



CAPÍTULO 2

EQUAÇÕES DE RENDIMENTOS: QUESTÕES METODOLÓGICAS

Naércio Menezes-Filho
Da FEA/USP

1 INTRODUÇÃO

Este capítulo visa discutir questões metodológicas associadas à estimação de equações de rendimentos. Por equações de rendimentos, entende-se uma tentativa de relacionar os rendimentos auferidos por um trabalhador com suas características pessoais (educação, idade, sexo, cor) e com as características de seu trabalho, como região e setor de atividade. Vários trabalhos tentaram fazer essa associação no Brasil e no resto do mundo.¹ O objetivo deste capítulo não é criticar metodologicamente esses trabalhos, mas listar uma série de problemas associados a qualquer tentativa de estimar equações de rendimentos e discutir as possíveis soluções para tais problemas, assim como as limitações dessas soluções. Primeiramente, procuraremos estudar o problema de forma geral, para em seguida tentar examinar questões particulares envolvidas com cada variável, possivelmente presente em uma equação típica de salários.

O principal objetivo associado a um procedimento que visa estimar econometricamente uma equação de salários, especialmente se o objetivo ulterior é simulação, diz respeito à questão causal. A maneira ideal de garantir que a relação entre duas variáveis seja cau-

1. Para uma resenha, ver Card (1999) e Barros (1998).



sal seria trabalhar com contrafactuais [ver Angrist e Krueger (1999)], ou seja, observar o mesmo trabalhador em duas situações diferentes, controlando por todos os demais condicionantes salariais (inclusive o tempo). Para usar um exemplo clássico de identificação [ver Girshick e Haavelmo (1947)], o analista gostaria de ter em mãos um conjunto de dados resultante de um experimento envolvendo vários consumidores, que tenham tido que decidir diversas vezes quanto adquirir de um produto, defrontando-se com séries alternativas de preços relativos e rendas.

Na questão salarial, o analista gostaria de ter acesso a dados sobre os rendimentos de vários trabalhadores, cada um deles observado em duas situações diferentes no mesmo momento do tempo, por exemplo, com ensino médio completo e com ensino superior completo. De posse do experimento contrafactual, as diferenças entre os rendimentos nas duas situações podem ser entendidas como *causadas* por diferenças educacionais, e um teste simples de média poderia ser conduzido.

Nas ciências experimentais, a aproximação do contrafactual se dá por meio de experimentos aleatórios. São conhecidos os experimentos realizados na medicina em que uma parcela (aleatória) da amostra recebe o medicamento a ser avaliado e a outra parte recebe um placebo. Talvez um dos grandes problemas responsáveis pela falta de credibilidade dos estudos empíricos na economia seja a dificuldade de condução de experimentos aleatórios. Recentemente, porém, estão sendo realizados alguns estudos que utilizam esse tipo de experimento, em especial na área de políticas públicas [ver Heckman, Lalonde e Smith (1999)].

2 FORMA FUNCIONAL

Considerando que o analista possua somente dados observacionais (situação mais comum), vamos supor que a questão de interesse seja a relação entre o logaritmo do salário $(lw)^2$ e uma variável x :

2. É comum a utilização do logaritmo do salário como variável a ser explicada, porque em muitos casos a distribuição dessa variável aproxima-se bem de uma distribuição normal.

$$lw_i = f(x_i) \quad (1)$$

O primeiro ponto a ser levantado é a forma funcional que será admitida a respeito dessa relação. Geralmente, em trabalhos empíricos, não há modelos teóricos guiando a forma funcional a ser adotada e o pesquisador impõe uma série de restrições para prosseguir com o trabalho empírico. As mais comuns são: *a*) a relação entre as variáveis é linear (ou log-linear); e *b*) a relação é a mesma para todos os indivíduos.

Essas duas restrições dão origem a uma forma funcional do tipo:

$$lw_i = \beta_0 + \beta_1 x_i \quad (2)$$

É importante ressaltar que no caso da equação de salários há modelos teóricos que justificam a adoção dessa forma funcional [ver Mincer (1974)]. Entretanto, essas hipóteses já introduzem uma série de restrições no modelo que será levado aos dados. Assim, mesmo que uma estimativa da relação entre lw e x tenha uma interpretação causal, é preciso levar em conta que ela pode ser válida apenas para a amostra em questão ou que pode variar para diferentes valores de x na própria amostra. Por exemplo, nada garante que o adicional salarial correspondente a um ano a mais de estudo seja o mesmo tanto para quem está no começo do ciclo básico como para alguém no fim do ensino médio. Mais ainda, nada garante que o diferencial salarial associado aos diferentes ramos de atividade seja o mesmo para trabalhadores nessas duas situações distintas.

Como relaxar essas hipóteses restritivas? A maneira mais simples de lidar com relações possivelmente não-lineares é através do uso de variáveis interativas, de termos quadráticos ou através de variáveis *dummies* correspondendo a cada grupo em particular. Por exemplo, em uma equação de rendimentos, é praxe incluir um termo quadrático em idade, pois os retornos à experiência tendem a variar ao longo do ciclo de vida. Além disso, pode-se incluir uma *dummy* para cada ano de estudo concluído, ao invés de tratar anos

de estudo como uma variável contínua. Finalmente, seria interessante avaliar os coeficientes estimados de interações das *dummies* educacionais com a idade (e idade ao quadrado), pois é sabido que os retornos à educação variam ao longo do ciclo de vida e que os retornos à experiência dependem do nível educacional [ver Lam e Levinson (1990)].

Procedimentos mais sofisticados de tratar o problema de heterogeneidade da relação entre salários e características pessoais são os modelos de *matching* e regressões quantílicas [ver Koenker e Basset (1978)]. Os modelos de *matching* explicitam o processo de agregação de efeitos causais para diferentes valores das co-variadas, que está implícito no procedimento de Mínimos Quadrados Ordinários (MQO), levando em conta a distribuição amostral dessas co-variadas [ver Angrist e Krueger (1999)]. O método de regressão quantílica, por sua vez, permite a estimação de efeitos que variam ao longo da distribuição condicional da variável dependente [ver Menezes-Filho, Fernandes e Picchetti (1999)].

3 CAUSALIDADE

Para estimar a equação (2), o pesquisador adiciona um termo estocástico com média populacional igual a 0, para levar em conta a aleatoriedade intrínseca ao mundo real. A equação resultante fica sendo:

$$lw_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (3)$$

O próximo passo diz respeito à estimação dos parâmetros β_0 e β_1 . O método usualmente empregado para estimá-los é o modelo de regressão clássico ou MQO:

$$\beta^{mqo} = [x'x]^{-1}[x'lw] \quad (4)$$

Esse modelo baseia-se, além das hipóteses mencionadas, na hipótese essencial de que:

$$E[\varepsilon_i / x_i] = 0 \quad (\text{hipótese H1})$$

Isso implica que não há informação sobre ε_i e embutida em x_i . Note-se que, como $\text{Cov}[x_i, \varepsilon_i] = \text{Cov}[x_i, E(\varepsilon_i/x_i)]$, a hipótese H1 implica que $\text{Cov}[x_i, \varepsilon_i] = 0$.

O principal problema com qualquer tentativa de estimação por MQO é a validade dessa hipótese, pois os coeficientes estimados só serão estimadores consistentes dos parâmetros que relacionam y_i a x_i na população se a hipótese H1 for válida, o que tornaria essa relação causal. Caso contrário, o problema de endogeneidade entra em cena e os coeficientes estimados não convergem em probabilidade para os parâmetros populacionais.

No caso da equação de salários, o problema é especialmente sério, uma vez que a variável resultante depende do “potencial de auferir rendimentos” de cada trabalhador, que é, em grande medida, não-observável. Esse potencial depende de fatores como habilidade, criatividade, capacidade de adaptação a mudanças, capacidade de relacionamento, ambição etc. É razoável admitir que esse “potencial” está correlacionado com alguns dos principais determinantes (observáveis) dos salários, como escolaridade, ocupação, região de domicílio, ramo de atividade etc. Isso faz com que a hipótese H1 não seja válida, o que invalidaria a interpretação causal dos coeficientes estimados.

Como exemplo de como a correlação entre o “erro” e os “regressores” pode afetar a interpretação dos coeficientes estimados, podemos considerar os diferenciais regionais de salários. Vamos supor que os rendimentos de um trabalhador dependam de sua região de residência, por fatores relacionados a oferta e demanda regional por trabalho qualificado:

$$lw_i = f(\text{região}_i, x_i) + f_i + \varepsilon_i \quad (5)$$

Nessa equação, o termo aleatório contém características não-observadas do trabalhador como as descritas anteriormente (f_i), e x_i engloba os outros determinantes salariais. Entretanto, a decisão de moradia do trabalhador leva em conta o fato de suas características

observáveis e não-observáveis terem sido mais bem remuneradas em algumas regiões, ou seja, que sua capacidade de auferir rendimentos depende da sua região de moradia. Dessa forma, algumas regiões terão uma concentração maior de trabalhadores mais “habilidosos” e assim:

$$E[f_i | região_i] \neq 0 \quad (6)$$

Isso faz com que o pesquisador, ao estimar o efeito da região de moradia sobre os salários ignorando as características não-observáveis, “capture” simultaneamente o efeito de f_i sobre os salários, o que afeta a interpretação dos coeficientes estimados. Assim, não se pode dizer com certeza que um trabalhador com determinadas características, ao migrar de um estado do Brasil para outro, irá obter o ganho ou perda salarial estimado através do método de MQO.

4 MÉTODOS PARA LIDAR COM O PROBLEMA DE ENDOGENEIDADE

Vamos agora discutir os principais procedimentos quase sempre empregados para tentar contornar o fato de a hipótese H1 geralmente não ter sido válida em modelos econométricos, em particular em equações salariais. Os principais métodos empregados são: *a*) seleção por observáveis (variáveis de controle); *b*) efeitos fixos; e *c*) variáveis instrumentais.

4.1 Seleção por observáveis (variáveis de controle)

O método mais fácil de lidar com o problema de causalidade é a inclusão de variáveis de controle observáveis que capturem a correlação entre os regressores de interesse e os componentes aleatórios. No caso do efeito da educação sobre os salários, por exemplo, vários estudos incluem resultados de testes de QI e *background* familiar na equação de rendimentos do trabalho [ver Card (1999)].³ Um exemplo clássico dessa metodologia é o artigo de Ashenfelter e Krueger

3. Os resultados para o caso americano surpreendentemente revelam que os controles parecem não afetar muito os retornos à educação.

(1994), em que os autores controlam o *background* familiar comparando os diferenciais salariais e os anos de estudo de gêmeos univitelinos. Particularmente relevante para o caso brasileiro é o estudo de Lam e Schoeni (1994), que inclui escolaridade dos pais como controle para o *background* familiar e encontra evidências de uma redução significativa nos retornos estimados à escolaridade.

A hipótese de identificação básica desses modelos é que a variável de controle é a única razão pela qual o regressor de interesse e o termo erro são correlacionados. Isso significa que o regressor é determinado independentemente do potencial para auferir salários, uma vez que o controle esteja presente. Quer dizer, se

$$lw_i = \beta_0 + \beta_1 x_i + \varepsilon_i \text{ e } E[\varepsilon_i / x_i] \neq 0 \quad (7)$$

Se decomposermos o termo erro em duas partes:

$$\varepsilon_i = \gamma w_i + u_i \quad (8)$$

A hipótese de identificação requer que:

$$E[u_i / x_i] \neq 0 \quad (9)$$

O grande problema dessa metodologia é a validade da hipótese de identificação. Dificilmente podemos admitir que uma variável de controle capture toda a correlação entre o regressor e o potencial para auferir rendimentos. Isso só é verdade quando o pesquisador conhece o processo que gera o regressor. Vamos supor, por exemplo, que o pesquisador está interessado na diferença salarial resultante do aprendizado em escolas públicas e privadas e que o processo de admissão nas escolas privadas seja conhecido. Sob essa hipótese, a inclusão das características que são levadas em conta no processo de admissão nas escolas privadas (renda dos pais, escolaridade dos pais, região de residência, cor, sexo etc.) poderia fazer com que o problema de endogeneidade fosse eliminado.

Os problemas adicionais dessa metodologia [ver Angrist e Krueger (1999)] são que, se a variável de interesse for medida com

erro (quase todas as variáveis microeconômicas o são), a inclusão de controles tenderá a agravar o viés atenuador decorrente desse erro. Finalmente, a endogeneidade do controle afeta a consistência do estimador de MQO.

4.2 Efeitos fixos

Outra maneira relativamente comum de lidar com o problema de endogeneidade (causalidade) é a metodologia de efeitos fixos. Essa metodologia envolve o controle por características não-observáveis da unidade observacional, desde que essas características sejam invariantes com relação ao tempo. A metodologia só pode ser aplicada quando o pesquisador dispõe de observações repetidas de cada unidade ao longo do tempo.

Vamos supor, por exemplo, que o trabalhador tem a possibilidade de trabalhar em dois setores: químico e têxtil. Seja Y_{it} o salário potencial de um trabalhador no setor têxtil e Y_{ci} o salário potencial do mesmo trabalhador no setor químico, de forma que:

$$lw_i = \beta_0 + \beta_1 x_i + \gamma q_i + \varepsilon_i \quad (10)$$

onde q_i é uma *dummy* para o trabalhador empregado no setor químico. O objetivo do pesquisador é estimar o ganho associado ao emprego no setor químico em comparação ao têxtil. O problema da estimação é que a escolha do setor está potencialmente correlacionada com o potencial de rendimentos do trabalhador, ou seja, dado que o setor químico tende a remunerar melhor as habilidades não-observadas, como criatividade e perseverança, o trabalhador com essas características será encaminhado para esse setor. Dessa forma, o q_i estimado não seria um estimador consistente da relação causal entre o emprego no setor químico e o salário de um trabalhador qualquer, pois:

$$E[\varepsilon_i / q_i] \neq 0 \quad (11)$$

Uma possível solução para esse problema é o controle por efeitos fixos, também chamados de “efeitos específicos” ou “heterogeneidade não-observada”:

$$lw_i = \beta_0 + \beta_1 x_i + \gamma q_i + f_i + \varepsilon_i \quad (12)$$

Nesse caso, a hipótese de identificação requer que:

$$E[\varepsilon_i | q_i, f_i] = 0 \quad (13)$$

ou seja, que toda correlação entre q_i e ε_i pode ser capturada por uma co-variada f_i que não varia entre períodos. A estimação desse modelo poderia se dar incluindo uma *dummy* para cada trabalhador, através de MQO (estimador *dentro de grupos* ou de *efeitos fixos*) ou em primeiras diferenças:

$$lw_{it} - lw_{it-1} = \beta_1 (x_{it} - x_{it-1}) + \gamma (q_{it} - q_{it-1}) + \varepsilon_{it} - \varepsilon_{it-1} \quad (14)$$

ou seja:

$$\Delta lw_{it} = \beta_1 \Delta x_{it} + \gamma \Delta q_{it} + \Delta \varepsilon_{it} \quad (15)$$

Nesse último caso, a heterogeneidade não-observada seria automaticamente eliminada. É importante ressaltar que para a identificação do efeito “emprego no setor químico” o analista necessita de variação no setor de atividade ao longo do tempo, ou seja, uma parcela significativa dos trabalhadores amostrados tem de ter migrado de um setor para o outro durante o período amostral.

Um dos principais problemas com o estimador de efeitos fixos pode acontecer se o pesquisador decidir incluir a variável dependente defasada (lw_{it-1}) no seu modelo de regressão original, ou seja:

$$lw_{it} = \beta_0 + \beta_1 x_{it} + \gamma q_{it} + \lambda lw_{it-1} + f_i + \varepsilon_{it} \quad (16)$$

Essa variável pode ser muito importante em alguns casos. Vamos supor, ainda segundo o exemplo anterior, que os trabalhadores que mudaram de setor o fizeram porque receberam uma diminuição *transitória* no seu salário real. Dessa forma, a análise em primeiras

diferenças iria imputar o aumento salarial à mudança de setor, quando na verdade o que aconteceu foi uma regressão para a média. O problema é que tanto o estimador de efeitos fixos como o de primeiras diferenças não são robustos à inclusão de uma variável dependente defasada [ver Nickell (1981)]. Por exemplo, no caso do estimador de primeiras diferenças, a transformação gera correlação automática entre o regressor e o erro:

$$\Delta lw_{it} = \beta_1 \Delta x_{it} + \gamma \Delta q_{it} + \Delta lw_{it-1} + \Delta \varepsilon_{it} \quad (17)$$

$$E[\Delta \varepsilon_{it} / \Delta lw_{it-1}] = 0 \quad (18)$$

pois:

$$E[\varepsilon_{it-1} / lw_{it-1}] \neq 0 \quad (19)$$

Uma possível alternativa seria utilizar na análise apenas trabalhadores que mudaram de setor involuntariamente, devido ao fechamento de seus locais de trabalho, por exemplo.⁴ Finalmente, um outro problema sério com as transformações usadas para estimação do modelo com efeitos específicos é que elas tendem a agravar erros de medida presentes nos dados originais [ver Griliches e Hausman (1986)].

4.3 Variáveis instrumentais

O método mais utilizado para corrigir o problema de endogeneidade é o de variáveis instrumentais. A idéia é utilizar variações exógenas para aproximar os experimentos aleatórios. Suponhamos, desta feita, que o objetivo do pesquisador é analisar os diferenciais de salário associados à cidade de residência no Brasil. Da mesma forma que nos casos anteriores, simplesmente comparar os rendimentos controlados dos trabalhadores nas duas cidades não seria suficiente para estimar o parâmetro da relação causal entre local de residência e rendimentos.

4. A referência clássica para diferenciais salariais associados a ramos de atividade é Krueger e Summers (1987).

Como visto, os determinantes do rendimento potencial do trabalhador estariam correlacionados com o local de residência, já que os trabalhadores com maior potencial tenderiam a se localizar em cidades onde esse potencial fosse mais valorizado:

$$lw_i = \beta_0 + \gamma c_i + \varepsilon_i \quad (20)$$

sendo que c_i indica a cidade de residência e:

$$E[\varepsilon_i / c_i] \neq 0 \quad (21)$$

Vamos supor que a variável c_i possa ser decomposta em dois componentes, um estocástico e um determinístico, de forma que:

$$lw_i = \beta_0 + \gamma(c_i^* + v_i) + \varepsilon_i \quad (22)$$

O problema de simultaneidade pode ser entendido como uma correlação entre v_i e ε_i ou seja:

$$E[\varepsilon_i / v_i] \neq 0 \quad (23)$$

De forma que o $\hat{\gamma}$ estimado por MQO incorpora o efeito de ε_i em lw_i , tornando a regressão espúria. O pesquisador necessita de instrumentos, z_i , ou seja, variáveis correlacionadas com c_i^* , mas não com v_i , quer dizer:

$$E[\varepsilon_i / z_i] = 0 \text{ e } E[c_i^* / z_i] \neq 0 \quad (24)$$

O grande problema com esse método é a procura por instrumentos que claramente satisfaçam as condições expressas em (24). No exemplo em questão, precisaríamos de alguma variável que fosse correlacionada com a cidade de residência, porém não correlacionada com a capacidade de obtenção de rendimentos do trabalho (atratividade da cidade, por exemplo). Grande parte da literatura americana trabalha com episódios esporádicos como sorteios para serviço militar [ver Angrist (1990)]⁵ ou mudanças nas regras que

5. Nesse estudo, Angrist (1990) analisou o efeito do serviço militar na Guerra do Vietnã sobre os salários, utilizando como instrumento sorteios que definiam os jovens americanos que iriam servir o Exército.

definem tamanho de classes para identificar seu efeito sobre o rendimento escolar [ver Angrist e Lavy (1998)], ou mudanças na legislação tributária para investigar oferta de trabalho [ver Blundell, Duncan e Meghir (1999)].

O estimador de variáveis instrumentais nesse caso é dado por:

$$\gamma^{iv} = [c'z(z'z)^{-1}z'c]^{-1}[c'z(z'z)^{-1}z'lw] \quad (25)$$

A hipótese básica do modelo de variáveis instrumentais é que a única razão pela qual lw_i varia com z_i é porque c_i varia com z_i . Esse estimador é equivalente ao de MQO em dois estágios, que usa os valores previstos de uma regressão de c_i em z_i no primeiro estágio como regressores no segundo estágio, em que a variável dependente é lw_i .

No caso em que o número de instrumentos é igual ao de regressores, o método de variáveis instrumentais gera o estimador:

$$\gamma^{iv} = [z'c]^{-1}[z'lw] \quad (26)$$

Finalmente, no caso em que os regressores são exógenos, temos o estimador de MQO.

Outro problema com o método de variáveis instrumentais é que em amostras finitas as propriedades do estimador não são confiáveis [ver Staiger e Stock (1997)]. Nos casos de correlação fraca entre os instrumentos e os regressores endógenos e de grande número de instrumentos, as estimativas de variáveis instrumentais podem se aproximar perigosamente das de MQO, dando a falsa impressão de exogeneidade dos regressores.

5 CONCLUSÕES

Neste capítulo, tentou-se entender melhor as questões metodológicas envolvidas na estimação de uma equação de rendimentos, ou seja, da relação entre o salário recebido por um trabalhador e várias características pessoais e de seu ambiente de trabalho. Vimos que o analis-

ta empírico usualmente impõe uma série de hipóteses no modelo a ser estimado e que essas hipóteses podem ter implicações muito fortes nos resultados obtidos.

Destacamos as hipóteses relacionadas à forma funcional da equação de rendimentos a ser estimada, principalmente aquelas relacionadas à log-linearidade da relação entre os salários e seus determinantes e da invariância dos efeitos com relação aos valores assumidos pelas co-variadas. Examinamos também algumas possibilidades de relaxamento dessas hipóteses, através de variáveis *dummies* e de interações.

Finalmente, examinamos em detalhe as implicações da hipótese de exogeneidade dos regressores, ou seja, de que os termos aleatórios e as variáveis independentes não são correlacionados. Vimos que essa hipótese é particularmente restritiva no caso da equação de rendimentos, uma vez que um dos principais determinantes do salário é sua capacidade de auferir rendimentos no mercado de trabalho e essa capacidade não pode ser observada pelo analista empírico. Procuramos examinar alguns métodos possíveis de lidar com esse problema, seja através de variáveis de controle, efeitos fixos ou variáveis instrumentais. Tentou-se ainda salientar as dificuldades envolvidas com cada um desses métodos.

Devemos salientar, entretanto, que o objetivo deste capítulo não é o de criticar metodologicamente os estudos existentes para o caso brasileiro, uma vez que o nosso mercado de trabalho ainda é relativamente pouco estudado. Assim, antes de nos concentrarmos na utilização de técnicas econométricas sofisticadas para lidar com os problemas aqui examinados, talvez valha a pena utilizar um aparato técnico mais simples, a fim de compreender melhor os vários aspectos que diferenciam nosso mercado de trabalho daqueles já exaustivamente estudados nas economias mais desenvolvidas.

BIBLIOGRAFIA

- ANGRIST, J. Lifetime earnings and the Vietnam era draft lottery: evidence from social security administrative records. *American Economic Review*, v. 80, 1990.
- ANGRIST, J., KRUEGER, A. Empirical strategies in labor economics. In: ASHENFELTER, O., CARD, D. *Handbook of Labor Economics*, v. 3A, Elsevier, 1999.
- ANGRIST, J., LAVY, V. Using maimonides rule to estimate the effects of class size on schooling achievement. *Quarterly Journal of Economics*, v. 114, p. 533-575, 1998.
- ASHENFELTER, O., KRUEGER, A. Estimates of returns to schooling using a new sample of Twins. *American Economic Review*, v. 84, n. 5, 1994.
- BARROS, R. P. de. *Os determinantes da desigualdade no Brasil*. Rio de Janeiro: IPEA, 1998, mimeo.
- BLUNDELL, R., DUNCAN, A., MEGHIR, C. Estimating labor supply responses using tax reforms. *Econometrica*, v. 66, p. 827-861, 1999.
- CARD, D. The causal effect of education on earnings. In: ASHENFELTER, O., CARD, D. *Handbook of Labor Economics*, v. 3A, Elsevier, 1999.
- GIRSHICK, M. A., HAAVELMO, T. Statistical analysis of the demand for food: examples of simultaneous estimation of structural equations. *Econometrica*, v. 15, n. 2, p. 79-110, 1947.
- GRILICHES, Z., HAUSMAN, J. Errors in variables in panel data. *Journal of Econometrics*, v. 31, n. 1, 1986.
- HECKMAN, J., LALONDE, R., SMITH, J. The economics and econometrics of active labor market programs. In: ASHENFELTER, O., CARD, D. *Handbook of Labor Economics*, v. 3A, Elsevier, 1999.
- KOENKER, R., BASSET, G. Regression quantiles. *Econometrica*, v. 46, p. 33-50, 1978.
- KRUEGER, A., SUMMERS, L. Efficiency wages and interindustry wage structure. *Econometrica*, v. 56, n. 2, 1987.
- LAM, D., LEVINSON, D. Idade, experiência e diferenciais de renda no Brasil. *Pesquisa e Planejamento Econômico*, Rio de Janeiro, v. 20, n. 2, 1990.

- LAM, D., SCHOENI, D. Effects of family background on earnings and returns to schooling: evidence from Brazil. *Journal of Political Economy*, v. 101, n. 4, 1994.
- MENEZES-FILHO, N., FERNANDES, R., PICCHETTI, P. Wage inequality in Brazil; some stylised facts. *Anais do XVI Encontro de Econometria*, Belém-PA, 1999.
- MINCER, J. *Schooling, experience and earnings*. Columbia University Press, 1974.
- NICKELL, S. Biases in dynamic models with fixed effects. *Econometrica*, v. 49, n. 6, 1981.
- STAIGER, D., STOCK, J. Instrumental variables regression with weak instruments. *Econometrica*, v. 65, n. 3, 1997.

