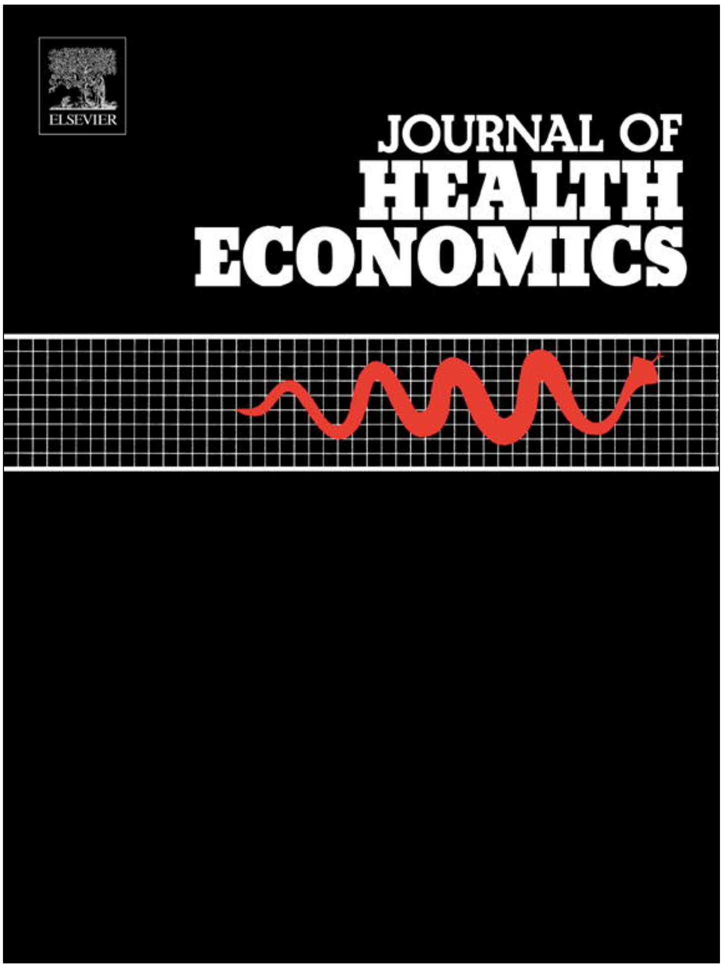
Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Journal of Health Economics 31 (2012) 584–598

Contents lists available at SciVerse ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

Does smoking affect schooling? Evidence from teenagers in rural China[☆]

Meng Zhao[a,*], Yoshifumi Konishi[b], Paul Glewwe[c]

[a] Waseda Institute for Advanced Study, Waseda University, Tokyo, Japan
[b] Faculty of Liberal Arts, Sophia University, Tokyo, Japan
[c] Department of Applied Economics, University of Minnesota, St. Paul, USA

ARTICLE INFO

Article history:
Received 16 November 2010
Received in revised form 9 April 2012
Accepted 25 April 2012
Available online xxx

JEL classification:
I18
I20
O15

Keywords:
Youth smoking
Educational achievement
Educational attainment
China

ABSTRACT

Youth smoking can biologically reduce learning productivity. It can also reduce youths' expected returns to education and lower their motivation to go to school, where smoking is forbidden. Using rich household survey data from rural China, this study investigates the effect of youth smoking on educational outcomes. Youth smoking is clearly an endogenous variable; to obtain consistent estimates of its impact, we use counts of registered alcohol vendors and a food price index as instrumental variables. Since the variable that measures smoking behavior is censored for non-smoking adolescents, we implement a two-step estimation strategy to account for the censored nature of this endogenous regressor. The estimates indicate that smoking one cigarette per day during adolescence can lower students' scores on mathematics tests by about 0.08 standard deviations. However, we find no significant effect of youth smoking on either Chinese test scores or total years of schooling.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

In many countries, consumption of addictive goods such as alcohol, marijuana, and tobacco is restricted or prohibited, particularly for adolescents. Parents often worry that addictive consumption at early ages may impair children's health and cognitive development, and may decrease their motivation to attend school via peer effects or prohibitions at their school, resulting in lower labor productivity and thus lower incomes throughout their lives. Over the last two decades, many economists have analyzed the causal effects of addictive consumption on educational outcomes (e.g. Cook and Moore, 1993; Bray et al., 2000; Register et al., 2001; Dee and Evans, 2003; McCaffrey et al., 2010). The present paper extends these efforts by investigating the effects of youth smoking in a developing country context.[1]

Unlike other abusive goods, such as alcohol and marijuana, the detrimental effects of smoking on learning abilities are less publicized. A large number of clinical studies, however, have clearly shown the negative impact of nicotine on the brain development and cognitive abilities of adolescent smokers, whose brains are particularly vulnerable to the neurotoxic effects of nicotine (Trauth et al., 2000; Jacobsen et al., 2005). The negative effect of smoking is more severe the earlier the age of the onset of smoking (Knott et al., 1999; Counotte et al., 2009). Adolescents who are daily smokers are found to have impairments in their working memory, and they perform poorer in various tests of cognitive abilities than their nonsmoking counterparts, irrespective of the recency of smoking. In addition, abstinence can have a much greater adverse impact on teens than on adults (Jacobsen et al., 2005). Moreover, youth smoking may also affect learning through its effects on health

☆ The paper is based on Essay 1 of the corresponding author's doctoral dissertation, which was in part supported by the Hueg-Harrison Fellowship at University of Minnesota. The authors gratefully acknowledge the helpful comments from seminar participants at the University of Minnesota, Waseda University, Williams College, and the 2011 American Economic Association meetings. We would also like to thank two anonymous referees and the editor for their constructive comments, which helped strengthen our paper substantially. All remaining errors are ours.
* Corresponding author. Tel.: +81 3 5286 2928; fax: +81 3 5286 2470.
E-mail addresses: zhaox057@umn.edu, konishi-moe@aoni.waseda.jp (M. Zhao).

[1] Approximately 80% of the world's smoking population lives in developing countries, with China alone accounting for more than 30%. Nonetheless, most of the existing literature studies youth smoking and substance use in developed countries. Teenagers in developing countries face rates of returns to education, working options, and social attitudes towards smoking that are substantially different from those in developed countries. The apparent shortage of such studies in developing countries is one of the main motivations for this study.

0167-6296/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jhealeco.2012.04.005

Author's personal copy

and nutrition. Cigarette smoking can cause serious health problems among children and teens, including coughing, respiratory illnesses, reduced physical fitness, poor lung growth and function, and worse overall health (USDHHS, 1994). Because smoking can interfere with the absorption of such vital nutrients as folate and vitamin B_{12} (Gabriel et al., 2006), it increases the risk of nutrition deficiency and anemia, which are known to lead to reduced learning (Glewwe et al., 2001).[2]

In addition to the biological effect, smoking may also reduce students' motivation to go to school and their study efforts. For example, in China, smoking is strictly prohibited in school, as required by law. However, because there is no law that specifies a legal minimum age for smoking *outside of school*, students have more freedom to smoke away from the school campus. Therefore, addicted teenage smokers may have a stronger incentive to skip classes or drop out of school than their non-smoking counterparts. Teen smokers may also reduce their investment in education based on their rational expectation of lower returns to education due to the adverse effect of smoking on their future health. Lastly, poor academic performance due to the biological effect can aggravate students' motivation to learn, via reduced interest in studying, reduced expected returns to education, and lower expectations from their parents regarding their future academic performance.

In contrast to the extensive clinical studies discussed above, little effort has been made to test whether the causal effects found in laboratories hold in observational data, and whether smoking indeed affects *educational outcomes rather than learning abilities* measured in a laboratory setting. On one hand, the negative effects of smoking may be worse in real life than in laboratories. Once teenagers start smoking, they may join a circle of peers who are less motivated to study, which may lead to a substantial reduction in their educational efforts. On the other hand, the negative effects of smoking on teenagers' learning abilities may not be large enough to reduce their school performance significantly. Moreover, human laboratory experiments are usually conducted with the subjects who volunteer to participate, and the smoking status of the subjects is often predetermined. Therefore, findings based on comparisons of the outcomes of smokers and non-smokers who volunteer for these studies are likely to suffer from bias due to self-selection of participants.

Health and education are two important forms of human capital, and both are endogenous. In recent years, a sizable economics literature has investigated the interrelationship between these two choice variables. On one hand, economists have long argued that healthy children learn more, and have used several different methods to empirically identify this causal relationship (e.g. Glewwe et al., 2001; Ding et al., 2009). On the other hand, others have investigated whether there is a causal relationship in the other direction, focusing on the impact of education on health outcomes. Such efforts are complicated by the existence of unobservable "third variables" such as preferences and abilities, which may influence both decisions simultaneously (e.g. Farrell and Fuchs, 1982). This endogenous interrelationship between health and education complicates our effort to identify the causal effect of youth smoking on educational outcomes.

This paper uses an instrumental variable (IV) approach to investigate the educational impacts of youth smoking, utilizing data from the Gansu Survey of Children and Families (GSCF). We explore the effects of youth smoking on two educational outcome variables: (i) "educational achievement", as measured by students' standardized test scores; and (ii) "educational attainment", as measured by total years of schooling. We exploit cross-sectional exogenous variation in the number of registered alcohol vendors and food prices to instrument the smoking decision. The GSCF data are less likely to suffer from bias due to omitted "third variables", because they contain rich information on various household and community characteristics, as well as school and teacher attributes, which were rarely available in previous studies. Furthermore, the GSCF data contain information on smoking intensity, as measured by the amount of cigarettes smoked per day over the previous month. Since we expect that regular smoking has more adverse effects on learning than experimental smoking, the information on smoking intensity should help to identify more accurately the impact of youth smoking on educational outcomes. Because smoking intensity is censored at zero, however, we need to correct for both the censoring and the endogeneity bias of the smoking intensity variable. For this, we employ a two-step IV estimator in the spirit of Heckman (1978) and Vella (1993): we first estimate a Tobit model of the smoking decision, and then estimate the second stage regression using the predicted smoking intensity.

The results provide support for a negative impact of youth smoking on educational achievement, particularly for the learning of mathematics. After accounting for endogeneity, smoking one additional cigarette per day for daily smokers aged 13–17 will lower their scores on the math exam by approximately 0.08 standard deviations. However, we find little effect of youth smoking on reading (Chinese) test scores. Moreover, we find no evidence of a causal effect of youth smoking on various measures of educational attainment, including total years of schooling, grade-for-age, and dropping out. Yet, consistent with previous studies (e.g. Farkas et al., 1999; Powell and Chaloupka, 2005), we do find strong empirical support for the effects of paternal smoking: children whose fathers smoke are significantly more likely to smoke, and to smoke more.[3]

To our knowledge, few studies have used observational data to investigate the causal effect of smoking on educational outcomes. However, a number of studies have used approaches similar to ours to investigate the effects of drinking and marijuana use on educational attainment. Cook and Moore (1993) used cross-state variation in the minimum legal drinking age (MLDA), while Dee and Evans (2003) exploited time variation in the MLDA as instrumental variables to control for the endogeneity of youth drinking. Bray et al. (2000) and Register et al. (2001) studied the impact of marijuana use on educational attainment in high schools in the U.S., using earlier use of marijuana and the residence in a decriminalized state at age 14, respectively, to instrument marijuana use. McCaffrey et al. (2010) used a two-step estimation approach to investigate the effects of marijuana use in grades 7–10 on dropout in grades 9 and 10.

In China, there is no law that specifies the minimum legal smoking age. Instead, we explore the exogenous variation in the supply of alcohol and the prices of food, both of which may influence

[2] Some smokers may believe smoking *enhances* learning, at least for a short period. Clinical studies appear inconclusive about this effect. Some studies have found that nicotine can reverse abstinence-induced declines in attention, memory and motor response to the levels before abstinence for nicotine-dependent individuals (Heishman et al., 1994). However, such enhancing effects usually happen within a short period immediate after smoking and the symptoms such as craving, anxiety, irritation, fatigue, headache, difficulty in concentration can occur as early as 30 minutes following smoking (Hendricks et al., 2006). Some previous studies have also observed short-term positive effects of nicotine on sustained attention and motor response for individuals who are not addicted to nicotine (Foulds et al., 1996). However, other studies have found null (Kleykamp et al., 2005) or negative effects (Poltavski and Petros, 2005) of nicotine among both nondeprived smokers and non-smokers.

[3] We cannot examine the impact of mothers' smoking because there are too few mothers who smoke in our sample. In general, the female smoking rate in rural China is extremely low, which is mainly due to cultural reasons.

Author's personal copy

the consumption of cigarettes. These aggregate-level factors are unlikely to be correlated with individual-level unobservables that affect both smoking and education decisions, especially after controlling for the grade fixed effects, school fixed effects and major regional characteristics, such as wage rates and school availability. The validity of our instrumental variables is also supported by various statistical tests.

The rest of the paper is organized as follows. Section 2 presents a conceptual framework for consumer's smoking and schooling decisions in the spirit of Becker and Murphy (1988). In Section 3, our identification and estimation strategies are discussed. Section 4 discusses the data and provides background information on youth smoking in China. Section 5 reports our results, and Section 6 concludes.

2. Empirical framework

In this section, we present a simple organizing framework for our empirical analysis. The framework presented here is an abbreviated version of the theoretical model developed in Appendix A.

Let S_t be consumption of cigarettes in period t. Past consumption of cigarettes can influence current and future consumption decisions through: (a) its effect on the marginal utility of smoking, and (b) its effect on current and future utility due to adverse health consequences or discomfort associated with addiction. The addictive stock in period $t+1$ depends on the amount of smoking and the addictive stock in period t:

$$A_{t+1} = f(s_t, A_t) \tag{1}$$

The more one smokes during the current period, or the more one has smoked in the past, the more addicted one is to tobacco in the next period: i.e. $f_S, f_A > 0$. Moreover, the addictive stock "depreciates" over time–the longer one abstains, the less addicted one is.

In addition, we extend the Becker–Murphy model to incorporate the consumer's educational decisions. The educational achievement (in knowledge and skills attained) at the beginning of period $t+1$, E_{t+1}, depends on the educational inputs in period t, e_t, and educational achievement at the beginning of period t, E_t:

$$E_{t+1} = \psi h(e_t, E_t) \tag{2}$$

where $\psi > 0$ is a parameter that describes productivity of educational inputs conditional on E_t, and h is an education production function with h_e, $h_E > 0$. The educational inputs include time and labor devoted to studying as well as material inputs. We emphasize here that, according to the finding of clinical studies, the learning productivity ψ is endogenous and indeed $d\psi/dA < 0$, but we assume that the consumer is unaware of this negative impact of smoking on learning. This assumption is plausible because the effects of smoking on cognitive abilities are seldom publicized, particularly in developing countries.

In Appendix A, we solve consumer's dynamic optimization problem subject to (1) and (2) and intertemporal budget constraints by transforming it to an equivalent dynamic programming problem. As a result, the optimal levels of current-period smoking and educational inputs are functions of the stock variables (A, E) up to period t and exogenous parameters of the model:

$$e_t^* = \varnothing_e(A_t, E_t; z_t, \omega_t) \tag{3a}$$

$$s_t^* = \varnothing_s(A_t, E_t; z_t, \omega_t) \tag{3b}$$

where z_t is a vector of current and future prices and ω_t is a vector of non-labor incomes and parameters describing preferences, health production, and educational production. The relationships derived in (3) make it clear that smoking and educational decisions are endogenous. First of all, since the optimal level of educational inputs e_t^* is a function of the addictive stock A_t (which depends on previous smoking decisions up to period t), it is clear that past smoking can affect current decisions on education. However, the reverse causality also exists. As shown in Eq. (3b), the optimal level of smoking in current period is affected by educational achievement E_t, which also depends on past educational decisions. Moreover, at least some factors in ω_t that can influence both s_t^* and e_t^* are likely to be unobserved. The objective of this study lies in identifying the effects of smoking on educational outcomes in the observational behavioral data, carefully addressing the reverse causality and the unobserved third factors.

The model enables us to obtain a clear theoretical prediction about the impact of a decrease in ψ due to smoking on educational outcomes e_t^* and E_{t+1}^*. According to the clinical studies, $d\psi/dA < 0$ and $dA/ds > 0$, which together imply $d\psi/ds < 0$. In the Appendix, we formally prove that if $d\psi/dA < 0$, an increase in smoking decreases both e_t^* and E_{t+1}^*, conditional on educational achievement E_t^*. The economic intuition behind this result is straightforward. Individuals with a higher level of addictive stock A_t are faced with lower returns to the educational inputs, which not only decreases educational achievement E_s for $s > t$ (for a given level of educational inputs) but also reduces their motivation to make further educational investments. Thus reduced learning ability due to smoking ($d\psi/ds < 0$) decreases both educational inputs and educational achievement.

A few caveats are in order. First, the effect of smoking on educational outcomes, $dE_{t+1}^*/ds_t^* < 0$, might arise either directly from reduced learning ability or indirectly from reduced demand for educational inputs, or both. Thus strictly speaking, the identified effect of smoking is a behavioral relationship, not the structural (clinical) relationship $d\psi/dA < 0$. Second, this model implicitly assumes that the individual makes decisions without information on $d\psi/dA < 0$. That is, the individual observes ψ, but is not aware of the effect of smoking on ψ. Once fully informed of this negative effect, the individual's demand for cigarettes would decrease because it would add to the (marginal) costs of smoking. Third, there may be another important pathway in which smoking affects educational outcomes.[4] Clinical studies also find that smoking may increase the chance of adverse health conditions such as lung cancer and respiratory diseases, which may result in a greater loss of work during working ages. Therefore, individuals with a higher level of addictive smoking stock may expect lower returns to education because they are faced with a higher risk of adverse health conditions, and thus may decrease educational investments. On the other hand, individuals with a higher level of educational achievement are faced with a higher cost of work loss, and thus may reduce smoking. Thus an analogous theoretical prediction could be derived even in the absence of the effect of smoking on learning ability. Unfortunately, neither our theoretical model nor empirical strategy can disentangle these multiple pathways.

In the empirical specification, educational achievement E_t^* is approximated by test scores in year t and educational attainment $\sum_{\tau=0}^{t} e_\tau^*$ by years of schooling up to year t. The next section will discuss the identification strategies to address the endogeneity problem discussed above.

3. Econometric model

This study attempts to identify empirically the causal effects of smoking on educational outcomes for teenagers, while taking

[4] We thank an anonymous referee for pointing this out.

Author's personal copy

into account the endogeneity of smoking choices. We focus on two types of educational outcomes; educational achievement and educational attainment.

3.1. Educational achievement (test scores)

To analyze the effect of smoking on educational achievement, we explore the cross-sectional variation in students' standardized scores on reading (Chinese) and mathematics tests. Standardized test scores are commonly used as measures of educational achievement in a given year. Since ψ is a function of s^*, we can rewrite Eq. (2) as $E^*_{t+1} = \varphi(s^*_t, e^*_t, E^*_t)$. Substituting (3a) and linearly approximating this equation, we obtain:

$$E_i = \boldsymbol{X}_i'\boldsymbol{\beta}_1 + \gamma_1 S_i + \varepsilon_{1i} \tag{4}$$

where S_i is the observed smoking behavior for the ith individual, and $\boldsymbol{X}_i$ denotes a vector of covariates (and a constant term) that can influence learning outcomes, such as academic inputs, learning efficiency, parental preferences, household economic resources, school characteristics, and the costs of schooling. Note that we do not include years of schooling in $\boldsymbol{X}_i$, because the test scores are grade-specific, standardized by the means and standard deviations of each grade level across schools, and, therefore, the identified effects in Eq. (4) are already implicitly conditional on years of schooling.

There are three empirical challenges to estimate Eq. (4). The first is the endogeneity of the smoking variable; S is likely to be correlated with ε_1 due to unobserved "third variables". For example, a "rebellious" child may take up smoking and skip class or drop out of school. Secondly, OLS estimates of Eq. (4) may also suffer from a downward bias because of measurement errors in the smoking behavior variable. Though there is no legal smoking age, parents and schools have legal obligation to prevent those under age of 18 from smoking. Therefore, teenagers are often restricted from smoking by their parents and schools and tend to under-report their smoking status. The reporting errors are likely to be more serious when parents or school authorities are present when the survey is administered.[5] Lastly, the smoking variable may suffer from a censoring problem. This study considers two smoking variables: (i) whether one has *ever* smoked; and (ii) the amount of cigarettes smoked per day in the most recent month. We anticipate that the latter offers more informative variation in smoking behavior, and thus it is our preferred variable. However, this variable equals zero for non-smokers and for light smokers who may have not smoked frequently enough to report smoking within the most recent month. All of these problems can lead to inconsistency of OLS estimates.

To address endogeneity and measurement error, we adopt an instrumental variable (IV) approach, using the number of registered alcohol vendors and a food price index at the community level as the exogenous instruments.[6] The instrumental variables may affect youth smoking through several mechanisms. First, teenagers' demand for cigarettes is mainly determined by their total budget, or pocket money. The supply of alcohol as well as food prices can affect household consumptions, resulting in a change in the household expenditures and thus a change in the amount of pocket money available for children.[7] Second, because drinking and smoking are related, either as substitutes or as complements, the supply of alcohol may be correlated with the demand for cigarettes negatively (as a substitute) or positively (as a complement). Third, although we do not have data on cigarette prices, it is possible that food prices reflect cigarette prices to a certain extent, therefore higher food prices imply higher cigarette prices and result in lower demand for cigarettes.

In order to qualify for a valid IV, the availability of alcohol and the food price index should not be correlated with the unobservables affecting educational achievement. The *aggregate-level* cross-sectional variation in the food price index and the alcohol supply are unlikely to be correlated with the *individual-level* or *household-level* unobservables. Of course, there remains some concern about the potential correlation between our IVs and the *community-level* unobservables that may affect educational achievement, such as unobservable school/teacher quality and some aspects of community environment. To address this concern, we control for the grade fixed effects, school fixed effects and some major community characteristics such as the availability of schools and the average wage rates in each village. Note that children from different communities may attend the same school, so that the instrumental variables can still predict the variation in youth smoking even within schools.

The validity of using the counts of the registered alcohol venders as an IV also relies on the assumption that alcohol consumption does not correlate with the error term in Eq. (4). This assumption will not hold if alcohol consumption has an additional effect on educational outcomes, i.e. through teen drinking and (or) parental drinking. Although some early studies show that *heavy* drinking during adolescence may reduce years of schooling (Cook and Moore, 1993; Koch and Ribar, 2001), more recent studies have criticized the methodological problems in this earlier literature and have provided evidence suggesting little causal effect of teen drinking on educational outcomes (Dee and Evans, 2003; Chatterji, 2006).[8] Given that youth drinking problem is much less severe in China compared to most developed countries,[9] teen drinking is unlikely to have a significant impact on educational outcomes. To examine parental drinking, we have done some preliminary checking and found little impact of the counts of alcohol venders on various factors that may mediate the effect of parental drinking on children's schooling: e.g. time spent by parents on helping children to study, children's mental health measured by the answers to a series of questions on child attitudes, and children's weight at birth as reported by mothers. To further examine the validity of the IVs, we rely on the statistical tests, as discussed in the result section.

To account for the large number of zero observations in the amount of cigarettes smoked per day variable (or the discrete nature of the ever-smoked variable) in conjunction with the IV strategy, we employ a two-step estimation strategy. For the ever-smoked variable, we first estimate a probit model against all the exogenous variables, including the excluded IVs. We then substitute the predicted smoking probability into the second-stage linear model for test scores. This two-step estimation provides consistent

[5] For example, the GSCF survey collected data on youth smoking behaviors in two ways. The first was by asking groups of teenagers to complete a questionnaire anonymously in a closed room without school officials or family members present, while the second way used a standard household survey questionnaire implemented at the teenager's home, in which anonymity is not guaranteed. These two different survey protocols generate considerably different rates of smoking among teenagers aged 13–17: about 12% using the former versus only 7% using the latter.

[6] Ideally, we would want to use the supply of cigarettes or cigarette prices as the instrumental variables. Unfortunately, reliable data are not available for these.

[7] We do not use the overall price index because it captures the prices of some educational inputs and can directly affect educational outcomes.

[8] To our knowledge, there are no rigorous studies on the educational effect of youth drinking in China. Most previous studies are based on data from developed countries.

[9] According to the WHO (2004), the prevalence of heavy episodic youth drinkers in China was 1.3% in 2001, compared to 10.7% in the United States, 15.3% in Canada, and 30% in the United Kingdom.

Author's personal copy

estimates and thus is recommended when the binary endogenous variable is determined by a continuous latent variable that crosses a threshold (Heckman, 1978). Yet, the two-step estimation is known to yield biased covariance estimates. Hence, we estimate the standard errors in the second stage via bootstrapping.

Because we have a large number of zero observations in the amount of cigarettes smoked per day in the most recent month, our preferred smoking variable, we use a Tobit specification in which the observed S is determined by the latent demand for cigarettes S^*:

$$S_i = \begin{cases} S_i^* & \text{if } S_i^* > 0 \\ 0 & \text{otherwise} \end{cases} \quad (5a)$$

$$S_i^* = \boldsymbol{Z}'_i \boldsymbol{\beta}_2 + \varepsilon_{2i} \quad (5b)$$

We assume that the error terms are normally distributed with zero means, variances $\sigma_{\varepsilon_1}, \sigma_{\varepsilon_2}$ and covariance $\sigma_{\varepsilon_1 \varepsilon_2}$. Since the variables in $\boldsymbol{X}$ may also affect youth smoking behaviors, the vector of $\boldsymbol{Z}$ includes all the explanatory variables in $\boldsymbol{X}$, in addition to the excluded instrumental variables that presumably affect only the smoking decision. Following Vella (1993), we first estimate a Tobit model in Eq. (5b) using all the instruments. The predicted amount of smoking is then used in the second stage linear model of test scores. As suggested by Vella (1993), we can also estimate the effect on test scores of latent smoking S^* as follows:

$$E(S_i^*|S_i) = I_i \boldsymbol{Z}'_i \tilde{\boldsymbol{\beta}}_2 + (1 - I_i)\{\boldsymbol{Z}'_i \tilde{\boldsymbol{\beta}}_2 - \tilde{\sigma}_{\varepsilon_2} \tilde{\phi}_i (1 - \tilde{\Phi}_i)^{-1}\} \quad (6)$$

where $\tilde{\boldsymbol{\beta}}_2$ and $\tilde{\sigma}_{\varepsilon_2}$ are the Tobit maximum likelihood estimates and I_i equals 1 if S_i is uncensored, and zero if otherwise. The P.D.F. and the C.D.F. of the standard normal distribution, $\tilde{\phi}_i$ and $\tilde{\Phi}_i$, are evaluated at $\boldsymbol{Z}'_i \tilde{\boldsymbol{\beta}}_2 / \tilde{\sigma}_{\varepsilon_2}$.

To control for heterogeneity in learning abilities and educational inputs, we include, as exogenous variables, parental education and smoking status, personal characteristics, and household income and land assets. Parental education and smoking may reflect the innate ability of children and parental preferences for children's education. Parents with higher education are also more likely to help their children with schoolwork. Parental smoking may expose children to secondhand smoking on a regular basis, which can have serious health effects on children, such as low birth weight, respiratory problems, and cognitive impairments. Household income is an indicator of resources allocated to children's education (e.g. richer parents can spend more on their children's schooling). It can also reflect some unobservable household characteristics, such as social status. Household land assets can serve as a measure of the household economic resources as well as an indicator of the household need for child labor, the latter can be an important factor for rural children in developing countries.

Educational achievement is also likely to be affected by unobservable school characteristics (e.g. unobservable quality of teachers and schools). Moreover, the exams are often different across schools. If the IVs are correlated with the unobservable school characteristics or the difficulty level of exams, the estimates of the smoking effects will be biased. Therefore, we generally include school fixed effects in the regressions. On the contrary, if the IVs are valid, there is no need to worry about such omitted variable bias. Moreover, since the correlation between any two students' test scores in the same school is likely to be different from that of two students from other schools, random effects estimates might be both asymptotically consistent and more efficient than fixed effects estimates. Hence, we also estimate the random effects model for comparison.

3.2. *Years of schooling*

Our model predicts that if youth smoking decreases the expected returns to education, it should also reduce the demand for education. Children (and parents) may be unaware of the detrimental cognitive effect of youth smoking (and hence, the effect on the education returns). However, they may still observe the signal from their lower school performance that they have the low returns to education.

To estimate Eq. (3a), we estimate a two-stage censored ordered probit model, with total years of schooling as the dependent variable. The censored ordered probit specification is used because: (1) observed year of schooling, which is a categorical variable, reflects a continuous latent demand for education; (2) our sample includes children who are currently enrolled in school, for whom the final years of schooling have yet to be observed. Thus their observed years of schooling are "right-censored" and provide only a lower bound of their final years of schooling. Failure to account for this censoring would yield parameter estimates that are both inconsistent and inefficient (see, for example, Vella, 1993; Glewwe and Jacoby, 1994; Zhao and Glewwe, 2010).

Let $y_i^* = \sum_{\tau=0}^{t} e_{\tau i}^*$ denote the latent continuous demand for educational inputs and let y_i be the observed years of schooling for ith child. Following Glewwe and Jacoby (1994), y_i^* and y_i are related to each other as follows:

$$y_i^* = \boldsymbol{X}'_i \boldsymbol{\beta}_3 + \gamma_3 S_i + \varepsilon_{3i} \quad (7a)$$

$$y_i = j, \quad \text{if} \quad \theta_{j-1} \le y_i^* < \theta_j \quad \text{for} \quad j = 1, \ldots, m \quad (7b)$$

where the elements of $\boldsymbol{\beta}_3$ are coefficients associated with all covariates in $\boldsymbol{X}$, and m is the highest level of y_i. Again, the smoking variable S is endogenous. As in Section 3.1, we use two alternative measures of smoking behavior. When the current amount of smoking per day is used, the observed variable S is related to the latent demand for smoking S^* as in the system (5).Assuming that ε_3 is *i.i.d.* and follows the standard normal distribution, the probability that $y_i = j$ is $Pr(y_i = j|\boldsymbol{X}_i) = \Phi(\theta_j - \boldsymbol{X}'_i \boldsymbol{\beta}_3 - \gamma_3 S_i) - \Phi(\theta_{j-1} - \boldsymbol{X}'_i \boldsymbol{\beta}_3 - \gamma_3 S_i)$ where Φ is the standard normal C.D.F. If person i is currently enrolled in year j, all we know is that her final years of schooling will be greater than or equal to j. Hence, the probability of observing j years of schooling should be $Pr(y_i = j|\boldsymbol{X}_i) = 1 - \Phi(\theta_{j-1} - \boldsymbol{X}'_i \boldsymbol{\beta}_3 - \gamma_3 S_i)$. Let $I_{ij} = 1$ if $y_i = j$ and $I_{ij} = 0$ otherwise. Furthermore, let $d_i = 1$ if y_i is censored and $d_i = 0$ otherwise. Then the log likelihood given the sample size N can be expressed as:

$$\ln L(\boldsymbol{\beta}_3, \gamma_3, \boldsymbol{\theta}) = \sum_{i=0}^{N} \sum_{j=0}^{m} I_{ij} \{\ln[\Phi(\theta_j - \boldsymbol{X}'_i \boldsymbol{\beta}_3 - \gamma_3 S_i)^{1-d_i} - \Phi(\theta_j - \boldsymbol{X}'_i \boldsymbol{\beta}_3 - \gamma_3 S_i)] \quad (8)$$

If S were exogenous and uncensored, maximizing the above log-likelihood function yields the consistent and efficient estimates of β_3, γ_3, and θ.

However, we have the same empirical challenges as discussed in Section 3.1. To address them, we adopt an IV approach based on the two-step estimation procedure employed in Rivers and Vuong (1988). We call the model a two-stage censored ordered probit model (2SCOP hereafter). As in Section 3.1, when we examine the effect of amount of cigarettes smoked per day, the procedure involves two steps: the first stage estimates a Tobit model and predicts the exogenous variation in smoking choice by instrumental variables, which is then substituted for S in the log-likelihood function (8); and we then estimate parameters using the standard

Author's personal copy

maximum-likelihood procedure. Again, the local availability of alcohol vendors and the food price index are used as IVs to correct for endogeneity and measurement error bias.

Previous studies on educational achievement often focus on the probability of dropping out of school. In fact, once we control for age and grade, an estimate of whether a child has dropped out or not would be getting at the similar underlying process that our censored ordered probit model does. However, estimating a dropout regression with a simple probit specification cannot examine *the timing of dropout* (or how many more years of education to acquire), even after controlling for age and grade in the right hand side of the probit equation. Therefore, the censored ordered probit uses more information than the simple probit does.

4. Data and background

The data used for this study are from the Gansu Survey of Families and Children (GSCF), which was conducted in the rural areas of Gansu, an underdeveloped northwestern province in China. In the year 2000, data were collected from a random sample of 2,000 children who were aged 9–13 years in that year. The sample was drawn from 20 counties that were randomly selected from all the major regions in Gansu. Within each of the counties, 100 children were randomly selected from the rural areas of those counties, yielding a sample of 1078 boys and 922 girls. Comprehensive data were collected through interviews of the sampled children, as well as interviews of their parents, teachers and school principals.

In 2004, the same children were interviewed again. Of the original 2000 children, 131 were not re-interviewed because of the following reasons: 108 children moved out of the county, 8 children died, 4 children were seriously ill, 2 children's parents were divorced, 1 household refused to be interviewed, and 8 for unknown reasons. Moreover, 24 observations were dropped due to the difficulty in matching data from the school survey and the household survey. Therefore, our study sample consists of 1845 teenagers aged 13–17 in 2004. Tests were not administered to the 204 sample children who had dropped out of school by 2004, which causes the sample size for the analysis of educational achievement to decrease to 1641. Although the GSCF was conducted in both 2000 and 2004, the first wave of the GSCF did not ask questions about youth smoking, so this study mainly uses the 2004 GSCF data, although the information on some baseline characteristics from 2000 GSCF are used.

One of the main educational outcomes of interest is educational achievement, as measured by scores on academic tests of math and Chinese skills, the two major subjects taught in primary and secondary schools in China. The GSCF collected comprehensive information on scores of tests administered by the school from the homeroom teacher of each sample child.[10] Homeroom teachers usually have accurate records of the previous test scores of the students in his or her homeroom class.

The test score variables used in our analysis are the averages of the scores on the mathematics and Chinese final exams in the last two semesters. In China, end of semester exams are usually given in the middle of January (end of fall semester) and the end of June (end of spring semester). As the GSCF surveys were conducted in the July of 2004, the test scores of the two most recent semesters are those from the exams given in January and June of 2004. There are two major reasons why we use the average scores of the last two semesters: (1) the majority of teen smokers started smoking well ahead of these exams and, therefore, their performance during these exams is likely to have been affected by their smoking; (2) averaged scores should reduce random errors in the test scores. Because the exams are usually different across grades, we standardize the test scores by the means and standard deviations of each grade level across schools to make the test scores comparable. Table 1 provides a comparison of the educational performance of smokers and non-smokers. Comparing the mean test scores at different percentiles for both math and Chinese scores, at lower percentiles, the mean standardized test scores of smokers are generally less than those of the non-smokers, but the differences are not statistically significant. However, as more students at higher percentiles are included in the comparisons, the test scores of smokers are significantly lower than those of non-smokers for both subjects.[11]

The other educational outcome investigated in this study is educational attainment, as measured by total years of schooling. The average years of formal schooling received by adults aged 15 or above in 2000 were about 7.1 in China, close to the average level in other developing countries in East Asia. The average years of schooling for the rural labor force in China were even lower, 6.1 years. The figure is much higher in the United States (12.7 years) and in other developed countries (10–12 years).[12]

The measure of years of schooling is defined as the current grade for those who were currently enrolled, because the survey was conducted at the end of the academic year. The highest grade attained is used for those who had dropped out of school by 2004. As discussed in Section 3, since the children in the sample are teenagers, we observe the total years of schooling only for those who have already left school. In our sample, 185 had left school by 2004; their average years of schooling were 6.7 years. These children's self-reported reasons for leaving school include unwillingness to attend school, financial difficulty, and academic difficulty. For those currently enrolled in school, the highest grades they will attain will be equal or greater than their grade attained in 2004. On average, the total years of schooling were 7.2 for those currently enrolled. Surprisingly, Table 1 shows that the average years of schooling of smokers were slightly higher than that of non-smokers. This may be attributed to the fact that older children are both more likely to smoke and more likely to have achieved greater years of schooling. However, we also find a lower dropout rate, although not quite statistically significant (p-value of 0.105). We thus suspect that these differences are also caused by measurement errors in the smoking variables for dropouts. Since the dropouts were interviewed at home, their smoking behaviors could be under-reported because of the presence of their parents during the interviews, as opposed to the interviews conducted at schools for the currently enrolled who completed the questionnaires without adults present.

Table 2 presents the descriptive statistics for the key variables used in the analysis. On average, 12% of the GSCF sample had smoked at least once. Among those who had smoked at least once, only 7 started to smoke after dropping out school. To avoid reverse causality, e.g. teenagers smoke due to lower educational attainment, these 7 observations are excluded from the analysis. About 25% of ever-smokers reported having smoked a positive amount

[10] In China, students are usually assigned to a homeroom class and stay in the same homeroom class until they graduate. A homeroom teacher is in charge of the administrative activities of a homeroom class, including keeping records of the students' profile, taking attendance, supervising students' overall performance, helping to solve students' problems, etc.

[11] We have also examined the test scores that are standardized by the means and standard deviations of each grade level in each school. The results are similar to those presented in Table 1.

[12] The statistics for the rural labor force in China are from de Brauw and Rozelle (2007). The other statistics are from the Barro and Lee Educational Attainment Dataset (Barro and Lee, 2010).

Author's personal copy

Table 1
Comparison of educational performance of smokers and non-smokers.

	All	Smokers		Non-smokers		$H_0: M_1 = M_2$; $H_A: M_1 < M_2$
	Mean	Obs.	$Mean_1$	Obs.	$Mean_2$	*P*-value
Math scores[a]						
Below 5th perc.	−2.43	11	−2.49	72	−2.42	0.33
Below 10th perc.	−2.03	26	−1.98	140	−2.04	0.73
Below 25th perc.	−1.36	57	−1.44	353	−1.35	0.18
Below 50th perc.	−0.78	111	−0.83	709	−0.77	0.22
Below 75th perc.	−0.38	155	−0.49	1077	−0.36	0.04**
Below 90th perc.	−0.16	181	−0.29	1298	−0.14	0.02**
Below 95th perc.	−0.08	192	−0.19	1367	−0.07	0.04**
For all	0.00	202	−0.11	1439	0.02	0.05**
Chinese scores[a]						
Below 5th perc.	−2.42	13	−2.42	70	−2.42	0.51
Below 10th perc.	−1.93	24	−1.95	141	−1.93	0.45
Below 25th perc.	−1.29	64	−1.29	347	−1.28	0.46
Below 50th perc.	−0.78	112	−0.86	706	−0.77	0.10*
Below 75th perc.	−0.39	155	−0.52	1077	−0.37	0.02**
Below 90th perc.	−0.18	185	−0.28	1288	−0.16	0.04**
Below 95th perc.	−0.09	198	−0.18	1359	−0.08	0.09*
For all	0.00	202	−0.13	1439	0.02	0.03**
Total years of schooling	7.15	222	7.77	1622	7.07	1.00
Dropout (1 = yes)	0.10	222	0.08	1623	0.10	0.10

[a] The test scores are standardized by grade-specific means and variances across schools.
* Significant at 10% level.
** Significant at 5% level.

Table 2
Descriptive statistics of key variables (2004).

	Obs.	Mean	S.D.	Min	Max
Standardized scores on Mathematics	1641	0.0	1.0	−4.5	2.4
Standardized scores on Chinese	1641	0.0	1.0	−5.1	2.4
Total years of schooling	1844	7.2	1.8	0	12
Ever smoked (1 = yes)	1845	0.12	0.33	0	1
If ever smoked					
Age started smoking	224	11.3	3.4	5	17
Currently smokes (1 = yes)	224	0.25	0.43	0	1
Cigarettes smoked per day last month[a]	224	3.5	3.1	0	30
Usually smokes at home (1 = yes)	224	0.28	0.27	0	1
Usually smokes at school (1 = yes)	224	0.31	0.27	0	1
Usually smokes at friends' places (1 = yes)	224	0.40	0.39	0	1
Usually smokes at social occasions (1 = yes)	224	0.17	0.17	0	1
Usually smokes at public (1 = yes)	224	0.20	0.21	0	1
Age	1845	14.6	1.2	13	17
Sex (1 = male)	1845	0.53	0.50	0	1
Father's years of schooling	1845	7.0	3.6	0	15
Mother's years of schooling	1845	4.3	3.5	0	13
Father smoking (1 = yes)	1845	0.77	0.42	0	1
Mother smoking (1 = yes)	1845	0.00	0.06	0	1
Household expenditures p.c. in 2000 (Yuan)	1845	1423	982	130	13,876
Log of household land assets (mu[b])	1839	2.0	0.8	−1.6	4.4
Distance from junior high school (km)	1845	3.7	4.2	0	30
Distance from senior high school (km)	1845	12.0	12.7	0.3	80
Average wage rate (Yuan)	1735	18.4	6.7	8	50
Counts of registered vendors of alcohol	716	21.6	30	0	99
Food price index	1788	113.4	2.7	108	118

[a] Calculated for only those who reported a positive amount of cigarettes smoked per day in the past 1 month.
[b] 1 mu = 667 m^2.

of cigarettes per day in the previous month. Of these, the average daily number of cigarettes smoked was 3.5. Approximately 40% of smoking teenagers reported that they smoked in their friends' houses, 31% smoked in school, 28% smoked at home, with about 20% smoking in public or at social events.[13] Note that, although smoking is forbidden in school, many students still secretly smoke in school at the risk of being caught and penalized by school authorities. The typical penalties for students who smoke include a verbal warning, a serious warning or a demerit record on the students' profiles. In more serious cases, the students may be placed on probation, asked to withdraw or expelled from the school. This suggests that many smoking teens may experience cravings for cigarettes that are too strong to resist, even during school hours.

[13] These percentages do not add up to 100% because multiple responses were permitted.

Author's personal copy

The proportion of fathers who smoked was 77%, 82% for teens who smoked and 76% for those who did not. The rate of smoking among mothers was very small compared to that of fathers. In fact, only 7 out of the 1845 mothers reported that they had ever smoked. This is consistent with the low prevalence of female smoking in many developing countries. The female smoking rate is slightly greater for teenagers, though: approximately 4.5% of ever-smokers were female, while the other 95.5% were male.

As household incomes are usually measured with substantial errors, we use household expenditures as a more reliable indicator of households' economic resources. However, there are still some concerns about endogeneity bias when using household expenditures as a regressor. For example, school dropouts may contribute to household expenditures. Since very few of the sample children had dropped out of school (and none reported that they were working) by the year 2000 (when they were 9–13 years old), we use household per capita expenditures in 2000 to measure household economic resources.

The counts of registered vendors of alcohol are calculated based on the registration records from the online database of China's Department of Commerce.[14] The data vary at the township level and are available for 13 townships in 8 counties, about only 40% of the GSCF sample.[15] The food price index is available at the county level and is available for 19 of 20 GSCF sample counties (Gansu Statistical Yearbook, 2008). It measures aggregate retail food prices in 2007 relative to the previous year.[16]

On average, there are 21.6 registered alcohol venders in the 13 townships for which data are available. In China, every alcohol vender is required to register at the local offices of the Department of Commerce. Failure to register can result in a fine up to a maximum of 2000 Yuan (approximately $310). Venders within the same province tend to carry similar types of alcohol and similar brands. The distribution of alcohol venders could vary widely. For example, there tend to be many alcohol venders if there is a local alcohol producer. Information on the food price index is available for 1788 children in 19 counties. The average level of the food price index is 113.4, indicating a general increasing trend in food prices.

5. Results

5.1. Determinants of youth smoking

The results of the first stage regressions are reported in Table 3. As discussed in Section 3, we report results for two measures of youth smoking: (i) whether one has ever smoked ("*ever-smoking*" henceforth); and (ii) the number of cigarettes smoked per day in the past month ("*current smoking intensity*"). The estimates of a probit regression for the first and a Tobit regression for the second are reported in columns (1) and (2), respectively. All the regressions control for all available covariates, distances to the closest lower and upper secondary schools, grade fixed effects, and school fixed effects. The robust standard errors are reported in parentheses.

The number of alcohol vendors and the food price index are negatively associated with both measures of youth smoking, and are significant at the 1% level. The negative correlation implies that the increase in the supply of alcohol and higher food prices may induce parents to spend more on alcohol and foods, and cut back on other things, such as pocket money for children, some of whom would spend it on cigarettes. It may also imply that greater alcohol supply induces less smoking, suggesting a substitute relationship between these two addictive goods. Lastly, the negative impact of food prices on youth smoking may also reflect the effect of greater cigarette prices. The estimated marginal effects are generally larger for the "*current smoking intensity*" regression than the "*ever-smoking*". For example, $(\partial E[S|S>0])/\partial vender = -0.006$ and $(\partial E[S=1])/\partial vender = -0.001$, which implies that participation in smoking is often experimental and is less responsive to teenagers' reduced budget or an increase in the cigarette prices.

Since our estimation hinges critically on the validity of our IVs, we conducted a likelihood ratio test for the explanatory power of our IVs, following Kan (2007). Under the null hypothesis that the IVs have no explanatory power to predict smoking, the test statistic follows a Chi-squared distribution with k degrees of freedom, where k is the number of excluded instruments, and follows an F distribution if divided by k. The calculated F-statistic should be close to or greater than 10 by the Staiger–Stock (1997) criterion. Since the Chi-squared statistic from the log likelihood ratio test is 23.77 for the "*ever-smoking*" regression and 17.13 for the "*current smoking intensity*", the F-statistics are 11.9 and 8.6, respectively, indicating that there is little reason to worry about the weak instruments problem.

We also find that paternal smoking has a significant impact on children's smoking behavior: teenagers whose fathers smoke are more likely to smoke, and smoke more per day if they smoke. If a father smokes, the probability that his child also smokes is 4% higher than those of non-smoking fathers. Moreover, his child smokes 0.23 more cigarettes per day than his counterparts. Unfortunately, since very few mothers reported that they smoked in the GSCF sample, we cannot estimate the effect of mothers' smoking on children's smoking choices. A possible explanation of the effect of fathers' smoking is that living in a household where a parent smokes makes it much easier for a teenager to obtain access to cigarettes. Moreover, children learn from their parents – observing their own parents smoke may make them underestimate the adverse health consequences of smoking. Our result is consistent with those from previous studies in a developed-country context: e.g. Powell and Chaloupka (2005) found a 6% higher probability of smoking for teenagers whose parents smoke and Farkas et al. (1999) reported a one-third decrease in youth smoking for children whose parents have quit smoking.

Interestingly, although not significant, we find that father's education is positively associated with youth smoking, while mothers' education has a negative coefficient in both regressions. These results are quite consistent in different specifications that are not reported in Table 3, which reflects the fact that mothers may have more say in children's education in China. In fact, according to the GSCF data, the probability for children to report that they have been informed of the harmfulness of smoking by parents is significantly higher if their mothers' education level is higher, which indicates that improving mothers' education may have a preventive effect on youth smoking.

Furthermore, per capita household expenditures have a negative and significant effect on youth smoking. The finding may appear contradictory to the intuition that teenage smoking is constrained by the budget for their pocket money. There are several possible reasons for this result. First, household expenditures

[14] Data are from http://jlscyxs.mofcom.gov.cn/vino/html/ (accessed on 02/05/2010).

[15] The online database of China's Department of Commerce is still under construction. Since some counties in Gansu have not yet joined the database, information on registered alcohol venders in those counties is missing. Note, however, that there are no systematic differences between the samples with and without data on alcohol vendors, which suggests little concern about sample selection bias due to missing information on alcohol vendors. To further confirm this, regression analyses are shown for both the full sample and the subsample with the alcohol vendor information. This issue is discussed in more detail when the results are discussed in Section 5.

[16] The county-level food price index was not available before 2007. However, since the trends in food prices in Gansu were relatively stable (Gansu Statistical Yearbook, 2007–2010), the food price index in 2007 is used as a proxy for that in 2004.

Author's personal copy

Table 3
First stage estimation of smoking choices.

	Ever-smoked, Probit[a]	Current smoking intensity, Tobit[a]
Instrumental variables		
Counts of registered vendors of alcohol	−0.014 (0.004)***	−0.056 (0.017)***
Food price index	−0.145 (0.028)***	−0.494 (0.148)***
Other explanatory variables		
Age	0.189 (0.082)**	0.585 (0.428)
Sex (1 = male)	1.814 (0.277)***	7.960 (1.287)***
Father smoking (1 = yes)[b]	0.475 (0.207)**	2.150 (0.953)**
Father's years of schooling	0.006 (0.024)	0.046 (0.114)
Mother's years of schooling	−0.011 (0.023)	−0.032 (0.116)
Log of household expenditures p.c. in 2000	−0.253 (0.151)*	−1.764 (0.709)**
Log of household land assets	−0.198 (0.164)	−0.248 (0.772)
Average wage rates (Yuan)	−0.015 (0.016)	−0.050 (0.077)
Dist. to the closest upper secondary school (km)	−0.008 (0.006)	−0.017 (0.029)
Dist. to the closest lower secondary school (km)	−0.054 (0.030)*	−0.058 (0.131)
Grade fixed effects[c]	Yes	Yes
School fixed effects[d]	Yes	Yes
Obs.	637	674
Log likelihood	−176	−375
Weak instruments test[e]	23.77	17.13
	[0.000]	[0.000]

[a] Robust standard errors are included in the parentheses and *p*-values in the square brackets.
[b] Mothers' smoking variable is automatically dropped because only four mothers reported smoking school.
[c] There are eight grade levels: no education, grades 4–6 in primary school and grades 1–4 in middle.
[d] Dummy variables for 32 schools are included in the regressions.
[e] Log likelihood ratio tests against the explanatory power of excluded IVs.
* Significant at 10% level.
** Significant at 5% level.
*** Significant at 1% level.

may be picking up the effects of other unobservable household or parental characteristics that are associated with household incomes, e.g. parents' preferences for child education and health, social status that may discourage teen smoking, or better information about the harmfulness of smoking. Second, household expenditures may be picking up the competing effect of other household consumptions on pocket money. For example, more food consumption may reduce the economic resources available for children. If that's the case, household expenditure will have a negative impact on youth smoking.

Lastly, age and sex are important predictors for both measures of smoking behavior. Boys are much more likely to smoke, and to smoke more. Among all ever-smokers, only 4.5% are girls. In general, the smoking rate increases with age, even after controlling for the grade fixed effects. Children who are older are significantly more likely to have ever smoked than younger children. The rate of smoking increases from 6.4% for youth aged 13 to 15.5% for youth aged 17.

5.2. Youth smoking and educational achievement

Table 4 presents estimates of the impact of youth smoking on educational achievement, as measured by test scores on math and Chinese (averaged over two semesters, using tests conducted in January and June of 2004).[17] The test scores are standardized by the means and standard deviations of each grade level *across schools*.[18] The top panel presents estimates of the effect of smoking on math scores, while the bottom panel provides estimates for Chinese scores. Seven regressions were estimated for each subject: columns 1–4 examine the effect on educational achievement of *ever-smoking* in an ordinary least squares specification. The first regression controls for the fixed effects of 46 schools in the full sample, and the second regression estimates the school random effects model. Regressions 3 and 4 are the same as the first two, except that they are based on the partial sample for which the information on the counts of alcohol venders is available. As discussed in Section 3, to correct for endogeneity and measurement error bias, we estimate the effect of smoking using a two-step estimation procedure, using the number of alcohol vendors and the food price index as instrumental variables. Because the information on alcohol vendors is missing for part of the sample, IV regressions can be estimated only for the subsample that has that information. For comparison, OLS regressions are shown for both the full and the partial sample. The IV estimates are reported in columns 5–7. The fifth and the sixth regressions use both the counts of alcohol venders and the food price index as the IVs, controlling for school fixed effects, with the latter controlling for the children's academic performance four years ago, as measured by the standardized test scores on math and Chinese reported by the home-room teachers in the first wave of the GSCF. The last column presents the IV estimates when only the food price index is used as the IV for the larger sample. The standard errors for all two-step IV estimations are obtained by non-overlapping block bootstrapping that draws the bootstrap sample at the township level for 300 times. All regressions include all the control variables used in Table 3 as well as grade fixed effects.

The OLS estimates of ever-smoking status are all negative but statistically insignificant. Results based on both full and partial samples indicate that the fixed effects estimates and random effects estimates differ by 0.006–0.016 for both subjects. As expected, the standard errors of the random effects estimates are smaller. However, the estimates are still statistically insignificant. In general, we prefer the fixed effects specifications because a Hausman test suggests that the fixed effects estimates are statistically different from

[17] Regressions that use only the January scores or only the June scores give similar, though slightly less precise, results.
[18] We have also estimated all the regressions using the test scores that are standardized by means and standard deviations of each grade level *within the same school*. The results are found to be very similar to those presented in Tables 4 and 5.

Author's personal copy

Table 4
Effects of smoking participation on educational achievement.

	OLS[a,b]			2-Step estimation[a,b]			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Math	−0.142	−0.126	−0.196	−0.190	−0.327	−0.322	0.259
	(0.089)	(0.082)	(0.137)	(0.121)	(0.367)	(0.336)	(0.537)
Grade fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes		Yes		Yes	Yes	Yes
School random effects		Yes		Yes			
Math scores in 2000						Yes	
Obs.	1535	1535	606	606	606	606	1478
R-squared	0.07	0.02	0.14	0.10	0.13	0.17	0.06
Overid. test[c]					9.89		
					[0.002]		
Chinese	−0.097	−0.087	−0.096	−0.086	−0.111	−0.255	0.087
	(0.092)	(0.082)	(0.144)	(0.125)	(0.433)	(0.473)	(0.591)
Grade fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes		Yes		Yes	Yes	Yes
School random effects		Yes		Yes			
Chinese scores in 2000						Yes	
Obs.	1535	1535	606	606	606	606	1478
R-squared	0.07	0.03	0.15	0.10	0.15	0.19	0.07
Overid. test[c]					17.15		
					[0.000]		

[a] Robust standard errors are included in the parentheses for the OLS estimations, bootstrapped standard errors that account for the clustering at the township level in the parenthesis for 2-step estimation, and *p*-values in the square brackets.
[b] All the regressions include all the explanatory variables other than the IVs in the first stage estimation.
[c] Overidentification tests are obtained by assuming the first stage estimation as linear.

random effects estimates. Comparing the estimates based on the full sample and the partial sample, the magnitude, sign, and statistical significance of the estimated effects of youth smoking do not differ significantly, even though the standard errors based on the partial sample get much larger. In fact, although not reported, the estimates of other variables in the regressions are also very similar regardless of the sample size. Hence, sample selection bias appears to be negligible.

The IV estimates are larger than the OLS estimates in absolute terms for both subjects when both IVs are used, but the estimates are not statistically significant. There is little difference in the estimates when the 2000 test scores are included.[19] When we use only the food price index to predict the probability of youth smoking, the IV estimates turn positive and are still statistically insignificant. The standard errors are even larger despite the much larger sample size. This may be due to a positive correlation between the food price index and the error terms in the ever-smoking specification. In fact, as shown in the fifth column, the IVs fail to pass the over-identification tests, suggesting that at least one of the IVs are probably invalid and one needs to be cautious in interpreting the IV estimates in Table 4.

These findings are in line with our expectations. As discussed in Section 4, most ever-smokers are experimental smokers – about 75% of them did not smoke in the last month before the interview. Experimental smokers do not smoke on a regular basis and thus are not addicted to cigarettes. Similarly, about 41% of the current-smokers do not smoke more than 1 cigarette per day. Some of these smokers may well be experimental smokers. We interpret the insignificance of the participation of smoking as suggesting that experimental smoking does not lead to regular smoking, that is to addiction to cigarettes. Therefore, it does not substantially affect either the amount of effort devoted to study or the cognitive learning ability.

Estimates of the effect of current smoking intensity on educational achievement are shown in Table 5. Again, the top panel is for mathematics and the bottom panel for Chinese. The results from eight regressions are reported. The OLS estimates are reported in the first four columns. The first two are based on the full sample, controlling for fixed effects and random effects, respectively. Regressions 3 and 4 are the same as the first two except the sample is limited to children only in the townships that have the alcohol vender information. The three regressions in columns 5–7 correspond to the IV regressions using both the counts of registered alcohol venders and the food prices as the IVs. The fifth regression uses the predicted observed smoking intensity variable, while the sixth uses the predicted latent smoking intensity variable. The seventh regression is the same as the fifth, except that the baseline test scores in 2000 are included for comparison. The last regression is also the same as the fifth, but uses only the food price index as the IV. The standard errors for all two-step IV estimations are obtained by bootstrapping, using 300 replications and controlling for the clustering at the county and the township levels.

All the OLS estimates of the impact of current smoking intensity are negative for both subjects, statistically significant at the 5% level for math and at the 1% level for Chinese. We do not observe any substantial difference between the fixed effects model and the random effects model. The efficiency gains from the random effects model are also very small, i.e. the difference in their standard errors ranges from 0.004 to 0.006. Therefore, we still prefer fixed effect specification. The magnitude and statistical significance of the estimates are also similar regardless of the sample size, indicating that sample selection bias would not be severe.

The IV estimate for math in column 5 is slightly smaller than the OLS estimates but more significant. According to the IV estimate in column 5, smoking one additional cigarette per day decreases the math test scores by approximately 0.076 standard deviations. Since the average teenage smoker smokes 3.5 cigarettes per day, this estimate translates into a large negative impact on math test scores:

[19] We have also estimated various OLS specifications that include the 2000 standardized test scores and generally find little difference from those presented in Table 4. We thus refrain from reporting them.

Author's personal copy

Table 5
Effects of smoking intensity on educational achievement.

	OLS[a,b]				2-Step estimation[a,b]			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Math	−0.081**	−0.074**	−0.081**	−0.074**	−0.076***	−0.074***	−0.081***	−0.058*
	(0.035)	(0.031)	(0.036)	(0.036)	(0.023)	(0.025)	(0.027)	(0.036)
Grade fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes		Yes		Yes	Yes	Yes	Yes
School random effects		Yes		Yes				
Math scores in 2000							Yes	
Obs.	1535	1535	641	641	641	641	641	1499
R-squared	0.07	0.03	0.16	0.11	0.17	0.18	0.20	0.08
Overid. test[c]					0.201		0.279	
					[0.654]		[0.598]	
Chinese	−0.103***	−0.101***	−0.097***	−0.099***	−0.018	−0.014	−0.032*	−0.043
	(0.037)	(0.031)	(0.037)	(0.037)	(0.038)	(0.028)	(0.024)	(0.044)
Grade fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes		Yes		Yes	Yes	Yes	Yes
School random effects		Yes		Yes				
Chinese scores in 2000							Yes	
Obs.	1535	1535	641	641	641	641	641	1499
R-squared	0.08	0.04	0.16	0.11	0.16	0.15	0.19	0.08
Overid. test[c]					1.832		0.713	
					[0.176]		[0.399]	

[a] Robust standard errors are included in the parentheses for the OLS estimations, bootstrapped standard errors that account for the clustering at the township level in the parenthesis for 2-step estimation, and p-values in the square brackets.
[b] All the regressions include all the explanatory variables other than the IVs in the first stage estimation.
[c] Overidentification tests are obtained by assuming the first stage estimation as linear.
* Significant at 10% level.
** Significant at 5% level.
*** Significant at 1% level.

approximately 0.267 standard deviations of the test score distribution. That the magnitude of the estimated coefficient decreases after controlling for the endogeneity bias implies that the OLS estimates suffer from an upward omitted variable bias: i.e. students who smoke may have some other characteristics that have negative effects on educational achievement.

In contrast to the math test scores, the estimated effect of smoking intensity on the Chinese test scores becomes much smaller and turns completely insignificant in the IV specification. Why does smoking affect the learning of math and Chinese differently? There are several possible reasons. For example, the learning of math and Chinese may require a different set of cognitive abilities, and nicotine affects them biologically differently. Another possibility is that learning of these two subjects may demand different amounts of effort and study time. In particular, Chinese is the students' native language. The learning of one's native language is usually influenced by many other factors that are less likely to be interfered by smoking, e.g. interest in reading Chinese novels. The IV estimates in column 6 indicate that the predicted latent smoking intensity has a similar effect on math test scores, but again no significant impact on Chinese test scores.

The IVs easily pass standard over-identification tests, failing to reject the hypothesis that the IVs are uncorrelated with the error term in the educational achievement specifications. Since it passes both the over-identification test and the weak instrument test, as discussed in Section 5.1, the fifth and the sixth regressions are our preferred specifications.

Our IV estimates are also robust to the inclusion of the students' baseline academic performance. When the test scores from four years ago are included in the seventh regression, the estimated effect of current smoking intensity increases slightly for math (still statistically significant at the 1% level). The estimate for Chinese even turns marginally significant, due to a greater magnitude and a smaller standard error. Because the sample children were mainly enrolled in grades 1–5 in primary schools four years ago, their test scores then may reflect their innate learning ability or other unobservable baseline characteristics. Our results then may be interpreted to suggest that smoking has a greater impact on learning when innate learning ability is controlled for, somewhat contrary to the conventional wisdom that students with poorer learning ability are more likely to self-select to smoke.[20] In fact, the estimated impact of the baseline test scores on smoking participation in the first stage is slightly positive, although it is not statistically significant.

In the last column, we have also conducted the IV estimation using only the food price index as the IV. The estimated effect turns slightly smaller for math and greater for Chinese, with a larger standard error. The decrease in precision is mainly due to a smaller explanatory power of the IV in the first stage estimation. Although we cannot test for correlation between the IV and the error term for the exactly identified model, the over-identification tests presented earlier imply that the estimates are not sensitive to the choice of the IVs, i.e. the cases when both IVs are used, or either of the two IVs is used.

5.3. Youth smoking and educational attainment

We now examine the effect of youth smoking on educational attainment, as measured by total years of schooling.[21] Table 6 presents the results of a censored ordered probit (COP) that estimates the impact of the participation in (top panel) and the

[20] There are 75 children who reported to have started smoking before 2000. Their academic performance in 2000 may have been already affected by smoking. To address this concern, we also estimated the regressions excluding these 75 children. The results are quite similar, both in magnitude and statistical significance.

[21] As explained in Section 4, the years of schooling used in the regression analysis are defined as the current grade or the highest grade attained before dropping out. One quarter of the sample children reported that they had repeated one or more grades, so our years of schooling variable undercounts years in school for these students. For robustness, we have also estimated all the regressions using total years of schooling that count repeated grades but found little difference in the results.

Author's personal copy

Table 6
Effects of youth smoking on educational attainment.

	COP			2SCOP		
	(1)	(2)	(3)	(4)	(5)	(6)
Ever-smoked[a,b]						
Smoking	0.368	0.362	0.383	−0.142	−0.101	−0.523
	(0.222)	(0.413)	(0.415)	(0.280)	(0.286)	(0.537)
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Test scores in 2000			Yes		Yes	
Obs.	1638	672	672	643	643	1600
Log likelihood	−477	−113	−119	−122	−120	−457
Overid. test[c]				1.60	0.94	
				[0.449]	[0.815]	
Current smoking intensity[a,b]						
Smoking	0.016	0.010	0.002	−0.032	−0.013	0.124
	(0.057)	(0.066)	(0.067)	(0.113)	(0.111)	(0.130)
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Test scores in 2000			Yes		Yes	
Obs.	1638	672	672	672	672	1600
Log likelihood	−478	−127	−127	−123	−122	−468
Overid. test[c]				0.32	0.81	
				[0.850]	[0.668]	

[a] Robust standard errors in the parentheses and p-values in the square brackets.
[b] Each regression has controlled for all the explanatory variables in the first stage estimation.
[c] Likelihood ratio test of the statistical significance of excluded IVs in the years of schooling equation.

intensity of (bottom panel) youth smoking, respectively, on years of schooling. For each measure of youth smoking, there are six regressions: columns 1–3 are the censored ordered probit estimates of the effect of the endogenous smoking variables without the IVs; regressions 4–6 are the estimates obtained from a two-step censored ordered probit that uses IVs to predict the exogenous variation in the smoking variables. Column 1 uses the full sample while column 2 uses the partial sample for comparison.[22] Regression 3 includes, as one of the explanatory variables, the math and Chinese test scores from four years ago using the partial sample. Column 4 presents the IV estimates using the count of alcohol vendors and the food price index as the excluded variables. Column 5 is the same as column 4 except that it uses test scores from four years ago as an explanatory variable in the second stage. The last column reports the IV estimates, but using the food price index as the only IV. All the regressions include school fixed effects and the same set of control variables as in the educational achievement specification.

The ordinary COP estimates of the effect of youth smoking, for both participation and intensity, are positive and insignificant. The positive coefficients reflect the surprising pattern observed in Table 1: smokers have more years of schooling than non-smokers. As discussed in Section 4, this likely reflects confounding factors, such as age (smokers tend to be older), and measurement errors in the smoking variables. Although the difference in Table 1 is statistically significant at the 1% level, the ordinary COP estimates are statistically insignificant because many confounding factors are controlled for in the regressions. As shown in columns 1 and 2, the estimates based on the full sample and the partial sample are similar, for both participation in and the intensity of smoking. Adding the baseline test scores in the regressions does not have much effect on both the magnitude and the statistical significance of the estimates.

The estimates from the 2SCOP in column 4 are also statistically insignificant. However, the coefficients turn negative, −0.142 for ever-smoking and −0.032 for currently smoking intensity, suggesting that the ordinary COP estimates suffer from some bias if they do not address endogeneity problem and measurement errors in the smoking variables. The estimates in column 5 (where the academic test scores from four years ago are included) are smaller, in absolute terms, but they remain statistically insignificant.

Lastly, the validity of the IVs used in columns 4 and 5 are supported by the over-identification tests, all of which indicate that the IVs are uncorrelated with the error terms in the educational attainment specifications. We have also estimated the 2SCOP using only the food price index as the IV for the full sample in column 6. The estimate of ever-smoking is much more negative but turns positive for current smoking intensity, yet both are statistically insignificant, which may be due to the lower explanatory power of the IV.

To confirm our findings from the 2SCOP, we have also estimated the effect of youth smoking on the grade-for-age variable, defined as the highest grade completed relative to the mean or median grade level for a child of that age. We have also estimated a probit model with a binary dependent variable indicating dropping out. The results from these attempts are consistent with the results in Table 6, although the standard errors are generally much larger. The larger standard errors are presumably because estimating a dropout regression with a simple probit specification or “grade for age” does not fully use all the information available, as discussed in Section 3.2.

There are several possible reasons for the insignificant results. First, the smoking variables may be subject to substantial sample selection bias in the years of schooling regressions, because a large number of dropouts could not be interviewed about their smoking behaviors and, when interviewed, the dropouts are likely to under-report smoking because anonymity was less likely. Since dropouts tend to under-report smoking behaviors, we may be observing a spurious “positive effect” – a large portion of the children who drop out are reported as non-smokers. With the two-step IV estimation, much of this spurious effect seems to disappear – the estimated coefficients on youth smoking generally turn negative. However, the spurious effect might not be completely removed. Second, we can observe the total years of schooling for only 10% of our sample and the rest 90% are right-censored, which implies a lack of precision in the left-hand side variables for 90% of the observations. Third, youth smoking may have adverse impacts on learning

[22] A comparison of the estimates based on the full sample and the partial sample confirms that there is little reason to worry about sample selection bias.

Author's personal copy

(i.e. its effect on test scores) but may have only minor impacts on years of schooling. Lastly, we lack dynamic data on smoking behavior over time. As discussed in the theoretical model, educational attainment, as measured by years of schooling, is the result of accumulated educational input decisions over time. Without detailed information on the exact timing and intensity of smoking over time, we may not be able to capture the real effect of youth smoking on the demand for education at each time period.

6. Conclusions

The detrimental effects of smoking on health have been both well documented and well publicized during the past several decades. Smoking is estimated to be responsible for 5.4 million global deaths annually (WHO, 2008). Over 80% of these deaths occur in developing countries. There are about one billion smokers in the world, of whom more than 80% live in developing countries and about 30% live in China. While adult smoking rates have slowly decreased in developed countries since the early 1990s, the rate of youth smoking has steadily increased in developing countries (Chaloupka et al., 2000).

This study has investigated the effects of youth smoking on educational outcomes. Using a rich dataset from China, this study has shown that youth smoking has adverse impacts on educational achievement. Smoking one cigarette per day at ages 13–17 is estimated to reduce standardized score on mathematics test by about 0.08 standard deviations. Interestingly, students' learning of Chinese is less affected by youth smoking. A possible reason for the smaller effect of smoking on learning Chinese may be that students generally need more time and effort to learn mathematics than to learn their native language. Moreover, the learning of Chinese and mathematics may also involve a different set of biological cognitive abilities, which may be affected by smoking differently.

Consistent with previous studies, our results also indicate substantial paternal effect on youth smoking. Fathers' smoking is one of the most important determinants of teenage smoking. This finding implies that a policy intervention targeted at parental smoking may be a cost-effective solution that "kills two birds with one stone" – it may improve the health and education of both parents and children.

Reduced learning per year during adolescence is an important addition to the real cost of smoking, in terms of productivity loss and possible lower life cycle welfare and income due to less educational achievement caused by youth smoking. Previous studies have considered the medical costs of smoking-caused diseases, financial costs of smoking-caused morbidity and mortality, property loss in smoking-caused fire, long-term special education care for low-birth-weight babies of smoking mothers, and expenditures on tobacco prevention and controls (Sloan et al., 2004). The present study argues that there is an additional cost to consider.

There are two caveats to the results of this study. First, the loss in learning could be underestimated since smoking may plausibly have additional adverse impacts on learning at the college level. In particular, smoking may not have a large impact on a decision to go to a college, but may affect the quality of colleges to which students who smoke are admitted. Second, since many children in our sample are still in school, we do not observe their total years of schooling. Though we use a censored ordered probit to control for this issue, the censored data can reduce the efficiency of our estimates. On the other hand, a sample consisting mainly of adults with completed years of schooling would suffer from substantial misreporting of smoking behaviors in their adolescence period. To address both of these concerns, future research may investigate the effect of youth smoking on high school graduates' college admissions.

Appendix A. The effects of smoking on educational outcomes

We model consumers' intertemporal smoking and educational decisions in the spirit of the rational addiction model of Becker and Murphy (1988) to translate the finding of recent clinical studies – that smoking negatively affects cognitive and learning abilities – into behavioral relationships that may be identified and estimated in observational data.

A consumer's preferences in each period are defined over a numeraire consumption good, x, and smoking, s. Following Becker and Murphy, it is assumed that the addictive good s contributes to an addictive stock, A, that also enters the consumer's utility. The one-period utility is thus given by $u(x_t, s_t, A_t)$.

Past consumption of cigarettes can influence current and future consumption decisions through: (a) its effect on the marginal utility of consuming s, and (b) its effect on current and future utility due to adverse health consequences or discomfort associated with addiction. More specifically, we assume $u_{SA} > 0$, which implies the marginal utility of smoking is higher if A is high, and $u_A < 0$, which means the marginal utility of addiction is negative. Our assumptions on transitional relationships, $A_{t+1} = f(s_t, A_t)$ and $E_{t+1} = \psi h(e_t, E_t)$, are described in Section 2.

It is assumed that the consumer is endowed with a constant amount of time in each time period, which is allocated between going to school and working. That is, if e_t increases, the time allocated to working will decrease and, therefore, income falls in that period. We thus assume that income I_t in each period decreases with educational inputs e_t and increases with educational achievement E_t. As in Becker and Murphy (1988) and Becker et al. (1994), the consumer lives infinitely and any effects of s or A on the consumer's length of life or other types of uncertainty are ignored.

The consumer chooses an optimal consumption path $\{x_t, s_t, e_t\}_{t=0}^{\infty}$, maximizing the discounted sum of utilities:

$$\sum_{t=0}^{\infty} \delta^t u(x_t, s_t, A_t) \tag{A1a}$$

subject to (1) and (2), and the intertemporal budget constraint:

$$x_t + p_t s_t + w_t e_t + (1+r)B_{t-1} \le I(e_t, E_t) + W_t + B_t \tag{A1b}$$

where δ is the consumer's time preference, p_t is the price of cigarettes, w_t is the price of educational inputs, r is the interest rate, and B_t is intertemporal borrowing. For simplicity, assume that r is constant and $\delta = 1/(1+r)$, as in Becker and Murphy. In earlier periods (e.g. teenage years), the consumer may obtain positive non-labor income $W_t > 0$, which is assumed to be exogenous. This budget balance condition is consistent with the idea that some families pay W_t to cover educational costs, living expenses, and basic leisure expenditures until children mature and attain sufficient skills to earn adequate incomes. Yet other poor families do not pay for these costs, and therefore their children may start working at an early age, before acquiring a high level of education.

Given certain regularity conditions,[23] the maximization problem (A1) can be reformulated as a recursive dynamic programming problem (Stokey et al., 1989):

$$v(A, E, B) = \max_{A', E', B'} [u(x, s, A) + \delta v(A', E', B')] \tag{A2}$$

[23] These conditions include (a) u is concave in x and s for every feasible A, (b) f and h are bounded, real-valued functions of s and e, respectively, for every feasible A and E, and (c) $\lim_{t\to\infty} \sum_{t=0}^{\infty} \delta^t u(x_t, s_t, A_t)$ exists for every feasible sequence of $\{x_t, s_t, e_t\}_{t=0}^{\infty}$. Condition (c) holds if u, f, h, and I are bounded and non-empty valued.

Author's personal copy

where primes indicate variables' values in the next period. Substituting the constraints, we can rewrite (A2) in terms of current period decision variables:

$$v(A, E, B) = \max_{x,s,e}\{u(x, s, A) + \delta v[f(s, A), \psi h(e, E), x + ps + we + (1 + r)B - I(e, E) - W]\} \tag{A3}$$

The first-order conditions are:

$$\varphi_x \equiv u_x + \delta v_B = 0 \tag{A4a}$$

$$\varphi_s \equiv u_s - pu_x + \delta v_A f_s = 0 \tag{A4b}$$

$$\varphi_e \equiv (\delta v_E/u_x)\psi h_e - w + I_e = 0 \tag{A4c}$$

Eq. (A4a) is the standard condition that the marginal utility of other consumption in each period equals the marginal utility (or shadow value) of money. Eq. (A4b) implies that the optimal cigarette consumption equates the marginal utility of cigarette consumption with the current price of cigarettes (multiplied by the shadow value of money) plus the discounted marginal effect on future utility from increased addiction. Similarly, Eq. (A4c) implies that the optimal educational input in each period equates the discounted marginal gain in future income streams from education with the costs of education.

The current period optimal decisions are thus functions of state variables (A, E, B) and exogenous parameters of the model:

$$e_t^* = \varnothing_e(A_t, E_t, B_t; \boldsymbol{p}_t, \boldsymbol{w}_t, \boldsymbol{W}_t, \psi, \delta, u, f, h, I) \tag{A5a}$$

$$s_t^* = \varnothing_s(A_t, E_t, B_t; \boldsymbol{p}_t, \boldsymbol{w}_t, \boldsymbol{W}_t, \psi, \delta, u, f, h, I) \tag{A5b}$$

where $\boldsymbol{p}_t$ and $\boldsymbol{w}_t$ are vectors of current and future prices and W_t is a vector of current and future non-labor incomes. Eq. (A5) is the formal version of the relationships in Eq. (3) discussed in Section 2.

Furthermore, the following proposition shows that if $d\psi/dA < 0$, an increase in smoking decreases both e_t^* and E_{t+1}^* conditional on educational achievement E_t^* up to period t.

Educational effects of smoking through reduced learning productivity: Suppose that the value function v of the recursive dynamic programming version of the model (A1) exists, is twice-differentiable, and is concave in endogenous arguments. Then conditional on educational achievement up to period t, E_t^*, the demand for both educational inputs e_t^* and educational achievement E_{t+1}^* decreases with a decrease in ψ. Because smoking decreases ψ, an increase in smoking has negative effects on both educational inputs and educational achievement.

Proof. Implicitly differentiate the system of Eq. (A4) with respect to ψ and e_t^*. By the implicit function theorem, we have:

$$\frac{de_t^*}{d\psi} = -\frac{1}{\Delta}\begin{vmatrix} \varphi_{xx} & \varphi_{xs} & \varphi_{x\psi} \\ \varphi_{sx} & \varphi_{ss} & \varphi_{s\psi} \\ \varphi_{ex} & \varphi_{es} & \varphi_{e\psi} \end{vmatrix}$$

where Δ is the determinant of the Hessian of the objective function (A3) and is ≤ 0 since the objective function is concave in endogenous arguments.

$$\begin{vmatrix} \varphi_{xx} & \varphi_{xs} & \varphi_{x\psi} \\ \varphi_{sx} & \varphi_{ss} & \varphi_{s\psi} \\ \varphi_{ex} & \varphi_{es} & \varphi_{e\psi} \end{vmatrix} = -\frac{\delta v_E}{u_x} h_e[u_{xs}^2 - u_{xx}u_{ss} - \delta v_A u_{xx} f_{ss}] \ge 0$$

By concavity of the utility function, $u_{xx}u_{ss} - u_{xs}^2 \ge 0$. For the production function of addictive stock, $f_{ss} \ge 0$ as a person gets more addicted to smoking when the consumption of cigarettes is higher. Because $(\delta v_E/u_x)h_e$ is the marginal benefit of educational inputs which is positive, the term in the brackets is non-positive. Thus we have $de_t^*/d\psi \ge 0$. Furthermore, educational achievement E_{t+1}^* also increases with ψ conditional on E_t^*:

$$\left.\frac{dE_{t+1}^*}{d\psi}\right|_{E_t} = h(e_t^*, E_t) + h_e\frac{de_t^*}{d\psi} \ge 0$$

□

References

Barro, R., Lee, J.W., 2010. A new data set of educational attainment in the world, 1950–2010. NBER Working Paper 15902.

Becker, G., Grossman, M., Murphy, K., 1994. An empirical analysis of cigarette addiction. American Economic Review 84 (3), 396–418.

Becker, G., Murphy, K., 1988. A theory of rational addiction. The Journal of Political Economy 96 (4), 675–700.

Bray, J., Zarkin, G., Ringwalt, C., Qi, J., 2000. The relationship between marijuana initiation and dropping out of high school. Health Economics 9, 9–18.

Chaloupka, F.J., Hu, T., Warner, K.E., Jacobs, R., Yurekli, A., 2000. Tobacco Control in Developing Countries. Oxford University Press, Oxford.

Chatterji, P., 2006. Does alcohol use during high school affect educational attainment? Evidence from the National Education Longitudinal Study. Economics of Education Review 25, 482–497.

Cook, P.J., Moore, M.J., 1993. Drinking and schooling. Journal of Health Economics 12, 411–429.

Counotte, D., Spijker, S., Van de Burgwal, L.H., Hogenboom, F., Schoffelmeer, A.N., De Vries, T.J., et al., 2009. Long-lasting cognitive deficits resulting from adolescent nicotine exposure in rats. Neuropsychopharmacology 34 (2), 299–306.

de Brauw, A., Rozelle, S., 2007. Returns to education in rural China. In: Hannum, E., Park, A. (Eds.), Education and Reform in China. Routledge, London.

Dee, T.S., Evans, W.N., 2003. Teen drinking and educational attainment: evidence from two-sample instrumental variables estimates. Journal of Labor Economics 21 (1), 178–210.

Ding, W., Lehrer, S., Rosenquist, J.N., Audrain-McGovern, J., 2009. The impact of poor health on academic performance: new evidence using genetic markers. Journal of Health Economics 28, 578–597.

Farkas, A.J., Distefan, J.M., Chi, W.S., Peirce, J.P., 1999. Does parental smoking cessation discourage adolescent smoking? Preventive Medicine 28 (3), 213–218.

Farrell, P., Fuchs, V.R., 1982. Schooling and health: the cigarette connection. Journal of Health Economics 1, 217–230.

Foulds, J., Stapleton, J., Swettenham, J., Bell, N., McSorley, K., Russell, M., 1996. Cognitive performance effects of subcutaneous nicotine in smokers and neversmokers. Psychopharmacology 127, 31–38.

Gabriel, H.E., Crott, J.W., Ghandour, H., Dallal, G.E., et al., 2006. Chronic cigarette smoking is associated with diminished folate status, altered folate form distribution, and increased genetic damage in the buccal mucosa of healthy adults. American Journal of Clinical Nutrition 83 (4), 835–841.

Gansu Statistical Yearbooks, 2008–2011. China Statistics Press, Beijing.

Glewwe, P., Jacoby, H., 1994. Student achievement and schooling choice in low income countries: evidence from Ghana. Journal of Human Resources 29 (3), 843–864.

Glewwe, P., Jacoby, H., King, E., 2001. Early childhood nutrition and academic achievement: a longitudinal analysis. Journal of Public Economics 81, 245–368.

Heckman, J., 1978. Dummy endogenous variables in a simultaneous equations system. Econometrica 46 (4), 931–959.

Heishman, S.J., Taylor, R.C., Henningfield, J.E., 1994. Nicotine and smoking: a review of effects on human performance. Experimental and Clinical Psychopharmacology 2 (4), 345–395.

Hendricks, P.S., Ditre, J.W., Drobes, D.J., Brandon, T.H., 2006. The early time course of smoking withdrawal effects. Psychopharmacology 187, 385–396.

Jacobsen, L.K., Krystal, J.H., Mencl, W.E., Westerveld, M., Frost, S.J., Pugh, K.R., 2005. Effects of smoking and smoking abstinence on cognition in adolescent tobacco smokers. Biological Psychiatry 57 (1), 56–66.

Kan, K., 2007. Cigarette smoking and self-control. Journal of Health Economics 26, 61–81.

Kleykamp, B.A., Jennings, J.M., Blank, M.D., Eissenberg, T., 2005. The effects of nicotine on attention and working memory in neversmokers. Psychology of Addictive Behaviors 19, 433–438.

Knott, V.J., Harr, A., Mahoney, C., 1999. Smoking history and aging-associated cognitive decline: an event-related brain potential study. Neuropsychobiology 40, 95–106.

Koch, S.F., Ribar, D.C., 2001. A siblings analysis of the effects of alcohol consumption onset on educational attainment. Contemporary Economic Policy 19 (2), 162–174.

McCaffrey, D.F., Pacula, R.L., Han, B., Ellickson, P., 2010. Marijuana use and high school dropout: the influence of unobservables. Health Economics 19 (11), 1281–1299.

Poltavski, D.V., Petros, T., 2005. Effects of transdermal nicotine on prose memory and attention in smokers and nonsmokers. Physiology and Behavior 83, 833–843.

Powell, L.M., Chaloupka, F.J., 2005. Parents, public policy, and youth smoking. Journal of Policy Analysis and Management 24 (1), 93–112.

Author's personal copy

Register, C.A., Williams, D.R., Grimes, P.W., 2001. Adolescent drug use and educational attainment. Education Economics 9, 1–8.
Rivers, D., Vuong, Q.H., 1988. Limited information estimators and exogeneity tests for simultaneous probit models. Journal of Econometrics 39, 347–366.
Sloan, F., Ostermann, J., Picone, G., Conover, C., Taylor, D., 2004. The Price of Smoking. The MIT Press, Cambridge, MA.
Staiger, D., Stock, J.H., 1997. Instrumental variables regression with weak instruments. Econometrica 65 (5), 557–586.
Stokey, N., Lucas, R., Prescott, E., 1989. Recursive Methods in Economic Dynamics. Harvard University Press, Cambridge, MA.
Trauth, J.A., McCook, E.C., Seidler, F.J., Slotkin, T.A., 2000. Modeling adolescent nicotine exposure: effects on cholinergic systems in rat brain regions. Brain Research 873, 18–25.
U.S. Department of Health and Human Services, 1994. Preventing Tobacco Use Among Young People: A Report of the Surgeon General. Atlanta, Georgia.
Vella, F., 1993. A simple estimator for simultaneous models with censored endogenous regressors. International Economic Review 34 (2), 441–457.
World Health Organization, 2004. WHO Global Status Report on Alcohol 2004. WHO, Geneva.
World Health Organization, 2008. WHO Report on the Global Tobacco Epidemic 2008: The Mpower Package. WHO, Geneva.
Zhao, M., Glewwe, P., 2010. What determines basic school attainment in developing countries? Evidence from rural China. Economics of Education Review 29 (3), 451–460.