

Anger, Silke; Kvasnicka, Michael

Working Paper

Biases in Estimates of the Smoking Wage Penalty

DIW Discussion Papers, No. 654

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Anger, Silke; Kvasnicka, Michael (2006) : Biases in Estimates of the Smoking Wage Penalty, DIW Discussion Papers, No. 654, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:

<https://hdl.handle.net/10419/18547>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Discussion Papers

DIW Berlin

German Institute
for Economic Research

654

**Silke Anger
Michael Kvasnicka**

Biases in Estimates of the Smoking Wage Penalty

Berlin, December 2006



Discussion Papers 654

Silke Anger ^a
Michael Kvasnicka ^b

Biases in Estimates of the Smoking Wage Penalty

Berlin, December 2006

^a DIW Berlin
^b Humboldt University Berlin

Opinions expressed in this paper are those of the author and do not necessarily reflect views of the institute.

IMPRESSUM

© DIW Berlin, 2006

DIW Berlin

German Institute for Economic Research

Königin-Luise-Str. 5

14195 Berlin

Tel. +49 (30) 897 89-0

Fax +49 (30) 897 89-200

<http://www.diw.de>

ISSN print edition 1433-0210

ISSN electronic edition 1619-4535

Available for free downloading from the DIW Berlin website.

Biases in Estimates of the Smoking Wage Penalty

Silke Anger ^{a,*}

^a *DIW Berlin*

German Institute for Economic Research
Koenigin-Luise-Str. 5, 14195 Berlin, Germany
E-mail: sanger@diw.de
Tel: +49-30-89789-526
Fax: +49-30-89789-109

Michael Kvasnicka ^b

^b *Humboldt University Berlin*

Department of Economics
Spandauer Str. 1, 10099 Berlin, Germany
E-mail: kvasnicka@wiwi.hu-berlin.de
Tel: +49-30-2093-5667
Fax: +49-30-2093-5696

* Corresponding author

Abstract

Empirical studies on the earnings effects of tobacco use have found significant wage penalties attached to smoking. We produce evidence that suggests that these estimates are significantly upward biased. The bias arises from a general failure in the literature to control for the past smoking behavior of individuals. 2SLS earnings estimates show that the smoking wage penalty is reduced by as much as a third, if past smoking of individuals is controlled for. Our results also point to significant wage gains for individuals that quit smoking, a finding that is of substantial interest, given the lack of evidence on the earnings effects of smoking cessation.

Keywords:

Smoking, wages, earnings regressions

JEL Classification:

J31, I19, C51

We would like to thank Bob Hart and Carsten Trenkler for helpful comments. All remaining errors are our own.

1. Introduction

Empirical studies on the earnings effects of tobacco use have found significant wage penalties attached to smoking, ranging from 2% to 24% (Levine, Gustafson, and Velenchik, 1997; Heineck and Schwarze, 2003; van Ours, 2004; Auld, 2005). Existing analyses, however, focus exclusively on the current smoking status of individuals.¹ This is likely to be a major shortcoming, as the comparison group includes not only individuals that have never smoked, but also former smokers. Unless the causal relationship between smoking and earnings is entirely contemporaneous, wage penalties calculated with reference to current smoking only will not reflect the true (overall) wage costs of smoking. For the magnitude of such estimates will also depend on the degree of contamination of the comparison group by past smokers, and on the strength of the effect of past smoking on current earnings.

For a number of reasons, past smoking is likely to matter for the current earnings of individuals. On the one hand, smoking may have persistent adverse effects on the earnings capacity of individuals. Not only may potential health damage caused by smoking be irreversible, but past discrimination by non-smokers might have led to irreparable career setbacks. In addition, lower productivity-enhancing human capital investments in the past due to higher rates of time preference of smokers may be difficult if not impossible to compensate later in life. On the other hand, however, smoking cessation may also have potential positive earnings effects (e.g. via motivation). Furthermore, former smokers may differ from individuals that have never smoked in unobserved individual characteristics, such as drive, or will power that render them intrinsically more productive. In sum, if past smoking has lasting adverse effects on the earnings

¹ An exception is the study by Lee (1999). However, it uses only very crude earnings information (the average earnings in the occupational category of an individual irrespective of her full- or part-time employment status).

of individuals, or if individuals that manage to quit exhibit characteristics that make them inherently more productive than either current smokers, no-time smokers, or both, then estimates of the wage effects of tobacco use that are based on data of current smoking status alone will be biased. In the former case, estimates will be downward biased, and in the latter case upward biased.

This paper provides evidence that existing estimates of the wage penalty attached to smoking are likely to be biased, a bias that is shown to arise from the failure to control for the past smoking status of individuals. It also produces first estimates of the wage effects of smoking cessation, which should be of substantial interest not only to smokers but also to policy makers.

2. Data and Summary Statistics

We use data from the German Socio-Economic Panel (SOEP), a representative longitudinal survey of individuals in Germany conducted annually since 1984 (see Haisken-DeNew and Frick, 2005, for a detailed description of this dataset). In 2002, extensive information on the current and past smoking behavior of individuals was sampled, including whether or not individuals had smoked regularly in the past, and at what age they had started to smoke. Unlike previous studies, we may hence distinguish between three groups of workers in our data: smokers, past smokers, and individuals that have never smoked.

Estimation Sample: We restrict the estimation sample to male workers in 2002 that are of German nationality, aged 27-55, work between 10 and 60 hours a week, earn a gross hourly wage of at least €4, and live in West Germany. The sex, age and hours restrictions are imposed to further comparability with existing studies on the wage effects of smoking. They serve to avoid entangling issues related to life-cycle labor supply such as child rearing, full-time education, and early retirement. The minimum hourly wage requirement in turn is set for plausibility reasons. In total, we have 1,968 individuals in our estimation sample.

Regression Variables: Our dependent variable is the log of gross hourly wages, which has been calculated from gross monthly earnings and actual weekly hours of work. The covariates included in all regression models are age and two sets of indicator variables for the respectively highest schooling and professional degree that individuals obtained. Instruments used in our 2SLS models are a dummy for early starting age of smoking (less than 16 years), which has been

used also in van Ours (2004), a dummy for co-residing with at least one no-time smoker, and a dummy for co-residing with at least one former smoker.²

Summary Statistics: Summary statistics on workers in the estimation sample are provided in Table 1. Three major insights may be gathered. First, past smokers do represent a significant share of current non-smokers. In fact, four out of ten individuals not smoking in 2002 had once regularly smoked (39.2%). Second, past smokers differ markedly both in their average age levels, and in their educational and professional qualifications from individuals that have never smoked: they are on average older than no-time smokers, and they have a lower incidence of higher education. Past smokers, therefore, do not only account for a sizeable fraction of current non-smokers but also differ in productivity-related characteristics from no-time smokers. Finally, smokers, former smokers, and individuals who have never smoked exhibit substantially different likelihoods of co-residing with either a no-time smoker or a past smoker. These differences will be exploited in our 2SLS estimations to instrument for the current and past smoking status of individuals. Note that average hourly wages of past smokers exceed not only those of smokers but also those of individuals that have never smoked.

- Table 1 about here –

² See Section 3 for a discussion of these instruments and alternative ones used in the literature.

3. Regression Results

Existing studies on the wage effects of tobacco have used a number of techniques to estimate potential wage penalties attached to smoking, including OLS, 2SLS, difference-in-differences, and hazard rates models. We use the first two types of analyses here, as these are the ones most commonly employed in the literature (cf. van Ours, 2004; Heineck and Schwarze, 2003; Levine, Gustafson, and Velenchik, 1997). Our regression models are kept deliberately parsimonious. We regress log hourly wages on indicator variables of current and past smoking status, controlling for the age, educational qualifications, and professional degrees of individuals.

Table 2 contains the regression output. Models 1 and 3 replicate existing studies in that both regression models only include an indicator variable for current smoking. Models 2 and 4, in contrast, control also for the past smoking status of individuals. In line with previous studies, a simple OLS regression (Model 1) reveals that current smokers experience a sizeable wage penalty relative to current non-smokers (4.5%), a wage discount that increases more than twofold to 9.9% if one controls for the endogeneity of current smoking by 2SLS (Model 3). When controlling also for past smoking (Model 2), however, the OLS estimate of the wage penalty of (current) smoking drops by as much as a third, to 3.1%. Moreover, and in contrast to current smoking, former smoking is associated with a wage premium of similar magnitude (3.5%) relative to no-time smokers. Instrumenting both current and past smoking status in Model 4 to account for potential endogeneity confirms these findings: a wage penalty (albeit now statistically insignificant) for current smoking, and a wage premium for past smoking. Past smoking therefore clearly matters for current earnings. As a consequence, wage penalties that are calculated with respect to current smoking status only, as in the existing literature (Models 1 and

3), tend to overestimate the true wage costs of smoking. Finally, Model 5 restricts the estimation sample to individuals that have ever smoked, i.e. to current and past regular smokers only. Rerunning our 2SLS earnings regression for this more homogenous group of individuals (all have at some time initiated smoking) confirms that quitting smoking is associated with a substantial wage gain for smokers.

- Table 2 about here -

Instruments used in 2SLS estimations: Existing 2SLS estimates of the wage penalty attached to smoking are often implausibly large, which may be attributed to the use of potentially inadequate instruments (see van Ours, 2004; Heineck and Schwarze, 2003). We experimented with alternative instruments used in the literature (being married, having children), but all of these were rejected by overidentification tests.³ Our novel co-residency instruments, in contrast, proved to be both strong (F-tests) and valid (overidentification test⁴), i.e. uncorrelated with the error term in the wage equations, and produced 2SLS estimates that are far more realistic in terms of their magnitude.

Robustness of results: Our results are robust to various changes of the estimation sample, including, among others, the expansion of the age cohort to older workers, the omission of the

³ Neither Heineck and Schwarze (2003) nor van Ours (2004) report overidentification tests for their preferred 2SLS specifications.

⁴ The tests of overidentifying restrictions have been carried out using the Stata ado file by Baum, Schaffer, and Stillman (2003).

minimum hourly wage restriction, or the increase of the lower threshold for weekly hours of work.⁵

⁵ Results of these regressions are available from the authors upon request.

4. Conclusion

This paper has shown that estimates of the wage penalty attached to tobacco use will be upward biased, if the past smoking status of individuals is not controlled for. Existing studies on the wage effects of smoking suffer from this defect, as they fail to differentiate between no-time smokers and past smokers among the group of individuals currently not smoking. We also find smoking cessation to be associated with sizeable wage gains for smokers. Further research, however, is required to uncover and quantify the underlying causes of these gains.

Literature

Auld, M.C., 2005, Smoking, drinking, and income, *Journal of Human Resources* 40(2), 505-518.

Baum C.F., M.E. Schaffer and S. Stillman, 2003, Instrumental variables and GMM: Estimation and testing, *Stata Journal* 3(1), 1-31.

Haisken-DeNew, J.P. and J.R. Frick, 2005, Desktop Companion to the German Socio-Economic Panel Study, German Institute for Economic Research (DIW), Berlin.

Heineck, G. and J. Schwarze (2003), Substance Use and Earnings: The Case of Smokers in Germany, IZA Discussion Paper No. 743.

Lee, Y., 1999, Wage effects of drinking and smoking: An analysis using Australian twins data, University of Western Australia Working Paper No. 22.

Levine P.B., T.A. Gustafson and A.D. Velenchik, 1997, More bad news for smokers? The effects of cigarette smoking on wages, *Industrial and Labor Relations Review* 50(3), 493-509.

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), 863-886.

Tables

Table 1: Summary statistics on workers by past and current smoking status

	Current Smokers:				Current Non-Smokers:			
			All non-smokers		No-time Smokers		Past Smokers	
	Mean (sd)		Mean (sd)		Mean (sd)		Mean (sd)	
<i>Dependent variable</i>								
Hourly gross wage (in €)	15.7	(5.9)	17.6	(6.6)	17.4	(6.8)	17.8	(6.2)
<i>Controls</i>								
Age (in years)	40.8	(7.3)	41.4	(7.4)	40.2	(7.3)	43.2	(7.2)
<i>Education (in %)</i>								
No school degree	0.02		0.01		0.00		0.01	
Primary	0.45		0.35		0.31		0.41	
Secondary	0.27		0.26		0.26		0.26	
Higher	0.22		0.36		0.40		0.30	
No Professional degree	0.11		0.06		0.05		0.06	
Apprenticeship	0.74		0.66		0.63		0.70	
University	0.15		0.29		0.32		0.24	
<i>Instruments</i>								
Co-residing with (in %)								
No-time smoker	0.33		0.55		0.61		0.46	
Past smoker	0.11		0.18		0.13		0.27	
Starting age less 16 (in %)	0.26		0.11		0.00		0.28	
N (Total = 1,968)	819		1,149		699		450	

Notes: Percentages of categories may not sum to one due to rounding differences.

Sample: SOEP 2002, West German male workers of German nationality, aged 27 to 55.

Table 2: Wage effects of current and past smoking

	OLS		IV		
	Model 1	Model 2	Model 3	Model 4	Model 5 (w/o no-time smokers)
Smoker	-0.045 (3.19) ^{***}	-0.031 (1.95) [*]	-0.099 (1.80) [*]	-0.064 (1.33)	-0.162 (2.22) [*]
Past Smoker		0.035 (1.90) [*]		0.116 (1.70) [*]	
Age	0.010 (10.51) ^{***}	0.010 (10.13) ^{***}	0.010 (10.06) ^{***}	0.009 (7.66) ^{***}	0.007 (4.94) ^{**}
<i>Education</i>					
Primary	0.077 (2.14) ^{**}	0.075 (2.10) ^{**}	0.071 (1.95) [*]	0.065 (1.77) [*]	0.057 (1.29)
Secondary	0.193 (5.24) ^{***}	0.192 (5.22) ^{***}	0.185 (4.88) ^{***}	0.181 (4.72) ^{***}	0.168 (3.7) ^{***}
Higher	0.288 (7.61) ^{***}	0.288 (7.61) ^{***}	0.278 (7.05) ^{***}	0.276 (6.95) ^{***}	0.253 (5.32) ^{***}
Apprenticeship	0.074 (2.89) ^{***}	0.075 (2.91) ^{***}	0.068 (2.59) ^{***}	0.069 (2.62) ^{***}	0.073 (2.41) [*]
University	0.266 (8.29) ^{***}	0.268 (8.35) ^{***}	0.254 (7.32) ^{***}	0.257 (7.49) ^{***}	0.252 (6.06) ^{***}
Constant	2.094 (38.06) ^{***}	2.091 (38.01) ^{***}	2.139 (30.28) ^{***}	2.137 (30.68) ^{***}	2.320 (20.19) ^{***}
R ² adjusted	0.26	0.26	0.25	0.24	0.21
N	1,968	1,968	1,968	1,968	1,269
<i>Instruments</i>			Early starting age Co-resident past smoker Co-resident no-time smoker	Early starting age Co-resident past smoker Co-resident no-time smoker	Early starting age Co-resident past smoker Co-resident no-time smoker

Notes: Absolute t-values in parentheses. F-tests of joint significance of the instruments are highly significant in all models. P-values of the Sargan's test statistic: 0.322 (Model 3), 0.968 (Model 4), 0.528 (Model 5).

Sample: SOEP 2002, West German male workers of German nationality, aged 27 to 55.