Distribuição e Dinâmica do Tamanho de Empresas Industriais

Eduardo Pontual Ribeiro[1]

Resumo: O estudo da distribuição e dinâmica do tamanho das empresas é tópico de interesse de pesquisadores em organização industrial e mercado de trabalho. A hipótese mais comum é a chamada Lei de Gibrat, que afirma que a taxa de crescimento de uma empresa é independente de seu tamanho. O objetivo deste trabalho é testar a hipótese de Gibrat através de uma metodologia nova na literatura e trazendo resultados para o caso inexplorado do Brasil. A análise condicional, via regressão, será baseada no método de regressão quantílica (*quantile regression*), que permite a estimação semi-paramétrica da distribuição condicional da variável em estudo, ao invés do primeiro momento apenas. O problema de inconsistência dos estimadores devido a termos individuais não observados também é levado em conta. Os resultados sugerem que empresas menores crescem mais rápido, sendo o crescimento pouco persistente.

Palavras chave: Indústria no Brasil, Lei de Gibrat, Emprego.

JEL: L11, C33

Abstract: The shape and evolution of the firm size distribution is a topic of interest for many researchers in industrial organization and labor economics. The most common hypothesis is Gibrat´s Law of Proportional Effect. It states that the size growth rate of firms is independent of size. The goal of this article is to test Gibrat´s Law using a nem empirical methodology and considering the unexplored Brazilian case. The regression method uses *quantile regression*. We estimate the evolution of the whole firm size disitrubion instead of previous studies that considered the conditional mean dynamics. Estimate inconsistency problems are addressed. The main issue is that common IV/GMM methods are innapropriate under the null but consistent under the alternative, while non-IV methods imposing exogeneity are consistend only under the alternative. Results suggest that smaller firms grow faster, but its growth is not much persistent.

Key Words: Gibrat´s Law, Quantile Regression, Brazil.

JEL:L11, C33

[1] PPGE/UFRGS e UFF e Pesquisador CNPq. Este trabalho foi escrito durante o período como Professor Visitante na Escola Nacional de Ciências Estatísticas/IBGE-Fundação Ford. As opiniões e estimativas apresentadas não representam nem são endossadas, necessariamente, pelo IBGE ou a Fundação Ford. Contato: eribeiro@ufrgs.br

1. Introdução

A compreensão de como as empresas crescem e como os mercados se organizam são pontos de considerável interesse em economia. Estimativas da distribuição do tamanho de empresas, seja este tamanho medido pelo número de empregados ou faturamento, são importantes para o acompanhamento macroeconômico e políticas públicas de geração de emprego e renda e de desenvolvimento industrial. Um dos primeiros estudos sistemáticos sobre o tamanho de empresas foi feito por Gibrat (1931) –veja síntese em recente em Sutton (1997)–, que identificou uma distribuição assimétrica para o tamanho das empresas industriais, para a França no início do século XX, e propôs um modelo de crescimento de empresas que seria independente de seu tamanho. Esta regularidade empírica recebeu o nome de *Lei de Gibrat*. Um modelo estocástico simples sugerido por Gibrat, com choques Gaussianos indicaria uma distribuição Log-Normal para o tamanho das empresas. Desta forma, uma medida de concentração seria obtida a partir dos parâmetros de dispersão da distribuição.

A Lei de Gibrat tem sido empregada por alguns autores, como Ijuri e Simon (1967) e Lucas (1978) para construir modelos mais sofisticados de comportamento da distribuição de empresas, modelos estes que tentam superar as deficiências do modelo original, como apontado por Kalecki (1945). Em particular Kalecki mostrou que se o modelo original estivesse correto, a dispersão de tamanhos cresceria sem limites o que não é verificado na prática. Isto por que o modelo de Gibrat impõe que o tamanho de uma empresa é um *random walk*.

Todavia vários estudos internacionais sugerem que a Lei de Gibrat, mesmo em versões modificadas, talvez não tenha suporte empírico. Um consenso inicial destes trabalhos, particularmente Evans (1987) e Hall (1987), seria de que a taxa de crescimento de uma empresa seria decrescente com o seu tamanho. Outros pesquisadores argumentam que a Lei pode ser válida para alguns casos, uma vez controlando um tamanho mínimo ou o problema de entrada e saída de empresas, por criação e falência.

Como a hipótese da Lei de Gibrat é extremamente popular na literatura internacional e a mesma não foi estudada no Brasil, o objetivo principal desta pesquisa será estimar a distribuição do tamanho das empresas industriais no Brasil, baseado nos dados da PIA para verificar a validade da hipótese da Lei de Gibrat para a distribuição de tamanho de empresas.

Do ponto de vista teórico e empírico a Lei de Gibrat apresenta relações com outras áreas além de organização industrial. Seu modelo dinâmico pode ser associado à evolução de distribuições de renda e tamanho (de países ou cidades), mobilidade e ao estudo da dinâmica do mercado de trabalho, em que a dispersão da criação e destruição de postos de trabalho (ou seja, o crescimento o emprego nas empresas) tem papel importante na compreensão na dinâmica do mercado de trabalho (ver por exemplo Davis, Haltiwanger and Shuh, 1996 e Hamermesh e Pfann, 1996).

Do ponto de vista econométrico o teste da lei de Gibrat apresenta particulariedades importantes. A primeira vista a hipótese implica que a evolução do tamanho das empresas é um passeio aleatório, encontrando contexto com a literatura de testes de raiz unitária. Todavia, os dados para testar a hipótese em geral se apresentam como um painel com reduzida dimensão temporal, inviabilizando o uso de testes que dependam da dimensão temporal, *T*, crescendo para derivações assintóticas. Da mesma forma, métodos usuais para estimação em modelos dinâmicos com dados em painel, como IV ou GMM impõem que os dados sejam estacionários. Desta forma, estimadores que seriam consistentes para o contexto de estimação, não o são sob a hipótese nula a ser testada, dificultando o teste da hipótese.

Apenas nos últimos anos há trabalhos que contribuem para este debate. De modo interessante, a conclusão de trabalhos como Bond, Nauges e Windmeijer (2002) e Hall e Mairesse (2002). Os autores sugerem o uso de estimadores como mínimos quadrados, com exogeneidade das explicativas em um modelo onde o tamanho é um passeio aleatório sob a hipótese nula e um processo autoregressivo com drift sob a hipótese alternativa. Testes de raiz unitária baseado em estimadores com IV/GMM não possuem propriedades melhores que os testes baseado em MQO.

Nossa idéia é explorar um outro aspecto da lei de Gibrat, que é o estudo da evolução da distribuição de tamanho das empresas considerando toda a distribuição, ao invés do foco na média condicional. Desta forma, nossa análise se aproxima, por exmeplo, do estudo do processo de convergência de rendas per capita dentro de uma análise markoviana (Quah, 1994, *e.g.*) e matrizes de transição, em contraposição a *growth regressions* popularizados por Barro e Sala-i-Martin (1994)[2].

Os resultados sugerem que a lei de Gibrat é rejeitada, seja com os estimadores de MQ e regressão quantílica seja com os estimadores de variáveis instrumentais. As menores empresas temdem a crescer mais rápido que as grandes, havendo uma forte assimetria no crescimento das empresas maiores. Em outras palavras, parece haver um limite superior para o crescimento das emrpesas maiores. Este fato não é identificado através de modelos de média condicional.

O artigo está organizado da seguinte forma. A segunda seção apresenta o referencial teórico e a terceira seção o modelo econométrico a ser empregado, ou seja, o modelo empírico e os métodos de estimação. A quarta seção apresenta uma visão geral dos dados empregados e a quinta seção os resultados empíricos. A última seção agrega os comentários finais do trabalho.

2. Referencial teórico

O estudo da distribuição do tamanho das empresas em uma economia sempre atraiu interesse dos pesquisadores em economia. Talvez a hipótese mais representativa destes estudos seja a *Lei de Gibrat* (Sutton, 1997). Gibrat identificou que a distribuição do tamanho das empresas (seja medido por pessoal empregado ou por vendas ou valor agregado) era bem aproximada por uma distribuição Lognormal, ou seja, eminentemente assimétrica. Para construir um modelo que especificasse a evolução do tamanho das empresas e ao mesmo tempo consistente com a verificação empírica, o autor propôs que o tamanho de um empresa i em um período t, Y_{it}, seja função do tamanho da empresa no período anterior (Y_{it-1}) multiplicado por um termo aleatório ($v'_{it} = v_{it}+a_i$) exponencial com média a_i (impomos que o valor esperado de v_{it} seja zero para normalização), isto é,

$$Y_{it} = Y_{it-1}\, exp(v_{it} + a_i) \ .$$

Desta forma, o comportamento do logaritmo do tamanho da empresa ($y_{it}= ln\ Y_{it}$) pode ser descrito como

$$y_{it} = y_{it-1} + a_i + v_{it} \qquad (1)$$

[2] para uma plicação de regressão quantílica em equações de crescimento, veja Porto Jr. e Ribeiro (2001).

onde v_{it} é o termo aleatório. O termo a_i pode ser similar entre empresas e pode ter valor zero. Supondo que os termos aleatórios a_i e v_{it} sejam Gaussianos, a distribuição do tamanho das empresas será Log-normal.

Este modelo apresenta algumas características interessantes. Primeiro, note que o modelo, omitindo *i*, é um modelo AR(1) com raíz unitária. Segundo, reescrevendo a equação em termos de diferenças de y_{it}, temos

$$\Delta y_{it} = (y_{it} - y_{it-1}) = a_i + v_{it}, \quad \text{ou}$$

$$\Delta y_{it} = a_i + v_{it} \,. \qquad (2)$$

Ou seja, as taxas de crescimento do tamanho das empresas (aproximadas por $ln(Y_{it}/Y_{it-1})$) seguem um passeio aleatório, independente do tamanho anterior da empresa. Por esta razão a Lei de Gibrat é dita "Lei do Efeito Proporcional" (*Law of proportional effect*).

Este tipo de "lei" para o crescimento do emprego gera uma distribuição de tamanho da forma Log-normal, ou sob certas transformações, da forma e Yule ou Pareto (Ijuri e Simon, 1958, Kwasnicky, 1998 e outros).

Teoricamente o modelo apresenta algumas dificuldades, em particular, a implicação de que a variância da distribuição de log-tamanho cresce sem limite ao longo do tempo. Por outro lado, se a_i for diferente de zero, a média de y_{it}, isto é, a média geométrica de Y_{it} muda ao longo do tempo. Estas características podem ser inferidas das propriedades de séries e visto pela substituição recursiva de $y_{iy-1}=y_{it-2}+v_{it-1}$ em (1)

$$y_{it}= a_i^t + y_{i0} + \Sigma_t\, v_{it} \,.$$

Desta forma o valor esperado do tamanho é dado por $E(y_{it}) = a_i^t + y_{i0}$ e a variância:

$$V(y_{it})= \Sigma_t\, V(v_{it}).$$

Todavia, as evidências para vários países sugerem que a distribuição de tamanho mantém-se constante ao longo do tempo, ou muda muito suavemente, dando pouca credibilidade a esta forma da "lei". Isto exige uma representação alternativa do processo gerador dos tamanho das empresas, que pode ser escrito como

$$y_{it} = \beta y_{it-1} + a_i + v_{it} \qquad (3)$$

onde $|\beta|<1$. Neste caso, a taxa de crescimento das empresas é inversamente proporcional ao seu tamanho, o que pode ser visto reescrevendo a forma da equação (2) como

$$\Delta y_{it} = (\beta - 1)\, y_{it-1} + a_i + v_{it}. \qquad (4)$$

Nesta forma modificada, a distribuição de empresas ainda é Log-normal mas garante-se que a variância não cresça sem limite ao longo do tempo. Em particular, supondo homocedasticidade em v_{it}, a dispersão do log do tamanho da empresa segue

$$V(y_{it})= V(v_{it})/(1 - \beta) = \sigma_v^2/(1 - \beta^2),$$

supondo que o processo tenha início há muito tempo (*t* "grande") e que β seja estável. Em relação ao valor esperado, ele será dado por

$$E(y_{it}) = a_i /(1 - \beta).$$

Aqui vale a pena um comentário. Note que o resultado acima representa a média *geométrica* do tamanho das empresas. Empregando as propriedades da distribuição lognormal, o valor esperado do tamanho da empresa, ou seja o correspondente teórico da média aritimética, é dado por

$$E(Y_{it})=exp(a_i/(1 - \beta) + ½\, \sigma_v^2/(1 - \beta^2)) = exp(a_i + ½\, \sigma_v^2/(1+\beta))\, exp(\beta - 1).$$

Se a variância dos choques aleatórios *v* mudar ao longo do tempo, o tamanho médio das empresas podem mudar, embora o *log* do tamanho esperado não mude.

A literatura para verificar a validade empírica de (1) em relação a (3) é extensa. Em vários casos, o modelo (1) é rejeitado em favor de (3), como em Evans(1987) e

Hall(1987). Estes dois trabalhos mostram que, para a indústria dos EUA, as pequenas empresas crescem mais rápido que as pequenas. Também foi claramente rejeitada a hipótese de homocedasticidade dos erros (a variância dos erros cai com o tamanho das empresas). Os autores também concluem que a taxa de variação do emprego depende da idade das empresas de modo inverso, *ceteris paribus*. Por fim, os autores tentam controlar o fato de que o grupo de empresas analisada ao longo do tempo não é o mesmo, devido à entrada e à falência de empresas (saída do mercado) e identificam que tal saída é também inversamente proporcional ao tamanho da empresa.

Desta forma a hipótese da Lei de Gibrat deve ser qualificada, de acordo com Mansfield(1962). Ela por ser válida para *(i)* todas as empresas, *(ii)* para as empresas sobreviventes (isto é, apenas para aquelas que Y_{it-1} e Y_{it}>0), ou *(iii)* para empresas "grandes" acima de uma escala eficiente mínima. Alternativamente, Dunne, Roberts e Samuelson (1989) ainda chamam a atenção de que a lei pode ser válida para todas as empresas em termos de taxas de crescimento potenciais (pós-entrada em operação), taxas estas que não são observados devido à censura de empresas por falência.

Vale a pena comentar que o modelo simples apresentado acima não controla por estas duas importantes características da demografia das empresas, ou seja, a entrada e saída de empresas. A entrada pode ser caracterizada por $Y_{it-1}=0$ e $Y_{it}>0$ e a saída por $Y_{it-1}>0$ e $Y_{it}=0$. Pelo emprego do logaritmo neperiano no modelo, $Y_{it} \in \Re^{*}_{+}$, o que impede a análise de entradas e saídas neste modelo. Para acomodar entradas e saídas de empresas, os tamanhos da empresas seriam caracterizadas por

$$ln\ Y_{it} \mid Y_{it}>0\ e\ Y_{it-1}<0 = \beta y_{it-1} + a_i + v_{it}\ . \qquad (3')$$

Das três situações indicadas por Mansfield, a mais promissora e menos problemática[3] para implementação empírica é a de que a lei pode ser válida para empresas "grandes" (Simon e Bonini, 1958). Esta será a forma que será estudada usando os dados da PIA para o Brasil.

3. Modelo Empírico e Métodos de Estimação

O modelo a ser estimado é baseado na equação (3´). A estimação pode ser feita empregando observações de empresas em dois períodos no tempo. Todavia há a questão da endogeneidade do emprego defasado em relação ao termo aleatório individual, a_i. Este termo agrega todos os fatores constantes no tempo que diferenciam as empresas, como por exemplo a tecnologia empregada, a habilidade gerencial dos proprietários, a qualidade das máquinas e outros. Estes fatores tem efeito sobre o tamanho da empresa.

[3] Menos problemática pois não exige o conhecimento da distribuições de taxas de crescimento potencial, que é desconhecido, nem a modelagem da seleção de amostras, que foi tentada por Hall(1987) e Evans(1987) de modo insatisfatório. Em particular, para uma correção da forma de Heckman, como a empregada pelos autores citados, é necessário identificar características que afetem a possibilidade de sobrevivência de uma empresa mas não seu tamanho. Como o processo de falência pode ser entendido como um "bad draw" da distribuição de v_{it}, não parece possível encontrar tal variável. Os autores citados exploraram não-linearidades do tamanho da empresa na seleção. Todavia Johnston e DiNardo (1995) argumenta fortemente que esta estratégia de identificação é fraca.

Assim, a estimação de um modelo de regressão que não considere a correlação deste termo com o emprego defasado, pode gerar estimativas viesadas.

De acordo com Angrist e Krueger (2000), existem várias formas de garantir que as estimativas de β sejam consistentes. Uma delas é evitar o viés por endogeneidade seria a inclusão de variáveis que aproximem a_i. Desta forma, faz-se necessário inserir em (3) características das empresas (como idade) e da indústria em que a empresa se insere para garantir que o termo aleatório tenha propriedades necessárias para que as propriedades dos estimadores sejam as desejadas para inferência. Desta forma o modelo estimado seria

$$y_{it} = \beta y_{it-1} + \mathbf{x}_{it}'\delta + v_{it} \tag{5}$$

onde $\mathbf{x}_{it}$ é um vetor k x1 de controles e δ um vetor k x 1 de coeficientes, incluindo uma constante.

Infelizmente não há garantias que estas "proxies" de a_i, representadas em $\mathbf{x}_{it}'$ tenham sucesso em fazer com que o termo aleatório da regressão tenha sido completamente "limpo" da correlação com o tamanho defasado. Este é um problema análogo à estimativa de modelos dinâmicos em dados de painel. Um modo alternativo de estimar a equação (3´) eliminando diretamente a endogeneidade, seria eliminando o efeito do termo aleatório através do uso de primeiras diferenças no tempo,

$$\Delta y_{it} = (y_{it} - y_{it-1}) = \beta\, \Delta y_{it-1} + (a_i - a_i) + v_{it} - v_{it-1}$$

ou seja,

$$\Delta y_{it} = \beta\, \Delta y_{it-1} + e_{it}, \tag{4}$$

onde, $e_{it} = v_{it} - v_{it-1}$.

O problema agora é a correlação entre Δy_{it-1} e o termo aleatório, e_{it}, pela presença de y_{it-1} e v_{it-1} em cada respectivamente. Pode-se demonstrar que a inconsistência dos parâmetros na estimação por métodos como mínimos quadrados mantém-se (Arellano e Bover, 1990, Johnston e DiNardo, 1994).

Em grandes amostras, pode-se demonstrar que a incosistência do estimador de mínimos quadrados b de β, é tal que uma regressão de (4) irá subestimar o real valor do parâmetro,

$$plim\, b - \beta = -(1 + \beta) / 2 < 0.$$

Ao contrário, a estimação de (3) através de MQO irá gerar estimativas que sobrestimam em grandes amostras o coeficiente β. O grau da sobrestimação depende da razão entre as variâncias de a_i e o termo aleatório v_{it}.

Para obter um estimador consistente para (4), Anderson and Hsiao (1984) notaram que o erro é uma média móvel de primeira ordem, ele não é correlacionado com termos defasados em dois ou mais períodos. Desta forma, pode-se considerar a estimação da equação em primeiras diferenças empregando o método de variáveis instrumentais, considerando como instrumentos todas as defasagens de y_{it} acima de dois períodos, ou seja, y_{it-2}, y_{it-3}, etc. Note que a menor amostra necessária para a estimação do modelo é um painel com três observações temporais.

O método de variáveis instrumentais consiste em estimar os coeficientes do modelo de regressão de modo indireto, empregando variáveis correlacionadas com as explicativas, mas independentes ao termo aleatório. No caso de estimação por mínimos quadrados ordinários, para o modelo acima, b, o estimador de variáveis instrumentais do parâmetro β é dado por

$$b_{IV} = \Sigma_i\, (\Delta y_{it}\, y_{it-2}) / \Sigma_i\, (\Delta y_{it-1}\, y_{it-2})$$

no caso de apenas um instrumento (y_{it-2}), e no caso de dois (ou mais) instrumentos,

$$b_{IV} = \Sigma_n\, (\Delta y_{it}\, P_Z\, \Delta y_{it-1}) / \Sigma_n (\Delta y_{it-1}\, P_Z\, \Delta\, y_{it-1}),$$

onde $P_Z = Z'(Z'Z)^{-1}Z$ e Z é uma matriz com linhas $(1\ y_{it-2}\ \ y_{it-3})$, $i=1, ..., n$. No segundo caso, o método de variáveis instrumentais é conhecido como mínimo quadrado em dois estágios (MQ2E), pois a matriz P_Z cria uma projeção de $\Delta\, y_{it-1}$ no espaço de Z, que é ortogonal a e_{it}. b_{IV} pode ser implementado em dois estágios: uma previsão de $\Delta\, y_{it-1}$ em função de y_{it-2} (e y_{it-3}); e após uma regressão de $\Delta\, y_{it}$ na previsão de $\Delta\, y_{it-1}$. As estimativas de b_{IV} são consistentes[4].

O cálculo da estimativa dos resíduos e da variância dos erros deve ser ajustada, no sentido de que sejam baseados em $e_{it} = \Delta y_{it} - b_{IV}\, \Delta\, y_{it-1}$ ao invés do uso da previsão de $\Delta\, y_{it-1}$. Por outro lado, medidas de qualidade de ajuste como o R^2 devem ser calculados como $R^2 = 1 - \Sigma_n\, \ddot{e}_{it}^{\,2} / \Sigma_n (\Delta y_{it} - (\Sigma_n \Delta y_{it}/n))^2$, onde $\ddot{e}_{it} = \Delta y_{it} - b_{IV}\, (P_2\, \Delta\, y_{it-1})$, ou seja, empregando como variável explicativa a previsão do primeiro estágio (Pesaran e Smith, 1994). Estes detalhes não são considerados na literatura.

O método de variáveis instrumentais apresentado baseia-se em estimadores para a média condicional da variável dependente. Todavia, da mesma forma que a média não representa completamente uma distribuição univariada, pode ser interessante estimar quantis condicionais. Isto pode ser levado a cabo através da técnica de regressão quantílica (Koenker e Bassett, 1978).

Este estimador permite a estimação dos percentis condicionais e não condicional da distribuição da variável dependente (em nosso caso, o log do tamanho da empresa), trazendo consigo mais informações que aquelas obtidas através de estimadores de médias condicionais. Em particular, pode ser possível verificar se a distribuição dos tamanhos é Log-normal ou não, na distribuição condicional. Ou seja, embora a distribuição do tamanho da empresa possa não ser Log-normal de modo não condicional, a distribuição condicional a certas características, talvez o seja.

Isto pode ser verificado construindo distribuições condicionais por célula de idade e outras características. Este procedimento pode ser muito trabalhoso por exigir muitos dados ou por ficar complexo rapidamente, ao serem incluídas outras características contínuas das empresas (Buchinsky, 1994). O método de regressão quantílica permite estudar as distribuições condicionais impondo poucas hipóteses sobre o comportamento dos percentis condicionais[5]

Para entender o estimador de regressão quantílica[6], considere y_i a variável de interesse ($i=1,...,n$). O modelo supõe que o θ'ésimo percentil da distribuição condicional de y, dado um vetor de variáveis explicativas (x_i), é linear em x_i, isto é,

$$Q_\theta\,(y|x) \equiv inf\{y|\ F_i(y|x) \geq \theta\,\} = x_i'\beta_\theta,$$

onde $F_i(\ )$ é função distribuição condicional, $Q_\theta(y|x)=F^{-1}(\theta)$ a respectiva função de percentis (*quantile function*) e β_θ é um vetor de coeficientes desconhecido, cuja estimação, para diferentes valores de θ no intervalo (0,1) é o objetivo do problema. Por exemplo, para x_i igual a zero, temos $u_i = y_i - \beta_0$, e se F_u é simétrico em zero, $Q_{0.5}\,(y)=0$, ou seja, a mediana de y (agora a distribuição não condicional) será igual a β_0 .

O vetor de coeficientes pode ser estimado através da minimização de

$$\Sigma_i\ \rho_\theta(y_i - x_i'\beta_\theta\,), \qquad \text{ou} \qquad \Sigma_i\ \rho_\theta(u_i\,),$$

[4] Kiviet(1995) indica que este estimador de Variáveis Instrumentais de Anderson e Hsiao pode ter melhores propriedades em pequenas amostras do que o estimador teoricamente mais eficiente de GMM de Arellano e Bond.

[5] Supões-se uma relação linear dos percentis para diferentes valores da variável explicativa. O trabalho com estimativas baseadas em médias condicionais (e respectivas variâncias) exige hipóteses mais fortes sobre a distribuição condicional, por exigir, por exemplo, que os percentis da distribuição condicional sejam parametrizados pela média e variância apenas. No extremo oposto há o trabalho com células, que seria não paramétrico.

[6] Maiores detalhes podem ser vistos em Koenker, 2000.

com

$$\rho_\theta = [\theta\, u_i\, I(u_i \geq 0) + (\theta - 1)\, u_i\, I(u_i < 0)\,],$$

onde *I( )* é a função indicador.

Note que os erros u_i são tomados em seu valor absoluto e ponderados de modo desigual, exceto para $\theta=1/2$, quando a mediana condicional está sendo estimada. O problema de minimização acima não possui forma fechada, mas pode ser estimado por métodos de programação linear (Koenker e Bassett, 1978). Sob hipóteses muito fracas sobre o termo aleatório, pode-se demonstrar que o vetor de coeficientes é consistente e assintoticamente normal. Assim como a mediana, em muitos casos, o estimador é mais eficiente que mínimos quadrados. Em termos de máxima verossimilhança, para $\theta=1/2$, o estimador pode ser entendido como o estimador de MV para a distribuição de Laplace (*double exponential*).

Algumas propriedades importantes do estimador podem ser comentadas. Primeiro, ele é robusto a outliers em y_i. Segundo, os coeficientes são equivariantes a transformações monotônicas (como log) ao contrário da média. Terceiro e mais importante, talvez, no caso de heterogeneidade dos dados (inclusive heterocedasticidade) os percentis condicionais irão variar ao longo da distribuição e assim o efeito das variáveis explicativas será diferente em diferentes pontos da distribuição. Em detalhe, supondo um modelo com erros *iid*, ou seja a distribuição condicional não varia em suas características para diferentes valores da variável explicativa, exceto pelo centro dela (média condicional) que segue uma relação linear com as explicativas. Da mesma forma, os coeficientes de regressão quantílica serão constantes para diferentes valores de *x* exceto pela constante, cujo coeficiente depende do percentil do erro,

$$Q_\theta(y|x) = \{\beta_0 + F^{-1}(u)\} + x'\beta,$$

onde *x* não inclui agora uma coluna de 1's. Todavia, se os percentis da distribuição condicional mudarem sistematicamente com as variáveis explicativas, os coeficientes angulares serão diferentes para diferentes quantis, ou seja,

$$Q_\theta(y|x) = \{\beta_0 + F^{-1}(u)\} + x'\beta_\theta.$$

Por exemplo, Hendricks e Koenker (1992) estimaram que embora haja um efeito-preço médio ao longo do dia para os consumidores de residenciais de energia elétrica em Chicago, tal efeito é estatisticamente nulo para pequenos consumidores e forte e significativo para grandes consumidores. O efeito médio esconde significativa heterogeneidade na sensibilidade ao preço.

Em recente aplicação da técnica de regressão quantílica para tratar a distribuição de tamanhos de empresas, temos Mata e Machado (2000). Os autores, através do método de regressão quantílica, especificam o efeito dos atributos dos setores nas medidas de localização, escala, curtose e assimetria da distribuição dos tamanhos das empresas é significativo, ou seja, os coeficientes de regressão quantílica são diferentes.

Para este estudo, faz-se necessário o uso de métodos de variáveis instrumentais. O uso de IV para a estimação de parâmetros dos quantis condicionais foi estudado por Powell (1983) e Chen e Portnoy(1996). Uma aplicação pode ser vista em Ribeiro (1998). De modo similar a MQ2E, no primeiro estágio a variável explicativa endógena é prevista, através de mínimos quadrados (ou regressão quantílica), em função dos instrumentos; no segundo estágio, a previsão é utilizada como variável explicativa em um modelo de regressão estimado por regressão quantílica. Este estimador em dois estágios é denominado na literatura regressão quantílica em dois estágios (RQ2E). Da mesma forma que o estimador de MQ2E, o cálculo da matriz de variância-covariância exige cuidados. Maiores detalhes podem ser vistos em Ribeiro (1998).

A interpretação dos coeficientes estimados por regressão quantílica para os modelos (3), com a variável dependente em nível, e (4), com a variável em primeiras diferenças deve ser feita com cuidado. Embora para a média condicional de (3) o coeficiente autoregressivo será o mesmo do modelo (4), para as estimativas de regressão quantílica os coeficientes podem diferir entre (3) e (4). Como a variável dependente e o erro tem uma distribuição que é uma função de outras variáveis aleatórias ($z_{it} = \Delta y_{it} = y_{it} - y_{it-1}$), pelo *convolution theorem*, não é possível obter os percentis de y_{it} a partir dos percentis de z_{it}. Por outro lado, note que a hipótese básica em estudo neste trabalho, isto é, a validade da Lei de Gibrat, pode ser testada em (4), com a hipótese de que o coeficiente autoregressivo é igual a zero. Não é válido que, no caso do coeficiente em (4) – que tem Δy_{it} como variável dependente – ser igual a zero para o percentil de 10%, afirmar que em (3) o coeficiente do percentil de 10% de y_{it} será igual a 1, respectivamente.

4. Dados empregados

O estudo baseia-se em dados de empresas da Pesquisa Industrial Anual – Empresa (PIA) do IBGE para os anos de 1996-1999. O corte temporal busca evitar as mudanças amostrais da PIA a partir de 1996. Para reduzir a heterogeneidade setorial e em linha com outros estudos, para garantir comparabilidade com a literatura, a medida de tamanho empregada é o pessoal ocupado em 31/12 de cada ano[7]. Os setores de extração e mineração foram excluídos por serem considerados muito diferentes de setores de manufatura em si e o setor de CNAE37 foi excluído por questões de confidencialidade.

O padrão amostral da PIA conjuntamente com o referencial teórico acima, indicou o uso do estrato certo da pesquisa e empresas que sobreviveram ao longo do período. Apenas empresas médias e grandes foram empresadas, sendo considerado o corte de 100 pessoas empregadas. Questões de confidencialidade exigiram a eliminação de empresas com emprego médio no período maior que 5000 pessoas ocupadas. Os ajustes na amostra e o próprio referencial teórico deixam claro que os resultados apresentados não serão representativos de toda a população de empresas industriais no Brasil. Tal cobertura exigiria o tratamento da questão da entrada e saída de empresas, que não é trivial (Bivar e Rodrigues, 2000).

As estatísticas descritivas dos dados estão na Tabela 1. O emprego médio nas empresas varia de 418.5 empregados em 1996 até 379.9 em 1999 A distribuição é claramente assimétrica a direita com a mediana variando de 218 empregados em 1996 até 206 empregados em 1999. A assimetria da distribuição do Log do tamanho pode ser identificada no coeficiente de assimetria apresentado e é visível no Gráfico 1, em que são destacas as distribuição de 1996 e 1999, o primeiro e último ano dos dados. No Gráfico 2 temos as distribuições de freqüência acumuladas para os quatro anos, em termos de log-emprego. Identifica-se um deslocamento para baixo das distribuições de tamanho de 1999/1998 em relação a 1996/1997, indicando uma retração em todos os níveis do emprego industrial e não apenas de partes das distribuições. Este resultado pode ser visto em mais detalhe no Gráfico 3. Entre 1996 e 1999, excetuando a grande diferença nos menores quantis, o emprego em 1999 era de 5 a 13% menor do que em 1996, dentro de cada quantil.

[7] O estudo foi replicado usando como medida de tamanho o emprego médio do ano, sem diferenças importantes nos resultados. Os resultados estão disponíveis com o autor.

Passamos agora aos resultados da análise do modelo de regressão.

5. Resultados Empíricos

Os resultados da análise de regressão usando um modelo simples (3), sem o uso de variáveis instrumentais, estão na Tabela 2 e Gráfico 4. A Tabela 2 apresenta os resultados da equação (3) empregando mínimos quadrados. Teoricamente, esperamos que os coeficientes sejam inconsistentes, sendo sobre-estimados. Eles são apresentados para comparação com a literatura. Nos três anos estudados, o R^2 é alto e o coeficiente autoregressivo é estatisticamente menor de 1, embora se aproxime do mesmo ao longo do tempo. Testes de especificação indicam a presença de heterocedasticidade exceto para 1997. A possibilidade de erros homocedásticos não é comum na literatura. Interpretando o coeficiente em relação a equação (4), temos que, em média, menores empresas crescem mais rápido, retirando o valor da Lei de Gibrat para a indústria no Brasil no período.

Todavia o efeito médio esconde uma heterogeneidade já detectada pelo teste de heterocedasticidade. Na Tabela 3 temos as estimativas de regressão quantílica para alguns percentis mais comuns e, no Gráfico 4, os coeficientes autoregressivos para os percentis no intervalo de 0.05 a 0.95, além de intervalos de confiança de 90%. Como mencionado acima, não esperariamos diferenças significativas para o coeficiente autoregressivo para diferentes percentis se não houvesse heterocedasticidade ou uma relação entre o tamanho da empresa e o termo aleatório, como indica Sosa *et.al.* (2001). Mas na Tabela 4 vemos que para 1999 o efeito do tamanho no crescimento (calculado como o coeficiente menos 1) é menor para as menores empresas na amostra e mais forte para as grandes empresas, pois o coeficiente autoregressivo cai a medida que os percentis aumentam. O Gráfico 4 deixa isto explicito. Até o percentil de 0.25 não podemos rejeitar a hipótese de que o tamanho da empresa não afeta seu crescimento, pois o coeficiente autoregressivo não é estatísticamente distinto de 1. Para percentis maiores do que 0.25, o coeficiente autoregressivo cai para menos que a unidade, tendo intervalos bem estreitos, sugerindo mudanças significativas no coeficiente, e chegando a menos de 0.89 para o percentil de 95%.

Por outro lado, para os anos de 1997 e 1998, o coeficiente autorregressivo parece ter um comportamento de U invertido sendo maior nos percentis medianos e menor nos extremos da distribuição condicional. Mesmo assim, em seus valores mais altos, próximos do percetil de 0.5, os intervalos de confiança não incluem a unidade, tirando credibilidade na hipótese de que o crescimento da empresa independe de seu tamanho.

O uso de regressão quantílica nos permite estudar a dinâmica do emprego de modo ímpar. Ao estimarmos os coeficientes para cada percentil, podemos estimar a distribuição condicional para cada valor da variável explicativa condicionante de modo semi-paramétrico. Para 1999, por exemplo, isto é apresentado no Gráfico 5a. Para facilitar a interpretação uma reta de 45^o é adicionada ao gráfico. O que as retas de regressão quantílica mediana e dos percentis de 0.10 e 0.90 nos dizem é que a distribuição condicional desloca-se para baixo da reta de 45^o a medida que o tamanho da empresa aumenta. No Gráfico 5b vemos o efeito de modo mais claro. No gráfico temos, para diferentes tamanhos de empresas no período anterior, a diferença do tamanho previsto nas retas de regressão quantílica. Por exemplo, devido a reta de regressão mediana passar por zero quando o tamanho é igual a 4.5, isto quer dizer que metade das empresas deste porte tem variações positivas em seu tamanho por estarem acima da reta.

Da mesma forma, para empresas de tamanho 4.5, apenas 10% das empresas aproximadamente, cresce mais que 0.30%[8]

Mudando o foco para o modelo (4), ou seja o modelo para a taxa de crescimento das empresas (primeiras diferenças) na Tabela 4 temos as estimativas de MQO, e na Tabela 5 temos os resultados do modelo estimado de modo consistente, ou seja, através de Variáveis Instrumentais. A Tabela 4 é apresentada para entendermos melhor o efeito das primeiras diferenças. Como esperado, os coeficientes são negativos e implicam valores teóricos para o coeficiente autoregressivo entre 0.80 e 0.85, de acordo com o viés assintótico explicitado acima. Estes valores teóricos indicam que os coeficientes estimados na Tabela 2 estão sobre-estimados, como esperado. De qualquer forma, mais uma vez transparece que o crescimento das empresas é negativamente relacionado com seu tamanho.

Passando para as estimativas obtidas por variáveis instrumentais (ou mínimos quadrados em dois estágios –MQ2E), na Tabela 5 observamos primeiro a qualidade dos instrumentos. Apesar dos R^2s serem baixos, os teste F das regressões com cada explicativa endógena em relação aos instrumentos sugerem que os instrumentos possam ser válidos (Bound *et al.*, 1995). As estimativas em si dos coeficientes autorregressivos são significativos e positivos, sendo bem menores que as estimativas da Tabela 2 e de sinal contrário ao da Tabela 4, como esperado. Mais uma vez temos rejeitada a hipótese da Lei de Gibrat e, interpretando o coeficiente como no modelo (2) temos que empresas menores crescem mais rápido.

Passando para as estimativas empregando regressão quantílica com variáveis instrumentais, os resultados da Tabela 6 mostram uma grande variabilidade dos coeficientes ao longo dos percentis. Isto pode ser devido ao efeito dos fracos instrumentos. Para 1999-1998, o coeficiente autoregressivo no percentil de 10% não é significativo. Quanto maior o quartil, maior o coeficiente, chegando a 0.75 para o percentil de 90%. O resultado muda para o ano de 1998, em que o coeficiente autoregressivo possui um formato de U para diferentes percentis, embora os intervalos de confiança sejam largos nos extremos.

Visualizando os coeficientes nos Gráficos 6a-6c para os diferentes percentis, confirmamos os resultados da Tabela 6. Para entender melhor a distribuição condicional ao longo dos quartis, temos no Gráfico 7 as retas de regressão quantílica para os percentis de 10, 50 e 90% das distribuições condicionais. Vemos que a medida que muda a taxa de variação do tamanho da empresa, mudam o comportamento das caudas é diferenciado do centro da distribuição, medido pela mediana. Empresas com reduções fortes em seu tamanho tem uma desaceleração de sua queda, pois quase toda a distribuição está acima da reta de 45^o no quadrante negativo-negativo. Por outro lado, empresas que estão expandindo fortemente (grande valor positivo no eixo horizontal) não estão com sua distribuição de taxas de crescimento em sua quase totalidade abaixo da reta de 45^o, em simetria ao verificado no quadrante negativo.

Há uma maior heterogeneidade das empresas neste quadrante. Vemos que a reta do percentil condicional de 90% não se aproxima da reta de 45^o, ao mesmo tempo que a reta do percentil condicional de 10% não se aproxima do eixo horizontal como assim o fez o percentil de de 90% no lado esquerdo do gráfico. Por exemplo, para as empresas com redução de 40% em seu emprego, a probabilidade de uma variação positiva era de aproximadamente 0.10, pois o percentil condicional de 90% neste ponto é quase de 0. Já para empresas com expansão de 40% em seu emprego, a probabilidade de uma variação negativa era de mais de 0.10. Na verdade, observando a reta do percentil condicional de

[8] A interpretação de variação percentual explora fato de que $(Y_{it} - Y_{it})/ Y_{it} \approx ln(Y_{it}) - ln(Y_{it-1}) = y_{it} - y_{it-1}$.

10%, o modelo sugere uma probabilidade de 0.10 que uma empresa com expansão de 40% tenha uma redução de seu tamanho maior de 12% (em valor absoluto). Esta heterogeneidade de dinâmica diferenciada entre empresas que crescem e contraem não era identificada em modelos de média condicional, como os da Tabela 5.

Em suma, os resultados das estimativas com variáveis instrumentais, seja para média condicional, seja para regressão quantílica, sugerem que o coeficiente autorregressivo do modelo generalizado da Lei de Gibrat (eq.4) é menor que a unidade e que empresas menores crescem mais rápido (embora empresas com grande crescimento tendam maior probabilidade de crescer mais lentamente no futuro).

6. Considerações Finais

O objetivo deste artigo foi o de estudar a distribuição do tamanho das empresas industriais no Brasil e sua dinâmica, verificando empiricamente a Lei de Gibrat para a segunda metade da década de 1990. Vimos que a distribuição de tamanho das empresas (medido pelo número de empregados no fim do ano) é assimétrica e mantém sua assimetria mesmo após a transformação logaritmica, dando pouca credibilidade à suposição de que a distribuição seria Log-Normal. Em termos de dinâmica, no período de 1996 a 1999, vemos que o emprego industrial caiu, tendo quase toda a distribuição de tamanho se desolcado para a esquerda.

A verificação da hipótese da Lei de Gibrat não pode se feita empregando modelos de regressão que supõe ortogonalidade entre o erro do modelo de regressão e o tamanho defasado devido a presença de componentes individuais não observados persistentes. Ao invés da busca de *proxies* para identificar este componente, explorando a característica de painel da base de dados (dados longitudinais), estimadores de variáveis instrumentais são empregados. Para uma análise mais completa do modelo, além do método de MQ2E, estimadores de regressão quantílica em dois estágios (RQ2E) foram empregados.

A análise de regressão preliminar baseado em estimadores inconsistentes indica que estimativas na literatura que empregam métodos de regressão como MQO ou similar são inconsistentes e sobre-estimam o coeficiente autoregressivo. Mesmo assim, nestes casos a Lei de Gibrat foi rejeitada, exceto para as menores empresas. As estimativas de regressão quantílica sugerem uma assimetria negativa e uma redução da dispersão da distribuição condicional a medida que o tamanho da empresa aumenta. A redução da dispersão dos dados com o aumento do tamanho já havia sido identificada como um fato estilizado na literatura, mas a assimetria negativa parece ter sido pela primeira vez identificada.

As estimativas de variáveis instrumentais sugerem que o coeficiente autorregressivo do modelo especificado está entre 0.3 e 0.6, seja para estimativas de média condicional ou de regressão quantílica, exceto para percentis maiores, quando o coeficiente chega a 0.75. Em suma, a Lei de Gibrat para o Brasil parece claramente rejeitada para as empresas industriais médias e grandes que estavam em operação durante todo o período em estudo.

Agradecimentos

Agradeço à Wasmalia Bivar, Naércio Menezes e Jorge Arbache por sugestões e ao apoio da DPE/DEIND pelo acesso aos dados e instalações, em particular Alexandre Brandão, e à Fundação Ford e CNPq pelo apoio financeiro.

Referências

Angrist, J. e Krueger, A. (1999). Empirical strategies in labor economics. *in Handbook of Labor Economics, v.3, O. Ashenfelter e D. Card (eds),* Amsterdan:Elsevier Science.

Arellano, M. e Bover, O. (1990). La econometria de los datos de panel. *Investigaciones Económicas* (2ª. época, Madrid), 14, 3-45.

Barro, R. e Sala-i-Martin, X. (1995) *Economic Growth*. New York:McGraw-Hill.

Bivar W. e Rodrigues, A. (2001) Aspectos demográficos das empresas industriais brasileiras. *mimeo,* IBGE/DEIND.

Bond, S., Nauges, C e Windmeijer, F (2002). Unit Roots and Identification in Autoregressive Panel Data Models: A Comparison of Alternative Tests *10th International Conference on Panel Data, Berlin, July 5-6, 2002*

Evans, D. (1987). Tests of alternative theories of firm growth. *Journal of Political Economy*, 95, 657-674.

Hall, B. (1987) The relationship between firm size and firm growth in the US manufacturing sector. *Journal of Industrial Economics*, 35, 583-606

Hall, B and Mairesse, J. (2002) Testing for unit roots in panel data: an exploration using real and simulated data. *mimeo*. <http://emlab.berkeley.edu/users/bhhall/>

Hendricks, W. e Koenker, R. (1992). Hierarchical spline models for conditional quantilees and the demand for electricity. *Journal of the American Statistical Association*, 87, 58-68.

Johnston, J. e DiNardo, J. (1995). *Econometric Methods, 4th. Ed.* New York:McGraw-Hill.

Kiviet, J. (1995). On bias, inconsistency and efficiency of various estimators in dynamic panel data models. *Journal of Econometrics*, 68, 53-78.

Koenker, R. (2000) *Quantile Regression* -book draft, University of Illinois at Urbana-Champaign, (www.econ.uiuc.edu).

Kwasnicki, W. (1998). Skewed distributions of firm sizes – an evolutionary approach. *Structural Change and Economic Dynamics*, 9, 135-158.

Machado, J. e Mata, J. (2000) Box-Cox quantile regression and the distribution of firm sizes. *Journal of Applied Econometrics*, 15, 253-274.

Mansfield, E. Entry, Gibrat´s Law and the growth of firms. *American Economic Review* 52, 1023-1051.

Pazello, E, Bivar, W. e Gonzaga, G. (2001). Criação e destruição de postos de trabalho por tamanho da empresa no Brasil.*Pesquisa e Planejamento Econômico, a sair*.

Pesaran, M.H e R.J. Smith (1994) A Generalized R^2 Criterion for Regression Models Estimated by the Instrumental Variables Method, *Econometrica*, 1994, Vol.62 No.3, 705-710.

Porto Jr. S. e Ribeiro, E.P. (2001) Dinâmica de Crescimento Regional – uma análise empírica para a região Sul. *Revista Econômica do Nordeste*, número especial.

Quah, D. (1996) Twin peaks: growth and convergence in models of distribution dynamics. *Economic Journal*.

Ribeiro, E. P. (1998). Small Sample Evidence of Quantile Regression Estimates for Structural Models: Estimation and testing. *Revista de Econometria* v.18(2), p.215-244, novembro.

Ribeiro, E. P.(2001) Asymmetric Labor Supply. *Empirical Economics*, 26:183-197.

Simon, H. e Bonini, C. The size distribution of business firms. *American Economic Review*, 48, 607-617.

Sutton, J. (1997). Gibrat´s Legacy. *Journal of Economic Literature*, 35, 40-59.

Tabela 1a - Estatísticas Descritivas do Tamanho das Empresas (em número de pessoas ocupadas em 31/12)

Variável	1o Quart.	Mediana	Média	3o. Quart.	Desv.Pad.	Assim.	Curt.
1999 (N99)	133.0	206.0	379.9	376.0	90.0	0.3992	2.8683
1998 (N98)	134.0	205.0	381.5	380.0	90.0	0.4228	2.7602
1997 (N97)	141.0	216.0	406.8	409.0	90.0	0.4403	2.7537
1996 (N96)	141.0	218.0	418.5	415.0	90.0	0.4380	2.7847

Tabela 1b - Estatísticas Descritivas do Log do Tamanho das Empresas (em no. de pessoas ocupadas em 31/12)

Variável	1o Quart.	Mediana	Média	3o. Quart.	Desv.Pad.	Assim.	Curt.
1999 (ln99)	4.890	5.328	5.481	5.930	0.873	0.1585	1.9880
1998 (ln98)	4.898	5.323	5.503	5.940	0.829	0.1842	1.9217
1997 (ln97)	4.949	5.375	5.567	6.014	0.821	0.1995	1.9092
1996 (ln96)	4.949	5.384	5.571	6.029	0.850	0.1936	1.9244

Fonte: calculos do autor baseados em dados da PIA/IBGE. N=5745

Medidas especiais empregadas: Assimetria=(Q.75+Q.25-2*Q.50)/(Q.75-Q.25), Curtose=(Q.90-Q.10)/(Q.75-Q.25), onde Q.z é tal que F(x)=z. Para a Distribuição Normal, Assim.=0 e Curt=1.9.

Tabela 2 - Modelo autoregressivo do tamanho das empresas - MQO

Variável	ln99	ln98	ln97
Intercept	0,1998*	0,3202*	0,5576*
	(0,048)	(0,029)	(0,037)
ln98	0,9597*		
	(0,009)		
ln97		0,9311*	
		(0,005)	
ln96			0,8991*
			(0,006)
Estatística-F-p-value	0	0	0
R-quadrado	0.8300	0.8515	0.8662
Test Het (p-value)	0.0005	0.0133	0.1717

Fonte: cálculos do autor empregando uma amostra da PIA/IBGE 1996-1999.

Notas: * - significativo a 5%. Teste Het (p-value) - valor *p* do teste de Heterocedasticidade de Koenker (1981). Desvios padrões robustos a heterocedasticidade entre parênteses.

5745 observações

Tabela 3 - Modelo autoregressivo do tamanho das empresas - Regressão quantílica

Painel a - 1999					
Variável	tau = 0,1	tau = 0,25	tau = ,05	tau = ,075	tau = 0,9
Intercept	-0.2681	-0.0658	0.0710	0.3096	0.6528
	(0,064)	(0,022)	(0,012)	(0,020)	(0,038)
ln98	0.9980	0.9921	0.9862	0.9623	0.9273
	(0,012)	(0,004)	(0,002)	(0,003)	(0,006)
R1	0.4772	0.6207	0.7089	0.7322	0.7130

Painel b - 1998					
Variável	tau = 0,1	tau = 0,25	tau = ,05	tau = ,075	tau = 0,9
Intercept	0.0634	0.0635	0.1419	0.2963	0.6127
	(0,071)	(0,030)	(0,017)	(0,017)	(0,027)
ln97	0.9253	0.9603	0.9694	0.9586	0.9254
	(0,013)	(0,005)	(0,003)	(0,003)	(0,004)
R1	0.4145	0.5899	0.7053	0.7468	0.7398

Painel c - 1997					
Variável	tau = 0,1	tau = 0,25	tau = ,05	tau = ,075	tau = 0,9
Intercept	0.3547	0.2597	0.2252	0.4147	0.7940
	(0,039)	(0,027)	(0,015)	(0,019)	(0,036)
ln96	0.8864	0.9322	0.9585	0.9445	0.9035
	(0,008)	(0,005)	(0,003)	(0,003)	(0,006)
R1	0.4893	0.6296	0.7124	0.7314	0.7063

Fonte: cálculos do autor empregando uma amostra da PIA/IBGE 1996-1999.
Notas: * - significativo a 5%. R1 - medida de qualidade de ajuste da regressão no quantil de Koenker e Machado (2000). Desvios padrões robustos a heterocedasticidade entre parênteses
5745 observações

Tabela 4 - Modelo autoregressivo do tamanho das empresas - Primeira Diferenças e MQO		
Variável	dn99	dn98
Intercept	-0,0285*	-0,0639*
	(0,005)	(0,004)
dn98	-0,1016*	
	(0,045)	
dn97		-0,0848*
		(0,029)
Estatística-F-p-value	0	0
R-quadrado	0.0083	0.0067
Test Het (p-value)	0	0.0807
Fonte: cálculos do autor empregando uma amostra da PIA/IBGE 1996-1999.		
Notas: * - significativo a 5%. Teste Het (p-value) - valor *p* do teste de Heterocedasticidade		
de Koenker (1981). Desvios padrões robustos a heterocedasticidade entre parênteses.		
5745 observações		

Tabela 5 - Modelo autoregressivo do tamanho das empresas - Primeira Dif. e Var. Instrumentais			
Variável	dn99	dn99	dn98
Intercept	0.0011	-0.0054	-0.0611*
	(0,008)	(0,008)	(0,005)
dn98a	0.3659*		
	(0,115)		
dn98a		0.2624*	
		(0,104)	
dn97a			0.524*
			(0,075)
R-quadrado	0,0033	0.0019	0.0192
Instrumentos	ln97	ln97, ln96	ln96
Teste F 1o. Estág.	180.6*	103.4*	468.6*
R-quadrado 1o. Estág.	0.0305	0.0348	0.0754
Fonte: cálculos do autor empregando uma amostra da PIA/IBGE 1996-1999.			
Notas: * - significativo a 5%. a - variável prevista pelos instrumentos listados			
Desvios padrões robustos a heterocedasticidade entre parênteses. 5745 observações			
5745 observações			

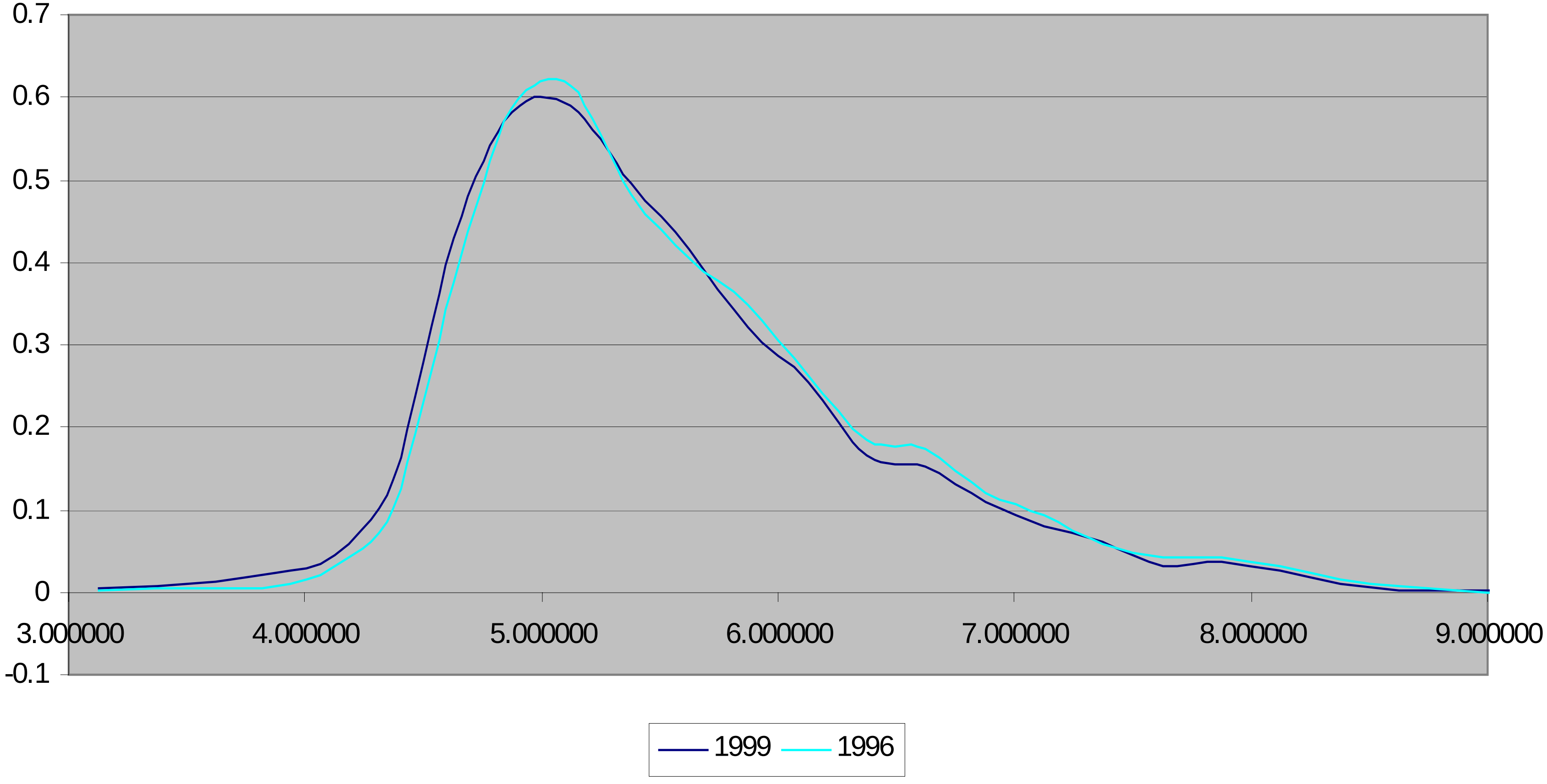
Tabela 6 - Modelo autoregressivo do tamanho das empresas - Primeira Diferenças e Regressão quantílica com Variáveis Instrumentais					
		Painel a - 1999			
Variável	tau = 0,1	tau = 0,25	tau = ,05	tau = ,075	tau = 0,9
Intercept	-0.2620	-0.0970	0.0087	0.1323	0.2955
	(0.0266)	(0.0066)	(0.0034)	(0.0077)	(0.2530)
dln98a	0.3520	0.1931	0.2084	0.4940	0.7571
	(0.3132)	(0.0770)	(0.0403)	(0.0901)	(0.2890)
Instrumentos	ln97				
		Painel b - 1999			
Variável	tau = 0,1	tau = 0,25	tau = ,05	tau = ,075	tau = 0,9
Intercept	-0.2668	-0.1011	0.0049	0.1246	0.2877
	(0.0257)	(0.0063)	(0.0033)	(0.0073)	(0.0238)
dln98a	0.2126	0.1394	0.1564	0.3812	0.6302
	(0.2932)	(0.0713)	(0.0379)	(0.0828)	(0.2713)
Instrumentos	ln97, ln96				
Painel c - 1997					
Variável	tau = 0,1	tau = 0,25	tau = ,05	tau = ,075	tau = 0,9
Intercept	-0.3494	-0.1549	-0.0279	0.0662	0.1980
	(0.0198)	(0.0055)	(0.0026)	(0.0042)	(0.0147)
dn97a	0.5832	0.3726	0.2874	0.3660	0.6636
	(0.2308)	(0.0641)	(0.0302)	(0.0487)	(0.1712)
Instrumentos	ln96				

Fonte: cálculos do autor empregando uma amostra da PIA/IBGE 1996-1999.

Notas: * - significativo a 5%. a - variável prevista pelos instrumentos listados

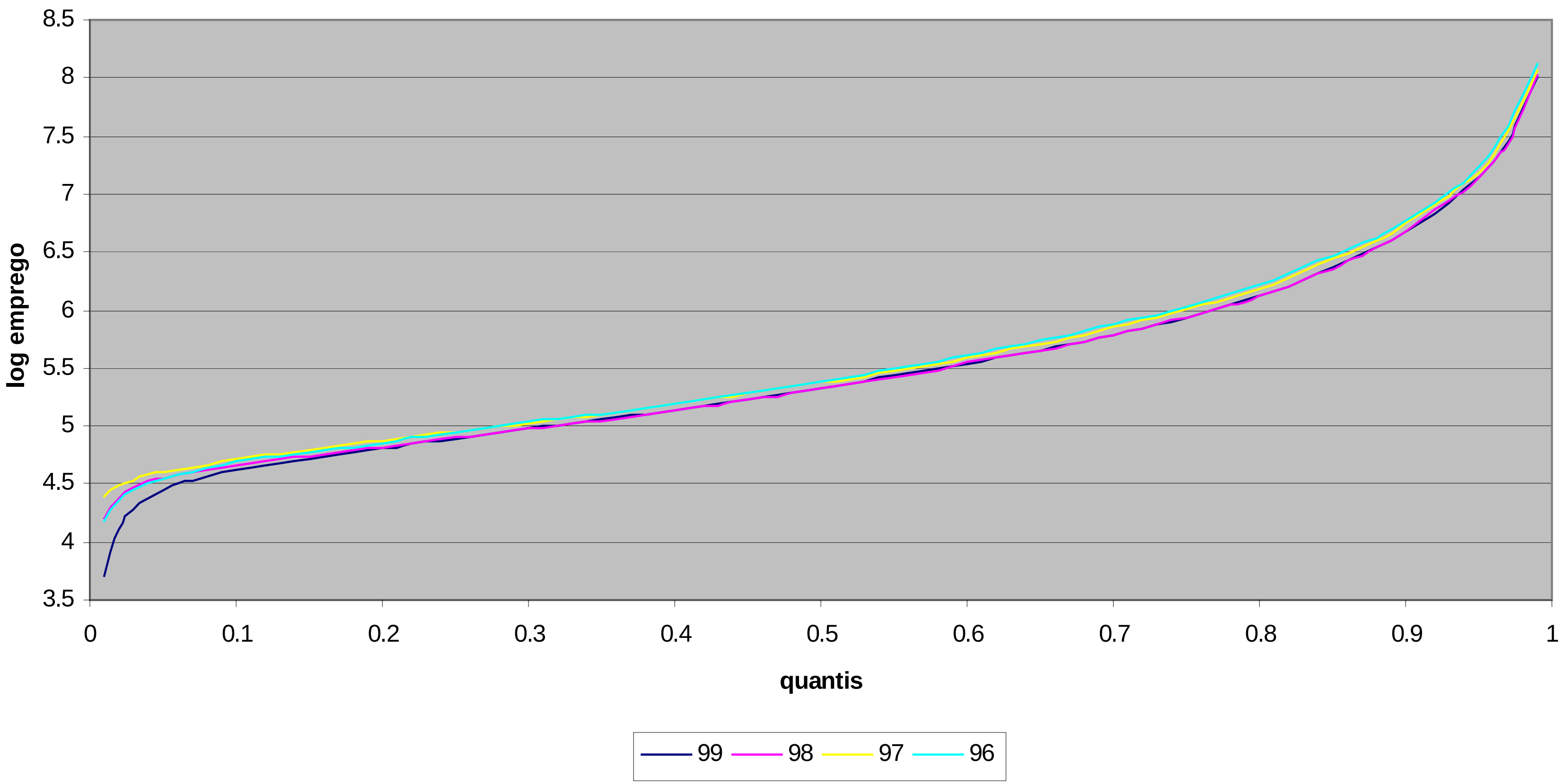
Desvios padrões robustos a heterocedasticidade entre parênteses. 5745 observações

Gráfico 1 - Histograma do Log do Tamanho de Empresas - 1999 e 1996



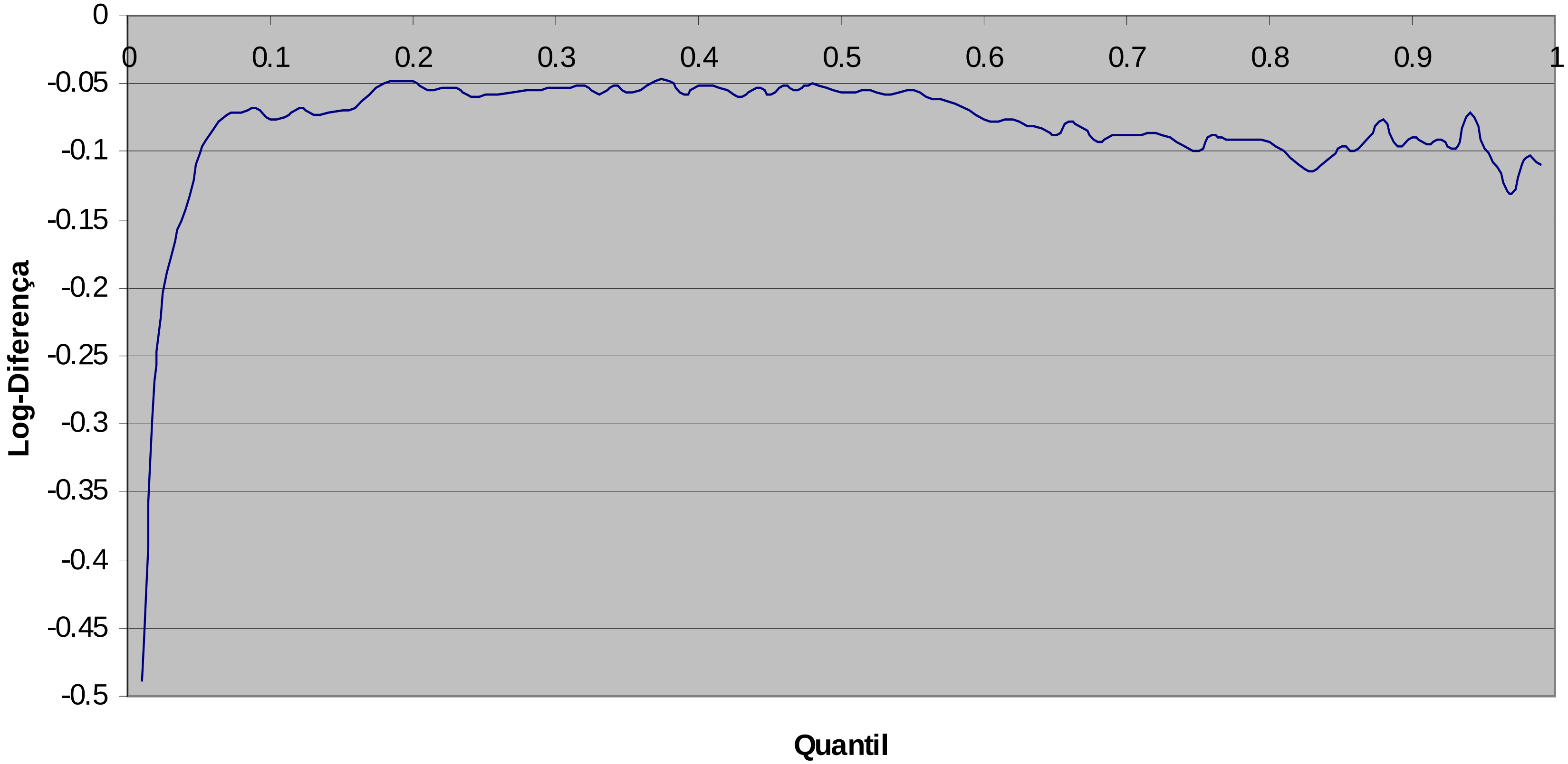
Fonte: Subamostra da PIA-EMP, 1996/1999

Gráfico 2 - Distribuição Acumulada do (Log) Tamanho das empresas, 1996-1999



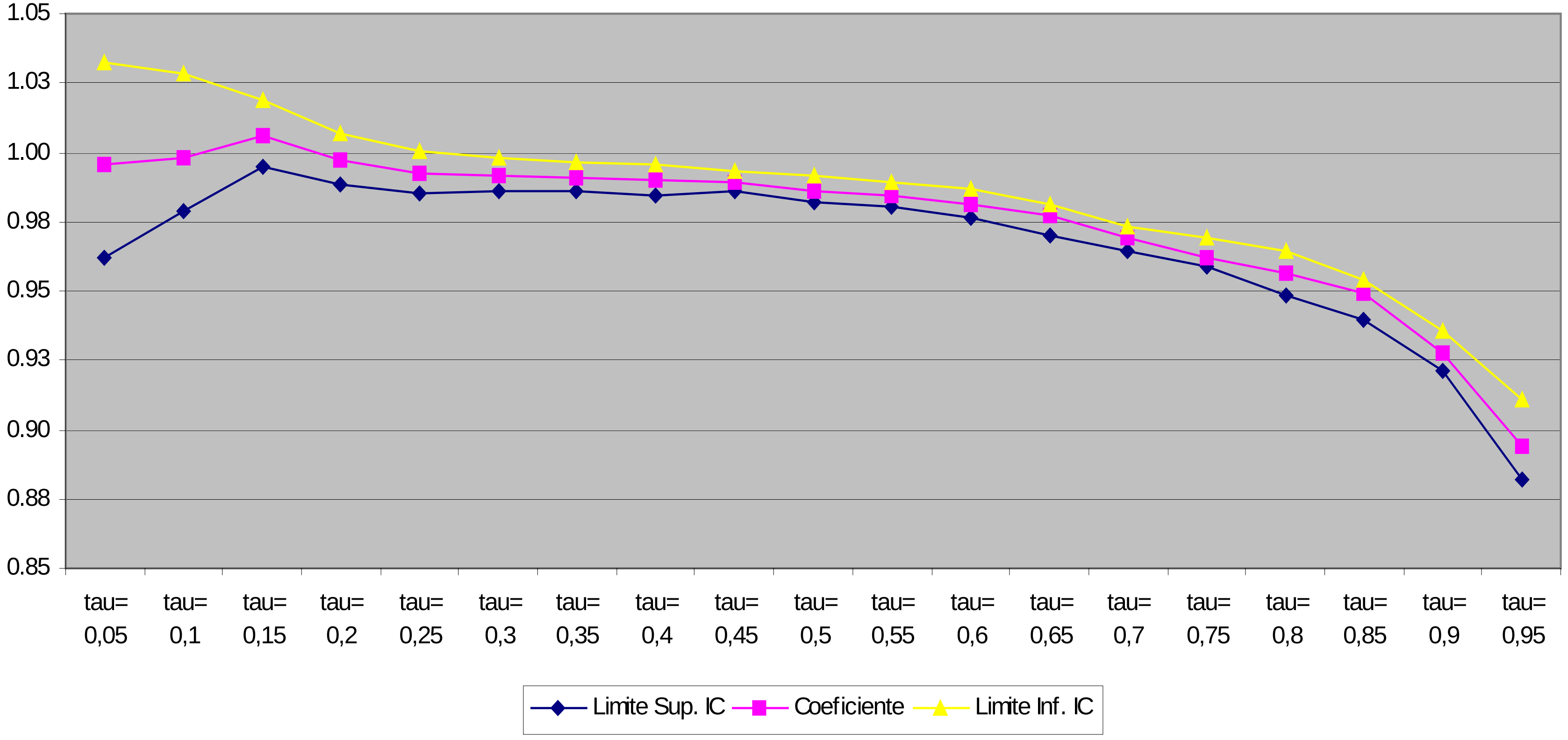
Fonte: Subamostra da PIA-EMP, 1996/1999

Gráfico 3 - Differença dos Quantis do (log) Tamanho das empresas 1996-1999



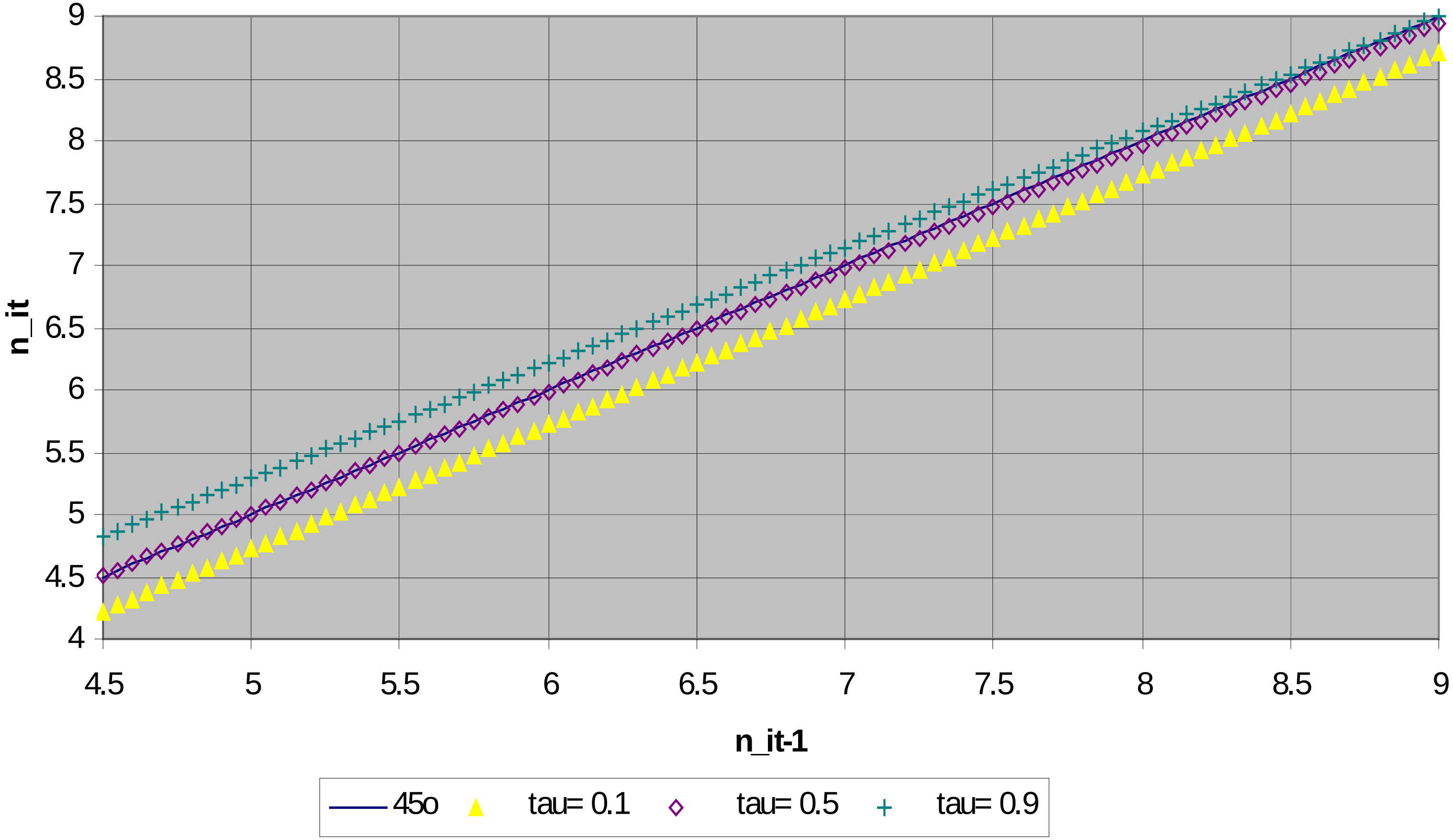
Fonte: Gráfico 2.

Gráfico 4 - Estimativa do Coeficiente Autoregressivo e Intervalo de Confiança (IC) do Modelo de Tamanho de Empresas, 1999-1998



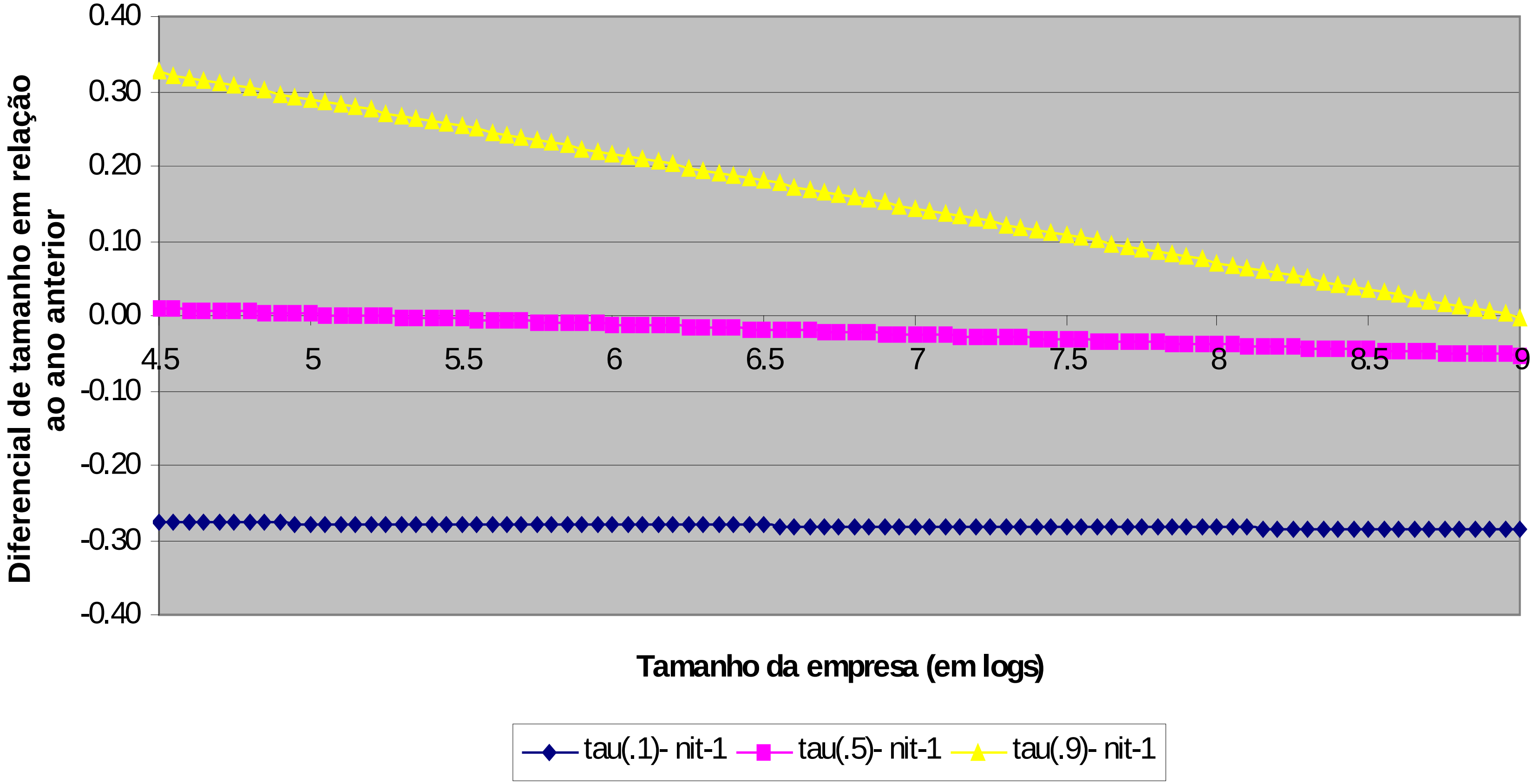
Fonte: cálculos do autor.

Gráfico 5a - Regressão Quantílica para o (log) Tamanho das empresas - 1999-1996



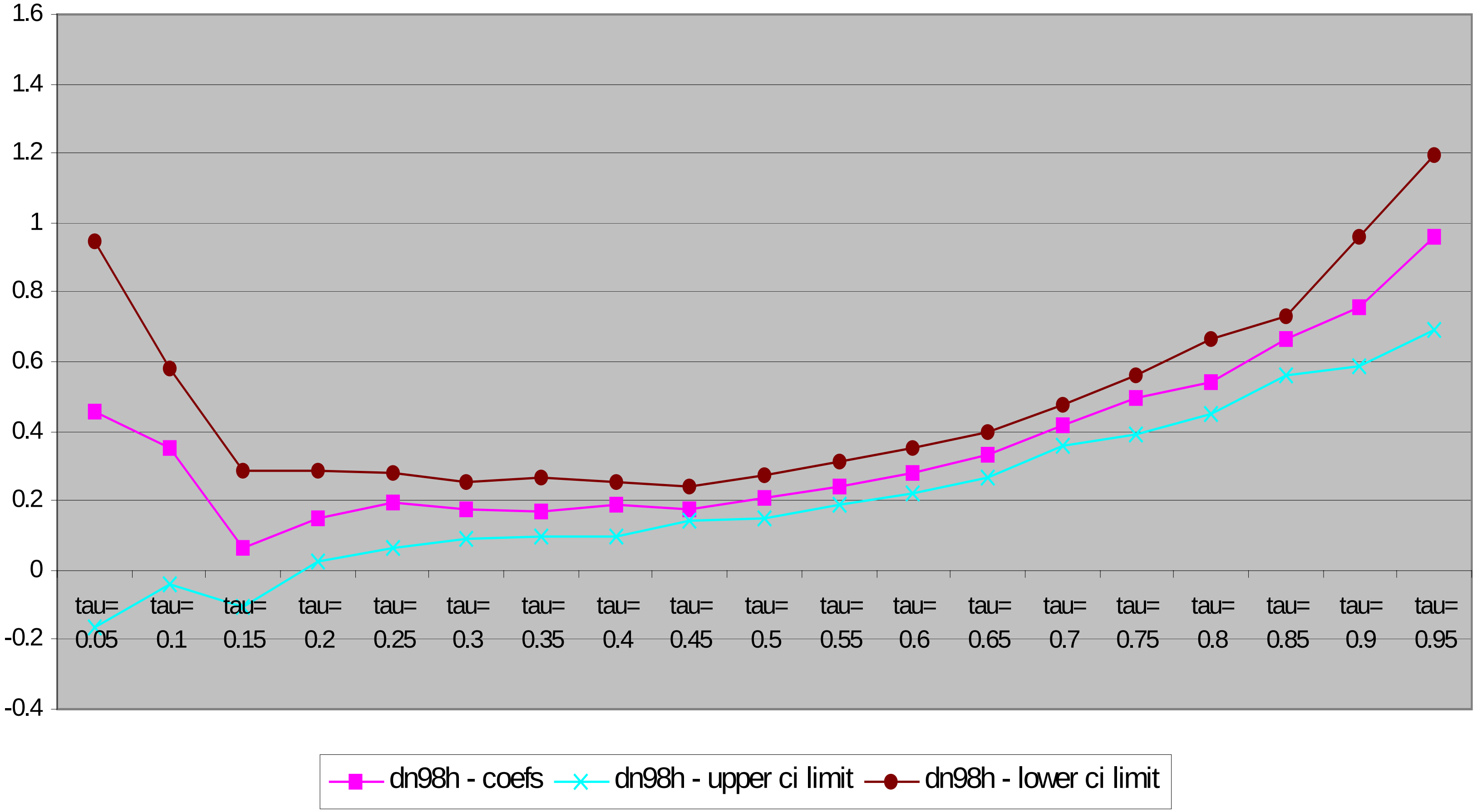
Fonte: cálculos do autor.

Gráfico 5b - Distribuições condicionais do tamanho do emprego, relativas ao tamanho no período anterior - 1999



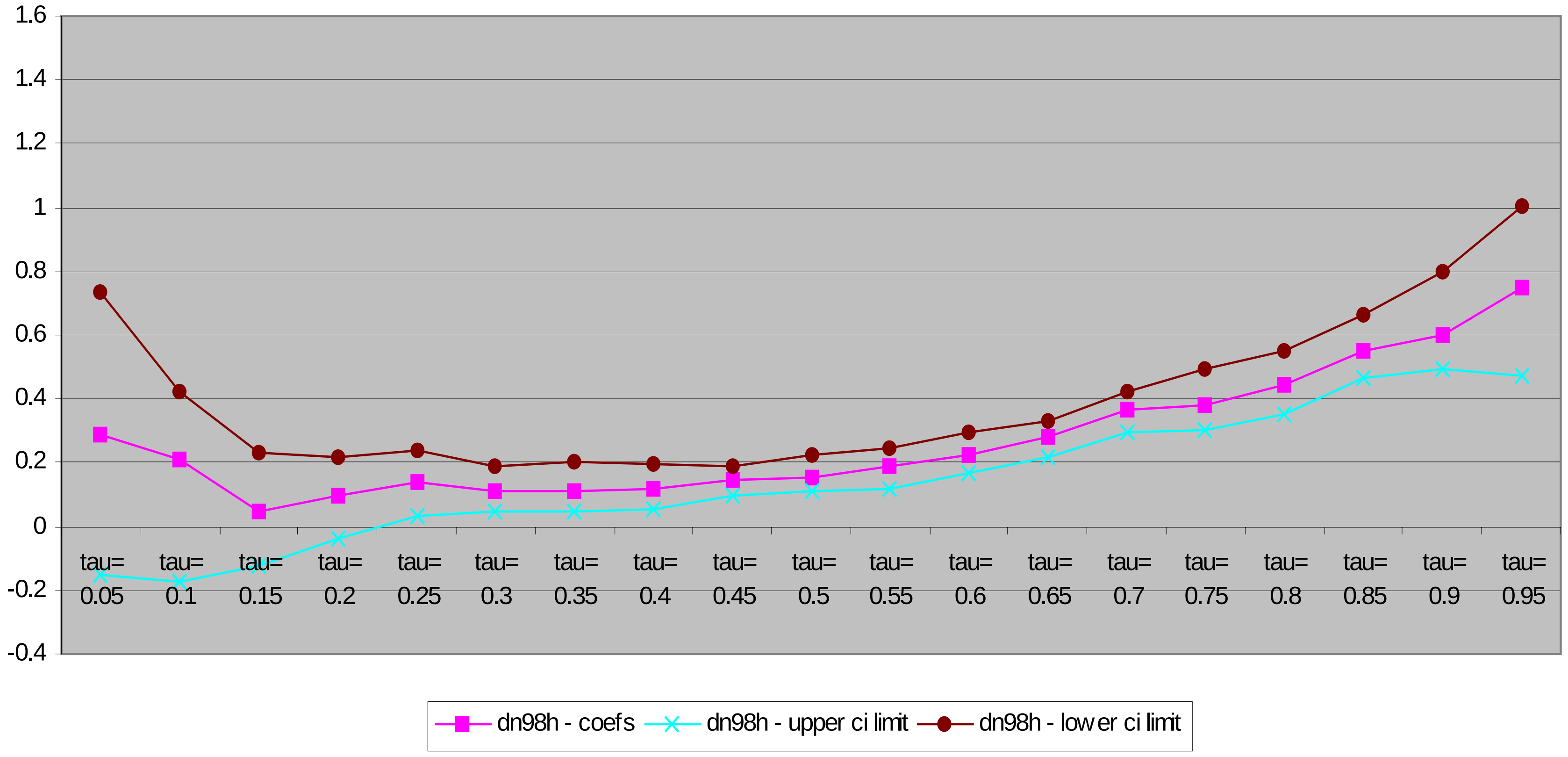
Fonte: cálculos do autor.

Gráfico 6a - Estimativa do Coeficiente Autoregressivo e Intervalo de Confiança (IC) do Modelo de Tamanho de Empresas, 1999-1998 (eq. 1)



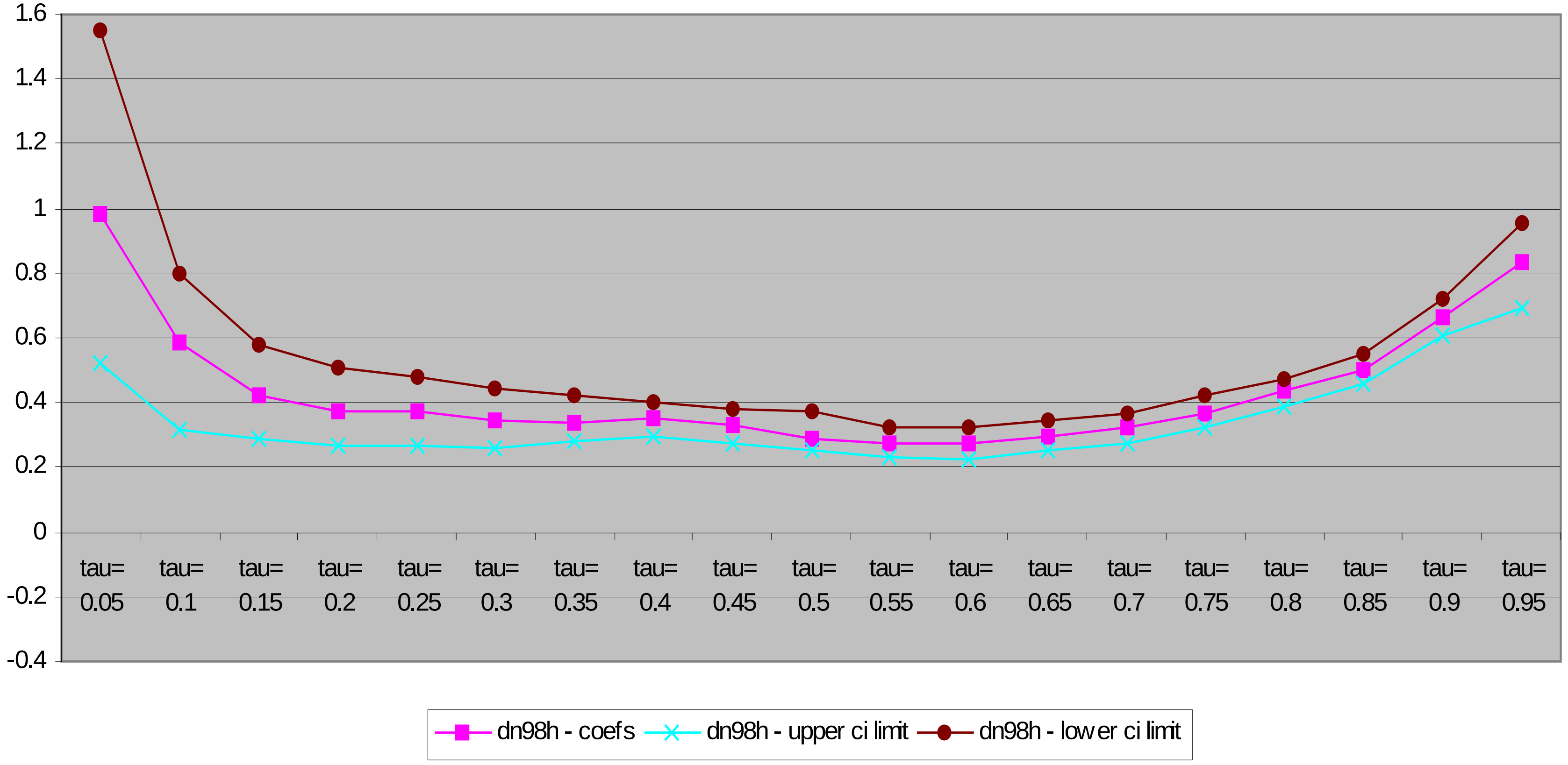
Fonte: cálculos do autor.

Gráfico 6b - Estimativa do Coeficiente Autoregressivo e Intervalo de Confiança (IC) do Modelo de Tamanho de Empresas, 1999-1998 (eq. 2)



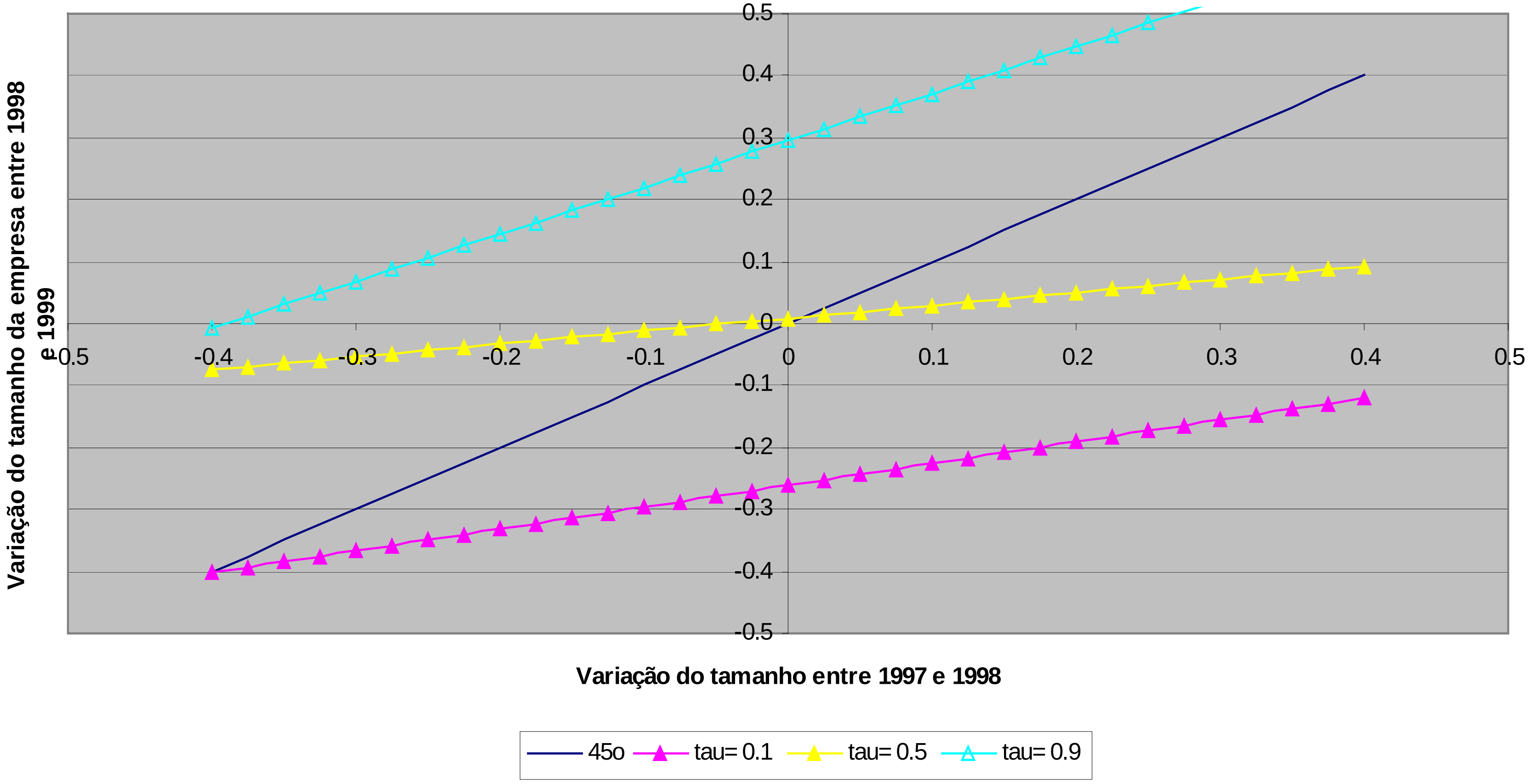
Fonte: Cálculos do autor

Gráfico 6c - Estimativa do Coeficiente Autoregressivo e Intervalo de Confiança (IC) do Modelo de Tamanho de Empresas, 1998-1997 (eq. 3)



Fonte: Cálculos do autor

Gráfico 7 - Quantis Condicionais da variação do tamanho da empresa - 1999



Fonte: Cálculos do autor, baseado no Gráfico 6[a].