

CONCENTRAÇÃO DE ATIVIDADES PRODUTIVAS NO BRASIL: DIFERENÇAS INTERSETORIAIS, REGIONAIS E SUA DINÂMICA

Ciro Biderman

FGV-EESP

Seminário nº 1/2004 – 11/03/2004

**São Paulo
2004**

CONCENTRAÇÃO DE ATIVIDADES PRODUTIVAS NO BRASIL: DIFERENÇAS INTERSETORIAIS, REGIONAIS E SUA DINÂMICA*

*Ciro Biderman***

Resumo

Nos últimos anos assistiu-se no Brasil a uma diminuição considerável na força de trabalho concentrada na indústria e a um aumento na concentração nos serviços. Este artigo pretende entender como este processo pode ser compreendido a partir de um modelo dinâmico onde as empresas decidem simultaneamente a sua localização. Os índices de concentração utilizados separam o grau de concentração de um setor devido à escala de produção do grau de concentração em função de vantagens regionais. Este tipo de análise exige a construção de índices que permitam esta decomposição. Neste artigo adotou-se o índice de Herfindahl-Hirshman, para a concentração devida à escala, e o índice de Ellison e Glaeser (1997) para a concentração devida às vantagens regionais.

Palavras-Chaves

economias de escala; vantagens aglomerativas; índice de Herfindahl-Hirshman; concentração setorial; setor de serviços no Brasil; setor industrial no Brasil.

Abstract

In the last decades it could be seen in Brazil a considerable decrease in the labor force in manufacturing and an increase in the service sector. This research attempts to understand that process in a dynamic model where firms choose simultaneously their location. The agglomeration index used decompose the agglomeration due to scale from the agglomeration due to regional advantages. For scale economies we adopt the Herfindahl-Hirshman index and for regional advantage the Ellison and Glaeser (1997) index.

Key words

scale economies; regional advantages; Herfindahl-Hirshman index; industrial concentration; service industry in Brazil; manufacturing industry in Brazil.

* Trabalho em andamento preparado especialmente para o seminário acadêmico da USP. Não citar sem consultar o autor. Agradeço a assistência de pesquisa de Fernando Coteló e Flávio Calife. Este trabalho foi financiado pelo Núcleo de Pesquisas e Publicações (NPP) da EAESP-FGV.

** Professor da EAESP-FGV e pesquisador do Laboratório de Urbanismo da Metrópole (Lume) da FAUUSP no projeto Cepid (CEM) apoiado pela Fapesp.

Introdução

Nos últimos anos assistiu-se no Brasil a uma diminuição considerável na força de trabalho concentrada na indústria e a um aumento na concentração nos serviços. Esta variação, no entanto, não foi uniforme nas regiões. Em 1977 as áreas urbanas brasileiras concentravam 23% da sua população em atividades industriais enquanto as regiões metropolitanas concentravam 26%. Em 1999 a concentração da força de trabalho neste setor de atividades era a mesma para o Brasil urbano e para as regiões metropolitanas (15%). O caso é ainda mais radical se tomarmos a maior região metropolitana do país. Em 1977, a região metropolitana de São Paulo (RMSP) concentrava 38% da sua população economicamente ativa na indústria de transformação, número reduzido pela metade (19%) em 1999.

Este artigo pretende compreender como este processo de desconcentração da força de trabalho na indústria de transformação e o movimento paralelo de concentração em serviços produtivos verificado com maior intensidade nas regiões metropolitanas do que nas demais regiões urbanas do país pode ser compreendida a partir de um modelo dinâmico onde as empresas decidem simultaneamente a sua localização. Em outras palavras, pretende-se verificar o que ocorreu ao longo das duas últimas décadas com as economias de aglomeração e de urbanização que acabaram induzindo a migração da indústria para regiões médias e ao aumento da concentração dos serviços produtivos em grandes regiões.

A concentração de determinados setores de atividade em certas regiões é um fator conhecido da literatura há muitos anos. Planejadores urbanos, geógrafos, sociólogos e economistas sempre procuraram explicar este fenômeno¹. Mais recentemente este padrão passou a ser explorado como uma fonte potencial de compreensão da natureza dos retornos crescentes e de externalidades que são a base das novas teorias de comércio internacional, desenvolvimento econômico e de economia regional e urbana. A motivação deste artigo é discutir qual deve ser o efeito sobre as grandes aglomerações desta mudança de tendência de aglomeração (da indústria para os serviços).

Micro-fundamentos para os índices de concentração das atividades

Um dos motivos econômicos para a concentração de um determinado setor em uma região está relacionado com as chamadas “economias de aglomeração”: as empresas se localizam próximas umas das outras para economizar em custos de transporte tanto na compra de matéria-prima como na entrega do produto aos clientes. Como os serviços, por definição, não apresentam um produto físico e considerando o declínio nos custos de transmissão da informação, bem superior ao declínio nos custos de transporte, pode-se concluir que o custo de entrega para os setores de serviço deve estar caindo em relação aos setores que produzem bens tangíveis. Esta diferença pode explicar o comportamento diverso dos dois grupos de atividade (indústria e serviços).

¹ Veja-se, por exemplo, Florence (1948), Hoover (1948), Fuchs (1962), Carlton (1983) e para uma literatura mais atual Henderson (1988), Enright (1990), Porter (1990), Krugman (1991) e Ellison e Glaeser (1997).

Para entender por que uma atividade produtiva está concentrada em uma região deve-se, antes de tudo, compreender por que existem cidades. As cidades só existem devido a alguma forma de economia de escala. Os motivos econômicos foram apontados pela primeira vez em Marshal (1896). Em um centro industrial as empresas economizam em transportes pois estão mais próximas dos clientes e fornecedores. Além do mais, trabalhadores especializados diminuem o risco de desemprego num grande centro se os choques de oferta não forem correlacionados. Finalmente, um grande centro facilita a circulação de informações, o que ficou conhecido na literatura como “technological spillover”.

Encontrar um índice que represente o grau de concentração de uma determinada indústria não é uma tarefa trivial. Uma medida “natural” seria:

$$g_j \equiv \sum_i (s_{i,j} - x_i)^2 \quad (1)$$

Onde x_i indica a proporção da sub-unidade geográfica na força de trabalho e s_{ij} representa a proporção do emprego no setor j nesta sub-unidade. Trata-se de uma medida do quanto o emprego em uma determinada indústria se distancia do padrão geral desta sub-unidade. É um índice de dispersão como qualquer outro porém mais fácil de trabalhar do que o índice de Gini, por exemplo². Assim como o índice de Gini, g_j será tanto menor quanto menor for a dispersão setorial. É interessante normalizar este índice, definindo-se um índice bruto de concentração setorial:

$$G_j \equiv \frac{\sum_i (s_{i,j} - x_i)^2}{1 - \sum_i x_i^2} \quad (2)$$

O grande problema deste índice é que ele não leva em conta o fato de que em alguns setores a escala ideal de produção é maior do que em outros. Independente da decisão das empresas, o setor de maior escala ideal seria “naturalmente” mais concentrado. Não se pretende considerar que a siderurgia é mais concentrada simplesmente porque a sua escala de produção exige que existam apenas poucas dezenas de plantas no país. Uma medida deste fator de escala é dada pelo Índice de Herfindahl-Hirshman (IHH). Existem diversas definições do IHH. Neste caso estamos considerando um índice baseado no emprego. Considere que o setor j possui N_j unidades de negócio cada qual com uma participação $z_{j,1}, z_{j,2}, \dots, z_{j,N_j}$ do emprego do setor. Neste caso o IHH pode ser definido como:

$$H_j \equiv \sum_{k=1}^{N_j} z_{j,k}^2 \quad (3)$$

Imagine que as empresas decidam a sua localização aleatoriamente. Se assim fosse, a probabilidade de uma sub-unidade ser escolhida dependeria do “tamanho” da sub-unidade. Uma boa medida de tamanho seria a proporção do emprego nesta sub-unidade³, ou seja, x_i . Dadas

² Trata-se de um índice antigo na literatura de economia regional. Florence (1948) utiliza um índice semelhante e fornece uma cuidadosa (e longa) justificativa para a sua utilização.

³ Como as empresas estão preocupadas com a presença de consumidores é natural pensar que a medida de tamanho não pode ser dada pela área da sub-unidade.

estas hipóteses, o valor esperado de G_j deveria ser igual ao IHH para a concentração de empresas⁴. Assumindo que o tamanho ideal da planta é uma variável exógena à tomada de decisão, a distância deste índice para a concentração de plantas é um indicador do grau de concentração desta indústria. Note-se que, se este índice não for estatisticamente diferente do índice de Herfindall, não há evidência de “vantagens naturais” ou de *spillovers* neste setor. Para entender esta última afirmação imagine que o lucro da empresa j ao se localizar na sub-unidade i é dado por:

$$\log \pi_{ji} \equiv \log \bar{\pi}_i + \varepsilon_{ji} \quad (4)$$

O que a equação (4) está dizendo é que o lucro da empresa que se localizar na região i será dado pelo retorno médio da região mais um fator aleatório. Assim, regiões com maiores “vantagens naturais” teriam um lucro médio maior e o simples fato de uma empresa se estabelecer nesta região aumentaria o seu retorno esperado. Assumindo empresas maximizadoras de lucro, a probabilidade de que uma empresa decida se instalar na região i é dada por:

$$p_i \equiv \frac{\bar{\pi}_i}{\sum_j \bar{\pi}_j} \quad (5)$$

Quanto maior a distância entre a probabilidade de se instalar em uma região i derivada a partir de um modelo de maximização e a proporção efetiva de empresas na região (x_i), maior o poder de explicação das vantagens naturais. A distribuição de probabilidades é Bernoulli, ou seja, a probabilidade de uma empresa se localizar na região i é p_i e a probabilidade de não se localizar em i é $1-p_i$. Uma maneira de fechar o modelo é assumir que o lucro médio da região é independente do desvio (ε_{jk}) e que a distribuição de probabilidades da decisão locacional das empresas (p_i) tem valor esperado igual à proporção efetiva (x_i) e variância dada por $\gamma_0 x_i(1-x_i)$ onde γ_0 é a correlação entre a decisão de localização de duas empresas no mesmo setor. Note que se γ_0 é igual a zero estamos na mesma situação anterior, ou seja, estamos no modelo definido em (3) onde não há vantagens naturais. Por outro lado, se γ_0 é igual a um, a componente idiossincrática é tão forte que todas as empresas deveriam se localizar na região com maior lucro médio⁵. Ellison e Glaeser (1997), proposição 3, demonstram que num modelo deste tipo o valor esperado de G_j será dado pela seguinte combinação linear⁶:

$$E(G_j) = \gamma_0 + (1 - \gamma_0)H_j \quad (6)$$

Dada esta construção, Ellison e Glaeser propõem um novo índice dado por:

$$\gamma_j \equiv \frac{G_j - H_j}{1 - H_j} \Rightarrow E(\gamma_j) = \frac{E(G_j) - H_j}{1 - H_j} = \gamma_0 \quad (7)$$

⁴ Ellison e Glaeser (1997) proposição 1.

⁵ Da forma como foi construído o modelo, a origem das vantagens naturais está na variância. Note-se também que a vantagem natural de uma região é igual para todos os setores, ou seja, não estamos preocupados com vantagens naturais específicas de um setor o que deve ser captado pelo índice de Herfindall.

⁶ Na realidade há um γ_0 para cada setor. Para simplificar a notação este fato não está explícito na equação (6).

Note que o índice de Ellison e Glaeser (EG de agora em diante) será zero se não houver vantagens naturais⁷ nas regiões. Assim, pode-se dizer que G_j é um índice bruto de concentração industrial no setor j enquanto γ_j é um índice ajustado para escala de produção. O índice bruto será zero apenas se não houver economias de escala entre os setores. O índice de EG será zero se não houver vantagens naturais nas regiões. Ou seja, computando-se dois índices básicos para a concentração de certos setores pode-se ter um retrato apurado de como se distribui a concentração setorial em uma dada região.

Base de dados

Os dados utilizados neste relatório foram obtidos a partir de quatro bases distintas: os microdados das amostras dos Censos Demográficos do IBGE para os anos de 1991 e 2000 e uma tabulação especial⁸ das Relações Anuais de Informações Sociais (RAIS) para os anos de 1991 e 2000.

A partir dos Censos Demográficos obtivemos o número de pessoas que se declararam ocupadas em cada um dos setores, em cada uma das microrregiões estudadas. O setor de atividade corresponde às variáveis V0347 no Censo de 1991 e V4462 no Censo de 2000.

A RAIS é uma pesquisa feita a partir de questionários enviados compulsoriamente ao Ministério do Trabalho e Emprego, preenchido pelos próprios estabelecimentos, contendo dados sobre suas atividades e folhas de pagamento. A partir das RAIS obtivemos o índice de Herfindahl-Hirschman para o número de empregados em cada setor de atividade.

A maior dificuldade observada na utilização dessas bases foi a necessidade de compatibilizar as classificações das atividades econômicas, que apresentam códigos bastante diferentes, cuja compatibilização é impossível sem que se reordenem as classificações, por vezes de forma relativamente arbitrária. A RAIS de 2000 utiliza a Classificação Nacional de Atividades Econômicas (CNAE), que é compatível com a classificação International Standard Industrial Classification of All Economic Activities/Classificación Internacional Industrial Uniforme (ISIC/CIU), proposta pela divisão de estatísticas das Nações Unidas. Esta é a classificação adotada oficialmente pelo Sistema Estatístico Nacional e pelos órgãos federais gestores de registros administrativos desde 1995.

A CNAE apresenta 564 classes de atividades, dispersas em quatro níveis de agregação particionados e indicados por uma letra seguida de um código decimal de cinco dígitos. Esses níveis de agregação denominam-se: Seção (indicado por uma letra do alfabeto), Divisão (dois primeiros dígitos), Grupo (três primeiros dígitos) e, finalmente, Classe (os cinco dígitos). Optamos por adaptar os códigos das outras bases a esta classificação, por sua compatibilidade com as bases internacionais que a seguem.

⁷ Os autores mostram adicionalmente que, sob determinadas condições, o resultado obtido em (6) também vale para um modelo com vantagens de localização dada pela maior circulação da informação (*technological spillover*).

⁸ Agradecemos ao Ministério do Trabalho e do Emprego (MTE) pela tabulação que engloba todas as empresas que preencheram a RAIS (negativa ou não) incluindo o total de empregados em dezembro do ano a que se refere e a massa salarial (para as empresas que apresentaram RAIS positiva).

O Censo Demográfico do IBGE de 2000 utiliza uma classificação chamada CNAE Domiciliar, que procura refletir a CNAE, mas só é compatível na grande maioria dos casos quando reduzida a dois dígitos, ou seja, no nível de agregação denominado Divisão. Ocorre que, em alguns casos, os códigos não são compatíveis nem no nível de Divisão, obrigando-nos a agrupar divisões para compatibilizá-las.

A RAIS de 1991 e o Censo Demográfico de 1991 utilizam um código próprio do IBGE, em uso até 1994, quando foi adotada a CNAE. O IBGE fornece uma tabela de correspondência. Há classes que utilizam o mesmo código segundo essa classificação, mas que aparecem em uma ou mais classes segundo a CNAE, ou seja, em geral a correspondência não é biunívoca. Nesses casos, o que fizemos foi agrupar as divisões onde classes do IBGE apareciam como correspondentes a mais de uma divisão da CNAE.

Assim, para que a comparação histórica pudesse ser feita com o mínimo de distorções, somente pudemos utilizar a CNAE no nível de Divisão (dois dígitos) e, ademais, tivemos que agregar algumas divisões.

Concentração setorial no Brasil e economias de localização

A questão mais essencial que este artigo deve responder é se as empresas estão de fato concentradas. Ainda que diversos autores tenham mostrado que há concentração setorial⁹, neste trabalho nos propomos a uma verificação mais formal desta afirmação, ou seja, podemos testar a hipótese de que o setor está mais concentrado do que se esperaria aleatoriamente.

A Tabela 1 traz os resultados médios obtidos para estes índices calculados para as micro-regiões. A média encontrada para G é de 0,5768 para um H calculado de 0,0173. O valor esperado de G, se não houvesse concentração, seria H_i . Como se pode observar na Tabela B1, o desvio padrão¹⁰ de G está na faixa de 0,005 o que significa que apenas confirmamos mais formalmente a hipótese de concentração setorial levantada por outros autores. A esperança média de G é significativamente maior do que o IHH.

⁹ Para o Brasil vide, por exemplo, os trabalhos de Azevedo e Toneto Jr. (2001), Azoni (1997), Cano (1997), Diniz e Crocco (1996), Pacheco (1999) ou Saboia (2000).

¹⁰ Note que a variância de G pode ser derivada do modelo de comportamento assumido e será dada por:

$$Var(G_i) = \frac{2}{(1 - \sum x_i^2)^2} \left(H_i^2 \left(\sum x_i^2 - 2 \sum x_i^2 + (\sum x_i^2)^2 \right) - \sum_j z_j^4 \left(\sum x_i^2 - 4 \sum x_i^2 + 3(\sum x_i^2)^2 \right) \right) \text{Onde}$$

as somatórias sem subscrito representam somatórias em i . Vide Ellison e Glaeser (1987) proposição 2.

Tabela 1: Indicadores de concentração calculados para microrregiões (1991 e 2000)

Variável	1991		2000		Correlação 2000-1991
	Média	Desvio	Média	Desvio	
H médio	0,0173	0,0402	0,0064	0,0095	0,6112
G médio	0,5768	1,6192	0,6811	1,8920	0,9914
% concentrada*	79,41%	0,4104	80,56%	0,4014	
γ médio	0,5589	1,6263	0,6766	1,8943	0,9917
Setores	34		36		

* porcentagem dos setores para os quais $G_i > H_i$ com 95% de segurança.

Fonte: Tabulação própria a partir dos Censos de 1991 e 2000 e Tabulação Especial da RAIS para 1991 e 2000.

Cerca de 80% dos setores apresentam concentração bruta significativamente acima do IHH. Apenas 7 setores (em 1991 ou em 2000) não apresentam concentração bruta acima do esperado. Os setores de fumo (16), petróleo (23) e os diversos setores de utilidades públicas¹¹ (40, 41 e 90). No caso do setor de fumo, em 1991 o IHH era de 0,11, muito acima da média. Ou seja, os dois primeiros setores (fumo e petróleo) poderiam ser consideradas classificações aleatórias se a escala das empresas não fosse tão elevada. O mesmo resultado se observa para os setores de extração de não metálicos em 1991 e de metálicos em 2000. A desconcentração observada para os serviços de utilidade pública provavelmente está menos ligada à escala das plantas. Com exceção da coleta de lixo em 1991 para a qual o IHH também foi de 0,11, os índices estão sempre abaixo de 0,03. Os serviços públicos são dispersos pois esta é a natureza do negócio. Este comportamento pode também ser observado para os Correios e Telecomunicações¹² em 1991.

Um problema desta amostra é que alguns setores apresentam concentração acima do que seria plausível. Lembre-se que a equação (6) implica que $E(G_j) = \gamma_o + (1 - \gamma_o)H_j$ onde γ_o foi estimado a partir do índice E-G e aparece na última linha da Tabela 1 acima. O parâmetro relevante em (6) será igual a 1 se as vantagens comparativas ou *spillovers* forem o fator relevante de concentração e igual a zero se não houver nenhuma economia de aglomeração efetiva (independente da origem). O problema é que alguns setores apresentaram um índice E-G acima de 1. Isto significa que a concentração devida à escala na realidade reduz a concentração *vis a vis* as economias de localização. É um pouco difícil de acreditar neste resultado.

¹¹ Eletricidade, Gás, Água Quente, Captação, Purificação e Distribuição de Água, Coleta de Lixo, Águas Residuais e Esgoto.

¹² Este setor pode ser considerado concentrado em 2000 devido à queda no IHH e não ao aumento da concentração bruta. A concentração por plantas não era elevada em 1991 (0,03) mas caiu consideravelmente na década para 0,008. Considerando que este setor foi altamente privatizado na década é interessante notar que houve redução considerável na escala das plantas. Este resultado certamente merece um estudo mais cuidadoso.

Em particular o setor de “agropecuária e exploração florestal” apresenta um grau de concentração acima de 110! Certamente isto não faz o menor sentido. Assim, provavelmente o índice não é eficiente para medir o grau de concentração da agropecuária. Por este motivo acabamos excluindo os setores agropecuários da análise ainda que a “pesca e aqüicultura” tenha apresentado índices de concentração razoáveis. Construção, comércio, administração pública e ensino também apresentam concentração acima de 1 (mas não valores tão díspares quanto a agropecuária). Também desconfiamos dos dois primeiros setores. Mais uma vez pode ser que o índice não seja eficiente para a análise do setor de construção. Especialmente se considerarmos que o IHH foi calculado a partir da localização da obra enquanto o G foi calculado a partir do local de moradia do trabalhador da construção. Desconfiamos do comércio pois a agregação de setores foi muito alta juntando todos os setores em um único compreendendo 23% da força de trabalho. Por este motivo, apresentamos as estatísticas da Tabela 1 excluindo os setores com índice E-G acima de 2.

Tabela 2: Indicadores de concentração calculados para microrregiões (1991 e 2000) com amostra restrita*

Variável	1991		2000		Correlação 2000-1991
	Média	Desvio	Média	Desvio	
H	0,0183	0,0413	0,0069	0,0098	0,6053
G	0,2513	0,3848	0,2536	0,4846	0,8816
% concentrada**	78,13%	0,4200	78,79	0,4151	
γ	0,2323	0,3999	0,2487	0,4897	0,8750
N	32		33		

* Retirou-se setores com $\gamma > 2$

** porcentagem para a qual $G > H$ com 95% de segurança.

Fonte: Tabulação própria a partir dos Censos de 1991 e 2000 e Tabulação Especial da RAIS para 1991 e 2000.

A amostra restrita apresenta G médio (0,25) bem abaixo da amostra ampla mas ainda assim significativamente acima de H. Ou seja, mais uma vez a hipótese de concentração sobre os setores em geral se confirma também na amostra restrita. Também para a amostra restrita quase 80% dos setores apresentam concentração significativa. Isto implica que o resultado de concentração significativa para os setores em geral sobrevive à eliminação de *outliers*. Uma forma mais robusta de testar a hipótese de concentração generalizada seria rodar a seguinte regressão:

$$G_{jt} = \alpha_0 + \alpha_1 H_{jt} + \varepsilon_{jt} \quad (8)$$

E testar a hipótese (conjunta) $\alpha_0 = 0$ e $\alpha_1 = 1$. As Tabelas 3 e 4 apresentam o resultado separado para 1991 e 2000 e mostram que a hipótese pode ser negada tranquilamente tanto pelo teste de cada um dos coeficientes quanto no teste conjunto. Empilhando os dois anos o resultado

(omitido por questão de espaço) é semelhante. Na realidade além do intercepto ser significativo, a inclinação (α_1) não é significativa sendo desnecessário o teste da hipótese conjunta. No entanto, a hipótese inversa de concentração total, quer dizer, $\alpha_0 = 1$ e $\alpha_1 = 0$ também não pode ser aceita, ainda que a hipótese não conjunta sobre o intercepto ($\alpha_1 = 0$) possa ser aceita. Isto implica que embora não seja possível negar a hipótese de concentração via economias de localização também não é possível aceitar a hipótese forte de concentração essencialmente por economias de aglomeração.

Tabela 3: Teste para aleatoriedade da localização (1991)

Variável	Coefficiente	Desvio Padrão	t
constante	0,286	0,074	3.850
H	-1,885	1,666	-1.130
$R^2 = 0,0409$			

Fonte: *Tabulação própria a partir dos Censos de 1991 e 2000 e Tabulação Especial da RAIS para 1991 e 2000.*

Tabela 4: Teste para aleatoriedade da localização (2000)

Variável	Coefficiente	Desvio Padrão	t
constante	0,308	0,106	2.900
H	-7,601	9,850	-0.770
$R^2 = 0,0409$			

Fonte: *Tabulação própria a partir dos Censos de 1991 e 2000 e Tabulação Especial da RAIS para 1991 e 2000.*

Conforme argumentamos, o índice E-G é a melhor maneira de avaliar o grau de concentração devido à economia de escala. As Figuras 1a e 1b apresentam o Histograma de γ em 1991 e 2000 para a amostra restrita ($\gamma < 2$). Como se pode observar a distribuição é assimétrica com uma grande concentração entre -0,05 e 0 tanto em 1991 (10 setores) como em 2000 (17 setores). Assumindo um desvio padrão¹³ entre 0,02 e 0,03, todos os setores com índice E-G abaixo de

¹³ O cálculo do desvio padrão de γ é bastante complexo. Isto porque o cálculo da variância depende da covariância entre a decisão das plantas. Assim, é necessário assumir alguma distribuição para a probabilidade de localização. Seguindo o procedimento proposto por Ellison e Glaeser (1997) que assumem que as probabilidades se distribuem

como Dirichlet com parâmetros $\left(\frac{1-\gamma_0}{\gamma_0} x_1, \frac{1-\gamma_0}{\gamma_0} x_2, \dots, \frac{1-\gamma_0}{\gamma_0} x_M \right)$ A variância foi estimada por simulação

“Monte Carlo” utilizando o script em matlab© gentilmente fornecido por Bruce Sacerdote. Os nossos resultados são semelhantes aos encontrados por Ellison e Glaeser, ou seja, entre 0,02 e 0,03 em 95% dos casos, ainda que valores bem mais elevados (até 0,1) tenham sido verificados em algumas simulações.

0,05 não podem ser considerados concentrados por motivos ligados a economias de escala. Isto significa que 15 setores em 1991 e 29 setores em 2000 não podem ser considerados concentrados por motivos ligados a economias de escala o que contradiz o resultado anterior e a maioria das análises usuais de concentração das atividades econômicas.

Figura 1a: Histograma de Gama para Microrregiões (1991)

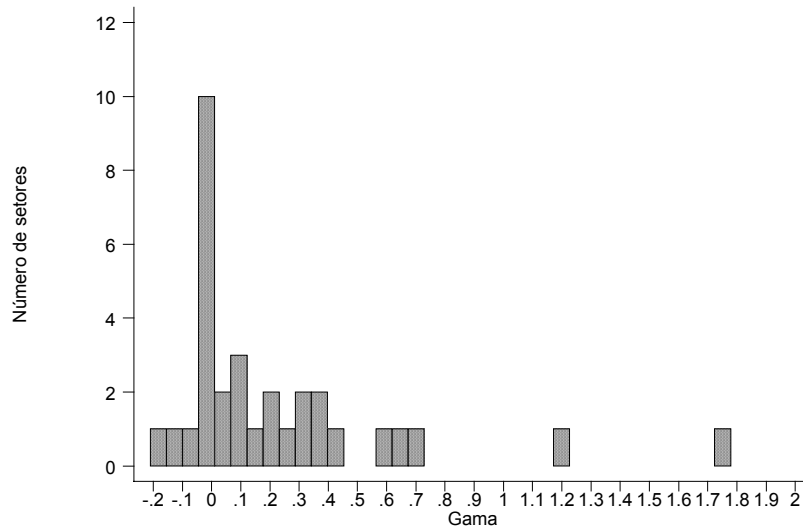
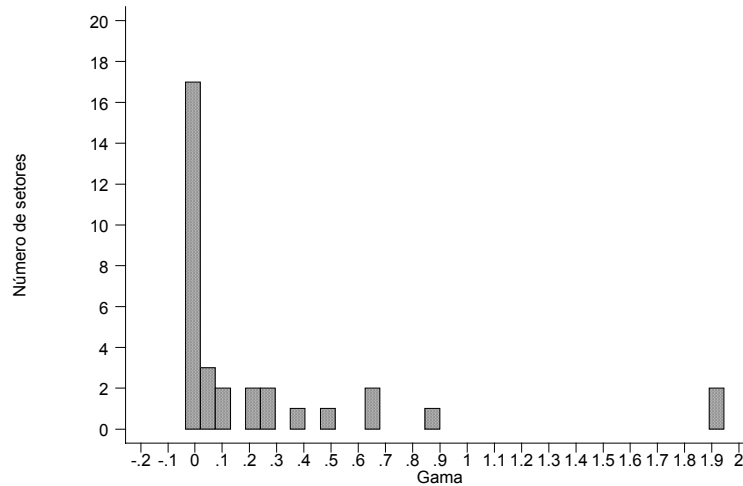


Figura 1b: Histograma de Gama para Microrregiões (2000)



Fonte: *Tabulação própria a partir dos Censos de 1991 e 2000 e Tabulação Especial da RAIS para 1991 e 2000.*

O resultado encontrado para o Brasil é muito semelhante ao encontrado para os Estados Unidos por Ellison e Glaeser (1997). Ainda que a grande maioria dos setores (80%) possa ser considerada concentrada em termos brutos, para 45% dos setores a concentração não está associada a economias de localização. Lembrando que o índice de E-G pode ser considerado uma medida da importância da localização e desconsiderando os *outliers*, apenas para cinco setores gama é maior do que 0,5. Isto implica que apenas para 5 setores as economias de aglomeração respondem por mais de 50% da decisão locacional da empresa.

Tabela 5: Índices médios de concentração por faixa de γ .

Faixa	H	G	γ	% setores	% concentradas
$\gamma \leq 0$	0.0426	0.0189	0.0275	26%	7%
$0 < \gamma \leq 0.05$	0.0028	0.0211	0.0183	19%	23%
$0.05 < \gamma \leq 1$	0.0024	0.3095	0.3081	43%	54%
$\gamma > 1$	0.0041	3.8435	3.8464	13%	16%

Fonte: Tabulação própria a partir dos Censos de 1991 e 2000 e Tabulação Especial da RAIS para 1991 e 2000.

A Tabela 5 apresenta os índices médios para diversas faixas de gama para a amostra empilhada. Como seria de se esperar, os setores com gama negativo apresentam IHH bastante elevado. Isto significa que as deseconomias de aglomeração fazem com que este setor seja menos concentrado do que deveria ser apenas em função das economias de escala que são relativamente elevadas. A última coluna indica a porcentagem de setores considerados concentrados quando se analisa apenas a concentração bruta que está em cada faixa para ilustrar a diferença entre os resultados obtidos a partir de um índice de concentração bruto *vis a vis* um índice de economia de localização. É verdade que 70% das empresas consideradas concentradas pelo critério bruto apresentam alguma vantagem locacional. No entanto, 30% destes setores concentrados em termos brutos não apresentam vantagens de localização significativa.

Escopo regional da concentração setorial

Uma pergunta relevante que precisa ser respondida refere-se ao escopo regional. Por um lado seria de se esperar que as vantagens locais diminuíssem conforme aumenta a distância. As Figuras 2a e 2b apresentam o histograma de gama considerando a distribuição de mesorregiões enquanto as Figuras 3a e 3b apresentam os histogramas para a divisão por Estado. Se não houver diferença entre as distribuições significa que toda a vantagem locacional se esgota na escala da microrregião. O que se observa, no entanto, é que as empresas são mais localizadas na escala da microrregião do que na escala da mesorregião ou na escala estadual. A concentração observada

na escala estadual é maior do que nas mesorregiões mas ainda menor do que a observada na escala da microrregião. Isto vale para a grande maioria dos setores.

Este resultado é de difícil interpretação mas indica que as vantagens locacionais se esgotam na escala microrregional. No entanto, consideramos ainda um pouco prematuro realizar qualquer afirmação mais categórica com relação ao escopo geográfico. Uma abordagem que provavelmente permitiria uma análise mais profunda seria considerar o efeito vizinhança entre as microrregiões utilizando técnicas de econometria espacial. Pretendemos realizar esta análise em trabalhos posterior.

Figura 2a: Histograma de Gama para Mesorregiões (1991)

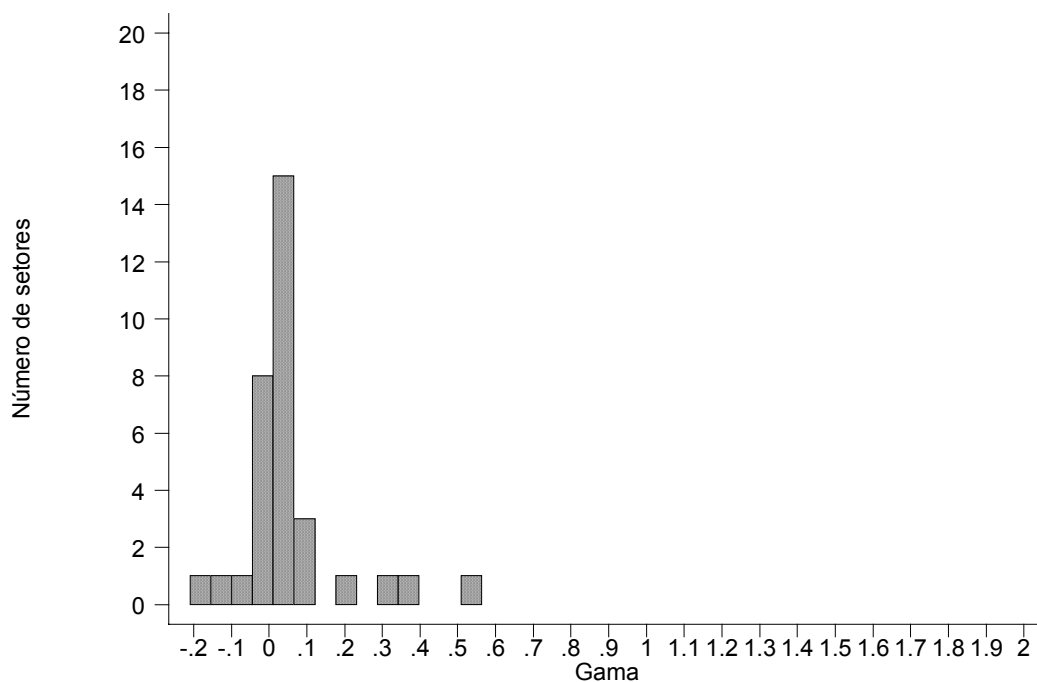
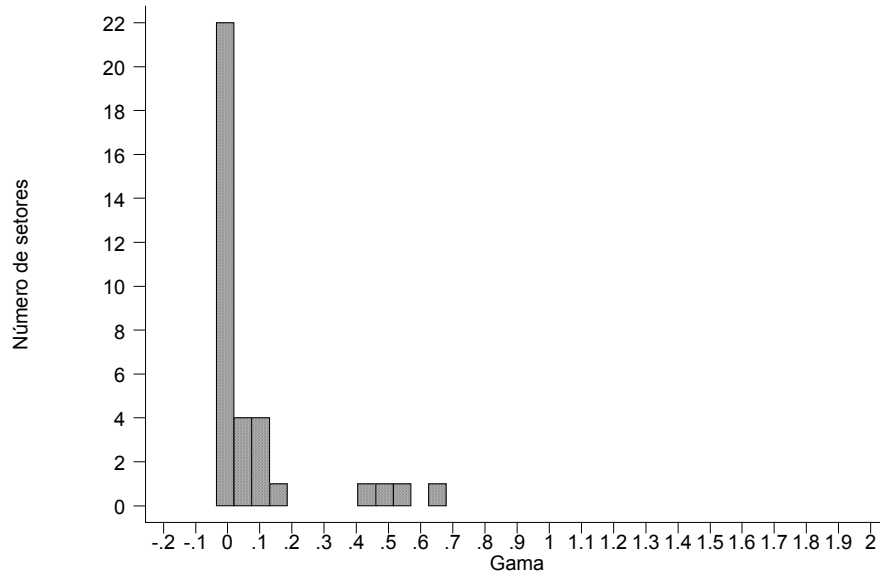


Figura 2b: Histograma de Gama para Mesorregiões (2000)



Fonte: *Tabulação própria a partir dos Censos de 1991 e 2000 e Tabulação Especial da RAIS para 1991 e 2000.*

Figura 3a: Histograma de Gama para Estados (1991)

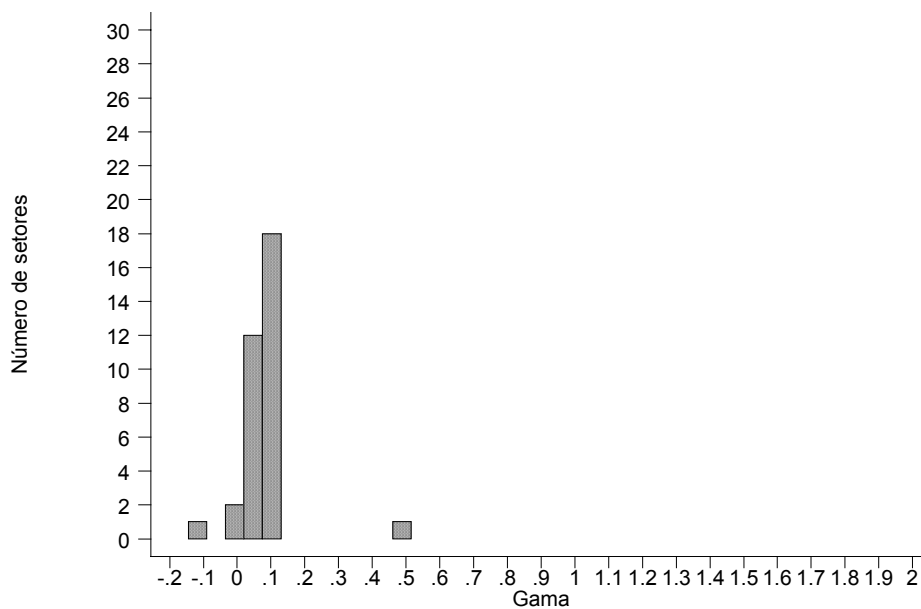
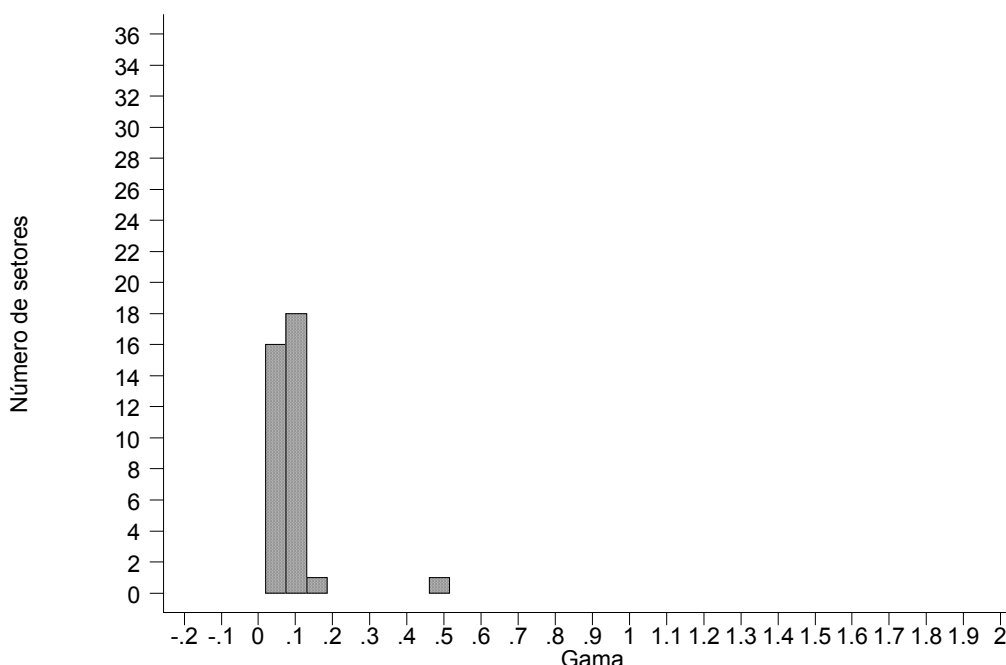


Figura 3b: Histograma de Gama para Estados (2000)



Fonte: Tabulação própria a partir dos Censos de 1991 e 2000 e Tabulação Especial da RAIS para 1991 e 2000.

A Dinâmica da Concentração nos anos 1990

Como se pode observar nas Tabelas 1 e 2, a correlação para G e γ entre 1991 e 2000 é muito alta. Para a amostra inteira a correlação é de mais de 99%. Os valores médios também não se alteram significativamente ao longo dos anos indicando que o grau de concentração regional e de concentração devido a economias de localização permaneceu essencialmente constante ao longo da década. Este resultado contrasta com a grande variação efetivamente observada nas empresas. Todos os anos milhares e milhares de empregos são criados ou destruídos. Como então é possível que a concentração permaneça a mesma?

A questão pode ser respondida se pensarmos que a destruição de empregos devido ao fechamento de empresas é compensada pela geração de empregos das novas empresas criadas. Este é um resultado conhecido em estatística. Se há grande variação de altura entre as gerações mas a altura média da população permanece a mesma, pais baixos devem ter filhos altos e vice-versa. A maneira que nos propomos a analisar a dinâmica da concentração industrial é justamente separando estas duas componentes da variação, ou seja, o quanto a variação na média é compensada por mudanças aleatórias. Considere a regressão:

$$s_{i,j,t+1} - s_{i,j,t} = \alpha + \beta(s_{i,j,t} - x_{i,t}) + \eta(x_{i,t+1} - x_{i,t}) + \varepsilon_{i,j,t} \quad (8)$$

Onde $x_{i,t}$ indica a proporção da sub-unidade geográfica i na força de trabalho no ano t e $s_{i,j,t}$ representa a proporção do emprego no setor j nesta sub-unidade no ano t ; α , β e η são parâmetros a serem estimados e $\varepsilon_{i,j,t}$ é o termo aleatório que, por construção, é ortogonal aos regressores. Note que, dada a especificação desta equação, se ela for estimada por mínimos quadrados ordinários, sempre deve gerar $\alpha = 0$ e $\eta = 1$. Definindo $g_t \equiv (1/J) \sum_j g_{j,t}$ como a média do grau de concentração bruto não normalizado entre as indústrias no ano t , a variação desta média no período pode ser escrita como:

$$\Delta g_t \equiv g_{t+1} - g_t = \frac{1}{J} \left[\sum_{i,j} (s_{i,j,t+1} - x_{i,t+1})^2 - \sum_{i,j} (s_{i,j,t} - x_{i,t})^2 \right] \quad (9)$$

Substituindo (8) em (9) temos que:

$$\Delta g_t = \frac{1}{J} \left[\sum_{i,j} (s_{i,j,t} + \alpha + \beta(s_{i,j,t} - x_{i,t}) + \eta(x_{i,t+1} - x_{i,t}) + \varepsilon_{i,j,t} - x_{i,t+1})^2 - \sum_{i,j} (s_{i,j,t} - x_{i,t})^2 \right] \quad (10)$$

Se lembrarmos que, por construção, $\alpha = 0$ e $\eta = 1$ com um pouco de álgebra é fácil de verificar que a variação da concentração bruta pode ser definida como:

$$\Delta g_t = \frac{1}{J} \left[\sum_{i,j} ((1 + \beta)(s_{i,j,t} - x_{i,t}) + \varepsilon_{i,j,t})^2 - \sum_{i,j} (s_{i,j,t} - x_{i,t})^2 \right] = (2\beta + \beta^2)g_t + \frac{1}{J} \sum_{i,j} \varepsilon_{i,j,t}^2 \quad (11)$$

Esta equação decompõe a variação na concentração média em dois termos. O primeiro termo, $(2\beta + \beta^2)g_t$, depende da correlação no tempo entre a representação relativa do setor na região. Por isto denominamos este termo de “efeito médio”. Se este efeito é negativo, o emprego em microrregiões inicialmente relativamente concentradas em um determinado setor tendem a diminuir enquanto microrregiões inicialmente desconcentradas tendem a aumentar o emprego relativo. O segundo termo, $\frac{1}{J} \sum_{i,j} \varepsilon_{i,j,t}^2$, será denominado “efeito de aleatoriedade” pois a sua magnitude reflete o grau de heterogeneidade entre as microrregiões que inicialmente têm grau de concentração semelhante em um determinado setor. Os resultados para a amostra de setores e para algumas agregações de setores são apresentados na Tabela 6 abaixo.

Tabela 6: Variação na concentração setorial bruta para diferentes grupos de setores

Grupo	N	Correlação entre as proporções setoriais	γ_{91}	β	Soma dos quadrados dos resíduos	Variação da Concentração bruta		
						Total	Reversão da média	Dispersão
Integral	34	0,907	0,559	-0,048	2,224	1,1%	-5,4%	6,5%
Restrita	32	0,776	0,232	-0,199	1,788	-3,4%	-9,0%	5,6%

Concentradas	8	0,772	0,751	-0,229	0,926	-18,9%	-30,5%	11,6%
Desconcentradas	8	0,629	-0,052	-0,297	0,020	-0,6%	-0,9%	0,3%
Não qualificados	5	0,740	0,318	-0,283	0,436	-7,9%	-16,6%	8,7%
Qualificados	6	0,638	0,087	-0,402	0,159	-3,2%	-5,9%	2,7%
Terc. Superior	4	0,387	0,045	-0,584	0,185	0,5%	-4,1%	4,6%
Serv. Públicos	8	0,911	0,386	0,017	0,292	5,0%	1,4%	3,7%
Serv. Pessoais	2	0,553	0,350	-0,205	0,188	-3,5%	-12,9%	9,4%

Fonte: Tabulação própria a partir dos Censos de 1991 e 2000 e Tabulação Especial da RAIS para 1991 e 2000.

A decomposição de g consegue explicar as duas evidências aparentemente inconsistentes: a grande variação na escala da empresa e a pequena variação na concentração agregada. A reversão da média compensada pela dispersão implica em variações muito maiores do que a observada no índice agregado. Note que este resultado poderia não ocorrer. O efeito da heterogeneidade é certamente positivo já que se trata da soma de termos quadrados. No entanto a reversão da média pode ser tanto positiva como negativa dependendo do sinal de β . Apenas para valores negativos do coeficiente a variação agregada será menor do que a variação observada na escala da empresa. As nossas estimativas indicam que β é quase sempre negativo. A única exceção são os serviços públicos para o qual a variação agregada é maior do que suas componentes. Isto significa que microrregiões originalmente mais concentradas em serviços públicos aumentaram esta concentração. Esta conclusão é reforçada pela correlação de 91% entre as proporções destes setores em 1991 e 2000. Os serviços públicos são os setores de atividade com maior inércia em termos de concentração. Partiram de um alto grau de concentração em 1991 que apenas aumentou na década de 1990.

É interessante observar o comportamento dos grupos. Para a amostra integral a concentração aumentou ligeiramente na última década. No entanto o efeito reversão da média foi cerca de 5 vezes o efeito total. Na realidade o que induziu ao aumento tímido da concentração foi o efeito heterogeneidade. Quando retiramos os dois setores com γ maior do que 2 nota-se que a concentração caiu mas, também neste caso, o efeito reversão foi muito superior ao efeito total. Os 8 setores originalmente concentrados reduziram consideravelmente sua concentração mas o efeito teria sido ainda maior não fosse o alto grau de heterogeneidade entre as microrregiões. De fato os setores originalmente concentrados apresentam uma forte tendência à desconcentração. O efeito reversão da média é de mais de 30%. É também o grupo de setores com maior efeito aleatório (11,6%).

O resultado para os setores mais concentrados é esperado. Em princípio as perdas de escala devem aumentar com a concentração. Aparentemente as perdas por excesso de concentração são bem elevadas, dada a reversão da média observada. O resultado para os setores originalmente desconcentrados é também esperado. Tais setores não variaram significativamente o grau de concentração ao longo da década. Quando observamos a decomposição notamos que nenhum dos termos é significativo. Este resultado é compatível com o que se espera de setores pouco

concentrados. Como não há economia de aglomeração significativa para estes setores, partindo de um baixo grau de concentração não há motivo para os setores se desconcentrarem ainda mais.

Quando analisamos a concentração em função da qualificação do setor o primeiro aspecto que chama a atenção é que os setores menos qualificados são originalmente (em 1991) concentrados enquanto os mais qualificados são originalmente desconcentrados. De qualquer maneira ambos estão reduzindo o grau de concentração ainda que o ritmo seja maior para os setores menos qualificados. Na realidade, a velocidade de desconcentração das indústrias menos qualificadas só é maior pois parte de um nível mais elevado em 1991. Note que o β calculado para os setores mais qualificados é um dos mais altos (em termos absolutos) entre os grupos ficando abaixo apenas dos setores do terciário superior.

O fato de setores menos qualificados serem concentrados não é nem esperado nem contraditórios com a teoria. Como vimos a origem da concentração pode ser tanto de spillovers quanto de vantagens naturais. Não há nada que diga que um setor de baixa qualificação não dependa de algum recurso natural. Tomando um sub-grupo dos setores de baixa qualificação, os serviços pessoais, o resultado se mantém. No entanto, o efeito reversão da média para os setores menos qualificados em geral é significativamente maior do que para os serviços pessoais enquanto os efeitos da heterogeneidade para os setores dos serviços pessoais são mais pronunciados do que para os demais setores de baixa qualificação. O resultado é que a desconcentração total é maior para o setor mais amplo.

A redução da concentração dos setores de baixa qualificação é compatível com a queda nos custos de transporte. Isto implicaria em uma queda nas vantagens de aglomeração e, portanto, diminuição do incentivo à concentração. Também é razoável um efeito menos pronunciado nos setores mais qualificados. Isto porque estes setores devem estar mais próximos da mão-de-obra qualificada que, em princípio, não foi afetada pela redução no custo de transportes. O resultado é ainda mais intenso para os serviços do terciário superior que aumentou a sua concentração na década. No entanto, o baixo grau de concentração dos setores qualificados não é perfeitamente compatível com a teoria. Setores de alta qualificação deveriam se localizar onde há concentração de mão-de-obra qualificada. Chama a atenção a correlação de apenas 39% entre as proporções setoriais no início e final da década. Sem dúvida esta questão merece pesquisas posteriores que permitam compreender a natureza do processo que está ocorrendo nestes setores fundamentais para a dinâmica econômica do país.

Referências

- Azevedo, P.F. e R. Toneto Jr. (2001) “Relocalização do emprego industrial formal no Brasil na década de 90”. *Pesquisa e Planejamento Econômico*, 31 (1).
- Azoni, C.R. (1997) “Concentração regional e dispersão das rendas *per capita* estaduais: análise a partir das séries históricas estaduais de PIB, 1939-1995.” *Estudos Econômicos*, 27 (3).
- Baer, W., E. Haddad e G. Hewings (1998) “The regional impact of neo-liberal policies in Brasil”. *Economia Aplicada*, 2 (2).

- Bonelli, R. (2000) “Emprego industrial e produtividade: novos resultados, velhas controvérsias”. *Revista Brasileira de Comércio Exterior*, 62.
- Cano, W. (1997) “Auge e inflexão da desconcentração econômica regional” in: Affonso, R.B.A. e P.L.B. Silva (1997) *A Federação em perspectiva: ensaios selecionados*. São Paulo: Hucitec.
- Carlton, D. (1983) “The location and employment decision of New Firms: An Econometric Model with Discrete and Continuous Exogenous Variables”. *Review of Economic and Statistics*, 65.
- Diniz, C.C. e M.A. Crocco (1996) “Reestruturação econômica e impacto regional: o novo mapa da indústria brasileira”. *Nova Economia*, 6 (1).
- Ellison, G. e E. Glaeser (1997) “Geographic Concentration in the U.S. Manufacturing Industries: A Dartboard Approach”. *Journal of Political Economy* 105(5).
- Enright, M. (1990) *Geography Concentration and Industrial Organization*. Ph.D. thesis, Harvard University.
- Florence, P.S. (1948) *Investment, Location and Size of Plant*. London: Cambridge University Press.
- Fuchs, V. (1962) *Changes in the Location of Manufacturing in the US since 1929*. New Haven: Yale University Press.
- Henderson, J.V. (1988) *Urban Development: Theory, Fact and Illusion*. New York: Oxford University Press.
- Hoover, E.M. (1948) *The Location of Economic Activity*. New York: McGraw Hill.
- Jaffe, A., M. Trajtenberg e R. Henderson (1993) “Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations”. *Quarterly Journal of Economics*, 108.
- Kim, S. (1995) “Expansion of Markets and the Geographic Distribution of Economic Activities: the Trends in U.S. Regional Manufacturing Structure, 1860-1987”. *Quarterly Journal of Economics* 110(4).
- Krugman, P. (1991) “Increasing Returns and Economic Geography”. *Journal of Political Economy*, 99.
- Marshall, A. (1896) *Principles of Economics*. London: Macmillan.
- Oliveira, C.W. e Guimarães Neto, L. (1997) “Emprego organizado e regiões nos anos 90: quem perdeu mais?” *Estudos Econômicos*, 27 (número especial).
- Pacheco, C.A. (1999) “Novos padrões de localização industrial? Tendências recentes dos indicadores de produção e do investimento industrial”. *Textos para Discussão IPEA*, 633.
- Porter (1990) *The Competitive Advantage of Nations*. New York: The Free Press.

- Saboia, J. (2000) “Desconcentração industrial no Brasil nos anos 90: um enfoque regional”. *Pesquisa e Planejamento Econômico*, 30 (1).
- Schmalensee, R. (1977) “Using the H-index of Concentration with Published Data”. *Review of Economics and Statistics*, 59.
- Suzigan, W., J. Furtado, R. Garcia e S.E.K. Sampaio (2001) “Aglomerações industriais no Estado de São Paulo”. *Economia Aplicada*, 5 (4).
- Zini Jr., A.A. (1998) “Regional income convergence in Brasil and its socio-economic determinants”. *Economia Aplicada*, 2 (2).