Schooling, cognitive ability, and health

M. Christopher Auld
auld@ucalgary.ca

and

Nirmal Sidhu
sidhu@ucalgary.ca

University of Calgary
2500 University Dr NW
Calgary, AB, T2N 1N4

Abstract

A large literature documents a strong correlation between health and educational outcomes. In this paper we investigate the role of cognitive ability in the health–education nexus. Using NLSY data, we show that cognitive ability accounts for roughly one quarter of the association between schooling and health. Both schooling and ability are strongly associated with health at low levels but less related or unrelated at high levels. Estimates treating schooling as endogenous to health suggest that most of the correlation between schooling and health is attributable to unobserved heterogeneity, except possibly at low levels of schooling for individuals with low cognitive ability. An implication is that policies which increase schooling will only increase health to the extent that they increase the education of poorly-educated individuals; subsidies to college education, for example, are unlikely to increase population health.

Auld is Assistant Professor and Sidhu a graduate student at the University of Calgary. We thank Dan Gordon, Marjon van der Pol, and Catherine Worthington for helpful comments. Auld thanks the Alberta Heritage Foundation for Medical Research for financial support.

First version: May 2004 **This version:** May 2004

This is a preliminary version, comments welcome.

Internet location of latest version of this document:
http://jerry.ss.ucalgary.ca/cog.pdf

Schooling, cognitive ability, and health

1 Introduction.

Does higher intelligence lead to greater health? How much of the well-documented positive association between schooling and health can be attributed to intelligence? In this paper we present estimates of models of health status focusing on schooling and cognitive ability as key explanatory variables. These estimates tie together two strands of the literature:

- Non–pecuniary effects of education are often considered in both the population health and health economics literatures to be at least as important as effects on labor market outcomes.[1] Perhaps the most commonly discussed and important non-pecuniary effect of education is improvements in health.

- Labor economists have examined the effect of of controlling for “ability bias” in wage regressions by including measures of cognitive ability (Blackburn and Neumark, 1993; Card, 1995). Yet we have only been able to find one paper in the economics literature which reports both a measure of schooling and a measure of cognitive ability in a health equation. It is striking that such bias is rarely discussed in the literature on non-pecuniary effects of education. Just as estimates of the effect of schooling on wages may reflect unobserved ability, estimates of the effect of schooling on health may reflect unobserved ability.

The correlation between health and education is very well-known but largely unexplained. Individuals who are observed to have higher health tend to be better educated, even conditional on other observable sociodemographic characteristics and regardless

[1]See for example Kitagawa and Hauser (1973), Deaton and Paxson (2001), Lochner and Moretti (2001), and Elías (2003).

of how health is measured.[2] The mechanisms through which health and education are related are topics of current research. Kenkel (1991) reports that changes in health behavior associated with higher education cannot explain the major part of the improvement in health associated with higher education, and Lleras-Muney and Lichtenberg (2003) report that more educated individuals are more likely to use recently developed pharmaceuticals. Thus, some but not all of the large association between schooling and health can be attributed to differences in health-related behavior, leaving the remainder of the correlation unexplained. A possible reconciliation is that the association between health and education is not primarily causal but rather reflects unobserved causes of both outcomes. For example, Fuchs (1982) argued that individuals with high discount rates will tend to invest in both less health and less education. Cognitive ability could be another such third factor.

Failure to control for cognitive ability in health equations biases the estimated effect of schooling. Further, the effect of cognitive ability is of direct interest in part because it provides another test of Grossman's (1972) hypothesis that the correlation between schooling and health obtains because schooling improves health production efficiency. If that hypothesis is correct we should also observe more cognitively able individuals to be healthier, all else including schooling equal, because these individuals ought to be better able to process diverse information on the relationships between various behaviors and treatments and likely health outcomes. Alternately, if the efficiency argument fails and the observed correlation between health and education obtains because of factors such as the discount rate or genetic endowments, then we should not expect to see more cognitively able individuals to also be healthier. In this case we will observe a correlation between schooling and health, but an exogenous increase in schooling will not lead to an increase in health.

The well-documented correlation between education and health could then be spurious and policies which increase education levels will fail to improve population health. To the best of our knowledge the only paper presenting a multivariate analysis of health which reports estimates including both schooling and measures of cognitive ability is Hartog and Oosterbeek (1998). Hartog and Oosterbeek display ordered probit estimates

[2]See Grossman and Kaestner (1997) for a review of the extensive literature documenting the association between health and education.

of self-reported health status which suggest that mathematical (but not verbal) ability is correlated with better health, holding schooling constant. Other evidence consistent with a causal effect of cognitive ability on health includes an association between cognitive ability in old age and greater life expectancy (Neale *et al.*, 2001) and an association between childhood IQ scores and life expectancy (Hart *et al.*, 2003; Whalley and Deary, 2001). Snowden *et al.* (1999) show that idea density in Nuns' autobiographies written in the 1930s and 40s predicted mortality in the 1990s. These studies treat schooling as exogenous or do not control for schooling.

Instrumental variable methods can recover the causal effect of schooling on health even when cognitive ability is not observed so long as cognitive ability is uncorrelated with the instruments. Berger and Leigh (1989) report a large effect when schooling is instrumented with selected parental characteristics. Similarly large effects are discovered by Arendt (2001, 2004), Adams (2001), and Lleras-Muney (2001), all of whom use changes in compulsory education laws as instruments. Arkes (2001) uses local unemployment rates as instruments and also reports that the effect of schooling on health survives when schooling is treated as endogenous to health. Thus, a small literature consistently finds that the correlation between health and education is mostly causal.

We use data from the National Longitudinal Survey of Youth (NLSY) to investigate the role of cognitive ability in the health–education nexus. Our results suggest that cognitive ability accounts for about one–quarter of the relationship between schooling and health, holding a variety of sociodemographic characteristics constant. Both schooling and cognitive ability are more highly correlated with health at low levels; increases in schooling for individuals with low schooling and cognitive ability are associated with significant improvements in health, whereas individuals in the top half of the distribution of cognitive ability who obtain at least a high school education benefit little from further increases in schooling. When we instrument for schooling we find that cognitive ability but not schooling appears to cause better health outcomes, but we are not confident that our instruments are valid and offer these results somewhat speculatively. We suggest these results are consistent with a lower causal effect of schooling on health than is widely believed. We show that Berger and Leigh's (1989) results hinge on their assumption that cognitive ability cannot directly effect health: When we relax that assumption the causal effect of schooling is much smaller. Since changes in compulsory schooling laws induce variation in schooling only for individuals likely to obtain low levels of schooling, using such changes as instruments recovers local causal effects for individuals with low levels of schooling. The results presented in this paper suggest that it may be misleading to generalize these results to individuals with higher levels of education.

2 Data.

We use data from the NLSY, a large longitudinal dataset which oversamples minority and economically disadvantaged youths. The complete dataset includes information on 6,111 randomly sampled youths, 5,295 oversampled youths, and 1,280 youths in active military service. As described below, we draw information from primarily the 1979 and 2000 surveys, only including respondents who completed the 2000 survey. After removing the military subsample, attrition and deleting responses missing critical items, our sample consists of 6,385 respondents. In some specifications we use finely meshed schooling–intelligence cells and drop 448 more observations from cells with fewer than 50 individuals.

Summary statistics for our sample are displayed in Table 1. Our main health measure is the response to the question, "Are you limited in the kind of work you are able to do by your health?" This is not an ideal measure. It is binary and thus intrinsically obscures much of the variation in health status, and it is self–reported and may be subject to systematic reporting error. On the other hand, it is a commonly used measure in applied work and thus affords comparability with other studies. Further, Bound (1991) has argued in the context of labor supply that the positive and negative biases roughly balance, and Baker *et al.* (2001) suggest objective but self-reported measures are also subject to large measurement error. We also use the related question, "Are you limited in the *amount* of work..." and also, for the subset of respondents for whom the questions have been posed, self-reported general health from the SF12 battery.

Our primary measure of cognitive ability is Armed Forces Qualification Test (AFQT) scores, adjusted for age and for years of schooling at time of testing. Hansen *et al.* (2003) estimate that a year of schooling increases AFQT score by 0.17 standard deviations. We adopt this estimate and use as our measure of cognitive ability the residuals from the regression of (AFQT $-$ $0.17S$) on cohort dummies, where AFQT is standardized AFQT score and S is years of schooling at the time of testing. Our measure should then reflect the innate characteristic and not intelligence produced by schooling. We also report alternate specifications using the constituent scores from the subscales of the Armed Services Vocational Aptitude Battery (ASVAB) and the first principal component thereof which is often interpreted as a measure of "g," general intelligence, in the psychometric literature (Carroll, 1997).

3 Analytical and empirical framework.

3.1 Theory.

In this section we briefly discuss theoretical issues in the relationship between health and schooling. We begin by observing that the canonical Grossman (1972) health demand model can be easily extended to include cognitive ability. Grossman assumes that the amount of health produced I for a given level of inputs (x) may depend on schooling S,

$$I = f(x; S). \tag{1}$$

Grossman shows that an individual with more schooling optimally maintains a higher stock of health capital. The idea is that schooling may lead to enhanced ability to acquire and understand diverse information on the relationships between various behaviors and health outcomes. The same argument can be made with respect to cognitive ability: If two individuals have the same schooling, we should expect under this hypothesis that the more able of the two will be healthier. We can express idea this by modifying Grossman's process,

$$I = g[x; \mu(S, C)], \tag{2}$$

where C is cognitive ability and $\mu(\cdot)$, a function increasing in both its arguments, indexes efficiency in producing health. All of Grossman's comparative dynamics arguments with respect to schooling then carry through to cognitive ability.

3.2 An empirical model with endogenous schooling.

For our empirical work we are interested in developing these arguments further in a stochastic setting. The literature on the closely related issue of the relationship between schooling and earnings is well-developed and can be readily modified to model the causal effect of schooling on health. Consider the econometric framework presented by Card (1995) and developed further by Heckman and Vytlacil (1998). Card assumes individuals solve

$$\max_{S} \quad U(Y(S), S) = \log Y(S) - \phi(S) \tag{3}$$

where Y is earnings and $\phi(S)$ is the cost of schooling. Generalize to the case where health (H) enters the utility function,

$$\max_{S} \quad U(Y(S), H(S), \phi(S)) = \log Y + \alpha \log H - \phi(S). \tag{4}$$

Optimal schooling satisfies the condition,

$$\frac{Y'(S)}{Y(S)} + \frac{H'(S)}{H(S)} = \phi'(S). \tag{5}$$

Linearize marginal costs and benefits,

$$\frac{Y'(S)}{Y(S)} = a_t + \eta_1 C_t - k_1 S \tag{6}$$

$$\frac{H'(S)}{H(S)} = b_{1t} + \eta_2 C_t - k_2 S \tag{7}$$

$$\phi'(S) = r_t + \eta_3 C_t + k_3 S \tag{8}$$

where a_t and b_{1t} are random variables with some joint distribution across individuals, C_t is cognitive ability and η are parameters specifying how ability affects the marginal benefits and costs of an additional year of schooling, and k_i are non-negative constants which sum to k. The observed joint distribution of schooling and health is then given by,

$$S_t = \frac{1}{k}[(a_t + \alpha b_t - r_t) + \eta C_t] \tag{9}$$

$$H_t = b_{0t} + \eta_2 C_t + b_{1t} S_t \tag{10}$$

where $\eta = (\eta_1 + \alpha\eta_2 + \eta_3)$ and b_{0t} is an individual-specific constant of integration. Following Heckman and Vytlacil (1998), a quadratic term in the health equation has been suppressed (by setting $k_2 = 0$)

3.3 Estimation issues.

Consider the regression of log health status H on schooling,

$$H = \beta_0 + \beta_1 S + \text{noise}. \tag{11}$$

Suppose that cognitive ability is *not* held constant in this regression. Observe that (10) is a correlated random coefficients model for health: the random intercept is $(b_{0t} + \eta_2 C_t)$ and the random slope coefficient is b_{1t}. Express (10) in terms of deviations from mean population values,

$$H_t = (\bar{b}_0 + \epsilon_{0t} + \eta_2 C_t) + (\bar{b}_1 + \epsilon_{1t})S_t, \tag{12}$$

to observe that single-equation methods will generally not recover a structural parameter. OLS estimates of the coefficient on schooling in regression (11) are centered on

$$\begin{aligned} E\beta_1 &= \bar{b}_1 + \frac{1}{\sigma_S^2}[\eta_2 \operatorname{Cov}(S_t, C_t) + \operatorname{Cov}(S_t, \epsilon_{0t}) + \operatorname{Cov}(S_t, \epsilon_{1t} S_t)] && (13) \\ &= \bar{b}_1 + \lambda_{SC} + \lambda_{S0} + \lambda_{S1}. && (14) \end{aligned}$$

where σ_{ij} is the covariance between i and j and

$$\begin{aligned} \lambda_{SC} &= \frac{\eta_2}{k\sigma_S^2}[\sigma_{Ca} + \alpha\sigma_{C1} - \sigma_{Cr} + \eta\sigma_C^2] && (15) \\ \lambda_{S0} &= \frac{1}{k\sigma_S^2}[\sigma_{a0} + \sigma_{01} - \sigma_{0r} + \eta\sigma_{C0}] && (16) \\ \lambda_{S1} &= \frac{1}{\sigma_S^2}[E(S_t^2 \epsilon_{1t})]. && (17) \end{aligned}$$

The OLS estimate diverges from the mean causal effect of schooling across the population ($\bar{b}_1$) when the λ's are not zero. Bias arises for three reasons. First, cognitive ability will generally be correlated with unobserved determinants of income and health benefits to schooling, and cognitive ability may be associated with higher or lower opportunity costs of schooling ($\lambda_{SC} \neq 0$). Second, unobserved determinants of health (ϵ_0) will generally be correlated with unobserved determinants of income, idiosyncratic returns to health, the opportunity cost of schooling, and cognitive ability ($\lambda_{S0} \neq 0$). Finally, further bias is introduced if idiosyncratic returns to schooling ϵ_1 are stochastically dependent on any of the determinants of schooling ($\lambda_{S1} \neq 0$). Generally we might expect all of the λ's to be positive, so long as the net effect of higher unobserved ability is to increase schooling. Notice that health regressions generally do not recover structural effects of schooling on health even if unobserved determinants of health and health returns to schooling are independent of schooling, because they will nonetheless generally be stochastically dependent on unobserved determinants of income ($\sigma_{a0} \neq 0$, $\sigma_{a1} \neq 0$). The terms involving a are the only difference between this model and the wage model considered by Card (1995).

If cognitive ability is held constant then the coefficient on schooling is still generally biased, but by an amount purged of the covariances between cognitive ability and the health and income costs and benefits of schooling. The bias terms in this case become,

$$\lambda'_{SC} = 0 \tag{18}$$

$$\lambda'_{S0} = \frac{1}{k\sigma_S^2}[\sigma_{a0} + \sigma_{01} - \sigma_{0r}] \tag{19}$$

$$\lambda'_{S1} = \frac{1}{\sigma_S^2}[E(S_t^2 \epsilon_{1t}|C)]. \tag{20}$$

We expect the bias to be smaller when cognitive ability is held constant. However, single-equation methods still do not recover structural parameters because (1) unobserved determinants of health levels may be correlated with unobserved determinants of the costs or benefits of schooling and (2) unobserved returns to schooling may covary with unobserved determints of health levels.

3.4 Instrumental variables.

We conclude that single-equation estimates of health regressions will generally fail to recover a mean causal effect of schooling on health, even when cognitive ability is held constant. Suppose we have available a vector of instrumental variable Z_t which affects the marginal cost of schooling,

$$r_t = Z_t\gamma + V_t \tag{21}$$

where $\gamma \neq 0$ and $E[V_t|Z_t] = 0$ such that optimal schooling is linear in Z

$$S_t = Z_t\pi + \omega_t \tag{22}$$

where $Z\pi = Z\gamma/k$ and $\omega_t = (a_t + \alpha b_t - V_t)/k$. If heterogeneity in returns to schooling is ignored ($\text{Var}(\epsilon_{1t}) = 0$) then the health equation (10) can be consistently estimated under the usual assumption that

$$E[\epsilon_0 + \eta_2 C|Z] = 0 \tag{23}$$

if cognitive ability is not held constant or the weaker condition

$$E[\epsilon_0|Z] = 0 \tag{24}$$

if cognitive ability is conditioned out, along with the usual rank conditions.

Estimation is more problematic when we allow for heterogeneity in health returns to schooling because $(\epsilon_{1t}S_t)$ may be correlated with Z_t even if Z_t and ϵ_{1t} are independent. We discuss estimation strategies for this case in the following subsection. It is worth briefly discussing the interpretation of conventional instrumental variable models when health returns to schooling vary across observationally identical individuals. Following Imbens and Angrist (1994), suppose that heterogeneity can be grouped into G categories with individuals within a category having the same preference and ability parameters. If an exogenous shock causes schooling in group $g \in \{1, ..., G\}$ to change by ΔS_g, then instrumental variables estimates of "the" causal effect of schooling on health converge to

$$\operatorname{plim} \tilde{\beta}_1 = \frac{E(b_{1g}\Delta S_g)}{E(\Delta S_g)}, \tag{25}$$

where b_{1g} is the causal effect of a unit change in schooling on health for individuals in group g. Thus, the IV estimate recovers a weighted average of causal effects and the weights (ΔS_g) reflect how much schooling changes in response to the exogenous change for each individual.

Finding instruments for schooling is difficult. We follow Arkes (2001) and use local unemployment rates as instruments, but we found that unemployment rates had extremely little explanatory power after controlling for our rich set of characteristics (F=0.54). We require more instruments to generate estimates with reasonable properties. We follow Berger and Leigh (1989) and many other papers in the labor economics literature and use parental education and, in some specifications, occupation as excluded instruments. However, we differ from some previous papers in that we do not assume that cognitive ability can have no direct effect on health.

3.5 Econometric models.

We examine the NLSY cohort at two times: once in 1979 when they are aged 15 through 22, and again in 2000 when they are 36 through 43. We also use the ability test information administered in 1980 and some health survey responses from 1998. We choose the earliest and latest dates available at the time of this writing because we wish to condition on early experiences and because we would like to use the more extensive health information available only in the latest (1998 and 2000) waves of the survey. Let H_i^* denote a latent measure of health status in period i, $i \in \{0, 1\}$, where period 0 denotes the individual's experience in early adulthood and period 1 denotes middle–age. The principal equation we estimate takes the form

$$H_1^* = X_1\beta_1 + X_0\beta_2 + \beta_3 H_0 + \mu(S, C; \theta) + u. \tag{26}$$

We assume current health H_1 depends on current and past characteristics, on past health, a function of schooling S and intelligence C denoted $\mu(\cdot)$, and on a disturbance term u. We consider several forms for $\mu(\cdot)$, from the linear and separable case to flexible forms imposing very little structure. Robustness to functional form is important to consider because of very strong sorting across schooling levels by cognitive ability Heckman and Vytlacil (2001). We include past health and past characteristics to remove covariation between adult health and schooling which arises because of genetic or early childhood influences. For example, low birth weight affects both cognitive and physical development, which may lead to both low schooling and poor adult health (Friedlandera *et al.*, 2003).

In some specifications we include current and past family income because we wish to ascertain whether the mechanism through which schooling or cognitive ability affects health is through increased income. We do not attempt to confront the difficulty that both income and past health may be endogenous except inasmuch as we present estimates conditioning on and not conditioning on these variables.

We present estimates that treat schooling as exogenous, conditional on (X_0, X_1, S, C), and estimates which treat schooling as an endogenous regressor. Treating schooling as exogenous has the advantages that we can characterize the relationship between schooling, intelligence, and health flexibly. But if unobserved determinants of the rate of decay of health are correlated with unobserved determinants of schooling then these associations do not recover causal relationships. The major disadvantage of attempting to control for potential endogeneity problems is the need to impose more structure on $\mu(\cdot)$ and in the form of debatable exclusion restrictions.

As discussed in section 3.4, we use certain parental characteristics and local unemployment rates as excluded instruments when estimating (26). We first present standard two-step estimates. The reduced form for schooling is assumed to be linear in the exogenous covariates,

$$S_t = X_t\gamma_0 + Z_t\gamma_1 + \omega_t, \tag{27}$$

where $X = (X_0, X_1, H_0, C)$ and Z are the excluded instruments. We estimate linear probability specifications of (26) using a feasible two-step generalized method of moments (GMM) approach. Robustness of the linear probability specification is assessed by comparing these estimates with those from an instrumental variable probit approach Newey (1987) which account for the nonlinearity induced by the binary outcome in (26). These estimates recover the causal effect of schooling on health ignoring heterogeneity in returns to schooling (*i.e.*, $b_{1t} = b_1\ \forall t$) under assumption (23) or (24) depending on whether cognitive ability is conditioned out.

We allow for parameter instability using correlated random coefficient models. Garen (1984) invokes the assumption that the random slope and intercept in the health equation are linear in unobserved determinants of schooling,

$$E[\epsilon_{0t}|S_t, Z_t] = \delta_0\omega_t \tag{28}$$

$$E[\epsilon_{1t}|S_t, Z_t] = \delta_1\omega_t \tag{29}$$

such that

$$E[H_{1t}|S_t, Z_t] = \bar{b}_0 + \bar{b}_1 S_t + \delta_0\omega_t + \delta_1 S_t\omega_t \tag{30}$$

Estimation proceeds by applying OLS to the equation above after replacing ω_t with $\hat{\omega}_t$, the residuals from OLS estimation of (27).

Wooldridge (2003) presents an alternate estimation strategy which has the advantage that it allows characterization of the effect of covariates on the mean causal effect of interest. Specify

$$b_{1t} = \bar{b}_1 + \epsilon_{1t} \tag{31}$$

$$= \bar{b}_1 + (X_t - \psi)\theta + v_t, \tag{32}$$

where $\bar{b}_1$ is the average partial effect of schooling on health for an individual with average characteristics, ψ is a vector of unconditional means $\psi_j = E(X_j)$, and θ is a vector describing how the effect of schooling on health varies X. In addition to the standard assumptions for consistency of estimates of (26) using (X, Z) as instruments, we require the assumptions that $E[v_t|X, Z_t]$ does not depend on Z and that $E(v_t S_t|X_t, Z_t)$ does not depend on (X, Z). The first condition implies that we can vary schooling exogenously while holding the causal effect of schooling constant. The second condition implies that heterogeneity only affects estimation of the constant, not the slope parameters, in (26). Under this condition the constant and correlation between ω and v are not jointly identified. It is important to see that the health return to schooling and the level of schooling can be arbitrarily correlated, but this correlation cannot itself depend on observed characteristics. We estimate this model by applying feasible two-step GMM to the equation

$$H_1 = X\beta + \bar{b}_1 S + S(X - \bar{X})\theta + \text{noise} \tag{33}$$

using $(X, \hat{S}, \hat{S}X)$ as instruments, where $\hat{S}$ are the predicted values from OLS estimation of (27).

The second step in the procedure due to Garen (1984) does not generate a consistent estimate of the covariance matrix, so we base inference on a nonparametric bootstrap with 1,000 replications. The second-step covariance matrix produced by Wooldridge's (2003) estimator requires only correction for heteroskedasticity of unknown form when estimated by standard two-stage least squares, such that valid inference can be based on our GMM estimates without resampling or corrections to the covariance matrix.

4 Econometric results.

In this section we discuss the relationships between the health, schooling, and cognitive ability of the NLSY respondents.

4.1 Descriptive statistics.

We begin by examining simple cross tabulations. Table 2 shows that respondents who are unhealthy have on average about one year less schooling, are about a third of a standard deviation lower in cognitive ability, and have less than half the family income of healthy respondents. Nothing can be inferred about causality from these results, but clearly poor health is unconditionally correlated with low schooling, low cognitive ability, and low income. Table 3 breaks these results down by years of schooling. Within each level of schooling health tends to increase with cognitive ability, and withing each quartile of ability health tends to increase with years of schooling. The table also illustrates the "ability sorting" problem emphasized by Heckman and Vytlacil (2001). High ability and low schooling cells are sparsely populated or unpopulated, similarly there are few respondents with low ability and high levels of schooling. It is difficult to disentangle the effect of ability from the effect of schooling because of this problem.

4.2 Parametric models treating schooling as exogenous.

Turning to regression models, Table 4 displays estimates of probit models of health status as we vary the set of included covariates.[3] Including only cognitive ability shows that a one standard deviation increase in ability is associated with 4.5% points lower probability of a health limitation. This is a very large effect given that the proportion of respondents with a limitation is 10%. One year of schooling is unconditionally associated with about 2% lower probability of a limitation. When we include both cognitive ability and schooling, the effect of cognitive ability falls to just under 3% and the effect of a year of schooling to 1.4%. A standard deviation change in ability has roughly the same effect on health as a two year change in schooling.

Model (4) in Table 4 reproduces the empirical regularity that schooling is associated with better health even after holding constant a wide variety of characteristics. Comparing columns (2) and (4) shows that holding age, background characteristics, past health, and current and past marital status and family characteristics constant has roughly the same effect on the schooling coefficient as holding cognitive ability but nothing else

[3] In Table 4 and subsequent tables we display only selected parameter estimates, and in models in which we treat schooling as endogenous we do not display first-stage results. Complete estimation results are available from the authors on request.

constant. When cognitive ability is also held constant, the effect of schooling falls modestly to 1.2% and the effect of a standard deviation change in ability remains roughly equivalent to a two year change in schooling.

Model (6) in Table 4 relaxes the assumption that the effects of schooling and cognitive ability are separable (in the probit index). The interaction term is positive and significant, suggesting that the effect of schooling (ability) is lower for individuals with high levels of ability (schooling).[4] Model (7) shows that the estimates are not sensitive to whether past health, which is potentially endogenous, is included as a covariate. Model (8) shows this result changes little when condition on past and present characteristics. Finally, model (9) also conditions on past and present family income. If the mechanism through which schooling or ability affects health is through their effect on income, then we would expect the coefficients on schooling and ability to fall to zero when we hold income constant. Since we find that these coefficients fall only very modestly when we condition on income, we conclude that it is not the case that the more able are healthier only, or even importantly, because they fare better in the labor market.

4.3 Semiparametric models.

We then estimated a number of models in which we make no parametric assumptions over the partial relationship between health and schooling and ability. We included a full set of dummies for years of schooling and approximated an unknown form for ability using a step function with steps at each quartile. Figure 1 shows results from a model including years of schooling dummies fully interacted with the ability quartile dummies. The effects of both schooling and ability are revealed to by highly nonlinear. Schooling and health are highly associated for individuals in the lowest ability quartile, but the effect decreases as we move to higher ability levels. Low ability individuals greatly benefit from increases in ability, and similarly increasing years of schooling at low levels substantially increases health. The effect of either schooling or ability for able and highly educated individuals is essentially zero.

[4]Note that in Table 4 and elsewhere we have calculated marginal effects on interaction terms in nonlinear models (such as Probit) using methods and software discussed by Ai and Norton (2003) and Norton *et al.* (2004). The estimates in Table 4 show numerically calculated cross-derivatives of the predicted probabilities, $\partial \Pr(y=1|X,S,C)/\partial S \partial C$ as opposed to the marginal effect of the interaction term $\partial \Pr(y=1|X,S,C)/\partial(S*C)$. These two expressions are not equivalent except in linear models.

4.4 Stratification by sex and race.

Lochner *et al.* (1999) show that the wage returns to ability vary substantially across men and women and across whites and minorities. The estimates presented in Table 4 may then be misleading because they average over sex and racial groups. Table 5 shows estimates of probit regressions of health limitation on schooling, ability, their interaction, and all past and present characteristics but income separately by gender and by race. The health returns to ability and schooling are strikingly similar across these strata. Formally, we tested and failed to reject the hypotheses that the coefficients on ability, schooling, and their interaction were equal across men and women (p=0.58), equal across whites and non-whites (p=0.32), and jointly equal (p=0.47).

We conclude that, unlike in the context of wage regressions, we may pool men and women and whites and minorities when running health regressions.

4.5 The health and intelligence of high school graduates.

The results in Tables 4 and 5 are plagued by the ability sorting problem. Since cognitive ability and schooling are so highly correlated, the coefficients on these variables are also highly correlated, and if we are estimating the effect of schooling with bias we are also likely to find misleading results with respect to ability. Table 3 suggests one way to increase confidence in our results: More than 40% of the sample has exactly a high school education, so we may examine the relationship between health and ability in this subsample while retaining a reasonably large number of observations.

Table 6 shows probit models for health status estimated using only the high school subsample. Model (1) shows that a high school graduate with one standard deviation greater intelligence has 4.2 percentage points lower probability of reporting a health limitation (t=5.1). This estimate is only slightly reduced when the contemporaneous controls are included (model 2, t=3.7) or the full set of controls but income is included (model 3, t=3.5) and falls by about a third when income is also included (model 4, t=2.5).

Model (5) in Table 6 relaxes the assumption that health is linear in ability (in the probit index) by including three dummies indicating ability is in the second, third, or fourth quartiles. The results show that moving from the first to the second ability quartile or from the third to the fourth quartile reduces probability of a health limitation more so than moving from the second to the third quartile. Models (6) and (7) show results from a finer decomposition into deciles, with income not included and included. Figure 5 graphs these results.

In the subsample of respondents with no more and no less schooling than a high school diploma, more cognitively able individuals are substantially less likely to report a health limitation. The magnitude of the effect is similar to estimates from the entire sample. This result cannot be attributed to bias in estimating the effect of schooling on health.

4.6 Models treating schooling as endogenous.

The results we have reported so far treat schooling as exogenous to health. Unobserved characteristics, such as the discount rate, which are correlated with both schooling and the rate of decay of health in adulthood will bias our estimates. In this section we report on our efforts to purge the effect of schooling of such bias. These estimates are somewhat questionable because we do not have a randomized instrument. Instead, we follow much of the labor economics literature in using certain background characteristics as instruments for schooling. In most specifications we use mother's and father's years of schooling, a set of dummies indicating father's occupation, and local unemployment rate in 1979 as excluded instruments. As discussed in Section 2, our estimates consistently recover weighted averages of mean causal effects under the strong assumption that variation in these parental characteristics induce variation in the marginal cost of schooling and are uncorrelated with unobserved components of ability and idiosyncratic returns to schooling. We do not attempt to estimate models with interaction terms (which must also be treated as endogenous) in this section. Instead, we stratify into groups with high school or less education and those with greater than high school to investigate nonlinearities.

Table 7 shows feasible two-step GMM estimates. Model (1) shows estimates treating schooling as exogenous for comparison, in which case the GMM estimator is heteroskedastic OLS. Comparing model (5) in Table 4 and model (1) in Table 7, we observe that the GMM estimate of the linear probability model produces point estimates which are very similar to marginal effects from probit regression. Model (2) shows two-stage least squares estimates which do not account for heteroskedasticity induced by the binary dependent variable. These estimates are similar to the two-step feasible GMM estimates displayed in Model (3). When schooling is treated as endogenous, it is estimated to have neither an economically or statistically significant effect on health. Diagnostic tests for model (3) suggest the model is reasonably consistent with the data: the first-stage F-statistic is 31.2, suggesting the instruments explain adequate variation in schooling. A test of the overidentifying restrictions, taking the form of Hansen's J-statistic to account for the possibility of heteroskedasticity, yields a p-value of 0.597, indicating that the instruments are valid. A "differences" or "C" test testing the null that intelligence is exogenous yields a p-value of 0.087, suggesting, if somewhat tenuously, that it is valid to treat intelligence as exogenously assigned in these models.

In model (4) of Table 7 we remove father's occupation from the list of excluded instruments. The occupation dummies have modest explanatory power but use many degrees of freedom, which may produce low power in our overidentifying restrictions test. The results on the parameters of interest change little, we still fail to reject the exclusion restrictions (p=0.18), and the first–stage F statistic on the excluded instruments rises to 77.9. It does not seem the results are an artifact of the questionable occupation dummies.

We next replicated two specifications from previous research. In column (5) we report estimates of a model in which cognitive ability is used as an instrument for schooling, that is, it is assumed to have no effect on health except indirectly through schooling. This model is very similar to the specification of Berger and Leigh (1989). The results seemingly suggest that schooling causes substantial increases in health. One year of additional schooling decreases the probability of a health limitation by 2.3 percentage points (t=6.9), which is quite similar to the estimated effect of schooling in the single-equation probits reported in Table 4. Similarly, if we leave ability out of the model altogether as in specification (6), we apparently find that a year of schooling causes a 1.5 percentage point decrease in probability of a health limitation (t=2.9). Drawing on the arguments presented in Section 2, we believe that both of these specifications are highly misleading. When ability is used as an instrument the correlation between ability and health is attributed entirely to a causal effect of schooling on health and thus the causal effect of schooling is substantially biased upwards. When we ignore ability altogether, condition (23) requires that parent's education and occupation are uncorrelated (conditional on other covariates) with respondent's cognitive ability, which is implausible. Thus, we have reason to believe condition (23) fails and thus that the estimates are inconsistent. However, we note that the overidentifying restriction tests on these models fail to reject the null that the instruments are valid.

In the final three columns of Table 7 we show estimates for the high school and less subsample and the greater than high school subsample. Although imprecisely estimated, the effect of schooling is much larger in magnitude in the low education subsample, with a point estimate of one year of schooling reducing probability of a health limitation by 4.1 percentage points as opposed to 2.6 percentage points in the high education subsample. One standard deviation increase in ability reduces probability of a health limitation by 3.3 percentage points in the high school or less subsample (t=2.7) contrasted with 0.007 percentage points (t=0.6) in the greater than high school sample. However, note that the first-stage F-statistics for these models are only 4.3 and 3.7, such that our instruments are quite weak after stratification. Variation in parental characteristics is highly correlated with the decision to undertake post-secondary explanation but less correlated with infra-marginal schooling decisions. Further, the overidentifying restrictions are rejected in model (8). We re-estimated model (8) excluding the occupation dummies. Model (9) passes the overidentifying restrictions test and the first-stage F rises somewhat to 7.2, but the point estimate on schooling changes sign.

In Table 8 we replicate some of these models using a two-stage probit procedure which captures the nonlinearity induced by the binary health outcome. The results are similar to those using feasible GMM: The effect of schooling is smaller than in single equation estimates and is statistically insignificant, ignoring ability or using it as an instrument overstates the causal effect of schooling, and both schooling and ability have larger effects for the low ability/schooling subsample. The GMM results do not appear to be an artifact of the questionable linear probability specification.

Estimates from correlated random coefficient models are presented in Table 9. When we do not condition on cognitive ability, estimates of equation (33) suggest that an individual with average observed and unobserved characteristics experiences a 1.9% decrease in the probability of a health limitation when schooling is exogenously increased by one year. Similarly, estimates of (30) show a 1.6% decrease. Unobserved determinants of schooling ω are associated with better health (t=1.9) and diminish the effect of schooling on health (t=2.6). When we condition on cognitive ability, the effect of schooling on health for the average person is roughly halved and loses statistical significance. However, the average effect obscures substantial variation with cognitive ability: A one standard deviation decrease in cognitive ability increases the effect of schooling health by about 1% (t=2.0). In other words, the causal effect of schooling on health is substantial only for individuals with low cognitive ability. After conditioning on cognitive ability, estimates using Garen's method suggest that the effect of schooling on health is not statistically or economically significant for the average respondent (t=0.6). Even after conditioning on ability, respondents with higher than expected schooling are in better health (t=2.0) and experience a lower causal effect of schooling on health (t=2.4). An unexpected result from either estimation strategy is negative sorting into schooling with respect to health returns: Respondents who gain the most health from additional schooling are likely to obtain *lower* schooling. A possible explanation is that individuals who would gain the most from additional schooling also have the highest opportunity costs of obtaining additional schooling.

On the basis of the results in Tables 7, 8 and 9 we are skeptical that the large association between schooling and health we reported in Table 4 and oft-reported elsewhere in the literature largely reflects a causal effect of schooling on health. When schooling is instrumented *and* we allow for the possibility that intelligence affects both schooling and health, we fail to find a statistically significant impact of schooling on health. Of course, these results could attributed to invalid instruments, but note that for our IV estimates to be biased towards finding a causal effect which is too small it would have to be the case that parental schooling is *negatively* correlated with unobserved determinants of health. It seems more likely that our IV estimates are biased towards finding a causal effect which is too large.

An important caveat is that the evidence *is* consistent with a substantive causal effect of schooling on health that diminishes rapidly with both level of schooling and innate cognitive ability. We note that equation (25) implies that our estimate of "the" effect of schooling on health depends critically on which respondents are induced to change schooling choices because of variation in parental characteristics. If, as some of our results suggest, the instruments chiefly affect the decision over whether to obtain any post-secondary education, we ought to conclude that the decision to undertake schooling beyond the high school level has at most modest effects on health.

4.7 Further robustness checks.

In this section we report on models in which we have used different measures of health or different measures of cognitive ability.

Table 10 shows estimates of ordered probit models of self-reported general health status. General health status may be a better measure of health status than the indicator for health limitations: It obscures less of the variation in health status, and is not subject to the problem that individuals may consider themselves employment limited differentially depending on their occupation. Columns (1) and (2) show that when only ability or only schooling is included in the model, either is a highly significant predictor of general health. The relative magnitudes are similar to those reported in Table 4, with one standard deviation in ability producing roughly the same effect on health as two years of schooling. Adding a complete set of covariates and the interaction of schooling and health, model (3), recovers the same pattern as in Table 4, with low ability and schooling individuals benefiting more from incremental gains than high ability and schooling individuals. Controlling for past and present income, model (4), does not appreciably affect the estimates. In an analogous set of (unreported) models, we replicated Table 4 using "are you limited in the amount of work you can do..." rather than "are you limited in the type of work you can do..." as our health measure. The results were nearly identical.

Finally, in Table 11 we report on health models with varying our measure of cognitive ability. For each measure of ability, we report probit estimates for both separable and interacted models. Notice that the units of the ability measures are not comparable so that the magnitudes of the coefficients are not directly comparable, but the t–statistics may be interpreted as indexing the amount of residual variation explained. The first four models show it makes little difference to our estimates whether we adjust AFQT scores for age and schooling at time of testing (as we have done in all estimates up to this point), or simply use raw AFQT scores, or use the first principal component of ASVAB measures (“g”). This result is not surprising given that eventual schooling and many other covariates are in our model. Note that g is a slightly stronger predictor of health than AFQT scores.

The results are reasonably similar across the constituent scales of the ASVAB battery. Increased cognitive ability is always statistically significantly associated with better health and the interaction effect suggests schooling and ability affect health the most at low levels of schooling and ability. Amongst the subscales, Arithmetic Reasoning, Work Knowledge, Paragraph Comprehension, Numerical Operations, and Mathematical Knowledge are comparable to general intelligence in predicting health. The others, particularly Auto and Shop Information, are weaker.

We conclude from these exercises that our key results reported in the previous subsections are robust to these alternate measures of health and to these alternate measures of cognitive ability.

5 Conclusions.

Respondents to the National Longitudinal Survey of Youth who are more cognitively able are also healthier, holding constant a variety of past and present sociodemographic characteristics, including past health. An increase in cognitive ability of one standard deviation is associated with an increase in health comparable to about two years of schooling, and about one–quarter of the association between schooling and health can be attributed to cognitive ability. Both cognitive ability and schooling are highly associated with health at low levels but weakly related to health at high levels. Notably, years of schooling beyond high school contribute very little to health at the margin. These results are robust to different measures of health, to different measures of cognitive ability, to stratification by sex or by race, to econometric estimation strategy, and they cannot be substantially attributed to the effects of schooling and ability on labor market outcomes such as income.

When we treat schooling as endogenous to health, the effect of schooling diminishes and loses its statistical significance. Schooling is, however, much more strongly associated with health (albeit not statistically significantly) among individuals with no greater than high school education than among individuals with post-secondary education. Estimates of correlated random coefficient models suggest the causal effect of schooling on health is greatest for individuals with low cognitive ability and that much of the association between schooling and health can be attributed to unobserved traits, for example the discount rate, rather than a causal effect.

A key implication of the findings is that an exogenous increase in schooling will have an effect on health only for individuals who obtain low levels of schooling, particularly low ability individuals. Policies which further increase education among the relatively well-educated, for example policies which increase the probability an individual will complete a college degree, are unlikely to have substantial health effects.

Some of these results seem to conflict with previous results which suggest that the causal effect of schooling on health is large. We offer the following reconciliation: First, we showed that statistical models which either ignore ability or use it as an instrument for schooling dramatically over-estimate the causal effect of schooling on health. Second, our results suggest that the causal effect of schooling on health may be large for individuals with low ability and low levels of schooling. Papers which use changes in compulsory education laws as instruments for schooling recover local average effects mostly for such individuals. Thus, there is no conflict: An exogenous increase in schooling causes better health only among poorly educated individuals.

We close emphasizing an important limitation of some of our estimates. Our structural models are identified using family background characteristics as instruments for schooling. This strategy hinges on questionable assumptions, particularly in models in which we require that the instruments are uncorrelated with not only the level of health but also the idiosyncratic component of the health return to schooling.

References

Adams, J. (2001) "Educational attainment and health: Evidence from a sample of older adults." *Education Economics*, **10**(1):97–109.

Ai, C. and Norton, E. (2003) "Interaction terms in logit and probit models." *Economics Letters*, **80**(1):123–129.

Arendt, J. (2001) "Education effects on health: Causal or from unobserved components?" Working paper, University of Copenhagen.

——— (2004) "In sickness and in health 'till education do us part: education effects on hospitalization." http://www.hha.dk/nat/cim/roervigII/JNARENDT.PDF.

Arkes, J. (2001) "Does schooling improve adult health?" RAND working paper.

Baker, M., Stabile, M., and Deri, C. (2001) "What do self-reported, objective measures of health measure?" Forthcoming, Journal of Human Resources.

Berger, M. and Leigh, J. (1989) "Schooling, self-selection, and health." *Journal of Human Resources*, **24**(3):433–455.

Blackburn, M.L. and Neumark, D. (1993) "Are OLS Estimates of the Return to Schooling Biased Downward? Another Look." NBER #4259.

Bound, J. (1991) "Self-Reported versus Objective Measures of Health in Retirement Models." *Journal of Human Resources*, **26**:106–138.

Card, D. (1995) "Earnings, ability, and schooling revisited." In S. Polachek, editor, *Research in Labor Economics*, volume 14. JAI Press.

Carroll, J. (1997) "Theoretical and technical issues in identifying a factor of general intelligence." In B. Devlin, S. Fienberg, D. Resnick, and K. Roeder, editors, *Intelligence, Genes, and Success: Scientists Respond to the Bell Curve*, pages 125–156. New York: Springer-Verlag.

Deaton, A. and Paxson, C. (2001) "Mortality, education, income and inequality among American cohorts." In D. Wise, editor, *Themes in the Economics of Aging*. Chicago: University of Chicago Press.

Elías, J. (2003) "The education premium puzzle." Working paper, University of Chicago, http://economics.uchicago.edu/download/The_Education_Premium_Puzzle_El.pdf.

FRIEDLANDERA, Y., PALTIELA, O., DEUTSCHA, L., KNAANIEA, A., MASSALHAA, S., TIRAMA, E., AND HARLAPB, S. (2003) "Birthweight and relationship with infant, child and adult mortality in the Jerusalem perinatal study." *Pediatric and Perinatal Epidemiology*, **17**(4):398–406.

FUCHS, V. (1982) "Time preference and health: An exploratory study." In *Economic Aspects of Health*, pages 93–120. Chicago: University of Chicago Press.

GAREN, J. (1984) "The return to schooling: A selectivity bias approach with a continuous choice variable." *Econometrica*, **52**(5):1199–1218.

GROSSMAN, M. (1972) "On the concept of health capital and the demand for health." *Journal of Political Economy*, **80**(2):223–255.

GROSSMAN, M. AND KAESTNER, R. (1997) "Effects of education on health." In J. Berhman and N. Stacey, editors, *The Social Benefits of Education*. Ann Arbor: University of Michigan Press.

HANSEN, K., HECKMAN, J., AND MULLEN, K. (2003) "The effect of schooling and ability on achievement test scores." NBER #9881.

HART, C., TAYLOR, M., SMITH, G., WHALLEY, L., STARR, J., HOLE, D., WILSON, V., AND DEARY, I. (2003) "Childhood IQ, social class, deprivation, and their relationships with mortality and morbidity risk in later life: Prospective observational study linking the Scottish Mental Survey 1932 and the midspan studies." *Psychosomatic Medicine*.

HARTOG, J. AND OOSTERBEEK, H. (1998) "Health, wealth, and happiness: Why pursue a higher education." *Economics of Education Review*, **17**(3):245–256.

HECKMAN, J. AND VYTLACIL, E. (1998) "Instrumental variable methods for the correlated random coefficient model." *Journal of Human Resources*, **33**(4):974–87.

——— (2001) "Identifying the role of cognitive ability in explaining the level of and change in the return to schooling." *Review of Economics and Statistics*, **83**(1):1–12.

IMBENS, G. AND ANGRIST, J. (1994) "Identification and estimation of local average treatment effects." *Econometrica*, **62**(2):467–76.

KENKEL, D. (1991) "Health behavior, health, knowledge, and schooling." *Journal of Political Economy*, **99**(2):287–305.

KITAGAWA, E. AND HAUSER, P. (1973) *Differential mortality in the United States: A study in socioeconomic epidemiology*. Harvard University Press.

Lleras-Muney, A. (2001) "The relationship between education and adult mortality in the United States." Technical report, NBER #8605.

Lleras-Muney, A. and Lichtenberg, F. (2003) "The effect of education on medical technology adoption: Are the more educated likely to use new drugs?" Mimeo.

Lochner, L., Cawley, J., Heckman, J., and Vytlacil, E. (1999) "Understanding the role of cognitive ability in accounting for the recent rise in the economic return to education." In K. Arrow, S. Bowles, and S. Durlauf, editors, *Meritocracy and Economic Inequality*. Princeton: Princeton University Press.

Lochner, L. and Moretti, E. (2001) "The effect of education on crime: Evidence from prison inmates, arrests, and self-reports." Technical report, NBER #8605.

Neale, R., Brayne, C., and Johnson, A. (2001) "Cognition and survival: An exploration in a large multicenter study of the population aged 65 years and over." *International Journal of Epidemiology*, **30**:1383–88.

Newey, W. (1987) "Efficient estimation of limited dependent variable models with endogenous explanatory variables." *Journal of Ecometrics*, **36**:231–50.

Norton, E., Wang, H., and Ai, C. (2004) "Computing interaction effects and standard errors in logit and probit models." *The Stata Journal*, **4**(2):103–116.

Snowden, D., Greiner, L., Kemper, S., Nanayakkara, N., and Mortimer, J. (1999) "Linguistic ability in early life and longevity: Findings from the Nun Study." In J-M. Robine, B. Forette, C. Franceschi, and M. Allard, editors, *The Paradoxes of Longetivity*, pages 103–13. New York: Springer.

Whalley, L. and Deary, I. (2001) "Longitudinal cohort study of childhood IQ and survival up to age 76." *British Medical Journal*, **322**:819–22.

Wooldridge, J. (1997) "On two stage least squares estimation of the average treatment effect in a random coefficient model." *Economics Letters*, **56**:129–133.

——— (2003) "Further results on instrumenal variables estimation of average treatment effects in the correlated random coefficient model." *Economics Letters*, **79**:185–91.

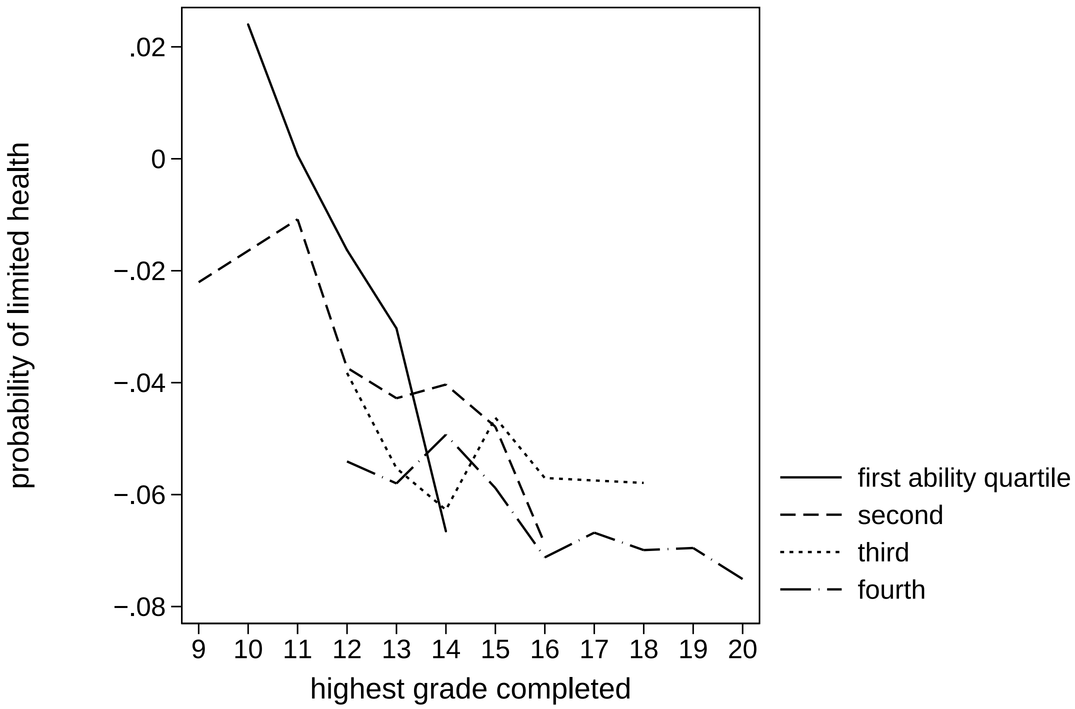

Figure 1: *Semiparametric estimates of effect of schooling and ability on health.* Marginal effects from probit models also including cohort dummies, time-invariant characteristics, and time-varying characteristics as described in Table 2. Baseline category is respondents in the first ability quartile with a grade 9 education.

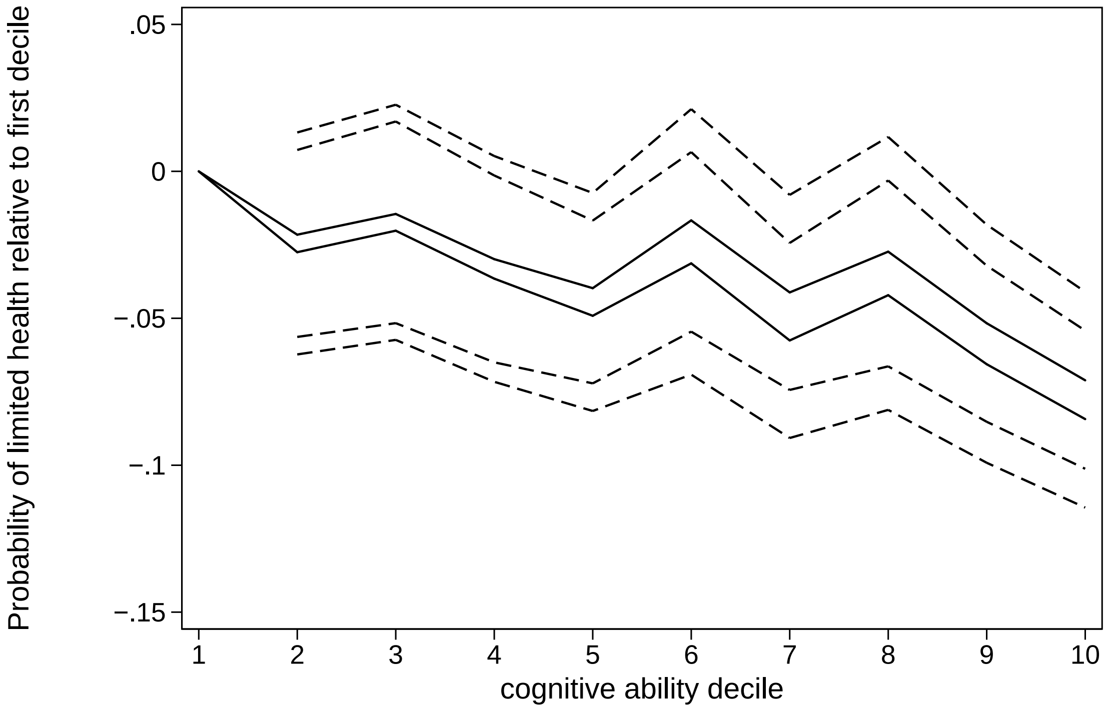

Figure 2: *Health limitations and intelligence among high school graduates.*
Figure shows probability of a health limitation against cognitive ability decile among respondents with exactly a grade 12 education. Upper solid line shows estimates holding income constant. Dashed lines are 95% confidence intervals. Cohort dummies, time-invariant characteristics, and time-varying characteristics as described in Table 2 have been held constant.

Table 1: **Descriptive statistics**

Variable	mean	std. dev.
<u>Endogenous outcomes</u>		
health limits type of employment (indicator)	0.100	0.300
health limits amount of employment (indicator)	0.042	0.201
SF12: general health (5=poor ... 1=excellent)	2.325	1.013
schooling (highest grade completed as of 2000)	13.227	2.383
<u>Cohort dummies</u>		
born in 1964 (indicator)	0.148	0.355
born in 1963 (indicator)	0.145	0.352
born in 1962 (indicator)	0.142	0.349
born in 1961 (indicator)	0.136	0.343
born in 1960 (indicator)	0.119	0.324
born in 1959 (indicator)	0.100	0.300
born in 1958 (indicator)	0.100	0.300
born in 1957 (indicator)	0.023	0.148
<u>Time-invariant characteristics</u>		
hispanic (indicator)	0.175	0.380
black (indicator)	0.297	0.457
male (indicator)	0.474	0.499
Southern residence at age 14 (indicator)	0.382	0.486
urban residence at age 14 (indicator)	0.791	0.407
household receive magazines at age 14 (indicator)	0.576	0.494
household receive newspapers at age 14 (indicator)	0.761	0.427
household member with library card at age 14 (indicator)	0.717	0.451
# of siblings	3.793	2.628

Table 1 continued

Variable	mean	std. dev.
Time-varying characteristics measured in 1979		
married (indicator)	0.078	0.268
divorced or widowed (indicator)	0.016	0.127
family size	4.774	2.191
SMSA residence (indicator)	0.690	0.463
urban residence (indicator)	0.787	0.410
health limits type of employment (indicator)	0.057	0.232
Time-varying characteristics measured in 2000		
married (indicator)	0.565	0.496
divorced or widowed (indicator)	0.236	0.425
family size	3.278	1.622
SMSA residence (indicator)	0.069	0.254
urban residence (indicator)	0.721	0.576
Parents' characteristics		
father's highest grade completed	9.605	5.521
father's education (missing indicator)	0.107	0.309
mother's highest grade completed	10.491	3.913
mother's education (missing indicator)	0.036	0.187
Father's occupation indicators:		
professional	0.193	0.395
clerk	0.069	0.254
farmer	0.019	0.137
craftsmen or foreman	0.336	0.472
laborer	0.078	0.268
service	0.060	0.237
armed forces	0.010	0.102
missing	0.235	0.424
local unemployment rate in 1979	6.196	2.219

N=6,385.

Table 2: **Selected outcomes by health limitation status**

health	years of schooling	standardized intelligence	family income
not limited	13.341	.039	47,899.59
limited	12.205	-.358	22,769.55
overall	13.227	-.000	45,384.62

Table 3: **Probability of health limitation by schooling and ability**

	Ability quartile				
schooling	first	second	third	fourth	Total
9	0.205	0.141	0.286	0.250	0.184
	88	64	7	4	163
10	0.224	0.265	0.000	0.000	0.208
	98	49	19	2	168
11	0.191	0.138	0.100	1.000	0.170
	115	58	20	1	194
12	0.157	0.109	0.103	0.072	0.118
	937	852	710	373	2872
13	0.131	0.085	0.066	0.057	0.083
	122	142	198	106	568
14	0.045	0.099	0.052	0.081	0.071
	88	152	192	185	617
15	0.061	0.089	0.083	0.053	0.072
	49	79	84	94	306
16	0.071	0.034	0.062	0.036	0.045
	42	87	225	447	801
17	0.200	0.000	0.089	0.038	0.054
	5	12	45	105	167
18	0.000	0.000	0.053	0.030	0.033
	5	19	57	132	213
19	0.000	0.000	0.091	0.026	0.032
	2	5	11	77	95
20	0.000	0.000	0.067	0.014	0.022
	2	4	15	70	91
Total	0.152	0.105	0.082	0.051	0.097
	1553	1523	1583	1596	6255

Notes: For each value of schooling, first row shows probability of an employment–limiting health problem and second row shows frequencies. Individuals with grade 8 or lower education excluded.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
intelligence	-0.045		-0.027		-0.022	-0.078	-0.077	-0.074	-0.074
	(10.44)		(5.49)		(4.22)	(3.04)	(3.03)	(2.96)	(3.03)
schooling		-0.019	-0.014	-0.015	-0.012	-0.014	-0.013	-0.012	-0.010
		(11.37)	(7.40)	(8.55)	(6.20)	(7.47)	(6.61)	(6.32)	(5.11)
intelligence*schooling						0.007	0.007	0.007	0.006
						(3.60)	(3.09)	(2.96)	(2.62)
cohort dummies	no	no	no	yes	yes	no	yes	yes	yes
fixed characteristics	no	no	no	yes	yes	no	yes	yes	yes
health in 1979	no	no	no	yes	yes	no	no	yes	yes
characteristics in 1979	no	no	no	yes	yes	no	no	yes	yes
characteristics in 2000	no	no	no	yes	yes	no	yes	yes	yes
family income	no	no	no	no	no	no	no	no	yes
Log-likelihood	-2,018.28	-2,005.84	-1,987.88	-1,888.76	-1,879.75	-1,987.88	-1,918.11	-1,877.57	-1,885.48

Notes: n=6,385. Marginal effects on interaction terms calculated using algorithms discussed in Norton *et al.* (2004).

Table 4: **Marginal effects from probit regressions**

Table 5: **Stratification by gender and ethnicity**

	men	women	white	nonwhite
intelligence	-0.083	-0.056	-0.045	-0.091
	0.040	0.032	0.031	0.060
	-2.10	-1.79	-1.45	-1.53
school	-0.013	-0.011	-0.010	-0.012
	0.003	0.002	0.003	0.004
	-4.52	-4.46	-3.59	-3.28
int*school	0.007	0.006	0.004	0.008
	0.003	0.003	0.003	0.005
	2.00	2.09	1.489	1.730
N	3358	3027	3366	3019
log-likelihood	-1115.621	-750.602	-912.309	-953.185

Notes: Table shows selected marginal effects from probit models. Dependent variable is unity when respondent reports a health limitation. All models include cohort dummies, time-invariant characteristics, and time-varying characteristics as described in Table 2. Marginal effects on interaction terms calculated using algoithms discussed in Norton *et al.* (2004).

Table 6: **The health and intelligence of high school graduates**

	1	2	3	4	5	6	7
intelligence	-0.042	-0.036	-0.033	-0.023			
	-5.09	-3.76	-3.52	-2.49			
quartile 2					-0.028		
					-1.92		
quartile 3					-0.031		
					-1.87		
quartile 4					-0.057		
					-2.96		
decile 2						-0.028	-0.021
						-1.42	-1.10
decile 3						-0.020	-0.014
						-1.00	-0.70
decile 4						-0.037	-0.030
						-1.81	-1.46
decile 5						-0.049	-0.040
						-2.50	-2.01
decile 6						-0.031	-0.017
						-1.46	-0.75
decile 7						-0.058	-0.041
						-2.68	-1.82
decile 8						-0.042	-0.028
						-1.80	-1.13
decile 9						-0.066	-0.051
						-2.74	-2.04
decile 10						-0.084	-0.072
						-2.87	-2.28
family income 1979				-0.0003			-0.0003
				-0.62			-0.60
family income 2000				-0.0011			-0.0011
				-5.17			-5.17
characteristics	no	yes	yes	yes	yes	yes	yes
past health	no	no	yes	yes	yes	yes	yes

Notes: Table shows selected marginal effects from probit models. Dependent variable is unity when respondent reports a health limitation.

	1	2	3	4	5	6	7	8	9
schooling	-0.011	-0.006	-0.005	-0.008	-0.023	-0.015	-0.041	-0.026	0.029
	0.002	0.008	0.008	0.010	0.003	0.005	0.051	0.019	0.029
	-5.781	-0.686	-0.662	-0.792	-6.853	-2.939	-0.789	-1.379	1.011
intelligence	-0.020	-0.027	-0.027	-0.024			-0.033	0.007	-0.038
	0.005	0.012	0.012	0.014			0.012	0.016	0.024
	-3.732	-2.258	-2.252	-1.679			-2.682	0.417	-1.544
estimator	HOLS	2SLS	GMM	GMM	GMM	GMM	GMM	GMM	GMM
instruments		occup	occup		occup, intel	occup.	occup	occup	
Sample	all	all	all	all	all	all	$S \leq 12$	$S > 12$	$S > 12$
N	6385	6385	6385	6385	6385	6385	3527	2858	2858
Hansen's J	7.796	7.388	7.388	3.422	12.342	8.840	6.499	21.297	3.129
Hansen p-value	0.649	0.597	0.597	0.181	0.263	0.452	0.689	0.011	0.209
Hansen d.f.	10	9	9	2	10	9	9	9	2
C-statistic			2.998	1.990	3.565		2.375	0.017	1.017
C p-value			0.083	0.158	0.059		0.123	0.897	0.313
First-stage F		31.2	31.2	77.9	181.2	68.4	4.31	3.70	7.20

Notes: All models include cohort dummies, time-invariant characteristics, and time-varying characteristics as described in Table 2. All models include father's and mother's education levels and local unemployment rates as instruments, some models also include **occup**, father's occupational category dummies, or respondent's **intelligence** as instruments. Reported C-statistics are exogeneity tests; the null hypothesis is that intelligence is exogenous.

Table 7: **GMM estimates of health status**

Table 8: **Marginal effects from instrumental variables probit models**

	1	2	3	4	5
schooling	-0.019	-0.009	-0.010	-0.050	-0.005
	(0.005)	(0.009)	(0.010)	(0.047)	(0.009)
	-3.50	-1.09	-1.06	-1.07	-0.54
intelligence		-0.026	-0.026	-0.036	-0.021
		(0.012)	(0.013)	(0.012)	(0.012)
		-2.13	-1.89	-2.89	-1.78
family income 1979					-0.0001
					(0.0003)
					-0.44
family income 2000					-0.0007
					(0.0001)
					-5.89
sample	all	all	all	$S \leq 12$	all
instruments	educ, occup	educ, occup	educ	educ, occup	educ, occup

Note: All models include cohort dummies, time-invariant characteristics, and time-varying characteristics as described in Table 2. Instrument set **educ** is father's and mother's education levels and local unemployment rates, **occup** is father's occupational category dummies

Table 9: **GMM estimates of correlated random coefficient models**

	Not conditioning on intelligence		Conditioning on intelligence	
estimator	Wooldridge (2003)	Garen (1984)	Wooldridge (2003)	Garen (1984)
schooling ($\bar{b}_1$)	-0.019	-0.016	-0.010	-0.006
	(2.439)	(2.966)	(1.019)	(0.640)
intelligence			-0.138	-0.029
			(2.562)	(2.281)
θ_C			0.008	
			(1.980)	
$\hat{\omega}$		-0.018		-0.025
		(1.882)		(2.008)
$\hat{\omega} *$ (schooling)		0.001		0.001
		(2.628)		(2.362)

Notes: Table shows selected estimates from equations (30) and (33). Dependent variable is unity when the respondent reports a health limitation. The parameter θ_C measures how the effect of schooling on health varies with cognitive ability. $\hat{\omega}$ denotes estimates of unobserved determinants of schooling. t-ratios based on bootstrapped standard errors in parentheses.

Table 10: **Ordered probit estimates of subjective health status**

	1	2	3	4
schooling		-0.105	-0.098	-0.093
		(0.011)	(0.012)	(0.012)
		-9.934	-8.215	-7.685
intelligence	-0.191		-0.335	-0.365
	(0.030)		(0.151)	(0.152)
	-6.474		-2.217	-2.404
sch.*intell.			0.018	0.021
			(0.011)	(0.011)
			1.642	1.911
family income 1979				0.000
				(0.000)
				0.236
family income 2000				-0.000
				(0.000)
				-2.980

Notes: Dependent variable is self-reported general health status from the SF12 battery (1=excellent ... 5=poor). All models include cohort dummies, time-invariant characteristics, and time-varying characteristics as described in Table 2.

Table 11: **Probit estimates varying cognitive ability measure**

	seperable		interaction		
	ability	school	ability	school	ability*school
AFQT	-0.027	-0.010	-0.064	-0.010	0.005
	-5.052	-5.179	-3.014	-5.126	2.772
Age and schooling adjusted AFQT	-0.021	-0.012	-0.069	-0.012	0.006
	-4.317	-6.372	-2.969	-6.482	2.806
g	-0.029	-0.010	-0.052	-0.010	0.004
	-5.464	-5.016	-2.680	-5.027	2.470
Science	-0.016	-0.013	-0.042	-0.013	0.004
	-3.414	-6.656	-2.167	-6.674	2.224
Arithmetic	-0.021	-0.012	-0.065	-0.012	0.006
	-4.495	-6.102	-3.070	-6.039	2.904
Word knowledge	-0.022	-0.011	-0.030	-0.012	0.003
	-4.622	-6.026	-1.600	-6.042	1.684
Paragraph comprehension	-0.019	-0.012	-0.030	-0.012	0.003
	-4.425	-6.339	-1.553	-6.363	1.661
Numerical operations	-0.021	-0.012	-0.030	-0.012	0.003
	-5.234	-6.354	-1.642	-6.355	1.756
Coding speed	-0.014	-0.013	-0.021	-0.013	0.002
	-3.351	-7.396	-1.148	-7.389	1.349
Auto and shop	-0.011	-0.014	-0.033	-0.014	0.003
	-2.277	-8.220	-1.541	-8.108	1.663
Math knowledge	-0.023	-0.011	-0.069	-0.011	0.006
	-4.786	-5.418	-3.075	-5.375	2.832
Mechanical	-0.015	-0.014	-0.047	-0.013	0.004
	-3.283	-7.602	-2.237	-7.501	2.310
Electronics	-0.010	-0.014	-0.036	-0.014	0.003
	-2.207	-7.787	-1.748	-7.756	1.826

Notes: Marginal effects from probit models of health limitation status. All models include cohort dummies, time-invariant characteristics, and time-varying characteristics as described in Table 2. "*g*" is general intelligence measured as the first principal component of ASVAB scores. Models run seperately for each measure of intelligence. All cognitive measures have been standardized. Marginal effects on interaction terms calculated using algorithms discussed in Norton *et al.* (2004).