

TEXTO PARA DISCUSSÃO Nº 707

**HETEROGENEIDADE ESPACIAL DA
PRODUTIVIDADE NA AGROPECUÁRIA:
BRASIL — 1970/96***

Ajax Moreira**
Hélio Migon***

Rio de Janeiro, fevereiro de 2000

* Os autores agradecem ao programa Pronex pelo suporte financeiro, a Ingreed Valdez e a Marina Paes pelo apoio na elaboração desta pesquisa, a Marcia Pimentel pelo apoio na construção do sistema geográfico e as sugestões de Steven Helfand. Naturalmente os erros são de exclusiva responsabilidade dos autores.

** Da Diretoria de Estudos Macroeconômicos do IPEA.

*** Do Instituto de Matemática e Estatística da UFRJ e da Coppe.

MINISTÉRIO DO PLANEJAMENTO, ORÇAMENTO E GESTÃO

Martus Tavares - Ministro

Guilherme Dias - Secretário Executivo



Presidente

Roberto Borges Martins

Diretoria

Eustáquio J. Reis

Gustavo Maia Gomes

Hubimaier Cantuária Santiago

Luís Fernando Tironi

Murilo Lôbo

Ricardo Paes de Barros

Fundação pública vinculada ao Ministério do Planejamento Orçamento e Gestão, o IPEA fornece suporte técnico e institucional às ações governamentais e disponibiliza, para a sociedade, elementos necessários ao conhecimento e à solução dos problemas econômicos e sociais dos países. Inúmeras políticas públicas e programas de desenvolvimento brasileiro são formulados a partir de estudos e pesquisas realizados pelas equipes de especialistas do IPEA.

TEXTO PARA DISCUSSÃO tem o objetivo de divulgar resultados de estudos desenvolvidos direta ou indiretamente pelo IPEA, bem como trabalhos considerados de relevância para disseminação pelo Instituto, para informar profissionais especializados e colher sugestões.

ISSN 1415-4765

SERVIÇO EDITORIAL

Rio de Janeiro – RJ

Av. Presidente Antônio Carlos, 51 – 14º andar – CEP 20020-010

Telefax: (21) 220-5533

E-mail: editrj@ipea.gov.br

Brasília – DF

SBS Q. 1 Bl. J, Ed. BNDES – 10º andar – CEP 70076-900

Telefax: (61) 315-5314

E-mail: editbsb@ipea.gov.br

© IPEA, 1998

É permitida a reprodução deste texto, desde que obrigatoriamente citada a fonte.

Reproduções para fins comerciais são rigorosamente proibidas.

SUMÁRIO

RESUMO

ABSTRACT

1 - INTRODUÇÃO	1
2 - ESPECIFICAÇÃO DA FUNÇÃO DE PRODUÇÃO.....	3
3 - MODELOS ADOTADOS.....	7
3.1 - Modelo com Parâmetro Fixo	7
3.2 - Modelo com Parâmetro Variando Espacialmente	8
4 - CRITÉRIO DE COMPARAÇÃO DOS MODELOS.....	12
5 - RESULTADOS	14
6 - CONCLUSÃO	20
BIBLIOGRAFIA	25

RESUMO

A produtividade dos fatores de produção na exploração agropecuária depende de muitos aspectos que não são observáveis ou mensurados. Esses aspectos são desigualmente distribuídos entre as regiões, o que implica a heterogeneidade da função de produção agropecuária.

Este trabalho utiliza um modelo que estima elasticidades dos fatores específicas para cada região, pressupondo que estas elasticidades apresentam uma tendência espacial, ou seja, que se modificam de forma suave no espaço. Essa característica é representada pelo princípio de similaridade com o mais próximo, ou seja, com as suas regiões vizinhas. Este modelo é uma generalização, para o contexto espacial, dos modelos de série temporal com parâmetro variando, onde a similaridade se dá segundo a seqüência temporal.

Este modelo foi estimado utilizando, separada e conjuntamente, as informações dos cinco censos agropecuários. O resultado obtido é uma descrição quantitativa da heterogeneidade da produtividade entre as microrregiões.

ABSTRACT

Factor productivity in agricultural activity depends on many non-observed or non-measured aspects. The intensity distribution of these elements are unequally distributed between the regions, which accounts for the heterogeneity of the production function.

This study uses a model which estimates specific elasticity for each region, supposing that it follows a spatial trend. This trend is represented by similarity principles between neighbourhood regions. This is a generalisation, for spatial models, of time series models with floating parameters.

This model was estimated using five information sets from the last agricultural census made in Brazil, separately and jointly. The results describe the productivity heterogeneity quantitatively.

1 - INTRODUÇÃO

A produtividade dos fatores de produção na exploração agropecuária depende de muitos aspectos que não são observáveis ou mensurados. Exemplos disso são os aspectos físicos, como a qualidade do solo, a incidência de chuvas e o grau de exaustão do solo, ou aspectos culturais, como o grau de instrução do pessoal ocupado no nível da execução ou do gerenciamento e a proporção de empresários que foi exposta ou que utiliza tecnologias mais eficientes. Esses aspectos são desigualmente distribuídas entre as regiões, o que implica a heterogeneidade regional da função de produção agropecuária.

Esses aspectos não-observáveis tendem a se alterar de forma suave no espaço, ou seja, apresentam uma tendência espacial. Exceto pela presença, pouco freqüente, de um obstáculo natural, os aspectos físicos e culturais são mais parecidos entre regiões próximas do que entre regiões distantes, seja porque a natureza tende a não dar saltos, ou porque os elementos culturais se difundem por contágio. Por exemplo, o regime de chuvas de uma região é, em geral, semelhante ao da região vizinha, ou a utilização de uma tecnologia agrícola ocorre em bolsões regionais.

No caso em que os dados são agregados por região, por exemplo, por microrregião homogênea (MR),¹ fica implícita a hipótese de homogeneidade entre os produtores de cada região, e o procedimento padrão para estimar a função de produção supõe que os seus parâmetros — elasticidades produto-fator, no caso de uma Cobb-Douglas — são iguais para todas as regiões. Ou seja, admite-se a hipótese de homogeneidade intra-regional e de igualdade inter-regional dos parâmetros, restrições suficientemente fortes para tornar questionável o resultado obtido. Uma abordagem alternativa seria utilizar os dados de cada produtor rural ou de subagregados de cada região para estimar uma função de produção para cada região.² Essa abordagem não considera a proximidade entre as MRs e utiliza um conjunto de informações diferente, portanto a influência das características das regiões vizinhas não afeta a estimativa da produtividade dos fatores das regiões, e os resultados não são diretamente comparáveis.

O modelo de parâmetros fixos com estrutura espacial descritos na Seção 3 [ver Anselin (1988)] utiliza as relações de vizinhança para eliminar a eventual correlação dos resíduos ou para introduzir como variável explicativa a média local da variável dependente. Esse modelo considera de alguma forma os aspectos não-observáveis mas, supondo os coeficientes iguais para todas as MRs, não tem como dar conta da heterogeneidade.

No caso em que a amostra é uma série temporal, em que a noção de similaridade é dada pelo encadeamento temporal, cada período está ligado aos períodos anterior e

¹ Microrregião homogênea é um critério de partição do espaço geográfico brasileiro adotado pelo IBGE que, em 1997, compreendia 558 regiões.

² No caso em que a região é uma microrregião homogênea, seria possível utilizar os dados dos municípios sempre que o número de municípios na microrregião fosse suficiente para a estimação do modelo.

posterior; os modelos de parâmetro variando de espaço ou de estado podem considerar os coeficientes dos regressores como variáveis cujo valor segue um passeio aleatório no tempo — o desvio em relação ao período anterior é estocástico — e cujo grau de variabilidade é controlado por variâncias que devem ser estimadas. Nos modelos desse tipo, os coeficientes dos regressores refletem características locais no sentido temporal.

Na literatura de análise da incidência espacial de doenças, autores como Knorr-Held e Besag (1998) transpõem para o contexto espacial o princípio de similaridade com o mais próximo utilizado pelos modelos de série temporal. Nesse modelo, os coeficientes dos regressores são definidos supondo que o desvio em relação ao coeficiente das regiões vizinhas é estocástico. Assim, ele pode estimar coeficientes específicos por região, mantendo o mesmo conjunto de informações de dados agregados regionalmente, e considerar a tendência espacial dos elementos não-observáveis que, supõe-se, estão refletidos na tendência espacial dos coeficientes do modelo.

Esse modelo será utilizado para estimar as elasticidades dos fatores de produção específicas para cada região, admitindo que a diferença entre as elasticidades das regiões vizinhas é estocástica, o que dá conta da diversidade das condições regionais e da tendência espacial dos aspectos não-observáveis.

Esse tipo de modelo foi utilizado pioneiramente por Assunção, Lemos e Assunção (1998) para analisar dados econômicos no Brasil, que estimaram a indução de alteração tecnológica na agricultura usando dados do Censo de 1985. A utilização desse tipo de modelo para avaliar a função de produção agropecuária suscita algumas questões: *a)* Serão os coeficientes das regiões realmente diferentes entre si, ou seja, em que medida é necessário considerar um modelo com parâmetros variando para estimar uma função de produção agropecuária para o Brasil? *b)* A distribuição da produtividade para os diferentes censos agropecuários é estável?³ *c)* Qual o aumento de precisão das estimativas devido à utilização de mais de uma observação por MR, o que é possibilitado pelo uso simultâneo de todos os censos? *d)* Qual o efeito de amostrar conjuntamente os coeficientes de todos os regressores de cada região, em vez de um esquema que sorteia os coeficientes por região e por regressor, como adotado no artigo citado?

Para discutir essas questões, consideramos, além do modelo de parâmetros variando, o modelo de parâmetros fixos com estrutura espacial. Os dois são estimados utilizando uma mesma abordagem estatística — Markov Chain Monte Carlo (MCMC) [ver Gamerman (1997)] — o que possibilita a utilização de critérios de comparação entre eles. O primeiro é uma extensão do modelo de parâmetros variando proposto por Assunção, Lemos e Assunção (1998), que considera conjuntamente as dimensões espacial e temporal e utiliza um outro esquema de amostragem. O outro é uma versão do modelo proposto por Anselin (1988) que, apesar de não admitir a heterogeneidade dos parâmetros, inclui, ainda

³ Para os anos de 1970, 1975, 1980, 1985 e 1996.

que de forma indireta, o efeito da vizinhança, e pode ser um substituto potencial do primeiro modelo.

A Seção 2 discute a especificação da função de produção, enquanto as Seções 3 e 4 apresentam, respectivamente, o modelo estatístico e o critério de comparação de modelos adotado. Finalmente, a Seção 5 apresenta os resultados obtidos e a Seção 6 as conclusões.

2 - ESPECIFICAÇÃO DA FUNÇÃO DE PRODUÇÃO

Este artigo focaliza a diversidade regional das tecnologias de produção. A complexidade da estimação do modelo com parâmetros variando recomenda, pelo menos numa primeira aproximação, que esta seja linear nos parâmetros. Uma forma funcional que atende a essa condição e que é freqüentemente utilizada na literatura pertinente é a Cobb-Douglas, que relaciona linearmente os logaritmos dos fatores de produção e do produto.

Os fatores de produção mão-de-obra e capital físico e humano, variáveis explicativas da função de produção, apresentam dificuldades específicas de mensuração. O capital humano é habitualmente considerado na literatura por medidas do grau de instrução dos trabalhadores envolvidos, como a escolaridade média dos trabalhadores ocupados na produção, e o capital físico é usualmente medido por um índice de *quantum* construído acumulando adequadamente os investimentos realizados.

Não existem, no nível regional, medidas da escolaridade da população ocupada nas atividades agropecuárias, ou do nível de investimento — medido em *quantum* — realizado em cada MR e em cada ano. De fato, mesmo as medidas de pessoal ocupado na atividade rural são imprecisas. As pessoas ocupadas podem envolver familiares cuja intensidade do trabalho é muito diferente, por exemplo, daquela dos trabalhadores temporários. Também não dispomos de informação sobre o grau de envolvimento das pessoas ocupadas. Por isso optamos pelo tratamento simplista de considerar apenas o número de pessoas ocupadas na atividade rural cujo logaritmo recebe o mnemônico (x_1).

Não é possível medir o estoque de capital envolvido na exploração agropecuária a partir da acumulação do valor dos investimentos realizados, seja porque medidas de valor monetário em uma economia que, no período em estudo, apresentou taxas elevadas e instáveis de inflação são precárias, ou porque só dispomos desse valor para os anos dos censos agropecuários. Gastos com investimentos medidos por agregados anuais deflacionados são uma medida extremamente imprecisa do investimento real, porque são distribuídos ao longo do ano sem um padrão esperado, o que torna impreciso o cálculo do deflator, particularmente para as altas taxas de inflação ocorridas. Dados disponíveis apenas para anos selecionados impedem a acumulação dos investimentos de forma coerente. Esses motivos nos levaram a medir o capital envolvido na produção agrícola por meio de indicadores que consideram quantidades de alguns tipos de capital.

Foram considerados quatro tipos de capital, a terra medida pela área utilizada (a) pelos estabelecimentos agropecuários cujo logaritmo recebe o mnemônico (x_2), e medidas de estoque de capital típicas do tipo de exploração agropecuária. No caso de cultura permanente, para alguns tipos de lavoura⁴ foi utilizada a soma do número de pés plantados (np); no caso de culturas temporárias foi utilizado número de tratores (nt); e no caso de pecuária foi adotado o número de bovinos (nb).⁵

As medidas do fator capital (nb , np , nt) assumem o valor zero para algumas das MRs em alguns censos, o que impede a sua utilização.⁶ Por outro lado, seria interessante distinguir de alguma forma a produção de subsistência da exploração destinada ao mercado. Esses dois aspectos foram considerados definindo uma densidade mínima de utilização desses fatores, e anulando o efeito dessas medidas nas MRs cuja densidade for inferior ao mínimo. Como densidade mínima foi escolhida a densidade da MR mediana. Esse critério corresponde a dar pesos iguais aos dois tipos de erro de classificação, considerar a MR como tendo o atributo — por exemplo, pecuária — quando de fato ela não é uma região de pecuária, e classificá-la como região de pecuária quando de fato não é.

Esse procedimento equivale à multiplicação de cada variável por uma variável indicadora que anula o efeito sempre que a densidade for inferior ao mínimo requerido. A utilização dessas variáveis indicadoras corresponde à especificação de uma função de produção para cada categoria de MR. A categoria é caracterizada pela existência de cada um dos tipos de exploração agropecuária, pela inexistência dos tipos, pela existência de todos os tipos, ou por alguma combinação dos tipos. A Tabela 1 apresenta, para cada ano, o número de MR e os percentis selecionados da distribuição das MRs segundo a intensidade de cada um dos tipos de capital.

Tabela 1

Distribuição entre as MRs da Razão por Tipo de Capital

Censo	MR	Bovinos por Área					Número de Pés por Área					Número de Tratores				
		5	25	50	75	95	5	25	50	75	95	5	25	50	75	95
1970	538	0,3	1,3	2,4	3,9	60	0,0	0,0	0,3	3,7	64,8	0	0	2	10	42
1975	538	0,3	1,5	2,9	4,6	72	0,0	0,3	1,0	8,3	72,6	0	1	4	19	75
1980	548	0,5	1,6	3,2	5,0	77	0,0	0,3	1,6	13,4	103,0	3	10	25	87	254
1985	553	0,6	1,8	3,2	5,0	78	0,0	0,3	1,8	15,3	101,6	2	10	30	107	362
1996	558	1,0	2,4	4,1	6,2	90	0,1	0,5	2,7	16,7	126,0	3	13	36	115	355

As variáveis referentes aos capitais específicos (x_3 = número de tratores, x_4 = intensidade de pecuária e x_5 = intensidade de lavoura permanente) foram construídas utilizando a função indicadora $I(.,.)$ ⁷ da seguinte forma:

⁴ Lavouras de café, cacau, laranja e banana.

⁵ Na falta de medidas confiáveis de reprodutores e matrizes fêmeas.

⁶ O logaritmo de zero não se define.

⁷ $I(a, b)$ é igual a 1 se $a > b$ e igual a zero em caso contrário.

$$x_3 = \text{Log}(nt) I(nt, 1) \quad x_4 = \text{Log}(nb) I(nb > 4,1a) \quad x_5 = \text{Log}(np) I(np, 2,7a)$$

Para estimar uma função de produção agregada para o conjunto das atividades da agropecuária, é necessária uma única medida do produto da MR. Numa primeira aproximação, foi adotado como medida do produto o valor da produção agregada da MR deflacionada pelo índice do INPC médio do ano, cujo logaritmo recebe o mnemônico (y). Essa medida tem as vantagens da simplicidade e cobertura, mas apresenta, ainda que de forma atenuada, os problemas decorrentes da agregação de variáveis anuais. De fato, o momento da geração da renda depende das safras dos produtos produzidos na MR, o que implica um padrão de deflacionamento intra-anual típico da MR.

Supondo-se que esse padrão é estável para todos os anos da amostra, fica definido um coeficiente para cada MR que corrige a desigualdade da distribuição da produção ao longo dos meses. Esse coeficiente, que é específico para cada MR, pode ser atribuído à constante específica para cada MR no modelo de parâmetros variando.

Uma medida alternativa seria o índice do *quantum* da produção agropecuária, construído a partir das quantidades de cada um dos itens produzidos na MR. Esse índice mede a produtividade dos fatores — relação entre a quantidade produzida e a intensidade de uso do fator —, ao contrário do valor da produção deflacionada, que mede o aumento de renda do produtor por unidade adicional de fator. O índice de *quantum* é menos afetado pela variação interanual e inter-regional dos preços relativos. Apesar dessa característica desejável, o índice não foi adotado nesta versão do texto, já que envolve a coleta e depuração de um número muito maior de informações. Ainda que as elasticidades estimadas não sejam propriamente medidas de produtividade, serão interpretadas como tal ao longo do texto.

As Tabelas 1 e 2 mostram que, pelo menos para 99% das MRs, os índices construídos são estáveis e têm uma tendência esperada.⁸ A Tabela 2 mostra a distribuição das MRs segundo as razões entre a área total ocupada pelos estabelecimentos e a área total da MR; o número de pessoas ocupadas e a área ocupada pelos estabelecimentos; e entre o valor da produção e a área ocupada. Apenas 1% das MRs apresentam taxa de ocupação maior do que a unidade. Naturalmente, as medidas referentes a essas MRs podem estar sujeitas a erros. No entanto, esse erro de medida será ignorado devido a sua pequena expressão no total da amostra.

⁸ Densidade aumentando com o tempo.

Tabela 2

Distribuição entre as MRs da Intensidade de Produção e Ocupação

Censo	MR	Produção R\$ 1.000/Área					Pessoal Ocupado/Área					Área Utilizada/Área MR				
		5	25	50	75	95	5	25	50	75	95	5	25	50	75	99
1970	538	0,01	0,02	0,05	0,10	0,20	0,01	0,04	0,08	0,14	0,28	0,07	0,50	0,76	0,88	1,05
1975	538	0,01	0,03	0,07	0,15	0,33	0,01	0,05	0,08	0,15	0,33	0,10	0,55	0,78	0,89	1,03
1980	548	0,01	0,03	0,09	0,19	0,41	0,01	0,04	0,08	0,14	0,27	0,13	0,57	0,79	0,90	1,06
1985	553	0,01	0,04	0,10	0,25	0,69	0,01	0,04	0,09	0,16	0,42	0,14	0,56	0,80	0,89	1,03
1996	558	0,02	0,06	0,15	0,35	0,81	0,01	0,04	0,08	0,14	0,29	0,08	0,47	0,68	0,82	0,95

O modelo proposto é, então:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \varepsilon = X\beta + \varepsilon \quad \varepsilon \sim (0, \sigma^2) \quad (1)$$

Na estimativa de uma função de produção típica, é necessário considerar que os fatores de produção são escolhidos em conjunto com o montante que será produzido, o que implica variáveis explicativas endógenas. Como consequência, temos que os estimadores de mínimos quadrados são viesados e inconsistentes. Nesse caso seria necessária a utilização de estimadores sistêmicos ou de variáveis instrumentais.

No nosso caso, o que está sendo estimado não é uma função de produção no sentido estrito, em que os agentes decidem em cada momento uma quantidade de fatores para obter uma quantidade de produto. De fato, os fatores pessoal ocupado e área utilizada são dotações de recursos que os agentes dispõem para escolher o montante produzido. A quantidade dos fatores de capital — máquinas, bovinos e número de pés de cultura permanente — não é necessariamente decidida ao mesmo tempo que a produção. Assim, consideramos exógenas todas as variáveis explicativas do modelo.⁹

Na literatura de economia agrícola, diversos autores propõem ampliar o conceito de função de produção incorporando como variáveis explicativas alguns dos aspectos não mensurados ou de difícil mensuração mencionados anteriormente. Como já foi dito, a abordagem de parâmetros variando espacialmente (PV) utilizada neste artigo considera esses aspectos de forma implícita, desde que estejam relacionados espacialmente. Assim, o modelo estimado tem, pelo menos de forma implícita, a mesma diversidade de aspectos dos outros modelos de parâmetro fixo (PF) mencionados. Naturalmente, a comparação dos resultados dessas duas abordagens é interessante.

O modelo em que os parâmetros podem variar espacialmente permite que sejam especificados como regressores apenas aquelas variáveis para as quais se deseja estimar o seu efeito. As demais variáveis de controle podem ser dispensadas à

⁹ Essa argumentação procura justificar a não-adoção de outros métodos, que tornariam mais complexos os procedimentos de estimação.

medida que estiverem correlacionadas espacialmente, pois estarão sendo consideradas pela estrutura espacial do modelo. Portanto, a exclusão de uma variável explicativa deve afetar pouco os resultados caso ela esteja correlacionada espacialmente.

No modelo (1), o intercepto (β_0) mede habitualmente a produtividade total dos fatores. No nosso caso, esse parâmetro não mede essa produtividade devido à forma como o modelo foi estimado.¹⁰ A soma dos parâmetros ($\phi = \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5$) é uma medida do grau de economia de escala da tecnologia de produção adotada em cada MR. Esse parâmetro pode ser visto como uma medida do grau de eficiência da tecnologia de produção adotada na MR ou de modernidade das técnicas de produção adotadas.

3 - MODELOS ADOTADOS

Esta seção discute os modelos estatísticos adotados, que serão apresentados com a notação do modelo linear geral. Em todos os casos, a informação espacial é introduzida por meio da matriz de vizinhança ($W = [w_{ij}]$), onde $w_{ij} = 1/n_i$ se j é vizinho de i e $w_{ij} = 0$ em caso contrário, e $n_i = \sum_j w_{ij}$. Vale mencionar que o modelo estatístico suporta outras especificações para (W), como por exemplo o efeito da distância entre os vizinhos, ou o efeito do tamanho da região vizinha.

3.1 - Modelo com Parâmetro Fixo

Os modelos de parâmetro fixo (PF) que consideram a informação contida em (W) foram propostos por Anselin (1988) e podem ter duas especificações. Na primeira admite-se que a variável dependente é auto-regressiva espacialmente e, na segunda, que o resíduo é auto-regressivo. Essas duas formas básicas são analogias diretas dos modelos AR e MA das séries temporais. Os modelos são:¹¹

$$y_r = \rho W_r y + X_r \beta + e_r \quad e_r \sim N(0, \sigma^2) \quad (2)$$

$$y_r = X_r \beta + u_r \quad u_r = \rho W_r u + e_r \quad e_r \sim N(0, \sigma^2) \quad (3)$$

Nos dois casos, o modelo é completado com as prioris:

$$p(\beta) \sim N(0, B) \quad p(\rho) \sim U(-1, 1)$$

¹⁰ A constante mede também o efeito do corretor da desigualdade da distribuição das safras. Naturalmente, essa deficiência não ocorre no caso de a variável explicada ser o índice de *quantum*.

¹¹ Na notação adotada, W_r corresponde à linha (r) de W e y_t indica as observações de todas as regiões para o período t . Variáveis indicadoras temporais podem ser incluídas em (X).

Conhecido o valor de (ρ) , transformações adequadas¹² fazem com que esses modelos possam ser estimados por mínimos quadrados, nos quais o logaritmo da verossimilhança marginal (LVM) desses modelos tem a seguinte expressão:

$$\text{LVM}(y_1, \dots, y_T | \rho) \propto -0,5 * \text{Lg}(s^2(\rho)) + \text{Lg}(|A(\rho)|) \quad (6)$$

onde $A(\rho) = (I - \rho W)$ e $s^2(\rho)$ é a soma dos quadrados das estimativas dos resíduos ($\hat{e} | \rho$).

Anselin propõe estimar (ρ) por máxima verossimilhança. Neste artigo utilizamos o método MCMC/Metropolis-Hastings [ver Gamerman (1997)], que pode ser aplicado sempre que o valor da LVM(ρ) é conhecido [ver Gamerman e Moreira (1998)]. Nesse método obtém-se uma amostra da distribuição *a posteriori* de $(\rho | y_1, \dots, y_n)$, e a distribuição de todos os parâmetros ou estatísticas pertinentes pode ser obtida por simulação. Vale notar que a forma (3) é equivalente a um caso particular do modelo de parâmetros variando apresentado na Subseção 3.2.¹³

3.2 - Modelo com Parâmetro Variando Espacialmente

Em um modelo de regressão múltipla ($y_r = x_r b + e_r$) em que os dados estão organizados espacialmente e os parâmetros (b) variam com cada região existem infinitas combinações de valores de (b) que reproduzem (y) com exatidão, e portanto o modelo não é identificável e muito menos estimável. A noção de similaridade espacial introduz uma restrição probabilística sobre os (b) de forma a tornar o modelo identificável e estimável. Essa noção é a semelhança entre coeficientes das regiões vizinhas, ($b_r \approx b_k$) $\forall k \sim r$, ou seja, (b_r) e (b_k) tendem a ser parecidos para toda região k vizinha da região r .

No caso em que os dados estão organizados com uma segunda dimensão, por exemplo o tempo, dispõe-se de mais do que uma observação por região, e o modelo é identificável desde que a diferença entre o número de parâmetros e o de observações incluídas seja positiva. Mesmo que essa diferença seja positiva, a noção de similaridade espacial é útil para melhorar a qualidade das estimativas, e para produzir estimativas de (b) que variem suavemente no espaço, o que é a hipótese básica deste modelo.

Naturalmente, a segunda dimensão possibilita a análise de problemas em que os coeficientes variam no tempo e no espaço, ou seja, quando se deseja estimar (b_{rt}), motivando a introdução de similaridade nas duas dimensões: no espaço entre

¹² No modelo (2) temos: $y - \rho W y = (I - \rho W)y = A(\rho)y = X\beta + e$.

No (3): $u = \rho W u + e \Rightarrow (I - \rho W)u = A(\rho)u = e \Rightarrow u = A(\rho)^{-1}e$, $y = X\beta + u = X\beta + A(\rho)^{-1}e \Rightarrow A(\rho)y = A(\rho)X\beta + e$.

¹³ No (PF) faça $(\rho = 1)$: $y_r = ax_r + u_r$, $u_r = Wu + e_r$, $e_r \sim (0, \sigma^2) \Rightarrow y_r = ax + [(I - W)^{-1}e]_r$; no (PV) faça o intercepto com estrutura espacial: $y_r = ax_r + b_r + e_r$, $e_r \sim (0, \sigma^2)$, $b_r = Wb + w_r$, $w_r \sim (0, s) \Rightarrow y_r = ax + [(I - W)^{-1}w]_r + e_r$ mostrando que o (PF) e o (PV) diferem por um termo estocástico não-correlacionado. Propriedade que não vale quando temos mais de um parâmetro flutuando ou parâmetros com regressores.

vizinhos e no tempo segundo o encadeamento temporal. Não analisaremos aqui simultaneamente a flutuação dos parâmetros nas duas dimensões, portanto consideraremos que alguns parâmetros flutuam no espaço e que os demais ou são constantes ou flutuam apenas no tempo.

Assim, se $(y_{rt} | \mu_{rt}, \sigma^2) \sim N(\mu_{rt}, \sigma^2)$ são observações independentes condicionais à média (μ_{rt}) e a variância (σ^2), onde r se refere às (n) regiões e t se refere por exemplo aos (T) períodos de tempo, admite-se que $(\mu_{rt} = x_{rt}\beta_r + z_{rt}\alpha)$ é determinado por dois blocos de componentes, o primeiro contendo os regressores (x_{rt}) com dimensão (Txp) — que têm coeficientes que flutuam no espaço segundo a restrição de similaridade espacial — e o segundo com os demais regressores (z_{rt}) com dimensão (Txq) .

O segundo bloco de regressores pode estar associado a parâmetros constantes, ou que variam com o tempo.¹⁴ Característico deste bloco é a falta de informação *a priori* sobre (α). Essa especificação possibilita que um mesmo regressor participe dos dois blocos, podendo portanto considerar os efeitos parciais das duas dimensões. Por exemplo, o regressor (x), que participa do primeiro bloco e do segundo variando com o tempo. Como $\{x_{rt}\beta_r + x_{rt}\alpha_t = x_{rt}(\beta_r + \alpha_t)\}$, que é uma representação de (b_{rt}) por meio dos efeitos parciais.

Para discutir o segundo bloco e introduzir a notação, vamos apresentá-lo num caso particular que será estendido posteriormente. Suponha que $(y_r | \beta_r, x_r, \sigma^2) \sim N(x_r\beta_r, \sigma^2)$, onde (x_r) é um vetor $(1xp)$ de regressores, e (β_r) um vetor de coeficientes da mesma dimensão. Admitindo que dispomos de dados para as (n) regiões condicionalmente independentes e com variância constante (σ^2), temos a função de verossimilhança:

$$L(\beta_1, \dots, \beta_r, \sigma^2 | y, x) \propto (\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_r (y_r - x_r\beta_r)^2\right\} \quad (7)$$

A informação espacial será introduzida por meio de uma *priori* para os coeficientes, que admite que os desvios em relação à média local dos coeficientes é estocástica, ou seja:

$$p(\beta_r | \beta_{-r}, \Sigma) \propto \exp\{-0,5 a_r (\beta_r - \bar{\beta}_r)' \Sigma^{-1} (\beta_r - \bar{\beta}_r)\} \quad (8)$$

onde (β_{-r}) indica os coeficientes de todas as regiões exceto (r), $\bar{\beta}_r = \sum_j w_{rj} \beta_j$ é a média local do coeficiente que depende de (β_{-r}), ($a_r > 0$) faz o papel de tamanho da amostra e (w_{rj}) é um peso refletindo a influência da região (j) sobre (r). Pode-se mostrar, como o fazem Waller *et alii* (1997) e Gamerman (1999), que a

¹⁴ A matriz (Z) pode conter regressores multiplicados por variáveis indicadoras do período de tempo, o que permite gerar uma estimativa do coeficiente para cada período de tempo da amostra.

distribuição conjunta de $\beta = \text{vec}(\beta_1, \dots, \beta_n)$ é dada por duas representações equivalentes:

$$P(\beta | \Sigma) \propto \exp\{-0,5\beta'(B \otimes \Sigma^{-1})\beta\} = \exp\{-0,5 \sum_{j < r} w_{jr}(\beta_j - \beta_r)' \Sigma^{-1}(\beta_j - \beta_r)\} \quad (9)$$

onde $(B_{ii} = a_i)$ e $(B_{ij} = a_i w_{ij})$. Assim, $\beta \sim N(0, B^{-1} \otimes \Sigma)$ é uma normal (np) variada, com densidade própria de média zero e variância $(B^{-1} \otimes \Sigma)$, quando (B) for simétrica e positiva definida. Uma especificação particular consiste em fazer $(B_{ii} = \sum_j w_{ij})$ e $(B_{ij} = -w_{ij})$. Nesse caso, (B) é singular e portanto a densidade da priori será imprópria. Isso não é necessariamente um problema, pois é possível mostrar que a posteriori pode ser própria.

Completando a especificação do modelo, vamos definir as prioris para (Σ) e (σ^2) , respectivamente uma Wishart invertida e uma Gama invertida, ou seja:

$$p(\Sigma | S; \nu) = k | \Sigma|^{(\nu + p + 1)/2} \exp\{-0,5tr(\Sigma^{-1}S)\} \quad (10)$$

$$p(\sigma^2 | s; \nu) = k (\sigma^2)^{-(\nu + 1 + 1)/2} \exp\{-s/(2\sigma^2)\} \quad (11)$$

O modelo especificado por (7), (8) (10) e (11) pode ser estendido pela inclusão de uma segunda dimensão. Assim, teríamos $(y_r | \alpha, \beta_r, x_{rt}, z_{rt}, \sigma^2) \sim N(x_{rt}\beta_r + \alpha z_{rt}, \sigma^2)$ onde (α) é o bloco de coeficientes que não varia com a região com dimensão $(1 \times q)$, Z são os regressores correspondentes com dimensão $(T \times n \times q)$, (β_r) tem dimensão $(p \times 1)$, (X_r) tem dimensão $(T \times p)$ e (y_r) é a variável dependente com dimensão $(T \times 1)$. Para completar o modelo, vamos acrescentar a distribuição *a priori* de (α) :

$$\alpha \sim N(a, A) \quad (11a)$$

Resumindo, o nosso modelo geral é dado pela verossimilhança:

$$L(\beta, \sigma^2, \alpha, \Sigma | y, x, z) \propto (\sigma^2)^{-nT/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{rt} (y_{rt} - z_{rt}\alpha + x_{rt}\beta_r)^2\right\} \quad (12)$$

e pelas prioris (8), (10), (11) e (11a).

3.2.1 - Estimação

Nesse modelo, a forma analítica das distribuições condicionais completas *a posteriori* de cada componente do vetor paramétrico $\varphi = (\beta, \sigma^2, \alpha, \Sigma)$ é conhecida, possibilitando a utilização do amostrador de Gibbs para a estimação do modelo. Nesse procedimento, os diversos componentes de (φ) são sorteados segundo (13-16) de forma seqüencial e repetida, até que os critérios de convergência mostrem que foi alcançada a distribuição de equilíbrio. Foram adotados a

amostragem por cadeias múltiplas e os critérios de convergência de Geweke (1992) e de Gelman e Rubin (1992). Denotando por (D) o conjunto dos dados disponíveis, temos:

1) Distribuição de $(\alpha | \beta, \sigma^2, \Sigma, D)$

$$p(\alpha | \beta, \sigma^2, \Sigma, D) \propto (\sigma^2)^{-nT/2} \exp\left\{\frac{1}{2\sigma^2} (y^\alpha - z\alpha)'(y^\alpha - z\alpha)\right\} \exp\{-0,5(\alpha - a)'A^{-1}(\alpha - a)\} =$$

$$= (\sigma^2)^{-nT/2} \exp\{-0,5[\alpha'(A^{-1} + z'z/\sigma^2)\alpha - 2\alpha'(z'y^\alpha/\sigma^2 + A^{-1}a)]\}$$

logo:

$$(\alpha | \beta, \sigma^2, \Sigma, D) \sim N(A^* (z'y^\alpha/\sigma^2 + A^{-1}a); A^*) \quad (13)$$

onde:

$$A^* = (A^{-1} + z'z/\sigma^2)^{-1} \quad e \quad y_{rt}^\alpha = y_{rt} - x_{rt}\beta_r$$

2) Distribuição de $(\beta_r | \alpha, \beta_{-r}, \sigma^2, \Sigma, D)$

$$p(\beta_r | \alpha, \beta_{-r}, \sigma^2, \Sigma, D) \propto$$

$$\propto (\sigma^2)^{-nT/2} \exp\left\{\frac{1}{2\sigma^2} (y^r - x_r\beta_r)'(y^r - x_r\beta_r)\right\} \exp\{-0,5(\beta_r - \bar{\beta}_r)' \Sigma^{-1} a_r (\beta_r - \bar{\beta}_r)\} =$$

$$= (\sigma^2)^{-nT/2} \exp\{-0,5[\beta_r' (\Sigma^{-1} a_r + x_r' x_r / \sigma^2) \beta_r] - 2\beta_r' (\Sigma^{-1} a_r \bar{\beta}_r + x_r' y^r / \sigma^2)\}$$

logo:

$$(\beta_r | \alpha, \beta_{-r}, \sigma^2, \Sigma, D) \sim N(D_r (\Sigma^{-1} a_r \bar{\beta}_r + x_r' y^r / \sigma^2); D_r) \quad (14)$$

onde:

$$\bar{\beta}_r = \sum_j w_{rj} \beta_j, \quad y^r = y_{rt} - z_t \alpha \quad e \quad D_r = (\Sigma^{-1} a_r + x_r' x_r / \sigma^2)^{-1}$$

3) Distribuição de $(\Sigma | \beta)$

$$p(\Sigma | \beta) \propto |\Sigma|^{-n/2} \exp\{-0,5 \sum_{j<r} w_{jr} (\beta_j - \beta_r)' \Sigma^{-1} (\beta_j - \beta_r)\} |\Sigma|^{(v+p+1)/2} \exp\{-0,5 \text{tr}(\Sigma^{-1} S)\} =$$

$$= |\Sigma|^{-(n+v+p+1)/2} \exp\{-0,5 \text{tr}[\Sigma^{-1} S + \sum_{j<r} w_{jr} (\beta_j - \beta_r)' \Sigma^{-1} (\beta_j - \beta_r)]\} =$$

$$= |\Sigma|^{-(n+v+p+1)/2} \exp\{-0,5 \text{tr}[\Sigma^{-1} (SS)]\}$$

logo:

$$(\Sigma | \beta) \sim WI(SS; n + v + p + 1) \quad (15)$$

onde:

$$SS = S + \sum_{j < r} w_{jr} (\beta_j - \beta_r)(\beta_j - \beta_r)'$$

4) Distribuição de (σ^2)

$$\begin{aligned} (\sigma^2 | \beta, \alpha, \Sigma, D) &\propto (\sigma^2)^{-nT/2} \exp\left\{ \frac{1}{2\sigma^2} \sum_{rt} (y_{rt} - z_t \alpha + x_r \beta_r)^2 \right\} |\sigma^2|^{(v+1+1)/2} \exp\{-s/(2\sigma^2)\} = \\ &= (\sigma^2)^{-(nT+v+1+1)/2} \exp\{-0,5[s + \sum_{rt} (y_{rt} - z_t \alpha + x_r \beta_r)^2] / \sigma^2\} \end{aligned}$$

logo:

$$(\sigma^2 | \beta, \alpha, \Sigma, D) \sim GI(s + \sum_{rt} (y_{rt} - z_t \alpha + x_r \beta_r)^2; nT + v + 1 + 1) \quad (16)$$

4 - CRITÉRIO DE COMPARAÇÃO DOS MODELOS

A seleção de modelos é uma questão de grande relevância, sobretudo quando passamos a considerar modelos mais complexos e os procedimentos de seleção usuais deixam a desejar: tanto a estatística clássica, que só compara modelos encaixados, quanto a abordagem bayesiana, onde o fator de Bayes fica maldefinido no caso de prioris impróprias. Nos dois casos, a seleção está associada a critérios derivados da razão entre as verossimilhanças, no caso clássico calculando esta razão no ponto de máximo, e no caso bayesiano tomando a razão entre as “médias” das densidades. Seja $p(D_T | \theta, M)$ a verossimilhança dado D_T , os parâmetros θ e o modelo M . Então, os critérios clássico e bayesiano são:

$$RV_c = \frac{\sup_{\theta} p(D_T | \theta, M_1)}{\sup_{\theta} p(D_T | \theta, M_2)} \quad RV_B = \frac{\int_{\theta} p(D_T | \theta, M_1) p(\theta | M_1) p(M_1) d\theta}{\int_{\theta} p(D_T | \theta, M_1) p(\theta | M_2) p(M_2) d\theta}$$

O critério de seleção tende a dar muito peso ao modelo sobreparametrizado. Por isso, na literatura clássica, introduz-se uma penalidade, baseada em teoria assintótica, que assume como fixa a dimensão do espaço de parâmetros. Essa abordagem dá origem aos critérios de informação de Akaike (AIC), Schwartz (BIC), ou Hanna-Quin. A hipótese de que a dimensão do espaço paramétrico é fixa não se aplica aos modelos de parâmetro variando, hierárquicos ou de efeitos aleatórios, onde o significado da dimensão do espaço de parâmetros não é evidente.

Na abordagem bayesiana, admitindo-se $p(M_1) = p(M_2)$, RV_B é conhecido como fator de Bayes, e pode estar maldefinido no caso em que a priori é vaga, como costuma acontecer nos modelos hierárquicos ou de efeitos aleatórios. Nos dois casos, os critérios de seleção correspondem à adoção de uma função de perda do tipo 0-1.

Um procedimento alternativo proposto por Gelfand e Ghosh (1998) consiste em adotar funções de perda mais gerais, que considerem um compromisso entre a capacidade preditiva e um fator de punição devido à dimensão paramétrica. Essa proposta está associada às funções de perda balanceadas de Zellner (1994).

Para cada modelo (M), denote a função de perda por $L(y_l^w, a_l; y_l)$, onde (y_l^w) é o valor replicado pelo modelo para (y_l) e (a_l) é o previsor de y_l . Então temos a perda (17), onde as componentes referem-se aos tipos de perda associadas ao modelo (M) de previsão, a primeira à imprecisão, e a segunda à inexatidão ou não-acurácia.

$$L_k(y_l^w, a_l | y_l) = L(y_l^w, a_l) + k L(a_l, y_l), k \geq 0 \quad (17)$$

Considere como critério de escolha de (a_l) a minimização da perda esperada, ou seja:

$$\mu_l = \arg \min_{a_l} E_{y_l^w | y_l} L_k(y_l^w, a_l | y_l)$$

Agregando-se (17) para todos os elementos da amostra, e substituindo (a) pela sua estimativa, temos:

$$L_k = \sum_{l=1}^n E_{y_l^w | y_l} L(y_l^w, \mu_l | y_l) = \sum_{l=1}^n E_{y_l^w | y_l} \{L(y_l^w, \mu_l) + kL(\mu_l, y_l)\}$$

Particularizando para o caso de perda quadrática, temos:

$$\begin{aligned} \mu_l &= E(y_l^w | D) \\ L_k &= k/(k+1) \sum_{l=1}^n (y_l - \mu_l)^2 + \sum_{l=1}^n (y_l^w - \mu_l)^2 \end{aligned} \quad (18)$$

onde a primeira parcela se refere à soma dos erros quadráticos e mede a variância do ajuste, e a segunda é um termo de penalidade que será chamado de variância preditiva. É fácil ver que $\lim_{k \rightarrow \infty} L_k = \sum_{l=1}^n (y_l^w - \mu_l)^2 + \sum_{l=1}^n (y_l - \mu_l)^2$, o que corresponde a atribuímos um peso muito grande ao erro de ajustamento, que será o critério adotado.

No caso em que o modelo é estimado por métodos de simulação, obtemos uma amostra $\{\theta^w \ w=1...R\}$ da posteriori dos parâmetros do modelo. Então, para cada

(w) podemos amostrar ($y_l^w \sim (y_l | \theta^w)$) e construir uma amostra das parcelas envolvidas em (18) utilizando Monte Carlo. Tomaremos a sua média como sumário das distribuições empíricas obtidas, obtendo estimativas da variância de ajustamento (VA), da variância preditiva (VP) e da variância total (VT), que são as estatísticas utilizadas para comparar os modelos.

$$\begin{aligned} m_l &= \sum_w y_l^w / R & VA &= \sum_l (y_l - m_l)^2 / n \\ VP &= \sum_l \sum_w (y_l^w - m_l)^2 / nR & DPT &= \sqrt{VP + VA} \end{aligned} \quad (19)$$

5 - RESULTADOS

Os modelos de parâmetro fixo (PF) e de parâmetro variando foram estimados para cada ano censitário $t \in \{1970, 1980, 1985, 1996\}$ e para todos os censos conjuntamente. O modelo (PF) foi estimado para cada período com a equação (20), e para o conjunto com a equação (21), onde foi imposta restrição de igualdade intertemporal dos coeficientes e incluídas as variáveis indicadoras (I_t) para medir o efeito do ano sobre o intercepto. Esses modelos foram estimados na forma (2), que considera auto-regressão espacial, e na forma (3), que considera o resíduo com estrutura espacial, sendo que a representação menos adequada foi omitida em algumas tabelas. Em todos os casos, (y_t) é ($n \times 1$), (ρ_t) é um escalar, (X_t) é ($n \times p$) e (β_t) é ($p \times 1$), onde n é o número de MR e p o número de regressores.

$$y_t = \rho_t W y_t + X_t \beta_t + e_t \quad e_t \sim N(0, I_n \sigma_t^2) \quad (20.2)$$

$$y_t = X_t \beta_t + u_t \quad u_t = \rho_t W u_t + e_t \quad e_t \sim N(0, I_n \sigma_t^2) \quad (20.3)$$

$$y_t = \sum \phi_t I_t + \rho W y_t + X_t \beta_0 + e_t \quad e_t \sim N(0, I_n \sigma^2) \quad (21.2)$$

$$y_t = \sum \phi_t I_t + X_t \beta + u_t \quad u_t = \rho W u_t + e_t \quad e_t \sim N(0, I_n \sigma^2) \quad (21.3)$$

De forma análoga, o modelo de parâmetro variando (PV) foi estimado respectivamente com as equações (22) e (23), onde (y_{rt}) é escalar e (X_{rt}) e (β_{rt}) são ($1 \times p$).

$$y_{rt} = X_{rt} \beta_{rt} + e_{rt} \quad e_{rt} \sim N(0, \sigma_t^2) \quad (22)$$

$$y_{rt} = \sum \phi_t I_t + X_{rt} \beta_r + e_{rt} \quad e_{rt} \sim N(0, \sigma^2) \quad (23)$$

A Tabela 3 apresenta para todos os casos o valor do critério de comparação entre modelos proposto na Seção 3. No modelo (PF), o número de parâmetros é ($npf = p + 1$) — número de regressores mais o coeficiente de auto-regressão espacial; no modelo (PV), esse número é ($npv = pN + p(p + 1)/2 + 1$), o coeficiente por regressor e por região, acrescido dos coeficientes da matriz (Σ) e

da variância observacional. É claro que $npv \gg npf$ e, portanto, é de se esperar que a variância preditiva dos modelos de PV seja maior do que no modelo (PF). Apesar disso, o desvio-padrão total dos modelos do tipo (PF) são, em todos os casos, maiores do que os correspondentes de PV, mostrando que estes últimos são mais adequados para representar o efeito espacial na produção agropecuária.

Tabela 3

Resultados dos Modelos por Período e Multiperíodo

Modelo	PV			PF		
	V. Ajuste	V. Preditiva	D. P. Total	V. Ajuste	V. Preditiva	D. P. Total
Censo 1970	0,0012	0,1262	0,357	0,0994	0,1109	0,459
Censo 1975	0,0009	0,1349	0,369	0,1306	0,1161	0,497
Censo 1980	0,0009	0,1398	0,375	0,1139	0,1047	0,467
Censo 1985	0,0008	0,1436	0,380	0,1142	0,1138	0,477
Censo 1996	0,0010	0,1530	0,393	0,1944	0,1802	0,612
Conjunto 1970/96	0,0259	0,0198	0,214	0,111	0,1210	0,483

A Tabela 4 compara as elasticidades estimadas para os modelos (PV) e (PF) obtidas para o Censo de 1996 e para o conjunto. No caso do modelo (PF), é apresentado o valor esperado de cada elasticidade — todas significativas ao nível de 5% — e, no caso de (PV), a média entre as regiões dos valores esperados estimados para cada região. A tabela mostra que, *grosso modo*, os resultados são comparáveis. No Gráfico 1 do Apêndice (conjunto de 10 gráficos), as elasticidades estimadas para cada microrregião são apresentadas justapostas, incluindo o intervalo de máxima densidade *a posteriori* ao nível de 65%. Excetuando as elasticidades referentes ao efeito da cultura permanente e da pecuária, não se rejeita a hipótese de que os coeficientes flutuam entre as regiões e são diferentes de zero, particularmente no caso da estimação conjunta de todos os censos. Como os modelos de parâmetro variando consideraram não-significativo o efeito da pecuária e da cultura permanente, o modelo (PV) foi reestimado eliminando esses regressores. Os resultados estão indicados com (*). O efeito da exclusão não foi importante qualitativamente.

A Tabela 5 apresenta a média entre as MRs, do desvio-padrão das elasticidades estimadas para cada censo, isoladamente e para o caso conjunto, que deixa claro que o acréscimo de informação reduziu a incerteza das estimativas. Finalmente, a Tabela 6 mostra o 5° e o 95° percentis da distribuição das elasticidades estimadas. O uso combinado das Tabelas 5 e 6 sugere em que medida se consegue de fato medir a heterogeneidade das elasticidades com o modelo (PV). A comparação da incerteza da estimativa, por exemplo a da média das elasticidades-área de 1996 (.285) com a dispersão das mesmas elasticidades [.05,.3] sugere que a heterogeneidade, nesse caso, não está bem medida. Já para o modelo multiperíodo temos, respectivamente, (.164) e [-.09,.405], o que claramente rejeita a hipótese de igualdade das elasticidades.

Tabela 4

Comparação (PV/PF) por Período e Multiperíodo

	PV 1970/96	PV* 1970/96	PF(2) 1970/96	PF(3) 1970/96	PV 1996	PV* 1996	PF(2) 1996	PF(3) 1996
Ro:Moda	–	–	0,727	0,269	–	–	0,639	0,245
Ro:IMDP	–	–	[0,70, 0,75]	[0,25, 0,29]	–	–	[0,60, 0,70]	[0,20, 0,30]
Intercepto	1,721	2,116	0,357	-0,050	0,863	1,382	1,060	1,368
Dummy 75	0,205	0,225	0,014	-0,017	–	–	–	–
Dummy 80	0,343	0,369	0,024	-0,096	–	–	–	–
Dummy 85	0,457	0,486	0,054	-0,039	–	–	–	–
Dummy 96	0,824	0,867	0,151	0,205	–	–	–	–
Área	0,239	0,202	0,148	0,142	0,254	0,184	0,092	0,092
P.Ocupado	0,360	0,376	0,490	0,384	0,384	0,402	0,384	0,309
N.Tratores	0,217	0,218	0,327	0,301	0,424	0,474	0,435	0,370
N.Bovinos	0,011	–	0,013	0,014	0,007	–	0,010	0,010
N.Pés	0,006	–	0,006	0,009	0,010	–	0,007	0,011
Escala	0,833	0,796	0,984	–	1,078	1,060	–	–

Tabela 5

Média nas MRs do Desvio-Padrão de cada Efeito no Modelo (PV)

Censos	1970	1975	1980	1985	1996	Conjunto
Área	0,234	0,244	0,251	0,268	0,285	0,164
P.Ocupado	0,291	0,298	0,301	0,313	0,343	0,186
Bovinos	0,220	0,235	0,261	0,267	0,290	0,096
Tratores	0,149	0,125	0,119	0,114	0,110	0,055
C.Permanente	0,104	0,082	0,074	0,072	0,068	0,042

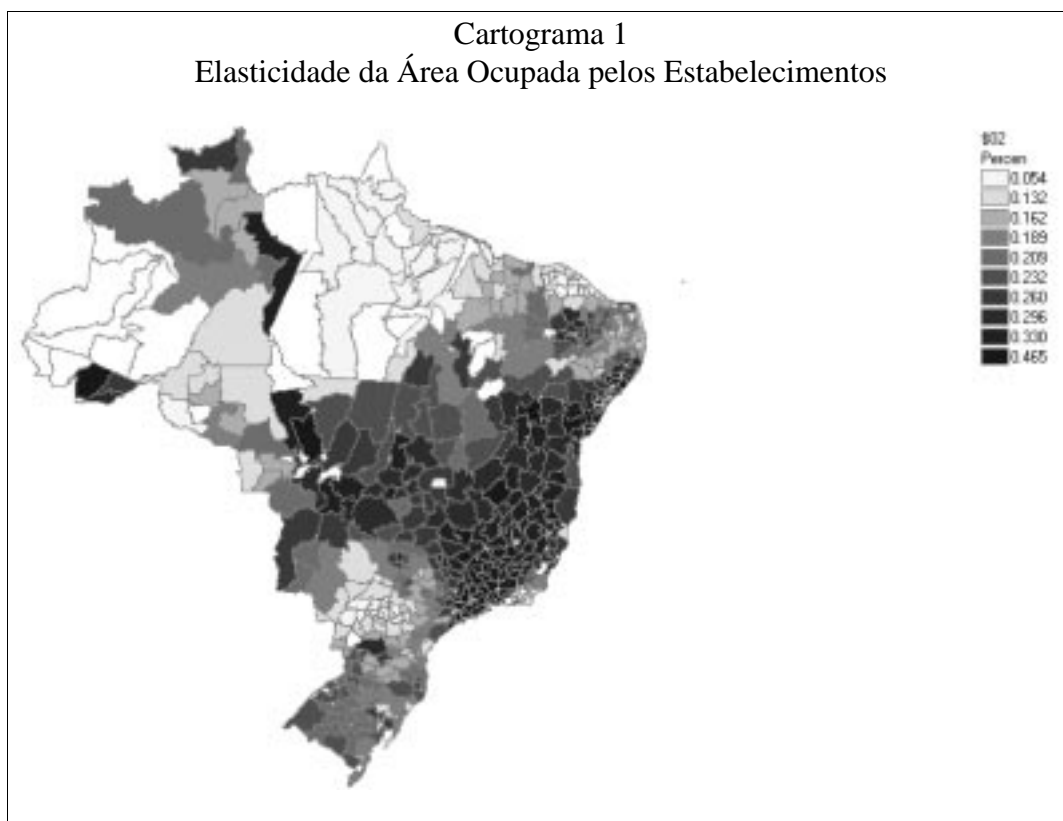
Tabela 6

Dispersão das Elasticidades entre as MRs

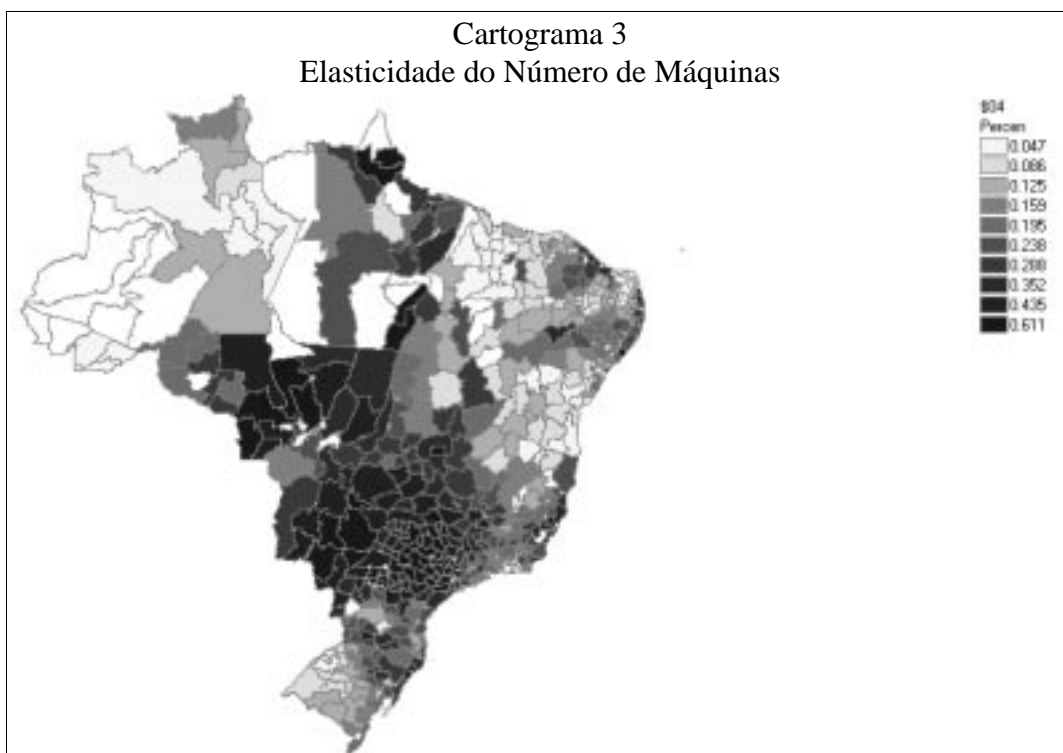
Percentil	5% 1970/96	95% 1970/96	5% 1970/96*	95% 1970/96*	5% 1996	95% 1996	5% 1996*	95% 1996*
Total	0,578	1,006	0,460	0,985	1,000	1,164	0,956	1,177
Área	0,035	0,374	-0,090	0,405	0,143	0,326	0,048	0,300
P.Ocupado	0,117	0,635	0,103	0,693	0,293	0,538	0,226	0,718
N.Tratores	-0,059	0,542	-0,099	0,568	0,239	0,505	0,182	0,607
N.Bovinos	-0,029	0,068	–	–	-0,033	0,053	–	–
N.Pés	-0,017	0,035	–	–	-0,014	0,044	–	–

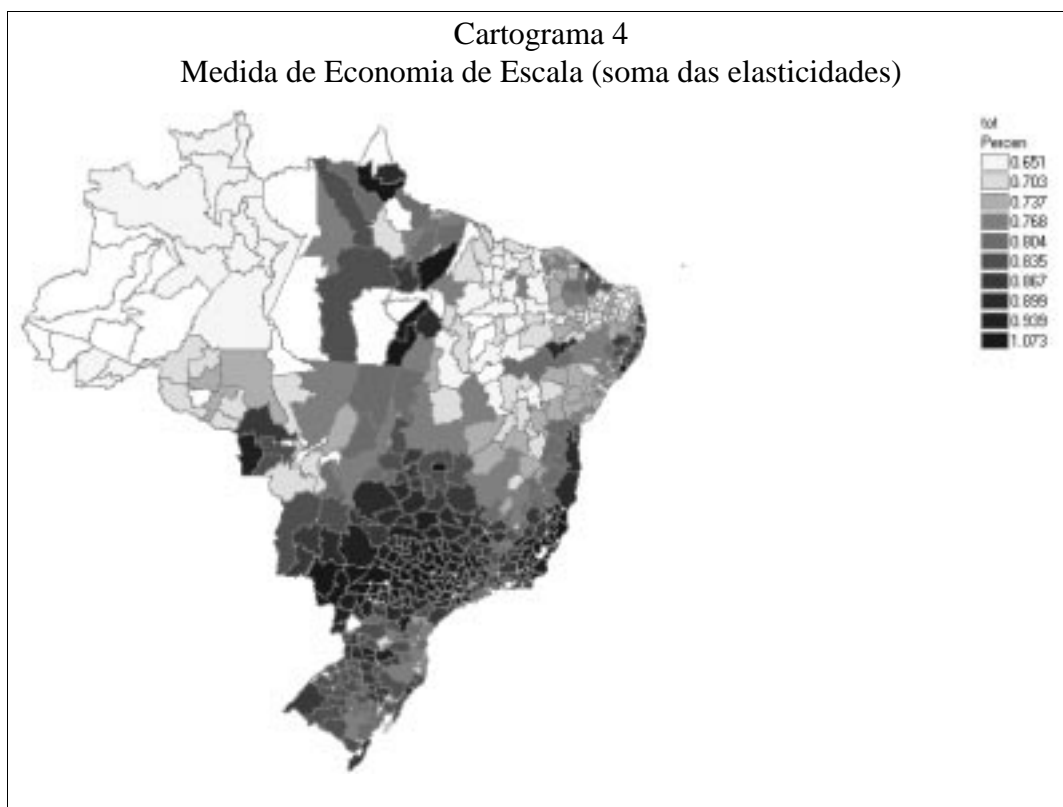
Os cartogramas apresentados a seguir resumem a distribuição espacial do valor de cada elasticidade estimada para o caso multiperíodo restrito, e para a medida do grau de economia de escala. Cada cartograma apresenta os resultados relativos a uma das variáveis, em que uma cor foi associada a cada percentil — como definido à direita do cartograma. As MRs foram então coloridas de acordo com o percentil associado a sua elasticidade. A tendência espacial é visualizada pela alteração suave do colorido obtido.

O Cartograma 1 mostra o efeito de área ocupada, em que os menores valores ocorrem nas regiões de abundância do fator terra, e os maiores nas regiões onde a fronteira agrícola já se estabilizou. O Cartograma 2 mostra o efeito do pessoal ocupado: os maiores valores ocorrem nas regiões novas, onde o fator mão-de-obra ainda é escasso; na região Sul o valor também é grande, mas não encontramos justificativa para o fato.



Finalmente, os Cartogramas 3 e 4 mostram o efeito do número de tratores, e os maiores valores ocorrem nas regiões no entorno de São Paulo/Paraná/Mato Grosso do Sul, que possivelmente adotam tecnologias mais modernas.





Vale mencionar que parte das diferenças de tonalidade, e portanto de heterogeneidade da produtividade, não pôde ser reproduzida numa apresentação em preto e branco.

Apresentamos a seguir um resumo dos resultados:

- A Tabela 3 mostra, como esperado, que o modelo de PF apresenta menor variância preditiva do que o modelo de PV, pois tem muito menos parâmetros.
- A Tabela 3 mostra, como esperado, que o modelo de PV apresenta menor variância de ajuste do que o modelo de PF.
- O modelo multiperíodo descreve os dados com mais parcimônia e os seus parâmetros são estimados com menor incerteza, como mostram a Tabela 5 e o Gráfico 1.
- O Gráfico 1 mostra que os efeitos do número de bovinos e do número de pés não são significativamente diferentes de zero, sugerindo a exclusão desses regressores. Coincidentemente, estes são os regressores que foram definidos com a restrição de intensidade mínima. Como, para o modelo (PF), estes regressores apresentaram coeficientes pequenos e significativos, não fica claro se a não-significância desses coeficientes no caso de (PV) decorre da natural maior incerteza dos estimadores desse modelo, ou se é consequência de uma

eventual incompatibilidade entre o uso de parâmetros fluando espacialmente e variáveis indicadoras.

- No Gráfico 1, particularmente no caso multiperíodo, observa-se que os parâmetros são diferentes entre as MRs, pois a hipótese de igualdade é rejeitada para a maioria dos subconjuntos de MR.
- No Gráfico 2 do Apêndice (conjunto de cinco gráficos) observam-se a semelhança das elasticidades estimadas entre os diversos anos e a divergência quando consideramos o caso multiperíodo, embora a hipótese de igualdade entre os parâmetros dos diversos censos seja rejeitada estatisticamente em muitos casos.¹⁵
- Vale notar que os parâmetros que flutuam no espaço são o resultado da mistura da informação contida nos dados com a priori espacial, e portanto, à medida que a quantidade de informação contida nos dados aumenta — o que ocorre no caso de multiperíodo —, o efeito relativo da priori se reduz. Portanto, não devemos esperar que o resultado multiperíodo seja aproximadamente a média dos uniperíodos, mesmo que todos os uniperíodos fossem iguais.
- Os cartogramas possibilitam uma visualização dos resultados, mostrando o efeito da restrição de similaridade espacial que resulta na tendência espacial da flutuação das elasticidades.

6 - CONCLUSÃO

A diversidade das condições rurais no Brasil sugere a necessidade da utilização de modelos de parâmetro variando espacialmente para estimar a função de produção agropecuária. No entanto, os resultados devem ser vistos com cautela, pois alguns pontos necessitam ser reconsiderados.

O fator humano foi medido apenas pelo número de pessoas ocupadas, sem considerar a intensidade do trabalho e a qualidade. A proporção dos trabalhadores temporários e das famílias traz informações sobre esses aspectos. O fator número de máquinas foi definido incluindo veículos de transporte, o fator número de bovinos deveria incluir apenas os reprodutores e as fêmeas matrizes e, finalmente, o fator relativo às culturas permanentes considera uma lista muito reduzida de produtos. Outra limitação que será superada numa próxima versão deste texto é a utilização de um índice de *quantum* do produto agropecuário em substituição ao valor da produção deflacionado.

¹⁵ O intervalo de confiança não foi apresentado por economia de espaço, e para não tornar as figuras do Gráfico 2 ainda mais confusas.

Na especificação do modelo, é interessante explorar outros critérios para medir a influência da vizinhança, como a lei do inverso do quadrado da distância, ou do tamanho medido pela área dos municípios vizinhos.

A literatura de economia agrícola e pecuária já acumulou uma vasta produção na estimativa de funções de produção para cada produto ou para o agregado, com formas funcionais mais complexas do que a adotada neste artigo, e muitas vezes utilizando um conjunto de informações mais rico. Seria interessante situar os resultados deste artigo comparando-os, quando possível, com aqueles obtidos com os modelos de parâmetro fixo mencionados.

No caso dos modelos de série temporal com parâmetros variando, Carter e Kohn (1994) mostram que os esquemas de amostragem conjuntos apresentam maior velocidade de convergência do que os esquemas que amostram componente a componente. Se essa propriedade for preservada para os modelos de parâmetro variando no espaço, podemos dizer que o procedimento proposto neste artigo é superior ao proposto por Assunção e Assunção (1998), e inferior aos esquemas que sorteiam conjuntamente os (β) , como proposto por Gamerman (1999). Como o sorteio conjunto envolve operações com matrizes muito grandes — posto igual ao produto entre o número de regiões e o número de regressores — a vantagem relativa dos esquemas deve ser avaliada empiricamente.

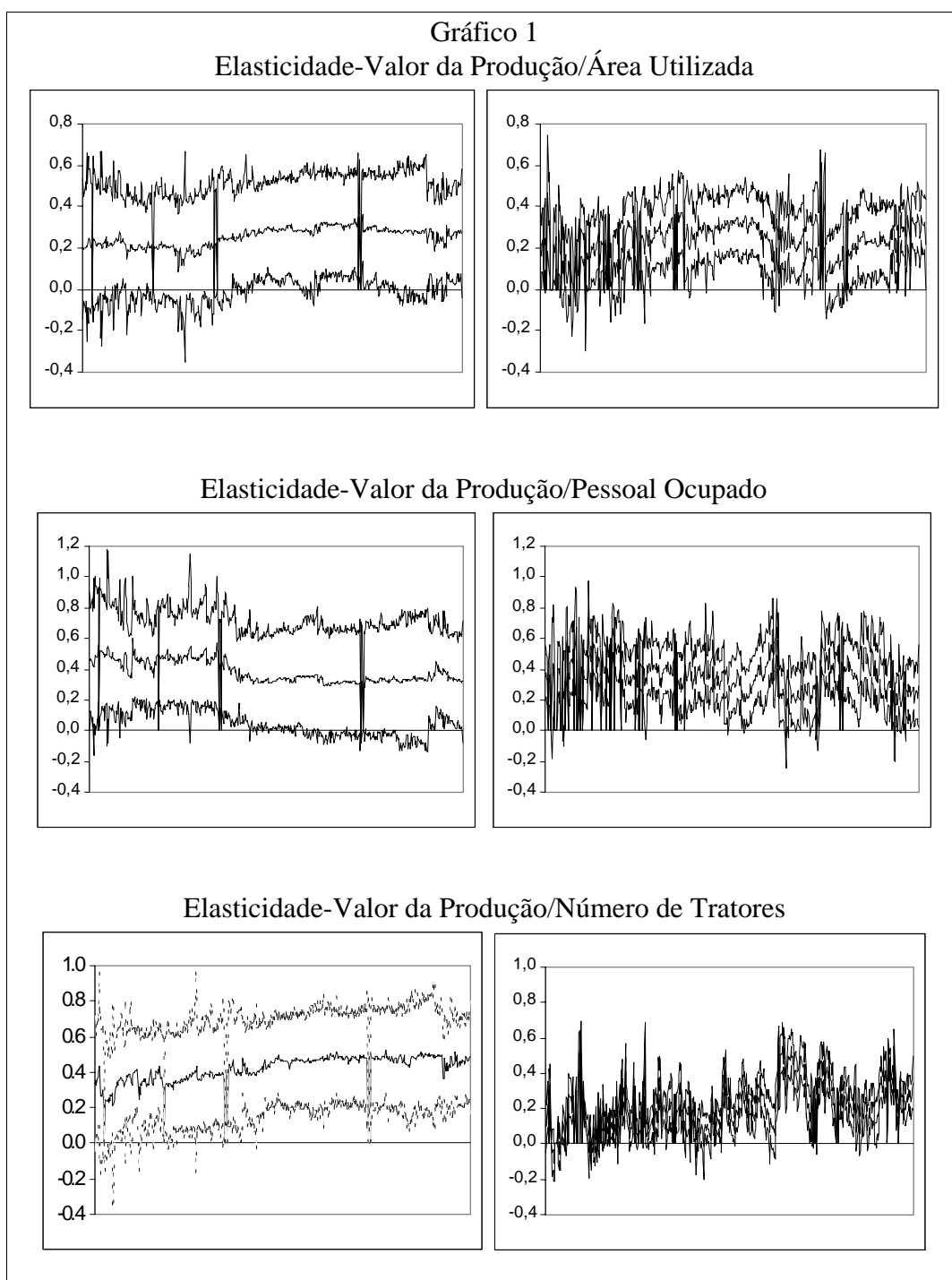
Em muitas questões a diversidade espacial é relevante, e portanto o modelo de parâmetros variando espacialmente parece ter um vasto campo de utilização, particularmente nos problemas demográficos de estimação de incidência de doenças e taxas de mortalidade/fecundidade. Nesse caso, a metodologia utilizada deve ser estendida. A hipótese de normalidade adotada é uma aproximação ruim para esses casos, devendo ser substituída pela distribuição de Poisson. Infelizmente, essa substituição torna desconhecidas algumas das distribuições condicionais, impedindo a utilização do amostrador de Gibbs.

Finalmente, vale registrar que esses resultados devem ser vistos como preliminares. Melhorias importantes podem ser apontadas desde já, nos aspectos econômicos, nos aspectos estatísticos de estimação e no entendimento mais completo das limitações e recursos dessa ferramenta. Apesar disso, acreditamos que os resultados permitem uma análise descritiva da heterogeneidade da produtividade.

O programa utilizado (PRVE) está disponível no seguinte endereço: www.ipea.gov.br/ftp/ipeaftp.html.

Apêndice

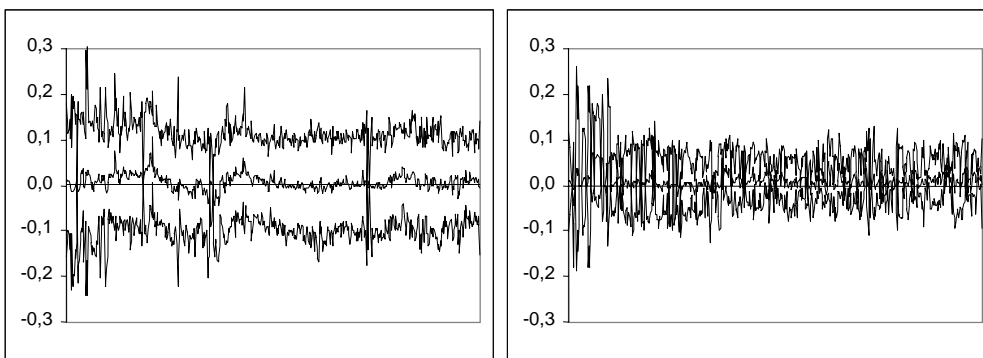
No Gráfico 1 (conjunto de 10 gráficos) estão apresentadas, na primeira coluna de figuras, as elasticidades entre cada um dos regressores e o valor da produção, e das MRs justapostas e seu respectivo intervalo de confiança (65%) estimados para o Censo de 1996. A segunda coluna apresenta as elasticidades para o conjunto dos censos {1970, 1975, 1980, 1985, 1996} na segunda coluna.



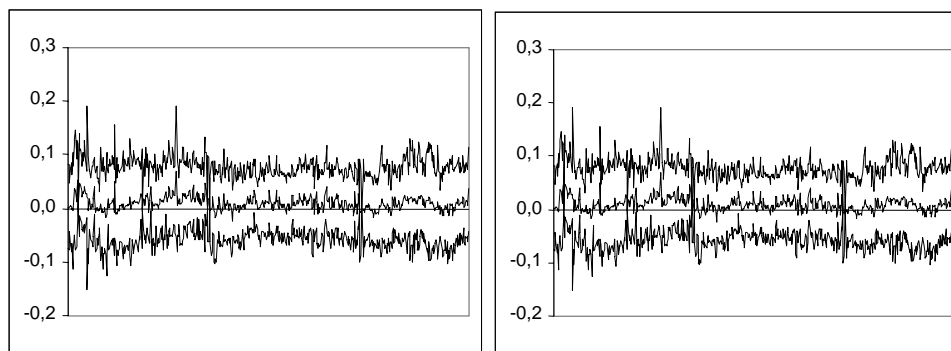
(continua)

(continuação)

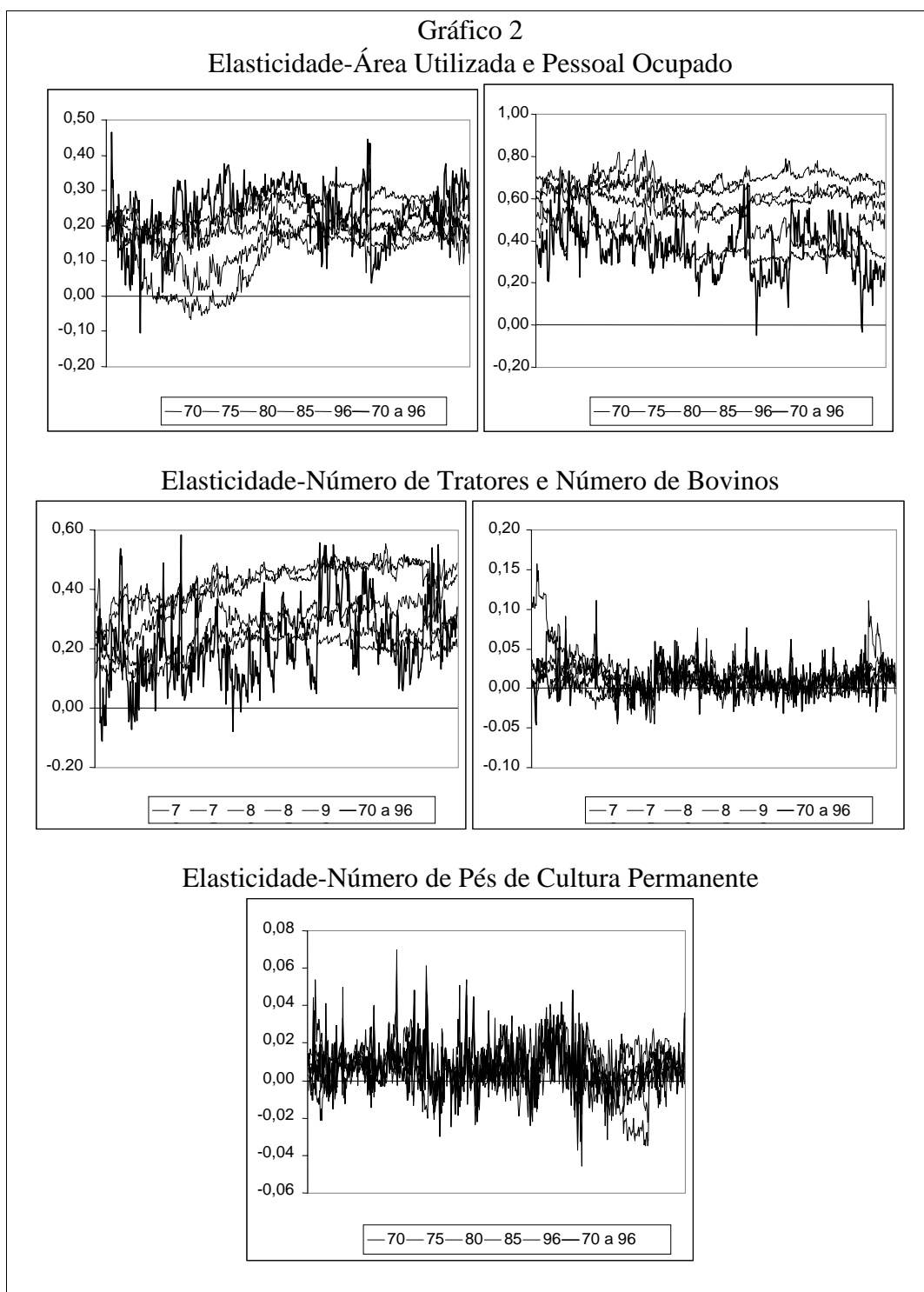
Elasticidade-Valor da Produção/Número de Bovinos



Elasticidade-Valor da Produção/Número de Pés



No Gráfico 2 (conjunto de cinco gráficos) estão apresentados o valor esperado das elasticidades entre cada um dos regressores e o valor da produção das MRs justapostas, estimadas para cada um dos diversos censos — linha mais fina — e para o conjunto dos censos — linha mais grossa.



BIBLIOGRAFIA

- ANSELIN, L. *Spatial econometrics: methods and models*. Kluwer Academic, 1988.
- ASSUNÇÃO, J., GAMERMAN, D., ASSUNÇÃO, R. Regional differences in factor productivities of Brazilian agriculture: a Bayesian space varying parameter approach. 1999, mimeo.
- ASSUNÇÃO, J., LEMOS, M., ASSUNÇÃO, R. Regional differences in factor productivities of Brazilian agriculture: a Bayesian space varying parameter approach. Anais da SBE, 1998.
- BAILEY, T. *Interactive spatial data analysis*. A. Longman, 1995.
- BRANDT, S., MARTINS, M., ARAÚJO, M., MELLO, A. Extensão rural e pesquisa agrícola: uma avaliação dos problemas de não-homoteticidade e viés tecnológico. *Pesquisa e Planejamento Econômico*, Rio de Janeiro, n. 18, p. 131-144, 1988.
- CARTER, C. K., KOHN, R. On gibbs sampling for state space models. *Biometrika*, n. 81, p. 541-553, 1994.
- GAMERMAN, D. *Markov Chain Monte Carlo: stochastic simulation for Bayesian inference*. Londres: Chapman & Hall, 1997.
- . *MCMC in econometrics*. 1999, mimeo.
- GAMERMAN, D., MOREIRA, A. *Bayesian analysis of econometric time series models using hybrid integration rules*. 1998, mimeo.
- GELFAND, A., GHOSH, S. Model choice: a minimum posterior predictive loss approach. *Biometrika*, n. 85, 1998.
- GELMAN, A., RUBIN, D. R. Inference from interactive simulation using multiple sequences. *Statistical Science*, n. 7, p. 457-511, 1992.
- GEWEKE, J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: BERNADO *et alii* (eds.). *Bayesian statistics*, Oxford: Oxford University Press, v. 4, p. 169-193, 1992.
- HOFFMAN, R., KASSOUF, L. Modernização e desigualdades na agricultura brasileira. *Revista Brasileira de Economia*, v. 43, n. 2, p. 273-303, 1989.
- KNORR-HELD, L., BESAG, J. Modelling risk from a disease in time and space. *Statistics in Medicine*, n. 17, 1998.
- WALLER, L., CARLIN, B. P., XIA, H., GELFAND, A. Hierarchical spatio-temporal mapping of disease rates. *JASA*, n. 92, p. 607-617, 1997.
- ZELNER, A. Bayesian and non-Bayesian estimation using balanced loss functions. In: GUPTA, S. S., BERGER, J. O. *Statistical decision theory and related topics*. New York: Springer-Verlag, v. 5, p. 377-390, 1994.