THE ROLES OF HIGH SCHOOL COMPLETION
AND GED RECEIPT IN SMOKING AND OBESITY

Donald S. Kenkel
Dean R. Lillard
Alan D. Mathios

The Roles of High School Completion and GED Receipt in Smoking and Obesity
Donald S. Kenkel, Dean R. Lillard, and Alan D. Mathios
NBER Working Paper No. 11990
January 2006
JEL No. I1, I2

ABSTRACT

We analyze data from the National Longitudinal Survey of Youth 1979 to explore the relationships between high school completion and the two leading preventable causes of death – smoking and obesity. We focus on three issues that have received a great deal of attention in research on the pecuniary returns to schooling. First, we investigate whether GED recipients differ from other high school graduates in their smoking and obesity behaviors. Second, we explore the extent to which the relationships between schooling and these health-related behaviors are sensitive to controlling for family background measures and cognitive ability. Third, we estimate instrumental variables (IV) models of the impact of schooling on smoking and obesity. Although our IV estimates are imprecise, both the OLS and IV results tend to suggest that the returns to high school completion include a reduction in smoking. We find little evidence that high school completion is associated with a lower probability of being overweight or obese for either men or women. The results also suggest that the health returns to GED receipt are much smaller than the returns to high school completion.

Donald S. Kenkel
Department of Policy Analysis and
Management
College of Human Ecology
Cornell University
Martha Van Rensselaer Hall
Ithaca, NY 14853-4401
and NBER
dsk10@cornell.edu

Dean R. Dillard
Department of Policy Analysis and
Management
College of Human Ecology
Cornell University
Martha Van Rensselaer Hall
Ithaca, NY 14853-4401
drl3@cornell.edu

Alan D. Mathios
Department of Policy Analysis and
Management
College of Human Ecology
Cornell University
Martha Van Rensselaer Hall
Ithaca, NY 14853-4401
adm5@cornell.edu

I. INTRODUCTION

In addition to higher earnings, the returns to schooling may also include better health (Grossman and Kaestner 1997, Grossman 2004). In this paper we explore the relationships between high school completion and the two leading preventable causes of death – smoking and obesity (Mokdad *et al* 2004). We focus on three issues similar to those that have received a great deal of attention in research on the pecuniary returns to schooling. First, we investigate whether GED recipients differ from traditional high school graduates in their smoking and obesity behaviors. Second, we explore the extent to which the relationships between schooling and these health-related behaviors are sensitive to controlling for family background measures and cognitive ability. Third, we estimate instrumental variables (IV) models of the impact of schooling on smoking and obesity.

The positive association between schooling and health has been called "one of the strongest generalizations to emerge from empirical research on health in the United States..." (Farrell and Fuchs 1982). Yet social scientists disagree on the importance of alternative pathways linking schooling to health. One causal pathway is through health information, where general schooling helps consumers respond to new information about the health consequences of behaviors such as smoking and obesity. Alternatively, the association between schooling and health may not be causal but may instead reflect hard-to-observe individual differences between people with different levels of schooling. Educational choices reflect complex decision-making and these choices are likely to be strongly correlated with hard-to-observe factors such as cognitive ability and time preference. These factors, in turn, are also likely to be correlated with health behaviors and ultimately, health outcomes.

We contribute to research on the health returns to schooling by using rich microdata from

the National Longitudinal Survey of Youth - 1979 (NLSY79) to explore the pathways from schooling to smoking and obesity. We also contribute to empirical research that uses the IV approach to estimate the returns to schooling by introducing a new set of potentially useful educational policy variables as IVs. Previous studies use compulsory schooling requirements as IVs for educational attainment (e.g., Angrist and Kreuger 1991), but there is less relevant variation in these laws for the NLSY79 cohort. Our new IVs – related to high school graduation requirements and GED policies – show more variation over the time period relevant to the NSLY79 cohort. We use variation in the state educational policy environment that individuals faced when they were teenagers to identify the treatment effects of high school completion and GED receipt on their smoking and obesity behaviors about 20 years later.

Section II of the paper provides background on the relationships between schooling, smoking, and obesity, and section III reviews research that uses instrumental variables methods to identify the causal impact of schooling on health. Section IV provides an overview of the data and describes the empirical approach. Section V presents the results, and Section VI concludes.

II. SCHOOLING, SMOKING, AND OBESITY

Table 1 provides descriptive evidence on the association of schooling with smoking status and weight. The evidence is from the reports by Schoenborn, Vickerie and Barnes (2003) and Schoenborn, Adams and Barnes (2002) that describe findings from the 1997-98 National Health Interview Survey. The estimates apply to the U.S. civilian noninstutionalized population, but have been age-adjusted to allow for comparisons across population subgroups with different age structures. The evidence shows a strong association between schooling and smoking among both men and women. For example, in 1997-1998, about 38 percent of men and 30 percent of women with less than a high school education were current smokers, compared to only 9 percent of men

2

with graduate degrees and 8 percent of women with graduate degrees. More schooling is generally associated with less smoking, but the patterns across schooling levels are complex. Most notably, men and women GED recipients are the most likely to currently smoke, about 10 percentage points more likely than high school dropouts. As another example, the gaps in smoking behavior between four-year college graduates and people with less than four years of college are substantially larger than the gaps in smoking behavior between people with less than four years of college and high school graduates. Using similar data from multiple waves of the National Health Interview Survey, Pierce *et al.* (1989) find that the gaps in smoking behavior across groups with different schooling levels have tended to widen, not narrow, over time.

Table 1 also documents a strong association between schooling and weight, at least among women. Fifty nine percent of women with less than a high school education are overweight. The fraction of overweight women drops fairly steadily with more schooling, reaching a low of 29 percent among women with a graduate degree. In contrast, for men the schooling-overweight gradient is non-monotonic and less steep overall. These general patterns are consistent with an earlier comprehensive review of research on the relationship between socioeconomic status and obesity in developed economies (Sobal and Stunkard 1989) and evidence from the World Health Organization's MONICA project (WHO 2000).

Research in health economics moves beyond descriptive evidence to explore causal pathways that could generate health returns to schooling. In Grossman's (1972, 2000) seminal model of the demand for health, schooling may play a causal role by increasing the efficiency of the household production of health. In particular, schooling may help people become better informed about the health consequences of behaviors like smoking and obesity, thus leading to healthier choices. The impact of health information on smoking behavior is documented in a

series of econometric studies that suggest that part of the large decrease in smoking prevalence after the 1964 Surgeon General's Report is due to improved consumer information about the risks of smoking (Hamilton 1972, Lewit, Coate and Grossman 1981, Schneider, Klein and Murphy 1981, Blaine and Reed 1994, Sloan, Smith and Taylor 2002).[1] Moreover, Farrell and Fuchs (1982) find that the strong negative correlation between schooling and smoking initiation developed only after the spread of information about the health hazards of smoking. More recently, DeWalque (2004) uses retrospective data from multiple years of the National Health Interview Survey to more thoroughly explore the evolution of the link between smoking and schooling. He finds that as information about the health hazards of smoking diffused, smoking prevalence declined earlier and more dramatically for college graduates.

Taking a somewhat different approach, several studies shed light on the importance of the causal pathway from schooling through health information by using microdata that provides direct measures of respondents' knowledge of the health consequences of their behaviors. Using data from the 1985 Health Interview Survey, Kenkel (1991) finds that both health knowledge and schooling are associated with a lower prevalence of smoking and more exercise. While part of the relationships between schooling and the health behaviors are explained by differences in health knowledge, most of the association with schooling remains after one controls for differences in knowledge. Nayga (2000) conducts similar analyses of the relationship between

_____

[1]These studies exploit information "shocks"in the U.S., including the 1964 Surgeon General's Report, the anti-smoking messages broadcast during the Fairness Doctrine era (1968-70), and the 1971 ban on television and radio advertising of cigarettes. Studies reviewed by Kenkel and Chen (2000) demonstrate that over the time period of the information shocks discussed above, people's knowledge about the health consequences of smoking sharply increased. For example, in a 1954 Gallup poll, only 41 percent of adults agreed that cigarette smoking causes lung cancer, but this proportion had increased to 66 percent in a 1964 survey and to over 90 percent in the 1990 National Health Interview Survey.

schooling, health knowledge and obesity. He finds that the relationship between schooling and obesity is no longer statistically significant once knowledge of diet-disease relationships is accounted for.

To the best of our knowledge, there are no studies of the health returns to schooling that distinguish between traditional high school completion versus GED receipt.   In his review of recent research on the earnings returns to the GED, Tyler (2003, p. 370) notes a "dearth of research" on the impact of the GED on outcomes outside the labor market: "Whether or not the GED affects outcomes such as parenting skills, health, citizenship, and involvement in crime are critical questions that have, unfortunately, not been addressed by the research community." Receipt of the GED appears to require relatively little investment in human capital, so large health returns to the GED seem unlikely. For example, while a typical year of high school involves over 400 hours in core curriculum classes, in 1989 the median examinee spent only 30 hours in preparation for the GED (Boesel, Alsalam and Smith 1998). The descriptive evidence in Table 1 that GED recipients are more likely to smoke than are high school dropouts suggests that any positive health returns to the GED are swamped by other factors. The pattern is consistent with evidence that while GED recipients have higher cognitive skills than dropouts they tend to be "...'wiseguys,' who lack the abilities to think ahead, to persist in tasks, or to adapt to their environments." (Heckman and Rubinstein 2001, p. 146).   Lower non-cognitive skills might explain the pattern in Table 1 that GED recipients are more likely to start smoking and less likely to successfully quit.

In sum, the evidence reviewed in this section is only partly consistent with a causal pathway from schooling through health information to smoking and obesity outcomes.   There is strong evidence that information about the health consequences of smoking matters and the link

between schooling and smoking only appeared after such information emerged. However, there is mixed evidence on whether the links between schooling and smoking and obesity outcomes persist after controlling for individual differences in health information. Moreover, with the increasingly wide dissemination of health information, the differences in smoking across schooling levels might be expected to narrow over time. Instead they have widened. Similarly, because the health consequences of smoking are now common knowledge (Kenkel and Chen 2000), it seems somewhat implausible that the informational advantage of college graduates over high school graduates explains the difference in their current smoking prevalence.[2] There may be other causal pathways linking schooling and smoking, such as peer effects. Or, the link may be at least partly non-causal. We turn next to studies that take an instrumental variables approach to shed more light on whether the associations between schooling and health reflect causation.

III. INSTRUMENTAL VARIABLES STUDIES OF SCHOOLING AND HEALTH

A number of health economics studies use the IV approach and its extensions to explore the causal impact of schooling on health. In an important study, Berger and Leigh (1989) explore the impact of schooling on blood pressure and health limitations. In one model of selection into schooling the identifying IVs include ancestry, per capita income and expenditures on education in state of residence in childhood; in the other model the IVs are the individual's IQ, Knowledge of Work test scores, and parents' schooling. Sander (1995a, b) estimates IV models of the impact of schooling on smoking and quitting smoking. His IVs include parents'

---

[2]One development in the past 20 years is the emergence of effective pharmaco-therapies for smoking cessation (USDHSS 2000). This raises the possibility that even if college graduates do not have an important informational advantage about the health consequences of smoking, they may be better informed about how to quit. Several recent studies find similar relationships between schooling and the use of other types of medical care (Goldman and Lakdawalla 2001, Goldman and Smith 2002, and Lichtenberg and Lleras-Muney 2002).

schooling, number of siblings, and rural residence/ region of residence at age 16. Leigh and Dhir

(1997) estimate IV models of the impact of schooling on disability and exercise. Their IVs

include parents' schooling, parents' income, and state of residence in childhood. Using a

somewhat different identification strategy, a second wave of IV studies of the schooling-health

link relies on variation in educational policies and other 'natural experiments.' In Lleras-

Muney's (2005) study of schooling and adult mortality, the IVs are compulsory attendance laws,

child labor laws, and state characteristics at age 14. Adams (2002) uses the same IVs as Lleras-

Muney in his analysis of the impact of schooling and functional ability and self-rated health.

Currie and Moretti (2003) use college openings in a woman's 17th year as an IV for women's

schooling to identify maternal schooling's impacts on birthweight, prenatal care, and smoking.

Internationally, Spasojevic (2003), Chou, *et al.* (2004) and Arendt (2004) use compulsory

education reforms in Sweden, Taiwan, and Denmark respectively as IVs to estimate the impact of

schooling on various health outcomes. Grossman (2004, p. 633) concludes that the results of the

recent studies "suggest causality from more schooling to better health" and notes that "the IV

estimates often exceed the OLS estimates...."

The IV studies of the schooling-health link face the same criticisms of the IV method in

general: the validity of the exclusion restriction and the weak instruments problem. The

exclusion restriction is the assumption that the instruments that are correlated with schooling can

be excluded from the health equation: *i.e.* the IVs are not direct determinants of health and are

not correlated with unobservable determinants of health. For example, in the earlier studies by

Berger and Leigh (1988), Sander (1995a,b), and Leigh and Dhir (1997), the authors assume that

variables like parents' schooling can be excluded from the health outcome equations. However,

if more educated parents invest more in their children's health and stock of health knowledge,

this exclusion restriction is invalid and the resulting estimates of the causal impact of schooling on health may be biased.

The exclusion restrictions invoked in the more recent studies by Lleras-Muney (2005), Currie and Moretti (2003), Chou, *et al.* (2004) and Spasojevic (2003), are easier to defend. Other than through their effect on schooling choices, it seems implausible to argue that educational policies and schooling reforms should be direct determinants of health outcomes. However, these more recent studies raise the weak instruments problem that occurs when the correlation between the instruments and schooling is low. In such cases, bias may remain in the estimate of the effect of schooling on health because, in fixed samples, the lower the correlation between the endogenous variable and the instrumental variables, the greater will be the inconsistency of the IV estimator (Staiger and Stock 1997). In fact, Staiger and Stock discuss this problem in the context of the use of the IV technique by Angrist and Kreuger (1991), who used compulsory schooling laws and the quarter in which a person was born to predict whether an individual dropped out of school. Because Angrist and Krueger use simple measures of the compulsory schooling laws there is little cross-state variation in the laws and therefore low correlation between schooling and the laws.

In sum, previous research establishes the IV approach as a potentially useful method to identify the causal impact of schooling on health outcomes. It should be stressed that although these IV estimates provide evidence of causal health returns to schooling, they provide less guidance about the specific causal pathways involved. Many of the studies tend to implicitly or explicitly assume that the causal pathway is through health information. However, schooling may influence health behavior through a variety of other channels including peer effects, occupation, and income. The IV estimates of the total causal effect of schooling includes both

the direct effect of schooling and the indirect, but still causal, effects through these other channels. Our study improves on previous IV studies because we will be able to both rely on defensible exclusion restrictions and bring to bear a richer set of instruments to predict schooling than has been previously used.   However, similar to previous IV studies, we will not be able to identify schooling's specific causal pathway; instead, we focus on identifying the total causal effect of schooling.

IV. DATA AND EMPIRICAL APPROACH

Data

We use data from the 1998 wave of the National Longitudinal Survey of Youth - 1979 (NLSY79). The NLSY79 includes youth aged 14 to 21 as of December 31, 1978. In the initial survey, 12,686 individuals were surveyed in 8,770 unique households. Retention for the total sample as of 1998 was 8,399, yielding potential samples of 4,100 men and 4,299 women. After deleting observations with missing information on the dependent or independent variables, our final samples of analysis consist of 3,205 men and 3,422 women.[3] Table 2 presents descriptive statistics for the dependent and independent variables.

Specification of OLS Models

In our first sets of models, we use ordinary least squares to estimate models of the general form:[4]

---

[3]Our sample sizes vary with the dependent variable we analyze; see Tables.

[4]All of our measures of health behavior are dichotomous, so we estimate linear probability models. These models can be directly compared to our IV two stage least squares models estimated below. Although we could use the IV approach and replace linear probability models with probit models (Maddala 1983), Angrist (2001) suggests that the use of two stage least squares in this type of application is often reasonable. He notes that the difficulties with endogenous dummy regressors in nonlinear models like probit "are usually more apparent than real."

(1) $\text{Health behavior}_{i,1998} = f(\text{schooling}_{i,1998}, X_{i,1998}, u_{i,1998})$

We estimate different versions of equation (1) corresponding to: four health behaviors; two different approaches to measuring schooling; and two different sets of control variables X. We also estimate each version separately for men and women, resulting in a total of 24 OLS models.

In terms of health behaviors, our first dependent variable measures self-reported current smoking participation. While smokers report less than their true consumption (Warner, 1978), a meta-analysis of studies that compared contemporaneous self-reported smoking with biochemical markers of smoking shows that people fairly accurately report *whether* they smoke (Patrick *et al.*, 1994). In our analysis samples, by this measure about 27 percent of men and 24 percent of women are current smokers. Current smoking participation reflects a series of past decisions about smoking initiation and cessation. To shed more light on this process, we also use a measure of former smoking status as our second dependent variable.[5] This measure is defined only for respondents who report ever smoking more than 100 cigarettes, and takes a value of one when such a respondent reports not smoking in 1998. About 41 percent of male ever smokers and 44 percent of female ever smokers are thus classified as former smokers in the 1998 NLSY79. Our NLSY79 measures can be compared to national data: in 1998 about 25 percent of US adults were smokers, and about 50 percent of ever smokers had quit (Schoenborn, Vickerie and Barnes 2003). Given the younger age of the NLSY79 respondents, it makes sense that our samples have somewhat more current smokers and somewhat fewer former smokers.

---

[5]Using a single cross-section from the NLSY79 we can not fully capture the dynamics of smoking decisions. In other work, we estimate discrete time hazard models of smoking cessation decisions over time (Kenkel, Lillard and Mathios 2002). A similar approach could be used to extend the current analysis by estimating discrete time hazard models of smoking initiation and cessation.

Our third and fourth dependent variables indicate whether the respondent is overweight or obese. Clinical guidelines define overweight and obese based on the body mass index, which is defined as weight in kilograms divided by height in meters squared (kg/m^2) (National Heart, Lung, and Blood Institute 1998). We calculate each NLSY79 respondents' body mass index based on self-reported weight in 1998 and self-reported height from an earlier wave of the NLSY79.[6] Our measure of overweight indicates that the respondent's body mass index is over 25 kg/m^2. Our measure of obese indicates that the respondent's body mass index is over 30 kg/m^2. The cutoffs we use follow standard guidelines, but it should be noted that instead of mutually exclusive categories, our measure of overweight includes obese. In our NLSY79 samples: about 72 percent of men are overweight and 25 percent are obese; about 55 percent of women are overweight and 27 percent are obese. These proportions are somewhat higher than national estimates from the 1997-1998 National Health Interview Survey (NHIS) (Schoenborn, Adams, and Barnes 2002). In the NHIS: 63 percent of men are overweight and 19 percent are obese; and 47 percent of women are overweight and 20 percent are obese. Because both our estimates and the NHIS estimates are based on self-reports, the measures of body mass index tend to be under-estimated on average, so the proportions overweight and obese also tend to be under-estimated.

Our key explanatory variables measure the respondent's schooling level. In the first set of models, we use a measure that indicates reported high school completion. The measure is based on whether the respondent reports having completed twelve or more years of schooling.

---

[6]In most cases the measure of height is from the 1985 wave of the NLSY79. For slightly more than 10 percent of the full sample, the 1985 measure is missing; in these cases we use self-reported height from the 1982 wave. Not all NLSY79 respondents had achieved their adult height by 1982 (Persico, Postlewaite, and Silverman 2004). As a result, we may underestimate the 1998 height of some respondents, causing us to overestimate their 1998 BMI.

We focus on high school completion because we believe, *a priori*, that this is the most relevant margin for a productive effect of schooling on health behaviors through health information. As noted above in section II, information about the health consequences of smoking is now widely disseminated. This suggests that the informational advantage associated with schooling past high school may be minimal. More generally, we agree with Grossman (2004, p. 633) that the health returns to schooling may be larger for persons with low levels of schooling, which in the U.S. context suggests a focus on the high school completion margin. Our measure of high school completion takes a value of one regardless of how those years were acquired, so it captures both traditional high school graduates and GED recipients.

In the remaining models we add a second schooling variable that indicates GED receipt. By our measures, about 87 percent of men report high school completion and about 16 percent of men report GED receipt. This suggests about 71 percent are more traditional high school graduates.[7] About 90 percent of women report high school completion and about 14 percent report GED receipt, suggesting that about 76 percent are more traditional high school graduates. By comparison, based on national data from the GED Testing Service we estimate that 13 percent of the NLSY respondents' age cohort took the GED test. Because the NLSY79 oversamples low income and minority youth – groups that disproportionately take the GED test – the proportion of GED recipients among the NLSY sample is in line with the national data.[8]

---

[7]Our measure of high school graduation mainly captures traditional graduates and GED recipients, but it also includes respondents who after dropping out of high school gained additional years of schooling through means other than the GED. In models not reported, the health behaviors of these less traditional high school completers seem to be more like traditional high school completers than they are like GED recipients.

[8]No data are available to allow us to estimate the ultimate GED pass rates by cohort. The average U.S. pass rate of people who took the GED test in a given year over this period was about 83 percent. Given that many test takers retake and ultimately pass the GED test, the cohort

Finally, note that though in equation (1) the subscript on the schooling variables reflect the schooling measured in 1998, the schooling was generally obtained many years earlier.

We use two different sets of additional explanatory control variables. The basic set of individual-level control variables measure age, age squared, race/ethnicity, and whether the respondent has a condition that limits the amount or kind of work he or she can do. The basic set also includes state fixed effects and two state-level variables that measure cigarette taxes and anti-smoking sentiment.[9] The state fixed effects reflect the respondent's state of residence in the year he first became age-eligible to drop out of high school. The tax variable is the average cigarette tax the respondent has faced since age 14, and the sentiment variable is the average anti-smoking sentiment in the states where the respondent resided in 1992, 1995, and 1998. The cigarette tax and anti-smoking sentiment variables are thus not perfectly collinear with the state fixed effects because of respondents who moved across states over the sample period. There is additional within-state variation in our measure of cigarette taxes among non-movers because respondents born in different years faced different tax histories.

Our second set of control variables adds measures of family background reflecting parental schooling and whether any member of the respondent's household in 1979 subscribed to a magazine. It also includes the respondent's schooling-corrected AFQT standardized test score as a measure of cognitive ability, and a version of the Rotter index of locus of control. The AFQT was administered to NLSY79 respondents in 1980, when some respondents had already completed high school but others had completed less schooling. To correct for this, we estimated

---

pass rate is higher. If 90 percent of GED test takers eventually pass, the national data suggest that about 11 percent of the NLSY79 respondents' age cohort passed the GED test.

[9]DeCicca, Kenkel, Mathios and Shin (2004) provide a more detailed discussion of the measure of state anti-smoking sentiment measure.

a regression of the extent to which schooling completed in 1979 is associated with the AFQT

score. We use the residual portion of the score which is not explained by 1979 schooling as our

schooling-corrected AFQT score to measure cognitive ability.[10] The Rotter index of locus of

control is based on four questions designed to measure the extent to which people believe they

have control over their own lives, as opposed to external forces such as chance or fate. Fuchs

(2004, p. 657) suggests that heterogeneity in such beliefs may help explain the correlation

between schooling and health, so we include it as a potentially important control variable

somewhat related to the idea of non-cognitive skills (Heckman and Rubenstein 2001).

Specification of IV Models

A problem with OLS estimation of equation (1) is that even with the rich set of controls

there still may be important unobserved factors that are driving both health behaviors and

schooling. In order to generate measures of schooling that will not be correlated with unobserved

individual factors we take advantage of the fact that high school completion and GED decisions

are influenced by the state educational policy environment these students faced earlier in their

lives. Specifically we will estimate two-stage least squares models that will incorporate

instrument equations of the following form:

(2) $\quad$ Schooling$_{is1998}$ = f(Graduation Requirements$_{ist}$, GED Policies$_{ist}$, Education

$\quad\quad$ Spending$_{ist}$, X$_{i1998}$, e$_{ist}$)

Our key identifying assumption is that the educational policy variables in equation (2) can

be excluded from the models of health behaviors described by equation (1). Because state fixed

effects are included in both equations (1) and (2), we rely on within-state variation in educational

---

[10]Hansen, Heckman and Mullen (2004) and Cascio and Lewis (2005) conduct more
sophisticated analyses to estimate the causal impact of schooling on the AFQT score.

policies for identification. As noted in section III, a number of recent studies of the health

returns to schooling use identification strategies that rely on similar exclusion restrictions.

Nevertheless, policy endogeneity is a potential problem: changes in state educational policies

result from the political process and are not likely to be randomly distributed across states.

However, in the context of this study, such policy endogeneity is only problematic if there are

within-state innovations that drive both changes in educational policies and smoking and obesity

outcomes. This would create correlations between the policy variables and the error term u in

equation (1).[11] One possibility is that 'progressive' states adopt different sets of both educational

policies and tobacco control policies that are potential influences on smoking behavior. To

control for this, equation (1) includes measures of state cigarette taxes and state anti-smoking

sentiment.

Furthermore, as suggested by the subscripts in equation (2), timing provides another

source of identification. Educational policy variables are assigned to individuals based on their

state of residence during the years in which they were making their schooling decisions. For

example, individuals are assigned the high school graduation requirements in force when they

entered 9th grade. State specific factors that drove innovations in those educational policies seem

unlikely to influence smoking and obesity behaviors in 1998, both because of the time elapsed

(about 20 years) and because many respondents no longer reside in the same state in which they

made their schooling decisions.[12]

---

[11]That is, in equation (2) we are not concerned with estimating the causal effect of state educational policies on schooling outcomes. So it is not necessarily a problem for us if the state policy variables are correlated with unobservable influences on schooling decisions captured by the error term e in equation (2).

[12]However, spurious correlations between state educational policies and teen health behaviors could persist to the extent that differences in teen smoking and body mass persist into

Equation (2) includes five IVs excluded from equation (1). They are variables related to the cost and difficulty of traditional high school graduation, the cost and difficulty of GED certification, and the general educational environment. The first IV is the number of courses the respondent was required to complete to graduate from high school The second IV is an indicator variable for respondents who lived in a state where local school districts (rather than the state) set minimum graduation requirements. The next two IVs are related to the ease of GED certification: the fraction of youth of the same age in each state that took the GED test in a particular year; and an index of GED policies. The index of GED policies is based on a factor analysis of nine variables measuring aspects of GED policies related to score requirements and fees.[13] Our fifth IV is the logarithm of per capita education spending, averaged over the twelve years the respondent was (or would have been) in school. We assign graduation policies, the GED variables, and per capita educational spending to each individual based on his or her state of residence in the relevant years. The Appendix provides additional details about these measures and the matching algorithm.

Because our IVs measure educational policies, another potential problem is that in states where it is easier than average to graduate high school or to obtain a GED, the average abilities of high school graduates and GED recipients will differ from the average abilities of their

adulthood.

[13]The state GED policies measured include: 1) Minimum score requirement; 2) Mean score requirement; 3) Whether must meet both the minimum and mean score requirements; 4) GED test fee (in real 2002 dollars); 5) GED certificate fee (in real 2002 dollars); 6) Whether state allows test fee to vary across GED testing centers; 7) Number of GED testing centers that reported fee information (in states with varying fees); 8) Whether fee data were imputed; and 9) Whether passing score data were imputed. We used factor analysis to reduce the dimension of the GED policy data. In all models we use only the first factor in our analysis. The factor loadings are available from the authors on request.

counterparts in other states. If ability is unobserved and is also a direct determinant of the

smoking and weight behaviors, the requirements for a good IV would be violated. However, in

our models of the smoking and weight behaviors we include explicit controls related to ability.

We include the schooling-corrected AFQT score as a measure of cognitive skills; and the Rotter

index of the locus of control as proxy for non-cognitive skills. To the extent these measures

control for differences in the abilities of high school graduates and GED recipients across states,

the state educational policy variables should be valid IVs.

In an alternative version of equation (2), we include measures of parental schooling as

additional identifying IVs that are excluded from equation (1). This specification follows the

identification approach used in earlier studies including Berger and Leigh (1988), Sander

(1995a,b), and Leigh and Dhir (1997). As noted above, on *a priori* grounds we argue that this

exclusion restriction may be invalid because parental schooling might influence a child's later

health behaviors directly, not just through the child's schooling. However, if most of the effect

is through the child's schooling, the parental schooling variables could still be useful IVs.

Because our model is over-identified, we supplement our *a priori* argument about the usefulness

of the alternative sets of IVs by conducting formal tests of the exclusion restrictions.

V. RESULTS

OLS Results

We organize our results by creating separate Tables (3 - 6) of results by health behavior.

Column 1 of each Table provides the estimates that correspond to the most basic model that

includes only the basic control variables and does not distinguish GED recipients from those who

completed a more traditional high school education. The results of the basic models suggest that

high school yields large apparent health returns in the form of less smoking. Both men and

17

women who report completing high school are much less likely to currently smoke (Table 3) . To illustrate the magnitude of the estimated relationships, where about 47 percent of male high school dropouts in our sample are current smokers, from Table 3 smoking among male high school completers is about 23 percentage points lower. The column (1) models in Table 4 show that ever smokers who report completing high school are also much more likely to be former smokers by 1998. The column (1) results in Tables 5 and 6, however, do not show apparent weight-related returns to schooling. Male high school completers are actually about 7 percentage points more likely to be overweight (Table 5). Reported high school completion is not statistically significantly related to the probability of being overweight for women (Table 5), or for the probabilities of being obese for either men or women (Table 6).

Column 2 of each Table provides the results when the indicator for GED completion is added to the models with the basic set of controls. For the smoking behaviors in Tables 3 and 4, the pattern of results suggest larger apparent health returns to high school, but much lower returns to GED receipt. Because of the way we have defined the variables, the sum of the coefficients on 'reported high school completion' and 'GED completion' provides the estimate of the impact of GED receipt. The smoking behaviors of both male and female GED recipients lie between the behavior of high school dropouts and other high school graduates. GED recipients are somewhat less likely to be current smokers and somewhat more likely to be former smokers than are high school dropouts, but less so than other high school graduates. In Tables 5 and 6, the probabilities of being overweight or obese are similar for GED recipients and other high school graduates (recalling that there were no apparent weight-related health returns to high school in the first place).

Column 3 of each Table provides estimates from models that include a richer set of

18

control variables measuring family background. The apparent smoking-related returns to schooling implied by the column 3 models are substantially smaller than those implied by the column 2 models. But reported high school completion is still associated with lower current smoking prevalence and higher former smoking prevalence for both men and women. Interestingly, although the individual's high school completion and GED receipt are not statistically significantly associated with the probability of being overweight or obese, individuals whose parents completed more schooling tend to be lighter. This association is particularly strong for women's BMI when the mother has the higher schooling. Cognitive ability, as measured by the schooling-corrected AFQT score, has strong associations with the smoking behaviors, but weaker associations with weight. The results for the Rotter index of locus of control suggest that men who believe that what happens to them is outside their control are more likely to currently smoke and are less likely to be former smokers. Locus of control is more weakly associated with women's smoking, and not associated with the probability of being overweight or obese for either men or women. Overall, the pattern of results tends to suggest that some of the apparent health returns to schooling in the column (2) models of smoking behaviors are actually due to the higher cognitive ability of reported high school completers.

To sum up our OLS results, whether or not we additional control variables, we find apparent health returns to traditional high school completion in the form of less smoking. The results suggest that the apparent health returns to GED receipt or non-traditional high school completion are much smaller. We find little evidence of apparent weight-related returns to schooling. We turn now to IV models to further explore whether the apparent returns are real.

IV Results

Columns 4 and 5 of each Table provides the results of the two-stage least squares models.

The column (4) models include the same set of explanatory variables used in the column (3) models that control for individual- and family-background. The excluded IVs for the column (4) models are the educational policy variables. The IV results in column (4) generally show the same patterns as in the OLS models, but the estimated effects are usually somewhat smaller. The parameters are also less precisely estimated, so they tend to lack statistical significance. For example, high school completion is estimated to reduce the probability of current smoking by 17 percentage points among men and 10 percentage points among women, but only the result for men is statistically significant at the 90 percent confidence level. The lack of precision also makes it difficult to compare the smoking behaviors of GED recipients versus traditional high school completers. As in the OLS models, these IV models provide no strong evidence that traditional high school completion or GED receipt are important determinants of the probability of being overweight or obese.

The Appendix provides additional information on the performance of the 'first stage' reduced-form models of reported high school completion and GED receipt. The F-statistics for the joint significance of the instruments are always greater than 10 and are almost always greater than 20, so they are not suggestive of a weak IV problem (Staiger and Stock 1997). Because of the definition of the dependent variable, the coefficients from the reduced-form equations are somewhat hard to interpret, but they generally work in the expected direction.

Column 5 of each Table provides the results of an alternative specification of the two-stage least squares models that adds parental schooling to the list of identifying IVs. We conduct over-identification tests as described by Wooldrdige (2002, pp. 122-124). Based on the results, we reject the validity of the exclusion restriction for the parental schooling variables for the

20

obesity model for men and for all of the health behavior models for women.[14] We fail to reject

the validity of the exclusion restrictions for the state education policy variables. To explore the

practical importance of treating the parental schooling variables as excluded IVs, compare the

estimated effects of high school completion in columns (4) and (5). For example, in column (5)

of Table 6, the estimate suggests that high school completion reduces the probability that a

woman is obese; the estimated effect is statistically significant and implausibly large (22

percentage points compared to a sample proportion of 27 percent obese). However, the results in

column (4) show no effect of high school completion on obesity, but do show a direct effect of

parental schooling on obesity. This pattern suggests that improperly excluding parental

schooling may result in over-estimates of the true causal effect of the child's schooling on health

behaviors.

VI. DISCUSSION

Similar to the research findings on the earnings returns to schooling, when estimating the

health returns to schooling we find that it is important to distinguish GED recipients from other

high school completers, and that it is important to control for cognitive ability. Although our IV

estimates are imprecise, both the OLS and IV results tend to suggest that one component of the

return to high school completion is a lower probability of smoking. We find little evidence that

high school completion is associated with lower probability of being overweight or obese for

either men or women. The results also suggest that the health returns to GED receipt are much

smaller than the possible returns to high school completion. On *a priori* grounds, in this paper

---

[14]The test statistic is distributed χ^2 with two degrees of freedom.   At the 95 percent
confidence level, the critical value is 5.99. For the models of men's obesity, women's smoking,
women's ex-smoking, women's overweight, and women's obesity, the test statistic values are
44.17, 42.30, 8.65, 26.86, and 9.17, respectively.

we focus on high school completion and GED receipt as the margin on which schooling is most likely to affect smoking and obesity. In future work it would be useful to examine the schooling-health link at different schooling margins. Research along those lines would require the use of different sets of IVs, such as proximity to college (Card 1993), that operate at different margins of schooling.

One approach in previous studies of the schooling-health link is to focus on general measures of health outcomes. Our study instead focuses on key health behaviors as specific pathways by which schooling can lead to differences in health outcomes. Another useful direction for future work is to study pathways at even more detailed levels. For example, because of addiction, smoking prevalence at a point in time reflects a series of past decisions about smoking initiation and cessation (DeCicca, Kenkel and Mathios 2004). While we begin to explore such dynamics by looking at the prevalence of both current and former smoking, future work could undertake a more detailed examination of the dynamics of schooling and smoking initiation and cessation decisions. In terms of obesity, not only do eating and exercise behaviors persist because of addiction or habit, but physical weight also persists over time. This suggests that it may be even more important to explore the dynamic process by which a given weight is reached. For example, while we find little evidence that individuals' schooling is related their BMI, their parent's schooling is more strongly associated; this is especially true for mother's schooling and their daughter's BMI. This association may stem from both habits and weight gained earlier in life under their parents' influence. The individual's later investments in schooling can not undo the past, but future work could explore whether such investments pay off by changing the dynamics of adult weight losses and gains.

REFERENCES

Adams, Scott J. 2002. Educational attainment and health: Evidence from a sample of older adults. Education Economics 10, no. 1: 97-109.

Angrist, Joshua. 2001. Estimation of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice. Journal of Business & Economic Statistics 19, no. 1: 2-16.

Angrist, Joshua and Alan Krueger. 1991. Does compulsory schooling attendance affect schooling and earnings? Quarterly Journal of Economics 106, no. 4: 979-1014.

Arendt, Jacob Nielsen. 2004. Does education cause better health? A panel data analysis using school reforms for identification. Economics of Education Review 24, no. 2: 149-160.

Berger, Mark C. and J. Paul Leigh. 1989. Schooling, self-selection, and health. Journal of Human Resources 24, no. 3: 433-455.

Blaine, Thomas W and Michael R. Reed . 1994. U.S. cigarette smoking and health warnings: New evidence from post World War II data. Journal of Agricultural and Applied Economics, 26, no. 2:535-544.

Boesel, D., Alsalam, N. and Smith, T. M. 1998. Educational and labor market performance of GED recipients. National Library of Education, Office of Educational Research and Improvement, U.S. Department of Education.

Card, David. 1993. Using geographic variation in college proximity to estimate the return to schooling. Working Paper no. 4483, National Bureau of Economic Research, Cambridge, MA. (published as Aspects of Labour Economics: Essays in Honour of John Vanderkamp, edited by Louis Christofides, E. Kenneth Grant and Robert Swindinsky. University ofToronto Press, 1995)

Cascio, Elizabeth U., and Ethan G. Lewis. 2005. Schooling and the AFQT: Evidence from school entry laws. Working Paper No. 11113, National Bureau of Economic Research, Cambridge, MA.

Chou, Shin-Yi, Jin-Tan Liu, Michael Grossman, and Theodore Joyce. 2004. Parental education and child health: Evidence from a natural experiment in Taiwan." Working Paper, December.

Currie, Janet and Enrico Moretti. 2003. Mother's education and the intergenerational transmission of human capital: Evidence from college openings. Quarterly Journal of Economics 118, no.4: 1495-1332.

DeCicca, Philip, Donald S. Kenkel, and Alan D. Mathios. 2004. Cigarette taxes and the transition from youth to adult smoking: Smoking initiation, cessation, and participation. Working Paper, Department of Policy Analysis and Management, Cornell University.

DeCicca, Philip, Donald S. Kenkel, Alan D. Mathios, and Justine Shin. 2002. Youth smoking, cigarette prices, and anti-smoking sentiment. Working Paper, Department of Policy Analysis and Management, Cornell University.

DeWalque, Damien. 2004. Education, information, and smoking decisions. Paper presented at the Lund Symposium on the Economics of Substance Abuse, Lund, Sweden, August 12-13, 2004.

Farrell, Phillip and Victor R. Fuchs. 1982. Schooling and health: The cigarette connection. Journal of Health Economics 1, no. 3: 217-230.

Fuchs, Victor R. 2004. Reflections on the socio-economic correlates of health. Journal of Health Economics 23, no. 4: 653-661.

Goldman, Dana P. And Darius Lakdawalla. 2001. Understanding health disparities across

education groups. Working Paper 8326, National Bureau of Economic Research, Cambridge, MA.

Goldman, Dana P. and James P. Smith 2002. Can patient self-management help explain the SES health gradient? Proceedings, National Academy of Sciences 99, no. 16: 10929-10934.

Grossman, Michael. 1972. On the concept of health capital and the demand for health. Journal of Political Economy 80, no. 2: 223-255.

Grossman, Michael. 2000. The human capital model. In the *Handbook of Health Economics*, ed. Anthony J. Cuyler and Joseph P. Newhouse. North-Holland, pp. 349-408.

Grossman, Michael. 2004. The demand for health, 30 years later: A very personal retrospective and prospective reflection. Journal of Health Economics 23, no. 4: 629-636.

Grossman, Michael and Robert Kaestner. 1997. Effects of education on health. In: The social benefits of education, ed. Jere R.Behrman and N. Stacey. Ann Arbor, MI: University of Michigan Press. pp. 69-123.

Hamilton, James L. 1972. The demand for cigarettes: Advertising, the health scare, and the cigarette advertising ban. Review of Economics and Statistics 54 (no. 4): 401-411.

Hansen, Karsten T., James J. Heckman, and Kathleen J. Mullen. 2004. The effect of schooling and ability on achievement test scores.   Journal of Econometrics 121, nos. 1-2: 39-98.

Heckman, James J. and Yona Rubinstein. 2001. The importance of noncognitive skills: Lessons from the GED testing program. American Economic Review Papers and Proceedings 91, no.2: 145-149.

Kenkel, Donald S. 1991. Health behavior, health knowledge, and schooling. Journal of Political Economy 99, no. 2:287-305.

Kenkel, Donald S. and Likwang Chen. 2000. Consumer information and tobacco use. In

    Tobacco control in developing countries, ed. Prabhat Jha and Frank J. Chaloupka.

    Oxdord: Oxford University Press, pp. 177-214.

Kenkel, Donald S., Dean Lillard and Alan Mathios. 2002. To quit or not to quit: An economic

    analysis of women's smoking cessation decisions. Working Paper, Department of Policy

    Analysis and Management, Cornell University.

Leigh, J. Paul and Rachna Dhir. 1997. Schooling and frailty among seniors. Economics of

    Education Review 16, no. 1: 45-57.

Lewit, Eugene, Douglas Coate and Michael Grossman. 1981. The effects of government

    regulation on teenage smoking. Journal of Law and Economics 24, no. 3: 545-69.

Lichtenberg, Frank and Adriana Lleras-Muney. 2002. The effect of education on medical

    technology adoption: Are the more educated more likely to use new drugs? Working

    Paper 9185, National Bureau of Economic Research, Cambridge, MA.

Lleras-Muney, Adriana. 2005. The relationship between education and adult mortality in the

    U.S. Review of Economic Studies 72, no. 250: 189-221.

Lleras-Muney, Adriana. 2002. Were compulsory attendance and child labor laws effective? An

    analysis from 1915 to 1939. Journal of Law and Economics 45, no. 2: 401-435.

Maddala, G.S. 1983. Limited dependent and qualitative variables in econometrics. Cambridge:

    Cambridge University Press.

Mokdad, Ali H., James S. Marks, Donna F. Stroup, and Julie L. Gerberding. 2004. Actual

    causes of death in the United States: 2000. JAMA 291, no. 10: 1238-1245.

National Heart, Lung, and Blood Institute. 1998. Clinical guidelines on the identification, evaluation,

    and treatment of overweight and obesity in adults – the evidence report. Obesity Research 6,

Supplement 2: 51S - 215S.

Nayga Rodolfo M. 2000. Schooling, health knowledge, and obesity. Applied Economics 32, no. 7: 815-822.

Patrick DL, Cheadle A, Thompson DC, Diehr P, Koepsell T, Kinne S 1994. The validity of self-reported smoking: A review and meta-analysis. American Journal of Public Health 84, no. 7: 1086-1093.

Persico, Nicola, Andrew Postlewaite, and Dan Silverman. 2004. The effect of adolescent experience on labor market outcomes: The case of height. Journal of Political Economy 112, no.5: 1019-1053.

Pierce, JP, M. C. Fiore, T. E. Novotny, E. J. Hatziandreu and R. M. Davis. 1989. Trends in cigarette consumption in the United States: Educational differences Are increasing. JAMA 261, no. 1: 56-60.

Sander, William. 1995a. Schooling and quitting smoking. Review of Economics and Statistics 77, no. 1: 191-199.

Sander, William. 1995b. Schooling and smoking. Economics of Education Review 14, no. 1: 23-33.

Schoenborn, Charlotte, Patricia F. Adams, and Patricia M. Barnes. 2002. Body weight status of adults: United States 1997-98. Advance Data from Vital and Health Statistics Number 330, September 6.

Schoenborn, Charlotte A, Jackline L. Vickerie, and Patricia M Barnes. 2003. Cigarette smoking behavior of adults: United States 1997-98. Advance Data from Vital and Health Statistics Number 331, February 7.

Schneider, Lynne, Benjamin Klein, and Kevin M. Murphy. 1981. Governmental regulation of cigarette health information. Journal of Law and Economics 24, no.3:575-612.

Sloan, Frank A., V. Kerry Smith, and Donald H. Taylor, Jr. 2002. Information, addiction, and 'bad

     choices': Lessons from a century of cigarettes. Economics Letters 77, no. 2: 147-155.

Sobal, Jeffrey and A.J. Stunkard. 1989. Socioeconomic status and obesity: A review of the literature.

     Psychological Bulletin 105, no. 2: 260-75.

Spasojevic, J. (2003). "Effects of Education on Adult Health in Sweden: Results from a Natural

     Experiment." Ph.D. Dissertation (City University of New York Graduate Center, New York).

Staiger, Douglas; Stock, James. 1997. Instrumental variables regression with weak instruments.

     Econometrica 65, no. 3: 557-586.

Tyler, John H. 2003. Economic benefits of the GED: Lessons from recent research. Review of

     Educational Research 73, no. 3: 369-403.

U.S. Department of Health and Human Services [USDHHS]. 2000. Clinical practice guideline:

     Treating tobacco use and dependence.

Warner, Kenneth E. 1978. Possible increases in the underreporting of cigarette consumption.

     Journal of the American Statistical Association 73, no. 362: 314 - 318.

WHO MONICA Project. 2000. Educational level, relative body weight, and changes in their

     association over 10 years: An international perspective from the WHO MONICA Project.

     American Journal of Public Health 90, no. 8: 1260-1268.

Wooldridge, Jeffrey M. 2002. Econometric analysis of cross section and panel data. Cambridge

     MA: the MIT Press.

Table 1. Smoking and overweight status by schooling level				
Schooling	**Current Smoker**		**Overweight (BMI > 25)**	
	Men	Women	Men	Women
< high school graduate	38%	30%	62%	59%
GED diploma	48%	40%	64%	56%
High school graduate	31%	26%	65%	50%
Some college - no degree	26%	22%	66%	46%
Associate degree	23%	22%	66%	46%
Bachelor's degree	14%	11%	59%	36%
Graduate degree	9%	8%	54%	29%
Sources: Schoenborn, Vickerie and Barnes (2003), Schoenborn, Adams and Barnes (2002).				

Table 2 Sample statistics				
	Men		Women	
Variable	Mean	Std.Dev.	Mean	Std.Dev.
Current smoker*	0.27	(0.45)	0.26	(0.44)
Former smoker*	0.41	(0.49)	0.43	(0.50)
Body mass index (BMI)	27.67	(4.69)	27.07	(6.39)
Overweight (BMI>25)	0.72	(0.45)	0.55	(0.50)
Obese (BMI>30)	0.25	(0.43)	0.27	(0.44)
High school graduate	0.88	(0.33)	0.90	(0.29)
GED recipient	0.16	(0.36)	0.14	(0.35)
Age	36.69	(2.24)	36.80	(2.22)
Black	0.29	(0.45)	0.29	(0.45)
Hispanic	0.18	(0.38)	0.19	(0.39)
Other race	0.05	(0.21)	0.06	(0.23)
Has health limitation	0.08	(0.28)	0.12	(0.32)
Schooling of parent with highest observed level	11.80	(3.34)	11.60	(3.37)
Observed schooling is mother's	0.67	(0.47)	0.67	(0.47)
Magazines in home when age 14	0.59	(0.49)	0.56	(0.50)
AFQT percentile score	41.10	(29.64)	38.62	(27.42)
Rotter index	3.60	(1.54)	3.61	(1.48)
State policy variables				
Average state anti-smoking sentiment	-0.01	(0.15)	-0.02	(0.15)
Average state cigarette tax	0.21	(0.08)	0.21	(0.08)
High school course graduation requirements	12.21	(7.85)	12.42	(7.79)
Graduation requirements set by local districts	0.30	(0.45)	0.29	(0.45)
Fraction of GED test takers of similar age	0.51	(0.61)	0.42	(0.54)
First factor GED policies	-0.03	(1.35)	-0.04	(1.36)
Log average per capita education spending	8.30	(0.23)	8.27	(0.23)
N	3248		3275	
Source: authors' compilations from NLSY79 and own policy compilations. *Sample sizes for the current smokers and ex smokers are 3174 and 1496 respectively for men and 3157 and 1491 respectively for women. The means of the other variables do not differ across the current smoker, former smoker, overweight, and obese samples.				

Table 3 Current smoker

Variable	Men 1	Men 2	Men 3	Men 4	Men 5	Women 1	Women 2	Women 3	Women 4	Women 5
High school graduate	-.226 ***	-.258 ***	-.209 ***	-.167 *	-.139 *	-.213 ***	-.243 ***	-.219 ***	-.103	.062
	(.027)	(.027)	(.029)	(.088)	(.080)	(.029)	(.030)	(.030)	(.130)	(.095)
GED		.156 ***	.125 ***	.203	.152		.227 ***	.210 ***	.265 *	.090
		(.024)	(.026)	(.126)	(.112)		(.025)	(.024)	(.139)	(.108)
Schooling of parent			-.002	.001				.008 ***	.007 **	
			(.003)	(.003)				(.002)	(.003)	
Observed schooling is mother's			.020	.008				-.009	-.006	
			(.016)	(.016)				(.015)	(.016)	
Magazines in home when age 14			.031 *	.015	.012			-.027 *	-.027	-.034 *
			(.018)	(.019)	(.019)			(.016)	(.020)	(.020)
AFQT percentile score			-.002 ***	-.002 ***	-.002 ***			-.004 ***	-.004 ***	-.004 ***
			(.000)	(.001)	(.001)			(.000)	(.000)	(.000)
Rotter index			.016 ***	.013 ***	.013 ***			.002	.003	.006
			(.005)	(.005)	(.005)			(.004)	(.005)	(.005)
N	3174	3174	3174	3174	3174	3157	3157	3157	3157	3157
R^2	.292	.299	.311	---	---	.282	.291	.299	---	---
F (k, n-k) on 1st stage IVs										
High school graduate				114.4	88.8				62.3	58.5
GED				37.7	29.5				37.6	32.1

Notes: Columns (1) - (3) contain results from linear probability models. Columns (4) and (5) contains results from two-stage least squares regressions. Column (4) only uses state policies as instruments. Column (5) adds the two parental schooling variables to the instrument set. Standard errors (in parentheses) are adjusted for heteroskedasticity. Coefficients denoted by ***, **, and * are statistically different from zero with p-values of <.01, <.05, and .10 respectively. All models include state fixed effects, indicators for race (black, hispanic, other), whether the respondent has a health limitation, age, age squared, a measure of state anti-smoking sentiment, and the average state cigarette tax.

Table 4 Former smoker																				
					Men										Women					
Variable	1		2		3		4		5		1		2		3		4		5	
High school graduate	.119	***	.142	***	.073	**	.063		.063		.190	***	.239	***	.164	***	.189		.203	*
	(.032)		(.033)		(.036)		(.117)		(.101)		(.035)		(.036)		(.038)		(.130)		(.108)	
GED			-.074	**	-.042		-.033		-.030				-.178	***	-.134	***	-.090		-.106	
			(.033)		(.034)		(.167)		(.148)				(.032)		(.032)		(.109)		(.101)	
Schooling of parent					.000		.000								.002		.002			
					(.005)		(.006)								(.005)		(.006)			
Observed schooling is mother's					.005		.005								.003		.003			
					(.028)		(.028)								(.027)		(.028)			
Magazines in home when age 14					-.020		-.019		-.019						.065	**	.064	*	.068	**
					(.028)		(.028)		(.028)						(.028)		(.033)		(.033)	
AFQT percentile score					.003	***	.003	***	.003	***					.004	***	.004	***	.004	***
					(.001)		(.001)		(.001)						(.001)		(.001)		(.001)	
Rotter index					-.016	**	-.016	**	-.016	**					-.003		-.004		-.004	
					(.008)		(.008)		(.008)						(.009)		(.009)		(.009)	
N	1496		1496		1496		1496		1496		1491		1491		1491		1491		1491	
R^2	.441		.443		.453		.		.		.469		.478		.496		.		.	
F (k, n-k) on 1st stage IVs																				
High school graduate							74.4		58.1								49.2		42.6	
GED							27.6		20.4								52.8		39.7	

Notes: Columns (1) - (3) contain results from linear probability models. Columns (4) and (5) contains results from two-stage least squares regressions. Column (4) only uses state policies as instruments. Column (5) adds the two parental schooling variables to the instrument set. Standard errors (in parentheses) are adjusted for heteroskedasticity. Coefficients denoted by ***, **, and * are statistically different from zero with p-values of <.01, <.05, and <.10 respectively. All models include state fixed effects, indicators for race (black, hispanic, other), whether the respondent has a health limitation, age, age squared, a measure of state anti-smoking sentiment, and the average state cigarette tax.

Table 5 Overweight																				
	Men										Women									
Variable	1		2		3		4		5		1		2		3		4		5	
High school graduate	.068	***	.067	***	.087	***	.110		.045		-.043		-.046		-.007		.098		-.119	
	(.025)		(.026)		(.027)		(.083)		(.076)		(.028)		(.029)		(.030)		(.120)		(.087)	
GED			.005		-.005		-.105		-.006				.016		-.006		-.018		.142	
			(.022)		(.023)		(.129)		(.105)				(.025)		(.025)		(.117)		(.093)	
Schooling of parent					-.005	*	-.005	*							-.011	***	-.012	***		
					(.003)		(.003)								(.003)		(.004)			
Observed schooling is mother's					-.002		.004								-.031	*	-.036	*		
					(.017)		(.017)								(.018)		(.020)			
Magazines in home when age 14					.029		.027		.023						-.024		-.031		-.018	
					(.018)		(.019)		(.019)						(.020)		(.022)		(.023)	
AFQT percentile score					-.001	**	-.001	**	-.001	**					-.001		-.001		.000	
					(.000)		(.000)		(.001)						(.000)		(.001)		(.001)	
Rotter index					.005		.003		.005						.003		.004		.003	
					(.005)		(.005)		(.005)						(.006)		(.006)		(.006)	
N	3248		3248		3248		3248		3248		3275		3275		3275		3275		3275	
R^2	.733		.733		.737		---		---		.601		.600		.606		---		---	
F (k, n-k) on 1st stage IVs																				
High school graduate							132.4		103.3								84.1		96.1	
GED							35.5		33.6								57.2		54.9	

Notes: Columns (1) - (3) contain results from linear probability models. Columns (4) and (5) contains results from two-stage least squares regressions. Column (4) only uses state policies as instruments. Column (5) adds the two parental schooling variables to the instrument set. Standard errors (in parentheses) are adjusted for heteroskedasticity. Coefficients denoted by ***, **, and * are statistically different from zero with p-values of <.01, <.05, and <.10 respectively. All models include state fixed effects, indicators for race (black, hispanic, other), whether the respondent has a health limitation, age, age squared, a measure of state anti-smoking sentiment, and the average state cigarette tax.

Table 6 Obese											
			Men						Women		
Variable	1	2	3	4	5	1	2	3	4	5	
High school graduate	-.017	-.014	.013	-.008	-.075	-.036	-.039	.005	-.022	-.215	**
	(.024)	(.024)	(.026)	(.082)	(.078)	(.028)	(.028)	(.029)	(.139)	(.108)	
GED		-.014	-.028	.041	.130		.015	-.005	.043	.197	*
		(.022)	(.022)	(.121)	(.123)		(.023)	(.022)	(.108)	(.109)	
Schooling of parent			-.006 **	-.005				-.014 ***	-.013 ***		
			(.003)	(.003)				(.003)	(.004)		
Observed schooling is mother's			.019	.018				-.028 *	-.027		
			(.016)	(.017)				(.016)	(.017)		
Magazines in home when age 14			.009	.009	.007			-.011	-.006	-.005	
			(.017)	(.019)	(.019)			(.018)	(.021)	(.020)	
AFQT percentile score			-.001 **	-.001	-.001			.000	.000	.000	
			(.000)	(.001)	(.000)			(.000)	(.001)	(.000)	
Rotter index			-.005	-.005	-.004			-.006	-.006	-.006	
			(.005)	(.005)	(.005)			(.005)	(.005)	(.005)	
N	3248	3248	3248	3248	3248	3275	3275	3275	3274	3275	
R^2	.264	.264	.265	---	---	.291	.291	.294	---	---	
F (k, n-k) on 1st stage IVs											
High school graduate				128.3	92.7				58.4	55.3	
GED				42.6	27.3				63.9	34.9	

Notes: Columns (1) - (3) contain results from linear probability models. Columns (4) and (5) contains results from two-stage least squares regressions. Column (4) only uses state policies as instruments. Column (5) adds the two parental schooling variables to the instrument set. Standard errors (in parentheses) are adjusted for heteroskedasticity. Coefficients denoted by ***, **, and * are statistically different from zero with p-values of <.01, <.05, and <.10 respectively. All models include state fixed effects, indicators for race (black, hispanic, other), whether the respondent has a health limitation, age, age squared, a measure of state anti-smoking sentiment, and the average state cigarette tax.

APPENDIX

This appendix describes how policy variables were matched to NLSY79 respondents. This matching process requires that we assign each respondent a state of residence for all years from the time he was first required to enter school until 1998. We describe the assignment of state of residence and the imputations that assignment requires. We also describe the policy variables we use as IVs. These policies include high school course graduation requirements, policies governing testing for and issuance of the certificate of General Educational Development, and the log of the average of annual per capita spending on public education in the states where the respondent lived while in grades 1-12. Annual education spending was averaged over the twelve years the respondent was (or would have been) in school. Finally, we describe how each type of policy is mapped to individuals.

State of residence algorithm

To assign a state of residence to each NLSY79 respondent we begin with the geocode data available under special arrangement with the Bureau of Labor Statistics. Those data identify each respondent's state of residence at the time of each survey. They also include information on the state in which he was born, the state he lived in at age 14, his parent's state of residence in 1979, and the state of residence at the time of each of up to six moves prior to various interview dates and the date (month and year) of those moves. For example, the base year survey asks respondents to report the state of residence at the time of up to five moves since January 1978. Although respondents were asked to report on the moves since January 1978 many respondents reported on moves that occurred prior to 1978. We code dates of moves and create an algorithm that compares states of residence at birth, age 14, and on the dates of the moves to assign a state of residence. When no information is available on dates of moves prior to 1978, our algorithm assigns states in the following ways. If the state of residence at birth, age 14 and in 1979 are all the same we assume the individual always lived

in the same state. When a person was born in a different state than he lived in at age 14, we assume he moved to that state in the year he turned 14. Similarly, when a person lived in a different state at age 14 than he did in 1979, we assume he moved to the latter state of residence in 1979.

To assign a state of residence in each year after 1979 we first code each respondent's state of residence in years he was interviewed. We then use information on the dates and states from which the respondent reported moving between interview dates. When a respondent moved across multiple state lines in a single calendar year, we assign the last state of residence in that calendar year as his state of residence. This algorithm assigns a state of residence to each individual for about 97 percent of all cases. In years that no survey was conducted in our sample period (1995 and 1997) and when no move data were available we assigned states as follows:

-when an individual lived in the same state in the bracketing years (1994 and 1996 for the missing 1995 data and 1996 and 1998 for the missing 1997 data) we assume the individual had not moved.

-when states in the bracketing years differed, we assume the individual moved in the first missing year to the state in which he resided when he completed the later survey (the later bracketing year).

We used the same algorithm for all other years for which data were missing. We followed the same imputation when data were missing for more than one year (these cases constituted less than 3 percent of assignments). The algorithm is available from the authors on request.

We then assign graduation policies, GED policies, and per capita spending on public education to each individual based on his assigned state of residence in each year. The algorithm we used is next described.

A. Grade 1-12 Education Policies

The analysis uses data on the number of high school courses his state of residence required him to complete to graduate from high school (using his state of residence in grade 9). A small number of individuals moved from one state to another starting in grade 9. For these individuals, we created the weighted average of the high school course graduation requirements during the years he would have been in high school. These course requirements are measured in terms of Carnegie units.[15] We also include an indicator for individuals who lived in a state where local school districts have primary responsibility for setting minimum graduation requirements (defined as states in which the state requires students complete five or fewer Carnegie units to graduate from high school).

B. General Educational Development Policies

The GED test is administered by the GED Testing Service (GEDTS) under the auspices of the American Council on Education. Since its inception in 1942, the GED has been advanced as a test that measures educational development at a level that is more or less equivalent to that possessed by a high school graduate. The GED test consists of a battery of exams in mathematics, social studies, science, literature and the arts, and writing skills. With the exception of the writing test, each test has a multiple choice format. The writing skills subtest has been required since 1988.

States determine the difficulty of meeting the GED score requirements by setting the minimum number of (standardized) points a test taker must score on each of the five subtests, the mean score he or she must achieve across all five subtests, and whether a test taker must meet both the minimum and mean score requirements or if he or she passes by meeting either the minimum or the mean score requirement. We term these policies as the GED minimum score requirement, the

---

[15]A Carnegie unit is earned upon successful completion of a class which meets on average fifty minutes per day, five days per week for 180 school days, the length of a typical school year

GED mean score requirement, and the GED score option respectively.[16]

The GED policy data also include the cost of taking the GED test and the cost of the GED certificate. These data consist of two types. In 27 states and the District of Columbia test fees are set at the state level and are uniform across GED testing centers. In 23 states, the state department of education allows test fees to vary across GED testing centers. For these states an average test fee and certificate fee was used, and the data include the number of GED testing centers used in the average.

To reduce the dimensions of the GED policy data we use factor analysis. The set of policies we use are listed below in section E. The estimated first factor is used as an IV.

Finally, for each person at risk to take the GED test (those who were eligible to drop out but had not yet completed a high school degree) we include a measure of the fraction of youth of the same age in each state that took the GED test in a particular year.

For each person we average the GED policies (first factor and fraction of youth who took the GED test) over the period from the year he was first eligible to drop out of school until either the year he turned nineteen or until he completed a high school degree, whichever occurred first. Although the choice to average over this period fails to capture the influence of GED policies in later years that a high school dropout was at risk to take the GED test, the empirical distribution of GED test taking shows that the hazard of taking the GED test falls dramatically for those ages 20 and older.

D. Data sources

All policy data we use were collected by Dean Lillard and his colleagues. The GED policy

---

[16]South Carolina, until 1991, and Tennessee, until 1978, required all test takers to meet only a GED mean requirement regardless of their minimum scores on each subtest. These observations are assigned a minimum of 20, the lowest standardized score assigned. At various times, some states only set a minimum score requirement. These states effectively set the mean requirement at the minimum and require test takers to meet both. In the few states, such as New Jersey and Missouri, that varied the minimum score requirement across one or more of the five tests, we assigned the minimum of the scores.

variables used in the study were compiled from a variety of sources. Most of the GED score

standards data are gathered from compilations of letters from each state department of education to

the GEDTS that were published by the American Council on Education under various titles in 1963,

1966, 1969, 1972, 1975, 1976, 1978, 1984, 1988, 1989, 1991, 1993, and 1997 and from the Annual

Statistical Report of the GED published by the GEDTS under a variety of similar titles. Lillard also

draws information from an unpublished set of letters collected by the GEDTS in 1995. These data

were supplemented by information gathered directly from web sites, administrative orders, state

statutes and memorandum from each state's department of education or agency responsible for the

GED testing program. Some of these data were transmitted in the form of personal correspondence to

Lillard. Lillard collected data on test fees in the 23 states that allow fees to vary across GED testing

centers in the state by directly surveying directors of existing GED testing centers. Gaps in policy

data were filled with data from bracketing years if the data in the two bracketing years were the same.

When changes had occurred, all data were carried forward until the year of the change and each state-

year observation was flagged with an imputation dummy.

Lillard compiled data on high school graduation requirements for each state from the Digest

of Education Statistics, from unpublished surveys of state departments of education conducted by the

Education Commission of the States, from publications from individual state departments of

education, from administrative law registers, and through personal communication between the author

and various officials of each state's department of education. These data include graduation

requirements for Iowa even though no state requirement is set there. Lillard creates an enrollment

weighted average of each school district's requirements. The data to calculate those averages were

provided to him by the Iowa State Department of Education. All data are available upon request.

Appendix Table 1 First stage equations

| | High school graduation | | | | | | | |
| | Current smoking status | | Ex smoker status | | Overweight | | Obese | |
Variable	Men	Women	Men	Women	Men	Women	Men	Women
High school course graduation requirements	-.010	-.009	-0.012	-0.011	-.010	-.006	-.009	-.008
	(.008)	(.008)	(0.015)	(0.020)	(.009)	(.010)	(.008)	(.009)
Graduation requirements set by local districts	-.024	-.080	-0.048	0.012	-.023	-.014	.004	.172
	(.133)	(.140)	(0.271)	(0.350)	(.151)	(.179)	(.148)	(.164)
Fraction of GED test takers of similar age	.112 ***	.102 ***	0.197 ***	0.188 ***	.125 ***	.127 ***	.131 ***	.124 ***
	(.008)	(.008)	(0.015)	(0.016)	(.008)	(.009)	(.009)	(.009)
First factor GED policies	-.262 ***	-.165 ***	-0.299 ***	-0.220 ***	-.285 ***	-.207 ***	-.272 ***	-.122 ***
	(.014)	(.014)	(0.022)	(0.023)	(.015)	(.015)	(.014)	(.012)
Log average per capita education spending	-.018	-.080	-0.027	-0.264	-.014	-.202 *	.024	.169
	(.115)	(.102)	(0.221)	(0.203)	(.127)	(.121)	(.118)	(.113)
Age	.142 *	.154 **	0.341 **	0.323 **	.099	.180 **	.133 *	.191 ***
	(.074)	(.067)	(0.144)	(0.139)	(.078)	(.077)	(.077)	(.073)
Age squared	-.002 *	-.002 **	-0.004 **	-0.004 **	-.001	-.002 **	-.002	-.002 **
	(.001)	(.001)	(0.002)	(0.002)	(.001)	(.001)	(.001)	(.001)
Black	.070 ***	.031 **	0.099 ***	0.056 **	.071 ***	.036 ***	.082 ***	.041 ***
	(.013)	(.012)	(0.025)	(0.024)	(.014)	(.014)	(.014)	(.014)
Hispanic	.010	.005	-0.034	0.031	-.009	.015	-.024	-.009
	(.014)	(.015)	(0.032)	(0.031)	(.017)	(.017)	(.018)	(.017)
Other race	-.022	-.305 ***	0.037	-0.019	.007	-.046 **	-.002	-.045 **
	(.020)	(.024)	(0.047)	(0.047)	(.021)	(.023)	(.025)	(.022)
Has health limitation	-.088 ***	-.059 ***	-0.035	-0.090 ***	-.076 ***	-.062 ***	-.071 ***	-.068 ***
	(.019)	(.014)	(0.027)	(0.025)	(.018)	(.014)	(.019)	(.016)
Schooling of parent	.009 ***	.012 ***	0.014 ***	0.015 ***	.007 ***	.014 ***	.010 ***	.013 ***
	(.002)	(.002)	(0.003)	(0.003)	(.002)	(.002)	(.002)	(.002)
Observed schooling is mother's	.015	.029 ***	0.038 *	0.045 **	.006	.037 ***	.017 *	.029 ***
	(.010)	(.009)	(0.020)	(0.018)	(.010)	(.010)	(.010)	(.010)
Magazines in home when age 14	.051 ***	.042 ***	0.078 ***	0.087 ***	.056 ***	.050 ***	.062 ***	.048 ***
	(.010)	(.010)	(0.019)	(0.019)	(.011)	(.011)	(.011)	(.011)
AFQT percentile score	.002 ***	.001 ***	0.004 ***	0.003 ***	.002 ***	.002 ***	.003 ***	.002 ***
	(.000)	(.000)	(0.000)	(0.000)	(.000)	(.000)	(.000)	(.000)
Rotter index	-.006 **	-.004	-0.010 *	0.001	-.004	-.004	-.006 *	-.001
	(.003)	(.003)	(0.006)	(0.006)	(.003)	(.003)	(.003)	(.003)
Average state anti-smoking sentiment	.088	.048	0.222 *	0.235 **	.136 **	.053	.134 **	.073
	(.055)	(.056)	(0.120)	(0.118)	(.063)	(.069)	(.059)	(.066)
Average state cigarette tax	-.121	.110	-0.098	-0.179	-.136	.105	-.116	.083
	(.124)	(.124)	(0.253)	(0.261)	(.139)	(.148)	(.134)	(.140)
Intercept	-1.689	-1.403	-5.551	-3.316	-.943	-.934	-1.987	-4.419 **
	(1.795)	(1.588)	(3.436)	(3.238)	(1.924)	(1.840)	(1.856)	(1.738)
N	3174	3157	1496	1491	3248	3275	3248	3274
R^2	.936	.949	.864	.882	.922	.924	.921	.934
F(k,N-k)	114.36	62.26	74.40	49.21	132.4	84.1	128.25	58.42
k	5	5	5	5	5	5	5	5
N-k	3108	3901	1434	1427	3182	3209	3182	3208

Notes: Standard errors (in parenthesis) are adjusted for heteroskedasticity. Coefficients denoted by ***, **, and * are statistically different from zero with p-values of <.01, <.05, and <.10 respectively. All models include state fixed effects.

Appendix Table 1 First stage equations

	GED							
	Current smoking status		Ex smoker status		Overweight		Obese	
Variable	Men	Women	Men	Women	Men	Women	Men	Women
High school course graduation requirements	.008	.001	0.011	0.015	.007	.009	.009	.008
	(.009)	(.009)	(0.017)	(0.023)	(.010)	(.012)	(.010)	(.011)
Graduation requirements set by local districts	.105	-.010	0.220	0.217	.092	.083	.128	.052
	(.163)	(.168)	(0.310)	(0.401)	(.186)	(.223)	(.175)	(.202)
Fraction of GED test takers of similar age	.136 ***	.139 ***	0.204 ***	0.296 ***	.133 ***	.194 ***	.149 ***	.206 ***
	(.010)	(.010)	(0.018)	(0.018)	(.010)	(.011)	(.010)	(.012)
First factor GED policies	.010	.018	-0.028	0.061 **	.010	.023	-.001	.037 **
	(.017)	(.017)	(0.025)	(0.027)	(.018)	(.019)	(.017)	(.015)
Log average per capita education spending	-.287 **	-.029	-0.304	-0.106	-.356 **	-.062	-.237 *	-.150
	(.140)	(.123)	(0.252)	(0.233)	(.157)	(.151)	(.139)	(.139)
Age	-.031	-.002	0.004	0.231	.080	.069	.079	.094
	(.091)	(.080)	(0.165)	(0.160)	(.096)	(.096)	(.092)	(.090)
Age squared	.000	.000	0.000	-0.003	-.001	-.001	-.001	-.001
	(.001)	(.001)	(0.002)	(0.002)	(.001)	(.001)	(.001)	(.001)
Black	.022	-.060 ***	0.063 **	-0.050 *	.048 ***	-.057 ***	.042 **	-.070 ***
	(.016)	(.015)	(0.028)	(0.028)	(.017)	(.017)	(.017)	(.017)
Hispanic	-.035 **	-.034 *	-0.051	-0.041	.006	.001	.012	-.015
	(.018)	(.018)	(0.037)	(0.036)	(.021)	(.022)	(.021)	(.020)
Other race	.016	.112 ***	0.040	-0.015	.000	-.039	.017	-.027
	(.024)	(.029)	(0.054)	(0.053)	(.026)	(.029)	(.030)	(.027)
Has health limitation	.083 ***	.033 **	0.069 **	0.020	.089 ***	.022	.077 ***	.010
	(.023)	(.017)	(0.031)	(0.028)	(.023)	(.017)	(.022)	(.020)
Schooling of parent	-.003	-.002	-0.007 *	-0.006	-.001	-.003	-.005 **	-.003
	(.002)	(.002)	(0.004)	(0.004)	(.002)	(.002)	(.002)	(.002)
Observed schooling is mother's	.000	-.015	0.041 *	-0.033	.010	-.017	.004	-.008
	(.012)	(.011)	(0.022)	(0.021)	(.013)	(.013)	(.012)	(.012)
Magazines in home when age 14	-.024 *	-.047 ***	0.015	-0.047 **	-.024 *	-.040 ***	-.016	-.043 ***
	(.013)	(.012)	(0.022)	(0.022)	(.013)	(.013)	(.013)	(.013)
AFQT percentile score	-.001 ***	-.001 ***	-0.001	0.000	-.001 ***	-.001 ***	-.001 ***	-.001 ***
	(.000)	(.000)	(0.001)	(0.001)	(.000)	(.000)	(.000)	(.000)
Rotter index	.002	.003	-0.003	0.004	-.002	-.001	-.002	.000
	(.004)	(.003)	(0.007)	(0.007)	(.004)	(.004)	(.004)	(.004)
Average state anti-smoking sentiment	-.148 **	-.131 **	-0.087	-0.273 **	-.158 **	-.187 **	-.126 *	-.164 **
	(.067)	(.067)	(0.137)	(0.135)	(.078)	(.086)	(.070)	(.081)
Average state cigarette tax	.292 *	.202	0.386	0.487	.294 *	.278	.266 *	.354 **
	(.152)	(.148)	(0.289)	(0.299)	(.172)	(.185)	(.159)	(.173)
Intercept	3.040	.490	2.633	-3.448	1.657	-.645	.562	-.364
	(2.191)	(1.908)	(3.930)	(3.710)	(2.373)	(2.294)	(2.197)	(2.145)
N	3174	3157	1496	1491	3248	3275	3248	3274
R^2	.233	.201	.341	.369	.242	.261	.249	.266
F(k,N-k)	37.70	37.61	27.64	52.79	35.5	57.2	42.57	63.89
k	5	5	5	5	5	5	5	5
N-k	3108	3901	1434	1427	3182	3209	3182	3208

Notes: Standard errors (in parenthesis) are adjusted for heteroskedasticity. Coefficients denoted by ***, **, and * are statistically different from zero with p-values of <.01, <.05, and <.10 respectively. All models include state fixed effects.

Appendix Table 2 Current smoker

Variable	Men 1	2	3	4	5	Women 1	2	3	4	5
High school graduate	-.226***	-.258***	-.209***	-.167*	-.139*	-.213***	-.243***	-.219***	-.103	.062
	(.027)	(.027)	(.029)	(.088)	(.080)	(.029)	(.030)	(.030)	(.130)	(.095)
GED		.156***	.125***	.203	.152		.227***	.210***	.265*	.090
		(.024)	(.026)	(.126)	(.112)		(.025)	(.024)	(.139)	(.108)
Age	-.032	-.122	.081	.015	.029	-.002	-.097	.018	.017	.017
	(.126)	(.123)	(.125)	(.124)	(.123)	(.122)	(.105)	(.107)	(.111)	(.115)
Age squared	.000	.002	-.001	.000	.000	.000	.001	.000	.000	.000
	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)	(.001)	(.001)	(.001)	(.002)
Black	.021	.008	-.019	-.035	-.041*	-.034*	-.031	-.138***	-.145***	-.163***
	(.021)	(.020)	(.023)	(.023)	(.023)	(.020)	(.019)	(.021)	(.024)	(.023)
Hispanic	-.065***	-.053**	-.046*	-.059**	-.070***	-.081***	-.080***	-.145***	-.158***	-.164***
	(.024)	(.023)	(.024)	(.025)	(.023)	(.024)	(.021)	(.023)	(.025)	(.024)
Other race	.006	-.029	.046	.015	.021	-.063*	-.012	-.133***	-.118*	-.049
	(.029)	(.034)	(.031)	(.034)	(.034)	(.036)	(.040)	(.032)	(.067)	(.057)
Has health limitation	.119***	.097***	.074**	.080**	.090***	.035	.022	-.002	.008	.019
	(.032)	(.031)	(.032)	(.036)	(.034)	(.024)	(.024)	(.023)	(.026)	(.024)
Schooling of parent			-.002	.001				.008***	.007**	
			(.003)	(.003)				(.002)	(.003)	
Observed schooling is mother's			.020	.008				-.009	-.006	
			(.016)	(.016)				(.015)	(.016)	
Magazines in home when age 14			.031*	.015	.012			-.027*	-.027	-.034*
			(.018)	(.019)	(.019)			(.016)	(.020)	(.020)
AFQT percentile score			-.002***	-.002***	-.002***			-.004***	-.004***	-.004***
			(.000)	(.001)	(.001)			(.000)	(.000)	(.000)
Rotter index			.016***	.013***	.013***			.002	.003	.006
			(.005)	(.005)	(.005)			(.004)	(.005)	(.005)
State anti-smoking sentiment	-.387***	-.342***	-.346***	-.330***	-.331***	-.260**	-.422***	-.264***	-.218**	-.243**
	(.098)	(.096)	(.094)	(.095)	(.094)	(.107)	(.097)	(.091)	(.096)	(.095)
Average cigarette tax	.303	.233	.355*	.249	.245	.028	.012	.178	.207	.222
	(.223)	(.218)	(.214)	(.214)	(.210)	(.242)	(.236)	(.207)	(.206)	(.211)
Intercept	.899	2.575	-1.298	-.072	-.301	.467	2.100	.125	.015	.001
	(2.314)	(2.262)	(2.303)	(2.282)	(2.271)	(2.252)	(1.943)	(1.982)	(2.045)	(2.115)
N	3174	3174	3174	3174	3174	3157	3157	3157	3157	3157
R^2	.292	.299	.311	---	---	.282	.291	.299	---	---
F (k, n-k) on 1st stage IVs										
High school graduate				114.4	88.8				62.3	58.5
GED				37.7	29.5				37.6	32.1

Notes: Columns (1) - (3) contain results from linear probability models. Columns (4) and (5) contains results from two-stage least squares regressions. Standard errors (adjusted for heteroskedasticity) appear in parentheses. Coefficients denoted by ***, **, and * are statistically different from zero with p-values of <.01, <.05, and <.10 respectively. All models include state fixed effects.

Appendix Table 3 Former smoker											
	Men					Women					
Variable	1	2	3	4	5	1	2	3	4	5	
High school graduate	.119 ***	.142 ***	.073 **	.063	.063	.190 ***	.239 ***	.164 ***	.189	.203 *	
	(.032)	(.033)	(.036)	(.117)	(.101)	(.035)	(.036)	(.038)	(.130)	(.108)	
GED		-.074 **	-.042	-.033	-.030		-.178 ***	-.134 ***	-.090	-.106	
		(.033)	(.034)	(.167)	(.148)		(.032)	(.032)	(.109)	(.101)	
Age	.177	.191	.225	.231	.231	.008	.083	.191	.165	.181	
	(.207)	(.206)	(.204)	(.206)	(.206)	(.206)	(.204)	(.201)	(.203)	(.201)	
Age squared	-.002	-.002	-.003	-.003	-.003	.000	-.001	-.002	-.002	-.002	
	(.003)	(.003)	(.003)	(.003)	(.003)	(.003)	(.003)	(.003)	(.003)	(.003)	
Black	-.026	-.015	.060 *	.061 *	.061 *	-.053 *	-.044	.057	.056	.060	
	(.032)	(.032)	(.036)	(.037)	(.037)	(.032)	(.032)	(.035)	(.038)	(.037)	
Hispanic	.088 **	.091 **	.113 **	.113 **	.113 **	.055	.067	.117 ***	.117 **	.118 ***	
	(.045)	(.045)	(.046)	(.046)	(.045)	(.045)	(.044)	(.045)	(.046)	(.044)	
Other race	.113 *	.120 *	.132 **	.131 *	.131 *	-.003	.002	.041	.043	.037	
	(.067)	(.067)	(.067)	(.067)	(.067)	(.069)	(.068)	(.068)	(.068)	(.067)	
Has health limitation	-.042	-.038	-.026	-.029	-.029	-.042	-.031	-.005	-.002	-.002	
	(.039)	(.039)	(.039)	(.042)	(.041)	(.037)	(.036)	(.036)	(.039)	(.038)	
Schooling of parent			.000	.000				.002	.002		
			(.005)	(.006)				(.005)	(.006)		
Observed schooling is mother's			.005	.005				.003	.003		
			(.028)	(.028)				(.027)	(.028)		
Magazines in home when age 14			-.020	-.019	-.019			.065 **	.064 *	.068 **	
			(.028)	(.028)	(.028)			(.028)	(.033)	(.033)	
AFQT percentile score			.003 ***	.003 ***	.003 ***			.004 ***	.004 ***	.004 ***	
			(.001)	(.001)	(.001)			(.001)	(.001)	(.001)	
Rotter index			-.016 **	-.016 **	-.016 **			-.003	-.004	-.004	
			(.008)	(.008)	(.008)			(.009)	(.009)	(.009)	
State anti-smoking sentiment	.387 **	.374 **	.352 **	.355 **	.355 **	.412 **	.349 **	.318 *	.328 *	.314 *	
	(.171)	(.170)	(.169)	(.172)	(.171)	(.174)	(.172)	(.170)	(.177)	(.176)	
Average cigarette tax	-.387	-.352	-.346	-.348	-.349	-.405	-.257	-.352	-.383	-.343	
	(.363)	(.362)	(.360)	(.367)	(.365)	(.392)	(.384)	(.376)	(.387)	(.383)	
Intercept	-3.002	-3.253	-3.871	-3.980	-3.971	-.094	-1.473	-3.602	-3.149	-3.431	
	(3.824)	(3.819)	(3.778)	(3.802)	(3.801)	(3.795)	(3.764)	(3.719)	(3.741)	(3.709)	
N	1496	1496	1496	1496	1496	1491	1491	1491	1491	1491	
R^2	0.441	0.443	0.453	---	---	0.469	0.478	0.496	---	---	
F (k, n-k) on 1st stage IVs											
High school graduate				74.4	58.1				49.2	42.6	
GED				27.6	20.4				52.8	39.7	

Notes: Columns (1) - (3) contain results from linear probability models. Columns (4) and (5) contains results from two-stage least squares regressions. Standard errors (adjusted for heteroskedasticity) appear in parentheses. Coefficients denoted by ***, **, and * are statistically different from zero with p-values of <.01, <.05, and <.10 respectively. All models include state fixed effects.

Appendix Table 4 Overweight																				
	Men										Women									
Variable	1		2		3		4		5		1		2		3		4		5	
High school graduate	.068	***	.067	***	.087	***	.110		.045		-.043		-.046		-.007		.098		-.119	
	(.025)		(.026)		(.027)		(.083)		(.076)		(.028)		(.029)		(.030)		(.120)		(.087)	
GED			.005		-.005		-.105		-.006				.016		-.006		-.018		.142	
			(.022)		(.023)		(.129)		(.105)				(.025)		(.025)		(.117)		(.093)	
Age	-.037		-.038		-.014		.014		-.031		.290	**	.285	**	.292	**	.285	**	.189	
	(.127)		(.127)		(.127)		(.129)		(.128)		(.137)		(.137)		(.137)		(.138)		(.140)	
Age squared	.001		.001		.000		.000		.000		-.004	**	-.004	**	-.004	**	-.004	**	-.003	
	(.002)		(.002)		(.002)		(.002)		(.002)		(.002)		(.002)		(.002)		(.002)		(.002)	
Black	.047	**	.047	**	.023		.026		.024		.249	***	.248	***	.220	***	.213	***	.205	***
	(.020)		(.020)		(.023)		(.024)		(.023)		(.021)		(.021)		(.024)		(.026)		(.024)	
Hispanic	.086	***	.086	***	.064	**	.063	**	.079	***	.163	***	.161	***	.107	***	.103	***	.120	***
	(.025)		(.026)		(.027)		(.028)		(.027)		(.029)		(.029)		(.031)		(.032)		(.031)	
Other race	.029		.028		.010		-.004		.020		.011		.011		-.005		-.001		.003	
	(.037)		(.037)		(.036)		(.036)		(.037)		(.041)		(.041)		(.041)		(.042)		(.042)	
Has health limitation	-.067	**	-.068	**	-.074	**	-.061	*	-.078	**	.104	***	.103	***	.091	***	.099	***	.075	**
	(.030)		(.031)		(.031)		(.035)		(.033)		(.025)		(.025)		(.025)		(.027)		(.030)	
Schooling of parent					-.005	*	-.005	*							-.011	***	-.012	***		
					(.003)		(.003)								(.003)		(.004)			
Observed schooling is mother's					-.002		.004								-.031	*	-.036	*		
					(.017)		(.017)								(.018)		(.020)			
Magazines in home when age 14					.029		.027		.023						-.024		-.031		-.018	
					(.018)		(.019)		(.019)						(.020)		(.022)		(.023)	
AFQT percentile score					-.001	**	-.001	**	-.001	**					-.001		-.001		.000	
					(.000)		(.000)		(.001)						(.000)		(.001)		(.001)	
Rotter index					.005		.003		.005						.003		.004		.003	
					(.005)		(.005)		(.005)						(.006)		(.006)		(.006)	
State anti-smoking sentiment	.055		.056		.077		.065		.081		.047		.054		.084		.076		.158	
	(.104)		(.104)		(.104)		(.109)		(.106)		(.124)		(.124)		(.124)		(.127)		(.130)	
Average cigarette tax	.342		.340		.322		.346		.326		-.400		-.410		-.396		-.402		-.492	*
	(.227)		(.227)		(.229)		(.235)		(.231)		(.266)		(.266)		(.266)		(.267)		(.270)	
Intercept	1.148		1.172		.805		.306		1.086		-4.817	*	-4.722	*	-4.680	*	-4.625	*	-2.914	
	(2.346)		(2.348)		(2.340)		(2.386)		(2.368)		(2.537)		(2.538)		(2.543)		(2.555)		(2.584)	
N	3248		3248		3248		3248		3248		3275		3275		3275		3275		3275	
R^2	.733		.733		.737		---		---		.601		.600		.606		---		---	
F (k, n-k) on 1st stage IVs																				
High school graduate							132.4		103.3								84.1		96.1	
GED							35.5		33.6								57.2		54.9	

Notes: Columns (1) - (3) contain results from linear probability models. Columns (4) and (5) contains results from two-stage least squares regressions. Standard errors (adjusted for heteroskedasticity) appear in parentheses. Coefficients denoted by ***, **, and * are statistically different from zero with p-values of <.01, <.05, and <.10 respectively. All models include state fixed effects.

44

Appendix Table 5 Obese																				
					Men										Women					
Variable	1		2		3		4		5		1		2		3		4		5	
High school graduate	-.017		-.014		.013		-.008		-.075		-.036		-.039		.005		-.022		-.215	**
	(.024)		(.024)		(.026)		(.082)		(.078)		(.028)		(.028)		(.029)		(.139)		(.108)	
GED			-.014		-.028		.041		.130				.015		-.005		.043		.197	*
			(.022)		(.022)		(.121)		(.123)				(.023)		(.022)		(.108)		(.109)	
Age	.285	**	.289	**	.275	**	.255	**	.245	**	.177		.172		.146		.146		.097	
	(.120)		(.120)		(.120)		(.122)		(.123)		(.120)		(.120)		(.117)		(.117)		(.121)	
Age squared	-.004	**	-.004	**	-.004	**	-.003	**	-.003	*	-.002		-.002		-.002		-.002		-.001	
	(.002)		(.002)		(.002)		(.002)		(.002)		(.002)		(.002)		(.002)		(.002)		(.002)	
Black	.075	***	.077	***	.052	**	.049	**	.051	**	.180	***	.180	***	.153	***	.156	***	.157	***
	(.020)		(.020)		(.022)		(.023)		(.023)		(.021)		(.021)		(.023)		(.026)		(.024)	
Hispanic	.110	***	.111	***	.089	***	.084	***	.085	***	.087	***	.085	***	.030		.028		.051	*
	(.026)		(.026)		(.028)		(.028)		(.028)		(.025)		(.026)		(.027)		(.027)		(.027)	
Other race	-.025		-.025		-.024		-.026		-.032		-.012		-.011		-.024		-.028		-.006	
	(.040)		(.040)		(.040)		(.040)		(.040)		(.036)		(.036)		(.036)		(.036)		(.037)	
Has health limitation	.037		.039		.032		.025		.016		.110	***	.109	***	.103	***	.099	***	.083	***
	(.029)		(.029)		(.029)		(.033)		(.033)		(.026)		(.026)		(.026)		(.028)		(.029)	
Schooling of parent					-.006	**	-.005								-.014	***	-.013	***		
					(.003)		(.003)								(.003)		(.004)			
Observed schooling is mother's					.019		.018								-.028	*	-.027			
					(.016)		(.017)								(.016)		(.017)			
Magazines in home when age 14					.009		.009		.007						-.011		-.006		-.005	
					(.017)		(.019)		(.019)						(.018)		(.021)		(.020)	
AFQT percentile score					-.001	**	-.001		-.001						.000		.000		.000	
					(.000)		(.001)		(.000)						(.000)		(.001)		(.000)	
Rotter index					-.005		-.005		-.004						-.006		-.006		-.006	
					(.005)		(.005)		(.005)						(.005)		(.005)		(.005)	
State anti-smoking sentiment	-.306	***	-.305	***	-.305	***	-.303	***	-.282	***	-.007		-.002		.012		.026		.052	
	(.095)		(.095)		(.093)		(.096)		(.099)		(.108)		(.108)		(.106)		(.109)		(.111)	
Average cigarette tax	.401	*	.401	*	.377	*	.384	*	.383	*	-.417	*	-.425	*	-.365		-.376	*	-.421	*
	(.213)		(.213)		(.210)		(.215)		(.218)		(.235)		(.236)		(.227)		(.228)		(.238)	
Intercept	-5.207	**	-5.276	**	-4.929	**	-4.559	**	-4.398	*	-2.957		-2.866		-2.191		-2.186		-1.305	
	(2.217)		(2.219)		(2.211)		(2.246)		(2.274)		(2.210)		(2.211)		(2.167)		(2.164)		(2.243)	
N	3248		3248		3248		3248		3248		3275		3275		3275		3274		3275	
R²	.264		.264		.265		---		---		.291		.291		.294		---		---	
F (k, n-k) on 1st stage IVs																				
High school graduate							128.3		92.7								58.4		55.3	
GED							42.6		27.3								63.9		34.9	

Notes: Columns (1) - (3) contain results from linear probability models. Columns (4) and (5) contains results from two-stage least squares regressions. Standard errors (adjusted for heteroskedasticity) appear in parentheses. Coefficients denoted by ***, **, and * are statistically different from zero with p-values of <.01, <.05, and <.10 respectively. All models include state fixed effects.