

O Diferencial de Salários Formal-Informal no Brasil: Segmentação ou Viés de Seleção?*

Naércio Aquino Menezes Filho **

Marcos Mendes ***

Eduardo Simões de Almeida ****

Sumário: 1. Introdução; 2. Metodologia e modelo econométrico; 3. Análise descritiva dos dados; 4. Resultados; 5. Conclusões e comentários finais.

Palavras-chave: salários; trabalho informal; cross-section repetidas; análise de coorte.

Códigos JEL: C23; J31.

Neste artigo são investigados os determinantes do diferencial de salários entre os mercados de trabalho formal e informal no Brasil. Utiliza-se um método econométrico de cross-section repetidas (pseudo-painéis), no qual o agrupamento dos dados por geração, tempo e escolaridade permite controlar o fenômeno estudado por características observáveis e não observáveis dos indivíduos. Há fortes evidências de viés de auto-seleção, indicando que os salários mais altos no setor formal decorrem dos melhores atributos individuais não observáveis dos empregados neste setor e não de características intrínsecas a este setor, como seria de se esperar pela hipótese de segmentação.

The paper examines the determinants of the wage differential between workers in the formal and informal sectors of the Brazilian labor market. An econometric method of repeated cross-sections (pseudo-panel) is used, that allows the researcher to control for observed and unobserved characteristics of the individuals in the sample. We present strong evidence of self-selection bias, meaning that the wage differentials are the consequence of better individual non-observable characteristics of the workers in the formal sector and not to the characteristics of this sector.

* Artigo recebido em abr. 2003 e aprovado em fev. 2004.

** Professor do Departamento de Economia da USP.

*** Doutor em Economia pela USP e Senado Federal.

**** Doutor em Economia pela USP.

1. Introdução

A questão do trabalho informal no Brasil ganha importância a cada dia. Além do grande número de trabalhadores informais existente hoje, este número vem crescendo ao longo do tempo. A cada ano é menor o percentual de trabalhadores que contribui para instituto de previdência e que possui carteira de trabalho assinada. Se for verdadeira a idéia de que os salários no setor informal são menores que os do setor formal, então teremos um contingente cada vez maior de trabalhadores com remuneração mais baixa. Além disto, os trabalhadores do setor formal, protegidos pela lei, conseguem vantagens trabalhistas não acessíveis aos trabalhadores do desregulamentado setor informal, o que resulta na ampliação do diferencial de salários entre os dois grupos.

Para que possamos atribuir uma relação de causalidade para correlação observada entre características do contrato de trabalho e o nível salarial, é necessário levar em conta vários fatores (Menezes-Filho, 2001). Em primeiro lugar, é preciso controlar pelo nível educacional dos indivíduos, que é positivamente correlacionado com o nível salarial e com a formalização. Além disso, gerações melhor preparadas, menos numerosas ou que entram no mercado de trabalho num período mais próspero, têm mais chances de serem melhor remuneradas. Finalmente, uma série de características não observáveis dos indivíduos pode fazer com que ele acabe trabalhando, por opção ou necessidade, no setor informal da economia.

Neste artigo, utilizamos um enfoque econométrico baseado no agrupamento dos dados por coorte-tempo-escolaridade, em conformidade com a literatura micro-econométrica recente,¹ para estudar um fenômeno tipicamente brasileiro: a informalidade no mercado de trabalho e os diferenciais de salários associados a ela. Utilizamos dados das Pesquisas Nacionais por Amostra de Domicílios (PNAD), que correspondem a *cross-sections* repetidas, entre 1981 a 2001, para todos os trabalhadores do sexo masculino.

O resultado central do trabalho é que o principal determinante do diferencial de salários entre trabalhadores dos dois setores é a diferença entre as características individuais observadas (principalmente educação) e não observadas entre eles, e não a formalidade ou informalidade do contrato de trabalho. Condicional ao nível de escolaridade, o salário no setor informal é maior que no setor formal; o que indica que as vantagens que a lei garante aos trabalhadores do setor formal (aviso prévio, adicional de férias, 13º salário, contribuição previdenciária patronal, etc.) são compensadas no setor informal por uma maior remuneração. Portanto, o

¹ Meghir e Whitehouse (1996), Blundell et alii (1998), Angrist (1991), Browning et alii (1985), Deaton (1985), Attanazio e Browning (1985).

diferencial observado nos dados desagregados decorre da existência de viés de auto-seleção, decorrente da correlação positiva entre características individuais não observáveis (habilidade, inteligência, qualidade da educação) e a probabilidade de emprego no setor formal. Desta forma, as evidências indicam que não existe segmentação no mercado de trabalho brasileiro.

Essas constatações são fundamentais para a formulação de política econômica. Elas indicam que para elevar salários o importante não é regulamentar o mercado de trabalho ou tentar coibir a informalização do emprego, mas sim elevar a quantidade e qualidade da educação dos trabalhadores brasileiros. A metodologia e o modelo econométrico adotados neste artigo são expostos na segunda seção. Uma análise descritiva da base de dados é empreendida na terceira seção. Os resultados da estimação e sua análise são apresentados na quarta seção. As conclusões estão na quinta e última parte.

2. Metodologia e Modelo Econométrico

O objetivo do trabalho é avaliar o diferencial dos salários entre os setores formal e informal. Assim, a equação básica a ser estimada é:

$$w_{i(t)t} = \beta_0 + \beta_1 x_{i(t)t} + u_{i(t)t} \quad (1)$$

na qual $w_{i(t)t}$ é o salário do indivíduo i observado no momento t , $x_{i(t)t}$ é uma variável *dummy* que indica se o indivíduo i observado no momento t pertence ao mercado informal ($x_{i(t)t} = 0$) ou ao mercado formal ($x_{i(t)t} = 1$) e $u_{i(t)t}$ é um termo estocástico.

O problema básico de estimar tal modelo é que $u_{i(t)t}$ pode ser decomposto em:

$$u_{i(t)t} = \epsilon_{i(t)t} + f_{i(t)}$$

em que $f_{i(t)}$ é o efeito fixo de cada indivíduo da *cross-section* observada em t e $\epsilon_{i(t)t}$ é o erro homocedástico e não-autocorrelacionado.

A variável $x_{i(t)t}$ pode estar correlacionada com $f_{i(t)}$ e, portanto, com $u_{i(t)t}$, de modo que β não seria um estimador consistente do diferencial de salário entre os setores formal e informal. Ou seja, características não observáveis dos indivíduos contidas em $f_{i(t)}$ (habilidade, força de vontade, determinação, criatividade etc.) e captadas por $u_{i(t)t}$ podem estar determinando a escolha entre trabalhar no mercado formal ou no mercado informal, ou ainda determinando o nível de salários que o indivíduo obtém em cada um dos dois mercados de trabalho.

Assim, para obtermos estimadores consistentes da relação entre formalidade e salários é necessário recorrer ao método de variáveis instrumentais.²

A classe mais importante de instrumentos consiste de funções do tempo, que, sob certas hipóteses (ver abaixo), não serão correlacionadas com $f_{i(t)}$. Além disso, se houver efeitos de coorte na população, como freqüentemente ocorre em estudos comportamentais, estimadores mais eficientes podem ser obtidos se também utilizarmos o ano de nascimento como instrumento. Consideremos a seguinte projeção linear de $x_{i(t)t}$, na qual g_{1k} são funções do tempo, g_{2m} são funções dos coortes e g_{3j} são interações entre tempo e coorte.

$$x_{i(t)t} = \sum_{k=1}^K \delta_{1k} g_{1k}(t) + \sum_{m=1}^M \delta_{2m} g_{2m}(c_{i(t)}) + \sum_{j=1}^J \delta_{3j} g_{3j}(c_{i(t)}, t) + u_{i(t)t} \quad (2)$$

Denotemos por \hat{x}_t os valores da equação (2) previstos por mínimos quadrados ordinários. A consistência do estimador de variáveis instrumentais requer que:

- \hat{x}_t varie com o tempo.
- $\text{plim} [(1/NT) \sum_{i,t} \hat{x}_t f_{i(t)}] = 0$.

Um instrumento baseado em efeitos de coorte, entretanto, provavelmente estará correlacionado com $f_{i(t)}$. Sendo assim, a segunda condição não é satisfeita. Definamos, todavia, $\bar{f}_c(t)$ como o efeito fixo médio dos indivíduos observados no instante t e nascidos no ano c . Considerando $v_{i(t)}$ como o desvio de $f_{i(t)}$ com relação à média amostral $\bar{f}_c(t)$ e $v'_{c(t)}$ como o desvio da média amostral com relação à média populacional \bar{f}_c^* , temos que:

$$f_{i(t)} = \bar{f}_c^* + v'_{c(t)} + v_{i(t)t} \quad (3)$$

Claramente, se supusermos que a população é fechada, \bar{f}_c^* será constante. Como \bar{f}_c^* pode ser representado explicitamente na equação (1) por dummies de coorte, a segunda condição requer simplesmente que \hat{x}_t não seja assintoticamente correlacionado com $v'_{c(t)}$ e $v_{i(t)t}$. Mas quando $N \rightarrow \infty$ e o número de coortes é mantido fixo, o erro amostral $v'_{c(t)}$ tende a zero e os desvios individuais $v_{i(t)t}$

²Uma discussão detalhada com respeito à identificação do modelo de efeitos fixos no caso de cross-sections repetidas é realizada por Moffit (1993).

passam a não depender de t e, portanto, de \hat{x}_t , pois a amostra torna-se idêntica à população.

A população de trabalhadores das PNADS não é fechada: algumas pessoas podem desistir de trabalhar e sair do mercado. Para solucionar esse problema, adotamos procedimento similar ao de Meghir e Whitehouse (1996), limitando a amostra a trabalhadores homens, sob a hipótese de que os homens tendem a permanecer no mercado de trabalho a despeito de condições adversas (baixo salário, alto desemprego, más condições de trabalho).

De modo geral, as funções g_{2m} não possuem, necessariamente, apenas as coortes como argumento. Qualquer conjunto de características invariantes no tempo como, por exemplo educação (se ignorarmos a possibilidade dos indivíduos retomarem os estudos), pode ser incluído em (2). Neste artigo utilizaremos coorte, tempo e educação como várias agrupadoras.

Angrist (1991) mostra que utilizar as interações entre coorte, tempo e educação como variáveis instrumentais equivale ao estimador de agrupamento proposto por Browning et alii (1985), que consiste em agrupar os dados por ano, coorte e escolaridade, calcular a média salarial e a porcentagem de formais para cada grupo e estimar o modelo com as médias dos dados agrupados. Assim, utilizar os instrumentos listados em (2) é o mesmo que transformar (1) em:

$$\bar{w}_{tce} = \beta_0 + \beta \bar{X}_{tce} + \gamma \mathbf{C} + k \mathbf{E} + \tau \mathbf{T} + \bar{u}_{tce} \quad (4)$$

em que a barra representa a média das variáveis observadas no momento t e pertencentes ao coorte c , com nível de escolaridade e , \mathbf{C} é um vetor de *dummies* de coorte, \mathbf{E} é um vetor de *dummies* de nível de escolaridade e \mathbf{T} é um vetor de *dummies* de tempo.

Esta equação pode ser estimada por Mínimos Quadrados Generalizados, atribuindo-se a cada par $(\bar{w}_{tce}, \bar{x}_{tce})$ um peso equivalente ao número de observações utilizadas para formar a média (ou seja, o número de observações em cada célula de coorte-tempo-escolaridade). É importante ressaltar que o procedimento de agrupamento por coorte-tempo-escolaridade só faz sentido se incluirmos as *dummies* de coorte, escolaridade e tempo na equação (4), pois o efeito dessas características tende a ser correlacionado com o efeito fixo. Se não incluirmos tais *dummies*, não eliminaremos os efeitos de coorte e escolaridade do componente de erro (\bar{u}_{tce}) , e a correlação entre \bar{x}_{tce} continuará existindo.

Conforme estabelecido em Blundell et alii (1998), uma forma alternativa de se estimar os parâmetros de interesse, e que produz resultados numericamente idênticos, é através da técnica de adição de resíduos. Para implementarmos essa técnica, estimamos por mínimos quadrados ordinários uma regressão (primeiro

estágio) de $x_{i(t)t}$ (formalidade) contra um conjunto de interações de coorte, tempo e educação, utilizando os dados individuais, e calculamos os resíduos desta regressão para cada indivíduo. Em seguida, fazemos uma regressão de $y_{i(t)t}$ (salário) em $x_{i(t)t}$ (formalidade), as dummies de coorte, ano e educação e os resíduos calculados no primeiro estágio, novamente com dados individuais. Esse método apresenta a vantagem de fornecer um teste direto de exogeneidade para a variável explicativa, que tem por base a estatística t calculada para o coeficiente dos resíduos.³ Este é o método que utilizaremos neste artigo.

É importante ressaltar que ao estimarmos a equação (4) estamos implicitamente fazendo a suposição de que o valor esperado do termo aleatório, condicional no regressor, pode ser representado de maneira adequada pelos efeitos coorte, tempo e educação, de forma aditiva:⁴

$$E[u_{i(t)}/X_{i(t)}, C, T, E] = C + T + E \quad (5)$$

Em outras palavras, nossa hipótese de identificação é que as diferenças não observadas entre os grupos, condicionais na formalidade, são invariantes no tempo, ou ainda, que as interações entre coorte, educação e tempo não estão presentes em (5). Além disso, o grau de formalidade deve variar de forma diferente entre os grupos ao longo do tempo, o que equivale à condição de posto para identificação:

$$E[\bar{x}_{tce}]^2 \neq 0 \quad (6)$$

3. Análise Descritiva dos Dados

Os dados utilizados foram retirados da Pesquisa Nacional por Amostra de Domicílios (PNAD) para o período 1981-2001. Não há informações para os anos de 1991 e 1994, nos quais não se realizou a PNAD. A amostra total é composta de 664.035 observações, sendo o ano de 1985 o que apresenta maior número de observações (55.087) e o ano de 1986 o que contém menos observações (30.506). Foram selecionados todos os homens com idade entre 24 e 57 anos,⁵ inclusive,

³Os desvios-padrões estimados dos resíduos precisam apenas ser corrigidos devido ao fato dos valores dessas variáveis utilizados no segundo estágio serem valores estimados.

⁴Ver Blundell et alii (1998).

⁵Ao impor a idade máxima de 57 anos estamos tentando excluir, ou minimizar, a presença de aposentados na amostra, dado que o interesse do trabalho é avaliar o salário de indivíduos em atividade.

cuja informação quanto à “posição na ocupação” era de empregado, trabalhador doméstico ou trabalhador na construção.⁶

As observações foram divididas em cinco coortes, de dez anos cada um e foram classificadas em dois níveis de educação: até sete anos de estudo, que corresponde ao ensino fundamental incompleto, e oito ou mais anos de estudo (pelo menos ensino fundamental completo), observados a cada ano. A tabela 1 apresenta algumas estatísticas descritivas, mostrando o período de nascimento de cada coorte, o número mínimo e médio de observações em cada célula, sua média e dispersão salarial.

Cabe notar que a maioria das células é composta por um grande número de elementos, o que revela a adequação do critério de agrupamento adotado a fim de se atingir as propriedades assintóticas definidas na seção anterior. De fato, não existe nenhuma célula com menos que 61 observações. Além disto, podemos notar o alto diferencial de salários associado ao nível educacional que chega a mais de 6 vezes entre os mais velhos. Podemos notar também que a dispersão salarial aumenta com a educação e com a idade.

Tabela 1
Características Gerais dos Dados Agrupados

Coorte	Nível de Educação	Tamanho Médio de Célula*	Tamanho Mínimo de Célula**	Média Salarial***	Dispersão Salarial****
1					
1924-33	Fundamental Incompleto	3705	253	339	392
	Fundamental Completo	926	61	2089	2230
2					
1934-43	Fundamental Incompleto	6601	246	347	403
	Fundamental Completo	2240	69	1860	1992
3					
1944-53	Fundamental Incompleto	8875	3141	334	362
	Fundamental Completo	4594	1485	1475	1824
4					
1954-63	Fundamental Incompleto	8306	5932	287	312
	Fundamental Completo	5845	4010	915	1078
5					
1964-73	Fundamental Incompleto	6831	912	249	307
	Fundamental Completo	5525	757	651	1088

(*) Média de observações em cada grupo de coorte/educação nos diferentes anos;

(**) Mínimo número de observações em cada grupo de coorte/educação;

(**) Valores em Reais de 1996; (****) Desvio-padrão.

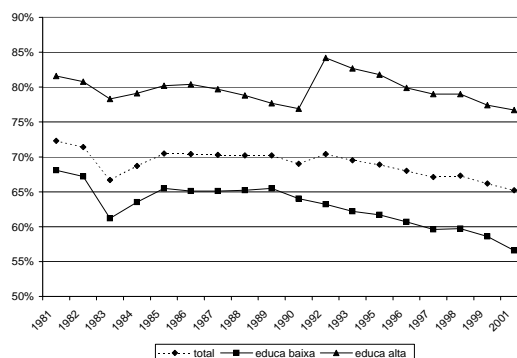
A variável dependente – salário real horário – foi construída a partir da in-

⁶Foram excluídos da análise os trabalhadores por conta própria.

formação relativa ao “rendimento mensal do trabalho principal” das PNADs.⁷ No que diz respeito à variável explicativa de interesse, adotou-se a posse de carteira de trabalho assinada para classificar os indivíduos como pertencentes ao mercado formal de trabalho. O grande problema com esta definição é que os funcionários públicos e militares eram considerados como trabalhadores sem carteira assinada até 1992, e não há como distingui-los dos demais trabalhadores até este ano. Desta forma, na seção empírica nós trabalharemos com duas amostras: 1981–90 e 1992–01.

A figura 1 ilustra o processo de informalização do mercado de trabalho brasileiro, mostrando que, em média, a porcentagem de trabalhadores com carteira assinada passou de 73% para 65% dos empregados homens com entre 24 e 57 anos de idade.⁸ Interessante notar que entre os menos educados o processo começou a partir de 1989, ocorrendo uma queda de cerca de 20 pontos percentuais até 2001. Para os trabalhadores mais educados, a queda parecia estar ocorrendo desde 1986, mas há uma elevação brusca na formalidade em 1992, que obscurece este processo. Esta elevação parece estar refletindo a mudança no critério de classificação dos funcionários públicos e militares, que tendem a ser mais educados, que até 1990 eram classificados como informais, deixando a amostra a partir de então.

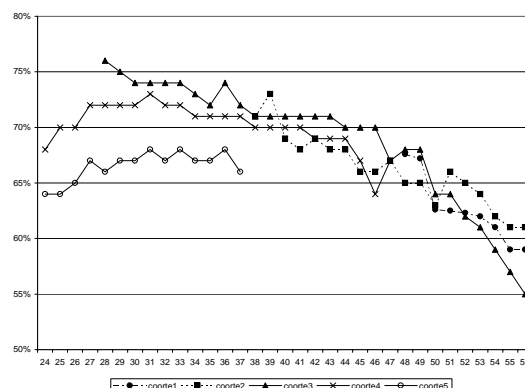
Figura 1
Porcentagem de trabalhadores com carteira assinada



⁷ Este valor foi deflacionado pelo índice INPC e, posteriormente, dividido por quatro (para se obter o rendimento semanal do trabalho) e dividido pelo número de “horas habitualmente trabalhadas por semana”. Como resultado, tem-se o salário real por hora de trabalho, para isolarmos o impacto da variação das horas trabalhadas sobre os salários.

⁸ É importante ressaltar que as figuras são a contrapartida empírica das distribuições não-condicionadas ao vetor de variáveis independentes, diferentemente das regressões abaixo, e devem portanto ser vistas com a devida cautela.

Figura 2
Trabalhadores no setor formal por coorte e idade

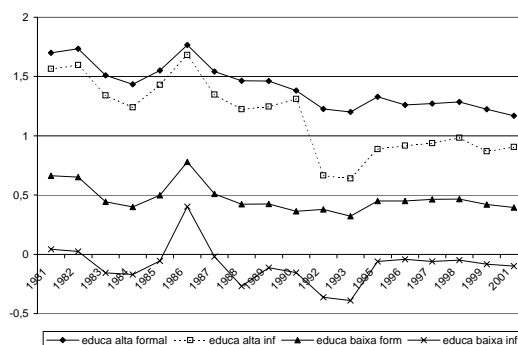


A figura 3 exibe a participação no setor formal para cada coorte e idade.⁹ Percebe-se claramente um efeito “ciclo de vida”, segundo o qual no começo da vida profissional os indivíduos tendem a trabalhar em empregos informais, no auge da vida produtiva têm maior espaço no setor formal e, no final da carreira voltam para o setor informal. O auge da formalidade ocorre cedo, por volta dos 30 anos. As diferenças no grau de formalidade entre gerações sucessivas observadas na mesma idade podem estar refletindo efeitos geracionais ou temporais. Interessante notar que não se observam diferenças significativas entre as gerações no que tange ao grau de formalização, a não ser para a geração mais nova (5), que parece estar encontrando sérias dificuldades de acesso ao mercado formal.

A figura 3 mostra que os maiores salários na nossa amostra estão com os trabalhadores de maior escolaridade empregados no setor formal. Logo em seguida vêm os de maior escolaridade empregados no setor informal, seguido pelos trabalhadores menos educados do setor formal e, por último, dos menos educados no setor informal. Vale ressaltar que o salário real médio dos trabalhadores vêm declinando no Brasil nos últimos 20 anos, especialmente para os trabalhadores mais escolarizados.

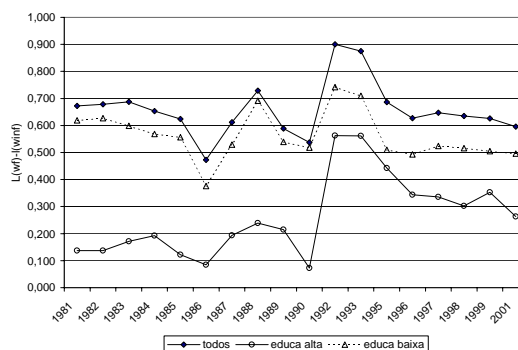
⁹ A partir dos dados agrupados por coorte, calculamos a idade média do coorte, da seguinte maneira: idade média = ano – ano médio de nascimento do coorte. Os anos médios de nascimento para cada coorte são, respectivamente: 1928, 1938, 1948, 1958 e 1958.

Figura 3
Log do salário médio por escolaridade e formalidade



A figura 4 mostra uma variável de interesse central para este estudo: a diferença de salários entre os setores formal e informal, para a amostra completa e entre os trabalhadores mais e menos educados separadamente. Até 1990 o diferencial de salários parece oscilar em torno de uma média. Em 1992 há um salto (dramatizado pelo fato de não haver informação para o ano de 1991) que coloca o diferencial em um novo patamar, tanto entre os mais como entre os menos educados. No caso dos mais educados, parte deste salto pode ser explicado pela saída dos funcionários públicos do grupo sem carteira, como descrito acima. Entretanto este salto ocorre também entre os trabalhadores menos escolarizados, o que sugere que este aumento pode não ser apenas o resultado de um problema de classificação. Desde então, a diferença salarial entre os trabalhadores dos dois setores vem declinando.

Figura 4
Diferença salarial entre trabalhadores formais e informais por escolaridade



4. Resultados

Os resultados das estimações são apresentados nas tabelas 2, 3 e 4.¹⁰ A tabela 2 utiliza a amostra completa. Na primeira coluna, apresentamos a correlação, condicionada nas dummies de ano, entre salários e formalidade, que mostra que estas variáveis são positivamente correlacionadas. A coluna (2) introduz a variável indicadora de escolaridade, que tem um efeito positivo sobre os salários e que reduz o coeficiente estimado da formalidade para 0,473. Na terceira coluna, introduzimos as dummies de coorte, que não afetam o coeficiente da formalidade, apesar de serem conjuntamente significantes.

Tabela 2
Amostra Completa

Log (salário)	(1)	(2)	(3)	(4)
Formal	0,658 (0,002)	0,473 (0,002)	0,470 (0,002)	-2,760 (0,034)
Educa Alta	–	1,042 (0,002)	1,077 (0,002)	1,559 (0,058)
Resíduo 1º Estágio	–	–	–	3,241 (0,034)
Dummies de Coorte	–	–	Sim	Sim
Dummies de Ano	Sim	Sim	Sim	Sim
Significância dos Instrumentos (p)	–	–	–	200,91 (0,000)
Número Observações	664.035	664.035	664.035	664.035

Nota: os erros padrões estão em parênteses.

Na quarta coluna, apresentamos o resultado central do artigo: quando introduzimos os resíduos da estimação de primeiro estágio para controlar pela endogeneidade da formalidade, o coeficiente muda de sinal, passando para $-2,760$. Isto significa que o trabalhador formal recebe, na verdade, um salário menor que o trabalhador do setor informal. Além disto, o coeficiente dos resíduos do primeiro estágio são altamente significantes, o que demonstra que a variável formalidade é endógena, ou seja, correlacionada com atributos individuais não observados. Interessante notar que a variável de educação continua atraindo um coeficiente positivo e estatisticamente diferente de zero.

¹⁰ A estimação das regressões foi feita usando o programa *STATA 5.0*. Em todas as regressões foi adotado o estimador robusto de erro padrão de White (1980).

Tabela 3
Década de 80

Log (salário)	(1)	(2)	(3)	(4)
Formal	0,634 (0,003)	0,457 (0,003)	0,458 (0,003)	-3,299 (0,046)
Educa Alta	–	1,144 (0,003)	1,185 (0,003)	1,642 (0,066)
Resíduo 1º Estágio	–	–	–	3,767 (0,047)
Dummies de Coorte	–	–	Sim	Sim
Dummies de Ano	Sim	Sim	Sim	Sim
Significância dos Instrumentos (p)	–	–	–	185,38 (0,000)
Número Observações	418.743	418.743	418.743	418.743

Nota: os erros padrões estão em parênteses.

Devido ao problema de classificação dos funcionários públicos, que eram classificados como sem carteira até 1990, é necessário refazer as estimações para as décadas de 80 e 90 separadamente. A tabela 4 apresenta os resultados para a década de 80 e mostra que o padrão dos coeficientes estimados não muda, apesar da grande redução do coeficiente associado à formalidade na coluna 4.

Tabela 4
Década de 90

Log (salário)	(1)	(2)	(3)	(4)
Formal	0,696 (0,004)	0,511 (0,003)	0,500 (0,003)	-1,610 (0,095)
Educa Alta	–	0,871 (0,004)	0,896 (0,003)	1,305 (0,019)
Resíduo 1º Estágio	–	–	–	2,113 (0,095)
Dummies de Coorte	–	–	Sim	Sim
Dummies de Ano	Sim	Sim	Sim	Sim
Significância dos Instrumentos (p)	–	–	–	211,55 (0,000)
Número Observações	245.292	245.292	245.292	245.292

Temos motivos para acreditar que os resultados mais confiáveis são os que se referem à década de 90, que exclui da amostra os funcionários públicos e militares. Na verdade, os resultados da tabela 4 confirmam as evidências anteriores, de que, quando controlamos pelo viés de auto-seleção dos trabalhadores no setor formal, eles na verdade recebem um salário inferior ao recebido pelos trabalhadores do setor informal. Novamente o coeficiente associado aos resíduos do primeiro estágio é estatisticamente diferente de zero, o que confirma a endogeneidade da posse de carteira assinada.

Este resultado contradiz o senso comum, segundo o qual a proteção legal gozada pelos trabalhadores formais lhes garantiria melhor remuneração. O que constatamos é que o fator escolaridade é mais importante na determinação do nível salarial que a formalidade ou não da relação de trabalho. O sinal negativo para a variável *formal* parece indicar que, condicional ao nível de escolaridade, os trabalhadores do setor informal recebem um adicional de salários para compensar benefícios (monetários ou não monetários) gozados pelos empregados no setor formal (adicional de férias, fundo de garantia, aviso prévio, contribuição patronal para previdência, etc.).

Essa constatação é de fundamental importância para formulação de política econômica. Ela nos diz que para elevar salários o fundamental não é regulamentar o mercado de trabalho ou coibir a informalização do emprego, mas sim elevar a qualidade educacional dos indivíduos.

5. Conclusões e Comentários Finais

Este trabalho estimou a equação de salários como função da formalidade, utilizando pseudo-painéis e uma metodologia econométrica relativamente recente para controlar pelo viés de auto-seleção na escolha do setor. O principal resultado do artigo é que, condicional ao nível de escolaridade, a remuneração do setor informal supera a do setor formal, indicando que a proteção legal gozada pelos trabalhadores do setor formal não lhes garante um maior nível salarial. Embora o setor formal garanta vantagens como férias, fundo de garantia, estabilidade no emprego e contribuição previdenciária patronal, tais vantagens parecem ser compensadas, no setor informal, por uma remuneração adicional.

A existência de viés de auto-seleção indica que, condicional ao nível de escolaridade, o diferencial de salários observados entre formais e informais decorre, na verdade, da melhor qualidade da força de trabalho empregada no setor formal. Isso significa que uma política eficiente de elevação dos salários reais passa pela melhor qualificação dos trabalhadores e não pela criação de leis que garantam benefícios trabalhistas.

Referências

Angrist, J. D. (1991). Grouped-data estimation and testing in simple labor-supply models. *Journal of Econometrics*, 47:243–266.

- Attanazio, O. P. & Browning, M. (1985). Consumption over the life cycle and over the business cycle. *The American Economic Review*, 85(5):1118–1137.
- Blundell, R., Duncan, A., & Meghir, C. (1998). Estimating labor supply responses using tax reforms. *Econometrica*, 66(4):827–861.
- Browning, M., Deaton, A., & Irish, M. (1985). A profitable approach to labor supply and commodity demands over the life-cycle. *Econometrica*, 53(3):503–543.
- Deaton, A. (1985). Panel data from time series of cross-sections. *Journal of Econometrics*, 30:109–126.
- IBGE (1995). Pesquisa nacional por amostra de domicílios – PNAD. Rio de Janeiro: Diretoria de Pesquisas.
- Meghir, C. & Whitehouse, E. (1996). The evolution of wages in the United Kingdom: Evidence from micro data. *Journal of labor Economics*, 14(1):1–25.
- Menezes-Filho, N. (2001). Microeconometria. In Lisboa, M. & Menezes-Filho, (editors) *Microeconomia e Sociedade*, 431–65.
- Moffit, R. (1993). Identification and estimation of dynamic models with a time series of repeated cross-sections. *Journal of Econometrics*, 59:99–123.
- White, H. (1980). A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica*, 48:817–838.