WIAS Discussion Paper No.2010-005

Does Smoking Make One Dumber?
Evidence from Teenagers in Rural China

December 3, 2010

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

[a] *Waseda Institute for Advanced Study, Waseda University*
[b] *Faculty of Liberal Arts, Sophia University*
[c] *Department of Applied Economics, University of Minnesota*

1-6-1 Nishiwaseda, Shinjuku-ku, Tokyo 169-8050, Japan
Tel: +81-3-5286-2460 ; Fax: +81-3-5286-2470

Does Smoking Make One Dumber? 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 55108, USA*

ABSTRACT

Youth smoking can biologically reduce learning productivity. It can also reduce youths' 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, conditional on years of schooling, smoking one cigarette per day during adolescence can lower students' scores on mathematics tests by about 0.1 standard deviations. However, we find no significant effect of youth smoking on either Chinese test scores or total years of schooling. This study also provides strong empirical support for "parental effects" – parental smoking has significant impacts on the probability and intensity of youth smoking.

JEL codes: I18, I20, O15

Key words: youth smoking, educational achievement, educational attainment, China

---

* Corresponding author. Tel.: +81 3 5286 2928; fax: +81 3 5286 2470.
E-mail address: konishi-moe@aoni.waseda.jp.§ 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, and Williams College. All remaining errors are ours.

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*., 2009). 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 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

[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 return 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.

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. 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 the 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

---

[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.

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. We explore the effects of youth smoking on two educational outcome variables: (1) "educational achievement", as measured by students' standardized test scores; and (2) "educational attainment", as measured by total years of schooling. We exploit cross-sectional exogenous variation in 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 censoring and 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 0.1 standard deviations. On the other hand, 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 either total years of schooling or dropping out. Yet we do find strong empirical support for the (differential) effects of parental smoking. Children whose fathers smoke are significantly more likely to smoke, and to smoke more.

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*. (2009) 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 price of food, both of which may influence 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 dynamic model of 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. Theoretical model

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.

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 this 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:

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

where $\psi > 0$ is a parameter that describes productivity of educational inputs conditional on E_t, and h is an education production function with $h_e, h_E > 0$. 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.

The educational input e_t includes time and labor devoted to studying as well as material inputs. 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 input e_t and increases with educational achievement E_t. As in Becker and Murphy (1988) and Becker, Grossman, and Murphy (1994), the consumer lives infinitely and any effects of s or A on the consumers' length of life or other types of uncertainty are ignored.

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

$$\sum_{t=0}^{\infty} \delta^t u(x_t, s_t, A_t) \quad (3a)$$

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

$$x_t + p_t s_t + w_t e_t + (1+r)B_{t-1} \le I(e_t, E_t) + W_t + B_t, \quad (3b)$$

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 (assumed constant, as in Becker and Murphy) and B_t is intertemporal borrowing. For simplicity, assume that $\delta=1/(1+r)$. 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[3], the maximization problem (3) can be reformulated as a recursive dynamic programming problem (Stokey, Lucas, and Prescott, 1989):

$$v(A,E,B) = max_{A',E',B'}[u(x,s,A) + \delta\, v(A',E',B')] \tag{4}$$

where primes indicate variables' values in the next period. Substituting the constraints, we can rewrite (4) in terms of current period decision variables:

$$\begin{aligned} v(A,E,B) = max_{x,s,e}\{u(x,s,A) + \delta v[f(s,A), \psi h(e,E), x + ps + we \\ +(1+r)B - I(e,E) - W]\} \end{aligned} \tag{5}$$

The first-order conditions are:

$$\varphi_x \equiv u_x + \delta v_B = 0 \tag{6a}$$
$$\varphi_s \equiv u_s - pu_x + \delta v_A f_s = 0 \tag{6b}$$
$$\varphi_e \equiv (\delta v_E/u_x)\psi h_e - w + I_e = 0. \tag{6c}$$

Equation (6a) is the standard condition that the marginal utility of other consumption in each period equals the marginal utility (or shadow value) of money. Equation (6b) 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.

[3] 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.

Similarly, equation (6c) 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:

$$s_t^* = \emptyset_s(A_t, E_t, B_t; p_t, w_t, W_t, \psi, \delta, u, f, h, I) \tag{7a}$$

$$e_t^* = \emptyset_e(A_t, E_t, B_t; p_t, w_t, W_t, \psi, \delta, u, f, h, I) \tag{7b}$$

The objective of this study lies in identifying the effects of smoking in the observational behavioral data. According to the clinical studies, $d\psi/dA < 0$ and $dA/ds > 0$, which together imply $d\psi/ds < 0$. The question then is, how a decrease in ψ due to smoking translates into educational inputs e_t^* and E_{t+1}^*. 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*. The poof appears in the Appendix.

Effects of Smoking on Education: *Suppose that the value function v of the recursive dynamic programming version of the model (1)-(3) exists, is twice-differentiable, and is concave in endogenous arguments. Then conditional on educational achievement up to period t,* E_t^**, both the demand for education input* e_t^* *and educational achievement* E_{t+1}^* *decrease with a decrease in* ψ*. Because smoking decreases* ψ*, an increase in smoking has negative effects on both education 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 in Eq. (6b).

In the empirical specification, E_t^* is approximated by test scores in year *t* and $\sum_{\tau=0}^{t} e_\tau^*$ by years of schooling up to year *t*. The obvious endogeneity arises because common factors affect both smoking s_t^* (and A_t^*) and educational input e_t^* (and E_t^*). The next section will discuss the identification strategies to address this problem.

3. Econometric Model

This study attempts to identify empirically the causal effect of smoking on educational outcomes for teenagers, while taking 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 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 equation (2) as $E_{t+1}^* = \varphi(s_t^*, e_t^*, E_t^*)$. Substituting (7b) and linearly approximating this equation, we obtain:

$$\mathrm{E_i} = \mathbf{X}_\mathrm{i}'\boldsymbol{\beta}_\mathbf{1} + \gamma \mathrm{S_i} + \varepsilon_{1\mathrm{i}} \tag{8}$$

where S_i is observed smoking behavior, and $\mathbf{X}_i$ denotes a vector of covariates, including the constant term, that can influence learning outcomes, such as academic inputs, years of schooling, and learning efficiency.

There are three empirical challenges to estimating equation (8). 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 drop out of school. Secondly, OLS estimates of equation (8) may also suffer from a downward bias because of measurement errors in the smoking behavior variable. Though there is no legal smoking age, smoking under age of 18 is strictly forbidden in China. Teenagers thus 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.[4] Lastly, the smoking variable may suffer from a censoring problem. This study considers two smoking variables: (i) whether one has *ever* smoked; (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 as the exogenous instruments. Teenagers' demand for cigarettes is mainly determined by their total budget, or pocket money. The supply of alcohol and food prices are expected to affect the household consumption of alcohol and foods, resulting in a change in the household expenditures and the budget available for children's pocket money. We do not use the overall price index because it captures the prices of some educational inputs and can directly affect educational outcomes.

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

[4] 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.

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 there is still variation in our IVs within school, because children from different communities may attend the same school.

To account for the large number of zero observations in the cigarettes 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 of 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 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^*:

$$\mathrm{S_i} = \mathrm{S_i^*} \text{ if } \mathrm{S_i^*} > 0; = 0 \text{ otherwise} \tag{9}$$

$$\mathrm{S_i^*} = \mathbf{Z}_\mathrm{i}'\boldsymbol{\beta}_\mathbf{2} + \varepsilon_{2\mathrm{i}}. \tag{10}$$

We assume that the error terms are normally distributed with zero means, variances σ_{ε_1}, σ_{ε_2} and covariance $\sigma_{\varepsilon_1\varepsilon_2}$. Since some of the variables in **X**, such as family background and personality, may also affect youth smoking behaviors, the vector of **Z** also includes all the explanatory variables in **X**, in addition to the excluded instrumental variables that affect only the smoking decision**.** Following Vella (1993), we first estimate a Tobit model in equation (10) 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:

$$E(S_i^*|S_i) = I_i \mathbf{Z}_i' \widetilde{\boldsymbol{\beta}}_2 + (1 - I_i)\{\mathbf{Z}_i' \widetilde{\boldsymbol{\beta}}_2 - \widetilde{\sigma}_{\varepsilon_2} \widetilde{\phi}_i (1 - \widetilde{\Phi}_i)^{-1}\}. \tag{11}$$

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

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 more likely to help their children with schoolwork. Parental smoking may expose children to secondhand smoke 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). Household land assets are both a measure of household economic resources and an indicator of the household need for child labor. Total years of schooling in the previous time period is also controlled for; however, since years of schooling may be correlated with some unobserved variables, the age variable is used instead to approximate years of schooling.[5]

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.

[5] More than 99% of the children in the sample used for the test score regressions were currently enrolled in school. Therefore, their ages can be used to approximate their years of schooling.

To estimate equation (7b), 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; and 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 i-th child. Following Glewwe and Jacoby (1994), y_i^* and y_i are related to each other as follows:

$$y_i^* = \mathbf{X}_\mathrm{i}'\boldsymbol{\beta}_\mathbf{3} + \delta \mathrm{S}_\mathrm{i} + \eta_{1i} \tag{12}$$

$$y_i = j, \text{ if } \theta_{j-1} \leq y_i^* < \theta_j \quad \text{for } j = 1, \ldots, m \tag{13}$$

where the elements of $\boldsymbol{\beta}_3$ are coefficients associated with all covariates in $\mathbf{X}$, and m is the highest level of y_i. Again, the smoking variable S is endogenous. As in Subsection 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 (9) and (10).

Assuming that η_1 is *i.i.d.* and follows the standard normal distribution, the probability that $y_i = j$ is $Pr\,(y_i = j|\mathbf{X}_\mathbf{i}) = \Phi(\theta_j - \mathbf{X}_\mathrm{i}'\boldsymbol{\beta}_\mathbf{3} - \delta\mathrm{S}_\mathrm{i}) - \Phi(\theta_{j-1} - \mathbf{X}_\mathrm{i}'\boldsymbol{\beta}_\mathbf{3} - \delta\mathrm{S}_\mathrm{i})$ where Φ is the standard normal CDF. 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 $Pr\,(y_i = j|\mathbf{X}_\mathbf{i}) = 1 - \Phi(\theta_{j-1} - \mathbf{X}_\mathrm{i}'\boldsymbol{\beta}_\mathbf{3} - \delta\mathrm{S}_\mathrm{i})$. 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:

$$\ln L(\boldsymbol{\beta_3}, \delta, \boldsymbol{\theta}) = \sum_{i=0}^{N} \sum_{j=0}^{m} I_{ij} \{\ln[\Phi(\theta_j - \mathbf{X}_i'\boldsymbol{\beta}_3 - \delta S_i)^{1-d_i} - \Phi(\theta_j - \mathbf{X}_i'\boldsymbol{\beta}_3 - \delta S_i)] \tag{14}$$

If S were exogenous and uncensored, maximizing the above log-likelihood function yields the consistent and efficient estimates of $\boldsymbol{\beta_3}$, δ, and $\boldsymbol{\theta}$.

However, we have the same empirical challenges as discussed in Subsection 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 Subsection 3.1, 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 (14); and we then estimate parameters using the standard 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.

4. Data and background

The first wave of the Gansu Survey of Families and Children (GSCF) was conducted in the year 2000. Data were collected from a random sample of 2,000 children in rural areas of Gansu province 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 1,078 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 2,000 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 some baseline characteristics are used from 2000 GSCF.

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. More specifically, the GSCF collected comprehensive information on scores of tests administered by the school from the homeroom teacher of each sample child.[6] Homeroom teachers usually have accurate records of 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 final exam scores in the last two semesters for math and Chinese. 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 averaged scores: (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 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 most of the percentiles, the mean standardized test scores of smokers are clearly lower than those of non-smokers, for both subjects.

---

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

The other educational outcome investigated in this study is educational attainment, measured by total years of schooling. 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 is 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, as the survey was conducted right after the end of the 2004 school year. On average, the total years of schooling is 7.2 for those currently enrolled. Surprisingly, Table 1 shows that the average years of schooling of smokers is slightly higher than that of non-smokers. This may be due to measurement errors in smoking variables for dropouts. Because some dropouts live at home, their smoking behaviors could be under-reported because their interviews were conducted at their homes, where anonymity was not guaranteed (as opposed to interviews conducted at schools, where questionnaires were completed without adults present).

Table 2 presents descriptive statistics for the key variables used in the analysis. On average, 12% of the GSCF sample have smoked at least once. Among those who have 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 report having smoked a positive amount 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.[7] 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 recording. 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

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

experience cravings for cigarettes that are too strong to resist, even during school hours.

The counts of registered vendors of alcohol are calculated based on registration records from the online database of China's Department of Commerce. Unfortunately, these data are available for only about 40% of the GSCF sample[8]. 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 for which the alcohol vendor information is available. This issue is discussed in more detail when the results are discussed below.

The proportion of fathers who smoked is 77%, 82% for teens who smoked and 76% for those who did not. The rate of smoking among mothers is very small compared to that of fathers. In fact, only 7 out of the 1845 mothers report that they have 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 are female, while the other 95.5% are 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.

[8] The online database of China's Department of Commerce is still under construction. Since some counties in Gansu haven't yet joined the database, information on registered alcohol venders in those counties is missing.

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. The estimated marginal effects are generally larger for the *"current smoking intensity"* regression than the *"ever-smoking"*. For example,$(\partial \mathrm{E}[\mathrm{S}|\mathrm{S}>0])/\partial vender = -0.006$ and$(\partial \mathrm{E}[\mathrm{S}=1])/\ \partial \mathrm{vender} = -0.001$, which implies that participation in smoking is often experimental and is less responsive to teenagers' reduced budget.

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 parental 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 report 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.

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 pretty 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, household economic resources have negative effects on youth smoking, suggesting that youth smoking is an inferior good. Although children from richer families are subject to looser budget constraints, they may be better informed of the harmfulness of smoking, as they may have more access to information resources such as the internet.

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 significantly 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 outcomes.

Table 4 presents estimates of the effect of youth smoking on educational achievement, as measured by standardized test scores on math and Chinese (averaged over two semesters, using tests conducted in January and June of 2004).[9] 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 1-3 examine the effect on educational achievement of *ever-smoking*; columns 4-7 investigate the effect of *current smoking intensity*. 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.[10] The IV estimates are reported in columns 3, 6, and 7, respectively, for each of the measures of smoking.[11] The two regressions in columns 6 and 7 correspond to the IV regressions using the predicted latent smoking intensity variable (column 6) and the predicted observed smoking intensity variable (column 7). The standard errors for all two-step IV estimations are obtained by bootstrapping, using 300 replications. All regressions include all the control variables reported in Table 3 as well as grade fixed effects and school fixed effects.

In both the OLS and the two-step IV regressions, ever-smoking status does not have a significant impact on students' academic performance on either Math or Chinese tests. The estimated coefficients do suggest, despite their lack of statistical significance, that smoking has a negative impact. In contrast, the estimated coefficient of current smoking intensity is significantly negative for Math in the two-step IV regressions. Although the OLS estimates of current smoking intensity are also significant for Chinese test scores, the coefficients are insignificant in the two-step

---

[9] Regressions that use only the January scores or only the June scores give similar, though slightly less precise, results.

[10] In general, the magnitude, sign, and statistical significance of the estimates for the two measures of youth smoking do not differ significantly between the full and the partial samples.

[11] We also estimate the regressions for the full sample, using only food price index as the IV. The results are similar but less robust, which is mainly because that the food price index alone appears a weak IV.

IV regressions. According to the IV estimates, smoking one additional cigarette per day decreases the test scores on math by approximately 0.093 standard deviations. Since the average teenager who smokes is smoking 3.5 cigarettes per day, smoking can have a large effect on math test scores: approximately 0.33 standard deviations of a test score. Comparing the IV estimates and those of the OLS, the magnitude of the impacts of smoking intensity increases for math, while it decreases for Chinese, after controlling for endogeneity and measurement error bias.

These findings are consistent with 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.

That the magnitude of the estimated coefficient increases for math but decreases for Chinese after controlling for the endogeneity bias implies that the measurement error bias dominates the omitted variable bias for Math while the opposite is true for Chinese. We conjecture that the difference is likely to come from the extent of the omitted variable bias, since the extent of the measurement error bias is likely to be similar for Math and Chinese.

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 which are biologically affected by nicotine 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 not likely to be interfered by smoking, e.g. interest in reading Chinese novels.

The IVs easily pass standard overidentification tests for the current smoking intensity regressions, but not for the ever-smoking regression. The problem is that there is something in the ever-smoking regression error term that is correlated with the IVs, but it is not in the error term in the other regression. Since the smoking intensity regression passes the overidentification test and the weak instrument test, as discussed in Section 5.1, it offers the most reliable estimate.

5.3 Years of schooling.

Table 5 presents the results of a censored ordered probit (COP) that estimates the impact of youth smoking on educational attainment, as measured by total years of schooling. As in Subsection 5.2, there are seven regressions: columns 1-3 examine the effect on educational achievement of *ever-smoking*; regressions 4-7 investigate the effect of *current smoking intensity*. Estimates are shown for both the full and the partial sample, for comparison.[12] The same IVs are used, namely, the count of alcohol vendors and the food price index. All the regressions include school fixed effects and the same set of control variables as in the educational achievement specification.

Although our IVs pass the overidentification tests, the estimated coefficients of youth smoking are insignificant for all specifications. There are several possible reasons for this result. 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 variable for 90% of observations. Third, youth smoking may have adverse impacts on learning (i.e. its effect on test scores) but may have only minor impacts on years of education. Lastly,

[12] 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.

we lack dynamic data on smoking behavior over time. As discussed in the theoretical model, 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 1990's, 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 test scores in math exams by about 0.1 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.

Our results also indicate substantial parental effects on youth smoking. Parental 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 total years of schooling for them. Though we use 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: Proof of the Effect of Smoking on Educational Outcomes

Implicitly differentiate the system of equations (6) with respect to ψ and e_t^*. By the implicit function theorem, we have:

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

where Δ is the determinant of the Hessian of the objective function (5) and is ≤0 since the objective function is concave in endogenous arguments.

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

By concavity of the utility function, $u_{xx}u_{ss} - u_{xs}^2 \geq 0$. For the production function of addictive stock, $f_{ss} \geq 0$ 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 input which is positive, the term in the brackets is non-positive. Thus we have $de_t^*/d\psi \geq 0$. Furthermore, educational achievement E_{t+1}^* also increases with ψ conditional on E_t^*:

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

References

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.

Cook, P.J., Moore, M.J., 1993. Drinking and schooling. Journal of Health Economics 12, 411-29.

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.

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.

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.

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), 345-95.

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.
Mccaffrey, D.F., Pacula, R.L., Han, B., Ellickson, P., 2009. Marijuana use and high school dropout: The influence of unobservables. Health Economics, forthcoming.
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.
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, Massachusetts.
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, Massachusetts.
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, 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.

Table 1: Comparison of Educational Performance of Smokers and Non-Smokers

	Smokers		Non-smokers	
	Obs.	Mean	Obs.	Mean
Standardized math scores below 5th percentile	11	-2.49	72	-2.42
Standardized math scores below 10th percentile	26	-1.98	140	-2.04
Standardized math scores below 25th percentile	57	-1.44	353	-1.35
Standardized math scores below 50th percentile	111	-0.83	709	-0.77
Standardized math scores below 75th percentile	155	-0.49	1,077	-0.36
Standardized math scores below 90th percentile	181	-0.29	1,298	-0.14
Standardized math scores below 95th percentile	192	-0.19	1,367	-0.07
Standardized math scores for all	202	-0.11	1,439	0.02
Standardized Chinese scores below 5th percentile	13	-2.42	70	-2.42
Standardized Chinese scores below 10th percentile	24	-1.95	141	-1.93
Standardized Chinese scores below 25th percentile	64	-1.29	347	-1.28
Standardized Chinese scores below 50th percentile	112	-0.86	706	-0.77
Standardized Chinese scores below 75th percentile	155	-0.52	1,077	-0.37
Standardized Chinese scores below 90th percentile	185	-0.28	1,288	-0.16
Standardized Chinese scores below 95th percentile	198	-0.18	1,359	-0.08
Standardized Chinese scores for all	202	-0.13	1,439	0.02
Total years of schooling	222	7.77	1,622	7.07
Dropout (1=yes)	222	0.08	1,623	0.10

Table 2: Descriptive Statistics of Key Variables (2004)

	Obs.	Mean	S.D.	Min	Max
Standardized scores on Mathematics	1,641	0.0	1.0	-4.5	2.4
Standardized scores on Chinese	1,641	0.0	1.0	-5.1	2.4
Total years of schooling	1,844	7.2	1.8	0	12
Ever smoked (1=yes)	1,845	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[1]	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	1,845	14.6	1.2	13	17
Sex (1=male)	1,845	0.53	0.50	0	1
Father's years of schooling	1,845	7.0	3.6	0	15
Mother's years of schooling	1,845	4.3	3.5	0	13
Father smoking (1=yes)	1,845	0.77	0.42	0	1
Mother smoking (1=yes)	1,845	0.00	0.06	0	1
Household expenditures p.c. in 2000 (yuan)	1,845	1,423	982	130	13,876
Log of household land assets (mu[2])	1,839	2.0	0.8	-1.6	4.4
Distance from junior high school (km)	1,845	3.7	4.2	0	30
Distance from senior high school (km)	1,845	12.0	12.7	0.3	80
Average wage rate (yuan)	1,735	18.4	6.7	8	50
Counts of registered vendors of alcohol	716	21.6	30	0	99
Fiid price index	1,788	113.4	2.7	108	118

Note: 1. Calculated for only those who reported a positive amount of cigarettes smoked per day in the past one month.
2. 1 mu = 667 square meters

Table 3: First Stage Estimation of Smoking Choices

	Ever smoked, probit[1,2]		Current smoking intensity, tobit	
Instrumental variables				
Counts of registered vendors of alcohol	-0.014 (0.004)	***	-0.056 (0.017)	***
Food price index	-0.143 (0.028)	***	-0.494 (0.148)	***
Other explanatory variables				
Age	0.185 (0.083)	**	0.585 (0.428)	
Sex(1=male)	1.820 (0.277)	***	7.960 (1.287)	***
Father smoking (1=yes)[3]	0.466 (0.206)	**	2.150 (0.953)	**
Father's years of schooling	0.006 (0.024)		0.046 (0.114)	
Mother's years of schooling	-0.008 (0.024)		-0.032 (0.116)	
Log of household expenditures p.c. in 2000	-0.273 (0.153)	*	-1.764 (0.709)	**
Log of household land assets	-0.194 (0.164)		-0.248 (0.772)	
Average wage rates (yuan)	-0.016 (0.016)		-0.050 (0.077)	
Distances to the closest upper secondary schoo	-0.008 (0.006)		-0.017 (0.029)	
Distances to the closest lower secondary schoo	-0.051 (0.030)	*	-0.058 (0.131)	
Grade fixed effects	Yes		Yes	
School fixed effects	Yes		Yes	
Obs.	636		674	
Log likelihood	-176		-375	
Weak instruments test[4]	23.77		17.13	
	[0.000]		[0.000]	

Note: 1. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level
2. Robust standard errors are included in the parentheses and p-values are included in square brackets.
3. Mothers' smoking status is automatically dropped because only four mothers in the GSCF sample smoked.
4 Log likelihood ratio tests against the explanatory power of excluded IVs.

Table 4: Effects of Youth Smoking on Educational Achievement

	Ever-smoked[1,2,3]			Current smoking intensity			
	OLS		2-Step	OLS		2-step (corr. for censoring)	
	Full	Partial		Full	Partial	Observed	Latent
Math	-0.143	-0.193	-0.314	-0.081 **	-0.081 **	-0.117 **	-0.086 ***
	(0.089)	(0.137)	(0.373)	(0.034)	(0.037)	(0.060)	(0.027)
Grade fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1535	605	605	1535	641	641	641
R-squared	0.14	0.14	0.13	0.08	0.15	0.17	0.16
Overidentification test[4]			9.89			0.010	
			[0.002]			[0.918]	
Chinese	-0.098	-0.091	-0.086	-0.102 **	-0.096 **	-0.073	-0.012
	(0.092)	(0.144)	(0.383)	(0.031)	(0.037)	(0.087)	(0.062)
Grade fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1535	605	605	1535	641	641	641
R-squared	0.15	0.14	0.14	0.08	0.16	0.15	0.15
Overidentification test[4]			17.15			2.174	
			[0.000]			[0.140]	

Note: 1. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level

2. Robust standard errors are included in the parentheses for OLS, bootstrapped standard errors are inclueded in the parentheses for 2-values are included in square brackets.

3. All the regressions include all the explanatory variables other than the IVs in the first stage estimation.

4. Overidentification tests obtained by assuming the first stage estimation as linear.

Table 5: Effects of Youth Smoking on Educational Attainment

	Ever-smoked[1,2,3]			Current smoking intensity			
	COP		2SCOP	COP		2SCOP	
	Full	Partial		Full	Partial	Observed	Latent
Years of schooling	0.350	0.473	-0.547	0.009	0.020	-0.265	-0.025
	(0.213)	(0.406)	(0.974)	0.057	(0.066)	(0.441)	(0.125)
School availability	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	1745	643	643	1745	672	672	672
Log likelihood	-514	-123	-121	-511.00	-127	-123	-120
Overidentification test[4]			0.18			0.32	
			[0.558]			[0.850]	

Note: 1. * Significant at 10% level, ** significant at 5% level, *** significant at 1% level

2. Robust standard errors are included in the parentheses and p-values are included in square brackets.

3. Each regression has controlled for all the explanatory variables in the first stage estimation.

4. Likelihood ratio test of the statistical significance of excluded IVs in the years of schooling equation.