

RELATÓRIO INTERNO / Nº 02

COORDENADORIA REGIONAL / RJ

Segmentação no Mercado de Trabalho: A Carteira de Trabalho na Construção Civil

Ricardo Paes de Barros
José Guilherme Almeida Reis
José Santos Rodriguez

ABRIL DE 1991

Instituto de Pesquisa Econômica Aplicada

O Instituto de Pesquisa Econômica Aplicada - IPEA
é uma Fundação vinculada ao Ministério da Economia,
Fazenda e Planejamento

PRESIDENTE

Antonio Kandir

DIRETORA TÉCNICA

Heloiza Camargos Moreira

DIRETOR TÉCNICO ADJUNTO

Marcos Reginaldo Panariello

DIRETOR DE ADMINISTRAÇÃO E FINANÇAS

Renato Moreira

COORDENADOR DE DIFUSÃO TÉCNICA E INFORMAÇÕES

Liscio Fábio de Brasil Camargo

COORDENADOR DE POLÍTICA AGRÍCOLA

Benedito Rosa do Espírito Santo

COORDENADOR DE POLÍTICA INDUSTRIAL E TECNOLÓGICA

Luis Fernando Tironi

COORDENADOR DE POLÍTICA MACROECONÔMICA

Eduardo Felipe Ohana

COORDENADOR DE POLÍTICA SOCIAL

Luiz Carlos Eichenberg Silva

COORDENADOR REGIONAL DO RIO DE JANEIRO

Ricardo Varsano

Tiragem: 80 exemplares

DIVISÃO DE EDITORAÇÃO E DIVULGAÇÃO

Brasília:

SGAN Q. 908 - MÓDULO E - Cx. Postal 040013

CEP 70.312

COORDENAÇÃO REGIONAL DO RIO DE JANEIRO

Av. Presidente Antonio Carlos, 51 - 17º andar

CEP 20.020

SEGMENTAÇÃO NO MERCADO DE TRABALHO:
A CARTEIRA DE TRABALHO NA CONSTRUÇÃO CIVIL

Ricardo Paes de Barros^{*}
José Guilherme Almeida Reis^{**}
José Santos Rodriguez^{***}

- I. INTRODUÇÃO
- II. CONCEITOS BÁSICOS
- III. PARÂMETRO DE INTERESSE
- IV. INTERPRETAÇÃO ECONÔMICA
- V. IDENTIFICAÇÃO
- VI. HIPÓTESES OPERACIONAIS
- VII. SENSIBILIDADE AOS ATRIBUTOS OBSERVÁVEIS
- VIII. UNIVERSO DE ANÁLISE
- IX. ATRIBUTOS OBSERVÁVEIS
- X. RESULTADOS
- XI. CONCLUSÕES
- XII. REFERÊNCIAS BIBLIOGRÁFICAS

* Do IPEA-RIO.

** Do IBGE.

Pesquisador-Visitante do IPEA-RIO.

I. INTRODUÇÃO

Além de uma elevada concentração de renda, a economia brasileira caracteriza-se por apresentar um alto grau de diferenciação salarial, superior ao encontrado em economias com níveis semelhantes de desenvolvimento. Em princípio, esta diferenciação poderia decorrer exclusivamente da heterogeneidade dos trabalhadores com respeito a determinados atributos produtivos. Entretanto, quando se controla por diferenças naqueles atributos produtivos que são observáveis, permanece inexplicada uma parte expressiva da desigualdade de salários, sugerindo a possibilidade de que existam trabalhadores igualmente produtivos recebendo remunerações distintas.¹

Duas possibilidades para o funcionamento do mercado de trabalho podem ocorrer: primeiro, trabalhadores igualmente produtivos podem estar indiferentes em ocupar postos de trabalho com remunerações distintas, caso em que os diferenciais de salários simplesmente compensam diferenças em outras características associadas aos postos de trabalho; e, segundo, trabalhadores idênticos em postos de trabalho distintos podem obter níveis de utilidade diferenciados, caso em que se diz que o mercado de trabalho se encontra segmentado.

Entre as razões mais comumente alegadas para a segmentação estão aspectos institucionais, aí incluídos a ação dos sindicatos e os efeitos da legislação trabalhista. Neste contexto, um indicador de segmentação que tem sido particularmente analisado é a posse da carteira de trabalho. Supõe-se que nos postos com carteira de trabalho os trabalhadores têm acesso a uma série de vantagens pecuniárias e não pecuniárias que não obteriam em postos de trabalho sem este vínculo legal. É importante mencionar que os postos de trabalho sem carteira constituem-se em uma expressiva fonte de absorção de mão-de-obra, tendo mesmo crescido, nos últimos anos, em termos relativos [ver Sedlacek (1989)].

Barros e Varandas (1987) constataram que o grupo de trabalhadores com carteira recebe salários em média bem mais elevados do que aqueles sem carteira. Estes resultados foram obtidos sem a imposição de controles para características dos indivíduos ou dos postos de trabalho ocupados. Em um

¹A outra possibilidade é que os trabalhadores se diferenciem com respeito a atributos produtivos não observáveis.

trabalho posterior Barros, Pontes e Varandas (1988) estudaram a sensibilidade dos diferenciais de salários à imposição de controles tais como sexo, região de residência (urbano/rural), condição na família (chefe/não chefe) e condição de ocupação (ligado ou não à produção).

O objetivo deste trabalho é continuar explorando a hipótese de que trabalhadores igualmente produtivos recebam salários mais elevados em postos de trabalho com carteira do que em postos sem carteira, procurando ressaltar alguns aspectos metodológicos que são, a nosso juízo, fundamentais para a avaliação dos resultados empíricos. A parte empírica deste trabalho difere dos anteriormente citados, de um lado, por se restringir a um universo mais homogêneo de trabalhadores e, de outro, por ampliar o conjunto de variáveis de controle e, ao mesmo tempo, refinar alguns dos controles utilizados tais como região de residência e ocupação.

II. CONCEITOS BÁSICOS

Seja P uma dada população de trabalhadores assalariados e (P_0, P_1) uma partição disjunta de P , de tal forma que os membros de P_0 ocupam postos de trabalho sem carteira de trabalho assinada enquanto os membros de P_1 ocupam postos com carteira assinada. Para cada trabalhador $p \in P$, seja $I(p)$ o indicador do tipo de posto de trabalho ocupado por p , de tal forma que $I(p)=0$ se $p \in P_0$ e $I(p)=1$ se $p \in P_1$.

Para cada trabalhador $p \in P_0$, seja $S_0(p)$ seu salário no posto de trabalho sem carteira que correntemente ocupa. Paralelamente para cada trabalhador $p \in P_1$, seja $S_1(p)$ seu salário no posto de trabalho com carteira que correntemente ocupa.

Suponhamos que "exista" um mercado por postos de trabalho sem carteira no sentido de que para cada trabalhador $p \in P_1$ exista um único salário que esteja correntemente sendo recebido por todos os trabalhadores em postos de trabalho sem carteira com "qualidade" idêntica à do trabalhador p . Para cada trabalhador $p \in P_1$, denotemos este salário potencial por $S_0(p)$. Em princípio, tal salário, como todos os outros, é função do estado da economia, em particular da divisão dos postos de trabalho entre sem e com carteira, i.é., da partição (P_0, P_1) . Ao longo de todo este trabalho estaremos sempre supondo

ser fixa esta partição.

Por questões de simetria, poderíamos procurar definir, para cada $p \in P_0$, $S_1(p)$ como o salário que trabalhadores idênticos a p estariam recebendo em postos de trabalho com carteira. A definição deste salário potencial envolve sérias dificuldades. Em primeiro lugar, é possível que para um grupo de trabalhadores em P_0 inexista demanda por seus serviços em postos de trabalho com carteira e portanto não existam correntemente trabalhadores similares em postos de trabalho com carteira que possam ser usados como comparação. Em segundo lugar, mesmo que existam trabalhadores similares em postos com carteira nada garante que o salário destes trabalhadores deva ser único; é possível que firmas distintas estejam dispostas a pagar salários distintos a trabalhadores idênticos. Em outras palavras, é possível que haja segmentação entre postos de trabalho com carteira. Dada estas dificuldades, desenhamos nossa análise de tal forma que fosse desnecessário definir $S_1(p)$ para $p \in P_0$. Esta opção certamente impõe limitações a análise. Uma destas limitações, a que julgamos mais importante, é discutida na próxima seção.

III. PARÂMETRO DE INTERESSE

Estamos agora em condições de definir o parâmetro de interesse. Seja $\Delta_0(p) = \log(S_0(p))$ e $\Delta_1(p) = \log(S_1(p))$. Assim, para cada trabalhador $p \in P_1$, $\Delta_1(p) - \Delta_0(p)$ mede o hiato salarial por estar ocupando um posto de trabalho com carteira ao invés de um sem carteira. O objetivo deste estudo é estimar o valor médio deste hiato entre aqueles trabalhadores correntemente ocupando postos de trabalho com carteira de trabalho assinada. Este hiato médio será denotado por Δ e nos referiremos a ele como o impacto da carteira de trabalho.² Assim:

$$\Delta \equiv E[\Delta_1 - \Delta_0 | I=1] = E[\Delta_1 | I=1] - E[\Delta_0 | I=1]$$

É importante explicitar por que escolhemos este parâmetro e não outras definições alternativas como, por exemplo, o ganho salarial médio que um

²Note-se que este diferencial não é necessariamente o que o trabalhador p ganharia se o setor formal fosse abolido. Trata-se da diferença entre wage gain e wage gap, tal como assinalada por Lewis (1986).

trabalhador correntemente sem carteira teria caso passasse a ocupar um posto de trabalho com carteira. Ocorre que, como anteriormente mencionado, o processo de seleção a postos com carteira pode ser tal que o custo de oportunidade dos trabalhadores atualmente sem carteira é zero, isto é, existem barreiras à entrada no setor com carteira.

Outra alternativa seria definir o impacto da carteira de trabalho como o ganho salarial médio daqueles que correntemente ocupam postos sem carteira, caso houvesse uma política no sentido da formalização das relações de trabalho, entendida aqui como "carteirização". Ora, como os salários dependem da alocação específica de todos os demais trabalhadores, isto é, da partição (P_0, P_1) , tal política gerará um movimento de trabalhadores de P_0 para P_1 , o que tenderá a provocar mudanças nos salários relativos de trabalhadores em postos com carteira e sem carteira, tema que está fora do objetivo deste trabalho. Portanto, consideraremos fixa a partição (P_0, P_1) .

IV. INTERPRETAÇÃO ECONÔMICA

O objetivo último deste estudo é testar a hipótese de que trabalhadores idênticos recebam salários mais elevados em postos de trabalho com carteira do que em postos sem carteira. Pelo menos dois tipos de fatores, com consequências radicalmente distintas sobre o bem-estar dos trabalhadores, podem dar origem a estes diferenciais de salário: i) fatores compensatórios; e ii) segmentação do mercado de trabalho.

Note-se que a comparação é entre salários brutos, de forma que os salários reportados para trabalhadores com carteira de trabalho assinada incluem todos os descontos, como o imposto de renda e a contribuição a previdência social. O caráter compulsório destes descontos pode ter implicações sobre o significado dos diferenciais de salários. Por exemplo, se o trabalhador tivesse a opção, e o fizesse por contribuições em menor escala, em equilíbrio era de se esperar que trabalhadores idênticos recebessem salários mais elevados em postos de trabalho com carteira do que em postos sem carteira.

Por outro lado, caso o trabalhador valorize os benefícios da previdência social acima de sua contribuição então em equilíbrio deve-se esperar salários mais elevados associados aos postos de trabalho sem carteira. Este também

seria o caso se aos postos de trabalho com carteira estiver associada a maior quantidade de benefícios não pecuniários.

Em todos os casos acima as diferenças salariais não geram desigualdades sendo, ao contrário, compensatórias, de tal forma a garantir que trabalhadores idênticos em postos com ou sem carteira atinjam o mesmo nível de bem-estar.

Embora não tenhamos empiricamente como separar as diferenças salariais de natureza compensatória daquelas que geram desigualdades de bem-estar entre trabalhadores idênticos, é sem dúvida a busca de evidências sobre diferenças salariais não compensatórias que norteiam e motivam este trabalho.

V. IDENTIFICAÇÃO

A definição do parâmetro de interesse, Δ , envolve aspectos contra-factuais: o salário que cada trabalhador com carteira teria caso ocupasse um posto de trabalho sem carteira. Como aspectos contra-factuais não são observáveis, surge a questão de como expressar Δ em função de aspectos que sejam observáveis, de forma a torná-lo passível de estimação. Em outras palavras, cumpre discutir como Δ será identificado. Antes, no entanto, temos que explicitar que aspectos neste trabalho são considerados observáveis.

5.1. Aspectos Observáveis

Em função da natureza das informações a que este estudo tem acesso, considera-se como variáveis observáveis: i) o indicador se o trabalhador possui ou não carteira de trabalho assinada, I ; ii) o logaritmo do salário percebido, Δ ; e iii) um conjunto de atributos pessoais, X . Note-se que o salário percebido observado é dado por:

$$\Delta = \Delta_0 \cdot (1-I) + \Delta_1 \cdot I$$

Em outras palavras, se $I=0$ observa-se Δ_0 e se $I=1$ observa-se Δ_1 . Portanto, a informação disponível consiste do vetor (I, Δ, X) para uma amostra de trabalhadores em P . A partir desta amostra é possível construir estimadores consistentes para a distribuição conjunta de (I, Δ, X) . Assim, consideraremos que Δ é um parâmetro identificado caso possa ser univocamente obtido a partir

desta distribuição conjunta.

5.2 A Diferenciação Salarial Observada e o Impacto da Carteira de Trabalho

A diferença entre a média do logaritmo do salário corrente dos trabalhadores em postos de trabalho com carteira e a média do logaritmo do salário corrente daqueles em postos sem carteira, D , é dada por:

$$D = E[\Delta | I=1] - E[\Delta | I=0]$$

ou, equivalentemente, por

$$D = E[\Delta_1 | I=1] - E[\Delta_0 | I=0]$$

Uma vez que o impacto da carteira é definido como

$$\Delta = E[\Delta_1 | I=1] - E[\Delta_0 | I=1]$$

tem-se que

$$D - \Delta = E[\Delta_0 | I=1] - E[\Delta_0 | I=0] = \frac{1}{p \cdot (1-p)} \cdot \text{Cov}(\Delta_0, I)^3$$

onde $p = P[I=1]$

Assim, $\Delta = D$ se, e somente se, Δ_0 e I forem independentes⁴ (que denotaremos por $\Delta_0 \perp I$), ou seja, se a escolha de trabalhadores a postos de trabalho com carteira é feita de forma independente de seus respectivos salários potenciais em postos sem carteira. Neste caso, $\Delta = D$ e como D pode ser unívocamente obtido a partir da distribuição conjunta de (I, Δ) temos que Δ é identificado. No caso em que $\text{Cov}(\Delta_0, I) \neq 0$, teremos $\Delta \neq D$ e a identificação de Δ

³ $\text{Cov}(\Delta_0, I) = E[\Delta_0 \cdot I] - E[\Delta_0] \cdot E[I] = E[\Delta_0 | I=1] \cdot p - E[\Delta_0] \cdot p =$
 $= p \cdot (E[\Delta_0 | I=1] - [E[\Delta_0 | I=0] \cdot (1-p) + E[\Delta_0 | I=1] \cdot p]) =$
 $= p \cdot (1-p) \cdot (E[\Delta_0 | I=1] - E[\Delta_0 | I=0])$

⁴ Porque I é uma variável binária, Δ_0 e I são independentes se, e somente se, são não correlacionados.

deverá ser obtida de outra forma.

5.3 Seleção nos Atributos Observáveis

A fim de que Δ possa ser identificado é necessário algum tipo de informação sobre o processo de seleção. No caso acima contemplamos a possibilidade que a seleção a postos de trabalho com carteira é feita de forma independente dos salários potenciais em postos sem carteira. Nesta seção consideramos uma generalização desta hipótese de independência, a qual será a fonte de identificação para a análise empírica que se segue.

Suponhamos que os trabalhadores sejam divididos em grupos homogêneos com respeito a um conjunto de atributos observáveis, Z , e que dentro de cada grupo os trabalhadores sejam alocados a postos de trabalho com carteira independentemente dos seus salários potenciais em postos sem carteira. Neste caso, diz-se que houve seleção baseada apenas nos atributos observáveis Z . Formalmente, esta hipótese equivale a supor que as distribuições de I e Δ_0 condicionadas nos atributos observáveis, Z , são independentes. Portanto:

$$E[\Delta_0 | I=0, Z] = E[\Delta_0 | I=1, Z]$$

Logo,

$$\begin{aligned} \Delta &= E[\Delta_1 | I=1] - E[\Delta_0 | I=1] = \\ &= E[E[\Delta_1 | I=1, Z] - E[\Delta_0 | I=1, Z] | I=1] = \\ &= E[E[\Delta_1 | I=1, Z] - E[\Delta_0 | I=0, Z] | I=1] = \\ &= E[E[\Delta | I=1, Z] - E[\Delta | I=0, Z] | I=1] = E[D(Z) | I=1], \end{aligned}$$

onde $D(Z) = E[\Delta | I=1, Z] - E[\Delta | I=0, Z]$

Uma vez que $D(Z)$ e $E[D(Z) | I=1]$ podem ser univocamente obtidos a partir da distribuição conjunta de (Δ, I, Z) , tem-se que Δ é identificado caso Δ_0 e I estejam independentemente distribuídos condicionado a Z , e Z seja um subconjunto conhecido de X .

Note-se que, se i) $P\{I=1|Z\}=P\{I=1\}$, ou ii) $E[\Delta_0 | I=0, Z]=E[\Delta_0 | I=0]$ - isto é, se i) a probabilidade de seleção a postos com carteira é a mesma para todos grupos, caso em que I e Z estão independentemente distribuídos, ou ii) o

conjunto de atributos Z não afeta o nível salarial em postos de trabalho sem carteira - então:

$$E[E[\Delta_0 | I=0, Z] | I=1] = E[E[\Delta_0 | I=0, Z] | I=0] = E[\Delta_0 | I=0]$$

Portanto:

$$\begin{aligned} \Delta &= E[E[\Delta_1 | I=1, Z] - E[\Delta_0 | I=0, Z] | I=1] = \\ &E[E[\Delta_1 | I=1, Z] | I=1] - E[E[\Delta_0 | I=0, Z] | I=1] = \\ &E[\Delta_1 | I=1] - E[\Delta_0 | I=0] = D \end{aligned}$$

Em suma, se i) $P[I=1|Z]=P[I=1]$, ou ii) $E[\Delta_0 | I=0, Z]=E[\Delta_0 | I=0]$ for satisfeita, então a hipótese de seleção nos observáveis implica que Δ_0 e I estejam independentemente distribuídos e, portanto, que $\Delta=D$. Em geral, no entanto, $\Delta_0 \perp I$ não implica nem é implicado por $\Delta_0 \perp I|Z$. Mais geralmente, $\Delta_0 \perp I|Z_1$ não implica nem é implicado por $\Delta_0 \perp I|Z_2$, qualquer que seja a relação entre Z_1 e Z_2 . Por exemplo, nada se pode dizer caso $Z_1 \subset Z_2$ ou se $Z_2 \subset Z_1$. Assim, a idéia de seleção nos observáveis gera na verdade uma família de hipóteses não relacionadas. Cada membro da família corresponde a uma escolha de $Z \subset X$. É conveniente considerar a hipótese $\Delta_0 \perp I$ como o membro desta família quando $Z=\emptyset$.

O fato que $\Delta_0 \perp I|X$ não implica que $\Delta_0 \perp I|Z$, leva a que, para obter identificação, seja necessário supor não só que a seleção se baseia em atributos observáveis como também conhecer o conjunto de atributos efetivamente utilizados no processo de seleção.

VI. HIPÓTESES OPERACIONAIS

A fim de facilitar o processo de estimação duas hipóteses adicionais foram introduzidas. Nesta seção estas hipóteses são introduzidas e suas conseqüências investigadas.

Em primeiro lugar, supomos que as regressões são lineares,⁵ isto é, que existem vetores β e β^* tais que:

$$E[\Delta | I=1, Z] = Z' \cdot \beta^* \quad (A11)$$

e:

$$E[\Delta | I=0, Z] = Z' \cdot \beta \quad (A10)$$

Em segundo lugar, supomos que as duas regressões são paralelas, isto é, que:

$$Z' \cdot \beta^* = \alpha + Z' \cdot \beta \quad (A2)$$

Estas hipóteses simplificam sobremaneira o processo de estimação. Conforme já visto, se $\Delta_0 \perp I | Z$, então:

$$\Delta = E[D(Z) | I=1]$$

$$\text{onde } D(Z) = E[\Delta | I=1, Z] - E[\Delta | I=0, Z]$$

Sob as condições acima, $D(Z)=\alpha$ e, conseqüentemente, $\Delta=\alpha$. Logo, pode-se escrever que:

$$E[\Delta | I, Z] = I \cdot \Delta + Z' \cdot \beta$$

VII. SENSIBILIDADE AOS ATRIBUTOS OBSERVÁVEIS

Neste estudo partimos da hipótese de seleção baseada nos atributos observáveis e procuramos analisar e interpretar a sensibilidade dos resultados à escolha de Z . Isto será feito considerando uma seqüência, Z_0, \dots, Z_n , crescente, i.é., $Z_0 \subset Z_1 \dots \subset Z_n$, tal que $Z_0 = \emptyset$ e $Z_n = X$.

Suponhamos que as hipóteses operacionais (A1)-(A2) se verifiquem para todas as escolhas de $Z \in \{Z_0, \dots, Z_n\}$. Desta forma a única razão para sensibilidade nos resultados seria a não-equivalência das diversas hipóteses:

⁵Note-se que na regressão o conjunto de características pessoais é um vetor, ou, mais apropriadamente, uma matriz. Para marcar esta distinção, as letras em negrito denotam matrizes.

de seleção nos observáveis sendo utilizadas. Cumpre, no entanto, mostrar que é compatível termos (A1)-(A2) se verificando para todas as escolhas de $Z \in \{Z_0, \dots, Z_m\}$. Isto será levado a cabo via uma caracterização alternativa deste conjunto de hipóteses. Suponhamos que (A1)-(A2) se verifiquem para Z_1 . Mostraremos que, se além disso tivermos:

$$E[Z_1 | I, Z_{1-1}] = \gamma_1 \circ I + \Lambda_1 \circ Z_{1-1} \quad (A3_1)$$

então:

$$\begin{aligned} E[\Delta | I, Z_{1-1}] &= E[E[\Delta | I, Z_1] | I, Z_{1-1}] = \\ &= E[I \circ \Delta_1 + Z_1' \circ \beta_1 | I, Z_{1-1}] = \\ &= I \circ \Delta_1 + E[Z_1' | I, Z_{1-1}] \circ \beta_1 = \\ &= I \circ (\Delta_1 + \gamma_1' \beta_1) + Z_{1-1}' \circ \Lambda_1' \circ \beta_1 = \\ &= I \circ \Delta_{1-1} + Z_{1-1}' \circ \beta_{1-1} \end{aligned}$$

onde $\beta_{1-1} = \Lambda_1' \circ \beta_1$ e $\Delta_{1-1} = \Delta_1 + \gamma_1' \beta_1$

Definamos W_1 tal que $Z_1 = (Z_{1-1}, W_1)$. A sequência W_0, \dots, W_m estará associada a sequência $\gamma_0' \beta_0, \dots, \gamma_m' \beta_m$, onde $\gamma_1' \beta_1$ representará o impacto marginal da inclusão do conjunto de variáveis W_1 sobre a estimativa do impacto da carteira de trabalho. A fim de padronizar os resultados optamos por reportar os valores para $\bar{\gamma}_1 = \gamma_1' \beta_1 / \Delta_0$.

Uma forma alternativa de investigar a sensibilidade do impacto da carteira à inclusão dos diversos atributos baseia-se na seguinte expressão:

$$\begin{aligned} \Delta_0 &= E[\Delta | I=1] - E[\Delta | I=0] = \\ &= E[E[\Delta | I=1, Z_1] | I=1] - E[E[\Delta | I=0, Z_1] | I=0] = \\ &= \Delta_1 + E[Z_1' | I=1] \circ \beta_1 - E[Z_1' | I=0] \circ \beta_1 = \\ &= \Delta_1 + \{E[Z_1' | I=1] - E[Z_1' | I=0]\} \circ \beta_1 = \\ &= \Delta_1 + \sum_{k=1}^1 \{E[W_k' | I=1] - E[W_k' | I=0]\} \circ \beta_{1,k} \end{aligned}$$

onde $\beta_1' = (\beta_{1,1}', \dots, \beta_{1,1}')$

Note-se que $\beta_{1,k}$ é o vetor de coeficientes do conjunto de atributos W_k quando a regressão inclui o conjunto de atributos $Z_1 = (W_1, \dots, W_1)$, $1 \leq k$. Portanto:

$$(\Delta_0 - \Delta_1) / \Delta_0 = \sum_{k=1}^1 Q_{1,k}$$

onde $Q_{1,k} = \{E[W'_k | I=1] - E[W'_k | I=0]\} \cdot \beta_{1,k} / \Delta_0$

Assim, para cada $k=1, \dots, m$, $\{Q_{1,k} : i=k, \dots, m\}$ mede o impacto da inclusão do conjunto de atributos W_k sobre o impacto da carteira quando o conjunto de atributos incluídos na regressão é Z_1 , $i=k, \dots, m$.

VIII. UNIVERSO DE ANÁLISE

Caso todos os trabalhadores com idênticos atributos observáveis, X , possuam o mesmo salário, i.é., $\text{Var}[\alpha_0 | X] = 0$, teremos certamente que $\alpha_0 \perp I | X$. Portanto, quanto menor a heterogeneidade dos trabalhadores maiores serão as chances da hipótese de seleção nos observáveis se verificar. Com base nesta observação procuramos escolher uma população de trabalhadores a mais homogênea possível. Porque é de se esperar que trabalhadores de baixa qualificação sejam mais homogêneos e porque é entre trabalhadores de baixa renda que acreditamos ser mais importante estudar o impacto da carteira de trabalho, optamos por restringir nossa análise a trabalhadores de baixa qualificação em ocupações típicas da construção civil.

Especificamente, o universo considerado inclui apenas os empregados com e sem carteira que:

- a) são homens;
- b) residem na área urbana de alguma das nove regiões metropolitanas do país;
- c) são chefes de família;
- d) exercem ocupações típicas da construção civil;
- e) não trabalham na administração pública;
- f) não possuem ensino superior completo ou incompleto;
- g) possuem no máximo 70 anos; e
- h) possuem salário e horas de trabalho maiores que zero.

Com exceção de (e) e (h) as demais restrições têm por objetivo gerar uma população de trabalhadores de baixa qualificação bastante homogênea. A condição (e) tem por objetivo excluir os estatutários que apesar de não

possuírem carteira de trabalho possuem relações de trabalho bastante peculiares. Finalmente, a condição (h) é necessária uma vez que a variável dependente nas regressões é o logaritmo dos salários. O Anexo A mostra os efeitos da introdução dos filtros no tamanho da amostra.

A análise que segue foi realizada para três anos: 1978, 1981, 1985. O tamanho da amostra selecionada para cada um destes anos foi respectivamente de 1986, 2590 e 2422 observações. A estrutura da amostra para cada um dos anos segundo a posse ou não da carteira assinada encontra-se no Anexo B.

IX. ATRIBUTOS OBSERVÁVEIS

Quatro conjuntos de atributos foram consideradas: Idade- W_1 ; Educação- W_2 ; Região Metropolitana- W_3 ; e Ocupação- W_4 . W_1 inclui duas variáveis: a idade em anos e seu quadrado.⁶ W_2 inclui também duas variáveis: última série primária concluída e um indicador se o trabalhador cursou alguma série do nível secundário. W_3 consiste de oito dummies indicando as regiões metropolitanas.⁷ W_4 consiste de sete dummies indicando as seis principais ocupações e o agrupamento das demais.⁸

As Tabelas 1a a 1c apresentam as médias e desvios padrões destas variáveis para cada um dos anos da análise e para os segmentos com e sem carteira. Destas tabelas observa-se que: i) os salários médios encontram-se em torno de dois salários mínimos sendo de 16 a 27% mais elevados entre os trabalhadores com carteira do que entre os sem carteira; ii) a idade média é aproximadamente igual nos dois grupos de trabalhadores sendo ligeiramente inferior a 38 anos; iii) os dois grupos possuem também aproximadamente o mesmo

⁶ Como W é o primeiro grupo de variáveis a ser introduzido na regressão, devemos¹ incluir em W_1 também a variável $e'=(1,1,\dots,1)$ cujo respectivo coeficiente será a constante da regressão.

⁷ São Paulo foi a região omitida. Assim, os coeficientes das dummies medem os salários em relação a São Paulo.

⁸ As seis principais ocupações são: servente de pedreiro, pedreiro, carpinteiro, mestre da construção civil, pintor, e encanador. Sete outras ocupações foram agrupadas como "outras", a saber: armador de concreto, ladrilheiro, operador de máquina, estucador, calceteiro, calafate e vidraceiro. Em conjunto, as 13 ocupações representam cerca de 70% dos assalariados da construção civil.

Tabela 1a

Estatísticas Básicas

Ano: 1978

Variável	Empregados Com Carteira	Empregados Sem Carteira	Diferencial Médio
Log-Salário	0,674 (.542)	0,510 (.577)	0,164
Salário	2,32 (1.80)	1,98 (1.29)	0,34
<i>Idade-W₁:</i>			
Idade(anos)	37,5 (11.0)	34,9 (11.7)	2,4
Idade Quadrada((anos/10) ²)	15,1	13,5	1,6
<i>Educação-W₂:</i>			
Primária(anos)	2,52 (2.16)	2,65 (2.40)	-0,13
Secundária(%)	1,30	0,55	0,75
<i>Região Metropolitana-W₃:</i>			
Belém(%)	1,83	3,27	-1,44
Fortaleza(%)	3,44	6,59	-3,15
Recife(%)	4,93	9,31	-4,38
Salvador(%)	7,08	2,75	4,33
Belo Horizonte(%)	11,09	7,38	3,71
Rio de Janeiro(%)	29,76	25,87	3,89
São Paulo(%)	29,47	37,66	-8,19
Curitiba(%)	5,92	2,96	2,96
Porto Alegre(%)	6,50	4,20	2,30
<i>Ocupação-W₄:</i>			
Mestre(%)	8,68	9,29	-0,61
Pedreiro(%)	29,30	28,99	0,31
Servente(%)	16,22	37,72	-21,50
Pintor(%)	7,08	8,44	-1,36
Encanador(%)	7,96	3,02	4,94
Carpinteiro(%)	18,81	5,62	13,19
Outros(%)	11,94	6,92	5,02

Tabela 1b

Estatísticas Básicas

Ano: 1981

Variável	Empregados Com Carteira	Empregados Sem Carteira	Diferencial Médio
Log-Salário	0,405 (.533)	0,154 (.568)	0,251
Salário	1,76 (1.48)	1,36 (.83)	0,40
<i>Idade-W₁:</i>			
Idade(anos)	37,7 (10.9)	37,6 (12.3)	,1
Idade Quadrada(anos/10) ²	15,4	15,7	-,3
<i>Educação-W₂:</i>			
Primária(anos)	2,62 (2.18)	2,51 (2.30)	0,11
Secundária(x)	1,27	1,70	-,43
<i>Região Metropolitana-W₃:</i>			
Belém(x)	1,64	1,89	-0,25
Fortaleza(x)	5,01	8,98	-3,97
Recife(x)	6,33	6,59	-0,26
Salvador(x)	6,33	3,43	2,90
Belo Horizonte(x)	10,90	9,08	1,82
Rio de Janeiro(x)	24,20	29,58	-5,38
São Paulo(x)	29,08	32,39	-3,31
Curitiba(x)	7,31	2,69	4,62
Porto Alegre(x)	9,20	5,36	3,84
<i>Ocupação-W₄:</i>			
Mestre(x)	7,01	2,91	4,10
Pedreiro(x)	31,42	31,38	0,04
Servente(x)	19,29	40,67	-21,38
Pintor(x)	5,63	6,29	-0,66
Encanador(x)	8,05	5,08	2,97
Carpinteiro(x)	16,71	5,84	10,87
Outros(x)	11,87	7,83	4,04

Tabela 1c

Estatísticas Básicas

Ano: 1985

Variável	Empregados Com Carteira	Empregados Sem Carteira	Diferencial Médio
Log-Salário	0,732 (.539)	0,466 (.548)	0,266
Salário	2,44 (1.58)	1,87 (1.28)	0,57
<i>Idade-W₁ :</i>			
Idade(anos)	37,7 (10.6)	37,6 (11.7)	0,1
Idade Quadrada(anos/10) ²	15,3	15,5	-0,3
<i>Educação-W₂ :</i>			
Primária(anos)	2,98 (2.32)	2,55 (2.28)	0,43
Secundária(x)	2,32	1,93	0,39
<i>Região Metropolitana-W₃ :</i>			
Belém(x)	1,97	2,87	-0,90
Fortaleza(x)	4,82	10,68	-5,86
Recife(x)	5,65	6,49	-0,84
Salvador(x)	6,72	5,31	1,41
Belo Horizonte(x)	9,44	7,69	1,75
Rio de Janeiro(x)	26,61	29,83	-3,22
São Paulo(x)	32,40	29,28	3,12
Curitiba(x)	4,98	3,02	1,96
Porto Alegre(x)	7,40	4,83	2,57
<i>Ocupação-W₄ :</i>			
Mestre(x)	7,11	3,77	3,34
Pedreiro(x)	31,12	31,89	-0,77
Servente(x)	18,21	40,09	-21,88
Pintor(x)	6,28	7,39	-1,11
Encanador(x)	9,32	2,04	7,28
Carpinteiro(x)	15,86	7,12	8,74
Outros(x)	12,10	7,71	4,39

nível educacional médio - de 2,5 a 3,0 anos completos de ensino primário e menos de 2% de membros com algum ensino secundário; iv) cerca de 50% dos trabalhadores residem no Rio de Janeiro ou São Paulo; v) em Salvador, Belo Horizonte, Curitiba e Porto Alegre a proporção de trabalhadores com carteira é superior à média entre as nove regiões metropolitanas; e vi) quanto a distribuição ocupacional nota-se que servente de pedreiro é a única ocupação que se encontra super-representada no grupo sem carteira; as demais, com exceção de pedreiros, encontram-se super-representadas entre os trabalhadores com carteira.

A escolha da seqüência de atributos - idade > educação > região metropolitana > ocupação - seguiu o critério de incluir atributos na mesma ordem em que acreditamos em sua natureza exógena. É desnecessário enfatizar a natureza arbitrária deste critério; a sensibilidade dos resultados a outras combinações de atributos poderia, ou mesmo deveria, ser considerada.

Alguns comentários sobre a hipótese (A1) de linearidade são oportunos. Primeiro, que devido à especificação de W_1 a dependência na idade é quadrática e não linear. Segundo, que tanto região metropolitana como ocupação são variáveis categóricas e portanto a hipótese de linearidade é inócua. No caso de variáveis categóricas o que (A1) impõe é a inexistência de interações entre as variáveis.

X. RESULTADOS⁹

Na Tabela 2 apresentamos os resultados obtidos para $\{\bar{r}_1, \dots, \bar{r}_4\}$, enquanto que na Tabela 3 encontram-se os resultados para $\{Q_{1,k} : 1 \leq k \leq 4\}$. Ambos os resultados indicam que o impacto da carteira de trabalho é bem menos sensível à idade, educação e região metropolitana do que à ocupação.

A inclusão das três primeiras variáveis reduz o diferencial inicial em algo entre 2 e 7 pontos percentuais, o que significa uma redução de 8 a 25% do diferencial inicial, dependendo do ano considerado. Nada se pode dizer a respeito da importância relativa destas três variáveis: educação é a menos

⁹ Os resultados das regressões encontram-se no Anexo C.

Tabela 2

Sensibilidade do Impacto da Carteira de Trabalho
às Variáveis de Controle Empregadas
($\bar{\gamma}_1$)

Ano	Variáveis de Controle	i	Δ_1	$\bar{\gamma}_1$	$\sum_{k=1}^i \bar{\gamma}_k$
1978	0	.164
	$Z_1 = \text{Idade}(W_1)$	1	.154	6,2	6,2
	$Z_2 = Z_1 + \text{Educação}(W_2)$	2	.152	1,2	7,4
	$Z_3 = Z_2 + \text{Reg. Metr.}(W_3)$	3	.142	6,0	13,4
	$Z_4 = Z_3 + \text{Ocupação}(W_4)$	4	.048	57,0	70,4
1981	0	.251
	$Z_1 = \text{Idade}(W_1)$	1	.242	3,2	3,2
	$Z_2 = Z_1 + \text{Educação}(W_2)$	2	.234	3,2	6,4
	$Z_3 = Z_2 + \text{Reg. Metr.}(W_3)$	3	.230	1,8	8,0
	$Z_4 = Z_3 + \text{Ocupação}(W_4)$	4	.104	50,4	58,4
1985	0	.266
	$Z_1 = \text{Idade}(W_1)$	1	.251	5,6	5,6
	$Z_2 = Z_1 + \text{Educação}(W_2)$	2	.221	11,3	16,9
	$Z_3 = Z_2 + \text{Reg. Metr.}(W_3)$	3	.199	8,3	25,2
	$Z_4 = Z_3 + \text{Ocupação}(W_4)$	4	.074	47,0	72,2

Tabela 3

Sensibilidade do Impacto da Carteira de Trabalho
às Variáveis de Controle Empregadas
($Q_{1,k}$)

	Idade(W_1)	Educação(W_2)	Reg. Metr. (W_3)	Ocupação(W_4)
1978				
$Z_1 = \text{Idade}(W_1)$	6,2
$Z_2 = Z_1 + \text{Educação}(W_2)$	12,6	-5,1
$Z_3 = Z_2 + \text{Reg. Metr.}(W_3)$	12,8	-5,1	6,2
$Z_4 = Z_3 + \text{Ocupação}(W_4)$	3,5	-2,4	7,5	62,3
1981				
$Z_1 = \text{Idade}(W_1)$	3,2
$Z_2 = Z_1 + \text{Educação}(W_2)$	4,1	2,2
$Z_3 = Z_2 + \text{Reg. Metr.}(W_3)$	4,3	2,2	1,5
$Z_4 = Z_3 + \text{Ocupação}(W_4)$	1,9	1,0	2,2	53,2
1985				
$Z_1 = \text{Idade}(W_1)$	5,5
$Z_2 = Z_1 + \text{Educação}(W_2)$	6,2	10,5
$Z_3 = Z_2 + \text{Reg. Metr.}(W_3)$	6,4	11,0	7,6
$Z_4 = Z_3 + \text{Ocupação}(W_4)$	2,9	5,8	7,2	56,2

importante em 1978, região metropolitana o é em 1981, enquanto que idade ocupa esta posição em 1985.

Ocupação, por outro lado, tem um efeito substancial em todos os anos. Quando incluída a estimativa do impacto da carteira invariavelmente se reduz em mais da metade. Esta enorme sensibilidade ao controle por ocupação demonstra a extrema relevância de se determinar se ocupação deve ou não ser controlada quando comparamos os salários de empregados com e sem carteira.

Estes resultados, de certa forma, já eram previsíveis a partir das estatísticas básicas apresentadas nas Tabelas 1a a 1c. Ali se observa que as diferenças nas características dos empregados que ocupam postos com carteira assinada, com respeito à educação, idade e região, são relativamente menores quando comparadas com as diferenças na composição por ocupação.

XI. CONCLUSÕES

Os resultados obtidos neste trabalho revelam que, mesmo em um universo bastante restrito, se observa um hiato salarial que oscila entre 16 e 27% a favor dos empregados que ocupam postos de trabalho com carteira assinada.

Este hiato tende a diminuir na medida em que são introduzidas variáveis de controle, isto é, na medida em que o conjunto de variáveis observadas, Z , é reespecificado. Mesmo após controlar por educação, idade, região e ocupação, no entanto, permanece um diferencial entre 5 e 10%, a depender do ano considerado.

A ordem de grandeza do diferencial se aproxima do percentual representado pelas contribuições compulsórias descontadas sobre o salário bruto de trabalhadores de renda baixa. Em outras palavras, este diferencial pode simplesmente ser interpretado como o preço a ser pago pelos empregadores para que os trabalhadores aceitem ocupar postos com carteira assinada. Sem embargo, esta interpretação requer que se conheça a avaliação que os trabalhadores fazem de todos os aspectos pecuniários ou não relacionados aos postos com carteira.

De qualquer forma, ficou patente neste estudo a sensibilidade da

magnitude do parâmetro de interesse à especificação do conjunto de variáveis de controle. Em particular, as estimativas apresentaram grande sensibilidade à inclusão ou não da ocupação, demonstrando a relevância de se determinar se esta variável deve ser controlada quando se comparam salários de empregados com e sem carteira.

Em essência, a questão de controlar ou não por ocupação envolve saber se: a) a ocupação é uma característica dos indivíduos, caso em que a imposição deste controle é importante; ou b) ao contrário, a ocupação é determinada após a posse da carteira, em cujo caso controlar por ocupação leva a uma redução artificial do impacto da carteira.

XII. REFERÊNCIAS BIBLIOGRÁFICAS

- BARROS, R. & VARANDAS, S. A carteira de trabalho e as condições de trabalho e remuneração dos chefes de família no Brasil. Revista da ANPEC, Ano X, n.º 12, dez. 1987.
- BARROS, R.; PONTES, J. & VARANDAS, S. Diferenciais de salário: questões metodológicas e ilustrações. Revista de Econometria, Vol. VIII, n.º 2, jul./dez. 1988.
- LEWIS, H. G. Union relative wage effects. In: ASHENFELTER, O. & LAYARD, R. Handbook of labor economics. North-Holland, 1986, vol.2, chap. 20.
- SEDLACEK, G. L. Uma análise empírica do efeito de mudanças do nível de atividade econômica na proporção dos empregados com carteira assinada. Rio de Janeiro, 1987. Mimeo.

Anexo A: Seleção da Amostra

	1978	1981	1985
<i>Homens</i>	190.932	172.901	188.926
<i>Região Metropolitana</i>	61.909	66.902	72.146
<i>10 a 70 Anos</i>	60.723	65.516	70.561
<i>Chefes de Domicílio</i>	31.304	34.421	38.804
<i>Ocupados</i>	27.406	29.402	33.656
<i>Construção Civil</i>	3.236	4.087	3.924
<i>Empregador</i>	2.062	2.702	2.545
<i>Posse de Carteira Definida</i>	2.062	2.701	2.545
<i>Educ.def. Exclui. Superior</i>	2.053	2.692	2.544
<i>Exclui. Adm. Pública</i>	1.986	2.604	2.439
<i>Renda Positiva</i>	1.986	2.590	2.422

Anexo B: Estratificação da Amostra

	1978	1981	1985
<i>Total</i>	1.986	2.590	2.422
<i>Com Carteira Assinada</i>	1.644	2.153	1.762
<i>Sem Carteira Assinada</i>	342	437	660

Fonte: PNAD.

Anexo C: Resultados das Regressões

Tabela C.1

Variável	1978	1981	1985
<i>Idade-W₁</i> :			
Idade	.0189(.0072)	.0226(.0082)	.0505(.0067)
Idade Quadrada	-.0002(.0001)	-.0002(.0001)	-.0005(.0001)
\bar{R}^2	.0148	.0372	.0762

Tabela C.2

Variável	1978	1981	1985
<i>Idade-W₁</i> :			
Idade	.0303(.0088)	.0309(.0081)	.0582(.0085)
Idade Quadrada	-.0003(.0008)	-.0003(.0001)	-.0006(.0001)
<i>Educação-W₂</i> :			
Primária	.075 (.008)	.056 (.008)	.066 (.005)
Secundária	.212 (.113)	.222 (.093)	-.104 (.075)
\bar{R}^2	.108	.093	.144

Tabela C.3

Variável	1978	1981	1985
<i>Idade-W₁</i>			
Idade	.0029 (.0088)	.0316 (.0059)	.0598 (.0063)
Idade Quadrada	-.0003 (.0001)	-.0003 (.0001)	-.0006 (.0001)
<i>Educação-W₂</i>			
Primária	.075 (.005)	.058 (.005)	.069 (.005)
Secundária	.197 (.108)	.241 (.090)	-.072 (.073)
<i>Região Metropolitana-W₃</i>			
Belém	-.563 (.151)	-.554 (.140)	-.748 (.140)
Fortaleza	-.437 (.139)	-.690 (.123)	-.928 (.129)
Recife	-.504 (.137)	-.658 (.123)	-.867 (.130)
Salvador	-.209 (.135)	-.440 (.124)	-.703 (.130)
Belo Horizonte	-.186 (.134)	-.583 (.121)	-.928 (.129)
Rio de Janeiro	-.084 (.131)	-.465 (.118)	-.896 (.128)
São Paulo	.012 (.130)	-.261 (.118)	-.589 (.125)
Curitiba	-.271 (.137)	-.433 (.123)	-.746 (.133)
Porto Alegre	-.240 (.136)	-.443 (.122)	-.696 (.128)
\bar{R}^2	.6563	.4064	.6719

Tabela C.4

<i>Variável</i>	1978	1981	1985
<i>Idade-W₁</i>			
<i>Idade</i>	.0061 (.0057)	.0143 (.0051)	.0277 (.0055)
<i>Idade Quadrada</i>	-.0001 (.0001)	-.0001 (.0001)	-.0003 (.0001)
<i>Educação-W₂</i>			
<i>Primária</i>	.037 (.005)	.028 (.004)	.036 (.004)
<i>Secundária</i>	.115 (.092)	.130 (.078)	-.031 (.062)
<i>Região Metropolitana-W₃</i>			
<i>Belém</i>	-.511 (.068)	-.325 (.068)	-.106 (.060)
<i>Fortaleza</i>	-.483 (.050)	-.440 (.039)	-.338 (.038)
<i>Recife</i>	-.465 (.043)	-.346 (.037)	-.251 (.038)
<i>Salvador</i>	-.245 (.041)	-.215 (.039)	-.129 (.038)
<i>Belo Horizonte</i>	-.155 (.034)	-.274 (.031)	-.321 (.033)
<i>Rio de Janeiro</i>	-.106 (.025)	-.176 (.023)	-.261 (.023)
<i>São Paulo</i>	-----	-----	-----
<i>Curitiba</i>	-.214 (.044)	-.149 (.037)	-.130 (.044)
<i>Porto Alegre</i>	-.164 (.042)	-.128 (.034)	-.081 (.037)
<i>Ocupação-W₄</i>			
<i>Mestre</i>	1.244 (.120)	.816 (.111)	.770 (.119)
<i>Pedreiro</i>	.561 (.114)	.182 (.104)	.158 (.112)
<i>Servente</i>	.131 (.110)	-.256 (.102)	-.257 (.108)
<i>Pintor</i>	.529 (.117)	.225 (.108)	.080 (.116)
<i>Encanador</i>	.776 (.116)	.429 (.107)	.540 (.117)
<i>Carpinteiro</i>	.550 (.113)	.219 (.108)	.244 (.113)
<i>Outros</i>	.657 (.115)	.239 (.106)	.201 (.113)
<i>R²</i>	.755	.558	.758

TEXTOS PUBLICADOS

Série RELATÓRIO INTERNO
Coordenação Regional do Rio de Janeiro

1991

Nº 01 - "Medidas de Desigualdade", Lauro Ramos e Ricardo Paes de Barros, Abril
de 1991.