Schmollers Jahrbuch 131 (2011), 277 – 286
Duncker & Humblot, Berlin

Keeping it in the Family? If Parents Smoke Do Children Follow?

By Dean R. Lillard*

Abstract

I use retrospective data on smokers from the German Socio-Economic Panel to investigate whether children are more likely to smoke if their parents smoke(d). Despite intense policy interest, researchers have not established whether the well-established (positive) association is causal. I exploit panel data observations on smoking behavior of parents and children to develop instrumental variables that identify the causal relationship between parental smoking and youth initiation. The results suggest that children are not more or less likely to start smoking if their parents smoke. Failing to control for the endogenous choice of parents to smoke leads to incorrect inferences.

JEL Classification: I12

1. Introduction

Smoking is one of the leading preventable causes of death in every economically developed country in the world. Worldwide mortality from tobacco is likely to rise from about four million deaths a year in 1998 to about 10 million deaths a year by 2030 (World Health Organization, 1999). Others observe that: "Unless more is done to help the 200 million European adult smokers stop, the result will be 2 million European deaths a year by 2040" (European Partnership to Reduce Tobacco Dependence, 2001).

A commonly held belief – that children are more likely to smoke if their parents smoke – lurks in recent policy debates but its causal nature is not established. Plausible causal mechanisms that have been proposed but not tested include inheritance of a genetic predisposition to nicotine dependence (Batra et al., 2003), transmission of social norms that condone smoking (Hunter et al., 1982), and children who inadvertently become addicted to nicotine because

* I thank participants at the 9th International German Socio-Economic Panel User Conference and an anonymous referee for comments and feedback. This research was supported by Award # R01 AG030379 from the US National Institutes of Health.

they are exposed to secondhand smoke. A final mechanism, not directly studied in the literature, posits that youth effectively pay less for cigarettes if they can steal cigarettes from parents who smoke. Here I use instrumental variables (IV) to explore whether parental smoking causes youth to start smoking.

2. Background

Reviews of empirical studies on the relationship between parental smoking and adolescent smoking show no consistent relationship (Tyas and Pederson, 1998). One review concludes, "findings across (87) studies show weak and inconsistent associations between parent and adolescent smoking …" (Avenevoli/Merikangas, 2003). When estimated correlations are statistically significant, they are small. No evidence suggests that the correlation varies when both parents smoke rather than just one. A separate peer effects literature also finds mixed evidence. Some studies report large effects (Powell et al., 2005). Others find no or small effects (Krauth, 2007).

The much smaller economics literature on youth smoking initiation has ignored the role that might be played by parental smoking. Most published initiation models do not include measures of parental smoking because such measures are not present in the data they use (DeCicca et al., 2002), (Nicolas, 2002), (Cawley et al., 2004). A recent exception, (Göhlmann et al., 2010), is of particular interest because they use the same SOEP data I use. They find that youth are more likely to start to smoke if their mothers and fathers currently smoke and if their fathers used to smoke. However they treat parental smoking as exogenously assigned. As I show below, when one instruments for the parents' smoking status, no statistical association remains significant.

This study attempts to overcome several shortcomings in the broader empirical literature. These include the use of cross-sectional data or short panels; proxy reports for smoking behavior; small unrepresentative samples; and a failure to control for the endogenous parental smoking decision. I use retrospective data that spans more than fifty years. In the SOEP, parents and children answer the same survey question on smoking. The SOEP yields an analysis sample of 1,723 boys and 1,506 girls matched to parents (more than 17,000 person-life years). By the standards in the published literature, these are very large samples.[1] The household design of the SOEP and its following rules allows me to link smoking behavior of children and parents and develop the instruments I use to estimate the causal effect of parental smoking. One must account for the fact that parents chose (and choose) whether or not to smoke because their unobserved tastes, rate of time preference, or information sets

[1] The median sample is less than 1,000 and ranges from 39 to 16,996 (Avenevoli/Merikangas, 2003).

may also influence whether and how their children decide to smoke (and how children respond to cigarette prices).

3. Data

I use data on smoking behavior that is retrospectively reported in the 2002 survey of the German Socio-Economic Panel (SOEP). I briefly describe the retrospective smoking data and other control variables. An Appendix, available by request, provides more details.

– *Measures of smoking* The 2002 wave of the SOEP asks all respondents age 16 and older "How old were you when you started smoking regularly?" and "When did you give up smoking? (Please provide the year and, if possible, the month)." With these data I create an observation for each person in the 2002 SOEP for each year of life from birth to the age she was in 2002. For each year I code two indicator variables that identify whether a person had started to smoke or currently smokes. The smoking status variable equals "0" at ages a person doesn't smoke and "1" at ages she does. The start variable equals "0" at every age until a person starts to smoke. At that age it equals "1" (and is not defined at all subsequent ages). Although I define the start variable at all ages, in the models I restrict the sample to observations when people were ages 11 to 29 because the risk of starting is neglible up through age 10 and after age 29. I also create a time-invariant variable that indicates if a person has ever smoked at least 100 cigarettes (5 packs).[2] A person enters the sample if she responded to the 2002 survey and can be matched to at least one parent who also responded to the 2002 survey. I match children to parents with the SOEP variables that identify respondents' mothers and fathers who were ever surveyed.

– *Other measures* All models include controls for time-varying and time-invariant characteristics. The only time-varying data I include are age (own, father, mother), a quadratic time trend, an indicator for whether their father or mother has died, and an indicator for their state of residence.[3] I also include as controls time-invariant measures that indicate the quality of the relationship between children and their parents, the education of the father and mother, their religious affiliation, and their ethnic background. One of the key control variables is the price of cigarettes. Price is measured as the ave-

[2] Smoking prevalence rates calculated using retrospectively reported data match closely with contemporaneous measured rates, even up to twenty years in the past (Kenkel et al., 2003).

[3] In years before the person is observed, I assign them their state of residence in the first wave they are surveyed. For all practical purposes this variable is mostly time-invariant.

rage price of 20 cigarettes (in constant 2002 Euros) as reported in the German Statistical Office publication VI D 46/4-39. More details are in the Appendix.

Table 1 presents selected summary statistics. The top panel describes characteristics that vary over time. The bottom panel presents sample statistics for the time-invariant characteristics of the persons represented in the event history data. A key measure that supports the validity of the instrument is the age of fathers and mothers when sample members were born. In this sample, fathers were about 29 and mothers were about 26 when the average youth was born. This difference lies at the heart of the instrument.

4. Specification and Empirical Strategy

I modify a standard empirical model of smoking initiation to incorporate a role for parental smoking. This specification flows from a latent variable approach. Under the usual assumptions about the distribution of the error term, the probability that a youth starts is:

$$(1) \qquad Pr(Smoke_{ist}) = F(\beta_0 + \beta_1 S_{ist} + \beta_2 P_t + D_{ist}\beta_3 + \varepsilon_{ist} + \nu_s)$$

for individual i, living in state s, in age-year (or calendar year) t. Here S is a vector of smoking status of parents that includes the indicators described above, P is price, and the vector D_{ist} includes time varying and time invariant demographic characteristics. A youth is at risk to start smoking if he has not yet smoked. The dependent variable $Smoke_{ist} = 0$ until the year he starts. In the year he starts $Smoke_{ist} = 1$ and, in all subsequent years, he is dropped.

Control variables include indicators of the quality of the relationship between children and their parents, the education of the father and mother, their religious affiliation, and their ethnic background. The error structure includes υ_s to allow for a time-invariant but unobserved determinant of smoking common to individuals in the same state. In the US, υ_s may reflect state anti-smoking sentiment (DeCicca et al., 2008). Here it incorporates the idea that Germans might be more likely to smoke if they live nearer to countries with lower cigarette prices. Evidence suggests they do (Kvasnicka, 2010).

I estimate, by OLS and Generalized Method of Moments (GMM), three variants of this model that include smoking variables for mothers only, fathers only, and both parents. In all three variants I include an indicator for whether the parent(s) currently smoke and formerly smoked. In the OLS models I ignore the endogeneity of the relative's decisions to smoke. I then use IV to estimate the probability that a parent ever smokes and the probability that a parent currently smokes. F-statistics from the first-stage models appear in brackets under the coefficient estimates.

5. Identification Issues

Here, a valid instrument must predict the probability that a parent smokes, be orthogonal to the probability that a youth starts to smoke, and be uncorrelated with the error term of the initiation equation. These conditions are difficult to meet in the best of situations. The matter is further complicated by the fact that the decision to continue to smoke today is correlated with decisions in past periods about whether or not to quit. Consider the probability a person smokes in a given period (t):

$$(2)\quad Pr(Smk_t) = Pr(Start_{18}) \prod_{k=1}^{k=t+1} Pr(Smk_{18+t-k}) = Pr(Start_{18}) \prod_{k=1}^{k=t+1} 1 - Pr(Quit_{18+t-k})$$

where $Pr(Start_{18})$ represents the probability a person starts to smoke by age 18. (2) assumes away temporary quit behavior. Once a person starts, he continues to smoke until he permanently quits. (2) also highlights the challenge one faces in trying to find an instrument that predicts whether a person smokes at a given age and is independent of another person's decision to start to smoke that is taken in the same period. However, if quit decisions in adjacent periods are not too strongly correlated, then the older a parent is when she bears her first child, the more likely it is that one can use my identification strategy.

My two main instruments are the price of cigarettes and the log of a count of articles published in US magazines that warned about the health risks of smoking. For both, I average the values of the variables during the years the parent was age 15 – 19. I also use indicators for ten-year birth cohorts of the father and mother. These instruments are clearly valid. The cigarette price a parent faced at ages 15 – 19 obviously affects her decision to smoke but cannot affect her child's decision. On average, a parent decides whether or not to smoke two decades before her child decides (see Table 1). Further, cigarette prices and cigarette taxes vary over time. Figure 1 plots the average price of cigarettes in Germany from 1949 to 2008 and superimposes the distribution of the start ages of SOEP mothers and their sons. I have centered the distributions around the calendar year that each cohort was age 17, using the average year of birth in each cohort. The second main instrument proxies for the flow of information about the health risks of smoking. The anti-smoking US article count, which strongly predicts parental start decisions, measures a flow of information. It relies on the assumption that German journalists write and published articles that are based on articles published in the US. In other unpublished work (available on request), I show that in a year when more US articles are published, more articles are also published in Spanish and Portuguese magazines. As with price, the instrument is valid as long as a child's decision to smoke is uncorrelated with an information shock that happened twenty years before. The partial F statistics reported in Table 2 clearly indicate that the first-stage is precisely identified.

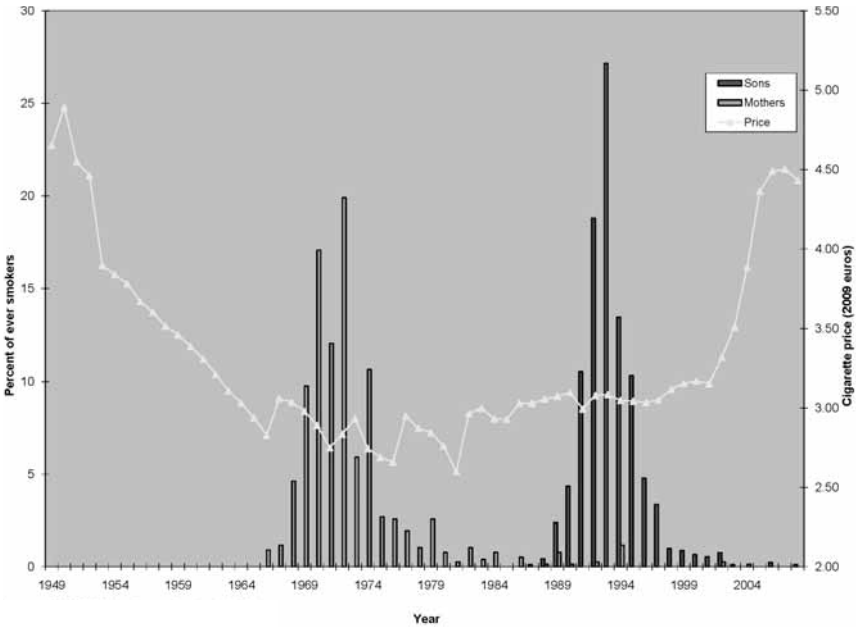
Table 1

Selected Sample Statistics

	Time-varying characteristics			
	Boys		Girls	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
Began regular smoking	0.045	(0.207)	0.037	(0.189)
Current smoker: father	0.319	(0.466)	0.321	(0.467)
Current smoker: mother	0.205	(0.404)	0.208	(0.406)
Price (2009 euros)	2.71	(0.23)	2.73	(0.21)
Age	16.74	(4.64)	16.85	(4.66)
Year	1993.0	(6.4)	1993.8	(5.90)
Years in sample (min/max)	1968	2002	1968	2002
N (person-year)	17359		15523	

	Time-varying characteristics			
	Boys		Girls	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
Ever smoked	0.392	(0.488)	0.323	(0.468)
Ever smoked father	0.651	(0.477)	0.658	(0.465)
Ever smoked mother	0.392	(0.488)	0.418	(0.493)
Cig. price when dad 15 – 18	2.30	(0.16)	2.29	(0.15)
Cig. price when mom 15 – 18	2.26	(0.12)	2.26	(0.10)
Age difference: father–child	28.83	(5.32)	29.06	(5.65)
Age difference: mother–child	25.90	(4.81)	26.11	(5.02)
Year of birth	1977.6	(6.1)	1978.5	(5.6)
N (persons)	1723		1506	

Source: 2002 SOEP. Cigarette prices calculated from German Statistical Office publication VI D 46/4-39.



Source: SOEP2002 & German Statistical Office, Statistical Yearbooks (various).

Figure 1: Average cigarette price and distribution of start ages of mothers and sons 1949 – 2008

6. Results

Table 2 reports coefficient estimates for the parental smoking variables and the cigarette price. In all models for both boys and girls, the probabiliy of starting is lower in years the cigarette price is higher. The effect is statistically significant at conventional levels.

Although my sample and specification differs from the models estimated in (Göhlmann et al., 2010), the naive associations are similar. For both boys and girls naive (OLS) results suggest that youth are more likely to start smoking if their mother or father currently smokes or if they used to smoke. Results in columns labeled (3) suggest that the current smoking behavior of mothers is correlated the current smoking behavior of fathers – the association between a youth's probability of starting and mother's current smoking status disappears when smoking behavior of both parents are included. In those models, a youth's start probability is strongly correlated with his father's current and former smoking status.

GMM-IV coefficient estimates are reported in the three columns on the right side of both tables. In general they are much larger than the OLS estimates (a

Table 2

Probability of Initiation

	Boys						Girls					
	OLS			GMM-IV			OLS			GMM-IV		
Variable	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Mother current smoker	0.014*		0.008	0.183**		0.067	0.020**		0.011	0.030		0.022
	(0.006)		(0.006)	(0.070)		(0.048)	(0.006)		(0.006)	(0.072)		(0.037)
First-stage F-stat				**[56.4]**		**[28.7]**				**[83.6]**		**[20.9]**
Mother ever-smoker	0.015**		0.012*	-0.079		0.023	0.020**		0.021**	0.023		0.051
	(0.004)		(0.005)	(0.069)		(0.057)	(0.005)		(0.006)	(0.078)		(0.044)
First-stage F-stat				**[41.9]**		**[12.0]**				**[50.8]**		**[10.2]**
Father current smoker		0.015**	0.012*		0.080	-0.002		0.013**	0.009*		0.062*	0.038
		(0.004)	(0.005)		(0.060)	(0.042)		(0.004)	(0.004)		(0.028)	(0.026)
First-stage F-stat					**[70.7]**	**[25.3]**					**[52.5]**	**[15.1]**
Father ever-smoker		0.014**	0.012*		-0.018	0.021		0.017**	0.011*		-0.063	-0.081*
		(0.004)	(0.005)		(0.081)	(0.038)		(0.004)	(0.005)		(0.034)	(0.034)
First-stage F-stat					**[34.2]**	**[29.5]**					**[25.2]**	**[16.6]**
Cig. price (2009 euros)	-0.070**	-0.061**	-0.069**	-0.059*	-0.062*	-0.064*	-0.104**	-0.096**	-0.098**	-0.105**	-0.101**	-0.107**
	(0.017)	(0.019)	(0.017)	(0.028)	(0.027)	(0.028)	(0.016)	(0.021)	(0.022)	(0.027)	(0.028)	(0.028)
N	18900	17359	16569	18900	17359	16569	17461	15523	15055	17461	15523	15055
R-squared	0.032	0.032	0.034				0.035	0.030	0.034			

Notes: Coefficients and robust standard errors (in parentheses). F-statistic on first-stage in brackets. * and ** denote coefficients that differ statistically from zero with p-values < 0.05 and < 0.01 respectively.

finding common in the literature). Based on these results we cannot conclude that a youth's decision to start smoking statistically depends on whether or not his mother or father currently or formerly smoked. Some evidence suggests that girls are less likely to start smoking if their father is a former smoker. Because the partial F-statistics on the set of variables used as instruments (in brackets and boldface type) verify that the instruments strongly predict the parental smoking measures, these results caution against attaching any causal interpretation to the OLS correlations.

7. Discussion and Conclusion

This study investigated the question, "Do parents who smoke induce their children to smoke?" These results suggest the answer is "No." When one treats parental smoking status as if it were exogenously assigned, it appears that youth are more likely to start smoking if their parents smoke. However, that result largely disappears when one accounts for systematic differences in the parents who start and quit smoking. This evidence means that anti-smoking interventions should not necessarily target the children of adult smokers. It appears that youth start to smoke independently of their parents' smoking habits.

References

Avenevoli, S. / *Merikangas,* K. R. (2003): Familial influences on adolescent smoking, Addiction 98 (Suppl. 1), 1 – 20.

Batra, V. / *Patkar,* A. A. / *Berrettini,* W. H., *Weinstein,* S. P. / *Leone,* F. T. (2003): The genetic determinants of smoking, Chest 123 (5), 1730 – 1739.

Cawley, J. / *Markowitz,* S. / *Tauras,* J. (2004): Lighting up and slimming down: The effects of body weight and cigarette prices on adolescent smoking initiation, Journal of Health Economics 23 (2), 293 – 311.

DeCicca, P. / *Kenkel,* D. S. / *Mathios,* A. D. (2002): Putting out the fires: Will higher taxes reduce the onset of youth smoking? Journal of Political Economy 110 (1), 144 – 169.

DeCicca, P. / *Kenkel,* D. S. / Mathios, A. D. / *Shin,* Y. / *Lim,* J. (2008): Youth smoking, cigarette prices, and anti-smoking sentiment. Journal of Health Economics 110 (1), 144 – 169.

European Partnership to Reduce Tobacco Dependence (2001): WHO evidence based recommendations on the treatment of tobacco dependence, World Health Organization, Geneva.

Göhlmann, S. / *Schmidt,* C. M. / *Tauchmann,* H. (2010): Smoking initiation in germany: The role of intergenerational transmission, Health Economics 19, 227 – 242.

Hunter, S. / *Baugh,* J. / *Webber,* L. / *Sklov,* M. / *Berenson,* G. (1982): Social learning effects on trial and adoption of cigarette smoking in children: the bogalus heart study, Preventative Medicine 11, 29 – 42.

Kenkel, D. S. / *Lillard,* D. R. / *Mathios,* A. (2003): Smoke or fire? are retrospective smoking data valid? Addiction 98 (9), 1307 – 1313.

Krauth, B. V. (2007). Peer and selection effects on youth smoking in california. Journal of Business and Economic Statistics 25 (3), 288 – 298.

Kvasnicka, M. (2010): Cigarette prices and tax evasion: Evidence from waste disposal contractors, Manuscript, RWI-Essen.

Nicolas, A. L. (2002): How important are tobacco prices in the propensity to start and quit smoking? an analysis of smoking histories from the spanish national health survey, Health Economics 11, 521 – 535.

Powell, L. / *Tauras,* J. / *Ross,* H. (2005): The importance of peer effects, cigarette prices, and tobacco control policies for youth smoking behavior, Journal of Health Economics 24 (5), 950 – 968.

Tyas, S. / *Pederson,* L. (1998): Psychosocial factors related to adolescent smoking: a critical review of the literature, Tobacco Control 7, 409 – 420.

World Health Organization (1999): The World Health Report 1999: Making a Difference, World Health Organization, Geneva.