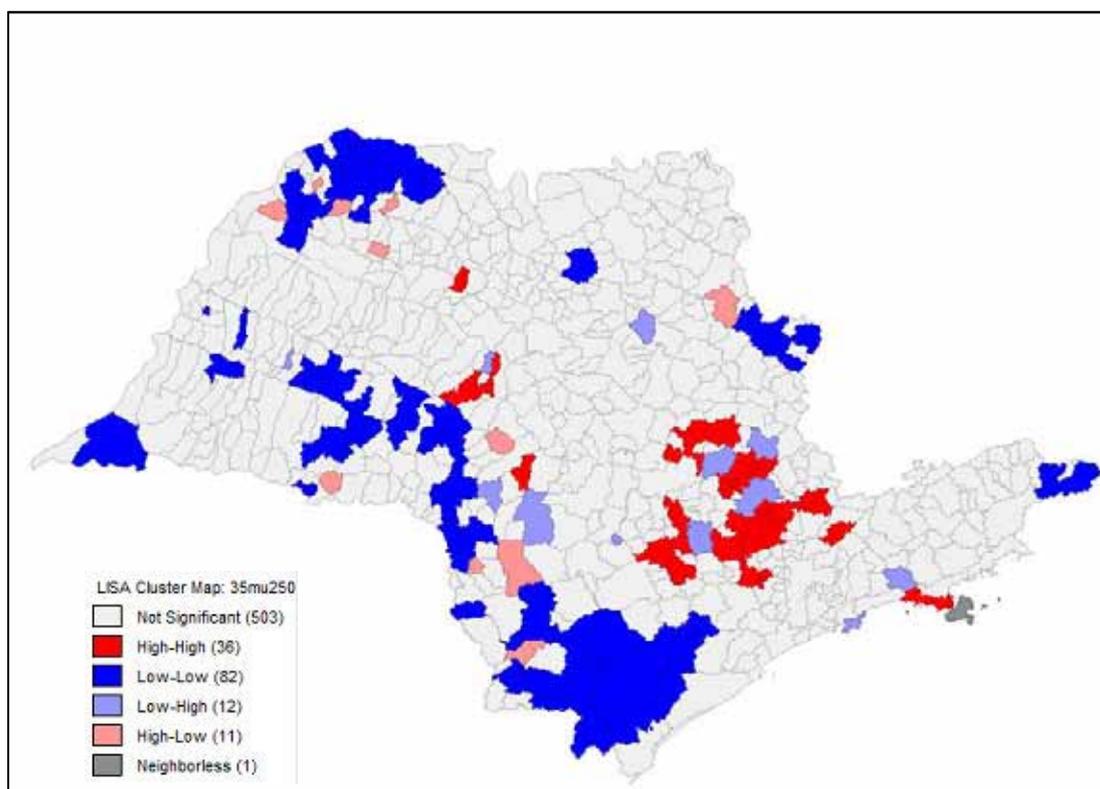


Figura 7: Mapeamento dos resultados da metodologia LISA

Os mapas analisados mostram que em certas regiões do estado, cidades próximas possuem o mesmo padrão de crescimento, o que foi corroborado pela metodologia LISA. Apesar do índice de Moran (I) não ser muito alto, uma estimativa dos dados analisados é importante para verificar se realmente existe uma dependência espacial e, caso exista, encontrar o modelo espacial que mais se adéqua ao padrão espacial apresentado por esses dados.

## 4. RESULTADOS

### 4.1. Modelo Empírico

Por meio dos métodos utilizados no capítulo anterior pode-se notar a possível relação espacial entre a variação populacional dos municípios paulistas. A partir dessa evidência será feita a estimação do modelo que mais se adéqua aos dados, adicionando variáveis que podem ter influência sobre o crescimento populacional. Essas variáveis foram descritas no capítulo anterior e são compostas por dados relacionados à renda, escolaridade, segurança pública, infraestrutura e emprego. Além disso, como indica Silva Junior (2007), é necessária a inclusão de variáveis como o logaritmo da população, para captar o surgimento de economia de aglomeração, e o logaritmo da população ao quadrado, pois a partir de certo ponto, essa aglomeração pode trazer malefícios à população.

O modelo utilizado será o mesmo de crescimento proposto por Glaeser et al. (1995) acrescido de parâmetros espaciais:

$$\begin{aligned} CR00 - 10 &= \rho W_1 CR00 - 10 + X'_{i,2000} \beta + \theta W_2 X'_{i,2000} + \mu \\ \mu &= \lambda W_3 \mu + \varepsilon \end{aligned} \quad (4.1)$$

$\varepsilon \sim N(0, \sigma^2 I_n)$

Onde CR00-10 é um vetor (645x1) que contém a taxa de crescimento dos municípios paulistas,  $X'_{i,2000}$  é um vetor (645xk) que contém as variáveis explicativas do modelo e uma coluna de 1s que corresponde ao termo do intercepto.  $W_1, W_2$  e  $W_3$  são matrizes de contiguidade normalizadas pelas linhas, já  $\mu$  representa o vetor com termos aleatórios, enquanto que  $\varepsilon$  é o termo do erro não correlacionado e  $\rho, \theta$  e  $\lambda$  são parâmetros. Este é o modelo completo, ou seja, possui dependência espacial na variável dependente, na variável explicativa e no termo de erro.

Neste trabalho será utilizada apenas a matriz de pesos espaciais  $W$  binária padrão, com  $w_{ij} = 1$ , caso os municípios  $i$  e  $j$  sejam vizinhos de fronteira e  $w_{ij} = 0$ , caso contrário. A matriz é normalizada dividindo-se cada elemento da linha pelo número de elementos diferente de zero.

### 4.2. Análise dos resultados econométricos

Na tabela 3 estão representados os resultados da estimação feita por MQO, utilizando todas as variáveis propostas anteriormente. Foram feitos testes da estatística LM para saber qual modelo é mais adequado a esses dados.

Tabela 3

Resultado da Estimação por MQO do modelo utilizando a matriz rainha			
Variável dependente: CR00-10			
Variáveis	MQO	R <sup>2</sup>	0.212
Constante	38.858** (17.071)	R <sup>2</sup> ajustado	0.193
Área municipal	0.0001 (0.0002)	Critério de Schwarz	2343.510
Logaritmo da população	-0.819 (1.080)	Estatística F	11.294***
Logaritmo da população ao quadrado	0.008 (0.123)	Log Likelihood	-1120
Esperança de vida	-0.363** (0.175)	LM-LAG	13.636
Mortalidade Infantil	-0.177* (0.093)	LM-LE	5.988
Homicídios	0.013*** (0.0037)	LM-ERR	9.214
Anos de escolaridade	0.815*** (0.180)	LM-EL	1.565
Percentual de analfabetismo	0.123*** (0.039)		
Percentual de casas com energia elétrica	0.0050 (0.040)		
Percentual de casas com água encanada	-0.034 (0.022)		
Renda <i>per capita</i>	0.007*** (0.001)		
Índice de Theil	-2.020** (0.850)		
Participação do setor industrial no PIB	2.257*** (0.485)		
Participação do emprego urbano	1.788*** (0.519)		
Distância à capital do estado	-0.001** (0.0004)		

Notas: (1) Desvio padrão dos parâmetros entre parênteses; (2)\* significativo a 10%, \*\* significativo a 5%, \*\*\* significativo a 1%.

Analisando os resultados podemos ver que a maioria das variáveis se mostrou significativa para o modelo, assim como a estatística F que foi significativa a 1%. O R<sup>2</sup> foi de 0,212 o que indica que 21,2% das variáveis são explicadas pelo modelo. Já o teste LM foi significativo tanto para o teste LM-LAG quanto para o LM-ERR, o que indica a

necessidade analisar o teste LM robusto. Esse teste robusto mostra que LM-LE é significativo, o que não ocorre com LM-EL, sendo assim, o modelo mais adequado a este conjunto de dados seria um com defasagem espacial (SAR).

Desta forma, na tabela 4 será estimado um modelo SAR, novamente com todas as variáveis propostas para o modelo na tabela anterior.

Tabela 4

Resultado da Estimação para um modelo SAR

Variável dependente: CR00-10

Variáveis	SAR	R <sup>2</sup>	0.2368
W_CR00-10	0.227*** (0.055)	Critério de Schwarz	2335.71
Constante	37.560** (16.594)	Log Likelihood	-1112.9
Área municipal	0.0001 (0.0002)		
Logaritmo da população	-0.967 (1.050)		
Logaritmo da população ao quadrado	0.019 (0.120)		
Esperança de vida	-0.345** (0.170)		
Mortalidade Infantil	-0.168* (0.090)		
Homicídios	0.013*** (0.004)		
Anos de escolaridade	-0.740*** (0.176)		
Percentual de analfabetismo	-0.113*** (0.038)		
Percentual de casas com energia elétrica	-0.001 (0.039)		
Percentual de casas com água encanada	-0.034 (0.022)		
Renda <i>per capita</i>	0.006*** (0.001)		
Índice de Theil	-1.531* (0.830)		
Participação do setor industrial no PIB	2.178*** (0.472)		
Participação do emprego urbano	1.885*** (0.504)		
Distância à capital do estado	-0.0009*		

(0.0005)

Notas: (1) Desvio padrão dos parâmetros entre parênteses; (2) \* significativo a 10%, \*\* significativo a 5%, \*\*\* significativo a 1%.

Nesta análise é possível notar que o  $R^2$  aumentou em relação à tabela 3 passando a ser de 0,2368, enquanto que o critério de Schwarz diminuiu para 2335,71, o que indica que este modelo se adequa melhor aos dados. Com relação às variáveis, algumas ainda não se mostraram significativas.

Procurando melhorar a relação entre o modelo e os dados, na tabela 5 será excluída a variável logaritmo da população ao quadrado.

Tabela 5

Resultado da Estimação para um modelo SAR			
Variável dependente: CR00-10			
Variáveis	SAR	R <sup>2</sup>	0.2367
W_CR00-10	0.227*** (0.055)	Critério de Schwarz	2329.26
Constante	36.631** (15.495)	Log Likelihood	-1112.9
Área municipal	0.0001 (0.0002)		
Logaritmo da população	-0.804*** (0.154)		
Esperança de vida	-0.339** (0.167)		
Mortalidade Infantil	-0.165* (0.088)		
Homicídios	0.013*** (0.003)		
Anos de escolaridade	-0.736*** (0.174)		
Percentual de analfabetismo	-0.112*** (0.037)		
Percentual de casas com energia elétrica	-0.0003 (0.038)		
Percentual de casas com água encanada	-0.034 (0.022)		
Renda <i>per capita</i>	0.006*** (0.001)		
Índice de Theil	-1.548* (0.823)		
Participação do setor industrial no	2.175***		

PIB	(0.471)
Participação do emprego urbano	1.870***
	(0.495)
Distância à capital do estado	-0.0009*
	(0.0005)

Notas: (1) Desvio padrão dos parâmetros entre parênteses; (2)\* significativo a 10%, \*\* significativo a 5%, \*\*\* significativo a 1%.

De acordo com os resultados o  $R^2$  diminuiu um pouco para 0,2367, o que não é uma diminuição relevante e o critério de Schwarz diminuiu mais uma vez para 2329,26 o que indica novamente a melhora do modelo em relação aos dados. A exclusão da variável logaritmo da população ao quadrado também foi benéfica para o modelo, pois permitiu que a variável logaritmo da população passasse a ser significativa.

Ainda assim, medidas para melhorar os resultados podem ser tomadas, e a exclusão de mais uma variável se faz necessária, portanto, na tabela 6 será excluída a variável percentual de casas com energia elétrica que pode ter alta correlação com a variável percentual de casas com água encanada.

Tabela 6

Resultado da Estimação para um modelo SAR			
Variável dependente: CR00-10			
Variáveis	SAR	R <sup>2</sup>	0.2367
W_CR00-10	0.227***	Critério de Schwarz	2322.79
	(0.055)	Log Likelihood	-1112.9
Constante	36.571***		
	(13.467)		
Área municipal	0.0001		
	(0.0002)		
Logaritmo da população	-0.804***		
	(0.154)		
Esperança de vida	-0.339**		
	(0.156)		
Mortalidade Infantil	-0.165**		
	(0.082)		
Homicídios	0.013***		
	(0.003)		
Anos de escolaridade	-0.735***		
	(0.172)		
Percentual de analfabetismo	-0.111***		
	(0.036)		
Percentual de casas com água encanada	-0.034*		
	(0.020)		

Renda <i>per capita</i>	0.006*** (0.001)
Índice de Theil	-1.548* (0.822)
Participação do setor industrial no PIB	2.175*** (0.471)
Participação do emprego urbano	1.869*** (0.490)
Distância à capital do estado	-0.0009* (0.0004)

Notas: (1) Desvio padrão dos parâmetros entre parênteses; (2)\* significativo a 10%, \*\* significativo a 5%, \*\*\* significativo a 1%.

Os resultados da tabela em análise mostram que o R<sup>2</sup> permaneceu constante em relação à tabela 5, mas o critério de Schwarz novamente diminuiu, indicando que esse novo modelo é melhor que o anterior. Os parâmetros mostraram-se todos significativos exceto no tangente à área municipal. Em virtude disso, na tabela 7 a estimativa será feita sem essa variável, a fim de melhorar ainda mais o modelo em questão.

Tabela 7

Resultado da Estimação para um modelo SAR			
Variável dependente: CR00-10			
Variáveis	SAR	R <sup>2</sup>	0.2366
W_CR00-10	0.226*** (0.055)	Critério de Schwarz	2316.42
Constante	36.393*** (13.455)		
Logaritmo da população	-0.784*** (0.140)	Log Likelihood	-1112.9
Esperança de vida	-0.337** (0.156)		
Mortalidade Infantil	-0.164** (0.082)		
Homicídios	0.013*** (0.003)		
Anos de escolaridade	-0.731*** (0.172)		
Percentual de analfabetismo	-0.112*** (0.036)		
Percentual de casas com água encanada	-0.034* (0.020)		
Renda <i>per capita</i>	0.006*** (0.001)		

Índice de Theil	-1.471*	(0.783)
Participação do setor industrial no PIB	2.164***	(0.470)
Participação do emprego urbano	1.847***	(0.485)
Distância à capital do estado	-0.0008*	(0.0004)

Notas: (1) Desvio padrão dos parâmetros entre parênteses; (2) \* significativo a 10%, \*\* significativo a 5%, \*\*\* significativo a 1%.

Nesta última tabela o  $R^2$  voltou a diminuir e seu novo valor é 0,2366, mais uma vez uma diminuição muito pequena, que não chega a ser relevante. Já o critério de Schwarz diminuiu para 2316,42 indicando que o modelo está mais adequado aos dados. Com a mudança feita da tabela 6 para a 7, todos os parâmetros passaram a ser significativos, o que indica que o modelo adequado foi encontrado.

A partir do momento que o modelo ideal é encontrado, torna-se possível a análise dos parâmetros. O logaritmo da população além de muito significativo possui um sinal negativo, o que indica que quanto maior a população de um município menos ele tende a crescer. Isso pode indicar que municípios grandes podem, devido à saturação da influência da aglomeração, estar gerando efeitos negativos sobre a população, o que pode levar muitos indivíduos a migrarem para cidades menores.

A taxa de homicídios tem um sinal positivo, mas isso não indica, necessariamente, que essa variável tem influência sobre o crescimento das cidades. A ideia mais plausível seria a de que cidades maiores, devido a vários fatores como a aglomeração de pessoas, por exemplo, apresentam taxas de homicídio maiores também. A variável esperança de vida com o sinal negativo deve ser analisada com a mesma linha de raciocínio que a anterior, ou seja, ela não estimula negativamente o crescimento, mas sim possui um valor menor para cidades que cresceram mais devido à influência do possível aumento da criminalidade e da poluição, por exemplo.

O percentual total de participação do emprego urbano sinalizado positivamente indica que cidades com maior número de pessoas trabalhando na área urbana têm maior taxa de crescimento, o que demonstra os benefícios da aglomeração. Outra variável com sinal positivo é a participação do setor industrial no PIB das cidades, o que demonstra a força da indústria auxiliando o crescimento das cidades.

Analisando as variáveis ligadas à educação, é possível notar que os anos de estudo e a taxa de analfabetismo têm sinais negativos, o que pode tornar difícil a interpretação desses parâmetros, pois ambos influenciam negativamente o crescimento. Um exemplo dessa dificuldade de interpretação foi demonstrado por Vieira (2008) quando propôs que pessoas com menor escolaridade vindas de outras regiões impulsionam o crescimento dos municípios, mas, como ele mesmo indica, o sinal negativo do analfabetismo induz a evidências contrárias.

É importante destacar também o sinal positivo da renda *per capita*, o que leva à conclusão de que quanto maior a renda da população mais o município tende a crescer. E a variável distância à capital sendo significativa demonstra que cidades próximas à capital tendem a crescer mais do que as cidades mais distantes.

#### **4.3.Comparação dos resultados com Vieira (2008)**

Este trabalho foi feito com o intuito de comparar os resultados da análise do crescimento econômico dos municípios paulistas no período de 2000 a 2010 com o trabalho de Vieira (2008) que fez uma análise similar para o crescimento dos municípios paulistas no período de 1980 a 2000.

Como no período de 1980 a 2000 houve um aumento no número de municípios no estado, Vieira optou por agrupar os territórios e manter o número de 571 municípios dos anos de 1980; o que torna o número de cidades analisadas diferente deste trabalho, qual seja, 645 localidades para os anos 2000. As variáveis explicativas utilizadas nos dois estudos são as mesmas diferenciando-se apenas no ano escolhido, no trabalho de Vieira, 1980 e neste, os anos 2000.

A primeira análise feita sobre a média, o desvio padrão, o valor máximo e o valor mínimo, demonstra que a maioria das variáveis tiveram uma notável melhora em seus índices entre esses dois períodos, como por exemplo a diminuição da taxa de analfabetismo, cujo valor médio era de 21,96% em 1980 e caiu para 11,18% em 2000. Já a esperança de vida ao nascer passou de uma média de 59,4 anos em 1980 para 71,66 em 2000, e a mortalidade infantil, que tinha uma média de 58,19 mortes para cada mil nascidos vivos em 1980, diminuiu para 15,45 mortes.

Apesar dessa melhora, alguns indicadores demonstraram o efeito inverso, como a taxa de homicídios, cuja média em 1980 era de 6,33 mortes para cada 100 mil habitantes e passou, em 2000, a ter 12,91 mortes para cada 100 mil habitantes. Ou seja, apesar do

padrão de vida dos paulistas ter aumento nos últimos vinte anos, índices como a violência aumentaram também.

Nas tabelas de correlações os resultados foram muito parecidos, havendo divergência apenas na variável percentual do PIB relativo à indústria. No trabalho de Vieira, esta variável teve alta correlação positiva com a escolaridade e correlação negativa com a taxa de analfabetismo, enquanto, no presente estudo, essa relação mostrou-se invertida, ou seja, a variável em questão teve correlação negativa com a escolaridade e positiva com a taxa de analfabetismo. Além disso, essa variável mostrou relação negativa com o crescimento populacional no presente trabalho, enquanto que no de Vieira mostrou uma alta correlação positiva quando relacionada à variação do crescimento populacional.

O índice de Moran (I) encontrado no trabalho de Vieira foi de 0,4591 e indicava maior probabilidade de dependência espacial para este índice, diferentemente do resultado de 0,1709 presentemente encontrado.

Já a metodologia LISA, demonstrada por meio de mapas, teve um resultado um pouco diferenciado. Em ambos os estudos, a região que concentra municípios com alto crescimento próximo a municípios com alto crescimento (High-High) é a região de Campinas, mas no de Vieira esse *cluster* estende-se também até a região metropolitana de São Paulo, para algumas cidades do litoral e mais para o interior, em direção a São Carlos, o que não ocorre no presente estudo. Já com relação a cidades com baixo crescimento próximas a cidades com baixo crescimento (Low-Low), em Vieira elas concentram-se principalmente na região Noroeste do estado; neste trabalho também há uma concentração de cidades com baixo crescimento no Noroeste do estado, mas a maior concentração está no sul do estado, mais especificamente na região de cidades como Itapeva e Capão Bonito.

Para comparar os resultados econométricos dos estudos em questão, será analisada apenas a regressão utilizando a matriz de pesos espaciais tradicional (rainha) de Vieira (2008), uma vez que no presente trabalho foi utilizada apenas a matriz citada. Sendo assim, os resultados da regressão por MQO mostraram, no trabalho de Vieira, que existem dependência espacial nos dados e o modelo mais adequado de acordo com os testes LM é o de defasagem espacial (SAR), o que ocorreu de maneira similar neste trabalho, indicando também o modelo de defasagem espacial como o mais adequado para dos dados.

Com relação à análise dos parâmetros, algumas variáveis explicativas escolhidas anteriormente foram excluídas de um ou de outro modelo para que a regressão ficasse mais adequada. No trabalho de Vieira foram escolhidas as variáveis logaritmo da população, logaritmo da população ao quadrado, área municipal, renda *per capita*, anos de

escolaridade, taxa de analfabetismo, percentual de casas energia elétrica, taxa de homicídios, mortalidade infantil, participação do setor industrial no PIB, participação do emprego urbano e distância à capital estadual. Já no presente estudo foram utilizadas as mesmas variáveis excluindo: logaritmo da população ao quadrado, percentual de casas com energia elétrica e área municipal; foram incluídas no modelo as variáveis percentual de casas com água encanada, esperança de vida ao nascer e índice de Theil.

Comparando os resultados da tabela 7 deste trabalho com os resultados de Vieira é possível notar uma grande semelhança nos sinais dos parâmetros escolhidos por ambos, como, por exemplo, anos de escolaridade e a taxa de analfabetismo, que nos dois estudos apresentaram sinais negativos, em contrapartida, a participação do setor industrial no PIB e a participação do emprego urbano demonstraram, em ambos, sinais positivos. Isso pode indicar que a influência que essas variáveis exercem sobre o crescimento econômico seguiu o mesmo padrão nos últimos trinta anos.

Quanto ao logaritmo da população no trabalho de Vieira demonstrou sinal negativo enquanto que no presente trabalho o sinal foi positivo. A interpretação que pode ser feita com relação a essa diferença é a de que entre os anos de 1980 e 2000 os municípios com maior população foram os que mais cresceram, enquanto que no período de 2000 a 2010 essa relação mostrou-se invertida, uma vez que os municípios menores foram os que mais cresceram. A renda *per capita* não foi significativa em Vieira, enquanto que no trabalho presente além de significativa, tal renda demonstra afetar positivamente o crescimento.

Outras variáveis também seguiram o mesmo padrão nos dois estudos, como a taxa de homicídios, que apresentou sinal positivo, porém, como foi explicado anteriormente, essa positividade não significa que o aumento no número de homicídios influencia o crescimento dos municípios, mas que quanto maior o município, maior tende a ser a taxa de homicídios além de, como já dito por Vieira, tratar-se de um problema de endogeneidade. A variável distância à capital estadual mostrou-se igualmente significativa em ambos os estudos, indicando que cidades mais próximas a São Paulo tendem a crescer mais do que as mais distantes.

Analisando genericamente os dois trabalhos, é possível notar que o padrão de crescimento nos dois períodos analisados foram muito parecidos. As variáveis explicativas utilizadas em comum também tiveram, em quase sua totalidade, a mesma influência sobre o crescimento econômico nos dois períodos. Já os mapas mostraram que a região de maior crescimento permaneceu a mesma enquanto que a região de menor crescimento aumentou no período de 2000 a 2010.

## 5. CONCLUSÃO

O objetivo deste estudo era analisar o crescimento econômico dos municípios do estado de São Paulo no período de 2000 a 2010, seguindo a linha do *mainstream* e utilizando as ferramentas da econometria espacial. Para escolher as variáveis que poderiam influenciar o crescimento, foram utilizados os estudos do crescimento econômico endógeno e da Nova Geografia Econômica. A econometria espacial foi empregada com o intuito de verificar em que medida o crescimento de um município afetou o crescimento de outros municípios próximos.

Foram escolhidas variáveis populacionais, educacionais, relativas à renda, emprego, infraestrutura, indicadores sociais e à distância da capital estadual. Calculando o índice de Moran (I) foi possível notar que existia algum padrão espacial na variação populacional do período. Então, utilizando a metodologia LISA tornou-se clara a existência de *clusters* de crescimento. A região que possui mais cidades com alto crescimento foi a de Campinas, enquanto duas regiões concentraram cidades com baixo crescimento, quais sejam, a região Noroeste do estado e, na parte sul do estado, a região que compreende as cidades de Itapeva e Capão Bonito.

A análise da correlação entre as variáveis demonstrou que a renda *per capita* está altamente correlacionada com emprego no setor urbano. Outra variável que mostrou alta correlação com o emprego no setor urbano é a escolaridade. A participação do setor industrial no PIB demonstrou uma notável correlação negativa com a renda *per capita* e com a escolaridade o que indica que quanto maior o nível de renda e mais anos de estudo da população, menor a participação do setor industrial no PIB municipal.

Para fazer essa análise foram utilizados os dados das variáveis nos anos 2000 e a variação populacional como *proxy* para o crescimento dos municípios. Os resultados da análise econométrica foram satisfatórios e proporcionaram uma interpretação adequada do tema proposto. Ficou claro que existia uma dependência espacial e concluiu-se, por meio do teste LM, que o modelo mais adequado aos dados é o modelo de defasagem espacial (SAR), ou seja, a dependência espacial está na variável endógena.

O resultado da regressão para um modelo SAR mostrou que as variáveis logaritmo da população, esperança de vida, mortalidade infantil, taxa de homicídios, anos de escolaridade, percentual de analfabetismo, percentual de casas com água encanada, renda *per capita*, índice de Theil, participação do setor industrial no PIB, participação do emprego urbano e distância à capital estadual foram todas significativas.

A maioria das variáveis apresentou o sinal esperado, exceto as variáveis escolaridade e percentual de casas com água encanada que apresentaram sinal negativo, o que indica uma relação inversa ao crescimento. Isso torna difícil interpretá-las, pois esperase uma influência positiva dessas variáveis sobre o crescimento.

A outra variável ligada à educação, a taxa de analfabetismo, demonstrou uma influência negativa, ou seja, municípios com alto percentual de analfabetos tendem a crescer menos. A mortalidade infantil também apresentou sinal negativo, o que era esperado, visto que cidades maiores têm, teoricamente, mais condições de oferecer acesso à rede de saúde.

As variáveis taxa de homicídios e esperança de vida apresentaram, respectivamente, sinais positivo e negativo, o que pode causar certo estranhamento a princípio. Desta forma, uma interpretação plausível seria a de que cidades maiores apresentam a variável taxa de homicídios também maior e a variável esperança de vida um pouco menor devido a fatores ligados a endogeneidade.

A participação do setor industrial no PIB mostrou sinal positivo, o que demonstra a importância desse setor para o crescimento dos municípios. Já a participação do emprego urbano, ao demonstrar sinal positivo, indica que a aglomeração é benéfica para o crescimento econômico municipal. Porém, é sempre interessante lembrar que esse benefício proporcionado pela aglomeração tem um limite, que quando ultrapassado passa a trazer malefícios à população.

O logaritmo da população indicou sinal negativo, o que leva à interpretação de que quanto menor a cidade mais ela tendeu a crescer. Isso pode indicar que muitas cidades grandes chegaram ao ponto de saturação da aglomeração, o que poderia ter desestimulado o crescimento nessas cidades e impulsionado o crescimento em cidades menores. É óbvio que esse não foi o único estímulo para o crescimento das cidades pequenas em detrimento das cidades maiores, outras variáveis provavelmente tiveram também forte influência sobre esse fenômeno.

A variável distância à capital estadual demonstrou que cidades mais próximas da capital tiveram maior crescimento do que cidades localizadas mais distantes. Isso fica evidente ao utilizar o mapa da metodologia LISA que indicou que a maioria dos *clusters* de baixo crescimento estavam localizados distantes da capital. Mas também é importante salientar que esse mapa mostrou que o *cluster* de alto crescimento deslocou-se da região metropolitana de São Paulo para a região de Campinas.

A comparação dos resultados com o trabalho de Vieira (2008) mostrou que o crescimento dos municípios paulistas não sofreu grandes alterações entre os períodos de 1980 a 2000 e de 2000 a 2010. A maioria das variáveis utilizadas tiveram sinais parecidos em ambos, sendo que a maior diferença foi observada na variável logaritmo da população, que no trabalho de Vieira demonstrou sinal negativo e no presente trabalho sinal positivo. Isso indica que no período de 1980 a 2000 os municípios maiores foram os que mais cresceram enquanto que no período de 2000 a 2010 foram os municípios menores que mais cresceram.

Após a análise dos resultados é possível concluir que para incentivar o crescimento dos municípios é necessário que os governos locais incentivem a educação e promovam melhorias na infraestrutura e na saúde pública. Outro ponto claramente observado é que incentivos à indústria também auxiliam no crescimento municipal. Por fim, a aglomeração foi uma das variáveis que mostrou-se benéfica para o crescimento, contudo, é necessário salientar que essa aglomeração tem um limite de saturação, e que depois de ultrapassado esse limite a aglomeração começa a causar malefícios, o que pode explicar o alto crescimento de cidades menores no período analisado.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ALONSO, W. **Location and land use: toward a general theory of land rent.** Cambridge: Harvard University, 1964.
- ANSELIN, L. **Spatial econometrics: methods and models.** Kluwer Academic Publishers: Dordrecht, 1988.
- ARROW, K. (1962) Economic welfare and the allocation of resources for invention. *In*: LAMBERTON, D. (ed.) **Economics of Information and Knowledge.** Harmondsworth: Penguin Books, 1971.
- BARRO, R & SALA-I-MARTIN, X. **Economic Growth.** McGraw-Hill: New York, 1995.
- BERA, A. & YOON, M. Specification testing with locally misspecified alternatives, **Econometric Theory**, n. 9, p. 649-658, 1993.
- CASS, D. Optimum Growth in an Aggregative Model of Capital Accumulation. **The Review of Economic Studies**, v. 32, n. 3, p. 233-240, jun. 1965.
- CAVALCANTI, L R. M. T. Produção teórica em economia regional: uma proposta de sistematização. **Revista Brasileira de Estudos Regionais e Urbanos**, v. 2, n. 1, 2008.
- CRUZ, B.O. et al. **Economia regional e urbana: teorias e métodos com ênfase no Brasil.** Brasília: IPEA, 2011.
- CHRISTALLER, W. (1933) **Central places in southern Germany.** London: Prentice Hall, 1966.
- DIXIT, A. K. & STIGLITZ, J. E. Monopolistic competition and optimum product diversity. **American Economic Review**, vol. 67, p. 297-308, 1977.
- FERREIRA, P. C. & ELLERY JR., J. Crescimento econômico, retornos crescentes e concorrência monopolística, *Revista de Economia Política*, vol. 16, n. 2, p. 86-104, abr./jun. 1996.
- FUNDAÇÃO SISTEMA ESTADUAL DE ANÁLISE DE DADOS – SEADE. Disponível em: <<http://www.seade.gov.br>>. Acesso em: 12 de abril de 2013.
- GLAESER, E. L. SCHEINCKMAN, J. A. & SCHLEIFER, A. Economic growth in a cross-section of cities, **Journal of Monetary Economics**, v. 36, p. 117-143, 1995.
- GOVERNO DO ESTADO DE SÃO PAULO. Secretaria de Economia e Planejamento. **Caracterização Regional do Estado de São Paulo a Partir dos Dados da RAIS 2008.** Disponível em: <[http://www.planejamento.sp.gov.br/noti\\_anexo/files/uam/trabalhos/RAIS-Estado.pdf](http://www.planejamento.sp.gov.br/noti_anexo/files/uam/trabalhos/RAIS-Estado.pdf)>. Acesso em: 27 de maio de 2013.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA – IPEADATA. Dados macroeconômicos e regionais. Disponível em: <<http://www.ipeadata.gov.br>>. Acesso em: 11 de abril de 2013.

JACOBS, J. **The Economy of Cities**. New York: Vintage, 1969.

KNIGHT, F. Diminishing Returns to Investment. **Journal of Political Economy**, n.12, p. 34-36, jun. 1944.

KOOPMANS, T. C. On concept of optimal economic growth. In: **The Economic Approach to Development Planning**. Amsterdan: North Holland, 1965.

KRUGMAN, P. Increasing returns and economic geography, **Journal of Political Economy**, v. 99, n. 3, p. 483-499, 1991.

LESAGE, J. P. (1999) **Spatial Econometrics**. Manuscrito não publicado disponível em <<http://rri.wvu.edu/regscweb.htm>>. Acesso em: 16 de maio de 2013

LÖSCH, A. **The Economics of Location**. New Haven, Conn: Yale University, 1954.

LUCAS, R. On the mechanics economic development, **Journal of Monetary Economics**, v.12, p. 3-42, 1988.

MALTHUS, T. R. **An Essay on the principle of population**, as it affects the future improvement of society, with remarks on the speculations of Mrs. Godwin, M. Condorcet and others Writers. J. Johnson, V IX, London, U.K, 1798.

MARQUES, H. A nova geografia econômica na perspectiva de Krugman: uma aplicação às regiões europeias. **Working Papers Centro de Estudos da União Europeia**, n. 7, Universidade de Coimbra, 2001. Disponível em: <[http://www4.fe.uc.pt/ceue/working\\_papers/ihelena.pdf](http://www4.fe.uc.pt/ceue/working_papers/ihelena.pdf)>. Acesso em: 4 de abril de 2013.

MARSHALL, A. (1920) **Princípios de Economia**. São Paulo: Abril Cultura, 1982.

OLIVEIRA, C. A. Externalidades espaciais e o crescimento econômico das cidades do Estado do Ceará, **Anais do X Encontro Regional de Economia do Nordeste**, Fortaleza, 2005.

RAMSEY, F.P. A mathematical theory of saving, **The Economic Journal**, v. 38, n. 152, p. 543-559, dec. 1928.

RESENDE, G. M. O crescimento econômico dos municípios mineiros: externalidades importam? **Cadernos BDMG**, n. 11, p. 1-96, Belo Horizonte, out. 2005.

RICARDO, D. (1817) **Princípios de Economia Política e Tributação**. São Paulo: Abril Cultura, 1982.

ROMER, P. Increasing Returns and Long-run Growth, **Journal of Political Economy**, v. 94, p. 1002-1037, 1986.

ROMER, P. Endogenous technological change. **Journal of Political Economy**, v. 98, n.5, p. S71-S102, 1990.

SCHUMPETER, J. (1911) **The Theory of Economic Development**. London: Oxford University Press, 1934

SHESHINSKI, E. Optimal accumulation with learning-by-doing. In: SHELL, K. (ed.) **Essays on the Theory of Optimal Economic Growth**. Cambridge; MIT Press, 1967.

SILVA, P.K. A evolução das modernas teorias do crescimento econômico. *A Economia em Revista*, v. 16, n. 1, p. 31-42, jun 2008.

SILVA JÚNIOR, D. **Aglomeração e desenvolvimento: evidências para municípios brasileiros**. Dissertação (mestrado), FEA/USP, São Paulo, 2007.

SMITH, A. (1776) **A Riqueza das Nações: investigação sobre sua natureza e suas causas**. (Trad. port.) São Paulo: Abril Cultura, 1983.

SOLOW, R. A contribution to the theory of economic growth. **The Quarterly Journal of Economics**, v. 70, n. 1, p. 65-94, feb. 1956.

VIEIRA, R. S. **Crescimento Econômico no Estado de São Paulo: Uma Análise Espacial**. Dissertação (mestrado), Universidade Estadual Paulista, Araraquara, 2008.

VON THÜNEN, J. H. (1826) **The Isolated State**. Oxford: Pergamon Press, 1966. Young, A. A. Increasing returns and economics progress. **The Economic Journal**, v. 38, p. 527-242, 1928.