

Convergência Condicional ou Clubes? Uma aplicação da econometria espacial para o setor agrícola na Amazônia Legal – 1996-2006

Resumo

Com base nos dados do Censo Agropecuário de 1995-96 e 2006, testou-se qual hipótese sobre convergência – se condicional ou clubes – melhor descreve a dinâmica de crescimento da produtividade agrícola entre 630 municípios da Amazônia Legal. Para tanto, incorporou-se ao modelo de Johnson e Takeyama (2003) a relação de dependência e heterogeneidade espacial, bem como elementos da Nova Geografia Econômica, como custo de transportes e aglomeração de unidades produtivas. Os resultados indicaram que não se pode descartar a hipótese de convergência clubes em favor da condicional. Os mesmos ainda indicaram a formação de seis clubes de convergência, embora somente quatro desses apresentem algum grau de dependência espacial positiva. Dois clubes ricos se destacam, embora estejam geograficamente distantes: o primeiro no extremo Norte (NO) e o segundo ao Sudoeste (SO) da Amazônia Legal. Em comum, ambos transmitem efeito transbordamento espacial pelo termo de erro, fator que tem estimulado a velocidade de convergência nesses clubes ao longo do período.

Palavra Chave: Convergência clube, produtividade agrícola, econometria espacial.

Abstract

On the basis of the data of the Farming Census of 1995-96 and 2006, which hypothesis on convergence - conditional or clubs was tested - better describes the dynamics of growth of the agricultural productivity enters 630 cities of the Amazon. For in such a way, it was become incorporated the model of Johnson and Takeyama (2003) the relation of dependence and space heterogeneities, as well as elements of New Economic Geography, as cost of transports and agglomeration of productive units. The results had indicated that if it cannot discard the convergence hypothesis clubs for the connective. The same ones had still indicated the formation of six clubs of convergence; even so four of these only present some degree of positive space dependence. Two rich clubs if detach, even so are geographically distant: first in the extremity the North (IN) and according to the Southwest (ONLY) of the Amazon. In common, both transmit effect space overflow for the error term, factor that has stimulated the speed of convergence in these clubs throughout the period.

Keyword: Convergence club, agricultural productivity, spatial econometrics

JEL Classification: C21, R11.

Área de Interesse: Área 10 - Economia Agrícola e do Meio Ambiente

Convergência Condicional ou Clubes? Uma aplicação da econometria espacial para o setor agrícola na Amazônia Legal – 1996-2006

1. Introdução

Arcar com explicações sobre a persistência de extrema pobreza em algumas regiões, em contraste com um ambiente provido com elevados níveis de renda e de crescimento sustentado, em outras é, certamente, uma das questões mais intrigante e controversa, por que se deparam adeptos da literatura do crescimento econômico regional. A economia regional difere da economia tradicional em muitos aspectos. Seu interesse maior está voltado à compreensão das relações entre agentes econômicos, em que o espaço geográfico assume papel central na análise. Barro *et al.* (1995) destacam algumas particularidades relacionadas à economia regional, como menor diferencial tecnológico, de preferências e de fatores institucionais. Segundo os autores, as regiões intranacionais têm Governo Central comum, logo, estruturas institucionais e sistemas legais similares. Dessa forma, aspectos legais, culturais, de lingüística e institucionais, tendem a não constituir barreiras à mobilidade dos fatores entre regiões.

Assim, negligenciar a hipótese de mobilidade de fatores (capital e mão-de-obra) talvez não seja tão agravante para os modelos de economia internacional ou para os determinantes do diferencial de crescimento entre países, do que para os modelos aplicados às ciências regionais. Até pouco tempo, a ciência regional era imaginada como um conjunto de métodos para programação de políticas, com modelos de natureza *ad hoc*, não conseguindo, portanto, se integrar a ciência tradicional. Foi exatamente nessa lacuna que surgiu uma nova concepção teórica chamada de Nova Geografia Econômica (NGE), a partir do trabalho de Krugman (1991). Os autores dessa corrente, entre eles, Fujita, Krugman e Venables (1999), argumentam que havendo mobilidade de fatores e retornos crescente, o diferencial da riqueza e pobreza entre regiões tende a se agravar, provocando aglomeração da atividade em uma região em detrimento das outras.

Neste contexto, o espaço tem implicações diretas sobre a localização das atividades produtivas e, em consequência, a aglomeração proporciona economias externas. De modo geral, as externalidades espaciais têm sido tratadas sob vários aspectos, no qual muitos desses estudos empíricos têm direcionado atenção para análise dos determinantes do crescimento regional, bem como sobre a análise do processo de convergência de renda. Entre esses, destacam-se os trabalhos de Mankiw *et al.* (1992), Durlauf e Johnson (1995), Galor (1996) e Johnson e Takeyama (2003).

Nesse sentido, diversas contribuições teóricas e empíricas têm revigorado a importância de se considerar o aspecto espacial como fator determinante do diferencial de crescimento entre regiões. Neste campo, incluem-se os modelos de crescimento neoclássico, os quais se apóiam sobre três hipóteses básicas de convergência, todas originárias de Solow (1956).

- A primeira, chamada de convergência absoluta (β -convergência), estabelece que as rendas *per capita* de um grupo de economias quaisquer, tendem incondicionalmente para o mesmo nível de longo prazo, o *steady-state*. A fragilidade desta versão consiste em supor que as economias experimentam dos mesmos padrões tecnológicos, preferências, de crescimento populacional, instituições e outros aspectos econômicos.
- A segunda sugere a noção de convergência condicional, em que um grupo de economias que possuem características comuns transitam para condição de equilíbrio de longo prazo, independentemente de suas condições iniciais.
- Por fim, sustentando a premissa de heterogeneidade espacial, diz-se que as economias somente transitaram para uma condição de equilíbrio, se dispuser de características estruturais comuns e de mesma condição inicial, gerando assim, múltiplos equilíbrios de longo prazo. É a partir dessa pressuposição que deriva o conceito de clube de convergência, em que as economias devem compartilhar de parâmetros estruturais semelhantes e baixa disparidade de renda inicial¹.

As três hipóteses podem ser testadas uma contra as outras, a fim de legitimar a prevalência da versão irrestrita, normalmente clubes, sobre as demais, consideradas restritas. No entanto, a forma metodológica com que os procedimentos são conduzidos é multivariada com resultado igualmente comum. Por exemplo, Durlauf e Johnson (1995) utilizam técnica não-paramétrica para classificação e identificação de clubes de convergência, chamada de *clustering*. Após a identificação a análise de significância segue por meio de *regression tree*. Noutro momento, utilizando a mesma técnica Johnson e Takeyama (2003) buscam determinar qual hipótese sobre convergência melhor descreve a dinâmica de renda dos estados norte-americanos, constatando a prevalência da hipótese clubes sobre as demais.

Bem verdade, que os testes realizados sobre convergência tem implicações bastante distintas quanto a forma e tratamento das hipóteses. Em muitos casos, como em Mankiw *et al.* (1992), a análise é conduzida por meio do modelo proposto por Barro (1991), em que evidências paramétricas de um coeficiente β negativo atestaria a convergência. Em outro turno, existem autores que defendem que a análise deva ser conduzida de forma intertemporal, com dados distribuídos em série de tempo ou em painel longo (QUAH, 1992; BERNARD and DURLAUF, 1994).

Vale ressaltar que várias críticas são direcionadas para modelos de convergência, sobretudo, se os dados forem obtidos em cortes transversais e indexados a unidades geográficas. Rey e Montoury (1999) advertem para a inconsistência dos resultados nos modelos de regressão do tipo de Barro (1991), que não incorporam na modelagem, a relação de dependência espacial, aspecto comum em economias com elevado grau de difusão tecnológica. Outra crítica padrão foi apresentada por Quah (1993); em que, segundo o autor, a obtenção de um coeficiente β negativo nada tem a dizer sobre convergência ou se o grau de dispersão de renda entre diversas regiões estaria diminuindo. Na realidade, indica apenas, numa dada amostra, uma tendência de ajuste dos dados em torno da média, supondo ainda, que os erros são bem comportados. Adverte ainda, para instabilidade estocástica do coeficiente β , ressaltando que teste sobre convergência baseado em Barro (1991), só seriam

¹ Além dessas versões existe a σ -convergência que analisa a dispersão intertemporal de renda *per capita* entre as economias. Para mais detalhes sobre hipóteses de convergência consulte Galor (1996).

validos se os determinantes do processo de crescimento fossem ‘estocasticamente’ invariantes ao longo do tempo. Porém, dados de renda e crescimento mostram que os parâmetros da referida regressão são variantes e obedecem a padrões estocásticos distintos (QUAH, 1993, p.02).

Vários são os estudos sobre convergência regional no Brasil, com destaque para os estudo de Ferreira e Diniz (1995) que detectaram uma reversão do processo de convergência renda regionais entre os anos de 1970 a 1995. Com dados disposto em larga série de tempo (1939-1995), Azzoni (2001) concluiu a favor de um processo de convergência entre 1970 a 1985, mas que passa a divergir posteriormente, no período de recuperação econômica. Artigos mais recentes, utilizam em sua metodologia instrumentos da econometria espacial, dentre eles destacam-se os estudos de Azzoni e Silveira Neto (2003), Mossi *et al.* (2003), Hewings *et al.* (2005). Outros incorporam explicitamente a mobilidade dos fatores capital e trabalho entre regiões, como por exemplo, Maciel, Andrade e Teles (2008).

Isto posto, pretende-se testar a hipótese de convergência aplicada sobre a produtividade agrícola regional. Mais especificamente, utilizando dados do Censo Agropecuário de 1995-96 e 2006 propõe-se testar qual hipótese sobre convergência – se condicional ou clubes – melhor descreve a dinâmica de crescimento da produtividade agrícola entre 630 municípios produtores estabelecidos na Amazônia Legal com o uso da econometria espacial e fundamentos da Nova Geografia Econômica. Ressalta-se a escolha do setor agrícola e da área de abrangência da Amazônia Legal como foco do trabalho devido sua importância na formação do PIB regional e nacional, bem como pela incipiência de estudos na área.

Inicialmente, serão construídos indicadores de autocorrelação espacial local (LISA), cujo objetivo é identificar regimes espaciais distintos dentro da amostra Amazônia Legal. Este procedimento permitirá a subdivisão amostral, gerando assim *clusters*, que segundo Anselin (1995), são bolsões de homogeneidade. Adicionalmente, serão feitos testes para diagnosticar a influência de efeitos decorrente da heterogeneidade e dependência espacial, inclusive testes de robustez². Com o resultado do diagnóstico, segue-se a etapa de seleção de modelos espaciais. Por fim, depois de efetivada a correção dos efeitos espaciais obtém-se a estimação dos parâmetros estruturais, seguido da análise dos resultados e validação (prognóstico) dos modelos selecionados. Espera-se que essas zonas de atração (*clusters*) sejam dotadas de características estruturais comuns dos fatores determinantes da convergência, com isso, atestar a prevalência da hipótese clubes sobre a condicional.

Além desta introdução, o estudo está dividido em outras quatro seções. A segunda aborda aspectos relacionados ao modelo teórico-empírico, baseado nos fundamentos da econometria espacial. A terceira seção destaca a construção das variáveis e fonte de dados. A quarta seção descreve os resultados dos testes de dependência espacial e análise dos resultados paramétricos dos modelos de convergência. Finalmente tem-se a conclusão, destacando os principais resultados encontrados.

2. Modelo Econômico - Empírico

Em estudos que visam investigar o processo de crescimento e distribuição de determinada variável-chave ao longo do espaço é comum fazer uso da análise β -convergência, seja ela estimada sob a forma incondicional, condicional ou clubes. Galor (1996) destaca que a diferença entre essas três hipóteses resume-se à relação de dependência e homogeneidade inerente às características estruturais, além da condição inicial. A hipótese de convergência condicional despreza o papel da condição inicial, mas requer que as economias tenham características estruturais comuns, para que atinjam suposta condição de múltiplos equilíbrios de longo prazo. Mas para a análise de clubes, a condição inicial somente tem importância, em grupos de economias com características semelhantes, dos determinantes da variável de interesse. Assim, a condição inicial atua como fator necessário, mas não suficiente, para o processo de convergência.

Segundo Johnson e Takeyama (2003), uma forma de testar à prevalência da hipótese clubes as demais é verificar a relevância da condição inicial, a influência de fatores condicionais, bem como a da proximidade espacial para a determinação dos níveis de produtividade de longo prazo, potencialmente observados entre grupos de economias com características comuns. Uma vez que o enfoque da investigação recai sobre as relações econômicas em nível regional (Amazônia Legal), não se pode deixar de considerar explicitamente questões importantes como a relação de dependência espacial, bem como a desigualdade quanto à distribuição dos níveis de produtividade média ao longo do espaço. Rey e Montoury (1999) argumentam que modelos β -convergência devem levar em consideração tais questões e estes devem ser interpretados como forma de expressão natural da interação entre os agentes econômicos.

Neste sentido, a dependência espacial, fenômeno observado entre grupos de economias vizinhas, reflete a idéia de que efeitos que derivam das relações econômicas estabelecidas entre produtores tendem a influir não somente noutros produtores instalados na mesma região (município), mas também naqueles instalados na vizinhança. Já a heterogeneidade espacial está relacionada com a não uniformidade quanto à distribuição da variável-chave ao longo do espaço, uma vez que é factível que existam diferenças entre combinações de fatores determinantes para a distribuição da produtividade agrícola³. Anselin e Bera (1998) destacam razões pelas quais se devem considerar a heterogeneidade espacial de forma explícita. Certamente, a principal delas está relacionada à falta de estabilidade dos parâmetros ao longo da estrutura espacial, o que significa que diferenças nas proporções de fatores que determinam a produtividade podem causar distorções na distribuição e, em consequência, podem ser gerados parâmetros variantes ao longo do espaço.

Formalmente, a relação de dependência espacial se expressa por meio da autocorrelação espacial e ocorre quando há correlação entre uma variável aleatória indexada em dada região i com seus vizinhos mais próximos j . Segundo Anselin

² Os testes seguem a lógica do estimador de Máxima Verossimilhança (LM), direcionados para correlação espacial nos erros (LM-error) e na variável dependente (LM-lag), além de suas versões robustas (LM-error* e LM-lag*). Ademais, utilizar-se-á o teste de Wald para atestar presença de heterocedasticidade nos resíduos. Para comparar a robustez dos resultados, todos os testes serão inferidos sobre duas estruturas espaciais distintas: uma construída com base na contigüidade dos municípios vizinhos (W_1) e outra construída com base no inverso da distancia euclidiana entre centróides (W_2).

³ A produtividade média agrícola depende em muitos aspectos de fatores externos como as características físicas do solo e seu grau de exaustão, a incidência de chuvas, a capacidade de gerenciamento do negócio, e outros fatores não observáveis. Esses aspectos são desigualmente distribuídos entre subespaços regionais, o que implica a heterogeneidade espacial.

(1988) a autocorrelação espacial sugere questionar o princípio da independência da distribuição de um conjunto de dados entre unidades. Ainda segundo o autor, a dependência espacial pode originar-se de problemas de mensuração, tal como a divisão da amostra em subgrupos ou ainda devido às externalidades espaciais. Já para Le Gallo e Erthur (2003) a dependência espacial pode ser entendida como a presença coincidente de similaridade entre valores mais próximos.

2.1 Modelo Convergência Clubes

A abordagem utilizada neste artigo considera testar separadamente a hipótese de convergência clubes contra a condicional. Para tanto, o modelo empírico adotado fundamenta-se na proposta de Durlauf e Johnson (1995) e Johnson e Takeyama (2003).

O modelo considera que o valor da produção agrícola (Y_{it}) de um dado município i no tempo t é determinada por uma função de produção Cobb-Douglas, que traz como argumentos, fatores como o grau de utilização do capital físico (K_{it}), o estoque de capital humano do pessoal ocupado na produção (H_{it}), o estoque de terra disponível para atividade (τT_{it}), com este último, potencializado por um coeficiente τ que exprime às características físicas e de fertilidade.

$$Y_{i,t} = A_0 K_{i,t}^{\alpha_1} H_{i,t}^{\alpha_2} (\tau T_{i,t})^{1-\alpha_1-\alpha_2}, \quad \alpha_1, \alpha_2 > 0 \quad (1)$$

O componente A_0 reflete o grau de produtividade média de outros fatores não incluídos no modelo, tais como fatores institucionais, fatores climáticos e outros.

A produtividade agrícola é uma medida de eficiência econômica, uma vez que expressa o nível de produto agrícola por unidade de terra requerida pela atividade. Com base nesta função de produção, pode-se obter a produtividade média agrícola dividindo ambos os lados da mesma pelo componente $\tau T_{i,t}$.

$$p_{i,t} = a_0 k_{i,t}^{\alpha_1} h_{i,t}^{\alpha_2}. \quad (2)$$

Assim, tem-se a produtividade média agrícola como função do estoque de capital físico e humano por unidade de terra eficiente. Os coeficientes α_1 e α_2 refletem as elasticidades parciais de seus respectivos fatores sobre a produtividade, sendo a_0 uma constante. A condição de convergência da produtividade somente será possível caso a função de produção apresente retornos decrescente de escala.

O modelo admite que uma fração fixa do produto é redirecionada na ampliação dos níveis de capital físico e humano (s^k e s^h), sendo que ambos se depreciam a uma taxa constante e equivalente a δ . Considerando que volume de trabalho é determinado exogenamente e cresce a taxa n_i e que o padrão tecnológico empregado na produção cresce a g_i , pode-se descrever a dinâmica de transição dos estoques de capital físico e humano *per capita* por meio das equações.

$$\begin{aligned} \dot{k}_{i,t} &= s^k p_{i,t} - (n_i + g_i + \delta)k_{i,t}; \\ \dot{h}_{i,t} &= s^h p_{i,t} - (n_i + g_i + \delta)h_{i,t}. \end{aligned} \quad (3)$$

observe que o sistema acima contempla a produtividade média agrícola em ambas as equações. Substituindo as equações do sistema em (2) obtém-se uma equação que contém somente as variáveis explicativas. Nesse sentido, equilíbrios de longo prazo somente serão possíveis caso as taxas de acumulação dos referidos estoques tendam a zero. Assim, com algumas manipulações algébricas a produtividade média agrícola de equilíbrio de longo prazo, sob forma logarítmica, é dada por.

$$\log p_{i,t}^* = \frac{\alpha_1}{1-\alpha_1-\alpha_2} \cdot \log \frac{s^k}{(n_i + g_i + \delta)} + \frac{\alpha_2}{1-\alpha_1-\alpha_2} \cdot \log \frac{s^h}{(n_i + g_i + \delta)} \quad (4)$$

Sendo $\log p_{i,t}^*$, o nível de produtividade agrícola de equilíbrio de longo prazo. Uma aproximação para o *steady-state* é possível desde que se diferenciem as equações do sistema (3) em relação a t , permitindo obter.

$$\frac{d \log p_{i,t}}{dt} \cong -\gamma (\log p_{i,t} - \log p_{i,t}^*), \quad \gamma = (1-\alpha_1-\alpha_2)(n_i + g_i + \delta) \quad (5)$$

Percebe-se que a variação contínua da produtividade no tempo, rumo ao *steady-state*, é ponderada pelo coeficiente γ , que expressa a velocidade de convergência. O coeficiente em epígrafe é resultante de combinações entre as elasticidades dos fatores mais a taxa de crescimento do pessoal ocupado, da tecnologia e a taxa de depreciação dos estoques. Além disso, retornos decrescentes restringiram γ ao intervalo [0,1]. Posto isto, a solução da equação (5) resulta em:

$$\log p_{i,t} = (1 - e^{-\gamma t}) \log p_{i,t}^* + e^{-\gamma t} \log p_{i,0} \quad (6)$$

Em que $\log p_{i,0}$ é o logaritmo produtividade média agrícola em determinado período inicial. Considerando qualquer intervalo de tempo, por exemplo, de 1996 a 2006, pode-se agora determinar a taxa de crescimento da produtividade média agrícola no período proposto. Para isso, deve-se subtrair o termo $\log p_{i,0}$ em ambos os lados de (6) e obter:

$$\log p_{i,t} - \log p_{i,0} = (1 - e^{-\gamma t}) \log p_{i,t}^* + (e^{-\gamma t} - 1) \log p_{i,0} \quad (7)$$

Finalmente, substituindo-se (4) em (7), chega-se a:

$$\Delta \log p_{i,t} = \alpha_i - \beta_i \log p_{i,0} + \phi_{1,i} s^k + \phi_{2,i} s^h - \phi_{3,i} (n_i + g_i + \delta), \quad \text{com} \quad (8)$$

$$\Delta \log p_{i,t} = \log p_{i,t} - \log p_{i,0}, \quad \alpha_i = a_0 (1 - e^{-\gamma t}), \quad \beta_i = (1 - e^{-\gamma t}), \quad \phi_{1,i} = \frac{\alpha_1 (1 - e^{-\gamma t})^{-1}}{(1 - \alpha_1 - \alpha_2)}, \quad \phi_{2,i} = \frac{\alpha_2 (1 - e^{-\gamma t})^{-1}}{(1 - \alpha_1 - \alpha_2)}, \quad \phi_{3,i} = \frac{(\alpha_1 + \alpha_2) (1 - e^{-\gamma t})^{-1}}{(1 - \alpha_1 - \alpha_2)}$$

A equação (7) expressa a taxa de crescimento da produtividade média agrícola em função da condição inicial mais o nível de produtividade de equilíbrio de longo prazo, sendo este último condicionado a fatores como a proporção do produto que é

reinvestida na ampliação do estoque de capital físico e humano (s^k e s^h), mais a taxa de crescimento do pessoal ocupado, a variação tecnológica imprimida na produção e taxa de depreciação dos fatores. Assim o modelo de econométrico será.

$$\Delta p_i = \alpha_i' \iota - \beta_i' p_0 + \phi_i' X + \varepsilon_i \quad (9)$$

Onde, Δp_i é um vetor ($N \times 1$) com a variável dependente, a taxa de crescimento da produtividade agrícola; p_0 é um vetor ($N \times 1$) com o logaritmo da produtividade média agrícola no ano de inicial; X é uma matriz ($N \times K$) de variáveis condicionais, sendo K o número de variáveis; ε_i é um termo de erro, cujas propriedades serão definidas mais adiante; ι é um vetor coluna unitário ($N \times 1$); por fim, α_i , β_i e ϕ_i são considerados parâmetros fixos, pelo menos entre grupos de economias com características comuns.

Autores como Mankiw *et al.* (1992), Durlauf e Johnson (1995) e recentemente Johnson e Takeyama (2003), destacam que cada uma das três hipóteses sobre convergência pode ser testada contra a versão restrita clubes. Caso se ignore a influência da heterogeneidade espacial na distribuição da produtividade média agrícola; mas, considerando a influência de fatores condicionais embutidos em X , permitindo que pelo menos um dos ϕ_i assumam valor diferente de zero. Então, evidências paramétricas de um sinal negativo para o coeficiente β e a significância de alguma variável de controle corroborariam com a hipótese de convergência condicional. Por outro lado, se relaxada tal restrição, com isso, assumindo a possibilidade de heterogeneidade espacial, é pertinente identificar subregiões dentro da amostra Amazônia Legal que experimentem de características estruturais comuns, no tocante à produção agrícola. Assim, o sinal negativo para o coeficiente β e a significância estatística de alguma variável de controle corroboraria com a hipótese de convergência clubes contra a condicional.

Assim, a questão requerida neste momento consiste em definir qual o critério adotado para identificar potenciais clubes. Como exemplo, Mora (2005) utilizou uma medida de desigualdade obtida a partir da decomposição do índice de Gini com objetivo de agrupar economias com características semelhantes. De outra forma, para investigar a hipótese de convergência clubes entre grupos de países no mundo, Durlauf e Johnson (1995) adotam o critério de particionamento da amostra com base na análise *regression tree*. O critério de agrupamento de subregiões adotado aqui considera explicitamente no modelo não só a influência da heterogeneidade, mas também a relação de dependência espacial, com base no índice de Moran local (ANSELIN, 1995). Este método permite identificar zonas de atração, consideradas bolsões de homogeneidade quanto à distribuição da produtividade média agrícola.

Antes de prosseguir para os critérios de seleção dos clubes de convergência é preciso definir medidas de ponderação de distância geográfica, as quais deverão ser aplicadas aos grupos de municípios vizinhos. Na literatura pertinente a econometria espacial (ANSELIN, 1988; ANSELIN e BERA, 1988; LeSAGE, 1999; ANSELIN, FLORAX e REY, 2004; e ARBIA, 2005) algumas medidas de ponderação espacial merecem destaque.

2.2 Peso Espacial

Pesos espaciais, medidas que ponderam distâncias geográficas, sugerem a seguinte premissa: “todo mundo influencia, porém quem está mais perto influencia mais” (ANSELIN, 1988). Pesos construídos a partir de distâncias geográficas são, em larga escala, os mais utilizados em trabalhos que se utilizam dos instrumentos da econometria espacial. A lógica consiste em estabelecer maior peso àquelas unidades mais próximas.

A matriz de peso espacial, também conhecida como matriz de proximidade espacial, é uma matriz quadrada de dimensão ($N \times N$), em que N é o número de unidades espaciais.

$$W = \begin{pmatrix} 0 & w_{12} & w_{13} & \dots & w_{1N} \\ w_{21} & 0 & w_{23} & \dots & w_{2N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ w_{N1} & w_{N2} & \dots & \dots & 0 \end{pmatrix} \quad (10)$$

Cada elemento w_{ij} estabelece uma relação de vizinhança entre determinado município i (na linha) com seus vizinhos j (na coluna). Assim, pesos espaciais podem ser definidos com base no seguinte critério.

$$w_{ij} = \begin{cases} 0, & \text{se } i = j; \\ 0, & \text{se } i \neq j, \text{ se } j \text{ não vizinho de } i; \\ 1, & \text{se } i \neq j, \text{ se } j \text{ vizinho de } i. \end{cases} \quad (11)$$

Tipicamente, há duas formas de se construir uma matriz de pesos espaciais. A primeira é construir pesos com base na contigüidade das fronteiras, como exposto em (11) e a segunda consiste em construir pesos com base na distância de centróides, calculando-os pelo inverso da distância euclidiana.

$$w_{ij} = \begin{cases} 0, & \text{se } i = j; \\ \frac{1}{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}, & \text{se } i \neq j. \end{cases} \quad (12)$$

Em que x_i , x_j , y_i e y_j são coordenadas de centróides das unidades i e j . Estas informações podem ser obtidas pela latitude e longitude geográfica das unidades.

4 Note o papel do subscrito i nos parâmetros α , β e ϕ do modelo, eles significam que esses devem ser fixos dentro de um mesmo clube, porém podem variar ao longo do espaço amostral (Amazônia Legal). Ou seja, sob hipótese clubes, é provável que o α , β e ϕ assumam valores distintos.

Usualmente, adota-se o mecanismo de normalização da matriz W em qualquer que seja o critério escolhido pelo pesquisador. A normalização garante que a soma dos pesos na linha resulte na unidade, $\sum_j w_{ij} = 1, i = 1, \dots, N$. Isto significa que após a normalização, cada elemento da matriz W assume valores equivalentes a $w_{ij}^* = w_{ij} / \sum_j w_{ij}$.

Anselin (1988) destaca que apesar de não haver restrições matemáticas para a normalização, este procedimento facilita a leitura dos coeficientes, uma vez que passam a capturar a influência relativa de um município i sobre seus vizinhos j . Ademais, tal procedimento assegura que o coeficiente de autocorrelação espacial restrinja-se ao intervalo $[0, 1]$.

Neste estudo, utilizar-se-á pesos espaciais calculados com base no critério de contigüidade, bem como pelo inverso da distância euclidiana calculada a partir de centróides. Como a matriz W é definida exogenamente, autores como Acosta *et al.* (2003) sugerem que se façam simulações a fim de atestar às propriedades assintóticas dos testes de dependência espacial e a robustez dos resultados obtidos pelo teste do Multiplicador de Lagrange (ML).

2.3 Critério de Seleção dos Clubes

Na tarefa de identificação de potenciais clubes, com base na hipótese de dependência espacial e também na similaridade de fatores estruturais, propõe-se fragmentar a amostra construindo-se índices de autocorrelação espacial local, chamados de estatística LISA (*Local Indicators of Spatial Association*). A estatística LISA detecta padrões de similaridade com base na relação de dependência espacial ao nível do município. Segundo Anselin (1995) esse índice deve satisfazer pelo menos duas propriedades básicas: *i*) um índice LISA deve possuir para cada observação uma indicação de *clusters* espaciais, cuja direção e magnitude devem ser testadas a partir de valores similares computados em torno da média; *ii*) o somatório dos índices LISAs, para todos municípios, deve ser proporcional ao índice de autocorrelação global, tipicamente conhecido como I de Moran global.

Com base em Anselin (1995), tais indicadores locais podem ser obtidos a partir da seguinte formulação.

$$I_{i,t} = \frac{(p_{i,t} - \mu_t) \sum_{j=1}^N w_{ij} \cdot (p_{j,t} - \mu_t)}{\sum_{j=1}^N (p_{j,t} - \mu_t)^2}, \quad -1 < I_{i,t} < 1. \quad (13)$$

Em que $(p_{i,t} - \mu_t)$ e $(p_{j,t} - \mu_t)$ são valores da produtividade agrícola computada como desvio em torno da média. Testa-se a hipótese nula de ausência de associação espacial local, com base na diferença entre o valor estimado da estatística com o valor de sua esperança matemática, $E(I) = -1/(N-1)$. Convém salientar, que a inferência estatística pode se constituir em um problema, uma vez que sua distribuição genérica é de difícil apuração. Para contornar esse problema, Anselin (1995) sugere que se trabalhe com resultados assintóticos, mas caso não seja possível uma alternativa é fazer uso de uma aleatorização condicional que permita auferir pseudo-níveis de significância⁵.

Ressalta-se que o índice em destaque resulta de uma relação entre valores ponderados (através dos w_{ij}) entre o desvio da produtividade em qualquer município i com seus vizinhos j , comparando-os com o quadrado da soma dos desvios dos vizinhos j . Caso o valor da produtividade de dado município i seja inferior a média da região, o índice assumirá valor negativo, indicando presença de autocorrelação espacial negativa. Valores negativos estão associados à existência de *clusters* espaciais com valores dissimilares da média da região ao ainda de seus vizinhos ricos e mais próximos.

Depois de obtidos os indicadores de autocorrelação local, convém plotá-los em um mapa da Amazônia Legal a fim de observar se há evidências de formação de *clusters* espaciais. Segundo Anselin (1995) esses *clusters* configuram bolsões de homogeneidade. Como mencionado em (13), os indicadores variam no intervalo fechado entre $[-1, +1]$, caracterizando diferenças entre regimes de transição espacial. Ainda segundo o autor, valores positivos e estimados acima da média devem ser interpretados como Alto-Alto (AA) e abaixo da média como Baixo-Baixo (BB). Valores estimados negativo são chamados de Alto-Baixo (AB) ou Baixo-Alto (BA). Comumente, observa-se concentração de municípios com valores positivos seja AA ou BB, porém margeados por municípios com valores negativos (AB ou BA), comportamento que pode refletir momentos de transição de tais economias.

Assim, a obtenção dos índices I de Moran local determinará subregiões consideradas zonas de homogeneidade dentro da amostra Amazônia Legal; com efeito, tem-se a subdivisão da mesma de forma criteriosa. Espera-se que essas zonas sejam dotadas de características comuns dos fatores determinantes do processo de convergência. Caso esta hipótese se confirme, então, não se pode descartar a hipótese de clubes em favor a condicional.

2.4 Modelo Econométrico

Com base em Anselin (1988), a dependência espacial pode ser operada de várias formas, pela variável dependente ($W\Delta p_i$), pelos termos de erros ($W\varepsilon$), bem como pelas variáveis explicativas (WX), sendo que as duas primeiras são as mais comumente utilizadas em trabalhos empíricos. Na literatura pertinente a econometria espacial (ANSELIN, 1988; LeSAGE, 1999) a forma genérica de modelos autorregressivos espaciais é chamada de SAC (*Spatial mixed regressive Autoregressive Complete*), combina a dependência espacial na variável dependente e no termo de erro. Porém, relaxando parcialmente uma das hipóteses, por exemplo, considerando ausência de dependência espacial nos erros, pode-se formalizar o chamado modelo semi-correlacionado autorregressivo, SAR (*Spatial Auto-Regressive*), tal como:

$$\Delta p_i = \rho W \Delta p_i + \alpha'_i t - \beta'_i p_0 + \phi'_i X + \varepsilon_i. \quad (14)$$

Com matriz de variância-covariância igual a, $\Omega = \sigma_\varepsilon^2 (I_N - \rho W)^{-1} [(I_N - \rho W)^{-1}]'$.

⁵ Pseudoníveis de significância sugerem aproximações marginais por intermédio da desigualdade de Bonferroni ou por meio da distribuição proposta por Sidák (1967) *apud* Anselin (1995). "This mean that when the overall significance associated with the multiple comparisons (correlated test) is set to α , and there are m comparisons, then the individual significance α_i should be set to either α/m (Bonferroni) or $1-(1-\alpha)^{1/m}$ (Sidák). Neste artigo será utilizada a análise de Bonferroni.

Em que W representa uma matriz ($N \times N$) de pesos espaciais definida anteriormente e I_N uma matriz identidade de mesma ordem. Nesta especificação, a dependência espacial incide exclusivamente através de um erro de medida associado à variável dependente, considerada tipicamente aleatória e latente. Por outro lado, a dependência espacial pode ser operada somente no termo de erro ε , e assim obter a especificação SEM (*Spatial Error Model*).

$$\begin{aligned}\Delta p_i &= \alpha'_i I - \beta'_i p_0 + \phi'_i X + \varepsilon_i, \\ \varepsilon_i &= \lambda W \varepsilon_i + \mu_i\end{aligned}\quad (15)$$

Com matriz de variância-covariância igual a $\Omega = \sigma_\varepsilon^2 (I_N - \lambda W)^{-1} [(I_N - \lambda W)^{-1}]'$ e esperando que $\mu \sim N(0, \sigma_\mu^2)$.

Tanto na especificação SAR quanto em SEM é importante observar o papel dos coeficientes ρ e λ , conhecidos como coeficientes autorregressivos espaciais. Esses escalares expressam a significância da dependência espacial, seja ela operada pela variável dependente ou pelo termo de erro, respectivamente.

Observe o argumento da matriz de variância-covariância em ambos os modelos espaciais (SAR e SEM), ele é composto por uma matriz identidade I_N , pelos coeficientes autorregressivos espaciais (ρ e λ), pela matriz W e pela variância populacional σ^2 . Havendo ausência de dependência espacial; isto é, assumindo que ρ e λ sejam iguais a zero, a matriz de variância-covariância resume-se a $\Omega = \sigma^2 I_N$. Neste caso, os resíduos da matriz Ω assumem propriedades desejáveis, porque seguem não correlacionados e identicamente distribuídos. Por outro lado, caso ρ e λ assumam valores diferente de zero, a estrutura de Ω segue correlacionada na covariância; em conseqüência, não se pode descartar a hipótese de dependência espacial.

Segundo Anselin (1988), quando a natureza da dependência espacial reside exclusivamente no termo de erro a variância de Ω torna-se não constante ao longo do arranjo espacial e este problema afeta a propriedade de eficiência dos estimadores. Caso a dependência espacial opere pela variável dependente, a autocorrelação espacial produz estimadores viesados e inconsistentes quando obtidos por OLS e o viés não desaparece mesmo aumentando o tamanho da amostra indefinidamente. Isto leva a concluir que sob presença de dependência espacial, o estimador clássico (OLS) não incorpora em si propriedades capazes de controlar a influência de tais efeitos, por isso não podem produzir estimativas consistentes quando a influência se dá através variável dependente, bem como perde em eficiência quando a influência ocorre por meio do termo de erro.

Uma saída plausível sugerida pelo Anselin (2001) consiste em conduzir o processo de estimação através do método de Máxima Verossimilhança (ML), uma vez que tal estimador dispõe de propriedades assintóticas desejáveis para controlar a influência dos efeitos espaciais. Segundo o autor, informações sobre o vetor de parâmetros $[\beta, \lambda, \sigma^2]$ do modelo SEM, podem obtidas com a otimização da seguinte função log-verossimilhança:

$$l(\beta, \lambda, \sigma^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} (\Delta p - X\beta)' (I_N - \lambda W)' (I_N - \lambda W) (\Delta p - X\beta) + \ln |I_N - \lambda W|. \quad (16)$$

Enquanto, para o modelo SAR, o vetor de parâmetros $[\beta, \rho, \sigma^2]$ pode ser estimado com base na otimização da função:

$$l(\beta, \rho, \sigma^2) = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} (\Delta p - \rho W y - X\beta)' (\Delta p - \rho W y - X\beta) + \ln |I_N - \rho W|. \quad (17)$$

Desde que as matrizes $(I_N - \lambda W)$ e $(I_N - \rho W)$ do modelo SEM e SAR, respectivamente, admitam inversas.

Outra questão que deve ser discutida e diz respeito à endogeneidade produzida pelo modelo SAR, porque se considera como regressor a variável dependente defasada. Isto viola a hipótese de que os mesmos devem se manter não-estocásticos. Por isso, Kelejian e Prucha (1998) sugerem que a estimação seja operada via variáveis instrumentais, conduzido por mínimos quadrados dois estágios (2SLS).

2.4.1 Testes de Dependência Espacial

Embora o estimador clássico não seja o mais indicado para estimar parâmetros consistentes em estruturas autocorrelacionadas, o estimador de OLS assume papel central no diagnóstico de dependência espacial, uma vez que sobre seus resíduos podem ser impostos testes indicados para detectar a presença de efeitos espaciais. Como mencionado anteriormente, um desses testes é o I de Moran formalizado aqui em sua versão global.

$$I = \frac{N}{S_0} \frac{\hat{\varepsilon}' W \hat{\varepsilon}}{\hat{\varepsilon}' \hat{\varepsilon}}. \quad (18)$$

Onde $\hat{\varepsilon}$ é um vetor de resíduos de ordem ($N \times 1$) extraído por OLS, o termo $S_0 = \sum_i \sum_j w_{ij}$ é um escalar e corresponde a soma de todos os elementos não-nulos da matriz W , N é número de observações. Embora de fácil aplicabilidade, o teste I de Moran precede da hipótese de normalidade do termo de erro e sua inferência estatística é validada através da significância do z -value, definido por:

$$z_I = \frac{I - E(I)}{SD(I)}. \quad (19)$$

Em que o termo $E(I)$ é a esperança matemática e $SD(I)$ é o desvio padrão da estatística.

Testa-se a hipótese nula de completa aleatoriedade quanto à distribuição espacial dos resíduos, em caso de aceitação de tal hipótese, então a variável se distribui ao acaso ao longo do espaço. Caso contrário, a distribuição dos resíduos é influenciada por características intrínsecas a estrutura espacial.

Todavia, a fragilidade da estatística I de Moran reside no fato de não especificar claramente a natureza da dependência espacial. Por esse motivo, Anselin (1988) e recentemente Floraz, Flomer e Rey (2003), destacam uma família de testes baseados em propriedades do estimador ML. Os testes, cujos acrônimos são LM-error e LM-lag, são capazes de definir a

natureza da autocorrelação espacial, se originária no termo de erro ou na variável dependente. Segundo Burridge (1981) o teste LM-error é proporcional ao quadrado da estatística I de Moran.

$$LM_{err} = \frac{\hat{\varepsilon}W\hat{\varepsilon} / \hat{\sigma}^2}{T} \sim \chi_1^2 \quad (20)$$

Com variância $(\hat{\sigma}^2 = \hat{\varepsilon}'\hat{\varepsilon} / N)$ estimada por máxima verossimilhança e com T equivalente ao traço da matriz de pesos espaciais, $T = tr(WW + W^2)$. O teste assintótico segue a distribuição χ^2 para um grau de liberdade.

O teste LM-lag direcionado para o caso em que a defasagem espacial opera-se sobre a variável dependente, é discutido por Anselin (1988) e formalizado por:

$$LM_{lag} = \frac{(\hat{\varepsilon}W\Delta p_i / \hat{\sigma}^2)^2}{NJ} \sim \chi_1^2, \text{ com } J = \frac{1}{N\hat{\sigma}^2} [(WX\hat{\beta})' [I_N - X(X'X)^{-1}X']^{-1} (WX\hat{\beta})]. \quad (21)$$

Em que, X é uma matriz de tamanho $(N \times K)$, $\hat{\beta}$ é um vetor de parâmetros estimado por OLS, $T = tr(WW + W^2)$ é o traço da matriz W . O teste LM-lag também segue a distribuição χ^2 para um grau de liberdade.

Ambos os testes visam atestar a hipótese nula de ausência de autocorrelação espacial, considerando estrutura de dependência espacial claramente definida, sob as especificações SEM ou SAR. Caso ambos os testes não sejam estatisticamente significantes; então, o estimador OLS é o mais apropriado. Porém, caso ambos os testes apresentem resultados significantes, aí se recorre às inferências se suas versões robustas, definidos em Anselin *et. all* (2004).

A versão robusta do teste LM-error é definida por:

$$LM_{err}^* = \frac{(\hat{\varepsilon}W\hat{\varepsilon} / \hat{\sigma}^2 - T(NJ)^{-1} \hat{\varepsilon}W\hat{y} / \hat{\sigma}^2)^2}{T[1 - T(NJ)]^{-1}} \sim \chi_1^2. \quad (22)$$

Enquanto a versão robusta do teste LM-lag é dada por:

$$LM_{lag}^* = \frac{(\hat{\varepsilon}W\hat{\varepsilon} - \hat{\varepsilon}W\hat{y} / \hat{\sigma}^2)^2}{NJ - T} \sim \chi_1^2. \quad (23)$$

Em Florax, Folmer e Rey (2003) pode-se encontrar uma regra de bolso que ajuda na tarefa de seleção de modelos. Esses autores explicam que os procedimentos necessários para identificar a melhor especificação de modelos espaciais, podem ser conferidos em seis passos:

1. Estima-se inicialmente o modelo proposto por OLS e extraem-se os resíduos.
2. Testa-se a hipótese ausência de dependência espacial, com base nas inferências das estatísticas LM-error e LM-lag, respectivamente.
3. Caso ambos os testes não sejam significantes, o modelo clássico estimado por OLS é o mais indicado. Caso contrário, segue-se ao passo seguinte.
4. Se ambos os testes sejam significantes, devem-se estimar suas versões robustas e comparar os resultados. Por exemplo: se $LM_{err}^* > LM_{lag}^*$ estima-se o modelo SEM; caso $LM_{lag}^* > LM_{err}^*$ estima-se o modelo SAR.
5. Obviamente, se LM-error for significativo e LM-lag não for, estima-se o modelo SEM, sem precisar fazer uso de suas versões robustas.
6. Mas se LM-lag for significativo e LM-error não for, estima-se o modelo SAR.

Este processo também é resumido em Anselin (2005) e pode ser analisado com auxílio da Figura 1 do anexo.

3. Base de Dados e Variáveis Utilizadas

Na implementação do modelo econométrico, empregam-se dados originários do Censo Agropecuário 1995-96 e 2006, pesquisa planejada e coordenada pelo Instituto de Brasileiro de Geografia e Estatística (IBGE) a nível municipal. Dos atuais 769 municípios que compõem o mapa da Amazônia Legal, somente 630 municípios informaram ter algum tipo de atividade agrícola, restringido o tamanho da amostra para 82% da população.

A variável-chave do modelo é a taxa de crescimento da produtividade média agrícola (YPMAGR9606) que foi calculada entre os anos de 1996⁶ a 2006, com base na diferença logarítmica do valor da produção (medido em R\$ e deflacionado para o ano de 2006) e a área cultivada medida em hectares. Com isso, tem-se a produtividade monetária média em termos reais por hectare (R\$/ha), a qual foi inspirada no trabalho de Freitas e Spolador (2006). Segundo esses autores, essa medida está diretamente associada à rentabilidade, portanto, a sustentabilidade econômica dos produtores. Ademais, a condição inicial foi obtida com base no logaritmo da produtividade média agrícola do ano de 1996 (PMAGR96).

A definição do intervalo temporal está condicionada a carência de informações a este nível de desagregação, de modo que permita construir uma amostra com abrangência mínima desejável, se comparada com o tamanho da população. Todos os municípios pesquisados foram fundados antes de 1996 e mantiveram suas áreas após 2006. Isto significa que municípios novos fundados entre 1996 e 2006, assim como aqueles que foram subdivididos no período, ficaram de fora da amostra. Outro fator importante que caracteriza esse período é o controle mais rigoroso dos níveis de preço, sobretudo, daqueles relativos aos custos de produção agrícola. Lembrando que no início do período, o cenário era de estabilidade dos preços, crescimento do nível de produto e de emprego.

⁶ Embora o Censo Agropecuário 1995-96 tenha transpassado de um ano ao outro, toma-se o ano de 1996 como referência inicial.

Para fins de comparação, é pertinente que a análise seja conduzida com base em medidas relativas. Assim, a matriz de regressores condicionais X foi composta pelas seguintes variáveis de controle, com suas respectivas denotações.

- A proporção de gastos com capital físico (PGCF06) utilizado na produção agrícola, calculada com base na soma dos seguintes gastos: gastos com adubos, corretivos e agrotóxicos; gastos com combustíveis e lubrificantes; e gastos com energia elétrica. Tudo isso em razão do valor da produção agrícola em 2006.
- Para o nível de capital humano (MAE06), utilizou-se uma *proxy* calculada a partir da média ponderada de anos de estudo de pessoas que dirigem o estabelecimento. Com base na metodologia empregada na Pesquisa Nacional por Amostra de Domicílio (PNAD) do IBGE, a classificação dos anos de estudo é função da série e do grau que o indivíduo estava frequentando, considerando a última série concluída em 2006. Por exemplo, nível fundamental completo, 8 anos de estudo; nível médio completo, 12 anos; e nível superior completo, 18 anos.
- Como *proxy* de custo de transporte, utilizou-se a proporção de gastos despendido com o transporte da produção (PGTR06).
- O valor do investimento realizado por estabelecimento (INVK06) em ativos móveis como aquisição de novas terras, novos prédios e novas instalações; assim como em ativos imóveis: aquisição de veículos, máquinas e tratores sejam novos ou usados. Ressalta-se que uma parcela significativa do investimento foi realizada via crédito agrícola.
- Índice de concentração de unidades produtivas (AGLM06), obtido por meio do índice T de Theil (THEIL, 1976 *apud* HOFFMANN, 2006). Para este caso, o conceito de *concentração* assemelha-se ao de *desigualdade*; logo, concentração máxima equivale dizer que único produtor responde por toda produção do município i . Segundo Hoffmann (2006), a vantagem dessa medida é que, como a produtividade agrícola pode ser agrupada por município, o índice T de Theil (T) pode ser decomposto ao nível do mesmo, desde que se obtenha a diferença entre o logaritmo de n e a entropia $H(p)$ da distribuição de probabilidade da produtividade agrícola.⁷

$$T = \log(n) - H(p), \quad 0 \leq T \leq \log(n) \quad (24)$$

Se $T = 0$ implica que o logaritmo natural do número de indivíduos iguala-se a entropia da distribuição de produtividade agrícola $\log(n) = H(p)$; em outras palavras, há perfeita igualdade da distribuição da produtividade. Havendo perfeita desigualdade, a entropia iguala-se a $H(p) = 0$, o que implica em $T = \log(n)$.

- Devido à relevância dos gastos com folha de pagamento dentre os gastos totais, utilizou-se a proporção dos gastos com salários pagos (PGSAL06) ao trabalhador empregado na atividade agrícola.
- Por fim, o termo $(n_i + g_i + \delta)$, é representado pela taxa de crescimento do pessoal ocupado, mais a taxa de variação tecnológica capturada pela variação do número de estabelecimentos que receberem orientação técnica no período e taxa de depreciação dos estoques dos fatores, que segundo Barro *et al.* (1995) cresce a 0.03 ao ano. Esta variável recebe a denotação TECP06 no modelo.

4. Resultados e Discussões

Esta seção está subdividida em outras três. A primeira trata da distribuição espacial dos níveis de produtividade média agrícola entre os anos de 1996 e 2006, bem como da detecção de *clusters* espaciais a partir dos resultados dos indicadores LISA's. A segunda dedica atenção ao diagnóstico de dependência espacial inferidos sobre a amostra Amazônia Legal e suas respectivas subdivisões amostrais. Finalmente, a terceira avalia a determinação de fatores condicionais ao processo de convergência clubes e condicional.

4.1 Distribuição e Clusters Espaciais

Primeiramente, pretende-se diagnosticar padrões de distribuição dos níveis de produtividade média agrícola ao longo do espaço. Para tanto, faz-se uso da Análise Exploratória de Dados Espaciais (AEDE), que reuni um conjunto de técnicas e procedimentos estatísticos apropriados para manipular dados georeferenciados. O método, permitir visualizar em um mapa a distribuição da variável de interesse ao longo do espaço e engloba ainda a estatística I de Moran global e local. Assim, sobre o mapa (ver Figura 2 do anexo) foram plotados níveis de produtividade média agrícola entre os 630 municípios que compreendem a amostra Amazônia Legal. Os valores individuais foram calculados em desvio em torno da média, para os anos de 1996 e 2006 (PMAGR96Z e PMAGR06Z) e estão computados em logaritmos.

Em primeira análise, nota-se, em 1996, que 287 e 178 municípios frequentavam as primeiras classes, imediatamente a baixo e acima da média amostral, respectivamente. Isto significa que 74% dos municípios (465) da amostra Amazônia Legal (AL) auferiam níveis de produtividade agrícola próximo da classe mediana. Mas em 2006, observa-se uma pequena alteração, 214 e 236 municípios frequentavam as mesmas classes, representando 71% (450) dos municípios que compõem a amostra. Contudo, convém comparar alterações dos níveis de produtividade média entre diferentes subespaços regionais.

Inicialmente, os maiores níveis de produtividade média agrícola concentravam-se entre grupos de municípios localizados no extremo Oeste da Amazônia Legal, especificamente no estado do Amazonas (AM), Acre (AC) e parte de Rondônia (RO). A maioria desses municípios apresentavam valores positivo e afastado da média amostral em 1996, com alguns frequentando classes com altos níveis de produtividade. Mas em 2006, alguns desses perdem eficiência produtiva e passam a frequentar classes abaixo da média amostral (Figura 2 do anexo).

No extremo norte da Amazônia Legal, também se observa comportamento semelhante. Alguns municípios do estado de Roraima (RR), Pará (PA) e Amapá (AP), que apresentavam valores positivos e acima da média (1996), passaram a frequentar classes abaixo da média no final do período. Contudo, no Pará (PA), alguns municípios situados no nordeste e no

⁷ Em 1967, no livro *Economics and Informations Theory*, Henry Theil criou medidas de desigualdade baseadas na teoria da informação. A preposição fundamental da teoria está associada à probabilidade de ocorrência de um evento com alto valor informativo. Assim, a entropia é esperança matemática de ocorrência de um evento com alto conteúdo informativo, que varia no intervalo $0 \leq H(x) \leq \log(n)$, sendo x a probabilidade. Ver Hoffmann (2006).

sudeste do estado obtiveram ganhos de eficiência, possibilitando migrar a classes daqueles situados a cima da média amostral.

Os municípios situados na região nordeste da amostra AL (Maranhão – MA) apresentavam padrões de distribuição bem semelhante, com baixíssimo grau de migração entre classes. Porém, alguns se destacam por conseguirem migrar de classes situadas abaixo da média para classes superiores. Semelhantemente a este cenário estão os produtores dos municípios do estado do Tocantins (TO), apresentando crescimento e distribuição dos níveis de produtividade agrícola de forma mais equilibrada, uma vez que não se observa grande disparidade sobre a distribuição e baixos fluxos migratórios entre classes (Figura 2).

O estado do Mato Grosso (MT) destaca-se pela forte aptidão agrícola, com atividades apresentando elevado nível tecnológico baseada no plantio mecanizado de grãos (com ênfase na soja). Em 1996, havia uma concentração excessiva de produtores em terras no extremo sul do estado, mas que aos poucos migram para o norte da região, tornando o processo de distribuição da produtividade agrícola menos concentrado. Aspectos semelhantes experimentam os agricultores do estado de Rondônia (RO), cujos mais ricos situavam-se em municípios ao oeste do estado. No entanto, em 2006 esse contexto modifica-se, em que os municípios mais ricos estão localizados mais ao leste do estado (Figura 2 do anexo).

Esses movimentos no processo de distribuição definem diferenças entre padrões de crescimento da produtividade agrícola em diferentes subregiões, aspecto que caracteriza a heterogeneidade espacial. Tal processo está relacionado à influência de fatores como redução de custo relativo do produtor (como o custo de transporte), maior disponibilidade de recursos e insumos (como preço da terra e mão-de-obra), o que gera conseqüentemente uma tendência de concentração da atividade, configurando os pólos mais dinâmicos de cada subregião (FUJITA, KRUGMAN e VENABLES, 1999).

Em linhas gerais, pode-se constatar (ver Figura 3 do anexo) um processo de acomodação e redistribuição da produtividade agrícola no conjunto dos municípios da Amazônia Legal, embora persista forte relação de dependência espacial a nível global, como mostram os resultados da estatística *I* de Moran.

O indicador *I* de Moran proporciona uma medida de dependência espacial em nível global e sua inferência estatística é validade comparando-o com a distribuição normal. Testa-se a hipótese nula de completa aleatoriedade quanto à distribuição do atributo ao longo do espaço, ou seja, de que o *I* de Moran é igual a zero.

Utilizando a matriz *W* (630×630), construída com pesos baseados na contigüidade das fronteiras, obtiveram-se os resultados dessa estatística a partir dos desvios em torno da média da produtividade agrícola nos anos de 1996 e 2006, respectivamente (PMAGR96Z e PMAGR06Z). Os resultados (*I* de Moran igual a 0.458 e 0.287 para anos de 1996 e 2006, respectivamente) revelaram a influência da dimensão espacial sobre a distribuição da referida variável ao longo do espaço, uma vez que se deve rejeitar a hipótese nula com uma probabilidade tão perto de zero (*prob*=0.01), em ambos os casos. Esses resultados revelam ainda uma redução do grau de dependência espacial ao longo dos anos – *I* de Moran (1996)=0.458 ~ *I* de Moran (2006)=0.287 – mas que é acompanhada por redução das desigualdades na distribuição da variável de interesse – Desvio padrão(1996)=0.673 ~ Desvio padrão (2006)=1.320.

No entanto, entre subregiões, em que a influência da dimensão espacial é marcada por padrões característicos da dependência espacial a nível local, o *I* de Moran Global pouco tem a dizer a respeito do problema. Por isso, foram estimados os indicadores de autocorrelação espacial local (estatística LISA), os quais permitem identificar subregiões cujo padrão de distribuição da produtividade agrícola é considerado mais homogêneo.

Na inferência da estatística LISA, foram feitas 99 permutações, resultando em uma pseudo-distribuição, que permiti comparar os resultados estimados para os índices LISAs com os *n* (630) possibilidades geradas a partir de uma distribuição aleatória (ANSELIN, 1995). Aqueles municípios com o indicador não significativo foram plotados em cinza no *Moran-Map*⁸. Os demais indicadores significativos (considerou-se probabilidade até 0.05), que apresentam padrões de dependência espacial distinto, estão relacionados às outras cores, ilustradas na legenda (ver Figura 4 do anexo).

A aplicação da estatística LISA revela a extensão das áreas de influência associadas à produtividade agrícola nos anos de 1996 e 2006, indicando tendência de se encontrar valores localmente similares, proporcionando assim, a subdivisão da amostra com base nos padrões de dependência espacial e nos níveis de produtividade no ano inicial. Assim, os *clusters* relativamente ricos, aqueles municípios com desvios positivos e a cima da média (Alto-Alto), encontravam-se situados ao Oeste da Amazônia Legal, especificamente nos estados do Acre (AC) e Amazonas (AM); e ao Norte, nos estados do Amapá (AP) e Pará (PA). Como resultado, tem-se a divisão amostral por subregiões, cuja denotação segue: OE (Oeste) e NO (norte), respectivamente. Por outro lado, agrupamentos relativamente pobres, uma vez que apresentam níveis de autocorrelação espacial positivo, porém abaixo da média (Baixo-Baixo), localizavam-se no eixo centro-sudeste da Amazônia Legal, envolvendo municípios do estado do Pará (PA), Tocantins (TO) e Mato Grosso (MT). Assim, este *cluster* compõe a subregião SE (Sudeste).

Destacam-se ainda outros *clusters* em processo de formação, localizados no SO (Sudoeste) da região Amazônica, envolvendo grupos de municípios do estado de Rondônia; mais ao SU (Sul) da região, com municípios do Mato-Grosso; e ao Nordeste (NE) da mesma, com produtores estabelecidos no estado Maranhão. Nestas subregiões identificou-se prevalência de ausência de autocorrelação espacial em nível local, computada no ano de 1996. Porém, cabe ressaltar que esses *clusters* passam por um momento de transição e mostram-se fortalecidos em 2006 (Figura 4).

4.2 Diagnóstico de Dependência Espacial

O diagnóstico de dependência espacial reúne uma família de testes originários do Multiplicador de Lagrange (LM-*error*, LM-*lag* e suas versões robustas), além do teste de Wald e *I* de Moran, direcionados aos erros da regressão. A vantagem dos testes LM é que apesar de ser baseado no processo de máxima verossimilhança, ele exige apenas o vetor de resíduos e o

⁸ Associado ao *Moran Map* está o Diagrama de dispersão de Moran (*Sactter Plot*), através dos quadrantes AA, BB, AB e BA.

vetor de parâmetros estimados por OLS, o que reduz bastante o custo do cálculo estatístico. No mais, os testes LM são estruturados de forma específica à natureza da dependência espacial, considerando se os efeitos são determinados pelos erros, através de validação do parâmetro λ , ou, pela variável dependente através da validação do ρ .

Assim, com base nos dados amostrais e utilizando duas estruturas espaciais distintas – a matriz de contigüidade (W_1) e matriz de distância euclidiana inversa (W_2) – testa-se a hipótese de ausência de dependência espacial originária nos erros ou na variável dependente, por meio do modelo β -convergência condicional e clube, conforme em (14) e (15). Ressalta-se, que o modelo mais apropriado para efetuar o controle dos efeitos espaciais, de acordo com os resultados dos testes de dependência espacial, está indicado pelo símbolo (*) no Quadro 1 do anexo.

Admitindo que a hipótese de ausência de dependência espacial deva ser rejeita a probabilidade de erro igual ou menor a 0.05. Então, de acordo com os resultados dos testes LM-error, corroborado por LM-error*, conclui-se a favor de algum grau de dependência espacial operada pelos erros da regressão das amostras AL, NO e SE. Com efeito, o modelo SEM seguramente é o mais apropriado para efetuar a correção dos efeitos espaciais e assim obter resultados estatisticamente desejáveis. Por sua vez, visto que a determinação dos efeitos espaciais origina-se na variável dependente. Então, não se pode aceitar a hipótese de ausência de dependência espacial, de acordo com os resultados dos testes LM-lag e LM-lag* para a amostra NE, indicando, que o modelo SAR é o mais adequado para acomodar a influência de tais efeitos na região.

A amostra SO, revela resultados estatisticamente significantes no que diz respeito aos testes LM-error e LM-lag. No entanto, os resultados de suas versões robustas mostram-se insignificantes. Neste caso, segundo Anselin (1988), pode-se operar o modelo geral chamado de SAC (*Spatial mixed regressive Autoregressive Complete*), que combina a dependência espacial na variável dependente e no termo de erro. Porém, caso não haja diferenças sistemática entre os parâmetros estimados nos dois modelos, torna-se facultada à utilização dos mesmos.

A amostra OE, onde há predomínio de valores Alto-Alto, apresenta clara ausência de dependência espacial em qualquer teste ou qualquer estrutura espacial que seja avaliada. Isto significa, que o modelo β -convergência clubes pode ser perfeitamente testado via OLS, sem que haja nenhum prejuízo a confiabilidade dos resultados.

Na tarefa de seleção de modelos, a amostra SU revela resultado inconclusivo, pelo menos a priori. Constata-se que não se pode rejeitar a hipótese de dependência espacial em todos os testes realizados, considerando a probabilidade usual de 0.05. Porém, ao nível de erro equivalente a 0.10 a hipótese de ausência de dependência espacial computada pelo teste LM-lag poderia ser rejeitada, se, e somente se, a estrutura espacial fosse construída com base na matriz W_2 de distância euclidiana inversa. Por esse motivo, autores como Acosta *et. all* (2003) sugerem que se façam simulações de Monte Carlo, a fim de testar às propriedades assintóticas dos testes de dependência espacial sob diversas estruturas, bem como conferir a robustez dos resultados paramétricos obtidos pelo teste de Multiplicador de Lagrange (LM).

4.3 Análise Convergência Clubes

Controlando o efeito da autocorrelação espacial, busca-se testar agora a prevalência da hipótese clubes (em que se destacam os *clusters*) sobre a condicional (cujo foco é o resultado obtido para a amostra AL). Além disso, cabe questionar se a influência da heterogeneidade espacial é forte o suficiente ao ponto de permitir a subdivisão da amostra AL, sobretudo, naquelas em que o modelo SEM é o indicado. Para tanto, sobre os resíduos extraídos por cada estimador selecionado, utilizou-se o teste de Wald, a fim de confirmar se a variância da matriz Ω de cada *cluster* (σ_1^2), bem como da amostra AL (σ^2) foi corrompida, após o processo de estimação (ANSELIN, 1988). Os resultados das estimações dos modelos estão descritos no Quadro 2 do anexo do trabalho.

Atentando para os resultados do teste de heterogeneidade espacial (Wald Test) observa-se que a variância da matriz Ω foi corrompida quando os resultados foram obtidos por OLS nas amostras AL, NO, SE, SO e NE, o que reforça a premissa de que o OLS é inadequado quando se introduz a hipótese de dependência espacial na modelagem. Contudo, efetuada a correção espacial, por meio da estimação do modelo SEM no que tange as amostras AL, NO, SE e SO e SAR para as amostras NE e também a SU, nota-se que a variância da matriz Ω segue constante ao longo da estrutura espacial. Nota-se ainda (Quadro 1), que a variância da amostra SU e OE segue constante, em qualquer modelo, inclusive o OLS, resultado que corrobora com o diagnóstico de dependência espacial.

Portanto, esse resultado que detecta diferencial entre padrões de crescimento e distribuição da produtividade agrícola em subregiões, é por si só, suficiente para justificar a subdivisão da amostra Amazônia Legal (AL), e, por conseguinte, corrobora com a hipótese de existência de grupos de municípios com características diferenciadas.

Posto isso, a condição inicial (PMAGR96) apresenta papel fundamental para o processo de convergência da produtividade agrícola, porque se mostra estatisticamente significativa em qualquer hipótese admitida (clubes ou condicional), em qualquer especificação selecionada (OLS, SEM ou SAR) e também em qualquer estrutura espacial construída – matriz de contigüidade (W_1) e matriz de distância euclidiana inversa (W_2) (ver Quadro 2 do anexo).

No entanto, observa-se diferenças paramétricas do β estimado entre amostras, em consequência, o processo de convergência evolui sob diferentes condições na transição. Por exemplo, a subregião NO destaca-se por ter alta condição inicial, a segunda maior taxa de crescimento (5,1% a.a) e com a maior velocidade de convergência ano⁹ (8,4%) entre os *clusters* identificados. A região SO também merece destaque, porque foi a que mais cresceu (6,2% a.a.) e seus municípios caminham para equilíbrio de longo prazo à velocidade próximo a 7,2% a.a. Estes resultados permitem reposicionar a região SO ao final do período, classificando-a entre aquelas com alto nível de produtividade agrícola e com maior taxa de crescimento, juntamente com a região NO (ver Quadro 3 do anexo).

Em média, os demais *clusters* apresentaram crescimento negativo dos níveis de produtividade. A região OE, a mais rica no início do período, foi a que mais decresceu (12,9% a.a), seguida pela região SE(3,1% a.a), NE(2,8% a.a) e SU(2,1% a.a),

⁹ A velocidade de convergência rumo ao *steade-state* é obtida por meio de, $\gamma = \ln(1 + |\beta|) / t$. Com t=10 anos, aproximadamente. Ver Barro (1995).

respectivamente. Em parte, isso se explica pela sobreposição da taxa de crescimento da área de cultivo em relação ao crescimento do valor da produção. Porém, existem fatores estruturais, em cada uma dessas regiões, que têm influenciado, sob diferentes medidas e direção, o crescimento da produtividade agrícola.

Fatores como a média de anos de estudo do produtor (MAE06) e a forte relação de dependência espacial experimentada entre produtores ($\lambda=0.437$) tem contribuído positivamente para redução das disparidades na subregião NO, fato constatado pela significância dos coeficientes associados a estas variáveis. Constata-se ainda, que os municípios com maiores taxas de crescimento da produtividade agrícola são aqueles que auferem maiores médias de anos de estudo do produtor¹⁰. Este resultado reforça a contribuição de Lucas (1988) e enfatiza a importância do investimento em capital humano. Indivíduos instruídos são capazes de gerar inovações, e, por conseguinte, externalidades positivas por meio fluxos informacionais, que segundo Lucas (1988), seriam os “motores” de crescimento econômico. Contudo, o próprio autor adverte para dificuldade de estimar a taxa de retorno do capital humano, a partir de dados diretamente observáveis, como os adotados aqui.

Contraopondo essa tendência, a segunda subregião com maior velocidade de convergência (SO), teve o crescimento da produtividade arrefecido pela distribuição dos níveis de capital humano. Constata-se, que os municípios com maiores médias de anos de estudo são aqueles que experimentaram as menores taxas de crescimento de produtividade no período de 1996 a 2006. Por outro lado, dentre as subregiões, a SO é a que apresenta maior taxa de retorno ($\lambda=0.578^{11}$) resultante da relação de dependência espacial estabelecida entre produtores vizinhos. Este efeito pode estar recompensando, a influência do capital humano e de outros fatores não favoráveis ao desencadeamento desse processo de redução das disparidades regionais.

A política de crédito agrícola pode ser considerada um desses fatores não favoráveis, porque tem freado a velocidade de convergência não só na região NO e SO, mas também em todas as outras. Em geral, o investimento realizado via crédito agrícola em 2006 (INVK06) tem sido mais frequentes entre os produtores que menos crescem, portanto, aqueles com alta condição inicial. Isto permite concluir, que menores frações de crédito têm sido direcionadas para os produtores relativamente pobres, embora sejam os que mais cresçam. Conclui-se ainda, que a elasticidade da produtividade agrícola em relação ao investimento realizado via crédito é mais alta exatamente nestas regiões NO e SO, com coeficiente de -0.553 e -0.913 , respectivamente. As demais regiões mostram-se menos sensíveis, com coeficientes estimados na ordem de -0.398 (NE), -0.310 (SE), -0.251 (OE) e -0.222 (SU), sucessivamente.

Por sua vez, a concentração de unidades produtivas (AGML06) foi considerada fator não estimulante à velocidade de convergência, sobretudo nos *clusters* NO e SO. Autores como Fujita, Krugman e Venables (1999) argumentam que a redução dos custos de transporte agrícola pode desencadear um processo de aglomeração, da mesma forma como ocorre na indústria. Argumentam ainda que a tendência de aglomeração resulta da presença de um ambiente de retornos crescente de escala, que é aquecida pela redução do custo de transporte. Sob esse enfoque, embora não se postule a hipótese de retornos crescentes decorrentes da produtividade agrícola, porque há evidências de convergência. Constata-se, que os municípios mais crescem, são os mais aglomerados, ou seja, menos concentrados, segundo o índice *T* de Theil (Quadro 2 e 3). A região NO, apresenta maior taxa de retorno, com média de -1.616 , seguida pelas regiões SO(-1.464) e SE(-1.274), com a mesma tendência. As demais regiões (OE, SU e SE) resumem-se a insignificância estatística desse fator.

Compatível com a linha teórica dos autores da Nova Geografia Econômica está à relação de crescimento dada entre o custo de transporte e a produtividade. Em média, maiores proporções de gastos com transporte da produção agrícola (PGTR06) são requeridos naqueles municípios com as menores taxas de crescimento da produtividade, e estão localizados nas regiões SO(-0.227), OE(-0.106), NE(-0.084) e SE(-0.070).

A proporção de gastos com capital físico (PGCF06) na produção agrícola (gastos com adubos, corretivos, agrotóxicos, combustíveis lubrificantes e energia elétrica), restringe-se a insignificância para a função auxiliar ao processo de convergência em quase todas as regiões, exceto a SU(-0.411), com a terceira maior velocidade de convergência (7.1% a.a.). Ademais, o sinal do coeficiente estimado para a referida região não confere com o esperado pela teoria: Mankiw, *et al.* (1992), Barro *et al.* (1996), Barro (1997), Durlauf e Johnson (1995) e Johnson e Takeyama (2003). Constata-se, que frações de uso do capital físico são frequentemente requeridas entre produtores com as menores taxas de crescimento, e no caso da região SU, 38 dos 56 municípios que compreendem o *cluster*, apresentaram retração dos níveis de produtividade, implicando em retorno negativo sobre a taxa de crescimento média agrícola.

Em parte, esse resultado não esperado pode decorrer da depreciação do termo de troca, ou seja, pela diminuição da diferença entre o preço recebido e pago pelo produtor, como concluem Freitas e Spolador (2006). Esses autores buscam entender os determinantes da relação de troca, sob a estrutura de um sistema VAR (Vetor Auto-Regressivo), considerando fatores pelo lado da demanda como taxa de câmbio, taxa de crescimento da economia mundial, preços internacionais e distância entre mercados; bem como pelo lado da oferta como, o custo de financiamento da atividade, custos dos fatores como mão-de-obra e terra, custo de insumos como fertilizantes, adubos e defensivos e outros. Constatam, que despesas com agroquímicos são um aspecto importante da estrutura da oferta e tendem apresentar certa rigidez para baixa, dada a natureza desses mercados, classificados como oligopólio diferenciado. Por outro lado, a produção de grãos (soja) é fortemente exógena, o que reforça a tese de que o produtor é um tomador de preço.

De certa forma, este resultado é compatível com o encontrado aqui. Como mencionado anteriormente, a taxa de crescimento do valor da produção é positiva em todas as subregiões. No entanto, a área de plantio nas subregiões NE, OE, SE e SU, tem crescido a velocidade superior ao valor da produção, repercutindo fortemente sobre o crescimento da produtividade agrícola, inclusive no SU, que apresentou crescimento médio negativo (Quadro 3 do anexo).

¹⁰ No Quadro 3 encontram-se informações adicionais sobre a média de anos de estudo do produtor por subregião, bem como informações das outras variáveis utilizadas no modelo (ANEXO).

¹¹ Este e os demais coeficientes mencionados no texto são médias dos coeficientes estimados nas duas estruturas espaciais (W_1 e W_2). Ver Quadro 2 (ANEXO).

Cabe ainda ressaltar algumas particularidades sobre a região SU. As variáveis condicionais discutidas neste estudo contemplam, em sua maioria, a estrutura da oferta e têm fundamentação microeconômica. A referida região é especializada no cultivo de grãos, com ênfase no plantio de soja, considerada uma *commodity* agrícola. Em geral, esses produtos homogêneos são comercializados em mercados organizados globalmente, em que fatores determinantes pelo lado da demanda (como os mencionados anteriormente), apresentam maior elasticidade sobre o crescimento do volume de produção, com efeitos recursivos, sobre o crescimento da produtividade agrícola, sobretudo, quando medida em termo monetário (R\$/ha). Estes argumentos ajudam a entender o baixo poder de explicação apresentado pelo modelo indicado para a região (Quadro 2 do anexo).

Outra variável importante na composição da estrutura de custo do produtor são os gastos com a remuneração da mão-de-obra (Quadro 3 do anexo). Embora a proporção de gastos com folha de pagamento (PGSAL06) seja expressiva e assegure o volume de produção, tal variável nada tem a dizer, enquanto taxa de retorno sobre a variável de interesse. Em geral, a variável revela-se insignificante para atuar como condicional ao processo de convergência em todas as amostras.

Paralelamente, o componente (TECP06), composto pela taxa de crescimento do pessoal ocupado, o diferencial tecnológico capturado pela variação do número de estabelecimentos que receberem orientação técnica no período e taxa de depreciação dos estoques dos fatores, também se revela insignificante em quase todas as amostras, a exceção da região NE. Este último resultado está de acordo com o esperado pela teoria. Segundo Barro *et.al* (1995), sob hipótese de mobilidade de mão-de-obra, qualquer diferencial de produtividade entre regiões deve ser eliminado. O autor adverte ainda que, aumentos da fração do estoque de capital que não é móvel retardam a velocidade de convergência.

A região NE foi a que apresentou a menor velocidade de convergência (4.7% a.a.), e teve condição inicial abaixo da média da região AL, mantendo-se assim até o final do período. Fatores como o elevado custo de transporte e principalmente o alto custo do crédito agrícola tem arrefecido o processo de convergência e contribuído para queda de produtividade. Por outro lado, a distribuição da produtividade agrícola mostra-se relativamente estável, com baixa sensibilidade a choques exógenos, embora não se descarte a hipótese de dependência espacial. Por esse motivo, o controle da autocorrelação espacial deve ser exercido através da variável dependente, cujo coeficiente foi estimado em $\rho=0.382$ (Quadros 2 e 3 do anexo).

A região SE se assemelha, em diversos aspectos, com a região NE. Ambas, obtiveram níveis iniciais abaixo da média regional no ano de 1996, e os resultados indicam que transitam para níveis de equilíbrio de longo prazo à velocidade de 4.8% e 4.7% ao ano, respectivamente. No entanto, a produtividade agrícola na região SE é mais suscetível a choques exógenos, revelando que o controle dos efeitos espaciais deve ser exercido pelo termo de erro ($\lambda=0.277$). Contudo, mantida a condição inicial e a influência de fatores estruturais como a relação de custo de transporte, a provisão de crédito agrícola e a baixa taxa de incremento tecnológico, as regiões NE e SE dificilmente alcançarão níveis de produtividade compatíveis a região NO, considerada a referência.

Não obstante a este cenário vivem os produtores da região OE. Embora apresente condição inicial relativamente alta, foi a que apresentou o maior declínio de produtividade ao longo do período. Os resultados indicam que os municípios produtores caminham para equilíbrio relativamente baixo de longo prazo, com velocidade perto de 5.9% ao ano. Entre fatores explicativos da eminente depreciação dos níveis de produtividade, somente o crédito agrícola mostra-se significativo, revelando que maiores proporções desse fator têm sido direcionadas para os produtores que menos têm crescido. Além disso, descarta-se totalmente a hipótese de dependência espacial, fator que poderia estimular a velocidade de convergência e ajudar a reverter à trajetória de retração da produtividade.

Em resumo, os resultados revelaram que aumentos marginais de capital físico tiveram retorno nulo sobre a produtividade agrícola, exceto na região SU, onde a atividade é mais mecanizada. Os estoques de capital humano estimularam o crescimento da produtividade agrícola somente na região NO. Mas este resultado se contrapôs na região SO. Uma explicação para este resultado pode ser atribuída à imigração interregional, hipótese que pode estar sendo motivada pelo diferencial da taxa de retorno do capital entre regiões. Lucas (1990) justifica a falta de mobilidade do capital (inclusive o capital humano) das regiões ricas em direção às pobres na medida em que diferenciais de taxa de retorno tendem a desaparecer.

Por fim, com base nas diferenças regionais supramencionadas e no que diz respeito aos fatores condicionais associados à dinâmica de convergência, não se deve descartar a hipótese clubes em favor a condicional. Nesta linha, identificou-se em quase todos os clubes, exceto o OE e SU, evidências de efeito transbordamento espacial. Somente em único caso, o clube NE, tal efeito decorre da defasagem espacial exprimida pela variável dependente, atestando a prevalência do modelo SAR. Em outros três clubes, *clusters* NO, SO e SE, o efeito transbordamento decorre dos choques exógenos exprimido pelo termo de erro.

Esses efeitos são transmitidos através de certos canais, sendo um deles é o canal do comércio. A comercialização de produtos agrícolas é bastante comum tanto entre mercados intra e inter regional. Outro canal está no fluxo de informação e troca de experiência, uma vez que na atividade agrícola os produtores têm certa facilidade de observar e copiar padrões de produção que sejam considerados bem sucedidos em regiões vizinhas.

5. Conclusão

Com base nos dados do Censo Agropecuário de 1995-96 e 2006, objetivou-se testar no trabalho em epigrafe qual hipótese sobre convergência – se condicional ou clubes – melhor descreveria a dinâmica de crescimento da produtividade média agrícola entre 630 municípios produtores na Amazônia Legal, com o uso da econometria espacial.

Os resultados dos testes baseados no Multiplicador de Lagrange (LM-error, LM-lag, LM-error*, LM-lag* e Wald-test) apontaram, em todos os casos, para forte heterogeneidade e dependência espacial inerente a produtividade agrícola dos 630 municípios da Amazônia Legal, revelando a falta de estabilidade dos parâmetros estruturas do modelo de β -convergência condicional. Em função disso, procedeu-se a subdivisão da amostra, baseado na relação de dependência espacial a nível

local da produtividade agrícola inicial. Os primeiros resultados revelaram a formação de seis *clusters* espaciais, sendo dois relativamente ricos localizados ao Norte (NO) e Oeste (OE) da Amazônia Legal e quatro pobres, localizados no Nordeste (NE), Sudeste (SE), Sudoeste (SO) e Sul (SU). Assim, sobre a subdivisão amostral, inferiram-se testes de heterogeneidade e de dependência espacial em que se constatou que não se pode descartar a hipótese de convergência clubes em detrimento da condicional, uma vez que existem diferenças sistemáticas dos parâmetros estruturais.

Já os resultados das estimações com correção para dependência espacial mostraram que, em geral, aumentos marginais de capital físico tiveram retorno nulo sobre a produtividade agrícola, exceto na região SU, onde a atividade é mais mecanizada. Mostraram, também, que o estoque de capital humano estimula o crescimento da produtividade agrícola somente na região NO, diferentemente do resultado para região SO, que apresentou resultados antagônicos. Uma explicação para este resultado pode ser atribuída à imigração interregional, hipótese que pode estar sendo motivada pelo diferencial da taxa de retorno do capital entre regiões. Lucas (1990) justifica a falta de mobilidade do capital (inclusive o capital humano) das regiões ricas em direção às pobres na medida em que diferenciais de taxa de retorno tendem a desaparecer.

Por fim, com base nas diferenças intra-regionais e no que diz respeito aos fatores condicionais associados à dinâmica de convergência, não se deve descartar a hipótese clubes em favor a condicional. Nesta linha, identificou-se em quase todos os clubes, exceto o OE e SU, evidências de efeito transbordamento espacial. Somente em um único caso, o clube NE, tal efeito decorre da defasagem espacial exprimida pela variável dependente, atestando a prevalência do modelo SAR. Em outros três clubes, *clusters* NO, SO e SE, o efeito transbordamento decorre dos choques exógenos exprimido pelo termo de erro.

Referências Bibliográficas

- ACOSTA, P., MATTOS, E. and FAVA, A. Monte Carlo Properties of Spatial Dependence LM test. Department of Economics, University of Illinois, Champaign, IL 61820, USA, 2003.
- ALMEIDA, E. S. Lei de Verdoorn Local Para a Agricultura. Anais do XXXV Encontro Nacional de Economia – ANPEC, Recife-PE, 2007. <Disponível em: http://www.anpec.org.br/encontro_2007.htm> Acesso em 08/07/2008.
- ALMEIDA, E. S.; e HADDAD, E. A. MEECA: Um Modelo Econométrico Espacial Para Projeção Consistente de Culturas Agropecuárias. <Disponível em: http://www.econ.fea.usp.br/nereus/td%5CNereus_02_03.pdf> Acesso em 26/07/2008.
- ANSELIN, L. Spatial Econometrics: methods and model. Kluwer Academic Publishers, Dordrecht, Netherlands, 1988.
- ANSELIN, L. Local indicators of spacial association – LISA. Geographical Analysis. v.27, n.2, p.93-115, 1995.
- ANSELIN, L. Spatial Econometrics In: Baltagi (eds). A Companion to the Theoretical Econometrics Oxford, Basil Blackwell, 2001.
- ANSELIN, L. and BERA, A. Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics In: A. Ullah and D. Giles (eds), Handbook of Applied Economic Statistics, New York: Marcel Dekker, 1998.
- ANSELIN, L.; FLORAX, R.J.G.M.; and REY, S.J. Advances in Spatial Econometrics: Methodology, Tools and Applications. Springer, Urbana, Amsterdam, San Diego, 2004.
- ARBIA, G. Spatial Econometrics: statistical foundations and applications to regional convergence. Springer-Urbana, Italy, Rome, 2005.
- AZZONI, C. Economic growth and regional income inequality in Brazil. The Annals of Regional Science, v. 35, p. 133-152, 2001.
- AZZONI, C. R.; SILVEIRA NETO, R. Location and growth in Brazilian states. Papers in Regional Science, 2003.
- BARRO, R. Economic growth in a cross section of countries. Quarterly Journal of Economics, v. 106, p. 407-444, 1991.
- BARRO, R.J. Determinants of Economic Growth: A Cross-Country Empirical Study. MIT Press, Cambridge-MA, 1997.
- BARRO, R.; MANKIW, N.; SALA-i-MARTIN, X. Capital mobility in neoclassical models of growth. The American Economic Review, v. 85, n. 1, p. 103-115, 1995.
- BERNARD, A. B.; DURLAUF, S. N. Interpreting Test of the Convergence Hypothesis. National Bureau of Economic Research, n° 159, 1994.
- BURRIDGE, P. Testing for a common factor in a spatial autoregression model Environment and Planning A, v. 13, p. 795-800, 1981.
- DURLAUF, S. N. and JOHNSON, P. Multiple Regimes and Cross-country Growth Behaviour. Journal of Applied Econometrics, v. 10, n° 4, pp. 365-384, 1995.
- EMPRESA BRASILEIRA DE PESQUISAS AGROPECUÁRIAS (EMBRAPA). V Plano Diretor da Embrapa (2008-20011-2023). Brasília (DF), Embrapa, 2008, p.44.
- FIGUEIREDO, L. e COELHO, R. Uma Análise da Hipótese de Convergência para os Municípios Brasileiros. RBE. Rio de Janeiro, v.61, n.3/p.331-352, 2007.
- FERREIRA, A. H.; DINIZ, C. Convergência entre as rendas per capita estaduais no Brasil. Revista de Economia Política, v. 15, n. 4, p. 38-55, 1995.
- FLORAX, R.J.G.M.; FOLMER, H. and REY, S.J. Specification searches in spatial econometrics: The relevance of Hendry's methodology, Regional Science and Urban Econometrics, vol. 33, p. 557-579, 2003.
- FREITAS, R. E. e SPOLADOR, H. F. S. Os Termos de Troca para Soja na Agricultura Brasileira. Texto para Discussão N° 1239 – Instituto de Pesquisa Economica Aplicada (IPEA). Brasília – DF, 2006.
- FUJITA, M.; KRUGMAN, P.; and VENABLES, A. The spatial economy: cities, regions and international trade. MIT Press, Cambridge, 1999.
- GALOR, O. Convergence? Inferences from Theoretical Model. Economic Journal, vol. 106, p. 1056-1080, 1996.
- JOHNSON, P.A. and TAKEYAMA, L.N. Convergence among the u.s. states: Absolute, conditional on clubs? 2003. Disponível em: <http://irving.vassar.edu/vcewp/vcewp50.pdf>. Acesso em 8 de maio de 2009.
- HEWINGS, A.; MAGALHÃES, A.; AZZONI, C. Spatial dependence and regional inequality in Brazil. Investigaciones Regionales, v. 6, 2005.
- HOFFMANN, Rodolfo. Estatística para Economistas. São Paulo: Thomson, 4ª Ed., 2006.
- KELEJIAN, H. H., and PRUCHA, I. R., A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances, Journal of Real Estates Finance and Economics, 17: 99-121, 1998.
- LE GALLO, J and ERTHUR, C. Exploratory spatial data analysis of the distribution of regional per capita. GDP in Europe, 1980-1995. Papers in Regional Science, vol. 82, n. 2, p. 175-201. 2003.
- LeSAGE, J.P. The Theory and Practice of Spatial Econometrics. Department of economic – University of Toledo, 1999.
- LUCAS, R. E. On the mechanics of economic development. Journal of Monetary Economics, v. 22, p. 3-42, 1988.

- MACIEL, P.J., ANDRADE, J. e TELES, V.K. Convergência Regional Brasileira Revisitada. Pesquisa e Planejamento Econômico (PPE). Brasília (DF), v.38, n.01, 2008.
- MANKIW, G.; ROMER, D.; WEIL, D. A contribution to the empirics of economic growth. Quarterly Journal of Economics, v. 107, n. 429, p. 407-438, 1992.
- MORA, T. Evidencing European regional convergence clubs with optimal grouping criteria. Applied Economics Letters, 12(15): 937-940. 2005. Disponível em: <http://ideas.repec.org/a/tarf/apeclt/v12y2005i15p937-940.html>.
- MARSHALL, A. P. Principles of economics. New York: Macmillan, 1920.
- MOSSI, M.; AROCA, P.; FERNÁNDEZ, I.; AZZONI, C. Growth dynamics and space in Brazil. International Regional Science Review, v. 26, n. 3, p. 393-418, 2003.
- QUAH, D. T. Empirical cross-section dynamics in economic growth. European Economic Review, v. 37, p. 426-434, 1993.
- REY, S.J. and MONTOURI, B.D., U.S. regional income convergence: A spatial econometric perspective, Regional Studies, vol. 33: 143-156, 1999.
- ROMER, P. Increasing returns and long-run growth. Journal of Political Economy, v. 98, n.5, p. 1002-1037, 1986

Anexo

Quadro 3: Classificação dos dados das variáveis por amostra (OE, NO, SE, SO, SU, NE e AL), por regime espacial, por nome da UF e por número de municípios, no período de 1996 e 2006.

CLUSTER	REGIME ESPACIAL	NOMEUF	N	YVPAGR9606	YAPAGR9606	YPMAGR9606	PMAGR96	PMAGR06	PGCF06	MAE06	PGTR06	INVK06	AGLM06	PGSAL06	TECP06
NE	Alto-Alto	MARANHAO	1	8,806	5,158	3,648	1.602,97	2.308,71	1,7134	6,9510	4,123	75,590	1,7399	80,130	1,945
	Alto-Baixo	MARANHAO	3	10,274	15,905	-5,630	894,84	509,60	0,1425	4,5008	3,286	18,155	1,8547	23,142	1,644
	Baixo-Alto	MARANHAO	1	15,829	27,102	-11,273	358,72	116,19	0,1548	6,2654	0,033	5,280	1,8429	5,140	10,215
	Baixo-Baixo	MARANHAO	2	17,195	18,449	-1,254	617,95	545,10	0,1388	3,5899	0,114	4,363	1,9952	11,108	3,590
	Não Significante	MARANHAO	111	6,641	9,558	-2,916	803,90	600,55	0,6806	4,4219	8,268	41,015	1,8488	19,089	2,200
NE Total			118	7,156	9,981	-2,825	806,06	607,67	0,6620	4,4469	7,899	39,803	1,8504	19,456	2,275
NO	Alto-Alto	AMAPA-PARA	33	15,139	11,742	3,396	1.554,99	2.183,82	5,8355	5,6143	15,565	99,932	1,6145	27,961	2,063
	Baixo-Alto	AMAPA	2	-1,601	5,667	-7,269	624,20	301,75	0,3372	5,6683	100,000	97,325	1,0814	66,421	1,620
	Baixo-Alto	PARA	3	19,910	-19,356	39,265	343,88	17.445,47	1,0202	6,8860	44,357	1515,017	1,1335	33,653	0,886
	Não Significante	AMAPA	5	3,234	4,251	-1,017	1.401,08	1.265,56	0,2302	5,3129	40,695	9,046	1,3466	27,532	2,077
	Não Significante	PARA	53	11,692	10,142	1,550	1.151,62	1.344,65	4,5779	5,0135	8,374	41,587	1,7548	23,037	3,253
NO Total			96	14,472	9,370	5,102	1.267,04	2.110,42	4,5843	5,3078	15,562	107,154	1,6518	26,199	2,675
OE	Alto-Alto	AMAZONAS-ACRE	56	2,474	15,458	-12,984	2.412,52	658,52	0,6398	4,4567	8,200	32,260	1,7782	22,627	4,914
	Baixo-Alto	ACRE	1	14,628	13,790	0,838	595,79	647,86	1,9079	5,9261	0,622	69,591	1,6105	17,526	2,729
	Baixo-Alto	AMAZONAS	6	2,340	13,775	-11,435	500,17	159,40	1,0429	4,9662	35,531	65,858	1,6207	13,399	6,082
	Baixo-Baixo	RORAIMA	1	-28,113	5,604	-33,717	681,82	23,41	0,6870	5,2416	100,000	15,809	1,5181	11,965	-0,603
	Não Significante	ACRE	12	10,821	24,547	-13,726	1.210,21	306,73	0,6988	4,5233	9,564	40,869	1,7628	13,405	5,074
	Não Significante	AMAZONAS	9	-1,600	11,535	-13,135	2.874,92	772,99	0,4872	4,7515	15,064	63,066	1,8178	16,352	4,818
	Não Significante	RORAIMA	7	-4,090	6,612	-10,702	840,42	288,21	1,6357	5,4243	2,215	284,387	1,5907	15,872	0,380
OE Total			92	2,462	15,352	-12,890	2.018,04	556,09	0,7489	4,6256	11,292	57,998	1,7509	19,523	4,573
SE	Alto-Baixo	TOCANTINS	6	5,816	15,847	-10,031	928,29	340,45	12,6473	6,1185	1,106	242,023	1,4232	25,253	2,302
	Baixo-Baixo	MATO GROSSO	29	21,834	21,958	-0,124	381,79	377,08	8,0049	6,5327	23,099	1255,484	1,4939	15,334	1,597
	Baixo-Baixo	PARA	13	2,555	10,920	-8,365	428,90	185,81	8,0186	5,3779	3,821	738,214	1,6747	27,425	2,336
	Baixo-Baixo	TOCANTINS	42	11,583	16,804	-5,221	436,18	258,78	4,1359	5,4205	31,245	298,330	1,4088	18,099	1,969
	Não Significante	MATO GROSSO	31	16,525	17,804	-1,279	557,59	490,66	4,1147	6,7618	12,181	250,836	1,5706	12,768	1,760
	Não Significante	PARA	32	12,722	10,512	2,211	605,05	754,75	6,3461	5,2962	24,582	967,398	1,5396	20,410	2,203
	Não Significante	TOCANTINS	75	12,339	19,002	-6,663	537,65	276,15	32,8634	5,6458	29,986	595,454	1,2775	14,941	1,813
SE Total			228	13,992	17,075	-3,083	515,39	378,67	14,8305	5,8169	23,911	628,857	1,4323	17,028	1,904
SO	Alto-Baixo	RONDONIA	1	-26,067	3,698	-29,766	732,19	37,32	3,9622	5,1074	5,083	813,132	1,2141	33,489	-1,215
	Baixo-Alto	RONDONIA	1	-0,507	0,939	-1,446	408,03	353,10	2,5792	5,3425	2,202	129,193	1,8432	12,905	2,702
	Baixo-Baixo	RONDONIA	5	6,084	-2,621	8,705	441,32	1.053,88	6,8069	5,3768	1,879	126,308	1,6189	15,835	1,754
	Não Significante	RONDONIA	33	5,970	-0,422	6,393	573,59	1.087,01	2,9560	5,6119	6,478	138,367	1,5856	12,323	1,367
SO Total			40	5,670	-0,560	6,230	556,89	1.038,27	3,4531	5,5631	5,761	153,499	1,5869	13,305	1,384
SU	Alto-Baixo	MATO GROSSO	2	13,826	14,691	-0,865	750,09	687,94	0,3543	7,3937	4,590	8,561	1,3730	12,472	1,873
	Baixo-Baixo	MATO GROSSO	1	20,020	12,400	7,620	559,10	1.197,94	1,4544	7,2539	239,723	1658,801	1,4174	8,218	2,135
	Não Significante	MATO GROSSO	53	8,568	11,000	-2,432	813,24	637,69	14,5036	6,8085	27,547	733,284	1,3992	19,651	1,189
SU Total			56	8,992	11,157	-2,165	806,45	649,49	13,7652	6,8374	30,516	723,928	1,3986	19,190	1,230
AMAZÔNIA LEGAL			630	10,856	12,675	-1,818	932,31	777,31	7,7420	5,3834	17,231	333,935	1,5974	19,200	2,388

Fonte: CENSO AGROPECUÁRIO 1995-96 e 2006 - IBGE.

DESCRIÇÃO DAS VARIÁVEIS - ELABORADO PELOS AUTORES.

DENOTAÇÃO	DESCRIÇÃO DA VARIÁVEL	UNIDADE
YVPAGR9606	Taxa de crescimento anual do valor da produção agrícola entre 1996 a 2006	%
YAPAGR9606	Taxa de crescimento anual da área de cultivo entre 1996 a 2006	%
YPMAGR9606	Taxa de crescimento anual da produtividade agrícola entre 1996 a 2006.	%
PMAGR96	Produtividade média agrícola medida no ano de 1996.	R\$/há
PMAGR06	Produtividade média agrícola medida no ano de 2006.	R\$/há
PGCF06	Proporção de gastos c/capital físico em relação ao valor da produção 2006.	%

Nota: elaborado pelos autores.

DESCRIÇÃO DAS VARIÁVEIS - ELABORADO PELOS AUTORES.

DENOTAÇÃO	DESCRIÇÃO DA VARIÁVEL	UNIDADE
MAE06	Média de anos de estudo do produtor em 2006.	ANO
PGTR06	Proporção de gastos com transporte da produção agrícola em 2006.	%
INVK06	Proporção do investimento realizado via crédito em termo valor da produção em 2006.	%
AGLM06	Índice de aglomeração de unidades produtivas - índice T de Theil no ano de 2006.	bit
PGSAL06	Proporção de gastos com salário em relação ao gasto total do produtor no ano de 2006.	%
TECP06	Variação da tecnologia, taxa de crescimento do pessoal ocupado e depreciação do capital.	%

Nota: elaborado pelos autores.

Figura 2: Distribuição espacial da produtividade agrícola entre municípios (630) da Amazônia Legal nos de 1996 e 2006

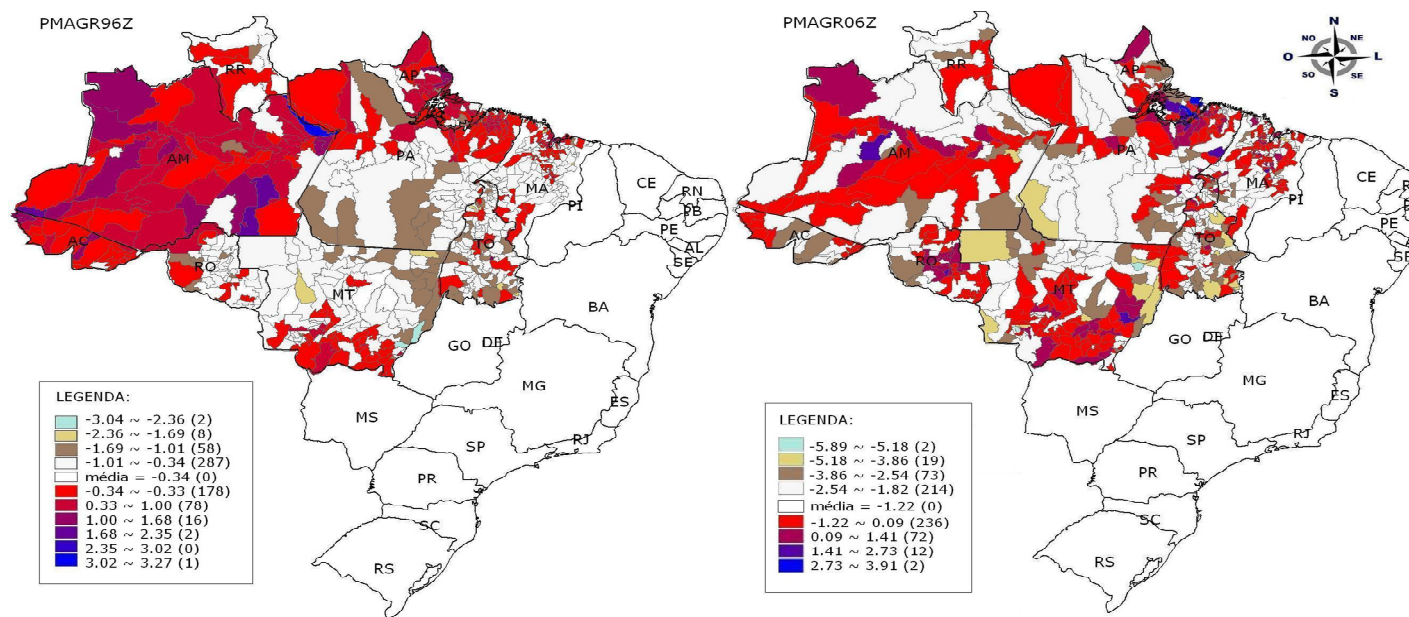


Figura 4: Estatística LISA aplicada a produtividade agrícola dos municípios da Amazônia Legal nos anos de 1996 e 2006.

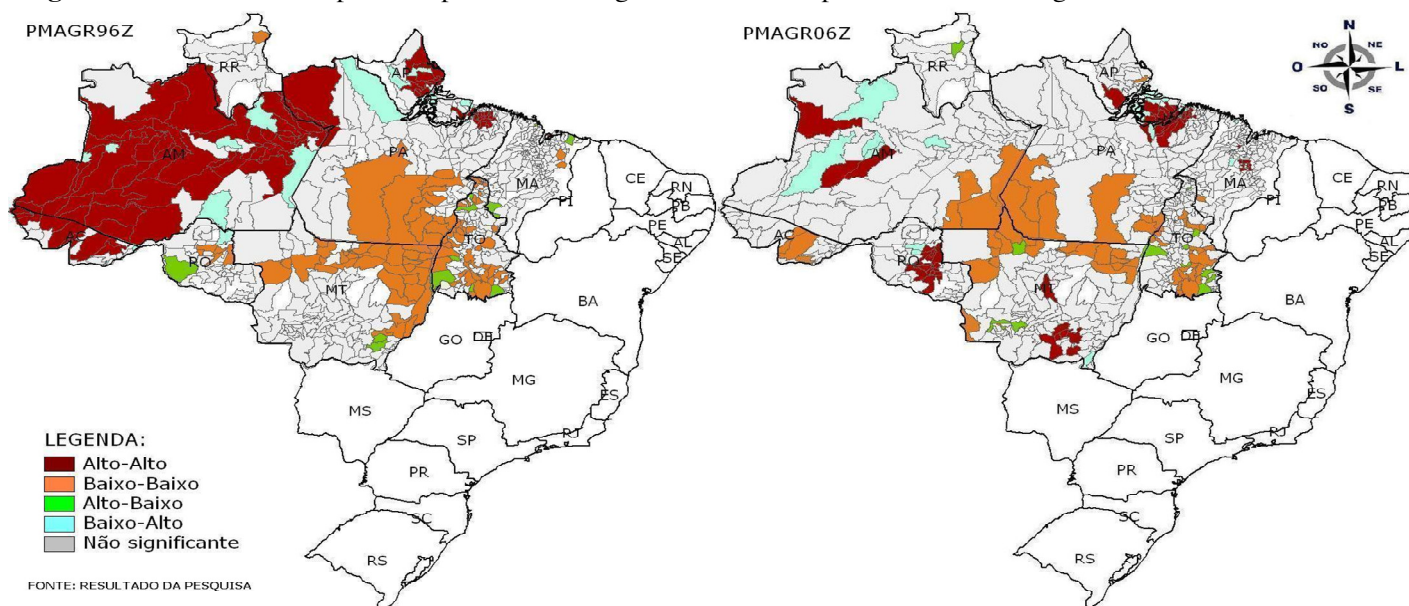


Figura 1: Regra de decisão para modelos espaciais

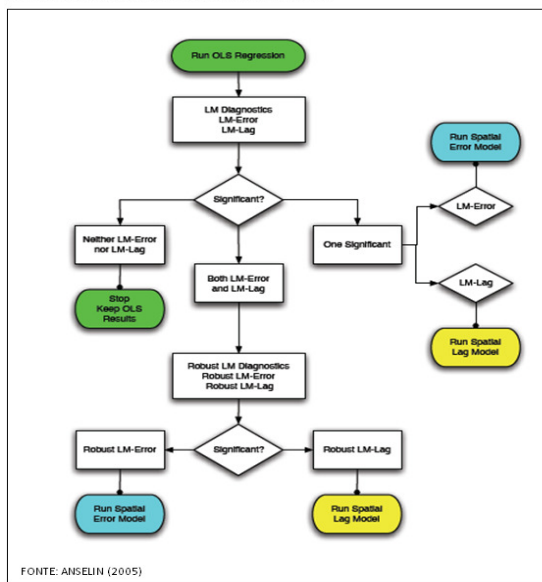


Figura 3: Estatística I de Moran global e diagrama de dispersão (Scatter Plot) aplicado a produtividade média agrícola dos municípios da Amazônia Legal nos anos de 1996 e 2006.

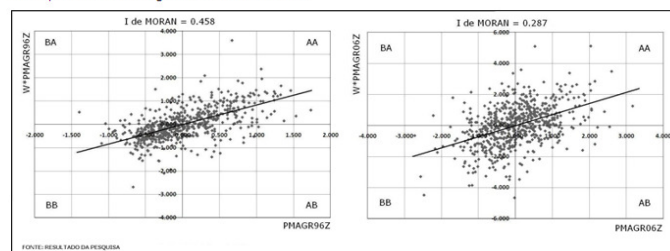


Figura 5: Subdivisão da amostra Amazônia Legal, com base relação de dependência espacial dos desvios da produtividade média agrícola

