

Escola de **ECONOMIA** de São Paulo

**Textos para
Discussão**

215

Outubro
de 2009



IMPACTO DA ALFABETIZAÇÃO DE ADULTOS SOBRE SALÁRIO E EMPREGO

**MAÚNA SOARES DE BALDINI ROCHA
VLADIMIR PONCZEK**

Os artigos dos *Textos para Discussão da Escola de Economia de São Paulo da Fundação Getúlio Vargas* são de inteira responsabilidade dos autores e não refletem necessariamente a opinião da FGV-EESP. É permitida a reprodução total ou parcial dos artigos, desde que creditada a fonte.

Escola de Economia de São Paulo da Fundação Getúlio Vargas FGV-EESP
www.fgvsp.br/economia

Impacto da alfabetização de adultos sobre salário e emprego

Maúna Soares de Baldini Rocha

Vladimir Ponczek

Resumo

Neste trabalho apresentamos evidência sobre os efeitos da alfabetização de adultos na renda e probabilidade de empregos dos indivíduos. Os dados longitudinais disponibilizados pela Pesquisa Mensal de Emprego (PME) permitem controlar os efeitos fixos não-observáveis dos indivíduos gerando uma estimação mais robusta e consistente dos efeitos analisados. De modo geral, os resultados deste trabalho indicam retorno de 9,3% em rendimentos para indivíduos que se alfabetizam. Mostramos evidência que sugerem que esse efeito no salário se deve a acréscimos de produtividade e não por um aumento na formalização dos trabalhadores alfabetizados. Ademais, observamos que o efeito na renda é mais proeminente em certos grupos de indivíduos: mulheres, residentes de Salvador e pessoas entre 45 e 60 anos.

Abstract

In this paper, we document some evidence about the impacts of adult alphabetization on earnings and labor force participation. The longitudinal data available in the Monthly Employment Survey (PME) allow us to control for unobservable fixed effects, which generate more consistent and robust estimators. On average, our results indicate an increase of 9.3% in the worker's hourly wage after getting alphabetized. Our results also suggest that the effects on wage are due to productivity gains and not caused by an increase of the workers' formalization. Moreover, we observe that the wage gains are more notorious on certain groups of individuals: women, people living in Salvador and people aged between 45 and 60 years old.

Introdução

A acumulação de capital humano é objeto de diversos estudos em diversas áreas da economia, as quais procuram analisar, entre outras coisas, o que motiva indivíduos na busca de melhores níveis educacionais e quais os impactos da educação sobre o desenvolvimento de uma sociedade, tanto em termos de distribuição de renda e pobreza como de sua contribuição para o crescimento de um país. No que se refere à prosperidade, a produtividade e a habilidade dos trabalhadores consistem nos principais meios pelos quais a educação contribui na determinação do nível de renda *per capita*. Por essa razão, a educação vem sendo cada vez mais importante em considerações de políticas públicas, as quais devem desempenhar importantes esforços para resolver defasagens educacionais, que ocorrem em maior intensidade em países subdesenvolvidos e em processo de desenvolvimento, ponto em que surge a figura de programas educacionais e profissionalizantes.

Sob a ótica da microeconomia, estudam-se as relações no âmbito de desagregado de cada agente econômico, como firmas, indústrias, indivíduos, e os determinantes de seu comportamento. Considerando que as decisões de educação são tomadas por indivíduos ou num nível domiciliar, o interesse consiste em focar a análise no comportamento nesses menores níveis. Consumidores tomam suas decisões em função dos bens que desejam consumir e do quanto trabalhar, considerando que o retorno obtido no desempenho da força de trabalho pode variar conforme o estoque de habilidade individual corrente. A educação se insere nesse último aspecto como variável determinante, seja como um fator de expansão das habilidades e/ou como seu sinalizador, o que fornece a oportunidade para maior captura de renda proveniente do trabalho (DENNY, HARMON & REDMOND, 2000). Dessa forma, a decisão do quanto acumular de capital humano se insere na busca de bem-estar pelos indivíduos de uma sociedade.

Nesse aspecto, a oferta gratuita de educação por gestores públicos assume um papel de extrema importância na equalização das oportunidades de aprimoramento das habilidades individuais, que permitam a todos os indivíduos maior capacidade de geração de renda e alcance de maiores níveis de bem-estar, raízes de uma sociedade mais igualitária. No entanto, as altas taxas de analfabetismo no mundo, sobretudo em países subdesenvolvidos, evidenciam o acesso não universalizado e o tratamento desigual da educação através dos países. Situação que tende a se perpetuar se nenhum esforço for desempenhado para sanar esse problema, pois gerações atuais pouco instruídas serão futuros pais pouco instruídos e com menores esforços quanto à educação dos filhos, ainda mais sob as condições de vida pouco favoráveis decorrentes da baixa escolaridade.

Atualmente no mundo, aproximadamente 18% da população de indivíduos com mais de 15 anos de idade não sabe ler e nem escrever (UNESCO), sendo essa proporção de 11% entre os brasileiros adultos, segundo dados da Pesquisa Nacional por Amostras de Domicílios de 2005. Existe, portanto, um grande contingente de pessoas com necessidades educacionais básicas, já defasadas para ocupar atividades das mais simples, quanto mais para atividades mais complexas, que são cada vez mais exigidas nos dias atuais. Reconhecendo essa carência, a partir da Conferência Mundial Sobre a Educação Para Todos, que ocorreu em 1990 na Tailândia, iniciou-se um movimento mundial comprometido com a universalização da educação básica, com grande ênfase na erradicação de desigualdades educacionais entre gêneros, raças e classes sociais, e da taxa de analfabetismo entre jovens e adultos.

Nesse sentido, diversos países implementaram programas para sanar as defasagens educacionais entre jovens e adultos no intuito de oferecer-lhes a oportunidade de aprimorar habilidades, e com isso elevar a capacidade de geração de renda e o bem-estar. Existem alguns trabalhos em âmbito internacional que procuraram avaliar os programas de alfabetização de adultos em termos de impactos no bem-estar dos envolvidos. Considera-se, além dos efeitos sobre ganhos salariais, a existência de externalidades sobre os demais indivíduos da família e os efeitos de uma extensão da rede de convivência em decorrência de participação no programa (BLUNCH & PÖRTNER, 2005).

Partindo de dados de uma grande pesquisa de domicílios de Gana, Blunch & Pörtner (2005) examinaram o efeito de programas de alfabetização de adultos sobre o padrão de vida dos domicílios, medido pela despesa *per capita*. Como a decisão de participação em tais programas sofre de endogeneidade, optou-se pelo uso do método de variáveis instrumentais, tendo encontrado como fortes instrumentos para a variável de participação variáveis como a existência de programas de alfabetização na comunidade e a ausência de algum indivíduo da família correntemente matriculado em programas de alfabetização (sinais positivo e negativo, respectivamente). Na estimação do modelo principal, encontrou-se que em domicílios em que nenhum membro completou algum tipo formal de alfabetização o efeito da participação em tais programas sobre as despesas (padrão de vida) é significativo e de maior impacto relativamente àqueles domicílios em que já existia alguém formalmente alfabetizado. Sobre essa evidência, os autores concluem que tais programas de fato ajudam aqueles que possivelmente não tiveram acesso ao sistema educacional formal, ou o tiveram sob um menor nível de qualidade.

Um papel igualmente importante desses programas de alfabetização é evidenciado pelos resultados encontrados por Green & Riddell (2001) para o Canadá. O trabalho baseou-se na *International Adult Literacy Survey*¹ (IALS) de 1994 e teve como proposta analisar os efeitos sobre os ganhos individuais no mercado de trabalho dos componentes observados da habilidade dos indivíduos. Além de encontrar um efeito parcial expressivo e estatisticamente significativo da alfabetização (medida pelo *score* médio resultante das questões respondidas no âmbito da pesquisa) sobre os ganhos salariais, as estimações ainda sugerem que essa tem importante papel no ajustamento de imigrantes no mercado de trabalho, os quais, com base nas estimativas de ganhos anuais, recebem salários 35% menores relativamente aos canadenses natos².

Seguindo essa mesma linha, Chiswick (2000) faz um estudo sobre o ajustamento de imigrantes em Israel usando o Censo de 1972. A proposta inicial do trabalho consistia em verificar o que determina a proficiência na língua hebraica por homens judeus imigrantes e suas conseqüências sobre os ganhos no mercado de trabalho. O autor encontra que alfabetização hebraica e a fluência nessa língua (tê-la como língua primária ou única) permitem um ganho 20% maior relativamente aos imigrantes que não possuem tais habilidades. Ainda, controlando imigrantes por país de origem, o adicional da fluência em língua inglesa permite um ganho adicional de 15% no mercado de trabalho.

Blunch & Verner (2000), partindo dos dados de 1991/92 de uma pesquisa domiciliar de Gana, *Ghana Living Standards Survey (GLSS3)*, procuraram identificar os determinantes da alfabetização e dos rendimentos dos indivíduos do país africano. Encontram que idade e alfabetização estão inversamente correlacionadas³, no sentido que gerações mais recentes possuem maior taxa de alfabetização relativamente às mais antigas, o que os faz sugerir que houve maior preocupação por parte dos governantes de Gana no sentido de melhorar o acesso a sistema educacional e os programas de alfabetização nos anos mais recentes. Ainda, a probabilidade de uma mulher ser alfabetizada era de quase 30% menor em comparação com homens de Gana, assim como a de trabalhadores no setor informal. A análise dos determinantes dos ganhos realizada por esse mesmo estudo indica que apenas a partir de um determinado nível de escolaridade que se começa a observar ganhos marginais significativos em decorrência de anos de estudos adicionais, fornecendo evidências de que um patamar mínimo de alfabetização é condição necessária para que indivíduos tenham acesso ao mercado de trabalho.

No Brasil, o governo instituiu o Programa Brasil Alfabetizado, em consonância com a ênfase dada pelo Plano Plurianual de 2003-2007 na erradicação do analfabetismo. Entre as diversas medidas implementadas, em termos gerenciais, houve aporte de recursos adicionais à educação de

¹ Pesquisa iniciada por sete países com o intuito de obter dados de alfabetização comparáveis entre países. Para maiores informações, consultar <http://www.statcan.ca/english/Dli/Data/Ftp/ials.htm>.

² Esses dados são referentes a imigrantes que chegaram ao Canadá entre 1980 e 1994. Aqueles imigrados anteriormente a esse intervalo de tempo, 1965-1979, não apresentaram diferenças significantes em ganhos salariais relativamente aos canadenses natos.

³ A amostra é composta por indivíduos com no mínimo 15 anos.

jovens e adultos (EJA), com maior repasse a estados e municípios, reformas institucionais para garantir gestão eficiente e contínua das ações na área, e criação de sistema integrado de informações para acompanhamento e avaliação do programa (UNESCO). No que se refere ao funcionamento do programa, elevaram-se os recursos para financiamento de formação de alfabetizadores em 50%, e maiores incentivos financeiros são fornecidos para alfabetizadores se direcionarem a regiões mais pobres e de menor densidade educacional.

Apesar dos grandes esforços desempenhados pela equipe de gestão do Programa Brasil Alfabetizado na avaliação para monitoramento e aprimoramento das atividades, ainda não houve avaliação dos resultados em termos de retorno da alfabetização em variáveis socioeconômicas diversas. Concentrando esforços nesse sentido, este trabalho procura encontrar os retornos médios em salário obtidos por indivíduos que se alfabetizam, além do efeito sobre a probabilidade de emprego. Estimar médias educacionais entre indivíduos alfabetizados e analfabetos não fornece um valor devidamente estimado do verdadeiro retorno da alfabetização buscado, uma vez que não toma os devidos cuidados com as características não-observáveis de cada indivíduo: aqueles que optam pela alfabetização podem ter motivação e esforço adicionais que os tornam, por si só, mais produtivos, de modo que a comparação de salários entre alfabetizados e analfabetos não gera o retorno decorrente somente da educação. Nesse sentido, a disponibilidade de dados longitudinais, da Pesquisa Mensal de Emprego do IBGE, permite que se controlem as características não observáveis constantes no tempo a fim de obter estimativas mais precisas do retorno da alfabetização entre indivíduos adultos.

Na ausência de variáveis constantes no tempo que atuem como controle nas estimações de rendimento e emprego, a estratégia utilizada foi estimar, além do impacto médio da alfabetização para a amostra geral trabalhada – analfabetos adultos com até um ano de escolaridade –, o retorno da alfabetização em sub-amostras definidas por gênero, por região metropolitana, por faixa etária e por condição de formalidade. De modo geral, os resultados da *cross-section* indicam que alfabetizados ganham, em média, 27,5% mais que analfabetos; no entanto, os resultados da estimação com dados longitudinais considerando efeitos fixos sugerem que indivíduos que se alfabetizam auferem, em média, rendimentos 9,3% maiores relativamente àqueles que permanecem analfabetos, e, ao controlar por formalidade, nota-se que a formalização não possui impacto sobre o salário, o que sugere que os ganhos salariais decorrentes da alfabetização se efetivam pelo aumento de produtividade, e não pelo fato de pessoas se formalizarem quando se alfabetizam. Ainda, mulheres alfabetizadas apresentaram retorno médio superior à média encontrada, 16,6%, enquanto os resultados entre os homens não se mostram significativos. Uma explicação seria o fato de que homens com baixa escolaridade costumam se dedicar mais a trabalhos que exigem força, de modo que a alfabetização em si pode não implicar em grandes retornos, enquanto simples afazeres domésticos, atividades da ocupação predominante entre mulheres de baixa escolaridade, exigem habilidades básicas de leitura e escrita. Em termos de região metropolitana, indivíduos de Salvador, onde o salário-hora médio é menor, apresentam retornos significativos de 21,7% provenientes da alfabetização.

A seção seguinte descreve os dados utilizados e a forma como foram organizados. Em seguida parte-se para a análise descritiva das principais variáveis disponíveis e envolvidas no processo de estimação, para posteriormente apresentar os principais resultados encontrados e, então, terminar com as considerações finais.

1 Organização dos dados

Este estudo utiliza os microdados da Pesquisa Mensal de Emprego (PME), realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE), para o período de janeiro de 2002 a dezembro de 2006. A PME possui dados que atualmente abrangem as regiões metropolitanas de Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo e Porto Alegre. Essa pesquisa fornece dados conjunturais do mercado de trabalho desde 1980, e sofreu uma reformulação metodológica em 2001, de forma que os dados sob a nova metodologia estão disponíveis somente a partir de 2002.

A PME estrutura-se em dados em painel, mantendo informações de um mesmo indivíduo para mais de um período de tempo, sendo tais informações mensais. Seleciona-se um conjunto de domicílios para investigação dentro de uma determinada área de abrangência da região metropolitana, mas a unidade mínima de interesse consiste no indivíduo, de forma que entrevista-se um ou mais indivíduos do domicílio para obter informações de todos os moradores. O questionário aplicado abrange informações sociodemográficas de todos os moradores do domicílio e características de educação e de trabalho no caso de indivíduos com dez ou mais anos de idade.

Cada domicílio permanece na amostra por dezesseis meses, sendo entrevistado nos quatro primeiros e nos quatro últimos meses consecutivos, de forma a manter um intervalo de oito meses entre esses dois ciclos de entrevistas. Com esse desenho, mantêm-se informações com um ano de intervalo para cada mês entrevistado: realiza-se a quinta entrevista um ano após a primeira; a sexta entrevista ocorre um ano após a segunda, e assim por diante, conforme ilustra a figura 1 abaixo.

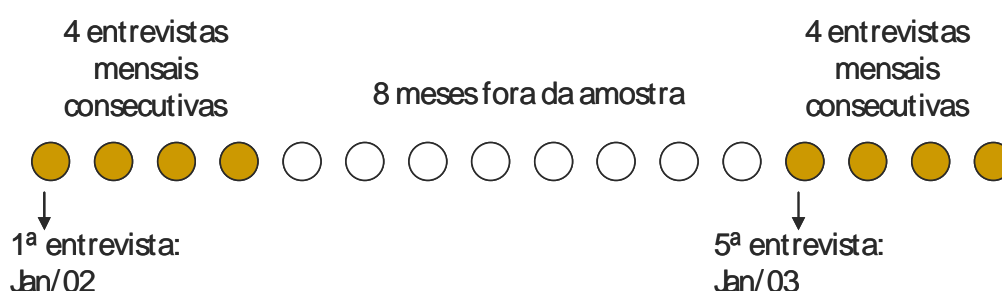


Figura 1 – Ilustração exemplo da seqüência de entrevistas da Pesquisa Mensal de Emprego

Cada mês da pesquisa utilizado no trabalho possui, em média, cerca de noventa e oito mil observações. No total, inicialmente, somam quase seis milhões de casos, cada qual representando uma entrevista, e abrangem informações de aproximadamente 1,6 milhões de indivíduos. Alguns dados relacionados ao banco de dados estão expostos na tabela 1.

Tabela 1 – Banco de dados: número de entrevistas, número de indivíduos, inconsistências da declaração de gênero e variáveis de entrevista própria e da PME (2002 – 2006)

Região Metropolitana	Entrevistas	Indivíduos	Inconsistentes		Variáveis de entrevistas		
	n°	n°	Indiv.	%	Própria	PME	Δ%
Recife	850.290	261.519	2.167	0,83%	832.084	850.290	-2,19%
Salvador	736.890	209.070	2.012	0,96%	722.926	736.890	-1,93%
Belo Horizonte	1.132.541	307.729	1.883	0,61%	1.109.988	1.132.541	-2,03%
Rio de Janeiro	1.063.087	254.403	2.455	0,97%	1.034.761	1.063.087	-2,74%
São Paulo	1.320.961	353.376	2.133	0,60%	1.290.590	1.320.961	-2,35%
Porto Alegre	852.362	234.468	1.213	0,52%	834.890	852.362	-2,09%
Total	5.956.131	1.620.565	11.863	0,73%	5.825.239	5.956.131	-2,25%

Elaborado a partir dos dados da Pesquisa Mensal de Emprego (2002 - 2006)

Para a construção da variável de identificação de cada indivíduo, além das variáveis de identificação que constam entre as variáveis originais da PME, utilizam-se a data de nascimento e o ano da primeira entrevista, ambas variáveis de elaboração própria criadas para garantir precisão no processo de identificação. Justifica-se a preocupação devido à repetição das variáveis de identificação da PME para diferentes domicílios e, conseqüentemente, para diferentes indivíduos, quando se juntam todos os anos num único arquivo. Isso ocorre porque um dado domicílio permanece na amostra por apenas dezesseis meses, de forma que o código até então atribuído a

esse, formado pela identificação da região metropolitana, setor censitário e sua ordem nesse setor⁴, pode identificar outro domicílio na amostra de domicílios selecionados no conjunto de entrevistas subsequente. Dessa forma, identificam-se os indivíduos pelas variáveis de controle do próprio banco de dados – número de controle, número de série e número de ordem – pela data de nascimento e pelo ano em que ocorre a primeira entrevista do indivíduo – criada com base nas variáveis de número da entrevista e da data em que essa ocorre.

O passo seguinte ao processo de identificação consistiu em procurar, dimensionar e eliminar algumas inconsistências que podem ser encontradas em banco de dados de grande porte, tanto devido ao grande contingente de observações administrado como pela potencial incompatibilidade entre as informações fornecidas pelo entrevistado e as informações reais sobre o morador cujas características são recolhidas, no caso do entrevistado não ser o indivíduo de interesse. Primeiramente, parte-se da variável que fornece o gênero de cada pessoa entrevistada e verifica-se se tal variável permanece constante para uma mesma pessoa ao longo de todas as entrevistas registradas. Houve variação para 11.863 indivíduos, ou seja, para menos de 1% da amostra (tabela 1).

Outro fato não esperado percebido no banco de dados está relacionada com a variável que fornece a ordem da entrevista no domicílio. Trata-se de entrevistas duplicadas para um mesmo indivíduo. Por exemplo, duas entrevistas classificadas como quarta entrevista na sequência, mas com datas diferentes, para um mesmo indivíduo. Esse fato foi notado quando da contagem do número de entrevistas (observações) para cada pessoa, tendo sido encontradas aquelas com até dezesseis registros de entrevista, o dobro do número máximo. Para contornar esse problema, criou-se uma variável de entrevista que pudesse identificar corretamente a ordem da entrevista de acordo com a data em que foi realizada. Essa variável, em muitos casos, não coincide com a variável equivalente da PME, uma vez que usa a primeira data de entrevista do indivíduo *registrada* no banco de dados como referência para ordenar as demais, ou seja, define-se esse primeiro registro de cada indivíduo como primeira entrevista, e, conforme a sequência das datas, numeram-se as demais.

A nova variável permite encontrar as entrevistas relevantes entre as informações de indivíduos que possuem mais de oito entrevistas. São alvos de interesse aquelas que seguem a sequência correta, no que diz respeito ao painel da PME, a partir do primeiro registro existente para o indivíduo. As quantidades de observações antes e depois desse passo estão expostas na tabela 1, assim como a perda relativa das informações disponíveis como consequência da redefinição das entrevistas.

Como já mencionado anteriormente nem todos os indivíduos possuem as oito entrevistas registradas e faz-se importante investigar a proporção de pessoas que possuem as informações para o segundo período das entrevistas, ou seja, a contrapartida de um ano. Para isso, identificam-se as entrevistas que apresentam um ano de intervalo uma da outra, tanto em relação à primeira como às demais entrevistas do primeiro bloco. Dado o processo de coleta de dados da PME, temos que os “pares” de entrevistas são: 1ª e 5ª, 2ª e 6ª, 3ª e 7ª, e 4ª e 8ª. Para identificar tais observações, basta dar o mesmo número para as entrevistas que formam um par. Ao todo, são quatro os pares de entrevistas. As entrevistas que não possuem informações de alfabetização em sua contrapartida de um ano (antes ou depois) são excluídas da amostra, uma vez que não possuem as informações necessárias.

Utiliza-se, em seguida, a variável que fornece a condição de alfabetização (“Sabe ler e escrever?”), cujos valores possíveis são “1”, para aqueles que declararam saber ler e escrever, e “2” caso a resposta para essa questão seja negativa. Vale explicitar alguns conceitos que serão utilizados a partir deste ponto. Definem-se dois tipos de transição que ocorrem entre as entrevistas: i) transição correta: mudança da condição de analfabeto para alfabetizado; e ii) transição incorreta: mudança da condição de alfabetizado para analfabeto. Cada uma dessas pode representar uma transição consistente ou inconsistente. As transições consistentes consistem nas transições cujas entrevistas anteriores condizem com a condição inicial de onde se transita e as posteriores condizem com a

⁴ Vide quadro X de descrição das variáveis.

condição final declarada: aquele que se diz analfabeto até uma dada entrevista e, a partir desta, se diz alfabetizado para todas as entrevistas seguintes caracteriza-se como um caso consistente. Indivíduos que apresentam transição consistente apresentam uma única transição entre as entrevistas: analfabetos para alfabetizados ou alfabetizados para analfabetos (essa no caso de transições incorretas). De modo oposto, as transições inconsistentes ocorrem quando há mais de uma mudança na variável de alfabetização para um mesmo indivíduo: até uma dada entrevista alguém se declara analfabeto, para então alfabetizar-se e posteriormente voltar à condição de analfabeto, por exemplo. A figura 2 abaixo facilita a compreensão dessas definições.

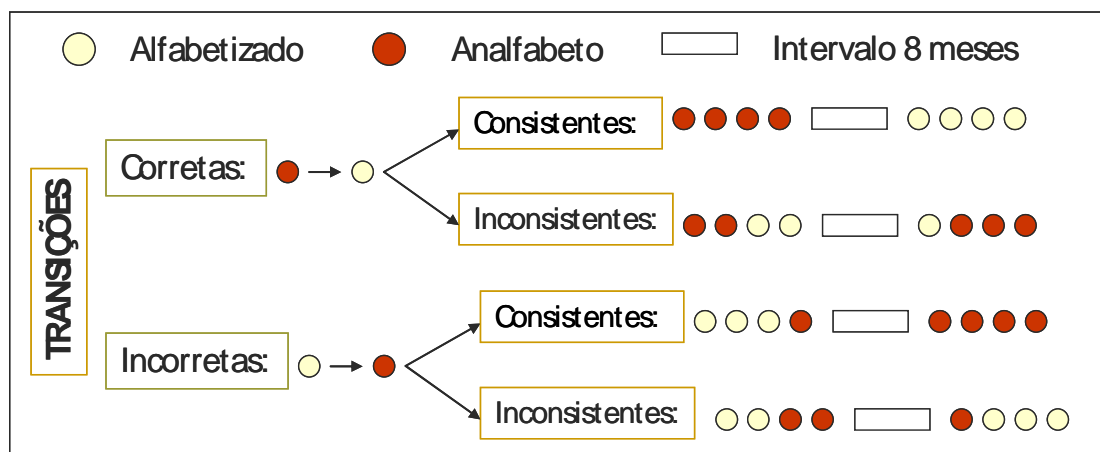


Figura 2 – Transições corretas, incorretas, consistentes e inconsistentes

Uma vez que existem transições incorretas e aquelas inconsistentes, resta verificar sua representatividade na amostra e comparar às transições corretas e às consistentes. Investiga-se, então, a proporção de transições ocorridas de cada tipo, condicional à situação inicial, ou seja, a proporção dos analfabetos que se alfabetizam e dos alfabetizados que se tornam analfabetos. As comparações entre esses tipos de transições se faz em três dimensões: i) transições entre quaisquer duas entrevistas consecutivas para cada indivíduo; ii) transições entre pares de entrevistas (intervalo de um ano); e iii) transições não redundantes. A figura 3 exibe, para um dado indivíduo (ID), cada uma dessas formas de transições calculadas para comparar corretas e incorretas e/ou consistentes e inconsistentes.

i) Transições totais			ii) Transições entre pares de entrevistas			iii) Transições consistentes		
Para cada indivíduo são marcadas as transições entre duas entrevistas seguidas quaisquer			Marcam-se apenas transições entre "pares" de entrevistas - 1 ano de intervalo			Consideram-se apenas entrevistas com indivíduos consistentes		
ID	Entrev	Trans	ID	Entrev	Trans	ID	Entrev	Trans
1	1	●	1	1	●	1	1	●
1	2	●	1	2	●	1	2	●
1	3	●	1	5	●	1	5	●
1	4	●	1	6	●	1	6	●

Figura 3 – Transições totais, transições entre pares de entrevistas e transições não redundantes

Verificam-se, primeiramente, as transições que podem ocorrer entre duas entrevistas consecutivas quaisquer para o mesmo indivíduo (quadro (i) da figura 3), as quais configuram as transições totais entre entrevistas. Encontram-se, aproximadamente, 15% de transições corretas, ou seja, dos registros que partem da situação inicial de analfabeto, cerca de 15% apresentam os registros subsequentes indicando que houve alfabetização, enquanto entre as entrevistas nas quais se declara alfabetizado 0,5% transitam para analfabetos em entrevista subsequente.

O segundo cálculo das transições faz em encontrar as mudanças na variável de analfabetismo entre os pares de entrevista, ou seja, para um intervalo de um ano (quadro (ii) da figura 3). De forma geral, entre as transições entre pares de entrevistas, como dado pela tabela 3, cerca de 12% correspondem a mudanças corretas na variável de analfabetismo, enquanto quase 0,4% correspondem às mudanças erradas, novamente relativamente à posição inicial de analfabeto e alfabetizado, respectivamente. No entanto, essa variável ainda agrega informações redundantes, uma vez que pode considerar mais de uma transição para cada indivíduo.

Tabela 2 - Número de transições corretas e incorretas - consistentes e não redundantes

Região Metropolitana	Corretas			Incorretas		
	Freq	Analfabetos	% cond	Freq	Alfabetizados	% cond
Recife	386	9.730	3,97	307	163.862	0,19
Salvador	230	4.618	4,98	194	136.084	0,14
Belo Horizonte	308	9.675	3,18	239	287.013	0,08
Rio de Janeiro	224	13.658	1,64	215	360.706	0,06
São Paulo	361	9.576	3,77	280	343.354	0,08
Porto Alegre	229	5.466	4,19	149	232.377	0,06
Total	1.738	52.723	3,30	1.384	1.523.396	0,09

Elaboração própria

A redundância ocorre pois a transição entre pares de entrevistas, quando consistentes, resulta na indicação de transição entre mais de um par de entrevista, como explicitado anteriormente, o que pode ser conferido na figura 3. Dessa forma, outras operações são necessárias para codificar apenas as transições não redundantes, e essas operações, da forma como elaboradas, ainda desconsideraram as transições corretas que ocorrem para um indivíduo que também apresentou transições incorretas. Ou seja, consideram-se, com essa variável, apenas os indivíduos com informações consistentes. A tabela 2 expõe a quantidade relativa dessas transições: as corretas consistentes somam a proporção de 3,3% daquelas em que se declara inicialmente analfabeto, enquanto as incorretas consistentes assumem a proporção de 0,09% das transições para as quais o estado inicial é alfabetizado.

Nota-se, portanto, que as transições corretas, para qualquer um dos conjuntos de transições investigadas, representam mais de trinta vezes a proporção das transições incorretas. Esse fato permite depositar um mínimo de confiança nos dados utilizados.

Um último ponto importante quanto à organização do banco de dados geral consiste no processo de deflação dos rendimentos individuais. Utilizou-se, para isso, o Índice Nacional de Preços ao Consumidor (INPC) específico de cada região metropolitana. As taxas de inflação foram recalculadas de tal forma que se adequasse à PME no sentido de captar o índice adequadamente para as informações de salário dessa pesquisa que, pelos questionários apresentados, registra informações de rendimentos recebidos no início de cada mês. Assim, recalcula-se para obter um índice centrado no dia primeiro de cada mês e utiliza-se o índice do mês subsequente para captar a variação dos preços ao longo do mês considerado⁵.

Os procedimentos até então apresentados, com respeito à organização do banco de dados, foram realizados para cada um dos bancos de dados agregados por região metropolitana, os quais iniciam com todas as informações de janeiro de 2002 a dezembro de 2006. Cada um dos conjuntos

⁵ Para obter informações mais detalhadas, consultar Corseuil e Foguel (2002).

de dados sofre um enxugamento ao longo desse processo, eliminando-se indivíduos com informações inconsistentes na declaração de gênero, indivíduos que não possuem nenhum par completo de entrevistas que possam ser comparadas (intervalo de um ano), indivíduos inconsistentes, além das entrevistas excedentes para indivíduos transitados e não transitados (uma vez que precisamos de apenas um par de entrevistas para comparação). Dessa forma, após a limpeza do banco de dados, as informações restantes somam 5.192 entrevistas correspondentes a 2.596 indivíduos, ou seja, exatamente duas entrevistas para cada indivíduo. A evolução do banco de dados ao longo de seu processo de organização pode ser observada na tabela 3, a qual a expõe as observações para a sub-amostra de transitados e não-transitados em termos de indivíduos, não de entrevistas, e cuja formação, juntamente com a formação das demais sub-amostras utilizadas, estão explicadas abaixo.

Tabela 3 – Número de entrevistas: evolução da amostra para criação da sub-amostra de transitados e não-transitados com até um ano de escolaridade

	Observações	% em relação a 1
1. N° inicial de indivíduos	1.620.565	100,00%
- indivíduos inconsistentes em gênero	11.863	0,73%
2. N° indivíduos consistentes	1.608.702	99,27%
- indivíduos com entrevistas sem contrapartida de 1 ano	1.322.983	81,64%
3. N° indivíduos consistentes e com contrapartida de 1 ano	285.719	17,63%
- indivíduos alfabetizados na primeira entrevista	235.702	14,54%
- indivíduos sem informações de alfabetização	41.512	2,56%
4. N° indivíduos analfabetos na primeira entrevista	8.505	0,52%
- indivíduos com mais de 1 ano de estudo	982	0,06%
5. N° indivíduos analfabetos na primeira entrevista e com até 1 ano de estudo	7.523	0,46%
- indivíduos fora da faixa de idade dos 25-60 anos na primeira entrevista	4.032	0,25%
6. N° indivíduos analfabetos com até 1 ano de estudo e na faixa dos 25-60 anos de idade	3.491	0,22%
- indivíduos com transição inconsistente	895	0,06%
7. N° de indivíduos da amostra final - transitados x não transitados - 1 ano de escolaridade	2.596	0,16%
8. com transição	666	0,04%

Diferentes sub-amostras são montadas para análise descritiva e estimações: i) alfabetizados *versus* analfabetos – primeira entrevista; ii) alfabetizados *versus* analfabetos – primeira entrevista e até um ano de escolaridade; iii) transitados *versus* não transitados – até um ano de escolaridade (tabelas 3).

Antes da criação de sub-amostras, foram criadas as variáveis julgadas necessárias para compor as bases de dados para estimação dos resultados, bases tais compostas pelas sub-amostras. Criaram-se *dummies* de gênero (homem=1), de raça (brancos=1), de setores de ocupação, de posição na ocupação (empregado, empregador, conta própria ou trabalhador não remunerado de membro da unidade domiciliar), de formalização, de emprego, de condição de alfabetização, e uma indicando o curso de alfabetização de adulto como curso mais elevado freqüentado pelo indivíduo. Além das *dummies*, criou-se uma variável que fornece a faixa etária na qual se encontra cada indivíduo, definidas para os seguintes intervalos: 25 a 35 anos, 36 a 45 anos e 46 a 60 anos. As informações utilizadas se restringem àquelas de indivíduos compreendidos nesse intervalo de idade (25 a 60 anos), uma vez que o interesse do trabalho jaz em encontrar o impacto da alfabetização de adultos sobre renda e emprego. Com base em tais faixas, criam-se *dummies* de idade. Ainda, definem-se faixas de horas trabalhadas por semana – até 35 horas, mais de 35 até 40 horas, mais de

40 até 45 horas, e mais de 45 horas -, além de calcular o salário-hora de cada pessoa, tanto do trabalho principal como dos demais trabalhos, com base no salário real encontrado ao deflacionar as informações de rendimentos nominais pelo INPC. O salário-hora foi calculado com base nas horas trabalhadas por mês – horas trabalhadas por semana multiplicadas por quatro.

Cada uma das sub-amostras geradas, como já mencionado, restringem-se a indivíduos entre os 25 e os 60 anos de idade na primeira entrevista. A primeira sub-amostra criada limita-se, ainda, a informações da primeira entrevista registrada. Parte-se dessa sub-amostra para análise descritiva em que se comparam analfabetos e alfabetizados.

O segundo conjunto de dados considerados origina-se dessa sub-amostra, mas limita-se a indivíduos com até um ano de escolaridade, de modo a comparar indivíduos relativamente mais homogêneos entre analfabetos e alfabetizados. A partir desses dados realiza-se uma nova análise descritiva e estima-se o impacto de alfabetização em salário e emprego (*cross-section*) controlando pelas variáveis sócio-econômicas disponíveis, por algumas variáveis relacionadas ao trabalho e por *dummies* de ano e região.

Finalmente, o conjunto de dados de interesse central abrange indivíduos analfabetos na primeira entrevista. Essa sub-amostra envolve aqueles que em entrevistas subseqüentes transitaram para alfabetizados e aqueles que permaneceram analfabetos. Desses últimos mantém-se apenas o primeiro par de entrevistas registrado, enquanto dos demais mantém-se apenas aqueles com transições consistente, e desses, apenas o primeiro par de entrevista em que se nota a transição. Enfim, parte-se dessa sub-amostra para descrever as características dos transitados e não-transitados, em termos de gênero, raça, salário médio, ocupação, formalidade, entre outros, além de realizar as estimações chaves do trabalho, que consistem na estimação do impacto da transição sobre salário e emprego, sob estrutura de painel, tanto considerando sob efeitos aleatórios como sob efeitos fixos.

2 Análise descritiva

2.1 Analfabetos versus alfabetizados: primeira entrevista

Os dados descritivos das amostras de alfabetizados e analfabetos na primeira entrevista constam na tabela 4. Considerando todos os indivíduos alfabetizados e analfabetos sob a faixa de idade considerada, encontra-se taxa de alfabetização de 97%. Os homens possuem maior proporção entre os alfabetizados e entre os analfabetos, proporções semelhantes à participação desse gênero na amostra, sendo a última pouco maior que a proporção de homens na amostra total (quase 58% de homens analfabetos contra 55% de homens na amostra). Ao considerar os indivíduos com até um ano de estudo, a proporção de analfabetos homens praticamente iguala a participação desses na amostra e, uma vez que se excluem indivíduos alfabetizados (com mais de um ano de estudo), a taxa de alfabetização se reduz para 27,68%.

No que se refere à raça, pouco mais de 95% dos indivíduos que se declaram não-brancos são alfabetizados, número inferior à taxa de alfabetização total, e que se reduz pra 25,6% quando calculada entre os indivíduos com até um ano de estudo. Entre os indivíduos que se declaram brancos, esses números são de 98,23% e 31,25%, respectivamente, valores sempre acima da taxa não condicional de alfabetização, refletindo o fato há muito tempo estabelecido de que os indivíduos de raça não-branca, em média, apresentam defasagem educacional relativamente aos indivíduos de raça branca, embora essa tendência tenha se reduzido ao longo do tempo.

Entre os alfabetizados, cerca de 90% dos indivíduos moram em domicílios com até cinco pessoas, enquanto 78,67% dos indivíduos analfabetos encontram-se nessa situação. Como à medida que se eleva o nível educacional considerado maior tende a ser a ocorrência de planejamento familiar, e, portanto, menor tende a ser o número de pessoas que moram num mesmo domicílio, ao se restringir a amostra àqueles com até um ano de estudo a proporção de indivíduos alfabetizados que moram com mais de cinco pessoas eleva-se de 9,78% para 15,59%. A diferença entre alfabetizados e analfabetos no que se refere ao convívio com indivíduos com menos de dez anos de idade apresenta-se muito menor: 96,6% dos alfabetizados convivem com até duas crianças do

domicílio contra 93% dos analfabetos, sendo essa proporção de pouco mais de 95% para os alfabetizados com até um ano de escolaridade.

Os alfabetizados distribuem-se quase igualmente entre as três faixas de idade consideradas: 25 a 34 anos, de 35 a 44 anos e de 45 a 60 anos. Mas os analfabetos estão em sua grande maioria na última faixa de idade, somando 59,81%, enquanto apenas 15,79% dos analfabetos possuem entre 25 e 34 anos de idade. Ao observar a proporção internamente a cada faixa etária, temos que 98,41% dos indivíduos na primeira faixa são alfabetizados, enquanto essa proporção é de 97,73% para a segunda faixa e de 95,12% entre os mais velhos considerados, ou seja, houve, ao longo das gerações, maior acesso aos serviços de educação básica.

Tabela 4 – Características da amostra de analfabetos e alfabetizados na primeira entrevista

	Primeira entrevista - todos		Primeira entrevista - até um ano de estudo	
	Alfabetizados	Analfabetos	Alfabetizados	Analfabetos
	0,97	0,03	0,28	0,72
Homens	0,55	0,58	0,57	0,58
Branca	0,57	0,34	0,42	0,35
Até 5 moradores	0,90	0,79	0,84	0,79
Mais de 5 moradores	0,10	0,21	0,16	0,21
Até 2 crianças moradoras ¹	0,97	0,93	0,95	0,94
Mais de 2 crianças moradoras	0,03	0,07	0,05	0,06
De 25 a 35 anos	0,31	0,16	0,21	0,15
De 36 a 45 anos	0,33	0,24	0,27	0,23
De 46 a 60 anos	0,36	0,60	0,53	0,62
Ocupados	0,67	0,42	0,53	0,42
Desocupados	0,06	0,05	0,05	0,05
Não Economicamente ativas	0,27	0,53	0,42	0,53
Privado ²	0,77	0,92	0,88	0,91
Público	0,23	0,08	0,12	0,09
Formal ³	0,81	0,67	0,72	0,67
Formal - setor privado	0,79	0,66	0,70	0,65
Menos que 35 horas	0,12	0,09	0,10	0,08
De 35 a 40 horas	0,25	0,12	0,17	0,13
De 40 a 45 horas	0,09	0,05	0,06	0,05
Mais de 45 horas	0,54	0,74	0,67	0,74
Salário-hora do trabalho principal ³	6,68	2,01	2,78	2,00

1. Indivíduos com menos de 10 anos de idade

2. Representativa apenas de um total de aproximadamente 16% dos analfabetos e de 42% dos alfabetizados da primeira amostra, enquanto de um total de cerca de 17% e 24% de analfabetos e alfabetizados, respectivamente, da segunda amostra.

3. Representativa apenas de um total de aproximadamente 40% dos analfabetos e de 62% dos alfabetizados da primeira amostra, enquanto de um total de cerca de 40% e 48% de analfabetos e alfabetizados, respectivamente, da segunda amostra.

No que se refere à condição de ocupação, os analfabetos possuem menor proporção de ocupados⁶ relativamente aos alfabetizados, 42,5% contra 66,95%, respectivamente, mas essa diferença é menor entre aqueles com até um ano de estudo, 52,5% contra 42,36%. Os empregos podem ser classificados pelo setor (público ou privado) e pela formalização (posse de carteira de trabalho assinada). No entanto, essas variáveis, de forma geral, possuem resposta para pouco mais de 40% dos analfabetos (na primeira entrevista), enquanto cerca de 65% dos alfabetizados

⁶ São classificadas como ocupadas na semana de referência as pessoas que exerceram trabalho, remunerado ou sem remuneração, durante pelo menos uma hora completa na semana de referência, ou que tinham trabalho remunerado do qual estavam temporariamente afastadas nessa semana.

respondem às perguntas que geram tais variáveis. Justamente por uma proporção até maior de não respostas na grande maioria das variáveis da PME referentes a características mais detalhadas de trabalho e de escolaridade o uso de tais variáveis faz-se inviável, tendo sido optado para este trabalho apenas as características então comentadas.

No que se refere ao setor de emprego, pouco mais de 90% dos analfabetos concentram-se entre os trabalhadores privados, enquanto 77,35% dos alfabetizados da amostra estão nesse setor. Isso evidencia maiores dificuldades inerentes ao processo de seleção de candidatos para ocupar empregos públicos, para o qual escolaridade atua como importante entrave. Corroborar-se esse fato ao notar a aproximação da participação de alfabetizados e analfabetos no emprego público entre aqueles com até um ano de escolaridade: 11,56% contra 8,84%. Restringindo-se ao setor privado, a taxa de formalização entre alfabetizados é de 78,8%, enquanto a de analfabetos, 65,82%. Essa diferença cai para 4,16% pontos percentuais ao considerar a sub-amostra de indivíduos com até um ano de escolaridade.

Em termos de horas trabalhadas, mais da metade dos indivíduos da amostra trabalham mais de quarenta e cinco horas semanais, mas a proporção é maior entre analfabetos do que entre alfabetizados, 73,91% e 54,12%, respectivamente. Por possuir menor estoque de capital humano analfabetos recebem menos que alfabetizados por hora trabalhada, em média, de modo que se faz necessário dedicar mais tempo ao trabalho para obter o mesmo rendimento. Confirma-se o fato ao considerar alfabetizados com um ano de escolaridade: 66,72% dedicam mais de quarenta e cinco horas semanais ao trabalho, proporção pouco menor que entre os analfabetos. Na questão de rendimentos, o salário-hora entre alfabetizados e analfabetos de modo geral, para o trabalho principal, difere substancialmente: em média, os primeiros auferem um valor, por hora, três vezes maior que relativamente aos segundos. No entanto, ao considerar aqueles com no máximo um ano de escolaridade a diferença se reduz bruscamente para aproximadamente 1,4 vezes – alfabetizados em relação a analfabetos.

2.2 Transitados *versus* não-transitados: um ano de estudo

Para contrastar dados entre transitados e não-transitados da condição de analfabetismo restringe-se a base de dados aos indivíduos analfabetos na primeira entrevista, e excluem-se aqueles cuja escolaridade excede um ano de estudo. Entre os 2.596 que se enquadram nessas condições 25,65% tornaram-se alfabetizados ao longo das entrevistas. Informações das variáveis analisadas podem ser conferidas na tabela 6. A proporção de homens e mulheres entre transitados e não-transitados apresenta-se praticamente a mesma, respeitando as proporções da amostra, cerca de 59% e 41%, respectivamente, assim como a participação de brancos e não-brancos entre esses grupos, de aproximadamente 65% e 35%, respectivamente.

A proporção de indivíduos que moram com até outros cinco não se diferencia entre transitados e não-transitados, havendo uma diferença muito pequena na quantidade relativa de pessoas que moram com no máximo duas crianças no mesmo domicílio, 2% a mais entre não-transitados, e que se reduz para 1% na média da segunda entrevista.

No que se refere às faixas de idade, como visto anteriormente para analfabetos e alfabetizados com até um ano de escolaridade, que configura o caso de transitados e de não-transitados, a maior parte dos indivíduos se situa na faixa dos 46 aos 60 anos de idade, algo em torno de 60%, enquanto apenas 15% possuem entre 25 e 35 anos de idade. Isso sugere que gerações mais recentes possuem menor taxa de analfabetismo relativamente às mais antigas, resultando em menos pessoas alfabetizadas quando adultas.

Em questão de ocupação, 39% dos não-transitados estão empregados na primeira entrevista contra 49% transitados, o que evidencia uma diferença já pré-existente entre os indivíduos que optam pela alfabetização relativamente àqueles que permanecem analfabetos. Para esses, a porcentagem de ocupados permanece a mesma na segunda entrevista, enquanto entre os que se alfabetizam a taxa de ocupação aumenta para 53%. Existe diferença prévia, ainda, nos rendimentos do salário principal entre esses grupos de indivíduos: salário-hora de R\$ 2,11, em média, para os analfabetos transitados contra R\$ 1,93 na média dos analfabetos que assim permanecem. Essa

diferença, inicialmente de quase 10% cai para aproximadamente 3% ao considerar os rendimentos médios da segunda entrevista.

Novamente, existe diferencial entre as taxas de formalização de transitados e não transitados logo na primeira entrevista, de uma taxa de 70% para os primeiros contra 66% dos segundos, diferença que se amplia para oito pontos percentuais na segunda entrevista, decorrente de elevação em dois pontos percentuais na taxa de formalização entre analfabetos que se alfabetizam e redução desse mesmo montante entre aqueles que se mantêm analfabetos. Restringindo a análise à formalização no setor privado, encontra-se o mesmo diferencial para a primeira entrevista, mas sob proporções de formais mais baixas para ambos os grupos, porém, a distância se amplia na segunda entrevista em nove pontos percentuais, com 71% de formalização entre os que se alfabetizam. A quantidade de trabalhadores nos setores privado e público não se diferencia muito entre transitados e não transitados, uma vez que a alfabetização por si só deve pouco afetar a probabilidade de ocupação de empregos públicos.

Tabela 6 – Características dos analfabetos transitados e não-transitados, na primeira e na segunda entrevistas

	Primeira entrevista		Segunda entrevista	
	Não - transitados	Transitados	Não - transitados	Transitados
	0,74	0,26	0,74	0,26
Homens	0,41	0,41	0,41	0,41
Branco	0,35	0,36	0,35	0,36
Até cinco moradores	0,79	0,79	0,81	0,80
Mais de cinco moradores	0,21	0,21	0,19	0,20
Até duas crianças moradoras ¹	0,95	0,93	0,95	0,94
Mais de duas crianças moradoras	0,05	0,07	0,05	0,06
De 25 a 35 anos	0,15	0,16	0,15	0,16
De 36 a 45 anos	0,23	0,24	0,23	0,24
De 46 a 60 anos	0,62	0,60	0,62	0,60
Ocupados	0,39	0,48	0,39	0,53
Desocupados	0,04	0,04	0,03	0,05
Não Economicamente ativas	0,57	0,47	0,58	0,42
Privado ²	0,93	0,91	0,93	0,92
Público	0,07	0,09	0,07	0,08
Formal ²	0,66	0,70	0,64	0,72
Formal - setor privado	0,64	0,68	0,62	0,71
Menos que 35 horas	0,08	0,10	0,08	0,11
De 35 a 40 horas	0,11	0,16	0,11	0,17
De 40 a 45 horas	0,05	0,06	0,06	0,06
Mais de 45 horas	0,76	0,68	0,75	0,66
Salário-hora do trabalho principal ³	1,93	2,11	2,07	2,14

1. Indivíduos com menos de 10 anos de idade.

2. Representativa apenas de um total de 14% e 19%, em média, dos não-transitados e dos transitados, respectivamente, da terceira amostra.

3. Representativa apenas de um total de 32% dos não-transitados e de 42% dos transitados da terceira amostra, em média.

Finalmente, no que se refere às horas trabalhadas, na primeira entrevista já se observa proporcionalmente mais pessoas trabalhando menos entre os transitados, 26% desses trabalham até 40 horas na semana contra 19% dos que não transitam, enquanto mais desses últimos trabalham mais de 40 horas semanais, com destaque àqueles que trabalham mais de 45 horas: 76% entre os não-transitados e 68% entre os que posteriormente se alfabetizam. Ao longo das entrevistas, os que

se alfabetizam passam a trabalhar menos horas, em média: 66% trabalham mais de 45 horas semanais e 28% que trabalham até de 40 horas na semana, enquanto a média permanece praticamente a mesma entre os que permanecem analfabetos através das faixas.

3. Impacto de alfabetização sobre salário e emprego

3.1 Cross section: alfabetizado versus analfabetos

As primeiras estimações procuram captar efeito da alfabetização sobre rendimento (logaritmo natural do salário-hora) e emprego utilizando informações dos indivíduos na primeira entrevista. O ano da primeira entrevista pode variar entre os indivíduos, uma vez que se juntam as bases de 2002 a 2006 e, de todas as informações, mantém-se apenas aquelas da primeira entrevista para cada indivíduo. Para controlar pelas diferenças específicas de cada ano usam-se, entre as variáveis de controle, *dummies* que os identificam.

Tabela 7 – Resultados: *Cross-section* de Rendimento e Emprego

Controles	Log salário-hora		Emprego	
	β /se	β /se	β /se	β /se
Alfabetizado	0.277*** (0.087)	0.224*** (0.084)	-0.011 (0.020)	-0.011 (0.021)
Gênero	0.188*** (0.062)	0.178*** (0.060)	0.022 (0.015)	0.021 (0.015)
Gênero x Alfabetização	-0.040 (0.101)	0.040 (0.097)	0.022 (0.027)	0.023 (0.027)
Raça	0.167*** (0.043)	0.041 (0.045)	0.028** (0.013)	0.023 (0.015)
Nº pessoas fam	-0.013 (0.014)	-0.009 (0.013)	-0.009** (0.004)	-0.008** (0.004)
De 35 a 44 anos ¹	0.095 (0.060)	0.103* (0.058)	0.024 (0.020)	0.025 (0.020)
De 45 a 60 anos	0.183*** (0.059)	0.192*** (0.057)	0.035* (0.019)	0.035* (0.019)
Formal	0.307*** (0.045)	0.302*** (0.044)		
Crianças	-0.017 (0.026)	-0.016 (0.025)	-0.006 (0.008)	-0.007 (0.008)
d_2003 ¹	0.238*** (0.057)	0.231*** (0.055)	-0.025 (0.017)	-0.029* (0.017)
d_2004	0.303*** (0.058)	0.282*** (0.056)	0.008 (0.017)	0.003 (0.017)
d_2005	0.286*** (0.059)	0.278*** (0.057)	0.008 (0.018)	0.004 (0.018)
Recife ¹		-0.451*** (0.070)		-0.021 (0.022)
Salvador		-0.522*** (0.079)		-0.010 (0.023)
Belo Horizonte		-0.354*** (0.063)		0.015 (0.021)
Rio de Janeiro		-0.129** (0.059)		0.020 (0.018)
Porto Alegre		-0.085 (0.071)		0.012 (0.023)
Constante	-0.055 (0.096)	0.191* (0.098)	0.897*** (0.026)	0.896*** (0.030)
Número de observações	804	804	2234	2234

* p<0.10, ** p<0.05, *** p<0.01

1. Dummies de referência (retiradas): de 25 a 35 anos; ano de 2002; RM de São Paulo

Os modelos estimados em *cross-section* incluem como controles, além das *dummies* de ano, variáveis binárias de alfabetização, de gênero, de interação gênero-alfabetização, de raça, de faixa etária, de formalidade de região metropolitana, e número de pessoas e de crianças na família. Os resultados estão expostos na tabela 7.

A primeira coluna de cada variável dependente são as estimações sem controlar por região metropolitana. Nota-se que não se encontra praticamente nenhum resultado estatisticamente significativo nas estimações de probabilidade de emprego. No entanto, praticamente todas as variáveis que atuam como controle nas estimações de salário sugerem forte impacto sob um nível de significância de 1%. De modo geral, alfabetizados ganham, em média, 27,7% mais que analfabetos, sem controlar por região metropolitana. Ao fazê-lo, o diferencial dos alfabetizados em relação aos analfabetos se reduz para 22,4%, sendo menor entre indivíduos de qualquer outra região metropolitana relativamente àqueles da região metropolitana de São Paulo. As demais variáveis apresentam resultados esperados em termos dos sinais dos coeficientes. Em média: i) homens auferem 17,8% mais em rendimentos de trabalho relativamente às mulheres; ii) entre os 25 e 60 anos de idade, quanto mais velho maior o salário, de modo que aqueles entre os 45 e 60 anos de idade ganham 19% mais que aqueles entre os 25 e 35 anos, evidenciando o efeito da experiência sobre o salário; iii) indivíduos formais recebem 30% mais por hora relativamente aos informais; e iv) indivíduos entrevistados em anos mais recentes apresentam maiores salários que aqueles entrevistados em 2002⁷.

3.2 Painel: transição *versus* não-transição para alfabetizado

Enquanto estimações *cross-section* resultam em coeficientes que fornecem efeitos médios da alfabetização sobre rendimento comparando resultados de indivíduos alfabetizados e de indivíduos analfabetos num determinado momento no tempo, estimações usando estrutura de painel, devido à existência de informações para mais de um período de tempo, permitem encontrar o efeito médio da alfabetização comparando indivíduos que se alfabetizam com aqueles que permanecem analfabetos.

A estimação do retorno médio da escolaridade, como explorado em muitos trabalhos, envolve variáveis individuais não observáveis relacionadas com escolaridade e que afetam a renda, como habilidade, de modo que métodos *cross-sections* simples como o OLS resultam em estimadores viesados. O uso de variáveis instrumentais faz-se necessário para contornar esse problema de endogeneidade. No entanto, encontrar um bom instrumento constitui uma etapa difícil de concluir, ainda mais sob escassez de respostas em uma grande quantidade de variáveis.

Nesse aspecto, o uso de dados longitudinais facilita a tarefa de estimação, uma vez que torna disponível o método de estimação de efeitos fixos que permite eliminar não observáveis individuais constantes no tempo. No que se refere a este trabalho, a disponibilidade de informação para dois períodos permite eliminar características não observáveis específicas ao indivíduo e constantes no tempo que afetariam a condição de alfabetização e, ao mesmo tempo, o salário e a probabilidade de emprego. Essas características são, de modo geral, identificadas como habilidade, mas envolvem qualificações como talento, esperteza, persistência, esforço, entre outras.

A dificuldade encontrada na utilização dos dados da PME jaz na ausência de variáveis que não sejam constantes ao longo do tempo, ou seja, variáveis além daquelas como gênero, raça e região metropolitana, uma vez que essas se anulam no próprio método de estimação utilizado com dados em painel que elimina os efeitos fixos individuais. As demais variáveis disponibilizadas por essa pesquisa, como já mencionado, possuem uma grande quantidade de não respostas. Devido a isso, restam como variável de controle apenas a *dummy* de alfabetização e o número de crianças no domicílio. Tenta-se também o uso da *dummy* de formalidade, apesar dessa apresentar grande proporção de não respostas, e interações da *dummy* de alfabetização com o segundo ano de entrevista.

Os primeiros resultados em dados em painel estão expostos nas tabelas 8 e 9. Sem controlar por ano (coluna 1.A), encontra-se que os indivíduos que se alfabetizaram ganham, em média, 9,3%

⁷ A *dummy* do ano de 2006 foi excluída nas regressões, não aparecendo em nenhum dos resultados estimados.

mais que aqueles que permanecem analfabetos, considerando os resultados de efeitos fixos (colunas azuis). Esse bônus de alfabetização aumenta três pontos percentuais quando se controla por formalidade (1B), cujo coeficiente aponta que indivíduos que se tornaram formais, em média, não apresentaram retorno em termos de rendimento. No entanto, ao estimar esse retorno apenas entre os formais (tabela 9 (2)) encontra-se valores maiores: 12,5% sobre o salário decorrente de aprender a ler e a escrever. Apesar do aumento da probabilidade de formalização dado pela alfabetização (tabela 9 (3)), 8,8%, esses resultados sugerem que os ganhos de se alfabetizar ocorrem mais pelo aumento da produtividade do que pela própria formalização.

Tabela 8 – Painel: resultados das estimações de salário – Efeitos aleatórios (RE) e efeitos fixos (FE)

Controles	(1) GERAL: LOG SALÁRIO-HORA							
	(A)		(B)		(C)		(D)	
	RE	FE	RE	FE	RE	FE	RE	FE
	β /se	β /se	β /se	β /se	β /se	β /se	β /se	β /se
Crianças no domicílio	-0.037** (0.015)	-0.031 (0.034)	-0.025 (0.016)	0.000 (0.030)	-0.037** (0.015)	-0.028 (0.034)	-0.025 (0.016)	-0.004 (0.030)
Alfabetizado	0.085*** (0.030)	0.093** (0.043)	0.095*** (0.032)	0.096** (0.047)	0.069 (0.055)	0.170** (0.078)	0.099* (0.051)	0.114* (0.065)
Formal			0.313*** (0.042)	0.139 (0.091)			-0.074 (0.065)	-0.073 (0.110)
Alfabx2004					-0.064 (0.078)	-0.213* (0.119)	0.103 (0.083)	0.140 (0.145)
Alfabx2005					0.047 (0.078)	-0.015 (0.116)	-0.015 (0.088)	-0.094 (0.129)
Alfabx2006					0.109 (0.083)	-0.087 (0.116)	0.311*** (0.042)	0.129 (0.087)
Constante	0.499*** (0.021)	0.522*** (0.025)	0.428*** (0.043)	0.535*** (0.074)	0.500*** (0.021)	0.520*** (0.025)	0.429*** (0.043)	0.543*** (0.071)
Nº observações	2075	2075	821	821	2075	2075	821	821

* p<0.10, ** p<0.05, *** p<0.01

(A) sem dummies de ano e sem formalidade, (B) sem dummies de ano, (C) sem formalidade, (D) com todos os controles

Considerando as interações de ano e alfabetização como controles, sem formalidade (1.C) os resultados da transição para alfabetizado são menores em todos os demais anos relativamente àqueles que se alfabetizam de 2002 para 2003, mas esse efeito apresenta-se significativo apenas para a interação de 2004 (sob p<10%), cujo coeficiente resultante indica que alfabetizados nesse ano tem retorno menor em alfabetização em 21,3%. Ao controlar tanto por ano como por formalidade o efeito da alfabetização passa para 11,4% de aumento sobre o salário, em média, no entanto perde-se significância dos efeitos ano-específicos e o coeficiente da formalidade apresenta-se negativo, mas com valor muito baixo e não significativo.

Uma vez que não se consegue controlar por variáveis constantes no tempo, a estratégia utilizada foi estimar o impacto da alfabetização sobre renda e emprego para cada uma das sub-amostras diferenciadas por gênero, por faixa etária (na primeira entrevista) por região metropolitana, e por condição de formalidade. Quebra-se, ainda, em níveis menores: por cada gênero em cada uma das regiões metropolitanas e por cada gênero em cada uma das faixas etárias. Serão expostos neste trabalho apenas os resultados significativos de cada estimação, devido ao grande número de resultados gerados com essa estratégia. Os resultados para formalidade dessas são expostos somente no caso da divisão em gênero, pois nas restantes a perda de informações é grande com o uso dessa variável devido a maiores divisões na amostra.

Tabela 9 – Painei: resultados das estimações de salário entre os formais e da probabilidade de formalização

Controles	(2) POR FORMALIDADE				(3) DE FORMALIDADE			
	FORMAL		FORMAL: SETOR PRIVADO		PÚBLICO E PRIVADO		PRIVADO	
	RE	FE	RE	FE	RE	FE	RE	FE
	β /se	β /se	β /se	β /se	β /se	β /se	β /se	β /se
Crianças no domicílio	-0.013 (0.017)	0.006 (0.027)	-0.007 (0.017)	0.008 (0.028)	0.002 (0.014)	-0.005 (0.026)	0.004 (0.015)	-0.006 (0.028)
Alfabetizado	0.116*** (0.036)	0.125*** (0.044)	0.119*** (0.039)	0.123** (0.049)	0.075*** (0.027)	0.088** (0.041)	0.095*** (0.029)	0.109** (0.043)
Constante	0.734*** (0.027)	0.722*** (0.024)	0.716*** (0.029)	0.707*** (0.025)	0.612*** (0.024)	0.658*** (0.023)	0.594*** (0.025)	0.639*** (0.025)
Nº observações	519	519	467	467	848	848	785	785

* p<0.10, ** p<0.05, *** p<0.01

De modo geral, entre as mulheres a alfabetização tem impacto maior e bastante significativo: recebem, em média, 16,6% mais aquelas que se alfabetizam (tabela 10). Ao controlar por formalidade o coeficiente da alfabetização mostra-se ainda maior, de 20,3%, e as formalizadas ganham um adicional médio de aproximadamente 23% sobre o salário. No entanto, ao controlar pelas interações entre transição e ano, os resultados indicam que as alfabetizadas de 2002 para 2003 recebem 41% mais relativamente às alfabetizadas em anos posteriores, em média. E de 2004 a 2006 o retorno da alfabetização oscila ao longo dos anos, tornando-se menos significativo. Junto com a formalidade como controle (3D), as *dummies* de ano e de formalidade perdem a significância, e o retorno da escolaridade para aquelas alfabetizadas em 2003 é de 28,8%, o qual é ainda maior e com muita significância para as alfabetizadas entre 2005 e 2006, 23,3% adicionais.

Tabela 10 – Painei: resultados das estimações de salário entre as mulheres

Controles	(3) POR GÊNERO: MULHER							
	(A)		(B)		(C)		(D)	
	RE	FE	RE	FE	RE	FE	RE	FE
	β /se	β /se	β /se	β /se	β /se	β /se	β /se	β /se
Crianças no domicílio	-0.065** (0.028)	-0.029 (0.057)	-0.060** (0.028)	-0.012 (0.060)	-0.042 (0.040)	0.030 (0.077)	-0.040 (0.042)	0.037 (0.081)
Alfabetizado	0.127*** (0.045)	0.166** (0.067)	0.214*** (0.079)	0.410*** (0.089)	0.186*** (0.046)	0.203** (0.077)	0.254*** (0.078)	0.288** (0.133)
Formal					0.362*** (0.067)	0.232** (0.098)	-0.143 (0.115)	-0.117 (0.208)
Alfabx2004			-0.160 (0.118)	-0.405** (0.196)			-0.077 (0.146)	-0.091 (0.231)
Alfabx2005			-0.189* (0.107)	-0.344** (0.144)			-0.116 (0.120)	-0.193 (0.193)
Alfabx2006			-0.019 (0.135)	-0.351* (0.181)			0.361*** (0.067)	0.232*** (0.085)
Constante	0.377*** (0.034)	0.390*** (0.038)	0.375*** (0.034)	0.380*** (0.039)	0.289*** (0.057)	0.343*** (0.071)	0.290*** (0.058)	0.339*** (0.074)
Nº observações	882	882	882	882	215	215	215	215

* p<0.10, ** p<0.05, *** p<0.01

(A) sem *dummies* de ano e sem formalidade, (B) sem *dummies* de ano, (C) sem formalidade, (D) com todos os controles

Não foram obtidos resultados para homens em praticamente nenhuma das estimações realizadas, apenas para alfabetizados na região metropolitana de Recife, como exposto em seguida. Isso indica que, de modo geral, os resultados significativos nas estimações de renda em alfabetização são, em grande parte, devido aos efeitos fortes que ocorrem entre as mulheres em termos de retorno sobre rendimentos. Uma vez que homens com baixa escolaridade costumam se dedicar mais a trabalhos que exigem força, a alfabetização em si pode não implicar em grandes retornos, uma vez que para determinadas ocupações pouco se exige em termos de escolaridade, até mesmo saber ler e escrever. Porém, faz sentido pensar que entre as mulheres a alfabetização por si só constitui uma qualificação importante, uma vez que praticamente nenhuma atividade é essencialmente mecânica. Simples afazeres domésticos, atividades da ocupação predominante entre mulheres de baixa escolaridade, exigem, em certa medida, que se saiba ler e escrever.

Tabela 11 - Painel: resultados das estimações de salário entre as regiões metropolitanas

Controles	(4) POR REGIÃO METROPOLITANA					
	(4.1) SALVADOR				(4.2) BELO HORIZONTE	
	(A)		(B)		(B)	
	RE	FE	RE	FE	RE	FE
	β/se	β/se	β/se	β/se	β/se	β/se
Crianças no domicílio	-0.085** (0.041)	-0.065 (0.111)	-0.083** (0.042)	-0.056 (0.105)	-0.023 (0.038)	-0.256** (0.111)
Alfabetizado	0.244*** (0.080)	0.217* (0.124)	0.354** (0.140)	0.445** (0.181)	0.074 (0.164)	0.458*** (0.174)
Alfabx2004			-0.258 (0.228)	-0.496 (0.434)	0.142 (0.210)	-0.397 (0.289)
Alfabx2005			-0.082 (0.197)	-0.220 (0.288)	0.046 (0.217)	-0.296 (0.402)
Alfabx2006			-0.242 (0.187)	-0.467* (0.277)	0.301* (0.176)	-0.694*** (0.250)
Constante	0.199*** (0.066)	0.223** (0.103)	0.198*** (0.066)	0.217** (0.097)	0.468*** (0.064)	0.624*** (0.082)
N observações	268	268	268	268	264	264

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

(A) sem dummies de ano e (B) com dummies de ano

A análise por região metropolitana, registrados na tabela 11, possui resultados significativos apenas para Salvador e Belo Horizonte, e essa última apenas quando controla-se por ano. Em Salvador, em média, alfabetizados obtiveram ganhos de quase 22% sobre o salário (4.1A) relativamente aos que permaneceram analfabetos. Controlando por ano (4.1B) observa-se que tais ganhos persistem apenas para aqueles que estão alfabetizados em 2003, enquanto nos restantes dos anos os resultados são relativamente menores, sendo significantes apenas para 2006, quando o retorno de alfabetizados apresentou-se 46,7% menor. O mesmo ocorre no caso de Belo Horizonte, retorno de 45,8% para os alfabetizados em 2003 enquanto aqueles em 2006 recebem 69,4% menos reativamente a esses.

Em praticamente todas as regiões metropolitanas analisadas, exceto na de São Paulo, os resultados são significativos quando se consideram os efeitos anos-específicos para mulheres que se alfabetizam em 2003, variando de 65% em Salvador a 30,4% no Rio de Janeiro, com resultados de efeitos negativos significativos em 2006 para Recife, Salvador e Belo Horizonte. Controlando apenas por alfabetização e pela quantidade de indivíduos com até 10 anos no domicílio, os coeficientes de alfabetização são significativos para homens em Recife, com ganhos em rendimento de 17% relativamente aos não alfabetizados, enquanto para as mulheres esses resultados são significativos apenas em Porto Alegre, onde os ganhos das alfabetizadas foi, em média, de 23,5%.

Tabela 12 - Painel: resultados das estimações de salário entre as faixas etárias e entre as mulheres – efeitos aleatórios (RE) e efeitos fixos (FE)

Controles	(5) POR FAIXA ETÁRIA		(6) POR FAIXA ETÁRIA: MULHER	
	De 46 a 60 anos		De 46 a 60 anos	
	RE	FE	RE	FE
	β/se	β/se	b/se	b/se
Crianças no domicílio	-0.043** (0.022)	-0.002 (0.052)	-0.034 (0.031)	0.027 (0.069)
Alfabetizado	0.140*** (0.038)	0.130** (0.059)	0.218*** (0.054)	0.243*** (0.082)
Constante	0.514*** (0.027)	0.527*** (0.027)	0.394*** (0.040)	0.401*** (0.037)
Nº de observações	1222	1222	592	592

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Finalmente, os resultados relacionados à faixa etária indicam, com significância ($p < 5\%$), que indivíduos entre os 45 e os 60 anos de idade que se alfabetizaram ganham, em média, 13% mais relativamente àqueles que permanecem analfabetos nessa mesma faixa etária. Uma vez que a quantidade de analfabetos nessa faixa etária é relativamente menor, o aprendizado de habilidades básicas de escrita e leitura atua como um diferencial em relação aos demais. Novamente, ao olhar em cada faixa etária por gênero, como exposto na tabela 12, os resultados para mulheres são maiores e bastante significativos. Sem *dummies* de ano (valores não apresentados), as estimações resultam em coeficientes significativos apenas para a última faixa etária: retorno de 24,3% da alfabetização para mulheres entre 45 e 60 anos de idade. Ao considerar as interações de ano com a *dummy* de alfabetização, observa-se resultados em todas as faixas etárias para mulheres, sendo esses maiores para as mais velhas, de 43,8% em média, e para as mais novas, de quase 40,1%, embora apenas para as alfabetizadas em 2003. Ressalta-se que o número de casos de transição é baixo principalmente na primeira faixa.

Como se nota, não foi exposto nenhum resultado em termos de probabilidade de emprego além das estimações com dados em *cross-section*. O motivo jaz na não obtenção de nenhum resultado significativo para essa variável em resposta à alfabetização entre os adultos.

4 Considerações finais

Diversos países no mundo têm sinalizado grande preocupação com as necessidades educacionais das respectivas populações, num momento em que se enfatiza a necessidade de equalizar as oportunidades individuais de acesso à educação para que todos possam formar capital humano para geração de uma renda mínima a fim de atingir padrões de vida satisfatórios. Inserem-se nesse contexto as iniciativas mundiais no sentido de prover habilidades básicas de escrita e leitura entre jovens e adultos que, por algum motivo, não puderam se educar quando crianças.

De modo geral, a alfabetização mostrou-se um importante meio de incremento salarial entre as populações analfabetas, e esforços no sentido dimensionar esse impacto foram empenhados em diversos estudos internacionais, tanto na estimação do impacto da proficiência na língua hebraica entre imigrantes de Israel como no retorno da alfabetização entre cidadãos de Gana. Este trabalho direcionou-se no mesmo sentido, procurando estimar os retornos médios em salário obtidos por indivíduos que se alfabetizam, além do efeito sobre a probabilidade de emprego. Estimar médias educacionais entre indivíduos alfabetizados e analfabetos não fornece um valor devidamente estimado do verdadeiro retorno da alfabetização buscado, uma vez que não toma os devidos cuidados com as características não-observáveis de cada indivíduo. Como os dados descritivos mostraram, preexistem diferenças salariais (média) entre aqueles que optaram pela alfabetização e os que permaneceram analfabetos. Nesse sentido, a disponibilidade de dados longitudinais, da Pesquisa Mensal de Emprego do IBGE, auxilia na tarefa de contornar esse problema ao permitir que

se controlem eliminem as características não observáveis constantes no tempo da estimação, obtendo estimativas mais precisas do retorno da alfabetização entre indivíduos adultos.

Entre aqueles que se alfabetizaram, o retorno médio apresentado foi de 9,3% sobre o rendimento. Como a formalização não apresenta coeficientes significativos em seus efeitos sobre salário, acredita-se que o retorno decorre essencialmente de aumento de produtividade. Entre as mulheres, essa transição resulta em aumentos ainda maiores, 16,6%, enquanto os coeficientes na estimação entre os homens não se mostram significativos. Esse fato sugere que as atividades de homens de baixa escolaridade, muitas vezes pela predominância no uso da força, não exige as habilidades básicas em escrita e leitura suprida pela alfabetização e requerida em simples atividades domésticas.

No tocante à idade de alfabetização, pessoas alfabetizadas entre os 45 e 60 anos de idade possuem coeficientes significativos de alfabetização em rendimento, 13%, sobretudo entre as mulheres, 24,3%, sugerindo que, em populações com baixa proporção de alfabetizados o retorno das habilidade básicas são maiores e significativos. Dos resultados por região metropolitana encontra-se que indivíduos de Salvador, região que possui o menor salário-hora médio, apresentam retornos significativos de 21,7% provenientes da alfabetização, enquanto homens de Recife ganham 17% mais e mulheres gaúchas, 23,4% adicionais. Nenhum dos resultados analisados em termos de probabilidade de emprego apresentou-se significante.

Esses são resultados importantes e contribuem para direcionar política públicas que pretendam estabelecer medidas nessa área. Como se observou, a alfabetização aumenta de fato a produtividade, sobretudo entre as mulheres, entre populações de baixo rendimento médio e de baixa taxa de alfabetização. No intuito de aumentar a capacidade de geração renda dos indivíduos desprivilegiados da sociedade para elevar os níveis de bem-estar, e diante de limites em termos orçamentários, esforços deveriam ser focados onde a alfabetização constituiria uma qualificação adicional, resultando em maiores benefícios.

Bibliografia

Blunch, N., Verner D. (2000) “Is Functional Literacy a Prerequisite for Entering the Labor Market? An Analysis of the Determinants of Adult Literacy and Earnings in Ghana”, Working Paper 00-05, Centre For Labour Market And Social Research.

Green, D.A, Riddell, W.C. (2001) “Labour Market Institutions and Outcomes: A Cross-National Study”, Discussion Paper n° 01-05, Department of Economics McMaster University March 2001

Chiswick, B.R., Repetto, C. (2000) “Immigrant Adjustment in Israel: Literacy and Fluency in Hebrew and Earnings”, Discussion Paper Series n° 177, Institute for the Study of Labor.

Denny, K., Harmon, C., Redmond, S. (2000) “Functional Literacy, Educational Attainment And Earnings: Evidence From The International Adult Literacy Survey”, Working Paper 00/09, The Institute For Fiscal Studies.