

Does Education Induce Healthy Lifestyle?*

Cheolsung Park
Department of Economics
National University of Singapore
Singapore

Changhui Kang[†]
Department of Economics
Chung-Ang University
Seoul, South Korea

Journal of Health Economics, forthcoming

Abstract

We investigate whether education induces individuals to have healthy lifestyle. To test for causality we instrument education by high school availability and birth order using data of Korean men; to account for correlations among health behaviors, we estimate a quadivariate probit model using simulated MLE. Our results indicate that an increase in education induces individuals to exercise regularly, and to get regular health checkups. We find, however, that education has little effect on smoking or drinking. We also discover that unobserved determinants of health behaviors are correlated, especially between smoking and drinking, and between exercising and getting health checkups.

JEL Classification: I10, C30

Keywords : Education, Health, Quadivariate probit, South Korea

* We thank the editor and two anonymous referees for helpful comments on previous versions of this paper. We also thank participants in seminars at National University of Singapore and Korea University for suggestions and comments on earlier versions. We are solely responsible for any remaining errors. The research of Changhui Kang was supported by the Chung-Ang University Research Grants in 2008.

[†] Corresponding author, Address: Department of Economics, Chung-Ang University, 221 Heukseok-Dong Dongjak-Gu, Seoul 156-756, South Korea; E-mail: ckang@cau.ac.kr, Phone: +82-2-820-5862, Fax: +82-2-812-9718.

1 Introduction

There is overwhelming evidence that an individual's educational attainment is strongly correlated with most health-related behaviors. It has long been observed that, for example, college graduates are less likely to smoke, less likely to be obese, and more likely to exercise than non-college graduates (e.g., Kenkel 1991, de Walque 2007). There is, however, a longstanding debate on the nature of the relationship. Does education induce individuals to adopt the healthier lifestyle? Or do unobserved third factors give rise to the correlation which is, in fact, spurious? For example, lower discount rate on the future may induce individuals to attain more education and to invest more in health by adopting healthy lifestyle. Parents with healthy lifestyle, of better health, and with higher income associated with better health may induce their children to attain more education, by providing them with more resources and better health, and to adopt healthy lifestyle through influence on the children's lifestyle choices. There is also a possibility of reverse causality. Healthy lifestyle leads to better health, and better health may lead to higher educational attainment.

During the past decade much research has been done to answer these questions. The methodology of choice has been to employ instrumental variables (IVs) for education which can, if correctly applied, filter out any non-causal effect of education. Earlier studies have used family background variables, such as parental education, number of siblings, and location of home, as IVs for education (Sander, 1995a, 1995b; Leigh and Dhir, 1997). These studies are among the first to tackle directly endogeneity of education in health-related behavior equations, but their choice of IVs for education is questioned (Grossman 2005). Most notably, parental education is likely to be correlated to an individual's health-related behaviors through, for example, informal interactions between parents and children.

More recent studies use institutional characteristics that are likely to affect individuals' educational attainment as IVs for education. Exclusion of such variables from health-related behavior equations is easier to defend than exclusion of family background characteristics. For example, Currie and Moretti (2003) use availability of colleges in the county of residence at age 17 as the IV for women's education to estimate its effect on women's choices of prenatal care and smoking during pregnancy, among others. They find evidence that education increases the former and decreases the latter. de Walque (2007) uses the risk of being drafted to the Vietnam

War as the IV for men’s college attendance to estimate its effects on smoking cessation and current smoking behavior. He finds that the estimated effect of college education on the current smoking behavior is negative and statistically significant only when Vietnam War veteran status—which should be endogenous according to the assumption of the paper—is included in the regression equation. Without the veteran status, the education coefficient is either positive or statistically insignificant. Furthermore, he finds little effect of education on smoking cessation. Kenkel et al. (2006) use state-wide educational policy variables as IVs for high school completion and general equivalency diploma (GED) receipt to estimate their effects on smoking and obesity. Most of their two-stage estimates of the education coefficients are smaller in magnitude than the OLS estimates and estimated imprecisely with large standard errors.

The aim of this paper is also to examine whether education has causal effects on health-related behaviors. We adopt the estimation strategy of recent studies, exploiting variations in policy to address endogeneity of education. The main difference of this research from previous research is that we examine choices of not one but multiple health-related behaviors, modeling them as correlated multivariate choices. The previous research either studies only one health-related behavior or assumes that one health-related behavior is adopted independently of another. The assumption is, however, questionable. If unobserved characteristics of individuals, such as the subjective discount rate or taste, are important determinants of health behaviors, an individual’s choices of health behaviors are likely to be correlated. If the correlations are indeed significant, restricting them to zeros may lead to erroneous inferences on the relationship between education and health behavior choices.

In this paper, we model and estimate jointly the determinants of four common health-related behaviors—receiving regular health checkups, doing exercise regularly, smoking and drinking,—while we also deal with potential endogeneity of education in the health-related behavior equations using IVs. Some studies have tried to estimate health-related behavior equations jointly as inputs to health production process (e.g. Contoyannis and Jones 2004), but they have treated education as exogenously given. In this paper we address both correlations among health-related behavioral choices and endogeneity of education.

We investigate effects of education on adult men’s health-related behaviors using data from South Korea. South Korea experienced a dramatic expansion of high school education during

the mid-1970s, which offers a good source of variation in an individual’s completed education. In addition, to our knowledge, this is the first study that investigates causality of the relationship between education and health behaviors with data from a country outside North America and Europe.

Our IV estimation results indicate that increased education induces individuals to exercise regularly. Education also induces individuals to get regular health checkups, but the effect of education is mainly through its effects on job characteristics. On smoking and drinking behavior, the effect of education is statistically insignificant in estimations using IVs. We also discover that unobserved determinants of health behaviors are correlated, especially between smoking and drinking, and between exercising and getting health checkups. In this paper ignoring such correlations leads to overestimation of education coefficients in IV estimations.

The balance of the paper is organized as follows. In section 2 we discuss an economic model of health behavior choices, our empirical model and the estimation method. In section 3 we explain the historical background of changes in high school availability in South Korea that we employ as an IV for education. In section 4 we describe the data and in section 5 we present the estimation results. We summarize and conclude the paper in section 6.

2 Model and Estimation Method

2.1 Economic Model of Health Behaviors and Estimation Issues

First let us consider a simple model of health-related behavioral choices that gives rise to a generic empirical model. More theoretical discussions on the model of health demand are found in, for example, Grossman (1972) and Rosenzweig and Schultz (1983). Think of an individual who derives utility from health, health-related behaviors, and consumption. The individual’s utility at time t is determined by the following function:

$$V_t = U(H_t, \mathbf{B}_t, C_t | \alpha_0, \mathbf{X}_t), \quad (1)$$

where H_t is health status, \mathbf{B}_t is a k -dimensional vector of health-related behavior choices that affect health status as well as utility, C_t is consumption, and α_0 and \mathbf{X}_t are time-invariant and time-variant utility shifters respectively. α_0 is unobserved. It is assumed to be formed

at the start of adulthood (time 0) after the individual finishes formal schooling. Knowledge of effects of health-related behaviors on health, and family background (e.g., parents' health-related behaviors), among others, are likely to determine α_0 . Vector \mathbf{X}_t includes the individual's observed personal characteristics.

We assume that an individual's health is produced as follows:

$$H_t = h(H_{t-1}, \mathbf{B}_t, \boldsymbol{\mu}_t), \quad (2)$$

where $\boldsymbol{\mu}_t$ is the vector of individual's characteristics that affect health in period t . It includes, for example, the individual's age, job characteristics, family characteristics, and so forth. The budget constraint is given by

$$C_t + A_{t+1} = y_t + (1 + r)A_t - \mathbf{p}'_t \mathbf{B}_t, \quad (3)$$

where y_t is income, A_t is asset holdings at time t , r is the rate of return on asset holdings, and \mathbf{p}_t is a k -dimensional vector of prices associated with health-related behaviors. The individual maximizes his or her lifetime utility, discounting the future by rate ρ . From this model, we can derive the following demand function for health-related behaviors:

$$\mathbf{B}_t^* = f(y_0, \dots, y_T, \mathbf{p}'_0, \dots, \mathbf{p}'_T, \mathbf{X}_0, \dots, \mathbf{X}_T, \boldsymbol{\mu}_0, \dots, \boldsymbol{\mu}_T, H_0, A_0, \rho, \alpha_0). \quad (4)$$

Equation (4) points to multiple channels through which the individual's education may be correlated, causally or non-causally, with choices of health-related behaviors. Those channels may be found in the taste side as well as in the health production technology side, since the demand function (4) reflects both taste and technology (Pollak and Wachter 1975, Rosenzweig and Schultz 1983).

Controlling for observable characteristics such as income and variables in \mathbf{X}_t , education will have causal effect on health-related behaviors, if education changes the individual's time preference ρ (Becker and Mulligan 1997), the utility shifter α_0 by augmenting the individual's knowledge on the effect of health-related behaviors on health, or $\boldsymbol{\mu}_t$ by influencing, for example, occupational choices or job characteristics. Writing explicitly the effect of education on those

variables, we can rewrite (4) into

$$\mathbf{B}_t^* = f [y_0, \dots, y_T, \mathbf{p}'_0, \dots, \mathbf{p}'_T, \mathbf{X}_0, \dots, \mathbf{X}_T, \boldsymbol{\mu}_0(s), \dots, \boldsymbol{\mu}_T(s), H_0, A_0, \rho(s), \alpha_0(s)], \quad (4')$$

where s is the years of education. Even if there is no causal effect, however, education is still likely to be correlated with health-related behavior choices, since education is correlated with H_0 , A_0 and the immutable part of time preference ρ non-causally. For example, health and wealth at the beginning of adulthood, H_0 and A_0 , are likely to be correlated with health during childhood and parental income and wealth which, in turn, affect the years of education.

From equation (4') the following health-related behavior equations for individual i can be derived:

$$\begin{aligned} b_{1it}^* &= \gamma_{11}s_i + \gamma'_{21}\mathbf{X}_{it} + \gamma'_{31}\mathbf{M}_{it} + \delta_{1i} + \varepsilon_{1it} \\ &\quad \vdots \\ b_{kit}^* &= \gamma_{1k}s_i + \gamma'_{2k}\mathbf{X}_{it} + \gamma'_{3k}\mathbf{M}_{it} + \delta_{ki} + \varepsilon_{kit}, \end{aligned} \quad (5)$$

where b_{jit}^* is i 's choice of health-related behavior j ($j = 1, \dots, k$); s_i is the years of education; \mathbf{X}_{it} is the vector of personal characteristics that affect preference shifters; \mathbf{M}_{it} is the vector of exogenous variables that determine i 's financial resources available for health choices; δ_{ji} is the unobserved heterogeneity which affects choice j derived from the composite of H_{0i} , A_{0i} , ρ_i , and α_{0i} ; and ε_{jit} is the random shock that affects choice j .¹

In estimating the set of equations (5) we have to deal with two issues— s_i is likely to be correlated with $\delta_{1i}, \dots, \delta_{ki}$, and choices $1, \dots, k$ can be inter-related. The correlation among choices is likely due to the correlation among $\delta_{1i}, \dots, \delta_{ki}$, but it is also possible that the shocks $\varepsilon_{1it}, \dots, \varepsilon_{kit}$ are correlated. While correlation between education and unobserved heterogeneity—endogeneity of education—can be dealt with using IVs for education, inter-related decisions require us to use a multiple-equation estimation method.

2.2 Empirical Model Specification and Estimation Method

In this paper we assume that individuals make joint binary decisions on each of four health-related behaviors—whether to receive regular health checkups (b_1), whether to exercise regularly

(b_2), whether to abstain from smoking (b_3) and whether to drink moderately as opposed to more than 5 times a month (b_4). We estimate the coefficients of the set of equations (5) at only one point of time due to data availability. For notational brevity, the time index t is dropped. We use a set of IVs, \mathbf{Z}_i for individual i , to deal with endogeneity of education.

We exploit two different sources for \mathbf{Z}_i that are unique to South Korea. The first variable is the number of high school classrooms per 1,000 middle school graduates disaggregated by the region and year of the individual's middle school graduation. As explained later in section 3 in greater detail, high school education expanded rapidly in the mid-1970s in Korea. We exploit such changes in the 1970s that yielded temporal as well as regional differentials in high school availability as a source of exogenous variation in an individual's completed education.

The second variable reflects traditional son preference of South Korean parents. Studies report that Korean parents favor sons, especially the first son, relative to daughters and later-born sons along various dimensions (Park and Cho, 1995). Education may be a dimension through which son preference is revealed: it is likely that Korean parents invest more time and financial resources in educating the first son than other siblings. Our second source of identification is whether or not the individual is the first or only son in the family. Since the variable considers birth order of sons alone, ignoring the presence of female siblings in the family, the first son is not necessarily the first child².

To estimate the parameters of five equations (four behaviors and education) and correlation coefficients together, we make a distributional assumption and use the simulated maximum likelihood (SML) estimation method. The model we estimate for individual i , $i = 1, \dots, N$, can be written as follows using notations from equation (5):

$$\begin{aligned}
 b_{1i}^* &= \gamma_{11}s_i + \gamma'_{21}\mathbf{X}_i + \gamma'_{31}\mathbf{M}_i + e_{1i} \\
 &\vdots \qquad \qquad \qquad \vdots \\
 b_{4i}^* &= \gamma_{14}s_i + \gamma'_{24}\mathbf{X}_i + \gamma'_{34}\mathbf{M}_i + e_{4i} \\
 s_i &= \zeta'_1\mathbf{Z}_i + \zeta'_2\mathbf{X}_i + \zeta'_3\mathbf{M}_i + u_i,
 \end{aligned} \tag{6}$$

where

$$e_{ji} = \delta_{ji} + \varepsilon_{ji}, \quad j = 1, \dots, 4,$$

$$(e_{1i}, e_{2i}, e_{3i}, e_{4i}, u_i) \sim N(0, \mathbf{\Sigma}).$$

b_{ji}^* is i 's desired demand level for health behavior j ($j = 1, \dots, 4$). We assume that $b_{ji} = I(b_{ji}^* > 0)$, where $I(\cdot)$ is the index function, and that the diagonal elements of $\mathbf{\Sigma}$ is 1 except for the last, since they are not identified separately from other parameters.

The model we estimate is a multivariate IV probit model. The estimation method is an extension of the full-information maximum likelihood estimation method of the univariate IV probit model (Wooldridge 2002, pp. 472-477). We maximize the following log likelihood function:

$$\log L = \sum_{i=1}^N \{ \log \Phi_4 [(2b_{1i} - 1) (\gamma_1 s_i + \gamma'_{21} \mathbf{X}_i + \gamma'_{31} \mathbf{M}_i), \dots,$$

$$(2b_{4i} - 1) (\gamma_4 s_i + \gamma'_{24} \mathbf{X}_i + \gamma'_{34} \mathbf{M}_i) | s_i, \mathbf{\Sigma}]$$

$$- 0.5 \log 2\pi - \log \sigma_u - (s_i - \zeta'_1 \mathbf{Z}_i - \zeta'_2 \mathbf{X}_i - \zeta'_3 \mathbf{M}_i)^2 / (2\sigma_u^2) \} \quad (7)$$

where $\Phi_4[\cdot | s_i, \mathbf{\Sigma}]$ is the quadivariate normal distribution function whose mean and variance are derived conditional on observed education level and the original variance-covariance matrix $\mathbf{\Sigma}$. The main difficulty of evaluating the likelihood function (7) lies in computation of the quadivariate joint normal probability. In this paper we use simulated probability using Geweke-Hajivassiliou-Keane (GHK) simulator. The GHK simulator is shown to be one of the best-performing simulators of multivariate normal probabilities developed so far (Hajivassiliou, McFadden, and Ruud 1996). The number of simulations is 100.

We control for the same set of explanatory variables for all the four health-related behaviors. \mathbf{X}_i includes i 's age, age squared, marital status, parental education level, and dummies for current residential region, residential area type at age 14. Initial health status upon the completion of education (H_0) is controlled for by a dummy variable indicating whether i has a physical limitation or chronic illness before 25 years of age. \mathbf{X}_i also contains a dummy variable for whether he served in military, since previous research shows serving in military affects some

health-related behaviors such as smoking (de Walque 2007; Bedard and Deschênes 2006). As determinants of financial resources available for health-related behaviors, \mathbf{M}_i includes the number of household members aged 0 to 6, 7 to 24, 25 to 55, and 56 and above, total monthly net household income (in a natural log) and broadly-defined occupations. By controlling for the number of household members of varying ages, we attempt to capture an individual's different budget constraints in different family structures. For example, if there are many children who consume a large amount of educational funds in the family, the household head will have a limited resource to invest in healthy behaviors.³

In an augmented specification, in addition to the explanatory variables listed above, we control for job characteristics by including wage-earner/self-employed dummy, union membership dummy, large/small firm dummy, and public/private sector dummy. Those characteristics are strongly correlated with education, and likely to affect health-related behaviors. For example, employees of large firms or union members may enjoy fringe benefits, such as more comprehensive health insurance and health checkups, that can affect their health behaviors.

3 Historical Changes of High School Availability in South Korea

During the late 1970s, South Korea saw a sudden and rapid expansion of high school education. For example, while high school enrollment rate increased by 11 percentage points from 1970 to 1975, it jumped up by 18 percentage points from 1975 to 1980 and by 15 percentage points from 1980 to 1985. The middle to high school advancement rate (= number of high school freshmen / number of middle school graduates) that had increased by 5 percentage points from 1970 to 1975 went up by 10 percentage points from 1975 to 1980.

The expansion owes to the policy implemented in 1974, known as the “leveling policy,” which changed the high school admission and allocation system in cities. It replaced fiercely competitive school-level entrance exams for general high school admissions with municipality-level standard exams in most major cities, starting from the 1974 academic year (March 1974 to February 1975) (Chung 1998, Kang et al. 2007). As any middle school graduate that satisfied the minimum standard was qualified for a general high school, there has been a big surge in the number of potential high school freshmen from 1974 onwards.

[Insert Figure 1 here.]

To accommodate the swelling number of new students, the government rushed to open new high schools, built more classrooms, and converted vocational high schools into general high schools. Between 1974 and 1975 the number of general high schools in the nation jumped by 36 percent and the number of general high school freshmen by 42 percent. Figure 1 shows such a dramatic increase in high school availability in terms of the number of high school classrooms per 1,000 middle school graduates. At the national level, the number of general high school classrooms per 1,000 middle school graduates gradually declined from 23.2 in 1965 to 14.2 in 1973. Then in 1974 and 1975 there were big jumps. The number reached 16.7 in 1974 and 19.4 in 1975.

If we disaggregate the number of classrooms by province, most provinces except Seoul, the capital city, experienced similar declines in the number of general high school classrooms between 1965 and 1973, but witnessed the dramatic increase between 1974 and 1975. The size of the increase ranged from 20 to 70 percent relative to pre-1974 figures.⁴ In contrast to rapid changes in 1974 and 1975, the national average number of general high school classrooms remained stable between 20 and 23 after 1975. The regional averages showed temporal patterns similar to the national average between 1975 and 1985. Unlike general high schools, vocational high schools did not go through such rapid changes in availability. An exception is the period between 1974 and 1975 when some vocational high schools were converted into general high schools.

It is notable that there was a sharp decrease in high school availability in early 1980s. It was due to increases of the cohort size caused by a baby boom. The government started collecting 'education taxes' in 1982 and used the tax revenue to increase investment in public education. As the result, the situation improved over time as Figure 1 shows.

The dramatic change in the number of general high schools and classrooms, along with the change to leveling policy in admission system, that happened in 1974 and subsequent changes in school availability suggest that across cohorts education opportunities differ. Changes in high school education opportunity are likely to affect not only high school education choices but also other types of education choices, as students (or their parents) make forward-looking decisions and because a high school diploma is needed to advance to college. Furthermore, the sudden nature of the change of 1974 suggests that the change is unlikely to be correlated strongly with

a shift in taste for education.

Therefore, we use high school availability by cohort, especially that of general high schools before and after the mid 1970s, as an exogenous source of variation in an individual’s completed education. Together with the temporal variations, we also exploit the regional variations in high school availability by provinces. In the subsequent analysis, as the measure of high school availability we use the number—disaggregated at the province level—of general and vocational high school classrooms per 1,000 middle school graduates by cohort. In some specifications we use the total sum, and in others we use two variables, one for general and the other for vocational high schools.

We restrict our sample to those who graduated from middle school between 1965 and 1985. We do not consider the period before 1965, because relevant educational statistics do not exist before 1965. We do not use the younger cohorts that entered high school after 1985, because unobserved heterogeneity across cohorts may increase with age differences. There is a trade-off between the sample size and exogeneity of changes in high school availability in mid 1970s when it comes to choosing a window size of analysis; the narrower the window, the more likely is the change to be exogenous but the smaller the sample size. We choose to use such a 20-year window for our main analysis. In unreported estimations, however, we narrow progressively the window of analysis to the years between 1968 and 1982, and to the years between 1970 and 1980; such results are not qualitatively different from those presented in the paper.

4 Data

For our empirical analysis we use the fourth wave of Korean Labor and Income Panel Study (KLIPS) fielded in 2001. KLIPS is a nationally representative longitudinal study of Korean households, modeled after the National Longitudinal Surveys (NLS) and the Panel Study of Income Dynamics (PSID) of the US. This survey has been used elsewhere to study microeconomic issues of South Korea (e.g., Cho and Keum, 2004; Kang et al., 2007; Lee and Tae, 2005). It is conducted annually by the Korea Labor Institute (KLI), a government-sponsored research institute (Korea Labor Institute 1998). The study started in 1998 with 5,000 households and 13,783 individuals aged 15 or older. We use the fourth wave of KLIPS data, because in that wave a supplementary survey on health, that contains information on respondents’ health-related be-

haviors, was conducted. A total of 11,051 individuals in 4,248 households were interviewed in the fourth wave.

KLIPS collects a wide range of information on individuals such as earnings, family background, and demographic characteristics. In addition to the years of completed education, it collects unusually detailed information on individuals' education history. For example, for those who have attended high school, KLIPS collects information on name, type, and location of the high school(s) that the respondent has attended, and the years of starting and finishing high school(s). Those pieces of information allow us to identify the degree of availability of high school education at the graduation of middle school.⁵ We merge the information on the year and region of middle school graduation with the total number of general and vocational high school classrooms per 1,000 middle school graduates that is available in an administrative database of the Korean Educational Development Institute (KEDI).

In order to construct the sample for our analysis, we have imposed the following restrictions. First, we exclude women. Korean women's education, labor market participation, and health-related behaviors during the past three decades have been heavily affected by the society's changing views on women's status and role in the society. By excluding women, we try to avoid biases caused by such unobserved social changes. 5,739 observations are removed by the exclusion. Second, we restrict the sample to men who graduated from middle school between 1965 and 1985, both inclusive, as explained earlier. This restriction further removes 3,140 observations from the sample. Third, we then remove 561 observations because of missing information on key variables. At the end, our sample consists of 1,611 men aged 31 to 54 in 2001.

[Insert Table 1 here.]

Table 1 shows the summary statistics of the analysis sample and the subsamples divided by the education level. Individuals educated over 12 years (college attendees) seem to maintain healthier lifestyle than those educated up to 12 years (high school graduates or below). The former are more likely to receive regular health checkups, do exercises regularly, refrain from smoking, and drink moderately (as opposed to heavily) than the latter are. In addition, college attendees are less likely to be handicapped with a physical limitation or chronic illness before 25 than high school graduates or below. All the preceding differences are statistically significant.

However, such differences do not necessarily imply causal effect of education on health-related behaviors or health.

More educated men are more likely to work in professional occupations, to have higher household income, to have been raised in metropolitan areas at 14 years of age, to be a union member, and to work in large firms (with more than 500 employees) or in the public sector than less educated men. The more educated are less likely to be self-employed.⁶ As expected from the previous section, the more educated are younger and are more likely to be the first son than the less educated.

5 Estimation Results

5.1 Non-IV Estimation Results

[Insert Table 2 here.]

Table 2 shows the estimation results of health-related behavior equations without using IVs for education. Columns (1) and (2) show the estimates under the restriction that error terms are uncorrelated, that is, estimates from equation-by-equation probit estimations, while columns (3) and (4) show the estimates jointly obtained with the correlation matrix. Columns (1) and (3) show the results without controlling for the job characteristics, and (2) and (4) controlling for them. In square brackets under the estimates of years of education we report the average partial effect of one extra year of education on each probability.

Education is significantly associated with all behaviors except drinking. According to column (1), one extra year of education increases (1) the probability of receiving regular checkups by 2.0 percentage points, doing regular exercises by 3.3 percentage points, and refraining from smoking by 2.3 percentage points on average. The values are significantly different from zero. In contrast, education does not appear to be correlated strongly with drinking behavior. One extra year of education is associated with a minimal change in the probability of heavy drinking (i.e., drinking more than 5 times a month), and the education coefficient is statistically insignificant. The finding in Table 2 is similar to what Kim (2004) has found using another data set from Korea, the National Health and Nutrition Survey of 2001.

If job characteristics are controlled for, the size of the education coefficient estimates slightly

falls for health checkups and rises for no smoking, while it changes little for exercises and drinking. Nevertheless, the coefficients are statistically significant at the 5% level, except for drinking. The zero correlation restriction seems to have little effect on the estimation results. With or without the restriction, the coefficient estimates are virtually identical.

5.2 Determinants of the Years of Completed Education

[Insert Table 3 here.]

Table 3 shows the results of the education equation in IV estimations. Columns (1) and (2) are results from separate OLS regressions of education on exogenous variables, and columns (3) and (4) are results from joint SML estimations with health-related behavior equations. Columns (1) and (3) do not control for job characteristics, while columns (2) and (4) do.

Increased high school availability measured by the total number of general and vocational high school classrooms per 1,000 middle school graduates leads to significantly higher years of education. The estimates suggest that 10 more high school classrooms increase the average years of education by 0.26 to 0.37 years. The effect is statistically significant at the 5 percent level.

Being the first son also has a positive effect on a man's years of completed education. In columns (1) and (3), the first-son coefficient is statistically significant. The estimate suggests that the first son receives 0.26 to 0.31 more years of education than other siblings. In columns (2) and (4), however, the coefficient is not statistically significant.

The F statistics for our two IVs, which are available for non-joint estimations, are 7.30 in column (1) and 6.65 in column (2). According to Stock, Wright and Yogo (2002, Table 1), the F-statistic threshold of potential weakness of IVs for linear IV models is 11.6 in the case of one endogenous variable and two independent IVs. Even if our F-statistics are not terribly smaller than the threshold, we need to exercise caution in interpreting the statistical test results. Our testing methods may over-reject the null hypothesis, since the confidence intervals based on weak-IV-robust methods are generally wider than those based on normal approximations. The χ^2 statistics for our two IVs, which are available in columns (3) and (4) for joint estimations, show similar strength of the IVs in terms of p-values. In section 5.5, we experiment with a specification that employs as a single IV the total number of general and vocational high school

classrooms alone, dropping a first-son dummy; in such a case the first-stage F-statistics (and χ^2 statistics) are closer to the threshold. Results based on such a specification are qualitatively similar, except that the IV estimates for checkups become smaller and less precise with greater standard errors.

5.3 IV Estimation Results

[Insert Table 4 here.]

Table 4 shows the IV estimation results of the health behavior equations. Results in columns (1) and (2) are obtained from equation-by-equation IV probit estimations, and those in columns (3) and (4) from the multivariate IV probit estimations. The results of the education equation are reported Table 3 in the same order.

Education appears to increase the probability of an individual's receiving regular health checkups. The estimates in columns (1)–(4) suggest that one more year of education increases the probability by 4 to 9 percentage points on average. In columns (1) and (3) where the job characteristics are not controlled for, the coefficient is statistically significant at the 5 percent level. However, once job characteristics are controlled for, the education coefficient size becomes smaller and the coefficient is statistically insignificant at any conventional level. Meanwhile, all the job characteristics variables but the public/private sector dummy are estimated to have strong and significant impacts on individuals' having regular health checkups. This implies that the causal effect of education on individuals' choice of getting regular health checkups is mainly indirect, that is, operating through its effect on job characteristics.

In all columns, education is estimated to increase the likelihood of a man's doing regular exercises. The estimates suggest that one extra year of education increases the probability of doing regular exercises by 7 to 11 percentage points on average. The effect is statistically significant at the 5 percent level in all columns but (3). This is strong evidence that education has a causal effect on individuals' exercise behavior.

In contrast to the preceding two behaviors, education seems to have little causal effect on a man's smoking and drinking behaviors. One extra year of education increases the likelihood of refraining from smoking by only 0.6 to 3.2 percentage points. Such an insignificant effect of education on smoking is in contrast to the findings of previous research for other countries (e.g.,

de Walque 2007; Grimard and Parent 2007). Moreover, one more year of education changes the likelihood of moderate drinking by mere -0.2 to 2.6 percentage points. In all columns of the smoking and drinking equations, the education coefficients are statistically insignificant. In fact, not only are the education coefficients insignificant but coefficients of other variables are hardly significant in the smoking and drinking equations. An exception is the coefficient for marital status in the drinking equation.

We suspect that Korean culture that approves, if not encourages, frequent drinking among men is responsible for negligible effects of education and other variables on drinking (Sharpe et al. 2001). According to the studies on alcohol consumption of Koreans, Koreans tend to underestimate significantly the harmful effects of alcohol and view alcohol positively as social lubricants. For example, Koreans often hold beliefs that alcohol helps people to get to know each other better, that alcohol is essential for creating a pleasant atmosphere, and not being able to drink can hurt the social and work lives of men (Lee et al. 2007). To the extent that such culture is widespread in different social groups and that a man's drinking behavior is strongly affected by externality arising in social relations, it is not surprising to find negligible effects of individual characteristics on drinking.

Regarding the weak effects of education and other individual characteristics on smoking, we conjecture that the following two factors are likely to explain the finding. First, military service of two years or more—the length has changed a few times in the past—required of Korean men in their early 20s may account for the absence of the effect of education. To the extent that military experience increases a man's likelihood to start smoking (Bedard and Deschênes, 2006) and that education plays a limited role in smoking cessation (Grimard and Parent 2007), education may have little influence on a man's current smoking behavior; neither can individual characteristics given that military service is completed by most ordinary Korean men. The military service coefficient is statistically significant at the 5% level in the smoking equation in Table 2, although it is not significant in Table 4. The estimated coefficient sizes suggest that military service raises a man's likelihood of smoking by 4.4 to 10.4 percentage points on average.

Second, Koreans may have a weak belief that cigarette smoking is harmful, just like Europeans. Cutler and Glaeser (2006) present that Europeans have different (weaker) beliefs about negative effects of smoking than Americans. The researchers ascribe the differential re-

relationships between income, education and smoking to different views about smoking between Europeans and Americans. The absence of the effect of education on smoking in Korea can also be explained by the possibility that Koreans are similar to Europeans with respect to a belief in smoking.

The results on weak effects of education and other demographic or socioeconomic factors on smoking and drinking reveal difficulties of anti-smoking and drinking policies in Korean society. Since the individual's choices of smoking and drinking are culturally rooted to a large extent, other measures than economic or legal ones may be needed to promote healthy behaviors on drinking and smoking of Korean men.

Comparing the education coefficient estimates across the columns, we find that the estimated effect of education on choice of getting health checkups and doing exercises is smaller by 12 to 39 percent and by 13 to 27 percent respectively, when the effect is estimated jointly with the correlation matrix than when it is estimated with the zero correlation assumption. In contrast, the estimated effects of education on smoking and drinking behaviors seem to be affected little by the restriction on the correlation matrix. The changes in coefficients suggest that imposing a zero correlation restriction may lead to overestimating the effect of education on the two behaviors that education is estimated to have a causal effect on.

5.4 Correlations between Different Health Behaviors

[Insert Table 5 here.]

Table 5 reports estimated correlation coefficients between unobserved determinants of health behaviors that are estimated by the IV method. Some correlations are significant, whether or not job characteristics are included in estimation. For example, unobservable factors that increase the propensity to do exercises also increase the propensity to receive health checkups. The estimated correlation coefficient is 0.35 to 0.38. In addition, what drives an individual to refrain from smoking induces him to refrain from drinking heavily. The estimated correlation coefficient is roughly 0.32, which is significantly different from zero. In contrast, other correlation coefficients remain insignificant at the 5 percent level. The results suggest that an unobservable common factor affects both health checkups and exercises, while another factor influences smoking and drinking behaviors. The two factors do not seem to be strongly correlated, as the

correlation between exercises and smoking is not statistically significant.

It is not clear what causes such a correlation pattern. We speculate that since both smoking and drinking are addictive behaviors to some degree, it is possible that some genetic factors are involved. It is also possible that since both smoking and drinking may happen in some social settings, a degree of “socialization” unobserved in the data may cause correlation between smoking and drinking. These genetic or socialization factors are unlikely to have any effect on decisions on doing exercise or receiving health checkups, which are less addictive and less social activities than smoking or drinking.

Using the IV method, we are also able to estimate correlations between the unobserved determinant of education and that of health behaviors. From the estimated correlation coefficients of the lower panel of Table 5, it is notable that those who are likely to pursue greater education are, everything else equal, less likely to get health checkups and, to the greater degree, less likely to do regular exercises. The correlation coefficient of unobservables is -0.22 between checkups and education, and it is -0.54 between exercises and education. In contrast, the correlation coefficients between education and smoking/drinking behavior are estimated to be quite small—less than 0.05 in absolute value.

It is not surprising that individuals of greater education tend to be sedentary, considering that the Korean education system puts most emphasis on rote learning; in order to excel in the system, the individuals should be able to spend long hours in studying. The strong positive effect of education on doing regular exercises shown in Table 4 suggest that education enables the individuals of greater education to overcome their sedentary tendency.

Tests for endogeneity of education that examine zero correlations of unobservables between education and health behaviors suggest that education is strongly endogenous in the exercises equation, while tests do not reject exogeneity of education in the other equations. A joint test for zero correlations between education and all the health behaviors also fails to reject overall exogeneity of education; a lower p-value, which is obtained in a specification that controls for job characteristics, barely supports a rejection of the exogeneity. Nonetheless, the preceding test statistics do not invalidate a use of IV models, as non-rejection of a hypothesis does not necessarily imply the hypothesis is true. Since there is evidence suggesting that education is endogenous at least in the exercises or checkups equation depending on the specification, there

exists a strong need for a use of IV models.

5.5 Robustness

Earlier we caution that our tests may over-reject the null hypothesis because of potential weakness of the IVs. From Table 3 we suspect that such weakness comes from the fact that the first-son status is a weak determinant of a man's completed education. To examine robustness of our results against weakness of the IVs, we here apply the same analysis using only the availability of high schools as the single IV. In such a case, the estimate of the number of total high school classrooms in the education equation is 0.030 (s.e. = 0.010) and the F-statistic is 9.45 if job characteristics are excluded. If they are included, the estimate is 0.035 (s.e. = 0.011) and the F-statistic is 11.0. Thus, according to the criterion proposed by Stock, Wright and Yogo (2002), the single IV does not seem to be weak; we can apply the usual testing method based on normal approximations. The IV results of the health behavior equations based on the single IV are reported in panel (B) of Table 6. For comparison, the results of Table 4 are summarized in panel (A). At the bottom of each panel we report F or χ^2 statistics (as well as p-values) for excluded IVs to examine their strength.

[Insert Table 6 here.]

Overall the estimated effects of education on health behaviors in panel (B) are similar to those in panel (A). An exception is that the estimated effect of education on health checkups in panel (B) is much smaller than that in panel (A), and none of the estimates for education is statistically significant. The effect of education on doing regular exercises is still estimated to be strong and statistically significant in panel (B). One extra year of education increases the likelihood of doing exercises by 10.2 to 11.5 percentage points on average.

As another check of robustness, we use the numbers of general and vocational high school classrooms separately as IVs in addition to the first-son dummy. Availability of general and vocational high schools may have different impacts on education, because general high school graduates are more likely to go to colleges than vocational high school graduates. The estimation results are reported in panel (C). In this case we find that overall strength of the IVs is somewhat reduced: the first-stage F statistics for the three IVs are 5.33 if job characteristics are not

included, and 4.92 if they are included in estimations. Nonetheless, the overall estimation results are similar to those in panel (A). The effect of education on getting health checkups is estimated to be a little stronger, while that on doing regular exercises is a little weaker.

As a third check of robustness of our results, we estimate linear probability models of health-related behaviors, employing the same sets of explanatory variables and IVs as in probit models. Although the linear probability models may be less efficient than our ML-based models and limited in estimating correlations between unobservable determinants of health-related behaviors, they are robust to functional form assumptions that may cause bias in ML-based models.⁷ The results of the linear probability models in Table 7 show little qualitative difference from the earlier findings.

[Insert Table 7 here.]

Similar to the earlier findings, education fails to show any significant effect on smoking and drinking behaviors. Although education generally increases the likelihood of refraining from smoking and drinking, the size of the effects is indistinguishable from zero. In contrast, education significantly increases the likelihood of doing exercises. One extra year of education increases the likelihood of doing exercises by 8.2 to 10.8 percentage points, which are generally comparable to the results of our ML estimates. Education also affects the likelihood of receiving regular checkups; however, the effect weakens if the job characteristics are controlled for.

The p-values of over-identification test statistics for IV exogeneity are reported in angle brackets of Table 7. The p-values of separate individual models are in columns (1) and (2); those of joint models are in columns (3) and (4). The majority of the cases in panels (A) and (C) support IV exogeneity at the 5 percent significance level. However, there are three notable exceptions. The first is the p-value for the effect on exercises with controls of job characteristics in panel (A); the test statistic, however, becomes insignificant if job characteristics are excluded or separate general/vocational classroom variables are employed. The second exception is the p-value for the effect on drinking without controls of job characteristics in panels (A) and (C); the test statistics, however, become insignificant if job characteristics are included. The third exception is p-values of joint models in panels (A) and (C). Three of four p-values of joint linear probability models are less than 0.05, suggesting some caution about the exogeneity of our IVs. Nonetheless, if we rely on the results of column (4) in panel (C) for which the p-value is in favor

of exogeneity of IVs, the overall patterns of the estimated effects of education remain similar to those reported earlier.

6 Summary and conclusion

In this paper we examine whether education induces individuals to have healthy lifestyle. In the estimation endogeneity of education is addressed by an IV estimation method in which high school availability and birth order of Korean men serve as IVs for education. In addition, to explore potential correlations between health behaviors, we employ a quadivariate probit model with an endogenous variable using simulated MLE.

The estimation results suggest that increased education induces individuals to exercise regularly. Education also induces individuals to get regular health checkups, but the effect of education is mainly through its effects on job characteristics. On smoking and drinking behavior, the causal effect of education is statistically insignificant. We also discover that unobserved determinants of health behaviors are correlated, especially between smoking and drinking and between exercising and getting health checkups. Ignoring such correlations leads to overstating effects of education on health behaviors in our case.

Given our current data set, it is difficult to explore further what explains the weak effects of education on smoking and drinking. The compulsory military service and cultural acceptance of smoking and drinking among Korean men are possible explanations. Further research is needed to address this issue in the future.

Notes

¹We assume that cross-sectional data are used and that prices do not vary across individuals. That is why prices are not included in equation (5).

²If parents favor the first son with respect to educational investments, they are also likely to support the same child more over other dimensions such as health. Thus if the effect of education on adulthood health is estimated on the basis of the first-son indicator as an IV, then a (probably upward) bias may arise. However, as we use adulthood health-related *behaviors* as the dependent variable in the current paper, such a bias will be relatively small.

³In order to integrate household income and household size, we also experiment with specifications that employ equivalent incomes proposed by OECD. The equivalent income is a household's income divided by the number of consumption units in the household, where the consumption unit differs by the age of household members. Estimation results under such specifications, however, are not qualitatively different from those reported in the paper. They are available upon request.

⁴In contrast to other regions, Seoul did not go through such a rapid change in the number of high school classrooms. However, it does not necessarily imply that middle school graduates in Seoul did not experience an expansion in high school availability. Before 1974 general high schools in Seoul were attended by middle school graduates not only from Seoul but from other regions of the nation. To the extent that the number of middle school graduates migrating from other regions to Seoul decreased after 1974, the availability of general high school for middle school graduates in Seoul must have similarly expanded. Kang, Park and Lee (2006, Table 2) show the decrease in migration into high schools in Seoul that happened with the leveling policy.

⁵For those whose education was completed below high school (below the 10th grade), we project the year of middle school graduation by adding 16 to the year of birth. 16 is chosen because Korean children enter elementary school at age 7 and the duration of elementary and middle school period is 6 and 3 years, respectively.

⁶In the subsequent analysis those self-employed are treated as non-union members who work at small firms in the private sector.

⁷For other merits and demerits of linear probability models over ML-based models, see Angrist (2001) and several comments about the paper.

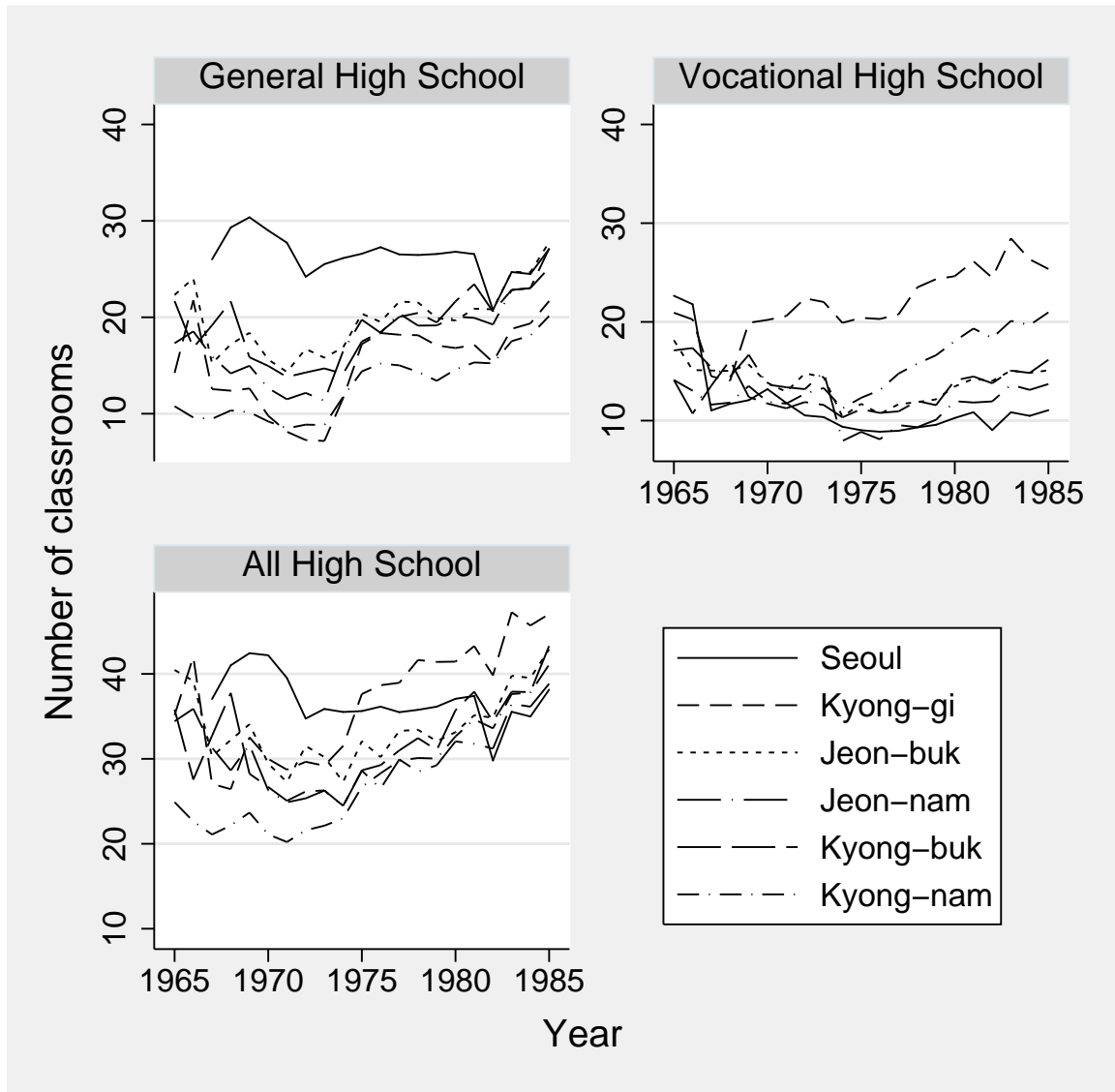
References

- Angrist, J. D. (2001). Estimation of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice. *Journal of Business & Economic Statistics* 19(1), 1–23.
- Becker, G. S. and C. B. Mulligan (1997). The endogenous determination of time preference. *Quarterly Journal of Economics* 112(3), 729–758.
- Bedard, K. and O. Deschênes (2006). The impact of military service on long-term health: Evidence from World War II and Korean War veterans. *American Economic Review* 96(1), 176–194.
- Cho, J. and J. Keum (2004). Job instability in the Korean labour market: Estimating the effects of the 1997 financial crisis. *International Labour Review* 143(4), 373–392.
- Chung, B.-G. (1998). *A Study of the High School Leveling Policy in the Republic of Korea: Its Genesis, Implementation, and Reforms, 1974–1995*. Ph. D. thesis, University of Hawaii.
- Contoyannis, P. and A. M. Jones (2004). Socio-economic status, health, and lifestyle. *Journal of Health Economics* 23, 965–995.
- Currie, J. and E. Moretti (2003). Mother’s education and the intergenerational transmission of human capital: Evidence from college openings. *Quarterly Journal of Economics* 118, 1495–1532.
- Cutler, D. M. and E. L. Glaeser (2006). Why do Europeans smoke more than Americans? NBER Working Paper 12124.
- de Walque, D. (2007). Does education affect smoking behaviors?: Evidence using the Vietnam draft as an instrument for college education. *Journal of Health Economics* 26(5), 877–895.
- Grimard, F. and D. Parent (2007). Education and smoking: Were Vietnam war draft avoiders also more likely to avoid smoking? *Journal of Health Economics* 26(5), 896–926.
- Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy* 80, 223–255.
- Grossman, M. (2005). Education and nonmarket outcomes. NBER Working Paper 11582.

- Hajivassiliou, V., D. McFadden, and P. Ruud (1996). Simulation of multivariate normal rect-angle probabilities and their derivatives: Theoretical and computational results. *Journal of Econometrics* 72, 85–134.
- Kang, C., C. Park, and M.-J. Lee (2007). Effects of ability mixing in high school on adult-hood earnings: Quasi-experimental evidence from South Korea. *Journal of Population Economics* 20(2), 269–297.
- Kenkel, D. S. (1991). Health behavior, health knowledge, and schooling. *Journal of Political Economy* 99, 287–305.
- Kenkel, D. S., D. R. Lillard, and A. D. Mathios (2006, January). The roles of high school completion and GED receipt in smoking and obesity. NBER Working Paper No. 11990.
- Kim, C.-Y. (2004, June). Socio-economic status, health, and health risk behavior. *Health and Well-being Forum* 92, 19–25. (in Korean).
- Korea Labor Institute (1998). *Reports on the Korean Labor and Income Panel Study (in Korean) (Various Years)*. Seoul: Korea Labor Institute.
- Lee, D. W., H. S. Park, T. S. Lee, M. K. Kim, and Y. H. Kim (2007). Functions of so-cial gatherings involving alcohol consumption with coworkers in Korea. *Asian Journal of Communication* 17(3), 266–285.
- Lee, M.-J. and Y. Tae (2005). Analysis of labour participation behaviour of Korean women with dynamic probit and conditional logit. *Oxford Bulletin of Economics and Statis-tics* 67(1), 71–91.
- Leigh, J. P. and R. Dhir (1997). Schooling and frailty among seniors. *Economics of Education Review* 16(1), 45–57.
- Park, C. B. and N.-H. Cho (1995). Consequences of son preference in a low-fertility society: Imbalance of the sex ratio at birth in Korea. *Population and Development Review* 21(1), 59–84.
- Pollak, R. A. and M. L. Wachter (1975). The relevance of the household production function and its implications for the allocation of time. *Journal of Political Economy* 83(2), 255–278.

- Rosenzweig, M. R. and T. P. Schultz (1983). Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight. *Journal of Political Economy* 91(5), 723–746.
- Sander, W. (1995a). Schooling and quitting smoking. *Review of Economics and Statistics* 77, 191–199.
- Sander, W. (1995b). Schooling and smoking. *Economics of Education Review* 14(1), 23–33.
- Sharpe, D. L., M. Abdel-Ghany, H.-Y. Kim, and G.-S. Hong (2001). Alcohol consumption decisions in Korea. *Journal of Family and Economic Issues* 22(1), 7–24.
- Stock, J. H., J. H. Wright, and M. Yogo (2002). A survey of weak instruments and weak identification in Generalized Method of Moments. *Journal of Business & Economic Statistics* 20(4), 518–529.
- Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press.

Figure 1: High school classrooms per 1,000 middle school graduates



Source: Database of the Korean Educational Development Institute (KEDI)

Table 1: Summary Statistics of the Main Sample

Variable	Total Sample	(1)	(2)	Difference (2)-(1)
		Education up to 12 years	Education over 12 years	
	Mean (S.D.)	Mean (S.D.)	Mean (S.D.)	Mean (S.E.)
Receives regular checkups ^a	0.413 (0.492)	0.338 (0.473)	0.556 (0.497)	0.218 (0.025)**
Does regular exercises ^a	0.294 (0.456)	0.215 (0.411)	0.446 (0.498)	0.231 (0.023)**
Refrain from smoking ^a	0.304 (0.460)	0.256 (0.437)	0.394 (0.489)	0.137 (0.024)**
Drink moderately ^a	0.593 (0.491)	0.566 (0.496)	0.646 (0.479)	0.080 (0.026)**
Handicapped ^a	0.028 (0.165)	0.038 (0.191)	0.009 (0.095)	-0.029 (0.009)**
Years of education	12.479 (3.072)	10.709 (2.067)	15.856 (1.383)	5.147 (0.098)**
Age	41.755 (5.884)	42.686 (5.803)	39.980 (5.630)	-2.706 (0.301)**
Married ^a	0.962 (0.191)	0.958 (0.200)	0.969 (0.173)	0.011 (0.010)
Household members aged 0-6	0.462 (0.735)	0.362 (0.670)	0.652 (0.813)	0.289 (0.038)**
Household members aged 7-24	1.361 (0.980)	1.504 (0.957)	1.087 (0.968)	-0.418 (0.050)**
Household members aged 25-55	1.978 (0.345)	1.973 (0.375)	1.987 (0.279)	0.015 (0.018)
Household members aged 56-	0.132 (0.387)	0.135 (0.376)	0.126 (0.406)	-0.009 (0.020)
Served in military ^a	0.777 (0.416)	0.715 (0.452)	0.895 (0.306)	0.180 (0.021)**
Parent's years of education	6.365 (4.698)	5.190 (4.199)	8.606 (4.789)	3.416 (0.231)**
Occupation ^a				
Professional	0.230 (0.421)	0.096 (0.295)	0.484 (0.500)	0.387 (0.020)**
Administrative and clerical	0.097 (0.297)	0.056 (0.230)	0.177 (0.382)	0.121 (0.015)**
Service and sales	0.160 (0.367)	0.167 (0.373)	0.148 (0.355)	-0.018 (0.019)
Manual labor	0.475 (0.500)	0.629 (0.483)	0.181 (0.385)	-0.449 (0.024)**
Agricultural	0.038 (0.191)	0.052 (0.222)	0.011 (0.104)	-0.041 (0.010)**
Household income $\times 10^{-6}$ (KRW)	2.068 (1.548)	1.730 (1.248)	2.711 (1.834)	0.980 (0.077)**
Job characteristics ^a :				
Self-employed	0.342 (0.475)	0.382 (0.486)	0.266 (0.442)	-0.116 (0.025)**
Union member	0.127 (0.333)	0.116 (0.320)	0.147 (0.354)	0.031 (0.018)*
Large firm	0.139 (0.346)	0.099 (0.299)	0.222 (0.416)	0.122 (0.019)**
Public sector	0.112 (0.316)	0.063 (0.243)	0.205 (0.404)	0.142 (0.016)**
Type of residential area at 14 ^a				
Metropolitan	0.291 (0.454)	0.234 (0.423)	0.399 (0.490)	0.165 (0.023)**
Other urban	0.242 (0.428)	0.231 (0.422)	0.264 (0.441)	0.033 (0.022)
Rural	0.467 (0.499)	0.535 (0.499)	0.338 (0.473)	-0.198 (0.026)**
General high school classrooms per 1,000 middle school graduates	19.3 (6.241)	18.4 (6.267)	20.9 (5.852)	2.505 (0.321)**
Vocational high school classrooms per 1,000 middle school graduates	14.4 (4.403)	14.4 (4.286)	14.4 (4.620)	-0.084 (0.231)
All high school classrooms per 1,000 middle school graduates	33.7 (7.165)	32.9 (7.424)	35.3 (6.352)	2.421 (0.371)**
First-son ^a	0.425 (0.495)	0.402 (0.491)	0.469 (0.500)	0.067 (0.026)**
Year of high school entrance	1975.4 (5.813)	1974.5 (5.764)	1977.1 (5.523)	2.591 (0.298)**
Sample size	1,611	1,057	554	

Note: ^a binary variables. * Statistically significant at the 10 % level; ** at the 5 % level.

Table 2: Health Behavior Equations: No IVs

Variables	Estimated separately		Estimated jointly	
	(1) Estimate (S.E.)	(2) Estimate (S.E.)	(3) Estimate (S.E.)	(4) Estimate (S.E.)
(I) Receives regular health checkups (1 if yes, 0 if no):				
Years of education	0.060 (0.015)** [0.020]	0.039 (0.018)** [0.011]	0.060 (0.015)** [0.020]	0.039 (0.018)** [0.011]
Handicapped	0.102 (0.208)	0.339 (0.235)	0.102 (0.205)	0.328 (0.231)
Age	-0.107 (0.090)	-0.020 (0.105)	-0.105 (0.090)	-0.024 (0.105)
Age ²	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Married	0.109 (0.235)	-0.049 (0.273)	0.103 (0.234)	-0.057 (0.271)
HH members aged 0-6	-0.071 (0.066)	-0.094 (0.077)	-0.067 (0.066)	-0.085 (0.077)
HH members aged 7-24	-0.012 (0.054)	-0.034 (0.063)	-0.010 (0.054)	-0.029 (0.063)
HH members aged 25-55	0.001 (0.118)	0.137 (0.135)	0.001 (0.119)	0.139 (0.135)
HH members aged 56-	-0.114 (0.093)	-0.091 (0.109)	-0.107 (0.093)	-0.081 (0.109)
Served in military	0.093 (0.086)	0.053 (0.100)	0.091 (0.086)	0.054 (0.100)
Parent's years of education	-0.003 (0.008)	0.002 (0.010)	-0.004 (0.008)	0.001 (0.010)
Ln(HH income)	0.349 (0.052)**	0.314 (0.061)**	0.341 (0.051)**	0.303 (0.061)**
Self-employed		-0.596 (0.098)**		-0.594 (0.098)**
Union member		0.929 (0.144)**		0.922 (0.143)**
Large firm		0.834 (0.141)**		0.823 (0.140)**
Public sector		0.486 (0.267)*		0.513 (0.268)*
(II) Does regular exercises (1 if yes, 0 if no):				
Years of education	0.104 (0.016)** [0.033]	0.105 (0.018)** [0.032]	0.103 (0.016)** [0.033]	0.104 (0.018)** [0.032]
Handicapped	-0.258 (0.239)	-0.183 (0.256)	-0.225 (0.233)	-0.145 (0.249)
Age	0.002 (0.093)	0.014 (0.102)	0.001 (0.093)	0.016 (0.101)
Age ²	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)
Married	-0.006 (0.245)	0.151 (0.266)	-0.018 (0.243)	0.141 (0.265)
HH members aged 0-6	-0.038 (0.068)	-0.055 (0.073)	-0.030 (0.068)	-0.050 (0.073)
HH members aged 7-24	0.015 (0.056)	0.000 (0.061)	0.013 (0.056)	-0.004 (0.062)
HH members aged 25-55	-0.020 (0.122)	-0.097 (0.133)	-0.020 (0.124)	-0.100 (0.136)
HH members aged 56-	-0.133 (0.095)	-0.062 (0.102)	-0.136 (0.095)	-0.067 (0.103)
Served in military	0.115 (0.091)	0.108 (0.100)	0.102 (0.090)	0.096 (0.100)
Parent's years of education	0.000 (0.008)	-0.007 (0.009)	0.000 (0.008)	-0.007 (0.009)
Ln(HH income)	0.139 (0.050)**	0.123 (0.054)**	0.126 (0.049)**	0.111 (0.053)**
Self-employed		0.217 (0.095)**		0.220 (0.095)**
Union member		0.370 (0.125)**		0.374 (0.125)**
Large firm		0.121 (0.123)		0.126 (0.123)
Public sector		0.384 (0.203)*		0.391 (0.204)*

(Continued to the next page.)

Table 2: Health Behavior Equations: No IVs (Continued)

Variables	Estimated separately		Estimated jointly	
	(1) Estimate (S.E.)	(2) Estimate (S.E.)	(3) Estimate (S.E.)	(4) Estimate (S.E.)
(III) Refrains from smoking (1 if yes, 0 if no):				
Years of education	0.069 (0.015)** [0.023]	0.076 (0.017)** [0.025]	0.068 (0.015)** [0.023]	0.077 (0.017)** [0.025]
Handicapped	-0.168 (0.217)	-0.185 (0.234)	-0.170 (0.217)	-0.182 (0.233)
Age	-0.050 (0.090)	-0.054 (0.097)	-0.051 (0.090)	-0.053 (0.097)
Age ²	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Married	-0.254 (0.230)	-0.246 (0.246)	-0.257 (0.230)	-0.255 (0.245)
HH members aged 0-6	-0.016 (0.066)	-0.046 (0.072)	-0.015 (0.066)	-0.046 (0.072)
HH members aged 7-24	-0.023 (0.053)	-0.030 (0.059)	-0.020 (0.053)	-0.029 (0.059)
HH members aged 25-55	0.168 (0.116)	0.167 (0.124)	0.166 (0.117)	0.162 (0.125)
HH members aged 56-	0.125 (0.088)	0.131 (0.096)	0.117 (0.089)	0.120 (0.096)
Served in military	-0.201 (0.085)**	-0.284 (0.093)**	-0.200 (0.085)**	-0.285 (0.093)**
Parent's years of education	0.015 (0.008)*	0.022 (0.009)**	0.015 (0.008)*	0.021 (0.009)**
Ln(HH income)	0.038 (0.047)	-0.011 (0.050)	0.039 (0.047)	-0.008 (0.050)
Self-employed		0.072 (0.092)		0.070 (0.092)
Union member		-0.010 (0.127)		0.000 (0.127)
Large firm		0.240 (0.123)*		0.235 (0.123)*
Public sector		-0.155 (0.202)		-0.170 (0.202)
(IV) Drink (1 if moderately, 0 if heavily):				
Years of education	0.017 (0.014) [0.006]	0.019 (0.016) [0.007]	0.017 (0.014) [0.006]	0.018 (0.016) [0.007]
Handicapped	0.133 (0.195)	0.254 (0.214)	0.137 (0.195)	0.258 (0.214)
Age	-0.112 (0.086)	-0.072 (0.093)	-0.110 (0.086)	-0.069 (0.093)
Age ²	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Married	0.594 (0.218)**	0.562 (0.233)**	0.583 (0.216)**	0.552 (0.231)**
HH members aged 0-6	-0.007 (0.063)	-0.008 (0.067)	-0.006 (0.063)	-0.008 (0.067)
HH members aged 7-24	-0.076 (0.051)	-0.062 (0.056)	-0.076 (0.050)	-0.062 (0.055)
HH members aged 25-55	-0.068 (0.110)	-0.099 (0.118)	-0.061 (0.109)	-0.093 (0.117)
HH members aged 56-	-0.044 (0.084)	-0.055 (0.091)	-0.042 (0.085)	-0.056 (0.091)
Served in military	-0.047 (0.081)	-0.101 (0.088)	-0.045 (0.081)	-0.098 (0.088)
Parent's years of education	0.001 (0.008)	0.005 (0.009)	0.002 (0.008)	0.006 (0.009)
Ln(HH income)	0.082 (0.044)*	0.086 (0.047)*	0.082 (0.045)*	0.086 (0.047)*
Self-employed		-0.120 (0.085)		-0.121 (0.085)
Union member		0.095 (0.123)		0.095 (0.123)
Large firm		0.006 (0.122)		0.002 (0.122)
Public sector		-0.276 (0.195)		-0.283 (0.195)
Number of sample	1,611	1,377	1,611	1,377

Note: Coefficients of the intercept, current residential region, residential area type at 14, and occupation dummies are suppressed. Asymptotic standard errors are in parentheses. Average partial effects of one extra year of education are in square brackets.

* Statistically significant at the 10 % level; ** at the 5 % level.

Table 3: Results of the Education Equation

Explanatory Variables	Dependent variable: Years of Education			
	Estimated separately		Estimated jointly	
	(1) Estimate (S.E.)	(2) Estimate (S.E.)	(3) Estimate (S.E.)	(4) Estimate (S.E.)
All HS classrooms	0.030 (0.010)**	0.035 (0.011)**	0.026 (0.015)*	0.037 (0.010)**
First-son	0.261 (0.116)**	0.185 (0.122)	0.309 (0.141)**	0.086 (0.181)
Handicapped	-0.549 (0.342)	-0.498 (0.359)	-0.555 (0.340)	-0.504 (0.356)
Age	0.314 (0.165)*	0.382 (0.173)**	0.289 (0.180)	0.388 (0.172)**
Age ²	-0.004 (0.002)**	-0.005 (0.002)**	-0.004 (0.002)*	-0.005 (0.002)**
Married	0.239 (0.377)	0.165 (0.394)	0.234 (0.374)	0.168 (0.390)
HH members aged 0-6	0.180 (0.109)*	0.211 (0.114)*	0.181 (0.108)*	0.210 (0.113)*
HH members aged 7-24	-0.153 (0.089)*	-0.110 (0.095)	-0.156 (0.088)*	-0.104 (0.094)
HH members aged 25-55	0.051 (0.193)	0.003 (0.200)	0.049 (0.191)	0.010 (0.199)
HH members aged 56-	-0.244 (0.149)	-0.260 (0.157)*	-0.255 (0.150)*	-0.240 (0.158)
Served in military	1.344 (0.137)**	1.341 (0.145)**	1.353 (0.138)**	1.331 (0.144)**
Parent's education	0.121 (0.013)**	0.129 (0.014)**	0.121 (0.013)**	0.130 (0.014)**
Occupation dummies (excluded: Professional)				
Admin. & clerical	-0.762 (0.214)**	-0.807 (0.251)**	-0.759 (0.212)**	-0.809 (0.248)**
Service and sales	-1.831 (0.187)**	-1.598 (0.207)**	-1.829 (0.186)**	-1.595 (0.205)**
Manual labor	-2.836 (0.152)**	-2.476 (0.162)**	-2.833 (0.151)**	-2.475 (0.160)**
Agricultural	-3.553 (0.326)**	-3.326 (0.337)**	-3.556 (0.323)**	-3.317 (0.333)**
Ln(HH income)	0.565 (0.076)**	0.503 (0.079)**	0.565 (0.076)**	0.502 (0.078)**
Self-employed		0.122 (0.146)		0.115 (0.144)
Union member		-0.076 (0.206)		-0.076 (0.204)
Large firm		0.513 (0.203)**		0.517 (0.201)**
Public sector		1.178 (0.332)**		1.172 (0.329)**
Intercept	5.817 (3.650)	3.968 (3.836)	6.484 (4.134)	3.790 (3.808)
Number of sample	1,611	1,377	1,611	1,377
F(Instruments excluded from the 2nd stage)	7.30	6.65		
Chi-square(Instruments)			14.47	12.87
P-value	0.0007	0.0013	0.0007	0.0016

Note: Coefficients of current residential region dummies and those of residential area type at 14 dummies are suppressed. Asymptotic standard errors are in parentheses.

* Statistically significant at the 10 % level; ** at the 5 % level.

Table 4: Results of the Health Behavior Equations: IVs

Variables	Estimated separately		Estimated jointly	
	(1) Estimate (S.E.)	(2) Estimate (S.E.)	(3) Estimate (S.E.)	(4) Estimate (S.E.)
(I) Receives regular health checkups (1 if yes, 0 if no):				
Years of education	0.286 (0.113)** [0.089]	0.227 (0.147) [0.064]	0.252 (0.104)** [0.081]	0.139 (0.178) [0.039]
Handicapped	0.232 (0.203)	0.417 (0.227)*	0.208 (0.183)	0.370 (0.228)
Age	-0.110 (0.085)	-0.042 (0.101)	-0.100 (0.079)	-0.036 (0.103)
Age ²	0.002 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Married	0.039 (0.224)	-0.069 (0.259)	0.037 (0.207)	-0.068 (0.261)
HH members aged 0-6	-0.102 (0.062)*	-0.128 (0.076)*	-0.089 (0.057)	-0.103 (0.079)
HH members aged 7-24	0.026 (0.055)	-0.010 (0.064)	0.023 (0.051)	-0.016 (0.066)
HH members aged 25-55	-0.015 (0.110)	0.125 (0.130)	-0.013 (0.102)	0.132 (0.132)
HH members aged 56-	-0.052 (0.096)	-0.040 (0.114)	-0.045 (0.093)	-0.055 (0.119)
Served in military	-0.241 (0.199)	-0.215 (0.236)	-0.205 (0.196)	-0.089 (0.279)
Parent's years of education	-0.033 (0.017)*	-0.024 (0.022)	-0.029 (0.016)*	-0.013 (0.026)
Ln(HH income)	0.168 (0.135)	0.192 (0.134)	0.158 (0.152)	0.238 (0.154)
Self-employed		-0.568 (0.116)**		-0.579 (0.123)**
Union member		0.858 (0.188)**		0.883 (0.204)**
Large firm		0.657 (0.245)**		0.729 (0.277)**
Public sector		0.223 (0.351)		0.372 (0.395)
(II) Does regular exercises (1 if yes, 0 if no):				
Years of education	0.289 (0.122)** [0.089]	0.368 (0.095)** [0.109]	0.212 (0.151) [0.068]	0.321 (0.080)** [0.100]
Handicapped	-0.109 (0.257)	0.020 (0.247)	-0.137 (0.268)	0.039 (0.219)
Age	-0.013 (0.089)	-0.024 (0.093)	-0.008 (0.089)	-0.019 (0.084)
Age ²	0.000 (0.001)	0.001 (0.001)	0.000 (0.001)	0.001 (0.001)
Married	-0.049 (0.235)	0.080 (0.242)	-0.043 (0.234)	0.066 (0.218)
HH members aged 0-6	-0.070 (0.067)	-0.105 (0.068)	-0.049 (0.069)	-0.090 (0.060)
HH members aged 7-24	0.044 (0.056)	0.033 (0.057)	0.030 (0.058)	0.025 (0.051)
HH members aged 25-55	-0.033 (0.116)	-0.078 (0.120)	-0.027 (0.118)	-0.072 (0.111)
HH members aged 56-	-0.082 (0.102)	0.016 (0.099)	-0.105 (0.108)	0.008 (0.092)
Served in military	-0.163 (0.220)	-0.307 (0.200)	-0.062 (0.266)	-0.272 (0.183)
Parent's years of education	-0.025 (0.019)	-0.044 (0.016)**	-0.015 (0.023)	-0.038 (0.014)**
Ln(HH income)	0.017 (0.107)	-0.043 (0.093)	0.053 (0.128)	-0.042 (0.088)
Self-employed		0.134 (0.102)		0.126 (0.101)
Union member		0.306 (0.132)**		0.279 (0.133)**
Large firm		-0.059 (0.141)		-0.043 (0.133)
Public sector		-0.020 (0.275)		0.001 (0.268)

(Continued to the next page.)

Table 4: Results of the Health Behavior Equations: IVs (Continued)

Variables	Estimated separately		Estimated jointly	
	(1) Estimate (S.E.)	(2) Estimate (S.E.)	(3) Estimate (S.E.)	(4) Estimate (S.E.)
(III) Refrains from smoking (1 if yes, 0 if no):				
Years of education	0.017 (0.158) [0.006]	0.088 (0.169) [0.029]	0.023 (0.160) [0.008]	0.096 (0.171) [0.032]
Handicapped	-0.199 (0.234)	-0.178 (0.253)	-0.196 (0.231)	-0.171 (0.255)
Age	-0.046 (0.091)	-0.055 (0.099)	-0.047 (0.090)	-0.055 (0.099)
Age ²	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Married	-0.240 (0.234)	-0.248 (0.247)	-0.244 (0.234)	-0.258 (0.245)
HH members aged 0-6	-0.007 (0.072)	-0.049 (0.080)	-0.007 (0.072)	-0.050 (0.080)
HH members aged 7-24	-0.031 (0.058)	-0.029 (0.062)	-0.026 (0.058)	-0.026 (0.062)
HH members aged 25-55	0.171 (0.116)	0.167 (0.124)	0.168 (0.116)	0.162 (0.125)
HH members aged 56-	0.115 (0.094)	0.134 (0.103)	0.108 (0.096)	0.124 (0.103)
Served in military	-0.129 (0.235)	-0.300 (0.248)	-0.137 (0.240)	-0.311 (0.247)
Parent's years of education	0.022 (0.021)	0.020 (0.024)	0.021 (0.020)	0.018 (0.025)
Ln(HH income)	0.066 (0.097)	-0.016 (0.098)	0.063 (0.096)	-0.018 (0.099)
Self-employed		0.071 (0.095)		0.067 (0.095)
Union member		-0.009 (0.127)		0.001 (0.127)
Large firm		0.234 (0.154)		0.224 (0.159)
Public sector		-0.168 (0.277)		-0.192 (0.278)
(IV) Drink (1 if moderately, 0 if heavily):				
Years of education	-0.006 (0.180) [-0.002]	0.070 (0.180) [0.026]	0.052 (0.203) [0.019]	0.003 (0.202) [0.001]
Handicapped	0.119 (0.225)	0.282 (0.232)	0.158 (0.227)	0.249 (0.246)
Age	-0.110 (0.088)	-0.078 (0.095)	-0.112 (0.086)	-0.066 (0.097)
Age ²	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Married	0.599 (0.219)**	0.551 (0.238)**	0.572 (0.232)**	0.554 (0.231)**
HH members aged 0-6	-0.003 (0.070)	-0.019 (0.077)	-0.012 (0.072)	-0.005 (0.080)
HH members aged 7-24	-0.080 (0.057)	-0.055 (0.060)	-0.070 (0.063)	-0.064 (0.059)
HH members aged 25-55	-0.066 (0.111)	-0.099 (0.118)	-0.064 (0.110)	-0.092 (0.117)
HH members aged 56-	-0.048 (0.091)	-0.043 (0.101)	-0.035 (0.095)	-0.060 (0.101)
Served in military	-0.015 (0.257)	-0.170 (0.258)	-0.092 (0.286)	-0.076 (0.292)
Parent's years of education	0.004 (0.024)	-0.002 (0.026)	-0.003 (0.027)	0.008 (0.028)
Ln(HH income)	0.095 (0.108)	0.060 (0.104)	0.062 (0.128)	0.093 (0.108)
Self-employed		-0.126 (0.087)		-0.119 (0.090)
Union member		0.097 (0.123)		0.094 (0.124)
Large firm		-0.022 (0.155)		0.010 (0.163)
Public sector		-0.332 (0.273)		-0.264 (0.307)
Number of sample	1,611	1,377	1,611	1,377

Note: Coefficients of the intercept, current residential region, residential area type at 14, and occupation dummies are suppressed. Asymptotic standard errors are in parentheses. Average partial effects of one extra year of education are in square brackets.

* Statistically significant at the 10 % level; ** at the 5 % level.

Table 5: Estimated Correlation Coefficients among Residuals

	Job characteristics not controlled for				
	(I)	(II)	(III)	(IV)	(V)
(I) Checkups	1.000	0.376** (0.144)	0.026 (0.166)	0.049 (0.217)	-0.455* (0.261)
(II) Exercises		1.000	0.217* (0.120)	-0.062 (0.117)	-0.259 (0.380)
(III) No smoking			1.000	0.317** (0.073)	0.098 (0.345)
(IV) Drink moderately				1.000	-0.078 (0.454)
(V) Education					1.000
LR test:					
$H_0: \rho_{15} = \rho_{25} = \rho_{35} = \rho_{45} = 0$				$\chi^2_{df=4}=4.65$ p-value=0.325	
	Job characteristics controlled for				
	(I)	(II)	(III)	(IV)	(V)
(I) Checkups	1.000	0.348* (0.193)	0.080 (0.097)	0.005 (0.105)	-0.221 (0.398)
(II) Exercises		1.000	0.230 (0.200)	-0.091 (0.234)	-0.536** (0.223)
(III) No smoking			1.000	0.324** (0.052)	-0.042 (0.375)
(IV) Drink moderately				1.000	0.034 (0.436)
(V) Education					1.000
LR test:					
$H_0: \rho_{15} = \rho_{25} = \rho_{35} = \rho_{45} = 0$				$\chi^2_{df=4}=7.53$ p-value=0.110	

Note: Asymptotic standard errors are in parentheses.

* Statistically significant at the 10 % level; ** at the 5 % level.

Table 6: Summary of IV Results

Estimated jointly	No	No	Yes	Yes
Job characteristics	No	Yes	No	Yes
(A) IVs: All High School Classrooms and a First-son Dummy				
Behaviors	(1)	(2)	(3)	(4)
(I) Checkups	0.286 (0.113)** [0.089]	0.227 (0.147) [0.064]	0.252 (0.104)** [0.081]	0.139 (0.178) [0.039]
(II) Exercises	0.289 (0.122)** [0.089]	0.368 (0.095)** [0.109]	0.212 (0.151) [0.068]	0.321 (0.080)** [0.100]
(III) No smoking	0.017 (0.158) [0.006]	0.088 (0.169) [0.029]	0.023 (0.160) [0.008]	0.096 (0.171) [0.032]
(IV) Drink moderately	-0.006 (0.180) [-0.002]	0.070 (0.180) [0.026]	0.052 (0.203) [0.019]	0.003 (0.202) [0.001]
F or Chi-sq value	7.30	6.65	14.47	12.87
P-value	0.0007	0.0013	0.0007	0.0016
(B) IVs: All High School Classrooms Alone				
Behaviors	(1)	(2)	(3)	(4)
(I) Checkups	0.041 (0.195) [0.013]	0.105 (0.187) [0.029]	0.037 (0.196) [0.013]	0.089 (0.186) [0.025]
(II) Exercises	0.369 (0.085)** [0.110]	0.391 (0.072)** [0.115]	0.325 (0.061)** [0.102]	0.344 (0.052)** [0.106]
(III) No smoking	0.008 (0.195) [0.003]	0.106 (0.183) [0.035]	0.002 (0.192) [0.001]	0.102 (0.181) [0.034]
(IV) Drink moderately	-0.204 (0.138) [-0.073]	-0.069 (0.171) [-0.026]	-0.182 (0.114) [-0.066]	-0.058 (0.165) [-0.022]
F or Chi-sq value	9.45	11.00	9.61	11.22
P-value	0.0021	0.0009	0.0019	0.0008
(C) IVs: Separate General and Vocational High School Classrooms and a First-son Dummy				
Behaviors	(1)	(2)	(3)	(4)
(I) Checkups	0.322 (0.091)** [0.098]	0.267 (0.123)** [0.075]	0.316 (0.065)** [0.100]	0.239 (0.114)** [0.070]
(II) Exercises	0.289 (0.115)** [0.089]	0.345 (0.107)** [0.103]	0.142 (0.207) [0.045]	0.233 (0.164) [0.073]
(III) No smoking	0.071 (0.157) [0.024]	0.149 (0.157) [0.049]	0.139 (0.165) [0.047]	0.160 (0.145) [0.053]
(IV) Drink moderately	0.108 (0.184) [0.039]	0.155 (0.162) [0.055]	0.229 (0.126)* [0.078]	0.165 (0.159) [0.058]
F or Chi-sq value	5.33	4.92	12.08	14.64
P-value	0.0012	0.0021	0.0071	0.0021

Note: Asymptotic standard errors are in parentheses. Average partial effects of one extra year of education are in square brackets.

* Statistically significant at the 10 % level; ** at the 5 % level.

Table 7: Summary of IV Results of Linear Probability Models

Estimated jointly	No	No	Yes	Yes
Job characteristics	No	Yes	No	Yes
(A) IVs: All High School Classrooms and a First-son Dummy				
Behaviors	(1)	(2)	(3)	(4)
(I) Checkups	0.089 (0.056) < 0.051 >	0.051 (0.052) < 0.206 >	0.089 (0.056)	0.050 (0.052)
(II) Exercises	0.082 (0.053) < 0.074 >	0.108 (0.058)* < 0.031 >	0.082 (0.052)	0.108 (0.057)*
(III) No smoking	0.009 (0.053) < 0.990 >	0.033 (0.056) < 0.717 >	0.009 (0.053)	0.033 (0.056)
(IV) Drink moderately	0.000 (0.057) < 0.032 >	0.022 (0.062) < 0.081 >	0.000 (0.057) < 0.010 >	0.022 (0.061) < 0.030 >
F or Chi-sq value	7.30	6.65	14.86	13.80
P-value	0.0007	0.0013	0.0006	0.0010
(B) IVs: All High School Classrooms Alone				
	(1)	(2)	(3)	(4)
(I) Checkups	0.009 (0.066)	0.021 (0.056)	0.009 (0.065)	0.021 (0.056)
(II) Exercises	0.151 (0.074)**	0.164 (0.072)**	0.151 (0.074)**	0.164 (0.071)**
(III) No smoking	0.008 (0.066)	0.042 (0.062)	0.008 (0.065)	0.042 (0.061)
(IV) Drink moderately	-0.089 (0.078)	-0.027 (0.068)	-0.089 (0.077)	-0.027 (0.068)
F or Chi-sq value	9.45	11.00	9.59	11.22
P-value	0.0021	0.0009	0.0020	0.0008
(C) IVs: Separate General and Vocational High School Classrooms and a First-son Dummy				
	(1)	(2)	(3)	(4)
(I) Checkups	0.106 (0.055)* < 0.088 >	0.065 (0.051) < 0.309 >	0.106 (0.055)*	0.065 (0.050)
(II) Exercises	0.082 (0.051)* < 0.203 >	0.094 (0.054)* < 0.064 >	0.082 (0.050)*	0.094 (0.053)*
(III) No smoking	0.025 (0.051) < 0.520 >	0.049 (0.054) < 0.605 >	0.025 (0.050)	0.049 (0.053)
(IV) Drink moderately	0.027 (0.055) < 0.023 >	0.046 (0.059) < 0.101 >	0.027 (0.055) < 0.023 >	0.046 (0.059) < 0.062 >
F or Chi-sq value	5.25	4.90	16.68	15.07
P-value	0.0013	0.0021	0.0008	0.0018

Note: Asymptotic standard errors are in parentheses; P-values of over-identification test statistics are in angle brackets. P-values of separate individual models are in columns (1) and (2); those of joint models are in columns (3) and (4).

* Statistically significant at the 10 % level; ** at the 5 % level.