

TABAGISMO E PENALIZAÇÃO SALARIAL NO MERCADO DE TRABALHO BRASILEIRO

ALÉSSIO TONY CAVALCANTI DE ALMEIDA*
IGNÁCIO TAVARES DE ARAÚJO JÚNIOR†

Resumo

Esse artigo investiga a heterogeneidade da repercussão de hábitos pessoais não saudáveis, expressa pelo uso do cigarro, sobre a produtividade do trabalho no Brasil. Baseado na Pesquisa Especial de Tabagismo integrante da Pesquisa Nacional por Amostra de Domicílios 2008, os principais modelos empíricos são desenvolvidos por meio de regressão quantílica com variáveis instrumentais. Os resultados encontrados explicitam que fumantes, independentemente dos modelos condicionados para a média ou por quantil com e sem variáveis instrumentais, apresentam menor rendimento do trabalho. Assim, a penalização salarial para fumantes com controle para endogeneidade varia de 15,2% a 36,5% ao longo da distribuição condicional dos rendimentos individuais.

Palavras-chave: Tabagismo; Mercado de Trabalho; Produtividade.

Abstract

This research investigates the heterogeneity of the repercussion of unhealthy personal behaviors, expressed by cigarette smoking, on labor productivity in Brazil. Based on the Special Smoking Survey included in the National Survey by Household Sample 2008, the central empirical models are developed by instrumental variable quantile regression. The findings evidence that smokers, regardless of conditioned models for average or quantile with and without instrumental variables, have lower labor income in Brazil. Thus, the smoking wage penalty with endogeneity control ranges from 15,2% to 36,5% over the conditional distribution of individual income.

Keywords: Smoking; Labor Market; Productivity.

JEL classification: I12, J24

DOI: <http://dx.doi.org/10.11606/1413-8050/ea146024>

* Programa de Pós-Graduação em Economia – PPGE/UFPB. E-mail: alessio@ccsa.ufpb.br

† Programa de Pós-Graduação em Economia – PPGE/UFPB. E-mail: ignacio.tavares@gmail.com

1 Introdução

As pesquisas dirigidas ao capital humano realçam a importância dos componentes educação e saúde para os retornos econômicos (Schultz 1961, Mincer 1974, Bartel & Taubman 1979, Deolalikar 1988). Conforme Grossman (1972), a saúde difere das outras formas de capital humano, pois o seu estoque determina a quantidade total de tempo e a intensidade que o agente se dedica às atividades produtivas e ao lazer, relacionando-se diretamente com os rendimentos individuais e com a sensação de bem-estar. Fatores não controláveis pelo indivíduo, como herança genética, e controláveis, como hábitos saudáveis, definem o seu estoque de saúde. Logo, a adoção de comportamentos prejudiciais à saúde podem exercer um impacto negativo sobre a produtividade.

O consumo de determinadas substâncias nocivas à saúde, como aquelas reunidas no cigarro, tem ganhado destaque nesse tipo de análise, motivado por suas consequências na redução da produtividade marginal do trabalho (Leigh & Berger 1989, Levine et al. 1997, Viscusi & Hersch 2001, Munasinghe & Sicherman 2006, Anger & Kvasnicka 2010, Bockerman et al. 2014). Para o caso específico do cigarro, os achados empíricos da literatura internacional suportam a hipótese da existência de um custo monetário indireto do tabagismo nos ganhos individuais como, por exemplo, em Levine et al. (1997), que estimam perdas salariais variando entre 4% e 8% para trabalhadores dos Estados Unidos. Bockerman et al. (2014) ressaltam ainda que o descasamento das maturidades envolvendo os custos (pecuniários e não pecuniários, a maior parte com maior prazo de maturação) e os benefícios (sensação de prazer imediata com dependência do estoque passado) implicam em uma visão míope dos indivíduos quanto à utilização do tabaco, dadas as perdas salariais no longo prazo.

No Brasil, os dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) 2008 expõem que aproximadamente 21% de indivíduos entre 25 e 60 anos de idade são atuais usuários do cigarro, sendo a maior parcela do sexo masculino (52,1%). Dentro do grupo dos não usuários, 18% são ex-fumantes, demonstrando que mais de 1/3 da população brasileira em idade produtiva consumiu ou consome cigarro. Considerando esse quantitativo de usuários e os aspectos relacionados às recentes mudanças demográficas vivenciadas pelo Brasil¹, que resultam em uma pressão crescente por seguridade social e demanda por serviços de saúde no Sistema Único de Saúde (SUS), torna-se relevante investigar as consequências do tabagismo sobre indicadores do mercado de trabalho brasileiro, sobretudo através do prisma microeconômico do retorno individual. Apesar de ser um tema já estudado em outros países, os efeitos do consumo de cigarro sobre a produtividade do trabalho é desconhecido no Brasil.

Para atenuar essa lacuna na literatura nacional, o presente estudo examina como hábitos prejudiciais à saúde, capturado pelo tabagismo, associam-se com os resultados no mercado de trabalho por parte de agentes em fase produtiva. Mais precisamente, investiga-se como o uso de cigarro repercute em termos de ganhos salariais no mercado de trabalho brasileiro. Na finalidade de responder a esse objetivo, estimadores quantílicos com controle para en-

¹Segundo os dados do Censo Demográfico do Instituto Brasileiro de Geografia e Estatística (IBGE), a mudança na composição da estrutura etária do país, com o envelhecimento da população brasileira, vem se intensificando nas últimas décadas (Instituto Brasileiro de Geografia e Estatística 2013).

dogeneidade e viés de seleção são empregados para examinar as estimativas ao longo de diferentes pontos da distribuição condicional do salário-hora de trabalhadores brasileiros.

A relevância de investigar essa temática nesse prisma, deve-se ao fato de que a diferença da penalização salarial entre indivíduos fumantes e não fumantes pode variar em função da distribuição condicional do salário-hora, no qual poderia existir uma sinalização que o custo do tabagismo fosse menor quanto maior fosse a produtividade marginal do indivíduo, ou vice-versa².

Além desta introdução, a pesquisa é composta por mais cinco partes. A Seção 2 faz uma breve revisão da literatura sobre os custos monetários indiretos do tabagismo no mercado de trabalho. As Seções 3 e 4 detalham o modelo empírico e a base de dados da pesquisa. Ao final, as Seções 5 e 6 apresentam os principais resultados e as considerações finais.

2 Revisão da literatura sobre tabagismo e retornos no mercado de trabalho

Tradicionalmente, como enfatizam Willis (1986) e Psacharopoulos & Patrinos (2004), as análises sobre os determinantes do retorno salarial do indivíduo se concentram no componente educação do capital humano, sendo a função de ganhos de Mincer (1974) a formulação que baliza grande parte das pesquisas econômicas na área. Estudos como Bartel & Taubman (1979), Deolalikar (1988), Hersch & Viscusi (1990), Levine et al. (1997), Thomas & Strauss (1997), Strauss & Thomas (1998), Zarkin et al. (1998), Viscusi & Hersch (2001), Auld (2005), Munasinghe & Sicherman (2006), Anger & Kvasnicka (2010), Bockerman et al. (2014), entre outros, acrescentam o componente saúde (\mathbf{K}_i^H) na expressão de rendimentos minceriana (Equação 1), sob a hipótese de que o estoque e comportamentos atrelados a saúde exercem efeitos na produtividade do trabalho:

$$W_i = w(\mathbf{K}_i^H, \mathbf{K}_i, \mathbf{X}_i) + \xi_i, \quad (1)$$

em que: W_i refere-se ao logaritmo do salário; \mathbf{K}_i^H é o vetor de medidas do componente saúde do capital humano; \mathbf{K}_i é o vetor de medidas não relativas à saúde do capital humano (como escolaridade e experiência); \mathbf{X}_i é o vetor de outras covariadas (idade, raça, gênero etc.); ξ_i é um termo de erro estocástico.

Tratando o estoque de saúde como um bem normal, Grossman (1972) assume que a saúde entra diretamente na função de preferência dos agentes (u), pois problemas de saúde geram uma redução no tempo disponível para o usufruto de atividades econômicas (como o trabalho) e não econômicas (como o lazer). A função de utilidade intertemporal de um agente típico é expressa pelo consumo de bens de saúde (G^H) e outros bens em geral (G), tal como evidenciado na Equação 2

$$u = u(\phi_0 G_0^H, \dots, \phi_n G_n^H, G_0, \dots, G_n), \quad (2)$$

em que: G_0^H = estoque herdado de saúde; G_t^H = estoque de saúde no t -ésimo período de tempo, ϕ_t = fluxo de serviço de saúde por unidade de estoque,

²As principais referências que dão suporte a este artigo são Levine et al. (1997), Auld (2005), Munasinghe & Sicherman (2006), Grafova & Stafford (2009), Fernandez & Nordman (2009), Anger & Kvasnicka (2010) e Bockerman et al. (2014).

$\phi_t G_t^H$ representa o consumo total dos serviços de saúde; G_t = consumo total de outros bens e serviços no t -ésimo período de tempo; o indexador n é uma variável endógena que reflete a duração da vida do indivíduo, que chega ao fim quando $G_t^H = G_{min}^H$.

Nessa direção, as escolhas dos indivíduos por um estilo de vida mais saudável afetam a produção do estoque de saúde, no qual o consumo do cigarro representa uma redução no nível de investimento líquido desse estoque, $G_{t+1}^H - G_t^H$ (Grossman 1972). Como resposta da redução do investimento no capital saúde e de outras indicações comportamentais reveladas pelo consumo do produto, Levine et al. (1997), Schwarze & Heineck (2003), Van Ours (2004), Auld (2005), Munasinghe & Sicherman (2006), Lokshin & Beegle (2006), Grafova & Stafford (2009), Anger & Kvasnicka (2010) e Bockerman et al. (2014), mostram empiricamente que o tabagismo provoca uma penalização na taxa de retorno individual no mercado de trabalho. A Tabela 1 sumariza os resultados dos citados trabalhos.

Agrupando os estudos por país ou região, tem-se inicialmente que Leigh & Berger (1989) são um dos primeiros a averiguar a relação entre tabagismo e salário, encontrando uma associação sem significância estatística. Neste estudo, os dados do *Quality of Employment Survey* (QES) de 1973 para trabalhadores dos Estados Unidos (EUA) são usados em conjunto com modelo clássico de regressão, no qual além do impacto do uso do cigarro, avalia-se também o papel do excesso de peso sobre os salários – que também não apresenta significância. Já em Levine et al. (1997), os fumantes, a partir dos dados do *National Longitudinal Survey of Youth* (NLSY) de 1984 e 1991 para os EUA, têm uma diminuição nos salários entre 4% e 8%, dependendo da especificação do modelo, quando comparado aos não fumantes. Os autores adotam diferentes estratégias, dentre elas, a utilização do modelo de efeito fixo do componente familiar na regressão de diferenças salariais entre os irmãos, que visa controlar características não observáveis da família. Além deles, Munasinghe & Sicherman (2006) e Grafova & Stafford (2009) desenvolvem aplicações para trabalhadores norte-americanos, dando suporte aos efeitos negativos do cigarro sobre a taxa de salário.

Grafova & Stafford (2009) examinam os diferentes grupos de fumantes em termos de ganhos salariais, concentrando-se nos usuários persistentes que consomem o cigarro por um longo período de tempo. Com base nos dados do *Panel Study of Income Dynamics* (PSID) nos anos de 1986, 1999 e 2001, os mesmos aferem diferenças salariais significativas entre fumantes e não fumantes. Por MQO agrupado, os fumantes recebem, em média, 7,4% a menos que os não fumantes. Considerando o histórico dos indivíduos, os consumidores persistentes quando confrontados com diferentes bases de especificação (indivíduos que nunca fumaram e ex-fumantes) apresentam variações salariais estatisticamente significativas variando de -8% a -12%, dependendo do período.

O trabalho de Auld (2005) tem por escopo aferir de forma conjunta o papel do consumo de bebidas alcoólicas e do cigarro sobre os salários de indivíduos canadenses a partir do *General Social Survey* (GSS) para os anos de 1985 e 1991. As suas principais descobertas dizem respeito à subestimação da penalização salarial do tabagismo, na abordagem sem controle para endogeneidade, em cerca de 1/3. Comparativamente, os efeitos do cigarro se mostram bem maiores do que o de consumo de bebidas alcoólicas sobre os salários, que em

Tabela 1: Resultados sobre penalização salarial do tabagismo na literatura especializada

Estudo	Local	Amostra	Método	T	N	Coef. (%)
Leigh & Berger (1989)	EUA	Ambos 18 ou mais	MQO	1973	1.142	-3,5 ^a
Levine et al. (1997)	EUA	Ambos	MQO	1984	2.817	-4,2
				1991	3.228	-6,9
			Diferenças entre irmãos	1984	151	-8,0
			1991	228	-8,1	
			Pooling		379	-7,9
Schwarze & Heineck (2003)	ALE	Homens 25-55 anos	MQ2E	1998	2.570	-2,5
			Painel	1998-2001	8.595	-0,7 ^a
Van Ours (2004)	HOL	Homens 26-55 anos	MQO	2001	503	-10,7
			MQ2E	2001	503	-16,6
Auld (2005)	CAN	Homens 25-59 anos	MV	1985 e 1991	3.891	-8,3
			FIMSL-VI	1985 e 1991	3.891	-24,0
Munasinghe & Sicherman (2006)	EUA	Ambos	Dinâmico(1)	1979-1994	9.396	-0,5
			Dinâmico(2)	1979-1994	9.396	-0,2
			Dinâmico(3)	1979-1994	9.396	-0,2

Fonte: Elaboração própria.

Legenda (Local): EUA = Estados Unidos; ALE = Alemanha; HOL = Holanda; CAN = Canadá; ALB = Albânia; RU = Reino Unido; FIN = Finlândia.
 Legenda (Método): MQO = Mínimos Quadrados Ordinários; MQ2E = Mínimos Quadrados em dois estágios; MV = Máxima Verossimilhança; FIMSL = Máxima verossimilhança simulada para informações completas (em inglês, *full information maximum simulated likelihood*); Gêmeos = modelo de diferenças entre gêmeos; DZ = gêmeos dizigóticos; MZ = gêmeos monozigóticos. Em Munasinghe & Sicherman (2006), Dinâmico(1), Dinâmico(2) e Dinâmico(3) representam a regressão sem controle, com controles limitados e controles completos.

^a = Não significativo estatisticamente; * Variável explicativa é o número de pacotes de cigarro consumidos.

Tabela 1: Resultados sobre penalização salarial do tabagismo na literatura especializada (continuação)

Estudo	Local	Amostra	Método	T	N	Coef. (%)
Lokshin & Beegle (2006)	ALB	Homens	MQO	2005	1.588	-4,8
		25-60 anos	MQ2E	2005	1.588	-25,6
Braakmann (2008)	RU	Homens	MQO	1991-2005	32.356	-3,1
			Painel-EF	1991-2005	32.356	-0,9 ^a
			Painel-VI	1991-2005	17.216	0,2 ^a
Grafova & Stafford (2009)	EUA	Homens	MQO	1986	3.056	-3,4
				1999	3.062	-9,2
				2001	2.898	-10,9
			MQO-pooled	1986-2001	9.016	-7,4
			MQO-EF	1986-2001	9.016	0,9
Anger & Kvasnicka (2010)	ALE	Homens	MQO	2002	1.968	-4,5
		27-55 anos	MQ2E	2002	1.968	-9,9
Bockerman et al. (2014)*	FIN	Ambos	MQO	1990-2004	3.914	-2,16
			Gêmeos	1990-2004	1.957	-1,41
			Gêmeos-DZ	1990-2004	1.350	-1,30
			Gêmeos-MZ	1990-2004	607	-1,85

Fonte: Elaboração própria.

Legenda (Local): EUA = Estados Unidos; ALE = Alemanha; HOL = Holanda; CAN = Canadá; ALB = Albânia; RU = Reino Unido; FIN = Finlândia.
 Legenda (Método): MQO = Mínimos Quadrados Ordinários; MQ2E = Mínimos Quadrados em dois estágios; MV = Máxima Verossimilhança;
 FIMSL = Máxima verossimilhança simulada para informações completas (em inglês, *full information maximum simulated likelihood*); Gêmeos =
 modelo de diferenças entre gêmeos; DZ = gêmeos dizigóticos; MZ = gêmeos monozigóticos. Em Munasinghe & Sichertman (2006), Dinâmico(1),
 Dinâmico(2) e Dinâmico(3) representam a regressão sem controle, com controles limitados e controles completos.

^a = Não significativo estatisticamente; * Variável explicativa é o número de pacotes de cigarro consumidos.

alguns casos os parâmetros referentes ao álcool não são diferentes de zero na perspectiva estatística.

Em relação aos trabalhos com dados europeus, Schwarze & Heineck (2003), através do *Socio-Economic Panel* (SOEP) para Alemanha, obtêm na abordagem com dados de corte que os fumantes do sexo masculino têm uma redução na taxa de salário de aproximadamente 2,5%, enquanto que no modelo com efeito fixo, controlando a heterogeneidade individual não observada, o impacto do cigarro sobre os salários não é significativo estatisticamente. Já Anger & Kvasnicka (2010), com a mesma base de dados, acham efeitos negativos do consumo do tabaco sobre os ganhos dos indivíduos, sendo os custos indiretos do cigarro na abordagem com variáveis instrumentais (VI) representado por uma redução em 0,099 do log do salário-hora.

Fazendo a varredura de estudos para outros países europeus, apenas o trabalho de Braakmann (2008), para dados do *British Household Panel Survey* (BHPS) de 1991 a 2005, não suporta a hipótese que relaciona a penalização salarial de fumar e menor produtividade dos fumantes nos modelos de efeito fixo e VI. Muito embora, identifica uma penalização de 3% com dados de corte. As demais pesquisas, como Van Ours (2004) para Holanda, Lokshin & Beegle (2006) para Albânia e Bockerman et al. (2014) para Finlândia, corroboram os achados da grande maioria da literatura sobre a diminuição nos rendimentos provocada pelo tabagismo.

Dentre as localidades reportadas na Tabela 1, a que possui o menor nível de renda *per capita* é a Albânia, onde Lokshin & Beegle (2006) fazendo uso do *Living Standards Monitoring Survey* (LSMS) de 2005, constatam uma penalização salarial do tabagismo superior a 20% no modelo com VI. Usando as pesquisas *Older Finnish Twin Cohort Study* (OFTCS) do Departamento de Saúde Pública da Universidade de Helsinque na Finlândia e *Finnish Longitudinal Employer-Employee Data* (FLEED) para gêmeos do sexo masculino nascidos entre 1945 e 1957, Bockerman et al. (2014) concluem que fumantes possuem menor rendimento de longo prazo, controlando por fatores ambientais, efeitos dos pares e carga genética (ao usar dados sobre gêmeos idênticos, monozigóticos).

Quanto aos motivos discutidos na literatura que podem explicar a penalização salarial dos fumantes no mercado de trabalho, enumeram-se os seguintes:

- Interferência do consumo do cigarro em tarefas de execução manual;
- Elevação dos custos relativos do empregador, devido, por exemplo, ao maior absenteísmo e custos do seguro-saúde;
- Discriminação no ambiente de trabalho causada por questões de atratividade física visual (possíveis danos na pele) ou olfativa (decorrente do cheiro de fumaça);
- Predisposição dos fumantes por emprego que disponibilizam seguro-saúde em detrimento de uma maior remuneração;
- Alta taxa de desconto intertemporal, indicando maior propensão ao consumo presente em relação ao consumo futuro e, assim, baixos investimentos no capital humano. (Levine et al. 1997, Grafova & Stafford 2009).

A literatura empírica sobre essa temática chama atenção, em especial, para a influência do último ponto acima na estimativa da penalização salarial, haja vista o valor não observado dessa taxa de desconto que pode influir simultaneamente na decisão de fumar e na taxa de salário. Nesse quadro, se uma pessoa com preferências imediatistas tem mais predisposição a fumar, o parâmetro do custo indireto do tabagismo seria tendencioso, porque as consequências negativas sobre os salários de tais preferências pelo consumo presente, dado o menor investimento no capital humano, seriam atribuídas ao ato de fumar. Assim, o indicador do status de fumante (H_i) é tratado como um regressor endógeno, $Cov(H_i, \epsilon_i) \neq 0$, o que demanda um tratamento para o problema de endogeneidade. Logo, a grande maioria dos estudos aplicados na área tenta contornar a questão dos fatores não observados, como fica demonstrado pelo crescente rigor metodológico das pesquisas mencionadas.

Muito embora as estimativas das penalizações salariais do tabagismo possam chegar em alguns casos até a 25,6% na Tabela 1, ainda não existem aferições dos efeitos do uso do cigarro sobre os resultados do mercado de trabalho brasileiro³. De uma forma global, a exploração do assunto está concentrada nos países da América do Norte e da Europa sem a incorporação da heterogeneidade da penalização salarial do uso do cigarro para trabalhadores, por exemplo, com baixos e altos níveis de produtividade do trabalho. Assim, considerando que a posição na distribuição dos rendimentos carrega certas peculiaridades do tipo de função desempenhada pelo trabalhador – por exemplo, ocupações com altas remunerações são mais propensas a requerer mais conhecimento técnico do que condicionamento físico –, o tabagismo pode ter diferentes repercussões ao longo da função quantílica condicional dos salários.

3 Estratégia empírica

A estratégia empírica baseia-se nos trabalhos de Levine et al. (1997), Schwarze & Heineck (2003), Van Ours (2004), Auld (2005), Lokshin & Beegle (2006), Anger & Kvasnicka (2010) e Bockerman et al. (2014). A abordagem mais recorrente na área é a de variáveis instrumentais (VI), que mensura o efeito do consumo do cigarro para um agente representativo sobre o valor condicional médio dos ganhos salariais, tratando essa variável explicativa como um regressor endógeno.

A avaliação do papel das condições de saúde e, em particular, dos hábitos não saudáveis sobre o retorno salarial possui problemas inerentes de mensuração do efeito causal em bases de dados, particularmente, sem caráter longitudinal (Levine et al. 1997, Bockerman et al. 2014). A dificuldade se deve a atributos não observáveis dos indivíduos que afetam concomitante a decisão de fumar e o salário-hora. Assim, as estimativas obtidas pelos métodos que supõem exogeneidade, como MQO e regressão quantílica (RQ), são tendenciosas.

Neste trabalho, estimadores quantílicos com regressores endógenos são utilizados para estimar a relação entre o uso do cigarro e ganhos salariais para os n -ésimos quantis (τ) da distribuição condicional do salário-hora, em que

³Grande parte da literatura nacional, balizada pela Equação 1, está voltada para o papel das condições nutricionais/antropométricas nesse processo, como é o caso de Thomas & Strauss (1997) e Kassouf (1997), ou faz uso de classificações do estado geral de saúde, como Alves & Andrade (2003).

$\tau \in (0, 1)$. A função de regressão quantílica condicional com variáveis instrumentais (RQVI), baseada em Chernozhukov & Hansen (2004, 2005, 2013) e Chernozhukov et al. (2007), é a abordagem escolhida para o cômputo das estimativas de interesse em contraponto à função condicional média com variáveis instrumentais, usada em análises de regressão linear como no método de mínimos quadrados em dois estágios (MQ2E). Além de verificar a heterogeneidade dos coeficientes ao longo da distribuição condicional, a regressão quantílica é mais robusta à presença de observações atípicas e de erros com distribuição não gaussiana do que os estimadores de mínimos quadrados (Koenker 2005). É válido salientar que na seção de resultados desta pesquisa, as estimativas inicialmente são calculadas pelos métodos MQO, MQ2E e RQ convencional (sem controle para VI), para em seguida apresentar os coeficientes obtidos por RQVI.

Os estimadores quantílicos, como destacado por Koenker & Bassett (1978), identificam o efeito das covariadas não apenas para a tendência central, mas também ao longo da distribuição condicional da variável de resultado. Com base na RQ convencional, a Equação 1 de rendimentos do indivíduo pode ser representada pela seguinte representação linear

$$W_i = \delta(\tau)H_i + \mathbf{X1}_i\beta(\tau) + \epsilon_{1i}(\tau), \text{ com } Q_\tau(W_i|H_i, \mathbf{X1}_i) = \delta(\tau)H_i + \mathbf{X1}_i\beta(\tau), \quad (3)$$

em que: W_i é o logaritmo do salário-hora do i -ésimo indivíduo; H_i é uma variável binária referente à condição de fumante do trabalhador; $\mathbf{X1}$ é um vetor formado por variáveis de controle, o que inclui idade, faixa de escolaridade, raça, fatores locacionais, outros atributos pessoais e o intercepto do modelo.

O τ -ésimo quantil condicional do logaritmo do salário-hora em relação às variáveis explicativas é denotado por $Q_\tau(W_i|H_i, \mathbf{X1}_i)$. Isto posto, a Equação 4 refere-se ao problema de minimização para a estimação dos parâmetros para um determinado quantil com $\theta = f(\delta, \beta)$:

$$\min_{\theta \in R^k} \left[\sum_{i:W_i \geq Q_\tau(\cdot)} \tau |W_i - \delta_\tau H_i - \mathbf{X1}_i \beta_\tau| + \sum_{i:W_i < Q_\tau(\cdot)} (1 - \tau) |W_i - \delta_\tau H_i - \mathbf{X1}_i \beta_\tau| \right]. \quad (4)$$

Nesse caso em estudo, ao averiguar o papel do regressor endógeno (H_i), o coeficiente $\delta(\tau)$ obtido a partir da regressão quantílica convencional se torna tendencioso, necessitando adotar alguma abordagem, como variáveis instrumentais, que permita tratar esse viés. Contudo, como apresentado em Chernozhukov & Hansen (2005), as técnicas padrões de variáveis instrumentais (como MQ2E) não são consistentes com o efeito de tratamento quantílico variando ao longo dos diferentes quantis da distribuição da variável dependente. Assim, no modelo RQVI, inicialmente, a relação entre o regressor binário endógeno $H_i \in \{0, 1\}$, que representa o status de fumante do trabalhador, e o salário-hora (W_i) pode ser representada pela Equação 5

$$W_i = \delta(\epsilon_{1i}^*)H_i + \mathbf{X1}_i\beta(\epsilon_{1i}^*), \text{ com } \epsilon_{1i}^*|H_i, \mathbf{X1}_i \sim \text{uniforme}(0, 1), \quad (5)$$

em que: ϵ_{1i}^* é uma variável aleatória não observada, representando alguma habilidade inata ou nível de preferência dos indivíduos, responsável pela heterogeneidade dos resultados para os indivíduos com idênticas características observáveis ($\mathbf{X1}$) e do tratamento (D); $\tau \rightarrow \delta(\tau)H_i + \mathbf{X1}\beta(\tau)$ é a função quantil

condicional de W_i . Por construção, $\delta(\tau)H_i + \mathbf{X1}\beta(\tau)$ é o τ -ésimo quantil do W_i condicionado as variáveis explicativas.

Como H_i é correlacionado com ϵ_{1i}^* , a Equação 5 é reescrita para responder o problema de endogeneidade, em que V_i é uma variável não observada estatisticamente dependente do termo de erro ϵ_{1i} e possui outros fatores omitidos que afetam a decisão de fumar,

$$W_i = \delta(\epsilon_{1i})H_i + \mathbf{X1}_i\beta(\epsilon_{1i}), \text{ com } \epsilon_{1i}|H_i, \mathbf{X1}_i, \mathbf{Z1}_i \sim \text{uniforme}(0, 1) \quad (6)$$

$$H_i = f(\mathbf{X1}_i, \mathbf{Z1}_i, V_i) \quad (7)$$

$$\tau \rightarrow \delta(\tau)H_i + \mathbf{X1}_i\beta(\tau), \quad (8)$$

em que essa decisão é determinada pela Equação 7. O vetor V_i é composto por variáveis não observadas correlacionadas com ϵ_{1i} , e $\mathbf{Z1}_i$ é um vetor de variáveis instrumentais independentes a ϵ_{1i} mas correlacionado com a decisão de fumar (H_i). As Equações 6 e 8, que balizam as condições da abordagem RQVI, implicam:

$$\Pr(W_i \leq \delta(\tau)H_i + \mathbf{X1}_i\beta(\tau)|\mathbf{X1}_i, \mathbf{Z1}_i) = \tau. \quad (9)$$

A Equação 9, segundo Chernozhukov & Hansen (2004), estabelece a conexão entre os parâmetros da RQVI para o conjunto de equações de momentos condicionais necessárias para a estimação e inferência. Para resolver a citada expressão é preciso achar a solução para regressão quantílica de $W_i - [\delta(\tau)H_i + \mathbf{X1}_i\beta(\tau)]$ em $(\mathbf{X1}, \mathbf{Z1})$, que equivale a:

$$Q(W_i \leq \delta(\tau)H_i + \mathbf{X1}_i\beta(\tau)|\mathbf{X1}_i, \mathbf{Z1}_i) = \mathbf{X1}_i\beta(\tau) + \mathbf{Z1}_i\gamma(\tau). \quad (10)$$

Dado o verdadeiro valor de $\delta(\tau)$, estima-se os demais parâmetros, $\beta(\tau)$ e $\gamma(\tau)$ por meio do seguinte problema de minimização:

$$\min_{\theta \in R^k} \sum_i \rho_\tau [W_i - \delta(\tau)H_i - \mathbf{X1}_i\beta(\tau) - \mathbf{Z1}_i\gamma(\tau)], \quad (11)$$

em que: ρ_τ é uma função definida como $\rho_\tau(\psi) = \tau\psi$ se $\psi \geq 0$ ou $\rho_\tau = (\tau-1)\psi$ se $\psi < 0$, com $\psi = [W_i - \delta(\tau)H_i - \mathbf{X1}_i\beta(\tau) - \mathbf{Z1}_i\gamma(\tau)]$. Como não se conhece o valor de $\delta(\tau)$, o problema exposto na Equação 11 é solucionado pela aproximação do valor de $\gamma(\tau)$ em torno de zero.

Segundo Van Ours (2004), as variáveis instrumentais frequentemente utilizadas para resolver o problema de endogeneidade da variável explanatória fumante são formadas, em geral, por fatores atrelados à religião, ao preço do produto, a auto avaliação e/ou ao comportamento da família. Dessa maneira, a validade de uma estimativa por VI depende do pressuposto de que os instrumentos não estão correlacionados com outras características latentes dos agentes que podem influir nos seus ganhos, não sendo portanto um exercício trivial encontrar bons instrumentos. Contudo, Auld (2005) ressalta que o preço do cigarro exerce influência na decisão de consumo do produto mas não se relaciona com a taxa de salário, sendo uma variável exógena consolidada nessa área. Além do mais, para Cutler & Glaeser (2005) e Anger & Kvasnicka

(2010), existem evidências na literatura sobre o papel da percepção dos indivíduos quanto ao malefício do cigarro e as características do ambiente familiar sobre o tabagismo no hábito de fumar, onde inexistente ou é pouco provável um canal direto dessas covariadas sobre os rendimentos atuais do indivíduo.

Com base nas variáveis instrumentais adotadas por Van Ours (2004), Auld (2005) e Anger & Kvasnicka (2010), o vetor $\mathbf{Z1}_i$ é composto pelo preço do cigarro, indicador binário para o indivíduo que reside com ao menos um ex-fumante, quantidade de pessoas fumantes na família (exclui o próprio indivíduo), indicadores sobre crença/percepção dos agentes quanto aos riscos do uso de cigarro para si e para terceiros e regras sobre o uso do cigarro no domicílio. Essas percepções, regras, preços do produto e características de outros moradores do domicílio são potenciais instrumentos por se relacionarem com a decisão de se fumar e serem condicionalmente não correlacionadas com os salários⁴. É válido acrescentar que as variáveis corresidentes, conforme Anger & Kvasnicka (2010), são usadas com a finalidade de medir possíveis influências sobre o comportamento do fumante corrente, tanto em termos das chances de sua decisão de fumar ter sido definida pela existência de outros familiares fumantes ou de pessoas na família com histórico bem-sucedido de cessação do tabagismo.

Ressalta-se que a estimação da Equação 6, seguindo os procedimentos de Buchinsky (1998, 2001) para modelos com regressores quantílicos, é realizada em dois estágios a fim de corrigir o problema de seleção amostral discutido em Heckman (1979), pois o termo de viés de autoseleção na força de trabalho tem uma forma desconhecida na RQ. Destarte, admite-se que o salário observado (S) pode ser escrito como

$$S_i = d_i \times W_i, \text{ com } d_i = I(W_i > W_i^R), \quad (12)$$

em que $I(\cdot)$ é uma função indicadora usual e W_i só é observado se exceder o salário de reserva (W_i^R), de modo que $d_i = 1$, caso o salário seja observável, e $d = 0$, caso contrário. Na presença do mecanismo de seleção, o quantil condicional do salário observado é expresso pela Equação 13

$$Q_\tau(S_i|H_i, \mathbf{X1}_i, \mathbf{Z1}_i) = \delta(\tau)H_i + \mathbf{X1}_i\beta(\tau) + \mathbf{Z1}_i\gamma(\tau) + h_\tau(g_0), \quad (13)$$

em que $h_\tau(g_0) = Q_\tau(\epsilon_{1i}|H_i, \mathbf{X1}_i, \mathbf{Z1}_i, d_i = 1)$ e $g_0 = g(\mathbf{J}_i, \kappa_0)$. Adianta-se que o vetor de variáveis explicativas, \mathbf{J} , é formado por covariadas sobre a decisão de oferta do trabalho (com $\mathbf{X1} \subset \mathbf{J}$), incluindo o rendimento familiar oriundo do não trabalho (aluguel, juros, transferências, doações etc.), estrutura familiar e dificuldades físicas para realização de trabalhos pesados⁵. Segundo Buchinsky (1998), a equação de participação na força de trabalho deve ser estimada por um modelo semiparamétrico. Por isso foi escolhido o modelo de resposta binária de estimadores de máxima verossimilhança semi não paramétricos⁶ (SNP) de Gallant & Nychka (1987).

Em seguida, a exemplo de Buchinsky (1998), a aproximação da razão inversa de Mills com dois termos é adicionada na equação dos salários obser-

⁴A fim de aferir a força dos instrumentos e a restrição de sobreidentificação, a Tabela A.2 reporta testes estatísticos para o ajustamento dos instrumentos e exogeneidade do regressor.

⁵A Tabela A.5 no Apêndice descreve as variáveis utilizadas para a equação de seleção.

⁶O comando semi-nonparametric (SNP) desenvolvido por De Luca (2008) é usado para a estimação desses coeficientes, supondo uma expansão polinomial univariada de terceira ordem de Hermite para fazer a aproximação da densidade desconhecida dos erros da regressão latente.

vados dos indivíduos⁷. Nesse ponto, o termo $h_{\tau}(g) = \rho'_{\tau} P_{3j}(g)$ controla a seleção amostral e $P_{3j}(g)$ é um polinômio de terceira ordem em g , cujo $h_j(g) = \lambda(\hat{\mu} + \hat{\sigma}g)^{j-1}$, com $j = 1, 2$ e $\lambda(\cdot) = \phi(\cdot)/\Phi(\cdot)$, que é a razão inversa de Mills. $\phi(\cdot)$ e $\Phi(\cdot)$ são respectivamente a função de densidade de probabilidade (fdp) e a função de distribuição acumulada (fda), bem como $\hat{\mu}$ e $\hat{\sigma}$ são parâmetros de localização e escala.

4 Base de dados e descrição das variáveis

A Pesquisa Nacional por Amostra de Domicílios (PNAD) 2008 é a fonte de informações usada por dispor da Pesquisa Especial de Tabagismo (PETab) e de um suplemento sobre acesso e utilização dos serviços, condições de saúde e fatores de risco e proteção à saúde. Segundo o Instituto Brasileiro de Geografia e Estatística (2009), a PETab, desenvolvida em parceria com o Instituto Nacional de Câncer (INCA) e a Agência Nacional de Vigilância Sanitária (ANVISA) vinculados ao Ministério da Saúde, tem o objetivo de levantar dados sobre uso dos produtos derivados do tabaco, das tentativas de abandono do hábito, da percepção das pessoas sobre esses riscos, dos aspectos relacionados à compra de cigarros etc. O questionário da PETab é aplicado a 1/3 da amostra de domicílios da PNAD para 39.425 indivíduos com 15 anos ou mais de idade.

Este estudo utiliza especificamente a Parte 28 do questionário da PNAD 2008 denominado de PETab, aplicado para uma subamostra de indivíduos da citada pesquisa, pela disponibilização de um conjunto mais amplo de variáveis referentes ao vício do cigarro que podem melhor instrumentalizar o modelo empírico. A PETab permite uma melhor análise sobre o uso do tabaco, inclusive considerando o conhecimento dos indivíduos sobre os riscos de saúde provocados pelo uso da nicotina e preços implícitos do produto. O número final de observações é de 6.979 pessoas (quando se realiza o controle amostral para ex-fumantes), o que corresponde em termos populacionais a 24,56 milhões de trabalhadores nas características descritas e a 58% do total da amostra da PETab para esse grupo etário do sexo masculino.

Com os microdados da PNAD 2008, que possui representatividade amostral para todo o território brasileiro, informações específicas são obtidas sobre os usuários do cigarro, atributos pessoais, aspectos locais e indicadores do mercado de trabalho, bem como é possível particionar a amostra em três grupos de trabalhadores: fumantes correntes, ex-fumantes e indivíduos sem experiência com o cigarro. Para Anger & Kvasnicka (2010), essa partição é importante para não contaminar o grupo de não fumantes com usuários que em algum momento da vida já fizeram uso do tabaco, pelos seguintes pontos: ex-fumantes podem ter um maior salário do que os recém consumidores do cigarro, admitindo que a produtividade é mais correlacionada com as características individuais do que com o uso recente do cigarro; como o tabagismo é viciante, o sucesso em parar de fumar requer alguns atributos que inclusive são relevantes para a produtividade do trabalho, como disciplina e força de vontade.

A exemplo dos outros trabalhos na área (Viscusi & Hersch 2001, Van Ours 2004, Auld 2005, Grafova & Stafford 2009, Anger & Kvasnicka 2010), a amos-

⁷No caso dos modelos de referência para a função condicional média, o cálculo da razão inversa de Mills tradicional (λ) usa o modelo probit para estimar a equação de participação.

tra é restringida para trabalhadores do sexo masculino com nacionalidade brasileira e faixa etária entre 25 e 60 anos de idade. Essas restrições têm o propósito de comparar as evidências para o Brasil com os de outros estudos (ver Tabela 1), assim como presume-se que aos 25 anos de idade a formação educacional já esteja praticamente maturada ou em fase final e com 60 anos ainda o indivíduo esteja em idade produtiva⁸. Além disso, aposentados e trabalhadores empregados no serviço público foram desconsiderados do desenho amostral. Para o caso dos servidores públicos, a produtividade do trabalho não implica, em regra, alterações na taxa de salário, bem como a existência do fator estabilidade no emprego dificulta o cotejo deste grupo de trabalhadores com os empregados no setor privado.

A Tabela 2 sumariza as estatísticas descritivas das variáveis usadas para os modelos de penalização para dois tipos de amostras, que se diferenciam pelo controle de ex-fumantes. As seguintes variáveis explicativas são utilizadas: status de fumante, faixa de escolaridade (menos que ensino fundamental, fundamental completo, médio completo e superior completo), raça⁹ (cor branca), idade (em nível e ao quadrado), interação entre anos de estudo e idade do indivíduo, indicador binário se o trabalhador é sindicalizado, estabilidade no emprego atual (*tenure*), grupos ocupacionais do trabalho principal e fatores locais (zona de localização, região metropolitana e macrorregiões de residência). A Tabela A.5 no Apêndice mostra a definição das variáveis utilizadas no estudo, inclusive as variáveis instrumentais e as variáveis usadas na equação de seleção.

Pela Tabela 2, a amostra extraída da PETab, para o modelo de penalização, dispõe de dois desenhos amostrais que se diferenciam pelo controle de ex-fumantes. Na Amostra (1), 21,3% dos indivíduos são ex-usuários do cigarro, enquanto que na Amostra (2) essas observações são excluídas. A comparação dessas duas amostras não revela grandes diferenças. Ambas são caracterizadas por pessoas residentes, em sua maior parte, na região Sudeste¹⁰, em áreas urbanas, com ensino fundamental incompleto, casadas e com idade em torno de 39 anos.

5 Resultados

Conforme Anger & Kvasnicka (2010), a verificação da hipótese de penalização salarial do tabagismo requer cuidados específicos com o grupo de não fumantes, porque os custos salariais efetivos de fumar podem ser influenciados pela existência de ex-usuários do cigarro neste grupo. Assim, para iniciar as discussões, a Tabela 3 mostra as características da amostra por grupos de fumantes e não fumantes de dois tipos, diferenciando-se pelo uso passado do cigarro. Dentre as diferenças de médias e proporções entre os grupos, apenas quatro não são estatisticamente diferentes de zero à pelo menos 5% de significância:

⁸Dada uma expectativa de vida de 69 anos para os homens, conforme dados do Banco Mundial para 2008 disponíveis em databank.worldbank.org/data, bem como o direito a aposentadoria masculina a partir apenas dos 65 anos de idade, de acordo com o art. 48 da lei nº 8.213/1991 (Redação dada pela Lei nº 9.032/1995).

⁹Na amostra deste trabalho (com controle amostral para ex-fumantes), verifica-se que 49% dos indivíduos se autodeclaram brancos e 42% se dizem pardos, de modo que os demais grupos étnicos correspondem a um percentual inferior a 10%.

¹⁰Como a amostra é dirigida apenas para trabalhadores vinculados ao setor privado, tem-se por decorrência que a maior parcela de indivíduos encontram-se na região Sudeste.

Tabela 2: Estatísticas descritivas das variáveis utilizadas nos modelos de regressão

Variáveis	Amostra (1)		Amostra (2)	
	Média	DP	Média	DP
Salário-hora (em log)	2,896	0,896	2,908	0,894
Sindicalizado	0,215	0,411	0,213	0,410
Grupos ocupacionais				
Dirigentes	0,070	0,254	0,070	0,255
Ciências e Artes	0,050	0,217	0,051	0,220
Técnicos	0,074	0,262	0,075	0,263
Serviços administrativos	0,051	0,220	0,055	0,228
Serviços*	0,113	0,317	0,115	0,319
Vendedores	0,073	0,260	0,072	0,258
Agrícola	0,168	0,374	0,162	0,368
Produção e manutenção	0,401	0,490	0,401	0,490
Estabilidade (tenure)	8,546	9,038	8,123	8,665
Faixa de escolaridade				
Menos que ensino fundamental (<EF)	0,474	0,499	0,453	0,498
Ensino fundamental completo (EF)	0,163	0,369	0,162	0,368
Ensino médio completo (EM)	0,275	0,446	0,294	0,456
Ensino superior completo* (ES)	0,088	0,284	0,092	0,289
Raça (cor branca)	0,486	0,500	0,490	0,500
Idade	39,747	9,640	38,589	9,368
Fumante	0,249	0,432	0,316	0,465
Ex-fumante	0,213	0,410	0,000	0,000
Chefe do domicílio	0,736	0,441	0,721	0,449
Tipo de família				
Casado sem filho	0,139	0,346	0,136	0,343
Casado com filho	0,688	0,463	0,682	0,466
Tamanho da família	3,588	1,494	3,580	1,494
Dificuldades físicas	0,156	0,363	0,137	0,344
Rendimento familiar do não trabalho				
Pensão (log)	0,555	1,835	0,565	1,860
Aluguel (log)	0,229	1,244	0,222	1,217
Outros (log)	1,029	2,528	0,992	2,483
Proibido fumar no domicílio	0,137	0,344	0,147	0,354
Crença/percepção				
Gera doenças aos outros	0,920	0,272	0,913	0,281
Gera doenças para si	0,966	0,182	0,962	0,191
Fumantes na família	0,123	0,404	0,149	0,442
Ex-fumantes na família	0,074	0,261	0,020	0,140
Preço do cigarro	2,648	0,476	2,646	0,473
Urbana	0,836	0,370	0,840	0,367
Região metropolitana	0,400	0,490	0,398	0,489
Macrorregião				
Norte	0,079	0,270	0,076	0,265
Nordeste	0,242	0,428	0,244	0,429
Sudeste*	0,450	0,498	0,453	0,498
Sul	0,157	0,364	0,157	0,363
Centro-Oeste	0,072	0,259	0,071	0,257
Controle: Ex-fumantes			X	
Observações (N)	8.808		6.979	
Observações expandidas** (P)	31.221.094		24.561.120	

Elaboração própria a partir dos microdados da PNAD/PETab 2008.

*Categoria omitida.

**As amostras possuem fatores de expansão distintos, dada às peculiaridades da PETab que é uma subamostra da PNAD com reponderação do peso amostral dos participantes.

renda do não trabalho e trabalhador casado (sem filhos) entre fumantes (A) e não fumantes do tipo 1 (B); horas de trabalho e morador chefe do domicílio entre fumantes (A) e ex-fumantes ou não fumantes do tipo 2 (C). Na Tabela 3 as diferenças de médias e de proporções são sempre examinadas em torno do grupo de fumantes.

Tabela 3: Características da amostra por grupos de fumantes e não-fumantes no Brasil, PETab 2008

Variáveis	Fumante	Não Fumante		Diferença	
		Tipo 1	Tipo 2	Coef.	Coef.
	(A)	(B)	(C)	(A)-(B)	(A)-(C)
Média					
Salário-hora	24,39	33,14	27,57	-8,74 ^c	-3,18 ^c
Horas de trabalho	44,45	44,55	45,13	-0,10 ^c	-0,67
Estabilidade (em anos)	8,15	7,26	9,18	0,89 ^c	-1,03 ^c
Anos de estudo	5,77	8,19	6,57	-2,41 ^c	-0,79 ^c
Idade	41,44	37,38	44,16	4,06 ^c	-2,71 ^c
Renda familiar do não trabalho	29,48	26,28	40,61	3,20 ^a	-11,13 ^c
Idade de 1º uso do cigarro	16,37		15,89		0,48 ^c
Número de cigarros ao dia	17,74				
Proporção (%)					
Ocupação	88,89	91,12	91,05	-2,23 ^c	-2,16 ^c
Chefe do domicílio	72,67	66,75	77,30	5,92 ^c	-4,63
Cor Branca	42,23	51,10	46,43	-8,87 ^c	-4,20 ^c
Ausência no trabalho*	9,32	5,09	9,62	4,23 ^c	-0,30 ^c
Dificuldades físicas	21,28	12,75	23,28	8,53 ^c	-2,00 ^c
Casado e sem filhos	14,51	12,61	15,09	1,90 ^a	-0,58 ^c
Casado e com filhos	63,92	68,72	69,35	-4,80 ^c	-5,43 ^c
Faixa de escolaridade					
<EF	63,54	38,10	55,46	25,44 ^c	8,08 ^c
EF	15,04	16,87	16,31	-1,83 ^c	-1,27 ^b
EM	16,73	33,72	20,35	-16,99 ^c	-3,62 ^c
ES	4,70	11,31	7,88	-6,61 ^c	-3,18 ^c

Fonte: Elaboração própria a partir dos microdados da PNAD/PETab 2008. Amostra restringida para homens nascidos no Brasil com faixa etária entre 25 e 60 anos, com exclusão de indivíduos que recebem aposentadoria e trabalhadores ocupados no serviço público.

Legenda: ^ap<0,10; ^bp<0,05; ^cp<0,01.

Não fumante Tipo 1: indivíduo que nunca teve experiência com o uso do cigarro.

Não fumante Tipo 2: indivíduo ex-fumante.

*Essa variável indica se nas duas últimas semanas, a pessoa não fez suas atividades habituais por motivo de saúde.

Pela Tabela 3, fumantes e não fumantes dispõem de certas características socioeconômicas distintas, principalmente nas variáveis anos de estudo¹¹ (e na sua distribuição por fase de ensino) e ganhos no mercado de trabalho. Essas diferenças são mais acentuadas quando se coteja a média desses indicadores entre fumantes e não fumantes sem experiência com o cigarro. Enquanto que variáveis como horas de trabalho, estabilidade no emprego (*tenure*), renda do não trabalho (exceto para os ex-fumantes), taxa de ocupação e outros atributos familiares relativamente não exibem diferenças tão marcantes como as

¹¹Existe uma linha de trabalhos que investigam as diferentes formas de causalidade envolvendo educação e saúde, visto que o hábito de fumar pode revelar altas taxas de descontos intertemporais, que por sua vez pode resultar em baixos investimentos na qualificação. Maiores detalhes sobre esse tipo de discussão, por exemplo, em Conti et al. (2010) e Cutler & Lleras-Muney (2010)

duas supracitadas variáveis (educação e salário). Observando o salário-hora, na média, os fumantes recebem menos que os dois tipos de não fumantes, com diferenças estatisticamente significativas a pelo menos 1%. Sem condicionar a outros controles, o diferencial salarial entre fumantes e não usuários do cigarro do tipo 1 é de R\$ 8,74 por hora de trabalho em favor dos não fumantes.

Em conformidade com as características da amostra, a média de idade que os fumantes e ex-fumantes consumiram o cigarro pela primeira vez é de cerca de 16 anos, bem como o número médio de cigarros consumidos diariamente pelos atuais fumantes é de quase um maço de cigarro. Logo, uma iniciação prematura ao fumo combinada com uma elevada quantidade consumida dele ao dia podem ter um impacto expressivo no estoque de saúde dos fumantes, que com apenas 25 anos de idade, com base nesses prognósticos para a média, já teriam consumidos pelo menos 53,97 mil cigarros. Assim, a Tabela 3 reporta que fumantes e ex-fumantes possuem uma maior taxa de ausência no trabalho por motivos de saúde, quase o dobro da registrada para os não fumantes que nunca consumiram o cigarro. Esse fato pode traduzir em uma maior tendência de absenteísmo do local do trabalho para esse grupo de pessoas, refletindo em um menor retorno salarial (Levine et al. 1997).

A existência de dificuldade física para levantar objetos, praticar esportes ou realizar trabalhos pesados é outro ponto de disparidade, dada uma taxa de 8,5 pontos percentuais (p.p.) maior para os fumantes no comparativo com os não fumantes tipo 1, tendo os ex-fumantes a maior proporção de indivíduos com essa dificuldade (2% a mais que os fumantes). Considerando uma possível contaminação no grupo de não fumantes, que contém indivíduos com experiência pretérita de uso do cigarro – que, em média, aproximam-se mais dos fumantes em características observadas na Tabela 3 –, os modelos econométricos fazem filtro amostral para ex-fumantes¹².

Na finalidade de adicionar a heterogeneidade da penalização salarial provocada pelo uso do cigarro sem perder de vista os problemas de endogeneidade, na Subseção 5.2 são apresentados os resultados da pesquisa por meio do método RQVI. Contudo, a Subseção 5.1, como ponto de partida, expõe as estimativas calculadas para a função condicional média dos salários.

5.1 Resultados por MQO e MQ2E

Os coeficientes do modelo de regressão para a penalização salarial do tabagismo estão exibidos na Tabela 4. Essas estimativas servem como referência para os resultados obtidos através dos modelos com estimadores quantílicos. Do modelo (1) ao (4) são expostos os coeficientes por MQO sem e com controle na amostra para ex-fumantes. As duas últimas colunas da tabela exibem as estimativas por MQ2E, diferenciando-se apenas no controle amostral para ex-fumantes.

Em termos de ajustamento dos modelos, as variáveis explicativas comumente utilizadas pela literatura da área, escolaridade, idade (que capta o efeito da experiência) e outros atributos individuais e locais, mostram-se com os sinais esperados e com significância estatística. Trabalhadores com menos escolaridade e situados nas regiões Norte e Nordeste apresentam uma menor remuneração, enquanto que indivíduos de cor branca, sindicalizados,

¹²Esse filtro não induz a mudanças drásticas nos achados de penalização, como pode ser visualizado na Tabela 4, mas é utilizado para dar um maior rigor na análise.

Tabela 4: Resultados da regressão para a penalização salarial do fumante corrente no Brasil, PETab 2008. Variável dependente: logaritmo do salário-hora

Covariadas	MQO			MQ2E		
	(1)	(2)	(3)	(4)	(5)	(6)
Fumante	-0,3450 ^c (0,0272)	-0,1258 ^c (0,0240)	-0,1053 ^c (0,0212)	-0,0885 ^c (0,0195)	-0,2160 ^c (0,0463)	-0,2479 ^c (0,0526)
Faixa de escolaridade						
<EF		-0,6193 ^c (0,0759)	-0,3778 ^c (0,0843)	-0,3751 ^c (0,0728)	-0,3658 ^c (0,0730)	-0,3648 ^c (0,0844)
EF		-0,6248 ^c (0,0582)	-0,3818 ^c (0,0682)	-0,3730 ^c (0,0593)	-0,3704 ^c (0,0596)	-0,3772 ^c (0,0686)
EM		-0,5223 ^c (0,0494)	-0,3300 ^c (0,0584)	-0,3180 ^c (0,0515)	-0,3203 ^c (0,0520)	-0,3334 ^c (0,0591)
Raça		0,2249 ^c (0,0213)	0,1020 ^c (0,0201)	0,1043 ^c (0,0179)	0,1014 ^c (0,0179)	0,0971 ^c (0,0201)
Idade		0,0409 ^c (0,0102)	0,0308 ^c (0,0096)	0,0281 ^c (0,0084)	0,0286 ^c (0,0084)	0,0306 ^c (0,0095)
idade ² /100		-0,0515 ^c (0,0124)	-0,0401 ^c (0,0116)	-0,0375 ^c (0,0100)	-0,0376 ^c (0,0100)	-0,0382 ^c (0,0116)
Interação: Educação e Idade		0,0020 ^c (0,0001)	0,0011 ^c (0,0001)	0,0011 ^c (0,0001)	0,0011 ^c (0,0001)	0,0011 ^c (0,0001)
Sindicalizado			0,1248 ^c (0,0238)	0,1198 ^c (0,0208)	0,1130 ^c (0,0208)	0,1156 ^c (0,0238)
Estabilidade (tenure)			0,0095 ^c (0,0014)	0,0099 ^c (0,0012)	0,0098 ^c (0,0012)	0,0094 ^c (0,0014)
Grupos ocupacionais						
Dirigentes			0,7921 ^c (0,0544)	0,7853 ^c (0,0470)	0,7813 ^c (0,0477)	0,7881 ^c (0,0553)
Ciências e Artes			0,7320 ^c (0,0809)	0,7311 ^c (0,0691)	0,7221 ^c (0,0699)	0,7221 ^c (0,0819)
Técnicos			0,5782 ^c (0,0440)	0,5788 ^c (0,0412)	0,5774 ^c (0,0413)	0,5789 ^c (0,0442)
Serviços administrativos			0,1672 ^c (0,0408)	0,1916 ^c (0,0366)	0,1869 ^c (0,0369)	0,1618 ^c (0,0411)
Vendedores			0,1679 ^c (0,0379)	0,1381 ^c (0,0377)	0,1313 ^c (0,0380)	0,1603 ^c (0,0384)
Agrícola			-0,2135 ^c (0,0396)	-0,2116 ^c (0,0356)	-0,2061 ^c (0,0358)	-0,2058 ^c (0,0398)
Produção e manutenção			0,1582 ^c (0,0243)	0,1797 ^c (0,0216)	0,1817 ^c (0,0217)	0,1616 ^c (0,0246)
Fatores locacionais						
Urbana			0,0985 ^b (0,0389)	0,1160 ^c (0,0340)	0,1175 ^c (0,0339)	0,1008 ^c (0,0387)
Região metropolitana			-0,0236 (0,0187)	-0,0244 (0,0169)	-0,0242 (0,0169)	-0,0228 (0,0187)
Norte			-0,1569 ^c (0,0323)	-0,1272 ^c (0,0281)	-0,1277 ^c (0,0281)	-0,1550 ^c (0,0325)
Nordeste			-0,4062 ^c (0,0254)	-0,3921 ^c (0,0225)	-0,3943 ^c (0,0225)	-0,4096 ^c (0,0254)
Sul			-0,0053 (0,0273)	0,0104 (0,0243)	0,0114 (0,0243)	-0,0035 (0,0272)
Centro-Oeste			-0,0085 (0,0288)	0,0274 (0,0273)	0,0241 (0,0274)	-0,0118 (0,0291)
λ			-0,4258 ^c (0,0822)	-0,4383 ^c (0,0748)	-0,4251 ^c (0,0749)	-0,4041 ^c (0,0822)
Constante	3,0164 ^c (0,0153)	2,0441 ^c (0,1957)	2,1590 ^c (0,1840)	2,1750 ^c (0,1648)	2,1867 ^c (0,1653)	2,1830 ^c (0,1847)
Controle: ex-fumantes	X	X	X			X
Instrumentos					X	X
N	6.979	6.979	6.979	8.808	8.808	6.979
P (em milhões)	24,56	24,56	24,56	31,22	31,22	24,56
R ² ajustado	0,03	0,37	0,49	0,47	0,47	0,49

Fonte: Elaboração própria a partir dos microdados da PNAD/PETab 2008. Erros-padrão robustos entre parênteses, corrigidos por correlação intragrupo e independência intergrupos.

Instrumentos: proibido fumar no domicílio, crença/percepção sobre as consequências do cigarro (gera doenças aos outros e para si), fumantes na família, existência de ex-fumantes na família e preço do cigarro.

Nota: O λ nesses modelos representa a razão inversa de Mills tradicional proposta por Heckman (1979).

Legenda: ^ap<0,10; ^bp<0,05; ^cp<0,01.

com maior tempo de serviço na mesma empresa, residentes em áreas urbanas e que ocupam, em especial, cargos de dirigentes e atuam como profissionais nas áreas das ciências e das artes possuem maiores taxas de salários. Além do mais, o grau de determinação do modelo, R^2 ajustado, é de pelo menos 47%, bem como a razão inversa de Mills tradicional (λ) tem significância estatística nas especificações em que ela é incluída, sinalizando a importância da correção do viés de seleção¹³. O $\lambda < 0$ expressa que os fatores latentes na equação de ganhos minceriana, que se relacionam com uma maior probabilidade de participação na força de trabalho, diminuem a taxa de retornos salariais.

O parâmetro associado aos indivíduos fumantes nos modelos sem VI é negativo e significativo estatisticamente para todas as especificações. Ademais, no modelo (1), que não controla os atributos pessoais, ocupacionais e locais dos indivíduos, o coeficiente do fumante é de -0,345, que é atenuado em termos absolutos ao ponderar os demais conjuntos de variáveis explicativas, situando-se em -0,105 no modelo (3). Muito embora, o acréscimo das características individuais, ocupacionais e locais reduzem os efeitos do tabagismo no logaritmo do salário-hora, os custos do tabagismo sobre os rendimentos individuais podem ser tendenciosos caso variáveis não observadas se correlacionem com os salários e com a decisão de fumar.

Após o controle de endogeneidade, nota-se que o fumante ganha em média 19,4% a menos que um não fumante¹⁴, uma taxa de desconto bem acima das estimativas desenvolvidas por MQO (respectivamente, 10% e 8,5% nos modelos 3 e 4), indicando a existência em valores absolutos de uma possível subestimação no citado efeito pela regressão clássica, inclusive do erro-padrão das estimativas. Tais achados seguem a tendência da literatura internacional, como ilustrado na Tabela 1. Em Auld (2005), por exemplo, após o controle para endogeneidade, o uso do cigarro provoca um impacto de -24% no salário de trabalhadores canadenses com baixa sensibilidade a escolha dos instrumentos. Na presente pesquisa, as diferentes especificações de VI, apresentadas na Tabela A.3 no Apêndice, mostram que o intervalo para a penalização para o caso nacional situa-se entre -0,30 e -0,22.

A amostra contaminada pelos ex-fumantes reduz também os efeitos da penalização salarial. Ao excluir da base amostral os ex-fumantes, ocorre uma ampliação da penalização salarial em 1,52 p.p. entre os modelos (3) e (4) e de 2,53 p.p. entre os modelos (6) e (5). Como os ex-fumantes recebem, em média, uma remuneração inferior do que aqueles que nunca fumaram, a base de comparação tende a subestimar os custos salariais do hábito de fumar. Desta maneira, a permanência dos ex-fumantes e o não controle para endogeneidade geram uma tendência para baixo no valor absoluto do coeficiente da penalização, distorcendo os custos monetários implícitos do uso do cigarro¹⁵.

¹³As estimativas usadas para calcular a razão inversa de Mills, por meio da equação de participação na força de trabalho, encontram-se na Tabela A.1 no Apêndice deste trabalho.

¹⁴Baseado em Halvorsen & Palmquist (1980), o valor da semielasticidade para o coeficiente do fumante (por exemplo, -0,216), que é um regressor binário, é calculado daqui em diante por: $(e^{-0,216} - 1) \times 100 = 19,4\%$.

¹⁵A Tabela A.4 no Apêndice reporta outra especificação do modelo de penalização salarial, incorporando covariadas sobre os ramos da atividade e ocupação informal para os modelos 3 e 6 da Tabela 4. Com base nesses resultados, o coeficiente para os fumantes permanecem negativos e com uma variação de -0,105 no modelo 3 para -0,094 no modelo 3' e de -0,248 no modelo 6 para -0,23 no modelo 6', o que de certa maneira mostra a robustez dos resultados frente a diferentes especificações.

Quanto ao MQ2E, como pode ser verificado na Tabela A.2, os resultados condicionados ao vetor de instrumentos exibem um ajuste aceitável na determinação da decisão de fumar do indivíduo, de acordo com Stock et al. (2002), Stock & Yogo (2005). No tratamento do problema de endogeneidade, três testes são comumente usados para mensurar o ajustamento da modelagem com VI: (i) teste de endogeneidade; (ii) condições de restrições de sobre-identificação dos instrumentos; e, (iii) resistência dos instrumentos. A Tabela A.2 no Apêndice reporta as estatísticas dos supracitados testes.

Inicialmente, embasado na amostra em estudo, a decisão de fumar é endógena? Caso ela seja exógena, o resultado por MQO seria mais eficiente do que a abordagem de VI (Wooldridge 2010). Pelo teste de endogeneidade, a hipótese de exogeneidade para a variável H_i não pode ser aceita, em ambos os testes (Durbin e Wu-Hausman). Com isso, os resultados dos modelos (1) ao (4), supondo a sua exogeneidade, são tendenciosos. Outro teste importante é o de análise de sobreidentificação das restrições, para situações em que a quantidade de instrumentos supera o número de covariadas endógenas. Esse teste admite a validade de um instrumento para em seguida observar a validade dos demais (ou seja, se os instrumentos não estão correlacionados com o termo de erro na segunda etapa). Sendo assim, a hipótese de sobreidentificação é rejeitada pelos valores estatísticos dos testes de restrição.

Para uma variável ser um instrumento válido, ela deve ser suficientemente correlacionada com o regressor endógeno mas não correlacionada com o termo de erro. Na Tabela A.2 estão disponíveis estatísticas sobre o poder explicativo dos instrumentos. Pelo ajustamento do primeiro estágio através do teste F, todos os instrumentos apresentam, de forma conjunta, significância estatística (visto o p-valor = 0), bem como o valor de $F > 11$ indica que os instrumentos são suficientemente confiáveis a pelo menos um viés relativo de 10% (Stock et al. 2002). Conforme os valores críticos calculados a partir do teste de identificação de Stock & Yogo (2005), as estatísticas de Cragg-Donald e Kleibergen-Paap estão bem acima dos pontos críticos, implicando que os instrumentos não podem ser taxados como fracos. Ademais, os valores para o R^2 constantes na Tabela A.2 apresentam valores maiores que 20%, ratificando o grau de ajuste dos instrumentos utilizados para identificação dos fumantes.

5.2 Resultados por Regressão quantílica e RQVI

Os efeitos do cigarro sobre o salário por MQO e MQ2E revelam apenas a tendência central da relação entre essas duas variáveis, não captando possíveis heterogeneidades do efeito ao longo da distribuição condicional dos salários. Os coeficientes da penalização salarial do fumante corrente através de estimadores quantílicos com controle da amostra para ex-fumantes estão reportados nas Tabelas 5 e 6, com utilização da abordagem RQVI na última tabela¹⁶. Os valores de τ estão dispostos para os quantis: 0,10; 0,25; 0,50; 0,75 e 0,90.

As variáveis de controles usuais se mostram com os sinais esperados e com significância estatística, tal como os achados para a função condicional média. Os estimadores quantílicos sem e com controle para endogeneidade da decisão de fumar sustentam a existência de um custo salarial para o uso do cigarro, independentemente da localização dos indivíduos em relação à distribuição do salário-hora. Em outras palavras, mesmo segmentando a função

¹⁶As estimativas usadas para realizar a aproximação da razão inversa de Mills de segunda ordem, seguindo a proposta de Buchinsky (1998), estão dispostas na Tabela A.1 no Apêndice.

Tabela 5: Resultados para a penalização salarial do fumante corrente por RQ no Brasil, PETab 2008. Variável dependente: logaritmo do salário-hora

Covariadas	Q(0,10)	Q(0,25)	Q(0,50)	Q(0,75)	Q(0,90)
Fumante	-0,1296 ^c (0,0263)	-0,1013 ^c (0,0189)	-0,0743 ^c (0,0166)	-0,0678 ^c (0,0176)	-0,1006 ^c (0,0238)
Faixa de escolaridade					
<EF	-0,2462 ^c (0,0849)	-0,3709 ^c (0,0716)	-0,5000 ^c (0,0729)	-0,5746 ^c (0,0702)	-0,3895 ^c (0,0949)
EF	-0,2551 ^c (0,0567)	-0,3765 ^c (0,0609)	-0,4980 ^c (0,0641)	-0,5972 ^c (0,0590)	-0,4081 ^c (0,0777)
EM	-0,2365 ^c (0,0411)	-0,3103 ^c (0,0544)	-0,4137 ^c (0,0599)	-0,4802 ^c (0,0541)	-0,2984 ^c (0,0683)
Raça	0,0335 (0,0246)	0,0414 ^b (0,0163)	0,0617 ^c (0,0160)	0,0678 ^c (0,0176)	0,1505 ^c (0,0296)
Idade	0,0197 ^a (0,0103)	0,0244 ^c (0,0084)	0,0266 ^c (0,0075)	0,0221 ^b (0,0086)	0,0283 ^b (0,0123)
Idade ² /100	-0,0385 ^c (0,0122)	-0,0361 ^c (0,0106)	-0,0353 ^c (0,0092)	-0,0285 ^c (0,0108)	-0,0352 ^b (0,0158)
Interação: Educação e Idade	0,0012 ^c (0,0002)	0,0009 ^c (0,0001)	0,0010 ^c (0,0001)	0,0011 ^c (0,0001)	0,0014 ^c (0,0002)
Sindicalizado	0,1345 ^c (0,0256)	0,1355 ^c (0,0196)	0,1301 ^c (0,0165)	0,1353 ^c (0,0231)	0,1473 ^c (0,0343)
Estabilidade (<i>tenure</i>)	0,0086 ^c (0,0015)	0,0066 ^c (0,0012)	0,0081 ^c (0,0011)	0,0115 ^c (0,0013)	0,0161 ^c (0,0018)
Grupos ocupacionais					
Dirigentes	0,3199 ^c (0,0744)	0,6331 ^c (0,0560)	0,7927 ^c (0,0407)	0,9953 ^c (0,0527)	1,1982 ^c (0,0758)
Ciências e Artes	0,4939 ^c (0,0741)	0,6645 ^c (0,0616)	0,7456 ^c (0,0544)	0,8331 ^c (0,0735)	1,0858 ^c (0,1139)
Técnicos	0,3021 ^c (0,0389)	0,4379 ^c (0,0416)	0,5636 ^c (0,0462)	0,6625 ^c (0,0444)	0,8549 ^c (0,0814)
Serviços administrativos	0,0675 ^a (0,0410)	0,1658 ^c (0,0388)	0,2056 ^c (0,0311)	0,1798 ^c (0,0388)	0,1642 ^c (0,0395)
Vendedores	0,0325 (0,0428)	0,0586 ^b (0,0241)	0,0840 ^c (0,0316)	0,2164 ^c (0,0607)	0,3329 ^c (0,0660)
Agrícola	-0,5108 ^c (0,0645)	-0,2848 ^c (0,0377)	-0,1832 ^c (0,0312)	-0,0550 ^a (0,0331)	0,0073 (0,0523)
Produção e manutenção	0,0349 (0,0358)	0,1225 ^c (0,0234)	0,1521 ^c (0,0203)	0,1699 ^c (0,0204)	0,1685 ^c (0,0349)
Fatores locacionais					
Urbana	0,1765 ^c (0,0461)	0,1092 ^c (0,0221)	0,0796 ^c (0,0255)	0,1102 ^c (0,0307)	0,1120 ^b (0,0455)
Região metropolitana	-0,0064 (0,0229)	-0,0033 (0,0157)	-0,0184 (0,0154)	-0,0443 ^c (0,0165)	-0,0577 ^b (0,0247)
Norte	-0,2059 ^c (0,0411)	-0,2158 ^c (0,0260)	-0,1697 ^c (0,0263)	-0,1281 ^c (0,0267)	-0,0873 ^a (0,0517)
Nordeste	-0,4869 ^c (0,0345)	-0,4162 ^c (0,0210)	-0,3473 ^c (0,0191)	-0,3286 ^c (0,0226)	-0,3084 ^c (0,0275)
Sul	0,0107 (0,0256)	0,0190 (0,0242)	0,0185 (0,0236)	0,0245 (0,0259)	0,0124 (0,0345)
Centro-Oeste	-0,0332 (0,0383)	-0,0605 ^b (0,0241)	0,0154 (0,0256)	0,0258 (0,0252)	-0,0004 (0,0514)
h_1	-0,0829 (0,1371)	-0,0772 ^b (0,0385)	-0,0995 (0,0689)	-0,0653 ^b (0,0284)	-0,0445 (0,0415)
h_2	-0,4396 ^b (0,2001)	-0,3783 ^c (0,0885)	-0,2678 ^b (0,1089)	-0,3030 ^c (0,0483)	-0,1137 (0,0836)
Constante	1,8803 ^c (0,2038)	2,1128 ^c (0,1663)	2,3600 ^c (0,1565)	2,6979 ^c (0,1721)	2,5489 ^c (0,2405)
Controle amostral: ex-fumantes	X	X	X	X	X
Instrumentos					
N	6.979	6.979	6.979	6.979	6.979
P (em milhões)	24,56	24,56	24,56	24,56	24,56
Pseudo-R ²	0,2634	0,2475	0,3033	0,3588	0,3718

Fonte: Elaboração própria a partir dos microdados da PNAD/PETab 2008. Erros-padrão robustos entre parênteses, corrigidos por correlação intragrupo e independência intergrupos.

Nota: h_1 e h_2 representam os dois termos de correção do viés de seleção para estimadores quantílicos.

Legenda: ^ap<0,10; ^bp<0,05; ^cp<0,01

Tabela 6: Resultados para a penalização salarial do fumante corrente por RQVI no Brasil, PETab 2008. Variável dependente: logaritmo do salário-hora

Covariadas	Q(0,10)	Q(0,25)	Q(0,50)	Q(0,75)	Q(0,90)
Fumante	-0,1644 ^c (0,0597)	-0,2310 ^c (0,0371)	-0,2019 ^c (0,0358)	-0,1766 ^c (0,0391)	-0,4547 ^c (0,0664)
Faixa de escolaridade					
<EF	-0,2866 ^c (0,0924)	-0,3376 ^c (0,0746)	-0,4669 ^c (0,0688)	-0,5014 ^c (0,0716)	-0,3641 ^c (0,1098)
EF	-0,2865 ^c (0,0699)	-0,3412 ^c (0,0645)	-0,4482 ^c (0,0584)	-0,5155 ^c (0,0618)	-0,4163 ^c (0,0975)
EM	-0,2383 ^c (0,0582)	-0,2867 ^c (0,0588)	-0,3847 ^c (0,0518)	-0,4364 ^c (0,0543)	-0,3476 ^c (0,0840)
Raça	0,0358 (0,0252)	0,0375 ^b (0,0173)	0,0785 ^c (0,0173)	0,0640 ^c (0,0195)	0,1525 ^c (0,0301)
Idade	0,0232 ^a (0,0120)	0,0208 ^b (0,0090)	0,0281 ^c (0,0074)	0,0262 ^c (0,0093)	0,0358 ^c (0,0126)
Idade ² /100	-0,0403 ^c (0,0149)	-0,0298 ^c (0,0113)	-0,0352 ^c (0,0092)	-0,0333 ^c (0,0114)	-0,0412 ^c (0,0155)
Interação: Educação e Idade	0,0010 ^c (0,0002)	0,0008 ^c (0,0001)	0,0008 ^c (0,0001)	0,0010 ^c (0,0001)	0,0012 ^c (0,0002)
Sindicalizado	0,1163 ^c (0,0300)	0,1294 ^c (0,0207)	0,1132 ^c (0,0190)	0,1292 ^c (0,0252)	0,1107 ^c (0,0357)
Estabilidade (<i>tenure</i>)	0,0085 ^c (0,0018)	0,0064 ^c (0,0013)	0,0081 ^c (0,0012)	0,0108 ^c (0,0014)	0,0114 ^c (0,0022)
Grupos ocupacionais					
Dirigentes	0,3274 ^c (0,0625)	0,6452 ^c (0,0557)	0,7639 ^c (0,0410)	0,9185 ^c (0,0548)	1,1071 ^c (0,0963)
Ciências e Artes	0,4925 ^c (0,0746)	0,6669 ^c (0,0664)	0,7427 ^c (0,0614)	0,8197 ^c (0,0775)	0,9251 ^c (0,1101)
Técnicos	0,3268 ^c (0,0482)	0,4372 ^c (0,0381)	0,5396 ^c (0,0446)	0,6334 ^c (0,0469)	0,7719 ^c (0,0738)
Serviços administrativos	0,0995 ^b (0,0459)	0,1554 ^c (0,0440)	0,2160 ^c (0,0357)	0,1853 ^c (0,0428)	0,1010 ^a (0,0597)
Vendedores	0,0375 (0,0458)	0,0372 (0,0298)	0,0740 ^b (0,0347)	0,2053 ^c (0,0511)	0,3633 ^c (0,0719)
Agrícola	-0,4976 ^c (0,0725)	-0,2867 ^c (0,0429)	-0,1611 ^c (0,0343)	-0,0631 ^a (0,0356)	0,0474 (0,0513)
Produção e manutenção	0,0541 (0,0343)	0,1191 ^c (0,0225)	0,1582 ^c (0,0214)	0,1809 ^c (0,0244)	0,1766 ^c (0,0399)
Fatores locacionais					
Urbana	0,1519 ^c (0,0521)	0,0786 ^b (0,0319)	0,0934 ^c (0,0288)	0,1097 ^c (0,0290)	0,1204 ^c (0,0391)
Região metropolitana	-0,0163 (0,0250)	-0,0079 (0,0170)	-0,0333 ^b (0,0159)	-0,0497 ^c (0,0187)	-0,0308 (0,0303)
Norte	-0,2432 ^c (0,0404)	-0,2173 ^c (0,0297)	-0,1490 ^c (0,0271)	-0,1464 ^c (0,0319)	-0,0795 ^a (0,0452)
Nordeste	-0,5032 ^c (0,0385)	-0,4046 ^c (0,0222)	-0,3493 ^c (0,0202)	-0,3211 ^c (0,0244)	-0,2837 ^c (0,0353)
Sul	0,0391 (0,0348)	0,0432 ^a (0,0242)	0,0740 ^c (0,0248)	0,0832 ^c (0,0271)	0,1228 ^c (0,0459)
Centro-Oeste	-0,0653 ^a (0,0371)	-0,0407 (0,0271)	0,0224 (0,0280)	0,0390 (0,0302)	0,0505 (0,0509)
h_1	-0,0091 (0,0426)	-0,0665 ^a (0,0356)	-0,0540 (0,0528)	-0,0473 (0,0910)	-0,0618 (0,1728)
h_2	-0,5403 ^c (0,0645)	-0,4032 ^c (0,0614)	-0,3081 ^c (0,0889)	-0,3448 ^b (0,1543)	-0,2669 (0,3062)
Constante	1,4591 ^c (0,2634)	1,7762 ^c (0,1967)	1,8458 ^c (0,1717)	1,9957 ^c (0,2074)	1,4620 ^c (0,2996)
Controle amostral: ex-fumantes	X	X	X	X	X
Instrumentos	X	X	X	X	X
N	6.979	6.979	6.979	6.979	6.979
P (em milhões)	24,56	24,56	24,56	24,56	24,56

Fonte: Elaboração própria a partir dos microdados da PNAD/PETab 2008. Erros-padrão robustos entre parênteses, corrigidos por correlação intragrupo e independência intergrupos. Instrumentos: proibido fumar no domicílio, crença/percepção sobre as consequências do cigarro (gera doenças aos outros e para si), fumantes na família, existência de ex-fumantes na família e preço do cigarro.

Nota: h_1 e h_2 representam os dois termos de correção do viés de seleção para estimadores quantílicos.

Legenda: ^ap<0,10; ^bp<0,05; ^cp<0,01.

condicional de ganhos para trabalhadores com baixo e elevado rendimento, os fumantes indistintamente da sua posição na distribuição possui uma menor taxa de salário quando se confronta com os não fumantes.

Admitindo a influência dos fatores não observados na decisão de fumar, as estimativas da RQ, mesmo considerando os vetores de variáveis explanatórias (individuais, ocupacionais e locacionais), ainda seriam tendenciosas. Na intenção de facilitar a visualização e comparação dos efeitos do tabagismo sobre os salários, a Figura 1 exibe as estimativas por RQ e RQVI, inclusive o tamanho do viés da estimativa da penalização salarial entre as duas abordagens¹⁷. Por isso, os resultados expressos na Tabela 6 mostram-se condizentes com as demandas técnicas requeridas nos trabalhos empíricos da área.

Pela Figura 1, os custos monetários indiretos do tabagismo são maiores pela abordagem com correção para endogeneidade, sugerindo que os efeitos não corrigidos para o problema da $Cov_{\tau}(H_i, \epsilon_{1i}) \neq 0$ são minorados. Resultado similar foi encontrado entre as estimativas para a função condicional média (MQO versus MQ2E). Enquanto o efeito por RQ varia de -12,2% a -9,6% entre os quantis 0,10 e 0,90, por RQVI estes valores variam de -15,2% a -36,5%.

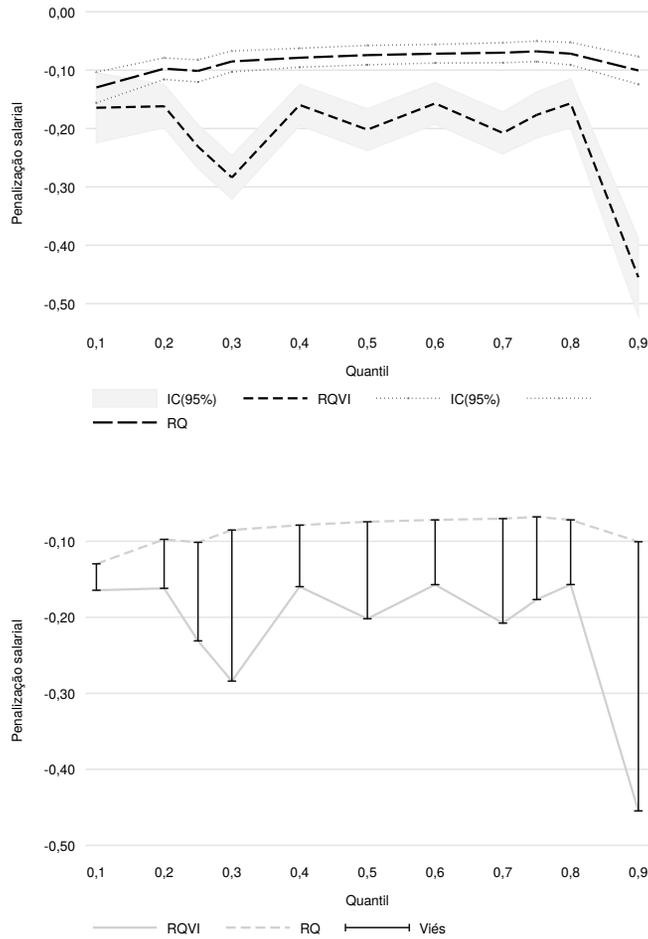
Também por meio da Figura 1, fica evidente que a maior parte das estimativas por RQ, praticamente, não apresentam interseções com as calculadas por RQVI, mesmo considerando o intervalo de confiança (IC) de 95% – a exceção fica por parte apenas do coeficiente no quantil 0,10 entre os limites de RQ e RQVI. O viés médio entre os dois métodos é superior a 15%, com destaque para a diferença registrada no quantil 0,90. Uma das possibilidades de explicação desse resultado pode estar atrelada aos maiores níveis de aptidão técnica demandados pelas atividades com maior taxa de retorno salarial no setor privado, no qual indivíduos com maior estoque de capital humano tendem a se sobressair. Então, a junção da perda relativa de capacidade física dos fumantes, como pode ser vista na Tabela 3, com a desvantagem da formação do capital humano motivado por menores investimentos sistemáticos no estoque de saúde são prováveis fatores explicativos.

Como pode ser verificado na Figura A.1 no Apêndice deste trabalho (que mostra a distribuição de usuários de cigarros por decil de renda), a maior parte dos trabalhadores fumantes encontram-se situados nos *loci* de menor salário-hora, muito embora existam frequência desses indivíduos em todo os agrupamentos. Enquanto no primeiro decil, 60% dos indivíduos são fumantes, no último decil de renda, apenas 13,1% dos trabalhadores usam cigarro. Pelas informações levantadas nesta pesquisa, verifica-se que os trabalhadores usuários de cigarros têm uma penalização salarial ao longo da distribuição condicional do rendimento individual, tendo uma distribuição de fumantes com maior concentração relativa na região de menor salário-hora.

Além do tabagismo ser um dos principais causadores de mortes dentre os fatores de risco modificáveis (World Health Organization 2009), os usuários do cigarro no Brasil, conforme sugerido pelos presentes resultados, arcam também com elevados custos monetários indiretos – visto que, independentemente da abordagem metodológica, existe uma associação negativa entre o tabagismo e o rendimento do trabalho. Com evidências semelhantes para a

¹⁷Através dos coeficientes obtidos pela RQ convencional, que supõem a exogeneidade da decisão de fumar, os usuários do cigarro recebem, na mediana, 7% a menos que os não fumantes, com significância estatística a pelo menos 1%. Para os percentis 10, 25, 75 e 90, as estimativas mantêm idêntico sinal da tendência central, mas com magnitudes maiores em termos absolutos, com exceção para o quantil 0,75.

Figura 1: Comparação dos efeitos do tabagismo por quantil da distribuição condicional dos rendimentos do trabalhador no Brasil entre RQ e RQVI, 2008



Nota: A amostra total é composta por 6.979 observações. O intervalo de confiança (IC) das estimativas considerado foi de 95%, usando erros-padrão robustos. As estimativas são reportadas para $\tau \in [0, 10; 0, 25; 0, 50; 0, 75; 0, 90]$.

Finlândia, Bockerman et al. (2014) questionam até que ponto os usuários do cigarro conseguem calcular de forma racional a relação de custo-benefício do produto como admitido em Becker & Murphy (1988).

6 Conclusões

As evidências do presente estudo suportam a existência de uma penalização salarial do tabagismo no mercado de trabalho brasileiro, que se sustentam mesmo após uma série de controles para endogeneidade, amostra, atributos observáveis individuais e posição do trabalhador na distribuição condicional dos rendimentos. Usando diferentes instrumentais econométricos, com ênfase no modelo RQVI, o usuário do tabaco no Brasil arca com um custo monetário indireto do produto, percebido pelos menores níveis da taxa de salário-hora. Complementando a literatura da área, as estimativas indicam que as penalizações salariais não se mostram constantes em diferentes partes da distribuição dos ganhos. Assim, as perdas, controlando para endogeneidade, variam de 15,2% a 36,5%, centrando em torno de 18,3% na mediana e alcançando 36,5% para os trabalhadores no extremo positivo da distribuição salarial.

A elevada penalização salarial encontrada no presente estudo mostra que as sequelas do tabagismo vão, em termos individuais, além das questões de saúde, algo já consolidado na literatura clínica (Newcomb & Carbone 1992, Hecht 2006). A menor produtividade do trabalho observada entre os fumantes repercute em um menor patamar de renda disponível, inclusive, para remédios e tratamentos de doenças decorrentes das próprias substâncias malélicas do produto ao organismo. Assim, políticas anti-tabagismo podem gerar impactos não apenas nas condições de saúde, como redução da taxa de mortalidade e aumento no tempo de sobrevivência, mas também em melhores indicadores econômicos no prisma individual e social.

Por fim, este estudo faz uma contribuição importante para a literatura nacional, pois é um dos primeiros a relacionar o tabagismo e os resultados no mercado de trabalho no Brasil, bem como incorpora a heterogeneidade dos efeitos do uso do cigarro nos salários. Embora os dados utilizados não sejam longitudinais (fato ainda recorrente nas pesquisas estatísticas nacionais), esse tipo de avaliação dos custos e das consequências do uso de substâncias nocivas à saúde deveriam ser mais enfatizadas no país, a fim de melhorar a alocação de recursos para combate aos problemas relacionados ao uso do cigarro e melhor compreender as relações envolvendo escolhas não saudáveis e resultados econômicos e não econômicos. Ademais, os achados desta pesquisa provocam novos questionamentos, como um melhor entendimento dos mecanismos causais no país entre o tabagismo e a perda de produtividade do trabalho, que poderiam ser investigados em pesquisas futuras.

Referências Bibliográficas

Alves, L. F. & Andrade, M. V. (2003), 'Impactos da saúde nos rendimentos individuais no Brasil', *Economia Aplicada* 7(2), 259–388.

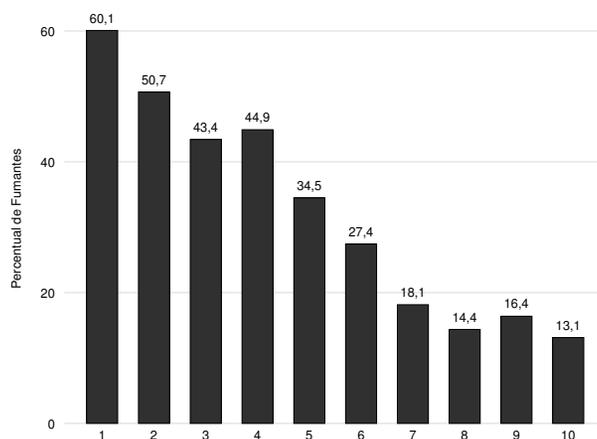
Anger, S. & Kvasnicka, M. (2010), 'Does Smoking Really Harm Your Earnings so Much? Biases in Current Estimates of the Smoking Wage Penalty', *Applied Economics Letters* 17(6), 561–564.

- Auld, M. C. (2005), 'Smoking, Drinking, and Income', *The Journal of Human Resources* **40**(2), 505–518.
- Bartel, A. & Taubman, P. (1979), 'Health and Labor Market Success: The Role of Various Diseases', *The Review of Economics and Statistics* **61**(1), 1–8.
- Becker, G. S. & Murphy, K. M. (1988), 'A Theory of Rational Addiction', *Journal of Political Economy* **96**(4), 675–700.
- Bockerman, P., Hyytinen, A. & Kaprio, J. (2014), 'Smoking and long-term labour market outcomes', *Tobacco control* pp. 1–7.
URL: <http://www.ncbi.nlm.nih.gov/pubmed/24570102>
- Braakmann, N. (2008), 'The smoking wage penalty in the United Kingdom: Regression and matching evidence from the British Household Panel Survey', *University of Luneburg Working Paper Series N. 96* pp. 1–33.
- Buchinsky, M. (1998), 'The dynamics of changes in the female wage distribution in the USA: a quantile regression approach', *Journal of Applied Econometrics* **13**(1), 1–30.
- Buchinsky, M. (2001), 'Quantile regression with sample selection: Estimating women's return to education in the U.S.', *Empirical Economics* **26**(1), 87–113.
- Chernozhukov, V. & Hansen, C. (2004), 'The Effects of 401(k) Participation on the Wealth Distribution: An Instrumental Quantile Regression Analysis', *The Review of Economics and Statistics* **86**(3), 735–751.
- Chernozhukov, V. & Hansen, C. (2005), 'Notes and comments an IV model of quantile treatment effects', *Econometrica* **73**(1), 245–261.
- Chernozhukov, V. & Hansen, C. (2013), 'Quantile Models with Endogeneity', *Annual Review of Economics* **5**(1), 57–81.
URL: <http://www.annualreviews.org/doi/abs/10.1146/annurev-economics-080511-110952>
- Chernozhukov, V., Hansen, C. & Jansson, M. (2007), 'Inference approaches for instrumental variable quantile regression', *Economics Letters* **95**(2), 272–277.
URL: <http://linkinghub.elsevier.com/retrieve/pii/S0165176506003703>
- Conti, G., Heckman, J. & Urzua, S. (2010), 'The Education-Health Gradient', *American Economic Review* **100**(2), 234–238.
- Cutler, D. & Glaeser, E. (2005), 'What explains differences in smoking, drinking and other health-related behaviors?', *NBER Working paper series n. 11100* pp. 1–20.
- Cutler, D. M. & Lleras-Muney, A. (2010), 'Understanding differences in health behaviors by education', *Journal of Health Economics* **29**(1), 1–28.
- De Luca, G. (2008), 'The Stata Journal', *The Stata Journal* **8**(2), 190–220.
- Deolalikar, A. B. (1988), 'Nutrition and Labor Productivity in Agriculture: Estimates for Rural South India', *The Review of Economics and Statistics* **70**(3), 406–413.

- Fernandez, R. M. & Nordman, C. J. (2009), 'Are there pecuniary compensations for working conditions?', *Labour Economics* **16**(2), 194–207.
URL: <http://linkinghub.elsevier.com/retrieve/pii/S0927537108000791>
- Gallant, A. R. & Nychka, D. W. (1987), 'Semi-Nonparametric Maximum Likelihood Estimation', *Econometrica* **55**(2), 363–390.
- Grafova, I. B. & Stafford, F. P. (2009), 'The Wage Effects of Personal Smoking History', *Industrial and Labor Relations Review* **62**(3), 381.
- Grossman, M. (1972), 'On the Concept of Health Capital and the Demand for Health', *Journal of Political Economy* **80**(2), 223–255.
- Halvorsen, R. & Palmquist, R. (1980), 'The Interpretation of Dummy Variables in Semilogarithmic Equations', *The American Economic Review* **70**(3), 474–475.
- Hecht, S. S. (2006), 'Cigarette smoking: cancer risks, carcinogens, and mechanisms', *Langenbeck's Archives of Surgery* **391**(6), 603–613.
- Heckman, J. J. (1979), 'Sample Selection Bias as a Specification Error', *Econometrica* **47**(1), 153–161.
- Hersch, J. & Viscusi, W. K. (1990), 'Cigarette Smoking, Seatbelt Use, and Differences in Wage-Risk Tradeoffs', *The Journal of Human Resources* **25**(2), 202–227.
- Instituto Brasileiro de Geografia e Estatística (2009), *Pesquisa Nacional por Amostra de Domicílios: Tabagismo*, IBGE, Rio de Janeiro.
- Instituto Brasileiro de Geografia e Estatística (2013), *Atlas do Censo Demográfico 2010*, IBGE, Rio de Janeiro.
- Kassouf, A. L. (1997), 'Saúde e mercado de trabalho', *Pesquisa e Planejamento Econômico* **27**(3), 587–610.
- Koenker, R. (2005), *Quantile Regression*, Cambridge University Press, Cambridge.
- Koenker, R. & Bassett, G. (1978), 'Regression Quantiles Roger', *Econometrica* **46**(1), 33–50.
- Leigh, J. P. & Berger, M. C. (1989), 'Effects of smoking and being overweight on current earnings', *American Journal of Preventive Medicine* **5**(1), 8–14.
- Levine, P. B., Gustafson, T. A. & Velenchik, A. D. (1997), 'More bad news for smokers? The effects of cigarette smoking on wages', *Industrial and Labor Relations Review* **50**(3), 493–509.
- Lokshin, M. & Beegle, K. (2006), 'Forgone earnings from smoking: evidence for a developing country', *World Bank Policy Research Working Paper n. 4018* pp. 1–27.
- Mincer, J. (1974), *Schooling, Experience, and Earnings*, Vol. I, National Bureau of Economic Research, Massachusetts.
URL: <http://www.nber.org/books/minc74-1>

- Munasinghe, L. & Sicherman, N. (2006), 'Why Do Dancers Smoke? Smoking, Time Preference, and Wage Dynamics', *Eastern Economic Journal* 32(4), 595–616.
- Newcomb, P. A. & Carbone, P. P. (1992), 'The health consequences of smoking, cancer', *The Medical clinics of North America* 76(2), 305–331.
- Psacharopoulos, G. & Patrinos, H. A. (2004), Human capital and rates of return, in G. Johnes & J. Johnes, eds, 'International Handbook on the Economics of Education', Edward Elgar Publishing Ltd, Cheltenham, pp. 1–57.
- Schultz, T. W. (1961), 'Investment in Human Capital', *The American Economic Review* 51(1), 1–17.
- Schwarze, J. & Heineck, G. (2003), 'Substance use and earnings: the case of smokers in Germany', *IZA Working Paper n. 743*.
- Stock, J. H., Wright, J. H. & Yogo, M. (2002), 'A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments', *Journal of Business & Economic Statistics* 20(4), 518–529.
URL: <http://www.tandfonline.com/doi/abs/10.1198/073500102288618658>
- Stock, J. H. & Yogo, M. (2005), Testing for Weak Instruments in Linear IV Regression, in D. W. K. Andrews & J. H. Stock, eds, 'Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg', Cambridge University Press, Cambridge, pp. 80–108.
- Strauss, J. & Thomas, D. (1998), 'Health, Nutrition, and Economic Development', *Journal of Economic Literature* 36(2), 766–817.
- Thomas, D. & Strauss, J. (1997), 'Health and wages: Evidence on men and women in urban Brazil', *Journal of Econometrics* 77, 159–185.
- Van Ours, J. C. (2004), 'A pint a day raises a man's pay; but smoking blows that gain away', *Journal of Health Economics* 23(5), 863–86.
URL: <http://www.ncbi.nlm.nih.gov/pubmed/15353183>
- Viscusi, W. K. & Hersch, J. (2001), 'Cigarette smokers as job risk takers', *The Review of Economics and Statistics* 83(2), 269–280.
- Willis, R. J. (1986), Wage determinants: a survey and reinterpretation of human capital earnings functions, in O. Ashenfelter & R. Layard, eds, 'Handbook of Labor Economics', v. 1 edn, Vol. L, North Holland, Amsterdam, chapter 10, pp. 525–602.
- Wooldridge, J. W. (2010), *Econometric Analysis of Cross Section and Panel Data*, 2 edn, MIT Press, Massachusetts.
- World Health Organization (2009), *Global health risks: Mortality attributable to selected major risks*, Health Statistics and Informatics Department, Geneva.
- Zarkin, G. A., French, M. T., Mroz, T. & Bray, J. W. (1998), 'Alcohol use and wages: new results from the National Household Survey on Drug Abuse', *Journal of Health Economics* 17, 53–68.

Apêndice A



Fonte: Elaboração própria a partir dos microdados da PNAD/PETab 2008.

Figura A.1: Distribuição do percentual de fumantes por decil do salário-hora, 2008

Tabela A.1: Determinantes da participação na força de trabalho – modelos de probabilidade de escolha binária

Covariadas	Probit(1)	Probit(2)	SNP(1)	SNP(2)
Faixa de escolaridade				
<EF	-0,0489	0,0024	-0,2615	-0,0826
EF	-0,2072	-0,1566	-0,4811 ^b	-0,3404
EM	-0,0669	-0,0001	-0,1401	0,0805
Raça (Branca)	-0,0159	-0,0135	0,0105	0,0199
Idade	0,0388 ^a	0,0423 ^a	0,0352	0,0056
Idade ² /100	-0,0559 ^b	-0,0628 ^b	-0,0534	-0,0249
Interação: Educação e idade	0,0004	0,0005	0,0003	0,0007
Fatores Locacionais				
Urbana	-0,3317 ^c	-0,3084 ^c	-0,6626 ^c	-0,6761 ^c
Região metropolitana	-0,0192	-0,0944 ^a	-0,0596	-0,1707 ^a
Norte	0,0994	0,0954	0,2383	0,3631 ^b
Nordeste	-0,1746 ^c	-0,1559 ^b	-0,2833 ^b	-0,313 ^b
Sul	0,0853	0,1123	0,1803	0,1751
Centro-Oeste	0,0986	0,0713	0,2157	0,187
Características da família e do indivíduo				
Chefe do domicílio	0,5521 ^c	0,6177 ^c	1,2645 ^c	1,6525 ^c
Casado sem filhos	0,1803 ^b	0,2088 ^b	0,2345 ^a	0,2606
Casado com filhos	0,224 ^c	0,1807 ^b	0,4334 ^c	0,4418 ^c
Tamanho da família	-0,0106	-0,0024	-0,047	-0,0197
Dificuldades físicas	-0,4466 ^c	-0,5279 ^c	-0,7326 ^c	-0,9686 ^c
Renda familiar do não-trabalho				
Pensão	-0,0419 ^c	-0,037 ^c	-0,3514 ^c	-0,3766 ^c
Aluguel	-0,0015	0,0027	0,0022	0,0399
Outros	-0,0325 ^c	-0,03 ^c	-0,3317 ^c	-0,3652 ^c
Constante	0,7178 ^a	0,572	0,8571	0,6972
Coefficientes SNP: polinômio de ordem 3				
1			0,4669 ^c	0,5554 ^a
2			0,3561 ^a	0,8218 ^a
3			-0,0033	0,1094
Controle: Ex-fumante				
N	10.139	8.040	10.139	8.040
P (em milhões)	36,30	28,65	36,30	28,65
Count-R2	0,998	0,998	0,595	0,561

Elaboração própria a partir dos microdados da PNAD/PETab 2008.

Tabela A.2: Testes para endogeneidade da variável H_i no MQ2E

Teste	Modelo (5)		Modelo (6)	
	Estatística	P-valor	Estatística	P-valor
Endogeneidade ($H_0: H_i$ exógeno)				
Escore robusto (Durbin) $\chi^2_{(1)}$	18,742	0,000	18,148	0,000
Regressão robusta (Wu-Hausman) F	10,464	0,001	9,927	0,002
Restrição de sobreidentificação				
Escore $\chi^2_{(5)}$	274,940	0,000	192,732	0,000
Estatística J de Hansen: $\chi^2_{(5)}$	232,570	0,000	165,203	0,000
Resistência dos instrumentos				
Ajustamento do primeiro estágio				
R^2	0,2663		0,287	
R^2 ajustado	0,264		0,284	
R^2 parcial	0,231		0,209	
F robusto	182,894	0,000	111,081	0,000
R^2 parcial de Shea	0,2308		0,2088	
R^2 parcial ajustado de Shea	0,228		0,205	
Teste de identificação de instrumentos fracos				
Cragg-Donald (Estatística F de Wald)	439,076		305,635	
Kleibergen-Paap (Estatística F de Wald)	182,894		111,081	
Valores críticos do teste de Stock & Yogo (2005)				
Viés relativo máximo de 5% na VI	19,28		19,28	
Viés relativo máximo de 10% na VI	11,12		11,12	
Viés relativo máximo de 20% na VI	6,76		6,76	
Viés relativo máximo de 30% na VI	5,15		5,15	
Tamanho máximo de 10% na VI	29,18		29,18	
Tamanho máximo de 15% na VI	16,23		16,23	
Tamanho máximo de 20% na VI	11,72		11,72	
Tamanho máximo de 25% na VI	9,38		9,38	

Fonte: Elaboração própria a partir dos microdados da PNAD/PETab 2008.

Tabela A.3: Análise de Robustez – diferentes especificações de instrumentos

Detalhamento	MQO	Variáveis Instrumentais (VI)				
		1	2	3	4	5
Coefficiente (Fumante)	-0,1053	-0,2211	-0,2312	-0,3001	-0,3016	-0,2479
Erro-padrão	0,0212	0,0599	0,0598	0,0851	0,0850	0,0526
R2-ajustado	0,4918	0,4885	0,4879	0,4826	0,4824	0,4869
Resistência dos instrumentos						
Ajustamento do primeiro estágio						
R2-ajustado		0,2477	0,2476	0,1609	0,1609	0,2838
R2-parcial		0,1685	0,1685	0,0726	0,0727	0,2088
Estatística-F robusta		190,21	126,81	166,42	111,32	111,08
P-valor(F)		0,0000	0,0000	0,0000	0,0000	0,0000
Instrumentos						
Proibido fumar no domicílio				X	X	X
Crença/percepção						
Gera doenças aos outros				X	X	X
Gera doenças para si					X	X
Fumantes na família		X	X			X
Ex-fumantes na família		X	X			X
Preço do cigarro			X			X
Atributos individuais	X	X	X	X	X	X
Fatores locais	X	X	X	X	X	X
Controle: ex-fumantes	X	X	X	X	X	X
Número de observações	6.979	6.979	6.979	6.979	6.979	6.979

Fonte: Elaboração própria a partir dos microdados da PNAD/PETab 2008.

Tabela A.4: Resultados para a penalização salarial do fumante corrente para especificações com setor informal e ramos da atividade – MQO e MQ2E. Variável dependente: logaritmo do salário-hora

Covariadas	MQO		MQ2E	
	(3)	(3')	(6)	(6')
Fumante	-0,1053 ^c (0,0212)	-0,0944 ^c (0,0212)	-0,2479 ^c (0,0526)	-0,2304 ^c (0,0532)
Faixa de escolaridade				
<EF	-0,3778 ^c (0,0843)	-0,3396 ^c (0,0837)	-0,3648 ^c (0,0844)	-0,3294 ^c (0,0839)
EF	-0,3818 ^c (0,0682)	-0,3516 ^c (0,0679)	-0,3772 ^c (0,0686)	-0,3481 ^c (0,0683)
EM	-0,3300 ^c (0,0584)	-0,3099 ^c (0,0581)	-0,3334 ^c (0,0591)	-0,3129 ^c (0,0588)
Raça	0,1020 ^c (0,0201)	0,0984 ^c (0,0200)	0,0971 ^c (0,0201)	0,0942 ^c (0,0199)
Idade	0,0308 ^c (0,0096)	0,0296 ^c (0,0095)	0,0306 ^c (0,0095)	0,0295 ^c (0,0095)
Idade ² /100	-0,0401 ^c (0,0116)	-0,0390 ^c (0,0115)	-0,0382 ^c (0,0116)	-0,0372 ^c (0,0115)
Interação: Educação e Idade	0,0011 ^c (0,0001)	0,0011 ^c (0,0001)	0,0011 ^c (0,0001)	0,0011 ^c (0,0001)
Sindicalizado	0,1248 ^c (0,0238)	0,1011 ^c (0,0243)	0,1156 ^c (0,0238)	0,0953 ^c (0,0242)
Estabilidade (tenure)	0,0095 ^c (0,0014)	0,0092 ^c (0,0014)	0,0094 ^c (0,0014)	0,0090 ^c (0,0014)
Grupos ocupacionais				
Dirigentes	0,7921 ^c (0,0544)	0,8625 ^c (0,0565)	0,7881 ^c (0,0553)	0,8587 ^c (0,0572)
Ciências e Artes	0,7320 ^c (0,0809)	0,7533 ^c (0,0797)	0,7221 ^c (0,0819)	0,7430 ^c (0,0806)
Técnicos	0,5782 ^c (0,0440)	0,6133 ^c (0,0451)	0,5789 ^c (0,0442)	0,6135 ^c (0,0451)
Serviços administrativos	0,1672 ^c (0,0408)	0,2072 ^c (0,0412)	0,1618 ^c (0,0411)	0,2016 ^c (0,0414)
Vendedores	0,1679 ^c (0,0379)	0,2794 ^c (0,0445)	0,1603 ^c (0,0384)	0,2713 ^c (0,0445)
Agrícola	-0,2135 ^c (0,0396)	-0,0516 (0,0775)	-0,2058 ^c (0,0398)	-0,0447 (0,0789)
Produção e manutenção	0,1582 ^c (0,0243)	0,2195 ^c (0,0308)	0,1616 ^c (0,0246)	0,2195 ^c (0,0307)
Setor Informal		-0,0953 ^c (0,0264)		-0,0835 ^c (0,0264)
Ramos da atividade				
Agrícola		-0,1749 ^b (0,0759)		-0,1764 ^b (0,0773)
Indústria de transformação e outras		-0,0362 (0,0338)		-0,0419 (0,0341)
Construção civil		-0,1248 ^c (0,0351)		-0,1130 ^c (0,0349)
Comércio e reparação		-0,1593 ^c (0,0335)		-0,1587 ^c (0,0334)
Outras atividades		0,0287 (0,0367)		0,0250 (0,0367)

Fonte: Elaboração própria a partir dos microdados da PNAD/PETab 2008.

Legenda: ^ap<0,10; ^bp<0,05; ^cp<0,01.

Tabela A.4: Resultados para a penalização salarial do fumante corrente para especificações com setor informal e ramos da atividade – MQO e MQ2E. Variável dependente: logaritmo do salário-hora (continuação)

Fatores locacionais				
Urbana	0,0985 ^b (0,0389)	0,0923 ^b (0,0388)	0,1008 ^c (0,0387)	0,0945 ^b (0,0386)
Região metropolitana	-0,0236 (0,0187)	-0,0190 (0,0187)	-0,0228 (0,0187)	-0,0187 (0,0187)
Norte	-0,1569 ^c (0,0323)	-0,1467 ^c (0,0323)	-0,1550 ^c (0,0325)	-0,1460 ^c (0,0323)
Nordeste	-0,4062 ^c (0,0254)	-0,3973 ^c (0,0253)	-0,4096 ^c (0,0254)	-0,4016 ^c (0,0253)
Sul	-0,0053 (0,0273)	-0,0058 (0,0271)	-0,0035 (0,0272)	-0,0036 (0,0271)
Centro-Oeste	-0,0085 (0,0288)	-0,0013 (0,0289)	-0,0118 (0,0291)	-0,0051 (0,0291)
λ	-0,4258 ^c (0,0822)	-0,4138 ^c (0,0816)	-0,4041 ^c (0,0822)	-0,3954 ^c (0,0815)
Constante	2,1590 ^c (0,1840)	2,1907 ^c (0,1836)	2,1830 ^c (0,1847)	2,2138 ^c (0,1844)
Observações (N)	6.979	6.979	6.979	6.979
N expandido (em milhões)	24,56	24,56	24,56	24,56
R2 ajustado	0,49	0,50	0,49	0,49

Fonte: Elaboração própria a partir dos microdados da PNAD/PETab 2008.

Legenda: ^ap<0,10; ^bp<0,05; ^cp<0,01.

Tabela A.5: Definição das variáveis usadas na pesquisa por tipo de modelo

Variáveis	Definição das variáveis	Modelos		
		M	λ	VI
Salário-hora	Renda do trabalho principal na semana de referência por horas trabalhadas (em log)	X		
Participação na força de trabalho	1 se o indivíduo possui rendimentos positivos e não nulos, 0 caso contrário (c.c.)		X	
Sindicalizado	1 se o trabalhador é sindicalizado, 0 c.c.	X		
Dirigentes	1 se dirigentes em geral no trabalho principal, 0 .c.c	X		
Ciências e Artes	1 se profissionais das ciências e das artes, 0 c.c.	X		
Técnicos	1 se técnicos de nível médio, 0 c.c.	X		
Serviços administrativos	1 se trabalhadores de serviços administrativos, 0 c.c.	X		
Serviços*	1 se trabalhadores dos serviços, 0 .c.c	X		
Vendedores	1 se vendedores e prestadores de serviço do comércio, 0 c.c.	X		
Agrícola	1 se trabalhadores agrícolas, 0 c.c.	X		
Produção e manutenção	1 se trabalhadores da produção de bens e serviços e de reparação e manutenção, 0 c.c.	X		
Estabilidade (<i>tenure</i>)	Tempo de serviço em anos no emprego atual	X		
<EF	1 se menos que ensino fundamental completo, 0 .c.c	X	X	
EF	1 se ensino fundamental completo ou médio incompleto, 0 .c.c	X	X	
EM	1 se ensino médio completo ou superior incompleto, 0 .c.c	X	X	
ES*	1 se ensino superior completo ou mais anos de estudo, 0 .c.c	X	X	
Raça	1 se o trabalhador é da cor branca, 0 c.c.	X	X	
Idade	Idade em anos	X	X	
Fumante	1 se o trabalhador é fumante, 0 c.c.	X		
Tamanho da família	Componentes da família (excl. pensionista, empregado doméstico ou parente do empregado)		X	
Dificuldades físicas	1 se tem dificuldade para levantar objetos, praticar esportes ou realizar trabalhos pesados; 0 c.c.		X	
Aposentadoria	Renda com aposentadoria de instituto de previdência ou do governo federal, ou outro tipo (em log)		X	
Pensão	Renda familiar com pensão de instituto de previdência ou do governo federal, ou outro tipo (em log)		X	
Aluguel	Rendimento familiar com aluguel (em log)		X	
Outros	Renda familiar com juros, dividendos, programas sociais ou outros rendimentos, abono ou doação (em log)		X	
Chefe do domicílio	1 se o indivíduo é o morador de referência do domicílio, 0 c.c.		X	
Casado sem filhos	1 se a pessoa é casada mas sem filhos, 0 c.c.		X	
Casado com filhos	1 se a pessoa é casada mas com filhos, 0 c.c.		X	
Urbana	1 se o indivíduo reside em área urbana, 0 c.c.	X	X	

Fonte: Elaboração própria.

Coluna M representa as variáveis utilizadas para as equações de rendimento, com ênfase na obtenção da penalização salarial promovida pelo uso do cigarro.

Coluna λ indica as variáveis usadas para a correção do viés de seleção amostral.

Coluna VI reporta a descrição das variáveis instrumentais adotadas para os modelos MQ2E e RQVI em M.

Tabela A.5: Definição das variáveis usadas na pesquisa por tipo de modelo (continuação)

Variáveis	Definição das variáveis	Modelos		
		M	λ	VI
Região metropolitana	1 se o indivíduo reside em região metropolitana, 0 c.c.	X	X	
Norte	1 se o indivíduo reside na região Norte, 0 c.c.	X	X	
Nordeste	1 se o indivíduo reside na região Nordeste, 0 c.c.	X	X	
Sudeste	1 se o indivíduo reside na região Sudeste, 0 c.c.	X	X	
Sul	1 se o indivíduo reside na região Sul, 0 c.c.	X	X	
Centro-Oeste	1 se o indivíduo reside na região Centro-Oeste, 0 c.c.	X	X	
Proibido fumar no domicílio	1 se é proibido fumar em qualquer lugar do domicílio, 0 c.c.			X
Gera doenças aos outros	1 se o indivíduo acredita que fumar gera doença para terceiros, 0 c.c.			X
Gera doenças para si	1 se o indivíduo acredita que fumar gera doenças graves para si, 0 c.c.			X
Fumantes na família	Quantidade de pessoas fumantes que residem com o indivíduo (exclusive a próprio pessoa)			X
Ex-fumantes na família	1 se existe algum ex-fumante no domicílio (exclusive a próprio pessoa), 0 c.c.			X
Preço do cigarro	Média do preço do pacote (maço) do cigarro no estado			X

Fonte: Elaboração própria.

Coluna M representa as variáveis utilizadas para as equações de rendimento, com ênfase na obtenção da penalização salarial promovida pelo uso do cigarro.

Coluna λ indica as variáveis usadas para a correção do viés de seleção amostral.

Coluna VI reporta a descrição das variáveis instrumentais adotadas para os modelos MQ2E e RQVI em M.