

# EVOLUÇÃO DO DIFERENCIAL DE RENDIMENTOS ENTRE SETOR FORMAL E INFORMAL NO BRASIL: O PAPEL DAS CARACTERÍSTICAS NÃO OBSERVADAS\*

*Ana Flávia Machado*<sup>\*\*</sup>

*Ana Maria Hermeto Camilo de Oliveira*<sup>\*\*\*</sup>

*Mariângela Antigo*<sup>\*\*\*\*</sup>

**RESUMO** Este artigo investiga o diferencial de rendimentos entre setor informal e formal no Brasil urbano nos anos 1992, 1998 e 2004. Utiliza duas definições para o setor informal e formal, decompondo os diferenciais obtidos por meio de regressões quantílicas, segundo o método padrão de Oaxaca-Blinder. Pretende-se assim, por um lado, recorrer a conceitos distintos de setor informal, avançando na compreensão da estratégia do trabalhador embutida na escolha ocupacional. Por outro lado, investigar a evolução dos diferenciais de renda, isolando os efeitos das mudanças nas dotações dos atributos e dos retornos dos mesmos sobre o diferencial entre os dois setores, por quantis, no tempo, é mais uma forma de analisar o papel das características não observadas, elucidando o quanto importante é a escolha ocupacional. A fonte de dados são as PNADs (Pesquisa Nacional de Amostra por Domicílio) de 1992, 1998 e 2004. Os resultados obtidos mostram que os ocupados do setor informal nos quantis da base da distribuição, particularmente na definição onde são considerados somente os “conta própria”, detêm retornos mais elevados por suas características não observadas, e essa tendência é crescente no período analisado.

\* Artigo recebido em 18 de setembro de 2006 e aprovado em 5 de março de 2008.

\*\* Professora adjunta do Cedeplar/UFMG, doutora pelo IE-UFRJ, e-mail: afmachad@cedeplar.ufmg.br

\*\*\* Professora adjunta, doutora pelo Cedeplar/UFMG, e-mail: ahermeto@cedeplar.ufmg.br

\*\*\*\* Doutoranda em Economia no Cedeplar/UFMG, e-mail: maantigo@cedeplar.ufmg.br

**Palavras-chave:** diferencial de rendimento; regressões quantílicas; setor informal; mercado de trabalho

**Código JEL:** J23-J31

### **EARNING GAP EVOLUTION BETWEEN FORMAL AND INFORMAL SECTOR IN BRAZIL: THE ROLE OF THE NON OBSERVABLE CHARACTERISTICS**

**ABSTRACT** This paper investigates the earning gap between informal and formal sector in urban Brazil in the year 1992, 1998 e 2004. We use quantile regression models to estimate the gap earnings of informal and formal workers and decompose the gap by Oaxaca-Blinder method. The aim is, on one side, to resort to distinct concepts of informal sector, moving on in the understanding of the worker's strategy embedded in the occupational choice. On the other side, to look into the evolution of the gap earnings, isolating the effects of the changes in the endowment of the attributes and of the returns of the same attributes between the two sectors, by quantiles, through time, is another way to analyze the role of the non observed characteristics, explaining how important the occupational choice is. The PNAD is the database. The results show that unobservable characteristics cause the gap reduction in the lower earnings quantiles in the period, especially to the self-employee. We also find that the earnings gaps between formal and informal workers are wider at low conditional quantiles than at high ones.

**Key words:** earnings gap; quantile regression; informal sector; labor market

## INTRODUÇÃO

A estrutura de rendimentos é um dos pontos mais abordados pela literatura sobre mercado de trabalho, dada a sua importância para a distribuição de renda em uma determinada economia. Muitos são os modelos teóricos que buscam explicar os diferenciais de rendimentos (capital humano, discriminação, segmentação, salário eficiência, barganha salarial etc.). Como se está interessado em analisar o diferencial de rendimentos entre setor informal e setor formal, recorre-se ao modelo de segmentação formulado por Dickens e Lang (1985). De forma resumida, o modelo de segmentação busca explicar o diferencial de rendimento entre trabalhadores igualmente produtivos por intermédio dos atributos do setor onde estão ocupados. Em outras palavras, mesmo que tenham atributos produtivos semelhantes ou idênticos, indivíduos podem ser diversamente remunerados em virtude do setor em que estão ocupados.

No Brasil, estudos como de Carneiro e Henley (2001), Menezes-Filho, Mendes e Almeida (2004) exploraram esse tema e mostraram que a hipótese da segmentação não é corroborada, uma vez que as diferenças entre os rendimentos nos dois setores são explicadas, em sua maior parte, pelo nível de escolaridade e pelas características não observadas. Com o mesmo objetivo, Pianto e Pianto (2002) evidenciam, por meio de aplicação de regressões quantílicas, que as características não observadas surtem efeito negativo sobre os trabalhadores inseridos no topo da distribuição. Concluem, também, que os diferenciais de rendimentos podem estar associados, em sua maior parte, às diferenças de atributos dos trabalhadores, com exceção daqueles de menor renda, para quem os retornos dos atributos também contam, corroborando a hipótese da segmentação.

Em que pesem as diferenças de métodos nesses artigos, o resultado é consensual quanto à relação entre o efeito das características não observadas e a escolha ocupacional do trabalhador para explicar os diferenciais de rendimentos entre o setor formal e informal, principalmente para os localizados no topo da distribuição.

A contribuição deste artigo é, portanto, decompor os diferenciais obtidos em cada quantil no tempo. Investigar a evolução dos diferenciais de renda, isolando os efeitos das mudanças nas dotações dos atributos e dos

retornos dos mesmos sobre o diferencial entre os dois setores, por quantis, no tempo, é mais uma forma de analisar o papel das características não observadas, elucidando o quão importante é a escolha ocupacional.

Ao se tratar da estratégia ocupacional, buscam-se, basicamente, duas associações. A primeira refere-se à falta de oportunidades de um emprego formal, o que levaria as pessoas a serem “empurradas” para o informal, o que não constituiria uma escolha ou estratégia propriamente dita. Por outro lado, se os trabalhadores, no caso por conta própria, valorizam o fato de não ter chefe, de estarem sujeitos a maior flexibilidade de jornada e, até mesmo, maior nível de rendimento, pode-se associar a inserção no setor informal como uma escolha.

Por essa razão, são definidas duas categorizações para a segmentação formal/informal. A primeira se associa ao conceito de regulamentação e a outra ao conceito de subordinação no mercado de trabalho, explicado na seção 1. A distinção minimiza a heterogeneidade presente no mercado de trabalho, em especial o informal, e, busca, assim, analisar, além do papel das características observadas, o papel que a heterogeneidade não observada exerce sobre as escolhas ocupacionais, por meio dos diferenciais de salários entre os dois setores.

Este artigo está, portanto, dividido em cinco seções, considerando esta introdução. Na próxima seção, apresenta-se breve revisão da literatura sobre conceitos de setor formal e informal. Na terceira e na quarta seções são descritos a metodologia e os resultados obtidos e, na última, descrevem-se algumas considerações gerais sobre o trabalho.

## **1. DA DEFINIÇÃO DE SETOR INFORMAL E DE ESTRATÉGIA OCUPACIONAL**

Na literatura nacional sobre mercado de trabalho, são utilizadas várias definições para o setor informal. Na verdade, há uma confusão entre relação de trabalho, atividade e setor. A mais utilizada associa a informalidade à regulamentação do trabalho. Nesse caso, trata-se como trabalhador formal os assalariados com carteira assinada e, como informal, os assalariados sem carteira assinada. É o caso dos trabalhos já citados (Carneiro e Henley, 2001; Menezes-Filho, Mendes e Almeida, 2004; Pianto e Pianto, 2002) e outros mais.

Há, no entanto, uma definição que considera a organização da produção de acordo com o proposto pelos trabalhos da Organização Internacional do Trabalho (OIT) em 1972. Segundo Cacciamali (1983, p. 18), o objetivo “era construir uma categoria de análise que descrevesse as atividades geradoras de uma renda relativamente baixa e aglutinasse os grupos de trabalhadores mais pobres no meio urbano”.

A concepção seminal de setor informal no trabalho da OIT (1972) ocasionou a definição da dicotomia formal/informal segundo uma abordagem denominada subordinação. As atividades do setor informal ocupariam nichos do mercado não preenchidos pela atividade tipicamente capitalista, uma vez que a lucratividade não é suficiente para operação da firma capitalista. Com o desenvolvimento da esfera produtiva, a atividade, até então espaço para o setor informal, se tornaria atraente para a atividade capitalista, levando-a a ocupar, por meio da concorrência, o lugar preenchido pelas unidades produtivas informais.

Seguindo, portanto, essa abordagem de subordinação, o setor informal compreenderia a pequena unidade de produção na qual não ocorre a dissociação do proprietário dos meios de produção do processo de trabalho. Em termos de uma tipologia, o informal englobaria as pequenas empresas familiares, o trabalhador autônomo (excluindo profissional liberal) e o serviço doméstico remunerado em moeda ou espécie. Marginalmente, o trabalho assalariado poderia ocorrer. Essa concepção está presente em estudos como os de Cacciamali (2000) e Abramovay et al. (2003), e permeia a construção da pesquisa Economia Informal Urbana (ECINF) realizada pelo IBGE, nos anos 1997 e 2003, em domicílios e estabelecimentos.<sup>1</sup>

Entretanto, a ECINF fica circunscrita ao setor informal, impossibilitando qualquer tipo de comparação entre os dois setores. Em razão dessa configuração, optamos por trabalhar com os microdados da Pesquisa Nacional de Amostra por Domicílio (PNAD), que, por intermédio da variável posição na ocupação, possibilita construir categorias da segmentação formal/informal.<sup>2</sup>

Recorre-se a duas definições para tal segmentação do mercado de trabalho. A primeira mescla a abordagem de regulação do trabalho com a de subordinação, pois considera como setor informal os trabalhadores que se declaram na posição de ocupação “trabalhador por conta própria”, excluindo

do as ocupações de profissionais liberais e os empregados sem carteira assinada. De acordo com essa categorização, o setor formal seria constituído pelos empregados assalariados com carteira assinada (domésticos ou não), funcionários públicos e militares, empregadores e profissionais liberais.

A segunda, mais atrelada à concepção de subordinação, abrangeria apenas os trabalhadores por conta própria no setor informal. O setor formal seria constituído pelos ocupados assalariados (domésticos ou não), funcionários públicos e militares, empregadores e profissionais liberais.

A adoção dessas duas definições se justifica pela preocupação em distinguir a heterogeneidade presente no setor informal, e, em boa parte, esta se reflete nas estratégias ocupacionais que, por sua vez, são resultado de possíveis diferenciais de rendimento entre os dois setores. Como bem lembram Hirata e Machado (2007, p. 2),

se os trabalhadores do setor informal estivessem todos atuando por meio de uma estratégia de sobrevivência, não haveria tanta controvérsia acerca da definição de informalidade, com uma análise sob o ponto de vista da posição. Isso porque seria um setor caracterizado por indivíduos que apresentariam variáveis indicativas de capital humano que apontariam para uma situação de desvantagem ou exclusão do mercado, tal como baixa escolaridade. No entanto, a controvérsia acerca da definição deste setor pode estar refletindo a presença de indivíduos que ingressam na informalidade por “opção”...

A decisão de se auto-empregar seria, portanto, fruto de uma estratégia. Os trabalhadores menos escolarizados podem vir a obter rendimentos mais elevados como autônomos do que como assalariados com carteira assinada, e até mesmo como sem carteira, pela ausência de qualificação requerida no setor formal, sendo, assim, preteridos na seleção por vagas no formal. Nesse caso, o setor informal estaria assumindo sua vocação em gerar postos de trabalho para os excluídos do processo de desenvolvimento econômico (Machado e Andrade, 1994).

Por outro lado, pode-se constituir em uma opção, porque o histórico familiar na atividade, a experiência profissional pregressa como assalariado, a flexibilidade da jornada de trabalho, a independência de uma hierarquia dentro de empresas e a liberdade quanto à escolha de um local de trabalho contribuem para definir essa inserção. Esses trabalhadores, portanto, esta-

riam no setor em decorrência, não apenas, de algum tipo de incompatibilidade entre sua qualificação e a ocupação, mas por outras razões (Machado, Penido e Oliveira, 2005). E são justamente essas outras razões não mensuráveis identificadas como características não observadas em modelos econométricos que ajudam a explicar os diferenciais de rendimentos entre os dois setores, justificando as duas definições para a segmentação formal/informal adotada neste artigo.

## **2. DIFERENCIAL DE RENDIMENTOS E DECOMPOSIÇÃO: O MODELO ECONOMÉTRICO**

Para analisar os diferenciais entre os segmentos informal e formal, são utilizadas regressões quantílicas, de modo a avaliar como se comportam os determinantes do rendimento tanto em termos de quantis como também ao longo do tempo. Não se negligencia a importância de uma possível seleção nos modelos. Entretanto, em geral, as estimativas dos diferenciais de rendimentos entre os setores são baseadas em pressupostos muito fortes sobre os processos de seleção subjacentes à escolha entre a ocupação em cada um dos setores. Argumenta-se que a heterogeneidade individual não mensurada leva a um viés nas estimativas dos diferenciais, ou seja, que as características produtivas daqueles que escolhem o setor formal diferem de forma não mensurada dos indivíduos idênticos (em termos das características observadas) ocupados no setor informal. Contudo, podemos afirmar que não há consenso sobre esse ponto, e mesmo quando se verifica que a seleção é importante, não há consenso sobre como o processo deve ser modelado. A identificação dos modelos de seleção exige variáveis que influenciam na decisão de seleção, mas que sejam independentes do salário. No entanto, a maioria das variáveis utilizadas na literatura pode ser igualmente interpretada como conectada diretamente aos rendimentos dos indivíduos, o que significa o questionamento de sua utilização como instrumentos. Adicionalmente, um fator complicador a ser considerado é que a composição do setor informal vem mudando muito ao longo do tempo, e, nesse sentido, muitos indivíduos “selecionados” para o setor formal podem ter se movido involuntariamente para o setor informal.

A regressão quantílica se refere, portanto, à distribuição dos rendimentos, condicional ao vetor de co-variáveis. É útil quando, ao invés da média,

se quer trabalhar com a mediana, e assumir que as medianas dos salários condicionais às co-variáveis são lineares nessas co-variáveis, ou ajustar uma função linear às medianas (regressão mediana ou regressão quantílica em 0.5). Em princípio, é possível fazer o mesmo para qualquer outro quantil da distribuição. Verificando diferentes regressões quantílicas, é possível explorar diferentes partes da distribuição condicional.

Na relação entre rendimentos e escolaridade, por exemplo, a um dado nível de escolaridade, há uma distribuição (condicional) de rendimentos, presumivelmente refletindo habilidade e outras qualificações para o mercado de trabalho não observadas. Em geral, não há razão para exigir que a taxa de retorno a um ano adicional de escolaridade seja a mesma em todos os pontos da distribuição de qualificação condicional na escolaridade, e a regressão quantílica capta essas diferenças. Usada dessa maneira, a regressão quantílica é uma técnica semiparamétrica que descreve a forma da distribuição empírica sem impor restrições anteriores, mas impondo uma forma funcional linear para os parâmetros do modelo. Se a distribuição condicional muda, a forma como uma ou mais variáveis explicativas atuam também se altera, por isso as regressões quantílicas nos vários quantis têm diferentes inclinações (Deaton, 1995). A estimação das regressões quantílicas se baseia em extensões do resultado de que a mediana é o ponto mais próximo aos dados no sentido de minimizar a soma dos desvios absolutos. Os parâmetros da regressão linear mediana são dados como o valor do vetor  $\beta$  que minimiza<sup>3</sup>

$$\sum_{i=1}^n |y_i - x_i' \beta| = \sum_{i=1}^n (0.5 - 1(y_i \geq x_i' \beta)) (y_i - x_i' \beta) \quad (1)$$

A abordagem padrão para explorar o diferencial de rendimentos entre os segmentos, decompondo-o em componentes “explicados” e “não explicados”, assume que o rendimento para o indivíduo no segmento formal (grupo 1) pode ser escrito como

$$W_{1i} = \beta_1 X_{1i} + v_{1i} \quad (2)$$

e o rendimento para o indivíduo no setor informal pode ser escrito como

$$W_{2j} = \beta_2 X_{2j} + v_{2j} \quad (3)$$



onde  $\beta_1$  e  $\beta_2$  são definidos tal que  $E(v_{1i} | X_{1i}) = 0$  e  $E(v_{2j} | X_{2j}) = 0$ . A diferença entre os rendimentos médios, em cada quantil, pode ser escrita como

$$W_1 - W_2 = (X_1 - X_2) \beta_1 + (\beta_1 - \beta_2) X_2 \quad (4)$$

onde  $W_g$  e  $X_g$  representam os rendimentos médios e as características de controle para todos os indivíduos no grupo  $g$ .<sup>4</sup> O primeiro termo dessa decomposição representa o componente explicado, devido a diferenças médias nas características produtivas dos indivíduos nos segmentos formal e informal. Ele indica o diferencial previsto usando o grupo 1 — setor formal — como referência. O segundo termo é o componente não explicado e representa diferenças nos coeficientes estimados, ou seja, diferenças nos retornos a características similares entre os indivíduos em cada segmento. A parte do diferencial salarial total devida a esse componente capta os efeitos das diferenças não observadas entre os grupos.<sup>5</sup>

Para decompor as mudanças dos diferenciais entre os grupos ao longo do tempo, é utilizada uma extensão da abordagem citada, incorporando os períodos (Altonji e Blank, 1999). Nesse sentido, a equação (4) é diferenciada entre períodos; sendo  $\Delta$  a diferença média entre grupo 1 e grupo 2 em um período, a mudança nos diferenciais salariais entre os períodos  $t'$  e  $t$  é

$$\begin{aligned} \Delta W_{t'} - \Delta W_t = & (\Delta X_{t'} - \Delta X_t) \beta_{1t} + \Delta X_{t'} (\beta_{1t'} - \beta_{1t}) + \\ & (\beta_{2t'} - \beta_{2t}) X_{2t} + (X_{2t'} - X_{2t}) \beta_{2t} \end{aligned} \quad (5)$$

onde o primeiro termo,  $(\Delta X_{t'} - \Delta X_t) \beta_{1t}$ , representa o efeito de mudanças relativas ao longo do tempo nas características observadas dos dois grupos; o segundo termo,  $\Delta X_{t'} (\beta_{1t'} - \beta_{1t})$ , representa o efeito de mudanças ao longo do tempo nos coeficientes para o grupo 1, mantendo fixas as diferenças nas características observadas. Esses dois componentes representam a mudança ao longo do tempo no diferencial salarial que seria esperado, dados as mudanças nas características dos dois grupos e os coeficientes dessas características para o grupo 1 nos períodos  $t$  e  $t'$ . O terceiro e o quarto termos captam a mudança no componente não explicado do hiato  $(\beta_{1t} - \beta_{2t}) X_{2t}$ . O terceiro termo,  $(\beta_{2t'} - \beta_{2t}) X_{2t}$ , é o efeito de mudanças ao longo do tempo nos coeficientes relativos entre os grupos; e o quarto termo,  $(X_{2t'} - X_{2t}) \beta_{2t}$ , capta o fato de que mudanças ao longo do tempo nas características do grupo 2 alteram as conseqüências de diferenças nos coeficientes dos grupos  $(\beta_{1t} - \beta_{2t})$ .

### 2.1 Fonte de dados e variáveis selecionadas

A fonte de dados básica empregada neste trabalho é a Pesquisa Nacional por Amostra de Domicílios (PNAD) do Instituto Brasileiro de Geografia e Estatística (IBGE). Essa pesquisa é realizada anualmente, exceto para os anos de realização dos Censos Demográficos e para o ano 1994, e abrange cerca de 300 mil indivíduos e 100 mil domicílios.

São utilizados os dados referentes à População Economicamente Ativa ocupada na semana de referência da pesquisa, com idade entre 18 e 65 anos, com rendimento e horas trabalhadas positivos no trabalho principal, e residente nos setores urbanos do país nos anos de 1992, 1998 e 2004. Os valores monetários são avaliados pelo INPC em setembro de 2004.

Para fins deste trabalho, são utilizadas duas definições para setor formal e informal. Na primeira definição o setor formal é formado pelo indivíduo ocupado com carteira de trabalho assinada, empregador, servidor público e profissional liberal, e o setor informal pelo trabalhador por conta própria sem o profissional liberal e pelo ocupado sem carteira de trabalho assinada.<sup>6</sup> Na segunda definição, considera-se, no setor informal, apenas o trabalhador por conta própria sem o profissional liberal.<sup>7</sup>

Dessa forma, são construídas quatro amostras para cada ano considerado, que são utilizadas para avaliar o diferencial salarial entre indivíduos do setor formal e informal em cada definição ao longo dos quantis e ao longo dos anos.

Os efeitos dos atributos produtivos e fatores do mercado de trabalho sobre os rendimentos dos indivíduos ocupados no setor formal e informal são analisados em termos da seguinte regressão condicional, aplicada aos quantis 10°, 25°, 50°, 75° e 90°:

$$y = \beta_1 + \beta_2 \text{Sexo} + \beta_3 \text{Raca} + \beta_4 \text{Idade} + \beta_5 \text{Idade}^2 + \beta_6 \text{Educacao} + \beta_7 \text{Rm} + \beta_7 \text{Casado} + \beta_8 \text{Contprev} + \beta_9 \text{Integral} + \beta_{10} \sum \text{Regiao} + \beta_{11} \sum \text{Categ} \quad (6)$$

em que:  $y$ : logaritmo do rendimento-hora real; *sexo*: *dummy* para o sexo do indivíduo (1 = homens; 0 = mulheres); *raça*: *dummy* para a raça/cor do indivíduo (1 = brancos; 0 = negros); *idade*: idade calculada do indivíduo; *educação*: anos de estudo; *rm*: *dummy* para região metropolitana; *casado*: *dummy* para indivíduos casados; *contprev*: contribuição à previdência no trabalho principal; *integral*: *dummy* para jornada de trabalho (1 = integral; 0 = par-

cial); *região*: conjunto de *dummies* para as regiões do país; e, por fim, *categ*: conjunto de *dummies* para categorias de ocupação do indivíduo. Estas são representadas em quatro *dummies*, que representam, por sua vez, respectivamente, as ocupações de nível superior, médio, manual e doméstico.<sup>8</sup>

A variável integral representa jornada de trabalho integral e parcial, sendo a jornada de trabalho parcial representada por até 30 horas trabalhadas. As regiões do país estão representadas em cinco *dummies* regionais: Norte, Nordeste, Sudeste, Sul e Centro-Oeste. Além disso, por indivíduo branco entendem-se os classificados como sendo de cor Branca e Amarela, e por indivíduo negro, aqueles classificados como de cor Preta e Parda. A variável *idade* é usada aqui como *proxy* de experiência, e seu termo ao quadrado busca captar a forma em U invertido no perfil de rendimento-experiência.

### 3. RESULTADOS

#### 3.1 Simulação dos diferenciais ao longo da distribuição dos rendimentos

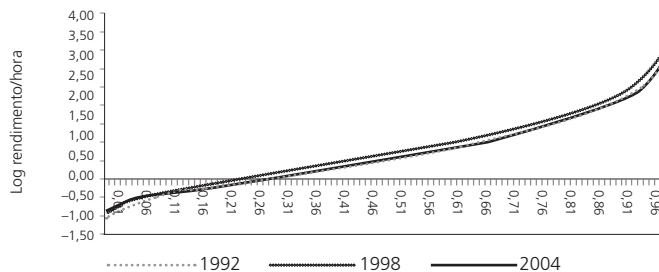
A distribuição do log do rendimento-hora dos trabalhadores do setor formal1 e informal1 (conta própria e sem carteira) é distinta nos três anos, identificando-se, nitidamente, no gráfico 1 e gráfico 2, todas as curvas referentes ao setor formal1 e informal1, respectivamente.

O mesmo não é visualizado na leitura do gráfico 3 e gráfico 4, onde ocorre superposição das distribuições para a segunda definição formal2 e informal2 (somente os trabalhadores por conta própria), em especial do 20º ao 60º quantil. Em 2004, o log do rendimento do formal2 é, praticamente, igual ao do informal2 em todos os percentis.

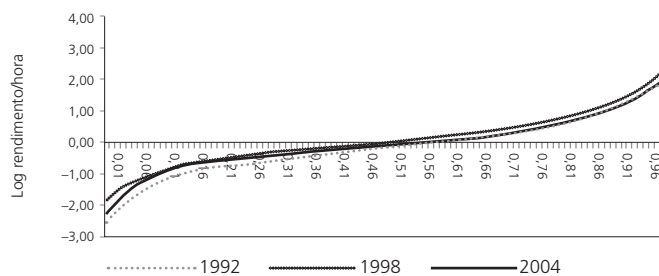
Como as distribuições são distintas para a primeira definição, os diferenciais<sup>9</sup> entre o formal e o informal tendem a ser maiores, particularmente entre os 20% mais pobres. No caso da base da distribuição, os diferenciais chegam, em 2004, a mais de 1,4 (gráfico 5). Depois do vigésimo quantil, há uma convergência entre os diferenciais nos anos 1992 e 2004, em torno de 0,60; em 1998, o diferencial se mantém pouco acima de 0,70.

Ao se trabalhar com a definição formal2-informal2, abordando apenas os autônomos sem o profissional liberal, observa-se que o nível do diferencial é menor: 0,20, em 1998, e 0,80, em 2004, para o 1% mais pobre (gráfico 6). Essa é uma evidência de que a retirada dos empregados sem carteira

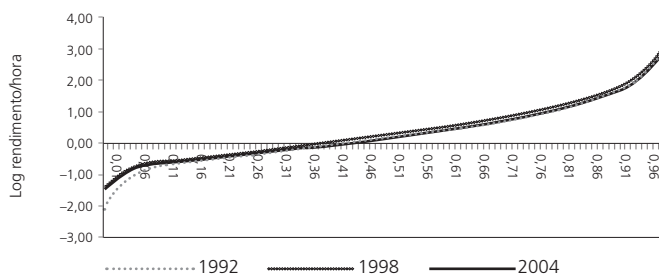
**Gráfico 1: Distribuições do log rendimento/hora, segundo setores formal1, Brasil, 1992, 1998 e 2004**



**Gráfico 2: Distribuições do log rendimento/hora, segundo setores informal1, Brasil, 1992, 1998 e 2004**

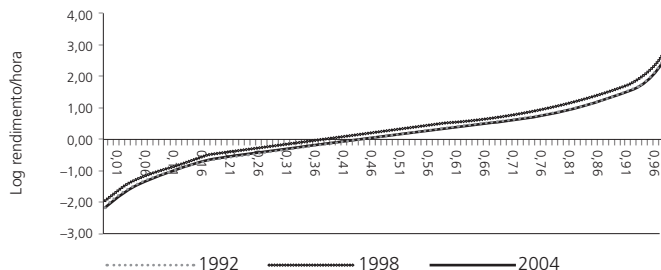


**Gráfico 3: Distribuições do log rendimento/hora, segundo setores formal2, Brasil, 1992, 1998 e 2004**

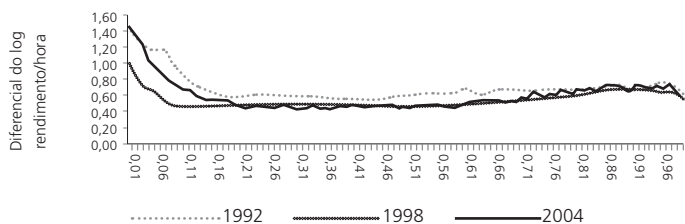


Fonte: Elaboração própria a partir dos microdados da PNAD de 1992, 1998 e 2004.

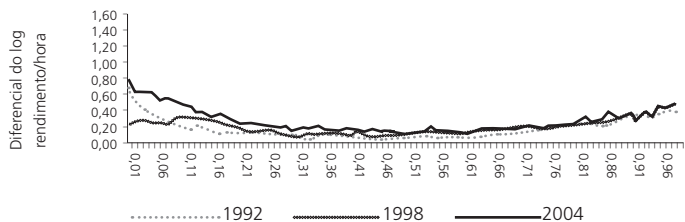
**Gráfico 4: Distribuições do log rendimento/hora, segundo setores informal2, Brasil, 1992, 1998 e 2004**



**Gráfico 5: Diferenciais por setores formal1 e informal1 ao longo das distribuições do log rendimento/hora, Brasil, 1992, 1998 e 2004**



**Gráfico 6: Diferenciais por setores formal2 e informal2 ao longo das distribuições do log rendimento/hora, Brasil, 1992, 1998 e 2004**

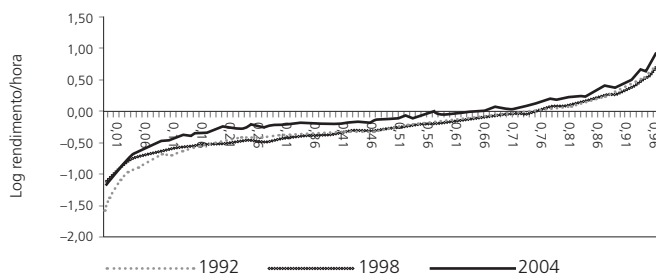


Fonte: Elaboração própria a partir dos microdados da PNAD de 1992, 1998 e 2004.

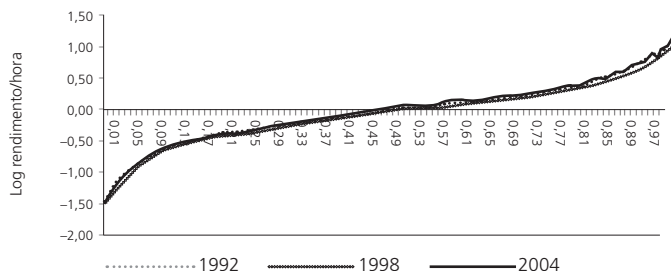
assinada na segunda definição de setor informal contribui para compensar as diferenças de rendimentos entre os dois setores, uma vez que tais ocupados tendem a se localizar nos estratos de renda mais baixos. No intervalo entre o 20º e o 60º quantil, como já verificado no gráfico da distribuição (gráfico 2), o *gap* de rendimento é quase inexistente.

Se for feita a suposição de que o setor informal percebe rendimentos idênticos ao do formal, controlando pelas características selecionadas para o trabalhador e o posto de trabalho, conforme descrito na equação 1, nota-se que a diferença é nula apenas no octogésimo quantil para os três anos quando se tem a primeira clivagem (gráfico 7), ao passo que, na segunda, esse resultado ocorre na mediana (gráfico 8).

**Gráfico 7: Diferenças entre curvas de quantil das distribuições log rendimento/hora observadas e simuladas dos setores formal1 e informal1, Brasil, 1992, 1998 e 2004**



**Gráfico 8: Diferenças entre curvas de quantil das distribuições log rendimento/hora observadas e simuladas dos setores formal2 e informal2, Brasil, 1992, 1998 e 2004**



Fonte: Elaboração própria a partir dos microdados da PNAD de 1992, 1998 e 2004.

### 3.2 Decomposição por setor

Entre os anos 1992 e 2004, decompondo os rendimentos nos quantis 10<sup>o</sup>, 25<sup>o</sup>, 50<sup>o</sup>, 75<sup>o</sup> e 90<sup>o</sup>, constata-se redução do efeito do componente não explicado sobre o rendimento em todos os quantis para a primeira definição formal-informal (tabela 1). No caso dos 10<sup>o</sup> e 25<sup>o</sup> mais pobres, essa diminuição não é monotônica, porque a parte não explicada da variação do rendimento aumenta em 1998, tornando-se negativa em 2004. Tal resultado mostra que as características não observadas entre os indivíduos passam a contribuir para estreitar o hiato de rendimentos entre os trabalhadores do setor formal e informal. Por outro lado, as diferenças de atributos pessoais, do posto de trabalho e local de residência acresceram o diferencial entre os dois setores, especialmente no 10<sup>o</sup> quantil da distribuição, em que a parte explicada retoma valores próximos aos do ano 1992.

Em termos da definição que considera informal apenas o conta própria sem o profissional liberal no setor informal (tabela 2), identifica-se reversão do diferencial no décimo mais pobre em favor dos informais no ano 2004, principalmente pela parte não explicada, posto que o componente devido às diferenças de atributos explicadas pela equação (1) é negativo nos três anos. Nos demais quantis, entre 1992 e 2004, o componente não explicado também diminui sobrepondo o aumento do diferencial entre formal e informal decorrente das diferenças entre as características observadas dos indivíduos situados nesses dois segmentos.

Embora Carneiro e Henley (2001), Menezes-Filho, Mendes e Almeida (2004) e Pianto e Pianto (2002) tenham definido setor informal de maneira diversa dessa segunda definição, é nessa clivagem que os resultados são mais

**Tabela 1: Decomposição, segundo setor formal1 e informal1, 1992, 1998 e 2004**

	Não explicada			Explicada			Total		
	1992	1998	2004	1992	1998	2004	1992	1998	2004
10 <sup>o</sup>	0,103	0,359	-0,244	0,571	0,330	0,594	0,674	0,689	0,350
25 <sup>o</sup>	0,265	0,370	-0,137	0,458	0,301	0,390	0,723	0,670	0,253
50 <sup>o</sup>	0,469	0,389	0,099	0,462	0,259	0,324	0,931	0,648	0,423
75 <sup>o</sup>	0,680	0,179	0,189	0,754	0,673	0,646	1,434	0,852	0,835
90 <sup>o</sup>	0,517	0,079	0,121	0,813	0,790	0,611	1,330	0,869	0,733

Fonte: Elaboração própria a partir dos microdados das PNADs de 1992, 1998 e 2004.

**Tabela 2: Decomposição, segundo setor formal2 e informal2, 1992, 1998 e 2004**

	Não explicada			Explicada			Total		
	1992	1998	2004	1992	1998	2004	1992	1998	2004
10 <sup>o</sup>	1,636	0,631	-0,111	-0,110	-0,127	-0,002	1,527	0,503	-0,113
25 <sup>o</sup>	1,552	1,052	0,365	0,019	-0,067	0,066	1,572	0,985	0,431
50 <sup>o</sup>	1,724	1,258	0,780	0,085	0,006	0,131	1,808	1,264	0,911
75 <sup>o</sup>	2,009	1,262	1,302	0,695	0,684	0,731	2,704	1,946	2,033
90 <sup>o</sup>	1,739	1,141	1,327	0,863	0,896	0,887	2,602	2,037	2,214

Fonte: Elaboração própria a partir dos microdados das PNADs de 1992, 1998 e 2004.

parecidos no que tange à maior dimensão do componente não explicado na determinação do diferencial de rendimentos. E, assim como em Pianto e Pianto (2002), os resultados dessa estimação evidenciam que os 10% mais ricos são afetados negativamente pelo componente não observado, ao passo que os mais pobres, positivamente.

### 3.3 Decomposição: componentes não explicado e total conforme as variáveis de controle

As tabelas 3, 4 e 5 apresentam os efeitos da decomposição do diferencial de rendimento entre setor formal1 e informal1 nos quantis selecionados conforme as características dos trabalhadores, do posto de trabalho e do local de residência nos anos 1992, 1998 e 2004, respectivamente.<sup>10</sup> Os componentes informados nas tabelas são o não explicado e o total, sendo o explicado obtido, portanto, através da subtração dos dois efeitos mostrados.

Em todos esses anos, são poucas as variáveis que chamam a atenção em termos da magnitude de seus efeitos. Assim como já evidenciado em Carneiro e Henley (2001) e Menezes-Filho, Mendes e Almeida (2004), a educação é uma característica que contribui para explicar o hiato de rendimento entre os dois setores, em especial para aqueles trabalhadores situados no topo da distribuição — o quarto e o décimo mais rico. Para os mais pobres, os do 10<sup>o</sup> quantil, o efeito é bem menor, uma vez que o componente não explicado atua no sentido de reduzir o efeito total, favorecendo os ocupados do setor informal. Esse é um achado muito semelhante ao de Maloney (1998), que aponta a escolha ocupacional dos mais pobres pela inserção no setor informal, dado que a baixa escolaridade não se constitui em barreira à entrada e, por conseguinte, em menor capacidade de auferir rendimentos



**Tabela 3: Decomposição, segundo setor formal1 e informal1, 1992**

Variáveis	Não explicada					Total				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	0,00	0,00	0,01	0,01	0,02	0,01	0,01	0,02	0,03	0,04
Integral	0,09	0,07	0,11	0,11	0,10	0,07	0,03	0,08	0,08	0,06
Idade	0,24	0,33	0,43	0,81	0,60	0,21	0,26	0,37	0,85	0,71
Idade2	-0,04	-0,07	-0,13	-0,30	-0,18	0,01	0,01	-0,05	-0,31	-0,24
Educação	-0,02	0,01	0,03	0,09	0,11	0,11	0,16	0,21	0,46	0,49
Rm	-0,02	-0,02	-0,01	-0,03	-0,03	0,00	0,01	0,01	-0,01	-0,02
Superior	0,00	0,00	0,00	0,01	0,01	0,03	0,03	0,05	0,39	0,58
Média	0,04	0,04	0,03	0,06	0,03	0,04	0,05	0,06	0,06	-0,08
Manual	0,04	0,07	0,07	0,08	0,05	0,06	0,09	0,08	-0,04	-0,09
Norte	-0,01	0,00	0,00	-0,01	-0,01	-0,02	-0,01	-0,01	-0,01	-0,01
Sudeste	-0,03	-0,02	-0,01	-0,02	-0,02	0,04	0,03	0,03	0,00	-0,01
Sul	-0,01	-0,01	-0,01	-0,01	-0,03	0,01	0,01	0,01	-0,01	-0,03
Co	-0,02	-0,02	-0,02	-0,01	-0,01	-0,02	-0,02	-0,02	-0,02	-0,01
Sexo	-0,12	-0,08	-0,01	0,00	0,01	-0,07	-0,04	0,02	0,00	0,01
Contprev	0,00	-0,01	-0,02	-0,08	-0,11	0,21	0,14	0,06	-0,01	-0,05
Casado	-0,03	-0,02	-0,01	-0,01	-0,02	-0,03	-0,02	-0,01	-0,01	-0,02
Total	0,10	0,26	0,47	0,68	0,52	0,67	0,72	0,93	1,43	1,33

Fonte: Elaboração própria a partir dos microdados da PNAD de 1992.

**Tabela 4: Decomposição, segundo setor formal1 e informal1, 1998**

Variáveis	Não explicada					Total				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	0,00	0,01	0,00	0,00	-0,01	0,00	0,02	0,02	0,02	0,01
Integral	0,01	0,05	0,13	0,15	0,16	-0,04	0,02	0,09	0,10	0,11
Idade	0,09	0,16	0,20	0,00	-0,04	-0,01	0,09	0,11	0,04	0,08
Idade2	0,05	0,02	0,00	0,10	0,16	0,14	0,08	0,07	0,09	0,09
Educação	0,02	0,03	0,06	0,10	0,15	0,13	0,17	0,23	0,46	0,50
Rm	-0,01	-0,01	-0,01	-0,02	-0,03	0,01	0,01	0,01	-0,01	-0,01
Superior	0,00	0,00	0,00	0,01	0,01	0,03	0,03	0,04	0,43	0,64
Média	0,11	0,07	0,04	0,03	-0,04	0,07	0,06	0,05	0,03	-0,13
Manual	0,17	0,10	0,03	-0,01	-0,03	0,20	0,12	0,03	-0,06	-0,08
Norte	0,00	0,00	0,00	0,00	0,00	-0,01	-0,01	-0,01	0,00	0,00
Sudeste	-0,01	-0,01	-0,01	-0,03	-0,07	0,04	0,04	0,03	-0,03	-0,07
Sul	0,00	0,00	0,00	-0,01	-0,02	0,01	0,02	0,02	-0,01	-0,03
Co	-0,01	-0,01	0,00	0,00	-0,01	0,00	0,00	-0,01	0,00	0,00
Sexo	-0,06	-0,03	-0,01	-0,02	0,01	-0,05	-0,02	0,00	-0,03	0,01
Contprev	-0,01	-0,01	-0,02	-0,13	-0,19	0,15	0,03	-0,06	-0,19	-0,26
Casado	0,01	0,00	0,00	0,00	0,01	0,01	0,00	0,00	0,00	0,01
Total	0,36	0,37	0,39	0,18	0,08	0,69	0,67	0,65	0,85	0,87

Fonte: Elaboração própria a partir dos microdados da PNAD de 1998.

**Tabela 5: Decomposição, segundo setor formal1 e informal1, 2004**

Variáveis	Não explicada					Total				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	-0,02	-0,01	-0,01	0,00	-0,03	-0,01	0,00	0,01	0,02	-0,01
Integral	-0,07	-0,04	0,01	0,06	0,06	-0,15	-0,13	-0,04	-0,01	-0,01
Idade	-0,44	-0,35	-0,11	0,11	0,24	-0,51	-0,42	-0,16	0,17	0,33
Idade2	0,29	0,25	0,15	0,07	0,00	0,35	0,30	0,20	0,04	-0,07
Educação	-0,04	0,03	0,09	0,13	0,17	0,05	0,14	0,22	0,41	0,45
Rm	0,00	0,00	0,00	0,01	0,01	0,01	0,00	0,00	0,02	0,02
Superior	0,00	0,00	0,00	0,02	0,02	0,04	0,04	0,05	0,43	0,52
Média	0,01	0,00	0,01	0,01	-0,01	0,03	0,04	0,06	0,04	-0,03
Manual	0,21	0,12	0,07	0,01	-0,04	0,20	0,11	0,06	-0,05	-0,11
Norte	-0,03	-0,03	-0,02	-0,01	-0,01	-0,03	-0,03	-0,02	-0,02	-0,02
Sudeste	-0,04	-0,03	-0,02	-0,04	-0,04	-0,02	0,00	0,00	-0,03	-0,04
Sul	-0,01	-0,01	-0,01	-0,02	-0,02	0,00	0,00	0,01	-0,02	-0,02
Co	-0,02	-0,02	-0,02	-0,01	-0,02	-0,01	-0,01	-0,01	-0,01	-0,01
Sexo	-0,08	-0,05	-0,03	-0,03	-0,02	-0,05	-0,02	-0,01	-0,04	-0,02
contprev	0,00	0,00	-0,01	-0,11	-0,19	0,44	0,22	0,06	-0,11	-0,24
Casado	0,01	0,00	-0,01	-0,01	-0,01	0,01	0,00	-0,01	-0,01	-0,01
Total	-0,24	-0,14	0,10	0,19	0,12	0,35	0,25	0,42	0,83	0,73

Fonte: Elaboração própria a partir dos microdados da PNAD de 2004.

nesse setor relativamente ao formal, que tem a educação formal como um dos critérios de entrada.

A idade é uma outra variável importante para explicar o diferencial de rendimentos, principalmente pelo componente não explicado. No entanto, os sinais dos efeitos se diferenciam ao longo dos anos. Em 1992, o efeito é positivo e recrudescer ao longo dos quantis até o 75°, inclusive (tabela 3). Os maiores valores dos retornos de idade são para os dois últimos quantis da distribuição. No ano 1998, idade perde importância para explicar o diferencial (tabela 4) e, em 2004, o efeito do componente não explicado passa a ser negativo, contribuindo para estreitar o hiato de rendimentos entre setor formal e informal. Contudo, o impacto da experiência descrito indiretamente pela variável idade ao quadrado (idade 2 nas tabelas) mostra ganhos de renda não explicados, ao longo do ciclo de vida, para o setor formal em face do informal entre os de menor renda (tabela 5).

As categorias de ocupação (Superior, Média e Manual) retratam efeitos distintos no tempo e pelos quantis em comparação à de referência, a das

ocupações domésticas. Essa categorização está fortemente associada à educação. O grupo Superior explica o diferencial nos quantis 75° e 90° em 1992 (tabela 3). A categoria Manual determina o diferencial positivo entre formal e informal, pelo componente não explicado entre os mais pobres em 1998 e 2004, respectivamente nas tabelas 4 e 5. Surpreende esse resultado porque é justamente nessas categorias que o setor informal pode, em termos de rendimentos, criar vantagens comparativas para os trabalhadores menos escolarizados, e isso é identificado pelo efeito da variável educação.

Contribuir para a previdência é uma característica importante para determinação do rendimento. O efeito do componente não explicado é negativo, em favor do setor informal, para os 10% mais ricos nos três anos. Todavia, para os situados nos quantis 10° e 25°, o componente explicado é maior e positivo, beneficiando os ocupados no setor formal. Ou seja, os mais ricos, por perceberem rendimentos mais altos, preferem a não-contribuição, posto que o custo seria mais elevado e, portanto, o rendimento líquido menor se participassem da seguridade social. Por outro lado, os trabalhadores situados na base da distribuição ficam mais protegidos no setor formal, porque grande parcela do encargo é de responsabilidade do empregador, o mesmo não ocorrendo com os informais.

Por fim, trabalhar em jornadas extensas (integral) favorece, pelo componente não explicado, os formais em 1992 (tabela 3), e ainda favorece os formais em 1998, somente a partir da mediana da distribuição (tabela 4), e, pelos componentes explicado e não explicado, os trabalhadores do informal no 10° e 25° quantil da distribuição de rendimentos em 2004 (tabela 5).

Nas tabelas 6, 7 e 8 têm-se as mesmas decomposições reportadas pelas tabelas 3, 4 e 5 para a segunda definição de formal-informal, a que considera apenas o trabalhador por conta própria no setor informal.

Ao contrário da outra definição de formal-informal, a escolaridade não é relevante para explicar o diferencial entre o formal e os trabalhadores por conta própria a não ser para os situados no último quarto e no último décimo da distribuição, em que um ano a mais de estudo favorece os formais. Desse modo, a hipótese da estratégia ocupacional por escolha de vida fica mais nítida para esses trabalhadores do informal, a exemplo do sugerido por Maloney (1998).

**Tabela 6: Decomposição, segundo setor formal2 e informal2, 1992**

Variáveis	Não explicada					Total				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	-0,01	-0,01	-0,01	-0,01	-0,01	-0,01	0,00	-0,01	0,00	0,00
Integral	0,03	0,04	0,05	0,06	0,06	-0,01	-0,01	0,01	0,03	0,01
Idade	0,99	0,69	0,71	0,85	0,48	0,50	0,28	0,32	0,75	0,49
Idade2	-0,43	-0,25	-0,26	-0,34	-0,13	-0,01	0,08	0,05	-0,24	-0,12
Educação	-0,02	0,00	0,02	0,07	0,11	0,02	0,05	0,10	0,38	0,49
Rm	0,02	0,01	0,02	0,01	0,00	0,01	0,02	0,02	0,03	0,02
Superior	0,00	0,00	0,00	0,01	0,01	0,01	0,02	0,03	0,32	0,50
Média	0,45	0,43	0,41	0,63	0,63	0,33	0,34	0,35	0,61	0,54
Manual	0,66	0,69	0,81	0,73	0,54	0,63	0,68	0,80	0,69	0,49
Norte	0,02	0,01	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,00
Sudeste	0,00	0,00	0,01	0,01	0,01	0,02	0,04	0,04	0,02	0,03
Sul	0,01	0,00	0,01	0,00	0,01	0,01	0,02	0,02	0,01	0,00
Co	0,00	-0,01	0,00	0,00	-0,01	0,02	0,01	0,00	0,00	0,00
Sexo	-0,12	-0,10	-0,04	0,01	0,04	-0,12	-0,10	-0,07	-0,01	0,03
Contprev	0,01	0,01	0,00	0,00	-0,01	0,08	0,13	0,15	0,14	0,13
Casado	0,04	0,03	0,01	0,00	0,00	0,04	0,03	0,01	0,00	0,00
Total	1,64	1,55	1,72	2,01	1,74	1,53	1,57	1,81	2,70	2,60

Fonte: Elaboração própria a partir dos microdados da PNAD de 1992.

**Tabela 7: Decomposição, segundo setor formal2 e informal2, 1998**

Variáveis	Não explicada					Total				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	-0,02	-0,01	-0,01	-0,02	-0,02	-0,02	-0,01	-0,01	-0,01	-0,01
Integral	-0,05	-0,06	0,00	0,02	0,01	-0,12	-0,12	-0,06	-0,02	-0,05
Idade	-0,03	0,36	0,63	0,61	0,62	-0,45	-0,04	0,24	0,50	0,59
Idade2	0,13	-0,09	-0,21	-0,18	-0,16	0,47	0,22	0,08	-0,09	-0,13
Educação	-0,01	-0,02	0,00	0,02	0,08	0,04	0,05	0,10	0,33	0,45
Rm	0,02	0,02	0,03	0,03	0,01	0,02	0,02	0,04	0,05	0,03
Superior	0,00	0,00	0,00	0,01	0,02	0,01	0,01	0,03	0,43	0,68
Média	0,36	0,41	0,33	0,41	0,33	0,24	0,31	0,27	0,38	0,23
Manual	0,33	0,51	0,51	0,39	0,26	0,33	0,51	0,51	0,36	0,23
Norte	-0,01	-0,01	0,00	0,00	0,00	-0,01	-0,01	-0,01	0,00	0,00
Sudeste	-0,01	-0,01	-0,01	0,00	-0,01	0,02	0,02	0,03	0,00	-0,01
Sul	0,00	0,00	0,00	0,01	0,00	0,00	0,00	0,01	0,00	-0,01
Co	-0,01	-0,01	-0,01	0,00	-0,01	0,01	0,01	0,00	0,01	0,00
Sexo	-0,09	-0,05	0,00	0,03	0,10	-0,12	-0,09	-0,05	0,00	0,07
Contprev	0,00	-0,01	-0,01	-0,07	-0,11	0,07	0,09	0,09	0,01	-0,05
Casado	0,01	0,02	0,01	0,01	0,02	0,01	0,02	0,01	0,01	0,02
Total	0,63	1,05	1,26	1,26	1,14	0,50	0,98	1,26	1,95	2,04

Fonte: Elaboração própria a partir dos microdados da PNAD de 1998.

**Tabela 8: Decomposição, segundo setor formal2 e informal2, 2004**

Variáveis	Não explicada					Total				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	-0,02	-0,03	-0,03	-0,01	-0,05	-0,02	-0,02	-0,02	0,00	-0,04
Integral	-0,08	-0,05	-0,03	-0,01	0,04	-0,15	-0,12	-0,09	-0,07	-0,05
Idade	-0,95	-0,51	-0,20	0,40	0,59	-1,25	-0,79	-0,50	0,30	0,58
Idade2	0,62	0,37	0,24	-0,04	-0,12	0,84	0,56	0,43	0,03	-0,10
Educação	-0,07	-0,05	-0,02	0,01	0,05	-0,01	0,02	0,08	0,26	0,35
Rm	0,04	0,03	0,04	0,05	0,05	0,03	0,03	0,04	0,06	0,06
Superior	0,00	0,00	0,00	0,03	0,05	0,02	0,03	0,04	0,49	0,71
Média	0,01	0,02	0,02	0,14	0,19	0,02	0,04	0,07	0,19	0,18
Manual	0,53	0,72	0,82	0,80	0,59	0,49	0,69	0,80	0,75	0,53
Norte	-0,03	-0,02	-0,01	-0,01	-0,01	-0,03	-0,03	-0,02	-0,01	-0,01
Sudeste	-0,03	-0,02	-0,01	-0,01	-0,01	-0,02	0,00	0,02	-0,01	-0,01
Sul	-0,01	-0,01	0,00	-0,01	-0,02	-0,02	-0,01	0,01	0,00	-0,02
Co	-0,02	-0,01	-0,01	-0,01	-0,01	0,00	0,00	0,00	0,00	0,00
Sexo	-0,12	-0,08	-0,03	0,03	0,07	-0,14	-0,10	-0,07	0,00	0,04
Contprev	0,00	0,00	-0,01	-0,06	-0,11	0,08	0,13	0,13	0,05	-0,02
Casado	0,03	0,01	0,00	0,01	0,01	0,03	0,01	0,00	0,01	0,01
Total	-0,11	0,37	0,78	1,30	1,33	-0,11	0,43	0,91	2,03	2,21

Fonte: Elaboração própria a partir dos microdados da PNAD de 2004.

Assim como para a definição formal1-informal1, a idade é a variável de maior importância e os sinais são distintos ao longo dos anos. O componente não explicado é positivo e crescente em 1992 (tabela 6), e negativo até a mediana em 2004 (tabela 7). No primeiro ano, contribui para ampliar o hiato de rendimentos entre formais e informais e, no último, sua contribuição é no sentido inverso. Ademais, o efeito da experiência medido pela idade 2 segue o seguinte comportamento: ganhos de renda não explicados, ao longo do ciclo de vida, para o setor informal em face do formal na base da distribuição em 1992 (tabela 6) e ganhos de renda não explicados para o formal nos estratos de renda mais pobres em 2004 (tabela 8).

Em relação às categorias ocupacionais, nota-se que o diferencial é positivo apenas naqueles condicionados ao 75° e 90° quantil, sendo irrelevante nos demais em todos os anos. As categorias Média e Manual apresentam um componente não explicado elevado e positivo em relação à de domésticas em 1998, responsável quase integralmente pelo efeito total do diferencial (tabela 7). Em 2004, somente a categoria Manual apresenta esse mesmo resultado (tabela 8). Tal evidência chama a atenção para o efeito não explicado

do sexo entre os 10% mais pobres. Particularmente nesse ano, trata-se de um componente relevante com sinal negativo, o que mostra que o diferencial nesse caso favorece os informais, como as ocupações domésticas integram a categoria de referência e são tipicamente femininas, a expressiva magnitude das categorias Média e, fundamentalmente, Manual pode estar sendo influenciada por um efeito indireto criado pelos autônomos do sexo feminino, que, provavelmente, encontram-se em grande parte na categoria doméstica, criando uma interação entre categorias ocupacionais e sexo.

No caso dos trabalhadores em jornada integral, outra variável relevante na determinação do diferencial, verifica-se que trabalhar nesse contexto favorece, do ponto de vista do rendimento, os trabalhadores por conta própria condicionados ao 10º e ao 25º quantil da distribuição em 1998 (tabela 7) e, em 2004 (tabela 8), pelos efeitos componentes explicado e não explicado.

### 3.4 Decomposição ao longo do tempo

As tabelas 9 e 10 reportam as informações referentes à decomposição dos coeficientes das variáveis controladas na equação de rendimentos (1),<sup>11</sup> a cada combinação de dois anos dos três que compõem o período selecionado: 1992-1998, 1992-2004, 1998-2004. A terceira coluna dessas tabelas —  $(DX_t^f - DX_t^i)bt$  — mostra o efeito das mudanças relativas entre os dois anos nas características observadas dos trabalhadores do formal e do informal pelos quantis. Na quarta, apresenta-se o efeito de mudanças no tempo nos coeficientes do setor formal, mantendo fixas as diferenças nas características observadas. A quinta coluna traz as alterações nos coeficientes relativos entre os dois setores nos dois anos e, por fim, a última coluna mostra como mudanças nas características do setor informal afetam os retornos nos dois setores. A soma da quinta e da sexta coluna informa o efeito de alterações do componente não explicado do hiato de rendimentos entre setor formal e informal.

Considerando a definição de informal que abrange empregado por conta própria e sem carteira (tabela 9), constata-se uma mudança composicional, favorecendo os trabalhadores do setor informal, com exceção apenas para os trabalhadores dos três primeiros quantis do período 1998-2004, em que os beneficiados são os trabalhadores do setor formal (terceira coluna).

**Tabela 9: Decomposição, segundo setor formal1 e informal1, ao longo do tempo**

		$(DXt' - DXt)b1t$	$DXt'(b1t' - b1t)$	$(Dbt' - Dbt)X2t$	$(X2t' - X2t)Dbt'$
1992-1998	10°	-0,142	-0,098	0,218	0,037
	25°	-0,055	-0,102	0,090	0,015
	50°	-0,074	-0,128	-0,102	0,022
	75°	-0,022	-0,058	-0,536	0,034
	90°	-0,013	-0,010	-0,488	0,050
1992-2004	10°	-0,063	0,086	-0,390	0,043
	25°	-0,050	-0,018	-0,440	0,039
	50°	-0,046	-0,092	-0,416	0,045
	75°	-0,034	-0,073	-0,556	0,064
	90°	-0,133	-0,069	-0,479	0,084
1998-2004	10°	0,078	0,185	-0,649	0,046
	25°	0,002	0,088	-0,546	0,039
	50°	0,035	0,031	-0,316	0,026
	75°	-0,016	-0,011	-0,011	0,021
	90°	-0,133	-0,046	0,015	0,027

Fonte: Elaboração própria a partir dos microdados das PNADs de 1992, 1998 e 2004.

**Tabela 10: Decomposição, segundo setor formal2 e informal2, ao longo do tempo**

		$(DXt' - DXt)b1t$	$DXt'(b1t' - b1t)$	$(Dbt' - Dbt)X2t$	$(X2t' - X2t)Dbt'$
1992-1998	10°	0,031	-0,049	-1,008	0,002
	25°	-0,007	-0,079	-0,504	0,003
	50°	-0,004	-0,075	-0,477	0,011
	75°	-0,005	-0,006	-0,777	0,030
	90°	0,007	0,026	-0,642	0,043
1992-2004	10°	0,128	-0,021	-1,778	0,031
	25°	0,124	-0,077	-1,224	0,036
	50°	0,133	-0,087	-1,001	0,057
	75°	0,107	-0,071	-0,784	0,076
	90°	0,047	-0,023	-0,495	0,082
1998-2004	10°	0,122	0,003	-0,812	0,071
	25°	0,136	-0,004	-0,739	0,053
	50°	0,149	-0,024	-0,533	0,054
	75°	0,121	-0,074	-0,007	0,047
	90°	0,039	-0,049	0,149	0,037

Fonte: Elaboração própria a partir dos microdados das PNADs de 1992, 1998 e 2004.

Além disso, observa-se que, tanto em 1992-1998 quanto em 1992-2004, para quase todos os quantis — exceto o 10º no período 1992-2004 —, o efeito da variação relativa das características observadas e dos retornos do setor formal, fixando a mudança dessas características, é negativo, ou seja, as mudanças respaldam os trabalhadores do setor informal.

O componente não explicado é, também, favorável aos trabalhadores do setor informal, ressaltando os quantis 10º e 25º no intervalo temporal 1992-1998. Já no período 1998-2004, as mudanças nas características observadas contribuem para ampliar o hiato de rendimentos formal-informal, com exceção dos quantis referentes aos trabalhadores no topo da distribuição de rendimentos (75º e 90º). O contrário ocorre com o efeito do componente não explicado, sendo, nos quantis inferiores, o valor de tal magnitude que compensa o efeito positivo das mudanças das características observadas entre os dois setores, possibilitando que o diferencial se torne favorável ao setor informal.

Utilizando, no entanto, a segunda definição de informal, a que considera apenas o trabalhador por conta própria, verifica-se que o efeito das mudanças relativas das características observadas favorece os trabalhadores do setor formal em todos os quantis dos períodos 1992-2004 e 1998-2004 (tabela 10). No intervalo 1992-1998, esse resultado só ocorre para o 10º quantil. Por outro lado, atribuir a variação dos retornos do formal aos trabalhadores do informal, fixando as mudanças das características observadas, reduz o hiato de rendimentos em favor dos informais. Entretanto, a soma desses efeitos —  $(DX_t^p - DX_t) b_{1t}$  mais  $DX_t'(b_{1t}' - b_{1t})$  — é positiva em todos os quantis dos períodos 1992-2004 e 1998-2004, com exceção do 90º. Esta é, também, uma exceção para o componente não explicado, porém, em valor positivo. Como o componente não explicado é negativo, e em uma magnitude que suplanta o explicado, há reversão do diferencial em favor dos trabalhadores por conta própria, assim como verificado para o informal<sup>1</sup>. Tal reversão seria em termos dos componentes não explicados, ou seja, dos diferenciais de retornos às características observadas, predominantemente em sua variação ao longo do tempo.



#### 4. CONSIDERAÇÕES FINAIS

Este artigo investigou a evolução do diferencial de rendimentos no Brasil entre o setor formal e informal. Foi utilizado um método semiparamétrico, regressão quantílica, que permite avaliar como se comportam os determinantes do rendimento, tanto em termos de quantis como também ao longo do tempo. Ao se estimarem diferentes regressões quantílicas, exploraram-se diferentes partes da distribuição condicional. Desse modo, captou-se o efeito de características observadas e não observadas, em cada ponto da distribuição, constatando-se diferenças expressivas.

Além disso, o recurso a duas tipologias para definição dos setores formal e informal, conforme a posição na ocupação do trabalhador, permitiu verificar comportamento diferenciado do hiato de rendimento entre a definição mais abrangente (empregado sem carteira e trabalhador por conta própria) e a mais restrita (trabalhador por conta própria).

Assim sendo, os diferenciais de rendimento entre o setor formal e informal tendem a ser maiores para a primeira definição, evidenciando que a exclusão dos empregados sem carteira assinada na segunda definição de setor informal contribui para compensar as diferenças de rendimentos entre os dois setores, uma vez que tais empregados tendem a se localizar nos estratos de renda mais baixos.

As características não observadas, englobadas pelo componente não explicado na determinação do diferencial de rendimentos, afetaram negativamente os 10% mais ricos e positivamente os mais pobres do informal nas duas definições, em especial na segunda, que apenas considera o trabalhador por conta própria nos três anos selecionados.

Somando-se a isso, ao se considerarem as variáveis de controle na análise, constatou-se que variáveis como educação, idade, experiência — descrita indiretamente pela variável idade ao quadrado —, contribuição para a previdência e trabalho em jornada extensiva (integral) são importantes para explicar o hiato de rendimento entre os dois setores. Por outro lado, idade e jornada extensa, pelo efeito do componente não explicado, reduziram esse hiato em 2004, uma vez mais de maneira mais intensa a favor dos informais na segunda definição.

A maior diferença, no entanto, ocorreu na variável escolaridade, que não se mostrou relevante para explicar o diferencial entre o formal<sup>2</sup> e informal<sup>2</sup> (trabalhadores por conta própria), a não ser para os situados no último quarto e no último décimo da distribuição, em que um ano a mais de estudo favorece os formais.

Na decomposição ao longo do tempo, no período 1998-2004, as mudanças nas características observadas contribuíram para ampliar o diferencial formal-informal, com exceção dos percentis referentes aos trabalhadores no topo da distribuição de rendimentos (75° e 90°). O contrário ocorreu com o efeito do componente não explicado. E esse efeito foi tão forte, nos percentis inferiores, que compensou o efeito positivo das mudanças das características observadas entre os dois setores, possibilitando que o diferencial se tornasse favorável ao setor informal em suas duas definições.

Concluindo, uma vez que os diferenciais de rendimentos foram controlados por diversas características ocupacionais, acredita-se que os ocupados do setor informal nos quantis da base da distribuição, particularmente na segunda definição, em que são considerados somente os “conta própria”, detenham, realmente, retornos mais elevados por suas características “subjetivas”. Ratificando a hipótese inicial, a decisão de se auto-empregar é, portanto, fruto de uma estratégia. Os trabalhadores menos escolarizados percebem rendimentos mais elevados como autônomos do que como assalariados com carteira assinada, devido a questões como o histórico familiar na atividade, a experiência profissional pregressa como assalariado, a flexibilidade da jornada de trabalho, entre outras, que contribuem para definir essa inserção. E essa estratégia identificada em características não observadas (com exceção da escolaridade) determina fortemente a definição do informal como um todo no Brasil hoje, haja vista a evolução no período.

## ANEXOS

Tabela A1: Média das variáveis, setores formal1 e informal1, Brasil, 1992

Variáveis	Formal1					Informal1				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	0,42	0,46	0,52	0,77	0,83	0,31	0,34	0,39	0,65	0,72
Integral	0,93	0,96	0,92	0,78	0,75	0,82	0,81	0,82	0,70	0,65
Idade	33,15	32,71	32,98	37,67	39,29	33,72	34,12	34,03	37,03	37,86
Idade2	1.236,15	1.201,40	1.210,77	1.507,58	1.629,86	1.320,84	1.334,41	1.319,85	1.495,24	1.550,31
Educação	5,00	5,32	6,00	11,32	12,49	3,11	3,40	3,85	7,20	8,41
Rm	0,45	0,47	0,49	0,60	0,62	0,29	0,32	0,36	0,51	0,54
Superior	0,06	0,05	0,06	0,43	0,59	0,00	0,00	0,00	0,05	0,08
Média	0,15	0,16	0,22	0,37	0,29	0,14	0,15	0,15	0,37	0,46
Manual	0,66	0,70	0,67	0,20	0,12	0,54	0,60	0,63	0,56	0,46
Norte	0,07	0,06	0,06	0,06	0,06	0,10	0,10	0,10	0,08	0,08
Sudeste	0,33	0,37	0,39	0,47	0,46	0,18	0,23	0,28	0,42	0,42
Sul	0,13	0,17	0,20	0,20	0,20	0,08	0,10	0,12	0,20	0,22
Co	0,10	0,10	0,10	0,11	0,12	0,09	0,11	0,12	0,12	0,12
Sexo	0,56	0,56	0,59	0,69	0,71	0,38	0,45	0,51	0,69	0,71
Contprev	0,90	0,94	0,95	0,92	0,92	0,03	0,04	0,07	0,34	0,43
Casado	0,27	0,29	0,27	0,21	0,20	0,38	0,36	0,33	0,24	0,22

Fonte: Elaboração própria a partir dos microdados da PNAD de 1992.

Tabela A2: Média das variáveis, setores formal1 e informal1, Brasil, 1998

Variáveis	Formal1					Informal1				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	0,37	0,43	0,50	0,78	0,84	0,30	0,33	0,38	0,67	0,73
Integral	0,99	0,95	0,93	0,78	0,74	0,83	0,86	0,82	0,66	0,63
Idade	33,08	32,84	33,16	39,12	41,08	34,97	34,18	34,62	38,52	39,26
Idade2	1.228,33	1.204,33	1.218,59	1.620,08	1.772,87	1.404,11	1.333,97	1.355,42	1.607,20	1.658,96
Educação	5,45	5,97	6,64	12,14	13,24	3,78	4,18	4,66	8,22	9,55
Rm	0,43	0,45	0,48	0,60	0,64	0,32	0,32	0,35	0,54	0,57
Superior	0,04	0,04	0,05	0,50	0,71	0,00	0,00	0,01	0,07	0,12
Média	0,10	0,13	0,19	0,34	0,21	0,18	0,16	0,16	0,36	0,44
Manual	0,68	0,71	0,68	0,15	0,07	0,57	0,60	0,64	0,53	0,41
Norte	0,07	0,07	0,06	0,06	0,06	0,13	0,13	0,12	0,08	0,08
Sudeste	0,29	0,33	0,37	0,43	0,42	0,15	0,20	0,25	0,42	0,43
Sul	0,10	0,15	0,19	0,21	0,20	0,09	0,09	0,12	0,21	0,22
Co	0,10	0,12	0,11	0,13	0,14	0,08	0,11	0,12	0,11	0,11
Sexo	0,49	0,52	0,56	0,64	0,68	0,44	0,49	0,53	0,67	0,69
Contprev	0,96	0,97	0,97	0,91	0,90	0,03	0,04	0,06	0,29	0,38
Casado	0,32	0,32	0,30	0,25	0,22	0,38	0,35	0,33	0,25	0,23

Fonte: Elaboração própria a partir dos microdados da PNAD de 1998.

**Tabela A3: Média das variáveis, setores formal1 e informal1, Brasil, 2004**

Variáveis	Formal1					Informal1				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	0,37	0,41	0,48	0,73	0,77	0,29	0,32	0,35	0,60	0,67
Integral	0,98	0,99	0,95	0,78	0,74	0,76	0,76	0,80	0,61	0,59
Idade	32,42	32,85	33,50	40,04	41,17	34,41	34,67	34,58	38,90	39,83
Idade2	1.169,75	1.198,88	1.240,87	1.700,57	1.828,77	1.348,27	1.358,32	1.344,97	1.640,51	1.709,40
Educação	7,01	7,32	7,86	12,47	13,42	5,14	5,47	5,91	9,01	10,28
Rm	0,45	0,44	0,45	0,53	0,58	0,38	0,39	0,39	0,48	0,50
Superior	0,08	0,06	0,07	0,58	0,71	0,00	0,01	0,01	0,14	0,22
Média	0,13	0,17	0,22	0,27	0,21	0,04	0,05	0,07	0,20	0,25
Manual	0,64	0,65	0,63	0,15	0,07	0,72	0,72	0,71	0,60	0,51
Norte	0,11	0,11	0,10	0,09	0,08	0,11	0,14	0,15	0,12	0,11
Sudeste	0,28	0,29	0,33	0,38	0,38	0,17	0,20	0,24	0,36	0,37
Sul	0,10	0,14	0,19	0,22	0,21	0,06	0,08	0,10	0,19	0,19
Co	0,10	0,11	0,11	0,14	0,16	0,06	0,08	0,10	0,14	0,14
Sexo	0,52	0,51	0,54	0,61	0,63	0,38	0,43	0,48	0,61	0,64
Contprev	0,93	0,96	0,96	0,91	0,90	0,01	0,02	0,05	0,28	0,37
Casado	0,33	0,35	0,33	0,27	0,25	0,42	0,39	0,37	0,29	0,26

Fonte: Elaboração própria a partir dos microdados da PNAD de 2004.

**Tabela A4: Média das variáveis, setores formal2 e informal2, Brasil, 1992**

Variáveis	Formal2					Informal2				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	0,34	0,40	0,46	0,75	0,81	0,33	0,38	0,43	0,68	0,73
Integral	0,87	0,90	0,90	0,78	0,75	0,68	0,71	0,76	0,70	0,63
Idade	31,50	31,78	32,20	37,01	38,77	40,38	39,18	38,81	38,56	38,64
Idade2	1.142,85	1.151,49	1.169,68	1.463,46	1.591,12	1.800,73	1.693,26	1.656,61	1.603,88	1.606,64
Educação	3,71	4,30	5,03	10,80	12,24	3,14	3,56	4,03	7,22	8,10
Rm	0,31	0,38	0,43	0,59	0,62	0,34	0,37	0,40	0,51	0,53
Superior	0,02	0,02	0,03	0,38	0,54	0,00	0,00	0,00	0,01	0,01
Média	0,07	0,09	0,14	0,38	0,32	0,33	0,32	0,30	0,44	0,52
Manual	0,56	0,63	0,65	0,24	0,14	0,66	0,68	0,70	0,56	0,47
Norte	0,09	0,08	0,07	0,06	0,06	0,12	0,11	0,11	0,07	0,07
Sudeste	0,23	0,31	0,35	0,46	0,46	0,18	0,22	0,27	0,43	0,41
Sul	0,10	0,13	0,17	0,21	0,20	0,09	0,11	0,13	0,21	0,23
Co	0,11	0,12	0,12	0,11	0,12	0,06	0,08	0,10	0,10	0,11
Sexo	0,44	0,51	0,54	0,68	0,71	0,44	0,53	0,60	0,72	0,72
Contprev	0,19	0,39	0,56	0,88	0,89	0,06	0,10	0,14	0,42	0,47
Casado	0,34	0,31	0,30	0,22	0,20	0,41	0,35	0,30	0,22	0,22

Fonte: Elaboração própria a partir dos microdados da PPNAD de 1992.

**Tabela A5: Média das variáveis, setores formal2 e informal2, Brasil, 1998**

Variáveis	Formal2					Informal2				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	0,32	0,37	0,44	0,75	0,82	0,31	0,37	0,42	0,70	0,74
Integral	0,93	0,94	0,90	0,77	0,73	0,75	0,78	0,76	0,67	0,63
Idade	31,43	31,89	32,38	38,39	40,23	40,50	39,92	39,53	40,26	40,72
Idade2	1.129,90	1.152,31	1.177,23	1.569,45	1.709,13	1.804,13	1.748,83	1.706,08	1.735,32	1.768,75
Educação	4,40	5,02	5,79	11,60	12,88	3,61	4,12	4,66	8,10	8,90
Rm	0,33	0,37	0,42	0,59	0,63	0,38	0,37	0,38	0,52	0,55
Superior	0,02	0,02	0,03	0,43	0,63	0,00	0,00	0,00	0,01	0,02
Média	0,04	0,08	0,13	0,36	0,26	0,38	0,34	0,30	0,43	0,51
Manual	0,57	0,65	0,66	0,20	0,10	0,62	0,66	0,70	0,56	0,47
Norte	0,12	0,10	0,09	0,06	0,06	0,13	0,13	0,12	0,08	0,08
Sudeste	0,23	0,27	0,32	0,43	0,43	0,14	0,19	0,24	0,43	0,43
Sul	0,08	0,11	0,16	0,21	0,20	0,09	0,11	0,13	0,22	0,22
Co	0,11	0,13	0,12	0,13	0,14	0,07	0,08	0,09	0,10	0,10
Sexo	0,44	0,48	0,52	0,64	0,67	0,53	0,58	0,62	0,72	0,73
Contprev	0,27	0,44	0,57	0,84	0,86	0,04	0,07	0,10	0,35	0,41
Casado	0,34	0,34	0,33	0,25	0,23	0,36	0,33	0,29	0,22	0,21

Fonte: Elaboração própria a partir dos microdados da PNAD de 1998.

**Tabela A6: Média das variáveis, setores formal2 e informal2, Brasil, 2004**

Variáveis	Formal2					Informal2				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	0,32	0,35	0,42	0,70	0,75	0,30	0,33	0,37	0,63	0,69
Integral	0,84	0,90	0,89	0,76	0,73	0,66	0,69	0,73	0,63	0,58
Idade	31,58	31,98	32,69	39,30	41,15	39,59	39,42	39,41	41,07	41,36
Idade2	1.126,70	1.147,80	1.192,21	1.647,84	1.790,72	1.729,84	1.707,85	1.696,52	1.803,42	1.828,10
Educação	5,93	6,53	7,14	12,06	13,19	4,79	5,15	5,56	8,61	9,51
Rm	0,37	0,39	0,42	0,52	0,56	0,44	0,44	0,43	0,46	0,49
Superior	0,04	0,04	0,05	0,52	0,68	0,00	0,00	0,00	0,03	0,05
Média	0,07	0,12	0,17	0,28	0,23	0,03	0,03	0,03	0,17	0,25
Manual	0,54	0,58	0,60	0,19	0,09	0,97	0,97	0,96	0,80	0,70
Norte	0,13	0,13	0,12	0,09	0,09	0,11	0,14	0,15	0,12	0,10
Sudeste	0,21	0,25	0,30	0,38	0,37	0,17	0,19	0,23	0,37	0,38
Sul	0,08	0,09	0,15	0,21	0,20	0,08	0,10	0,11	0,20	0,21
Co	0,09	0,10	0,12	0,14	0,16	0,05	0,07	0,09	0,13	0,13
Sexo	0,40	0,45	0,48	0,61	0,62	0,45	0,52	0,59	0,67	0,70
Contprev	0,18	0,42	0,57	0,85	0,86	0,02	0,04	0,06	0,29	0,36
Casado	0,39	0,37	0,36	0,28	0,26	0,43	0,37	0,32	0,26	0,23

Fonte: Elaboração própria a partir dos microdados da PNAD de 2004.

### Coeficientes das regressões quantílicas

**Tabela A7: Regressão quantílica, setores formal1 e informal1, Brasil, 1992**

Variáveis	Formal1					Informal1				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	0,131	0,123	0,139	0,158	0,197	0,136	0,124	0,124	0,136	0,166
Integral	-0,195	-0,246	-0,271	-0,322	-0,408	-0,309	-0,330	-0,402	-0,474	-0,566
Idade	0,054	0,056	0,062	0,073	0,074	0,047	0,046	0,049	0,051	0,058
Idade2	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001
Educação	0,067	0,075	0,084	0,090	0,093	0,072	0,071	0,075	0,077	0,080
Rm	0,141	0,153	0,187	0,182	0,190	0,225	0,213	0,214	0,244	0,247
Superior	0,576	0,702	0,855	0,998	1,099	0,574	0,608	0,709	0,876	0,929
Média	0,363	0,418	0,499	0,581	0,631	0,117	0,157	0,300	0,425	0,569
Manual	0,175	0,201	0,263	0,341	0,399	0,098	0,088	0,145	0,196	0,300
Norte	0,279	0,233	0,210	0,177	0,159	0,343	0,277	0,253	0,256	0,274
Sudeste	0,430	0,373	0,352	0,296	0,227	0,575	0,482	0,397	0,336	0,282
Sul	0,389	0,325	0,275	0,208	0,111	0,501	0,428	0,349	0,281	0,229
Co	0,350	0,265	0,228	0,177	0,144	0,546	0,425	0,356	0,284	0,258
Sexo	0,297	0,356	0,433	0,482	0,500	0,622	0,528	0,460	0,486	0,486
Contprev	0,250	0,164	0,092	0,112	0,121	0,364	0,376	0,358	0,351	0,387
Casado	-0,018*	-0,007*	-0,008*	-0,027*	-0,027*	0,067	0,049	0,015*	0,028*	0,060
_cons	-2,702	-2,388	-2,224	-2,143	-1,809	-3,032	-2,396	-1,935	-1,544	-1,312

Fonte: Elaboração própria a partir dos microdados da PNAD de 1992.

Nota: Significante no nível de: 10%; \*Não significantes.

**Tabela A8: Regressão quantílica, setores formal1 e informal1, Brasil, 1998**

Variáveis	Formal1					Informal1				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	0,132	0,143	0,141	0,155	0,169	0,145	0,127	0,138	0,148	0,181
Integral	-0,320	-0,329	-0,328	-0,369	-0,439	-0,336	-0,393	-0,482	-0,589	-0,698
Idade	0,049	0,051	0,057	0,058	0,063	0,046	0,047	0,051	0,058	0,064
Idade2	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001
Educação	0,070	0,078	0,088	0,092	0,096	0,065	0,070	0,075	0,080	0,080
Rm	0,148	0,170	0,184	0,188	0,210	0,172	0,205	0,224	0,225	0,259
Superior	0,731	0,805	0,883	0,974	1,070	0,498	0,576	0,740	0,890	0,981
Média	0,474	0,466	0,442	0,417	0,411	-0,144	0,042	0,185	0,324	0,502
Manual	0,236	0,196	0,156	0,134	0,149	-0,066	0,034	0,113	0,145	0,212
Norte	0,161	0,213	0,233	0,275	0,305	0,197	0,182	0,207	0,222	0,268
Sudeste	0,377	0,405	0,390	0,341	0,263	0,462	0,446	0,440	0,416	0,416
Sul	0,367	0,354	0,306	0,244	0,186	0,319	0,332	0,313	0,294	0,291
Co	0,252	0,288	0,290	0,282	0,276	0,377	0,347	0,327	0,295	0,327
Sexo	0,285	0,360	0,413	0,447	0,469	0,427	0,420	0,430	0,471	0,449
Contprev	0,165	0,049	-0,042	-0,096	-0,139	0,357	0,348	0,353	0,358	0,368
Casado	-0,038	-0,022	-0,025	-0,029	-0,004*	-0,054	-0,033	-0,022*	-0,023*	-0,063
_cons	-2,295	-2,075	-1,844	-1,473	-1,207	-2,185	-1,912	-1,670	-1,414	-1,140

Fonte: Elaboração própria a partir dos microdados da PNAD de 1998.

Nota: Significante no nível de: 10%; \*Não significantes.

**Tabela A9: Regressão quantílica, setores formal1 e informal1, Brasil, 2004**

Variáveis	Formal1					Informal1				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	0,101	0,118	0,132	0,162	0,159	0,159	0,138	0,157	0,159	0,199
Integral	-0,354	-0,371	-0,376	-0,410	-0,520	-0,259	-0,317	-0,393	-0,502	-0,626
Idade	0,035	0,037	0,047	0,058	0,065	0,047	0,047	0,050	0,055	0,059
Idade2	0,000	0,000	0,000	0,000	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001
Educação	0,050	0,060	0,070	0,080	0,089	0,058	0,055	0,055	0,066	0,072
Rm	0,073	0,089	0,112	0,133	0,140	0,066	0,092	0,115	0,112	0,122
Superior	0,601	0,721	0,828	0,917	1,015	0,366	0,480	0,624	0,740	0,940
Média	0,249	0,299	0,360	0,413	0,483	0,119	0,226	0,276	0,358	0,521
Manual	0,115	0,128	0,119	0,127	0,155	-0,180	-0,035	0,024	0,107	0,231
Norte	0,151	0,171	0,182	0,179	0,164	0,445	0,359	0,286	0,263	0,280
Sudeste	0,268	0,304	0,295	0,260	0,210	0,534	0,467	0,394	0,362	0,316
Sul	0,267	0,283	0,260	0,205	0,151	0,488	0,429	0,368	0,328	0,277
Co	0,246	0,271	0,278	0,303	0,277	0,577	0,488	0,429	0,390	0,386
Sexo	0,254	0,311	0,364	0,409	0,421	0,480	0,420	0,425	0,465	0,445
Contprev	0,476	0,244	0,077	-0,006*	-0,103	0,423	0,362	0,341	0,376	0,406
Casado	-0,003*	-0,007*	-0,024	-0,028	-0,032	-0,018*	-0,014*	0,001*	0,014*	0,004*
_cons	-2,077	-1,838	-1,749	-1,690	-1,395	-2,557	-2,151	-1,810	-1,614	-1,362

Fonte: Elaboração própria a partir dos microdados da PNAD de 2004.

Nota: Significante no nível de: 10%; \*Não significantes.

**Tabela A10: Regressão quantílica, setores formal2 e informal2, Brasil, 1992**

Variáveis	Formal2					Informal2				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	0,124	0,117	0,127	0,144	0,175	0,153	0,131	0,156	0,161	0,189
Integral	-0,224	-0,259	-0,314	-0,364	-0,452	-0,266	-0,311	-0,384	-0,446	-0,553
Idade	0,055	0,055	0,058	0,065	0,067	0,031	0,037	0,040	0,042	0,054
Idade2	-0,001	-0,001	-0,001	-0,001	-0,001	0,000	0,000	0,000	0,000	-0,001
Educação	0,066	0,073	0,081	0,087	0,091	0,072	0,074	0,077	0,077	0,078
Rm	0,177	0,184	0,207	0,210	0,217	0,118	0,147	0,163	0,192	0,212
Superior	0,686	0,715	0,799	0,862	0,927					
Média	0,448	0,397	0,404	0,402	0,429	-0,884	-0,935	-0,969	-1,038	-0,778
Manual	0,226	0,144	0,136	0,127	0,157	-0,765	-0,871	-1,017	-1,182	-0,996
Norte	0,335	0,275	0,246	0,219	0,223	0,208	0,206	0,213	0,235	0,226
Sudeste	0,507	0,421	0,384	0,330	0,271	0,502	0,423	0,361	0,312	0,237
Sul	0,469	0,379	0,321	0,253	0,174	0,410	0,334	0,274	0,231	0,140
Co	0,429	0,324	0,287	0,233	0,203	0,469	0,389	0,322	0,260	0,252
Sexo	0,362	0,382	0,432	0,487	0,505	0,638	0,563	0,499	0,476	0,448
contprev	0,523	0,422	0,341	0,315	0,325	0,338	0,349	0,330	0,322	0,342
Casado	0,033*	0,025	0,002*	-0,003*	-0,003*	-0,066*	-0,063	-0,025*	0,008*	0,001*
_cons	-3,170	-2,621	-2,253	-1,968	-1,618	-1,671	-1,104	-0,462	0,131*	0,290*

Fonte: Elaboração própria a partir dos microdados da PNAD de 1992.

Nota: Significante no nível de: 10%; \*Não significantes.

**Tabela A11: Regressão quantílica, setores formal2 e informal2, Brasil, 1998**

Variáveis	Formal2					Informal2				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	0,128	0,129	0,131	0,141	0,160	0,186	0,158	0,163	0,169	0,192
Integral	-0,352	-0,375	-0,408	-0,462	-0,572	-0,280	-0,301	-0,402	-0,497	-0,592
Idade	0,046	0,049	0,055	0,058	0,063	0,047	0,041	0,039	0,043	0,048
Idade2	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001	0,000	0,000	0,000	0,000
Educação	0,065	0,072	0,081	0,086	0,092	0,069	0,076	0,080	0,084	0,082
Rm	0,170	0,197	0,214	0,228	0,232	0,108	0,134	0,140	0,162	0,209
Superior	0,671	0,777	0,893	0,991	1,068					
Média	0,366	0,380	0,391	0,386	0,384	-0,585	-0,808	-0,716	-0,552	-0,259*
Manual	0,112	0,094	0,075	0,058	0,073	-0,425	-0,678	-0,662	-0,634	-0,468
Norte	0,154	0,180	0,209	0,271	0,307	0,214	0,247	0,237	0,219	0,294
Sudeste	0,383	0,401	0,413	0,377	0,334	0,469	0,477	0,437	0,383	0,349
Sul	0,348	0,344	0,319	0,279	0,247	0,307	0,323	0,298	0,256	0,249
Co	0,281	0,287	0,300	0,299	0,297	0,379	0,375	0,357	0,320	0,360
Sexo	0,297	0,362	0,421	0,465	0,480	0,469	0,451	0,423	0,416	0,347
Contprev	0,328	0,269	0,223	0,172	0,127	0,405	0,378	0,372	0,387	0,390
Casado	-0,038	-0,021	-0,016*	-0,020	-0,021*	-0,074	-0,074	-0,041*	-0,051	-0,121
_cons	-2,245	-2,066	-1,890	-1,585	-1,283	-1,904	-1,125	-0,596	-0,203*	0,053*

Fonte: Elaboração própria a partir dos microdados da PNAD de 1998.

Nota: Significante no nível de: 10%; \*Não significantes.

**Tabela A12: Regressão quantílica, setores formal2 e informal2, Brasil, 2004**

Variáveis	Formal2					Informal2				
	10°	25°	50°	75°	90°	10°	25°	50°	75°	90°
Raça	0,103	0,110	0,125	0,146	0,159	0,177	0,190	0,200	0,169	0,229
Integral	-0,349	-0,358	-0,392	-0,450	-0,552	-0,224	-0,287	-0,357	-0,442	-0,613
Idade	0,037	0,038	0,044	0,054	0,059	0,061	0,051	0,049	0,044	0,045
Idade2	0,000	0,000	0,000	0,000	-0,001	-0,001	-0,001	-0,001	0,000	0,000
Educação	0,052	0,056	0,062	0,073	0,082	0,067	0,067	0,066	0,072	0,077
Rm	0,087	0,105	0,131	0,145	0,155	-0,004*	0,029	0,044	0,039	0,049
Superior	0,571	0,701	0,831	0,935	1,042					
Média	0,242	0,274	0,318	0,382	0,465	-0,186*	-0,339	-0,367	-0,478	-0,309
Manual	0,088	0,087	0,060	0,069	0,099	-0,459	-0,658	-0,789	-0,925	-0,743
Norte	0,226	0,221	0,220	0,211	0,207	0,514	0,388	0,304	0,276	0,283
Sudeste	0,336	0,348	0,331	0,300	0,254	0,544	0,457	0,365	0,331	0,286
Sul	0,333	0,329	0,304	0,260	0,191	0,488	0,381	0,310	0,298	0,266
Co	0,318	0,324	0,334	0,338	0,311	0,654	0,513	0,420	0,391	0,391
Sexo	0,284	0,319	0,374	0,433	0,452	0,563	0,470	0,429	0,386	0,350
Contprev	0,492	0,348	0,263	0,198	0,168	0,423	0,385	0,382	0,417	0,463
Casado	0,006*	-0,003*	-0,007*	-0,008*	-0,012*	-0,062	-0,031*	-0,019*	-0,033	-0,072
_cons	-2,235	-1,949	-1,786	-1,715	-1,479	-2,707	-1,718	-0,979	-0,239	0,133*

Fonte: Elaboração própria a partir dos microdados da PNAD de 2004.

Nota: Significante no nível de: 10%; \*Não significantes.



## NOTAS

1. A pesquisa Economia Informal Urbana (Ecinf), do IBGE, considera o setor informal como composto por trabalhadores por conta própria e empregadores com até cinco empregados, independente do registro legal da empresa.
2. A definição por posição na ocupação na PNAD é permeada, também, por limitações na construção da tipologia se estamos buscando algo mais próximo à concepção de subordinação. Para tal, precisaríamos de informações advindas da unidade estabelecimento e não do indivíduo, de modo a caracterizar a pequena produção.
3. O estimador dos outros quantis pode ser calculado através da minimização de uma generalização desta expressão:

$$\tilde{\beta} = \arg \min \sum_{i=1}^n (p - 1) (y_i \geq x_i' \beta) (y_i - x_i' \beta)$$
 Apesar de essas expressões não permitirem soluções explícitas, os parâmetros podem ser obtidos através de métodos de programação linear.

4. A diferença entre os rendimentos médios é operacionalizada a partir da diferença entre as duas equações dos grupos (2) e (3), avaliadas nos pontos médios das variáveis. Somando e subtraindo  $\beta_1 X_2$ , de forma a não alterar a igualdade, obtemos a equação (4).
5. Essa decomposição, feita ao longo da distribuição de rendimentos, sugere como o diferencial de rendimentos é afetado pela distribuição de rendimentos geral (o que ocorre quando os retornos à qualificação são diferentes). Aumentos na dispersão dos rendimentos aumentariam o diferencial entre os segmentos, mesmo que essas mudanças não tivessem efeito sobre a localização das distribuições dos dois grupos.
6. Doravante denominado Setor Formal1 e Setor Informal1.
7. Doravante denominado Setor Formal2 e Setor Informal2.
8. As categorias ocupacionais foram construídas com base em uma escala de nível de qualificação da ocupação, caracterizado por seu nível educacional médio exigido e grau de hierarquia ocupacional. O detalhamento das ocupações que compõem cada categoria pode ser requisitado aos autores.
9. Diferencial é a razão entre os valores do logaritmo dos salários de um grupo (formal) e de outro (informal), ao longo de suas distribuições específicas, para cada ano analisado.
10. Os resultados das regressões e as médias das variáveis se encontram em tabelas no apêndice.
11. Os coeficientes estimados da equação são apresentados no anexo.

## REFERÊNCIAS BIBLIOGRÁFICAS

ABRAMOVAY, R.; SAES, S.; SOUZA, M. C.; MAGALHÃES, R. Mercados do empreendedorismo de pequeno porte no Brasil. In: *Pobreza e Mercados no Brasil, Uma Análise de Iniciativas de*

- Políticas Públicas*. Brasília: Department for International Development/CEPAL, Escritório no Brasil, 2003.
- ALTONJI, J. G.; BLANK, R. M. Race and gender in the labor market. In: ASHENFELTER, O.; CARD, D. (Ed.). *Handbook of Labor Economics*, Elsevier Science, v. 3, 1999.
- CACCIAMALI, M. C. *Setor informal urbano e formas de participação na produção*. (Tese de Doutorado em Economia) – Universidade de São Paulo, São Paulo, 172f. 1983.
- . Globalização e processo de informalidade. *Economia e Sociedade*, São Paulo, IE – UNICAMP, p. 153-175, jul. 2000.
- CARNEIRO, F. G. A.; A. HENLEY Modelling Formal vs. Informal employment and earnings: microeconomic evidence for Brazil. In: ENCONTRO NACIONAL DE ECONOMIA, 29. *Anais*, ANPEC, 2001.
- DEATON, A. Data and econometric tools for development analysis. In: BEHRMAN, J.; SRINIVASAN, T. N. (Ed.). *Handbook of Development Economics*. Amsterdam: Elsevier, 1995, v. 3A, p. 1.785-1.882.
- DICKENS, W.; LANG, K. A test of dual labour market theory. *American Economic Review*, v. 75, n. 4, p. 1-22, 1985.
- HIRATA, G.; MACHADO, A. F. Conceito de informalidade e formalidade e uma proposta de tipologia. In: *Boletim de Mercado de Trabalho Conjuntura e Análise*, n. 34, nov. 2007.
- IBGE. Pesquisa Nacional de Amostra por Domicílio, PNAD, Rio de Janeiro, 1992; 1998; 2004.
- MACHADO, A. F.; ANDRADE, M. V. Qualificação do excedente de mão-de-obra: estratégia de vida dos trabalhadores por conta própria. In: ENCONTRO NACIONAL DE ECONOMIA, 22. *Anais*. Florianópolis, SC: ANPEC 1994.
- MACHADO, A. F.; PENIDO, M.; OLIVEIRA, J. Análise de sobrevivência na posição de trabalhador por conta própria no Brasil metropolitano (1997-2001) Recife. In: ENCONTRO NACIONAL DA ASSOCIAÇÃO BRASILEIRA DE ESTUDOS DO TRABALHO – ABET, 9. *Anais*, ABET, 2005.
- MALONEY, W. Are LDC Labor Markets Dualistic? *Working paper*, World Bank, 1998.
- MENEZES-FILHO, N. A.; MENDES, M.; ALMEIDA, E. S. O. O diferencial de salários formal: informal – segmentação ou viés de seleção? *Revista Brasileira de Economia*, Rio de Janeiro, v. 58, n. 2, p. 235-248, 2004.
- PIANTO, M. E. T.; PIANTO, D. M. Informal employment in Brazil — a choice at the top and segmentation at the bottom: a quantile regression approach. *Texto para Discussão*, Brasília, n. 236, ago. 2002.
- TAYLOR, M. P. Survival of the fittest? An analysis of self-employment duration in Britain. *The Economic Journal*, v. 109, p. 140-155, mar. 1999.