Estimating Aggregate Human Capital Externalities

Junjie Guo, Nicolas Roys and Ananth Seshadri*

January 2018

Abstract

This paper estimates human capital externalities formulated in Lucas (1988). We incorporate externalities into an overlapping generations model of human capital accumulation with Compulsory Schooling Laws (CSL). The model implies that human capital externalities can be estimated from the effect of CSL affecting the schooling decision of one generation on the wage of other generations. Using an instrumental variable strategy deduced from the model, we find that one more year of average schooling at the U.S. state level raises individual wage by about 6-8%. Taking this reduced form estimate into account, we estimate that the elasticity of a typical firm's productivity with respect to the average human capital of an economy is 0.121.

JEL: E23; J24; O47

*Guo: Center for Research on the Wisconsin Economy (CROWE), University of Wisconsin-Madison, jguo27@wisc.edu; Roys: Royal Holloway, University of London, nicolas.roys@rhul.ac.uk; Seshadri: Department of Economics, University of Wisconsin-Madison, 1180 Observatory Drive, Madison, WI 53706, aseshadr@ssc.wisc.edu, Phone: 608-262-6196, Fax: 608-262-2033. We thank seminar participants at UCL, University of Oslo, Stockholm University, Stockholm School of Economics and UW-Madison as well as Gary Becker, Richard Blundell, Bob Lucas, Kjetil Storesletten and Chris Taber for very helpful comments. All errors are ours.

1 Introduction

Human capital externalities have long been thought to be important to rationalize the prevalence of educational policy as well as to understand cross-country income differences (Lucas, 1988).[1] The empirical evidence for externalities, however, is mixed. Rauch (1993) is one of the first attempts to estimate externalities. He finds externalities on the order of 3 to 5%. A major obstacle in estimating the external effect of education on income is the ability to identify causal effects. Acemoglu and Angrist (2001) use variation in compulsory schooling laws (CSL) over time in each state as an instrument for average schooling and find very little evidence in support of externalities. Moretti (2004a) uses the (lagged) city demographic structure and the presence of a land-grant college in a city as instruments for the share of college graduates in the city's labor force (college share) and finds that a one percentage point increase in a city's college share raises the wage of workers in that city by about 0.4-1.9%. Other important studies in this literature include Moretti (2004c) and Ciccone and Peri (2006). After summarizing the relevant literature, Moretti (2004b) concludes "empirical literature provides some intriguing evidence on the existence of human capital externalities, but we are still far from a consensus on the magnitude of such externalities. The empirical literature on the subject is still very young and the econometric challenges are difficult to overcome. More work is needed".

The objective of this paper is to estimate human capital externalities. Different from the existing literature where most of the studies are reduced form and use schooling as a direct measure of human capital, we take a structural approach and model individual human capital accumulation both in and out of school while allowing for human capital externalities. We model human capital accumulation via the seminal work of Ben-Porath (1967), which is extended to include a government that sets Compulsory Schooling Laws (CSL) in each period requiring a minimum number of years of schooling for each individual born in that

[1] We focus on the externalities related to the productivity of workers and firms. Human capital externalities may also appear as nonproduction externalities, for example, by reducing crime rates and increasing civic participation. Davies (2003) and Lochner (2011) provide reviews of this literature.

period. We embed the extended model into an overlapping generations framework where the production function features human capital externalities as in Lucas (1988). Specifically, the average human capital of an economy is allowed to affect the productivity of a typical firm in the economy.

The model implies that, with a positive human capital externality, a tightening of CSL for one generation would affect the wage of other generations. There are two channels. The first is a direct effect of human capital price. When some individuals in one generation are forced to stay in school for a longer period of time and acquire more human capital, the average human capital would rise when these individuals enter the labor market. With a positive human capital externality, the increase in average human capital would raise the productivity and thus the price of individual human capital. A higher price would raise the wage of other generations in the labor market at the same time.

The second channel is an indirect effect that works through individual human capital accumulation. Specifically, when CSL is tightened for one generation, in anticipation of the higher human capital price in the future, individuals will increase their investments in human capital so that they would have more human capital to supply to the market and earn a higher wage when the price is higher. This effect works for all individuals who would be working when the price is higher due to the tightening of the CSL, including those generations who are not directly affected by the tightening of the CSL.

Following the model's implication, we first provide some reduced form evidence supporting a positive human capital externality. We take each U.S. state as an economy, and approximate average human capital with state average schooling. To estimate the causal effect of state average schooling on individual wages, in light of the model, we use the CSL affecting other workers' schooling decisions as an instrumental variable (IV) for the average schooling faced by the worker whose wage is in consideration. Using decennial censuses 1960-1980, we find a positive and statistically significant effect of state average schooling on individual wages: other things equal, one more year of state average schooling would raise individual wages by about 6-8%. We also find the effect is larger when the exogenous variation in average schooling is driven by the tightening of CSL for older cohorts, and it is larger for less-educated workers. Our estimates are robust to different empirical specifications.

Taking the reduced form evidence into account, we estimate the model by indirect inference. We allow each state to have its own paths of CSL and total factor productivity (TFP), where the former is taken from data and the latter is estimated to match the path of per capita income over time. We also allow each state to have its own distribution of initial conditions across individuals. Given a guess of model parameters, we solve for the transitional dynamics for each state by iterating over the path of average human capital. We generate the joint distribution of schooling and earnings from the simulated data and match it with the actual data. Human capital production technologies in school and at work are estimated to match the schooling distribution and wages by schooling and age. To identify human capital externalities, we run regressions using simulated data to estimate the effect of state average schooling on individual wages and match the resulting estimates with those from actual data. We find an elasticity of firm productivity with respect to average human capital at around 0.121. We show that the model cannot match the empirical estimates of the effect of average schooling on individual wages in the absence of human capital externalities.

Our baseline model assumes that human capital is homogeneous and hence does not feature any complementarity across types. A sceptical reader might take the position that our results are suggestive of complementarities between different types of human capital. We argue that our finding of a sizable externality continues to hold in an environment featuring complementarity between two types of human capital (low skilled and high skilled). Because the direct effect of CSL is on the schooling decision of individuals who would not finish high school and are most likely to be low skilled workers, an increase in CSL raises the amount of low skilled human capital through more schooling. This, due to the complementarity between the two types of human capital, raises the relative price of high skilled human capital. Low skilled workers at the margin would choose to become high skilled, and this selection effect reduces the average ability of low skilled workers and leads to a negative correlation between CSL and individual wages conditional on schooling if there is no human capital externality. The positive effect of CSL on individual wages observed in the data thus requires a positive human capital externality even if different types of human capital are not perfect substitutes. Indeed, in the presence of complementarities, we would need stronger external effects than in the baseline case to match our empirical IV estimates.

Our findings are complementary to Gennaioli et al. (2013) who find overwhelming evidence that human capital fosters development through entrepreneurial education and human capital externalities. With significant human capital externalities, workers in rich countries would accumulate more human capital over the life cycle than those in poor countries. This prediction is consistent with the findings in Lagakos et al. (forthcominga,f). Our findings also complement Choi (2011) and Malley and Woitek (2017), both of which find significant learning externalities along the lines of Tamura (1991). An interesting direction for future research is to estimate the externalities in Lucas (1988) and Tamura (1991) jointly.

The rest of this paper proceeds as follows. Section 2 presents an overlapping generations model of human capital accumulation with externalities. Section 3 presents some reduced-form evidence suggesting a positive human capital externality. Section 4 estimates the structural model. Section 5 extends our baseline model to the case of heterogeneous human capital and demonstrates that our conclusion on human capital externalities holds with complementarity between two types of human capital. Section 6 concludes.

2 Model

This section presents an overlapping generations model of human capital accumulation with externalities. Following Lucas (1988), we model human capital externalities as the effect of average human capital of all workers on the productivity of a typical firm in an economy. We model human capital accumulation via Ben-Porath (1967), which we extend to include a government that sets Compulsory Schooling Laws (CSL) in each period requiring a minimum years of schooling for individuals born in that period.

Consider a small open economy which takes the world interest rate r as given. There are three types of agents in this economy, workers, firms and the government. Their problems are discussed in turn.

2.1 Workers

The economy is populated by overlapping generations. Each generation, indexed by time of birth t_b, consists of a measure one of individuals different from each other in terms of their innate ability to learn, z, and initial stock of human capital, h_0. We index each individual by the tuple (z, h_0, t_b).

Each individual lives for $R+1$ periods. An individual maximizes the present discounted value (PDV) of lifetime income by choosing the amount of money x_0 and time n_0 spent in school at age 0 and the amount of time n_a spent on training at each age $a > 0$ on the job. The dynamic programming problem for individual (z, h_0, t_b) is

$$V(z, h_0, t_b) = \max_{n_a \in [0,1], x_0 \geq 0} \left\{ \left[w_{t_b} h_0 (1 - n_0) - x_0\right] + \sum_{a=1}^{R} \frac{w_{t_b+a} h_a (1 - n_a)}{(1+r)^a} \right\} \tag{2.1}$$

subject to

$$\begin{aligned} h_1 &= z (n_0 h_0)^{\alpha_1} x_0^{\alpha_2} + h_0 (1 - \delta) \\ h_{a+1} &= cz (n_a h_a)^{\alpha_3} + h_a (1 - \delta) \text{ for } a \geq 1 \\ n_0 &\geq \overline{CSL}_{t_b} \end{aligned}$$

where w_t is the price (rental rate) of human capital at time t, δ is the human capital depreciation rate per period, and the individual's ability to produce human capital at work is allowed to be different from his innate ability to learn by a factor of c. The amount of human capital produced in school depends on both time n_0 and goods x_0 inputs, while human capital production at work is assumed to depend only on the time spent on training n_a. The elasticity of human capital production with respect to the time input depends on whether an individual is in school or at work. With time endowment normalized to 1 in each period, the amount of time spent at work is given by $(1 - n_a)$, and the resulting wage income is $w_{t_b+a}h_a(1 - n_a)$.

The mechanism of this model is well known. Other things equal, an increase in n_a (x_0) raises future human capital and income at the cost of a lower income from the current period, and the optimal n_a (x_0) is obtained by equating the marginal benefit from future income with the marginal cost from current income. However, as Compulsory Schooling Laws (CSL) are introduced, individual choice of n_0 may be constrained. In particular, for all individuals born at time t_b, their decisions must satisfy $n_0 \geq \overline{CSL}_{t_b}$.

2.2 Firms

Firms produce the final good using human capital h and physical capital k. The production function of a typical firm is given by

$$y_t = A_t k_t^{\alpha} h_t^{1-\alpha} \bar{H}_t^{\theta}$$

where y is output, A is productivity, and t is a time index. Following Lucas (1988), we assume that the effective productivity of a firm also depends on the average human capital of the economy $\bar{H}$. The strength of this effect is captured by θ, the key parameter of interest in this paper.

Markets are competitive. Taking interest rate r, price (rental rate) of human capital w_t and (effective) productivity $A_t\bar{H}_t^{\theta}$ as given, a firm maximizes profits

$$\pi_t = \max_{k_t, h_t} \left\{ A_t k_t^{\alpha} h_t^{1-\alpha} \bar{H}_t^{\theta} - r k_t - w_t h_t \right\} \tag{2.2}$$

by choosing k_t and h_t, where the price of final good is normalized to be 1. Profit maximization implies

$$w_t = (1-\alpha)\left(\frac{\alpha}{r}\right)^{\frac{\alpha}{1-\alpha}} A_t^{\frac{1}{1-\alpha}} \bar{H}_t^{\frac{\theta}{1-\alpha}} \tag{2.3}$$

2.3 Government

The government has only one role, making decisions in each period t on minimum amount of schooling time $\overline{CSL}_t$. As we are not trying to work out the optimal choice of $\overline{CSL}_t$, we make no assumption on the objective of the government. The only assumption is that choices of $\overline{CSL}_t$ are exogenous to both workers and firms. As shown below, workers' responses to exogenous variations in $\overline{CSL}_t$ are critical for the identification of human capital externalities θ.

2.4 Equilibrium

In equilibrium, both workers and firms make decisions optimally and all markets clear. In particular,

1. For each individual indexed by (z, h_0, t_b), x_0^* and $\{n_a^*\}_{a=0}^R$ are optimal solutions to the problem given by equation (2.1).
2. For each firm, k_t^* and h_t^* are optimal solutions to the problem given by equation (2.2).
3. The market for human capital clears in each period.

Let $\{h_a^*\}_{a=0}^R$ be the optimal path of human capital for individual (z, h_0, t_b) implied from condition 1. Assume the distribution of initial conditions (z, h_0) is the same across generations and given by $F(z, h_0)$. The average human capital of all workers involved in the production of final good at time t, denoted by $\bar{H}_t$, is given by

$$\bar{H}_t = \frac{\sum_{a=0}^R \int h_a^* (1-n_a^*)\, dF(z, h_0)}{\sum_{a=0}^R \int (1-n_a^*)\, dF(z, h_0)} \tag{2.4}$$

Condition 3 requires the $\bar{H}_t$ implied from equation (2.4) is equal to the $\bar{H}_t$ in equation (2.3).

2.5 Discussion

Other things equal, an increase in $\overline{CSL}_t$ forces individuals born at time t who would have chosen $n_0 < \overline{CSL}_t$ to stay in school for a longer period of time and acquire more human capital. This would raise their wages later when they enter the labor market. We call this the private return to schooling, and this is the only effect of CSL in the absence of human capital externalities $\theta = 0$.

With a positive human capital externality $\theta > 0$, an increase in $\overline{CSL}_t$ would also affect the wage of those whose schooling is not directly affected. To see this, suppose there is a temporary increase in $\overline{CSL}_t$ from s_1 to s_2, and there is no change in $\overline{CSL}$ for other generations. The individuals that are directly affected are those born at time t who would have chosen $n_0 < s_2$ if $\overline{CSL}_t = s_1$. We use D to represent this group of individuals, and use N to represent other individuals born at time t whose schooling is not directly affected by the increase in $\overline{CSL}_t$.

Due to the extra human capital accumulated by group D, the increase in $\overline{CSL}_t$ would raise the average human capital $\left\{\bar{H}_{t+\tau}\right\}_{\tau=1}^{R}$ when group D enters the labor market. With $\theta > 0$, equation (2.3) implies that the prices of human capital $\{w_{t+\tau}\}_{\tau=1}^{R}$ would also be higher. The higher prices would raise the wage income of all individuals working in those periods, including group N as well as individuals born before $(t - R + 1 \leq t_b \leq t - 1)$ and after $(t + 1 \leq t_b \leq t + R)$ time t.

In addition to this direct effect, there is another effect that works through human capital accumulation. Specifically, in anticipation of the higher prices $\{w_{t+\tau}\}_{\tau=1}^{R}$ which would raise the marginal benefit of human capital investments, group N and individuals born at $t_b \in [t - R + 1, t - 1]$ would re-optimize and invest more in human capital production at time t.[2]

[2] They would also invest more at time $t+1$ if the increase in $\{w_{t+\tau}\}_{\tau\geq 2}$ is much larger than the increase

The extra human capital accumulated at time t would raise their wages further.

In summary, the model predicts that human capital externalities can be estimated from the effect of CSL on the wage of individuals whose schooling is not directly affected. In particular, a positive impact of CSL for one generation on the wage of other generations is evidence for a positive human capital externality. In the next section, we provide some reduced-form evidence suggesting a positive effect of CSL on the wage of individuals whose schooling is not directly affected. The reduced-form estimates are then incorporated into the estimation of the structural model.

3 Reduced-Form Evidence

This section presents some reduced-form evidence suggesting a positive human capital externality. We treat each U.S. state as a small open economy, and approximate average human capital $\bar{H}_t$ with state average schooling. In light of the model, we use CSL that does not directly affect the schooling of the individual in consideration as an instrument for state average schooling. We find a large positive and statistically significant effect of state average schooling on individual wages, and the effect is larger when the increase in average schooling is due to more schooling attained by older cohorts. We also find that the effect is larger for less-educated workers.

3.1 Empirical Specification

Assume each U.S. state is a small open economy. According to the model, the log wage income, $Y_{ijt} \equiv \log\left[w_{jt}h_{ijt}\left(1-n_{ijt}\right)\right]$, of worker i who supplies $h_{ijt}\left(1-n_{ijt}\right)$ units of human capital to state j at time t is given by

$$Y_{ijt} = \beta_0 + \frac{1}{1-\alpha}\log A_{jt} + \frac{\theta}{1-\alpha}\log \bar{H}_{jt} + \log\left[\left(1-n_{ijt}\right)h_{ijt}\right]$$

in w_{t+1}, and same is true for periods after $t+1$. Actually, in those cases, even workers born after t may increase their investments in human capital.

where $\beta_0 = \frac{\alpha}{1-\alpha} \log \frac{\alpha}{r} + \log(1-\alpha)$ is a constant.

Without a direct measure of human capital, it is standard to approximate it using years of schooling and other individual characteristics. Assume

$$\log \bar{H}_{jt} = \delta_j^1 + \delta_t^1 + \gamma \overline{S}_{jt} + u_{jt}^1$$

where δ_j^1 and δ_t^1 are state and time fixed effects, $\bar{S}_{jt}$ is the average schooling of all workers in state j and time t, and u_{jt}^1 is the error term. We obtain, by scaling the variables by a factor of $\frac{\theta}{1-\alpha}$,

$$\frac{\theta}{1-\alpha} \log \bar{H}_{jt} = \delta_j + \delta_t + \gamma_1 \overline{S}_{jt} + u_{jt}^2$$

where $\gamma_1 = \frac{\gamma\theta}{1-\alpha}$ is a reduced-form measure of human capital externalities θ. As long as average human capital depends positively on average schooling ($\gamma > 0$), and human capital contributes positively to the final good production ($1-\alpha > 0$), γ_1 would have the same sign as θ. In other words, a positive γ_1, as estimated below in this section, is evidence of positive human capital externalities $\theta > 0$.

Assume further that

$$\log\left[(1-n_{ijt})\, h_{ijt}\right] = \gamma_2 s_i + X_i^{'} \mu + \epsilon_{ijt}$$

where s_i is the schooling of worker i, X_i is a vector of individual characteristics including state-of-birth (SOB) and year-of-birth (YOB) dummies, and ϵ_{ijt} is the error term. We can write the (log) wage equation as

$$Y_{ijt} = \delta_j + \delta_t + \gamma_1 \overline{S}_{jt} + \gamma_2 s_i + X_i^{'} \mu + u_{jt} + \epsilon_{ijt} \tag{3.1}$$

where u_{jt} is a state-time error component that accounts for both u_{jt}^2 and the unobservable productivity term $\frac{1}{1-\alpha} \log A_{jt}$.

3.2 Data

We use U.S. decennial censuses 1960-1980. Following Acemoglu and Angrist (2001) who have estimated equation (3.1) using the same data, the schooling variable s_i for an individual is defined to be the highest grade completed and capped at 17, and Y_{ijt} is the log weekly wage calculated by dividing annual wage and salary income by weeks worked in the previous year before taking the log. Average schooling $\overline{S}_{jt}$ is calculated as the weighted average years of schooling of all US-born persons aged 16-64 who lived in state j in census year t, where the weight is the product of the IPUMS weighting variable SLWT and weeks worked in the previous year. The main analysis is limited to US-born white men in their 40s with positive weekly wages reported in the censuses. All estimates are weighted by SLWT.[3]

We also use three measures of CSL. Acemoglu and Angrist (2001) construct two measures for each state in each year from 1914 to 1978, one for compulsory attendance laws (CA) and the other for child labor laws (CL), as follows

$$\begin{aligned} CA &= \max\left\{req\_sch, drop\_age - enroll\_age\right\} \\ CL &= \max\left\{work\_sch, work\_age - enroll\_age\right\} \end{aligned}$$

where req_sch is the minimum years of schooling required before leaving school, $drop\_age$ is the minimum dropout age, $enroll\_age$ is the maximum enrollment age, $work\_sch$ is the minimum years of schooling required before work was permitted and $work\_age$ is the minimum work age.

As noted by Goldin and Katz (2008), the first term in the max function for CA is an exception which allows the child to leave school before the dropout age, the correct calculation

[3] For comparison, we downloaded the data directly from Angrist's webpage `http://economics.mit.edu/faculty/angrist/data1/data/aceang00`. There are three datasets, one for compulsory schooling laws (CompSchoolLaws.rar), one for average schooling (average4.sas7bdat or average4.dta) and one for the micro data (three.rar). The micro data is only a subset of the original censuses (taken from the IPUMS system) in that it contains only white men aged 21-58. See appendix B of Acemoglu and Angrist (2001) for more details on data sources and the construction of variables.

of CA would use a min function

$$CA = \min\left\{req\_sch, drop\_age - enroll\_age\right\}$$

We use this corrected CA along with the original CL.

Our last measure of CSL comes from Stephens and Yang (2014). They construct a measure of required years of schooling RS which accounts for any changes to the compulsory attendance and child labor laws that may occur during the child's schooling years. For each individual born in year t in state j, RS is generated by iterating through ages 6 to 17 to determine whether the child was required to attend school at that age based on the law that was in place in that state in that year. By using this iterative process, the number of years of schooling the child would have been required to complete by each age is determined, which, in turn, is used to determined whether the child was eligible for any school attendance exceptions at each age. For each age between 6 and 17, if the child either has not reached his dropout age or is not eligible for an exception, RS is increased by one. Once the child either reaches the dropout age or meets the minimum age and/or years of schooling for an exception, RS is not increased unless there is a subsequent change in the schooling statutes. This measure is available for all individuals born in mainland U.S. between 1905 and 1961.

3.3 Identification Strategy

Because average schooling $\overline{S}_{jt}$ is likely to be correlated with u_{jt} via the unobserved productivity term A_{jt}, OLS estimates of γ_1 are likely to be biased. To address this, we use CSL as an instrument for average schooling $\overline{S}_{jt}$.

As discussed in section 2.5, the effect of CSL for one birth cohort ($\overline{CSL_t}$) on the wage of other birth cohorts ($t-R+1 \leq t_b \leq t-1$ and $t+1 \leq t_b \leq t+R$)) that works through average human capital $\left\{\bar{H}_{t+\tau}\right\}_{\tau=1}^{R}$ is evidence for human capital externalities. In light of this, given the individual level RS that potentially affects a worker's schooling,[4] we could calculate $\overline{RS}^a_{jt}$, the average RS among all workers at age a in state j in census year t. Naturally, $\overline{RS}^a_{jt}$ is a measure of the strength of CSL affecting schooling decisions of all workers at age a in state j in census year t, and it should potentially be correlated with $\overline{S}^a_{jt}$, the average years of schooling for this group of workers. Because $\overline{S}^a_{jt}$ is a part of $\overline{S}_{jt}$, $\overline{RS}^a_{jt}$ is a potential determinant of $\overline{S}_{jt}$. A positive effect of $\overline{RS}^a_{jt}$ on the wage of workers in state j in year t but not at age a would be evidence for a positive human capital externality. In other words, for worker i at age a in state j in year t, we could use $\left\{\overline{RS}^{a'}_{jt}\right\}_{a'\neq a}$ as instruments for the average schooling $\overline{S}_{jt}$ faced by this worker. A positive estimate of γ_1 from this strategy implies a positive effect of $\left\{\overline{RS}^{a'}_{jt}\right\}_{a'\neq a}$ on the wage of worker i via $\overline{S}_{jt}$. As $\left\{\overline{RS}^{a'}_{jt}\right\}_{a'\neq a}$ are not expected to affect worker i directly, a positive estimate of γ_1 implies a positive human capital externality $\theta > 0$.

Actually, if the schooling of worker i is not constrained by CSL, which would be the case if worker i belongs to group N discussed in section 2.5, we could also use $\overline{RS}^a_{jt}$ as an instrument for $\overline{S}_{jt}$. For simplicity, our baseline specification does not exclude $\overline{RS}^a_{jt}$ and uses the whole vector $\{\overline{RS}^a_{jt}\}$ as instruments for $\overline{S}_{jt}$. Our results, however, are robust to two alternative specifications.

In the first case, we include the individual level RS intended to affect the schooling of the individual whose wage is in consideration (own CSL) as an additional control variable into equation (3.1). With this additional control, the identification of γ_1 comes from workers faced with the same CSL, and they face different average schooling $\overline{S}_{jt}$ purely due to the CSL effective for other workers. Because the main analysis is limited to US-born white men in their 40s, in the second case, we exclude $\{\overline{RS}^a_{jt}\}^{50}_{a=40}$ and use $\{\overline{RS}^a_{jt}\}_{a<40}$ and $\{\overline{RS}^a_{jt}\}_{a>50}$ as instruments for $\overline{S}_{jt}$. In this case, the variation in $\overline{S}_{jt}$ comes purely from the CSL effective for other birth cohorts.

In addition to the instruments, we also control for either state- or region-specific trends

[4] RS is used here as an example. The same argument works for the other two measures CL and CA.

to further reduce the potential correlation between average schooling $\overline{S}_{jt}$ and the error term u_{jt}. As pointed out in Angrist and Pischke (2014),[5] "the principal threat to validity in this context is omitted state-specific trends. Specifically, we must worry that states in which compulsory schooling laws grew stricter simultaneously experienced unusually large wage growth across cohorts for reasons unrelated to schooling. ... a simple check for state-specific trends adds a linear time trend for each state to the model of interest. In this case, the relevant time dimension is year of birth, so the model with state-specific trends includes a separate linear year-of-birth variable for each state of birth in the sample". Angrist and Pischke (2014) conclude that the strategy used in Acemoglu and Angrist (2001) to estimate equation (3.1) is "a failed research design" because it does not control for state-specific trends.

Similarly, Stephens and Yang (2014) argue that the key to the identification strategy which exploits variation in the timing of law changes across states over time is the common trends assumption that "all other changes which occur across states during this period are uncorrelated with the law changes, educational improvements, and the outcomes under investigation". To examine the importance of this common trends assumption, Stephens and Yang (2014) use a specification where year of birth effects are allowed to vary across the four US census regions of birth. They find that significant estimates of the impact of schooling on a variety of outcomes including wages become insignificant once year of birth effects are allowed to vary by region of birth.

For brevity, we only report estimates of equation (3.1) augmented with region-specific trends as in Stephens and Yang (2014). Results with state-specific trends modeled along the lines of Angrist and Pischke (2014) are similar both qualitatively and quantitatively.

[5]Pages 226-227.

3.4 First Stage Estimates

We first estimate the impact of the proposed instruments on state average schooling $\overline{S}_{jt}$. Table 1 reports the estimates from three specifications where $\{\overline{CL}_{jt}^{a}\}$, $\{\overline{CA}_{jt}^{a}\}$ and $\{\overline{RS}_{jt}^{a}\}$ are used as the instruments, respectively. In all three cases, the dependent variable is $\overline{S}_{jt}$, and we control for all variables on the right hand side of equation (3.1) as well as region-specific trends as in Stephens and Yang (2014). All standard errors reported in this paper are clustered to the state-by-year level.

Each row of Table 1 reports the estimated impact of CSL for a particular age, with the starting and stopping ages determined by data availability.[6] For example, the first row reports the estimated impacts of $\overline{CL}_{jt}^{a=21}$, $\overline{CA}_{jt}^{a=21}$ and $\overline{RS}_{jt}^{a=21}$. As states don't change CSL very often, the age-specific instruments are highly correlated with each other. The age-specific estimates are thus not very meaningful. However, we can still look at the joint significance of all instruments summarized by the (partial) F-statistics reported in the last row of Table 1. The F-statistic is above 20 for CL and RS and above 10 in all three cases, indicating a strong and statistically significant effect of the instruments on state average schooling $\overline{S}_{jt}$.

3.5 Baseline Estimates of External Returns to Schooling

With a significant first stage, we now turn to estimates of external returns to schooling γ_1 obtained from regressions of equation (3.1) augmented with region-specific trends as in Stephens and Yang (2014).

Table 2 reports the baseline estimates. The first column of Panel A reports the OLS estimates. The private γ_2 and external γ_1 returns to schooling are estimated to be 7.2% and 6.9%, respectively, and both are statistically significant. The last three columns in panel A report estimates where average schooling $\overline{S}_{jt}$ is instrumented by $\{\overline{CL}_{jt}^{a}\}$, $\{\overline{CA}_{jt}^{a}\}$

[6] For example, as data for RS starts from 1905, we could not calculate $\overline{RS}_{jt}^{a}$ for $a \geq 56$ using the 1960 census. Consequently, $\overline{RS}_{jt}^{a}$ stops at $a = 55$.

and $\{\overline{RS}_{jt}^{a}\}$, respectively. All estimates with instrumental variables (IV) reported in this paper are obtained from Two-Stage-Least-Square (2SLS) regressions. Results from other methods like the limited information maximum likelihood (LIML) are very similar. Despite a small increase in the estimated standard error, all three IV estimates of external returns to schooling reported in panel A are positive and statistically significant. Although the estimates vary with the particular instruments used, they fall in a narrow range of 6-8% and are comparable to the OLS estimate.

While panel A treats individual schooling as exogenous, in panel B, we report estimates where individual schooling is taken as endogenous and instrumented with quarter of birth (QOB) dummies as in Angrist and Krueger (1991). Treating individual schooling as endogenous turns out to have little impact on estimates of external returns, which are still statistically significant and comparable to both the OLS estimate and the IV estimates with exogenous individual schooling. For instance, the estimated external returns when $\{\overline{CL}_{jt}^{a}\}$ are used as instruments is 6.9% when individual schooling is treated as exogenous, and it is 7.3% when individual schooling is treated as endogenous. The estimated external returns with $\{\overline{RS}_{jt}^{a}\}$ as instruments are 6.1% no matter whether individual schooling is taken as exogenous or endogenous.

The row named F-CSL reports the partial F-statistics of the joint significance of the CSL instruments for average schooling $\overline{S}_{jt}$, similar to those reported in the last row of Table 1 with the difference that now individual schooling s_i is replaced with QOB dummies. Clearly, the CSL instruments still have a strong impact on $\overline{S}_{jt}$ when individual schooling is endogenous.

The row named F-QOB reports the partial F-statistics of joint significance of QOB dummies in a regression of individual schooling s_i on all control variables in equation (3.1), region-specific trends, CSL instruments and QOB dummies. The F-statistics are only around 5, indicating that QOB dummies don't have a strong impact on individual schooling conditional on other variables like the region-specific trends. As our estimates of external returns are similar whether individual schooling is treated as exogenous not, in what follows we only report results treating individual schooling as exogenous.

3.6 Estimates from Alternative Specifications

Table 3 reports estimates from the two alternative specifications discussed in section 3.3. Panel A reports the estimates where own CSL, the CSL intended to affect the schooling of the individual whose wage is in consideration, is included as an additional control variable in equation (3.1), and the three columns in panel B report the estimates where $\{\overline{CL}_{jt}^{a}\}_{a=40}^{50}$, $\{\overline{CA}_{jt}^{a}\}_{a=40}^{50}$ and $\{\overline{RS}_{jt}^{a}\}_{a=40}^{50}$ are excluded from the set of instruments, respectively. In both cases, the estimated external returns are very close to our baseline estimates reported in Table 2. If anything, estimates from these alternative specifications are slightly larger. When $\{\overline{CL}_{jt}^{a}\}_{a=40}^{50}$, $\{\overline{CA}_{jt}^{a}\}_{a=40}^{50}$ and $\{\overline{RS}_{jt}^{a}\}_{a=40}^{50}$ are excluded, the strength of the instruments is reduced, as indicated by the smaller (partial) F-statistics of the joint significance of the instruments for $\overline{S}_{jt}$ reported in the last row of Table 3. However, these F-statistics are still around 10. As the estimates from this specification are similar to those from other specifications, we are not too concerned about the potential issues with weak instruments.

As pointed out in Acemoglu and Angrist (2001), states with higher wages due to the unobservable term u_{jt} may attract more better-educated workers from other states. Because better-educated workers are more likely to be from states with stricter CSLs, interstate migration may lead to a correlation between the error term u_{jt} and state of birth (SOB) CSLs, invalidating the use of SOB CSLs as instruments for average schooling $\overline{S}_{jt}$. State of residence (SOR) CSLs, defined as the CSLs that would have affected a worker's schooling were he born in his current SOR, on the other hand, is not subject to this critique based on migration. Because most workers were living in their SOB when surveyed by censuses, SOR CSLs should be correlated with and can be used as instruments for average schooling. We thus calculate instruments based on SOR CSLs for each birth cohort in each state in each census in the same way as we did for SOB CSLs. Table 4 reports the estimates using these SOR CSLs as instruments for average schooling. Again, the estimated external returns in

Table 4 are very close to corresponding estimates in Table 2, and all estimates are statistically significant. This suggests that, as in Acemoglu and Angrist (2001), our previous estimates of external returns to schooling using SOB CSLs as instruments for average schooling are not seriously biased by potential interstate migration.

As the estimated external returns are robust to different specifications, we proceed with our baseline specification in the following analysis.

3.7 Estimates from Different CSL Cohorts

Because younger workers have more time to reap the benefits of human capital investments, our model predicts that, other things equal, an (expected) increase in human capital price would have a larger impact on the investment decision of younger workers. One implication of this prediction is that, for a given individual, the effect of a change in CSL for older cohorts would be larger than the effect of a similar change for younger cohorts. This occurs because an individual should be relatively younger and have a stronger incentive to adjust human capital investments when the price of human capital changes due to a tightening of the CSL faced by older cohorts. On the other hand, if the tightening of CSL occurs when an individual is approaching retirement, he would have limited incentive to respond.

To see whether this prediction is borne out in data, we break the instruments, say $\{\overline{RS}_{jt}^a\}$, into two groups. The first group $\{\overline{RS}_{jt}^a\}_{a<40}$ includes CSLs affecting schooling decisions of workers younger than 40, and the second group $\{\overline{RS}_{jt}^a\}_{a\geq 40}$ includes CSLs affecting schooling decisions of workers at age 40 and beyond. We estimate the external return to schooling using each group as instruments for average schooling $\overline{S}_{jt}$.

Table 5 reports the estimates from this exercise. The top panel reports the estimates where average schooling is instrumented by $\{\overline{CL}_{jt}^a\}_{a\geq 40}$, $\{\overline{CA}_{jt}^a\}_{a\geq 40}$ and $\{\overline{RS}_{jt}^a\}_{a\geq 40}$, and the bottom panel reports similar estimates where the instruments are $\{\overline{CL}_{jt}^a\}_{a<40}$, $\{\overline{CA}_{jt}^a\}_{a<40}$ and $\{\overline{RS}_{jt}^a\}_{a<40}$. With a smaller number of instruments than the baseline, the strength of the first stage is reduced substantially. This can be seen from the relatively small F-statistics of the joint significance of the instruments for average schooling reported in rows named F-CSL. Most of the F-statistics are only slightly above 5, suggesting that results from this exercise should be interpreted with caution.

With this caveat in mind, we find that a higher average schooling due to older workers has a much larger impact than a higher average schooling due to younger workers in all specifications. For instance, external returns are estimated to be 9.1% and statistically significant when the increase in average schooling comes from the effect of child labor laws on older workers $\{\overline{CL}_{jt}^{a}\}_{a=40}^{58}$, while the estimate is 0.8% and insignificant when the increase in average schooling comes from the effect of child labor laws on young workers $\{\overline{CL}_{jt}^{a}\}_{a=21}^{39}$. Larger estimated effects from older cohorts are also observed with the other two CSL measures. Roughly speaking, results in Table 5 are consistent with the model's prediction that the effect of a change in CSL for older cohorts is larger than the effect of a similar change for younger cohorts.

3.8 Estimates by Years of Schooling

Until now we have been assuming that external returns to schooling are homogeneous across the population. This section investigates whether and how the external return varies with individual schooling. In the model, other things equal, the response of human capital investments to a price change is increasing in the ability to learn z and, with a diminishing return $\alpha_3 < 1$, decreasing in the stock of human capital h. As better-educated individuals tend to have both a higher ability to learn z and a larger stock of human capital h, with the model parameters yet to be estimated, it is not clear how the external return would vary with individual schooling. To investigate this empirically, we divide the workers used in the baseline regression (US-born white men in their 40s with positive earnings reported) into 3 groups based on years of schooling: 0-8, 9-11 and 12+. Table 6 reports estimated external returns to schooling for each group under different models.

According to Table 6, less-educated workers benefit more from an increase in average schooling than better-educated workers. In particular, for workers with 8 years of schooling or less, one year's increase in state average schooling leads to about $8-13\%$ increase in their wages. For workers with 9-11 years of schooling, the numbers are positive but much smaller and mostly insignificant. Finally, for workers with 12 years of schooling or more, the effect is essentially zero. Results are similar when workers are grouped in other ways. For example, for workers with 9-12 years of schooling, the four estimates are: OLS 0.020 (0.022), CL 0.004 (0.023), CA 0.035 (0.025) and RS 0.009 (0.027).

3.9 Estimates from Workers Between 30 and 39 Years Old

Previous results are estimated from workers in their 40s in each census. For further evidence, we run the same regressions using workers between 30 and 39 years old in each of the three censuses. Table 7 reports the results. Panel A reports the baseline estimates where the specification is the same as the one used for Panel A of Table 2. Panel B reports the estimated external returns by using difference CSL cohorts as instruments, and the specifications are the same as those used for Table 5. Finally, Panel C reports the estimated external returns by years of schooling, where the specifications are the same as those in Table 6.

Panel A shows that both the OLS and IV estimates of external returns are around 10% and statistically significant. These estimates are slightly larger than previous estimates from workers in their 40s. Consistent with the results in Table 5, estimates in Panel B suggest that the external returns are larger when the CSLs of older cohorts are used as instruments. The differences in the two sets of estimates by CSL cohorts are smaller in panel B than the differences in Table 5.[7] This arises from the larger estimates in panel B when the CSLs of younger cohorts are used as instruments. Because of the larger estimated coefficients, all six estimates in panel B are statistically significant. While in Table 5, only three of them are significant. However, as in Table 5, the partial F-statistics of the joint significance of the instruments for $\overline{S}_{jt}$ in panel B are relatively small, suggesting that the results should be

[7] For example, when CL measures are used as the instruments, the difference in the two estimates in panel B is 0.006 (0.111-0.105). The corresponding difference in Table 5 is 0.083 (0.091-0.008).

interpreted with caution. Finally, panel C shows the estimated external returns are larger for less-educated workers, consistent with previous estimates from workers in their 40s reported in Table 6. However, different from Table 6, all but one estimate in panel C are statistically significant at conventional significance levels.

Overall, the results in Table 7 are consistent with and reinforce previous estimates from workers in their 40s.

3.10 Discussion

In summary, we find a positive and statistically significant external return to schooling γ_1. There is evidence that the external return is larger when the variation in average schooling is due to older workers, and it is larger for less-educated workers.

Our results are in contrast to Acemoglu and Angrist (2001) who find no significant effect of state average schooling $\overline{S}_{jt}$ on individual wages. In addition to the inclusion of a trend term into equation (3.1) which is shown to be important by Angrist and Pischke (2014) and Stephens and Yang (2014), the main difference between our approach and Acemoglu and Angrist (2001) is in the choice of instruments. While Acemoglu and Angrist (2001) use individual level CSL as an instrument for average schooling, we choose to use CSL effective for other workers under the guidance of a model. Under the framework of local average treatment effect proposed by Imbens and Angrist (1994), it is well understood that different instruments could lead to different estimates if the treatment effect is heterogeneous and the instruments exploit different variations in the same endogenous variable. As suggested by the different estimates in Tables 5 and 6, CSLs for different birth cohorts do result in different variations in average schooling, and the effect of average schooling does seem to be heterogeneous across individuals. It is thus not surprising that we find a significant external return to schooling while Acemoglu and Angrist (2001) didn't.

In a related paper, Ciccone and Peri (2006) argue that the Mincerian approach as the one used in this section may overestimate the external returns to schooling. This could occur due to a downward sloping demand curve for human capital. For example, assume there is no externality, and some individuals who would have chosen 0-8 years of schooling were forced to acquire 9-11 years of schooling. The wage for individuals with 0-8 (9-11) years of schooling would rise (drop) due to a lower (larger) supply. If the pool of workers with 0-8 years of schooling is much larger, the Mincerian approach could result in an overall wage increase attributed to the increase in average schooling because of the large weight assigned to the wage increase incurred by the large number of workers in that pool.

As shown in Acemoglu and Angrist (2001), CSL primarily shifts the distribution of schooling in middle- and high-school grades, and has no effect on the proportion of the population attending college.[8] In other words, CSL reduces the fraction of workers with 0-8 years of schooling and raises the fraction of workers with 9-11 years of schooling as assumed in the above example. Without external returns to schooling, an increase in the fraction of workers with 9-11 years of schooling would reduce their wages in the case of a downward sloping demand curve. Both Tables 6 and 7, however, show a positive effect of average schooling on the wage of workers with 9-11 years of schooling. This suggests that the positive external returns estimated in this section are robust to the critique in Ciccone and Peri (2006). We will return to this later in section 5.

4 Estimation

This section estimates our model using indirect inference. Indirect inference works by the selection of a set of statistics of interest which the model is asked to reproduce.[9] These statistics $\hat{\Psi}$ include the reduced form estimates from the previous section, and the complete list is described below. For an arbitrary value of the vector of parameters to be estimated β, we use the model to generate the target moments $\Psi(\beta)$. The parameter estimate $\hat{\beta}$ is then

[8] For birth cohorts between 1905 and 1959 in the 49 states that we consider, there are 2695 observations of RS. Among them, 2.41% have a value between 0 and 5, 6.42% have a value of 6, 16.92% have a value of 7, 43.86% have a value of 8, 23.41% have a value of 9, 6.01% have a value of 10, and the rest 0.96% have a value of 11. The statistics for CL and CA are similar.

[9] See Gourieroux, Monfort, and Renault (1993) for a general discussion of indirect inference.

derived by searching over the parameter space to find the parameter vector that minimizes the criterion function,

$$\hat{\beta} = \arg\min_{\beta} \left(\hat{\Psi} - \Psi(\beta)\right)' W \left(\hat{\Psi} - \Psi(\beta)\right) \tag{4.1}$$

where W is a weighting matrix. This procedure generates a consistent estimate of β. The variance-covariance matrix of the estimated parameters is

$$(G'WG)^{-1} G'W\hat{V}WG (G'WG)^{-1}$$

where G is the jacobian of $\Psi(\beta)$, and $\hat{V}$ is the variance-covariance matrix of data moments estimated using bootstrap. As discussed below, $\hat{\Psi}$ includes both state-specific moments and moments involving all states. We use a weighting matrix W specified such that each state-specific moment has a weight of one while other moments have a weight of ten. One of the moments in $\hat{\Psi}$ is the share of educational expenditure in GDP. Because we do not have individual level information on educational expenditure, this moment is not included in the bootstrap estimation of $\hat{V}$. As this moment is mainly used to identify α_2, we ignore the standard error of α_2 in the following discussion.

4.1 Predetermined Parameters

One model period is taken to be 10 years. Human capital is assumed not to depreciate ($\delta = 0$) since most of the decline in earnings towards the end of the life cycle is due to the decline in hours worked. Assuming a positive depreciation ($\delta > 0$) gives us similar results. The annual interest rate is taken to be 3%. As a result, $r = (1 + 3\%)^{10} - 1 = 0.34$. The share of physical capital in the production of the final good is set to be $\alpha = 0.33$.

$R = 4$ so that each individual lives for five periods, with periods 0, 1, 2, 3 and 4 corresponding to real-life ages of 12-21, 22-31, 32-41, 42-51 and 52-61 respectively. We assume that individuals start schooling at age 6, and all individuals have to be in school for at least

6 years. As a result, the decision problem starts from age 12. Actual years of schooling s and the amount of time spent in school in the model n_0 is related to each other through $s = 6 + 10n_0$. The space of n_0 is discretized such that the set of possible values for schooling is $S = \{6, 7, 8, 9, 10, 11, 12, 14, 16\}$.

The model is estimated using data on U.S. states. The path of $\overline{CSL}_t$ for each state is calculated using required years of schooling RS from Stephens and Yang (2014). As explained earlier, RS accounts for any changes to the compulsory attendance and child labor laws that may occur during an individual's schooling years, while other measures (CL and CA) only account for specific laws at age 14. However, results are similar when alternative measures are used. The first model period is taken to be calendar years 1911-1920, because workers making schooling decisions in this period were born in 1899-1908 and data on RS started from the cohort born in 1905. As RS is available yearly in the data, while a model period is 10 years, we calculate the average RS over the 10 years covered by each period and use it as $\overline{CSL}_t$. As RS is available until 1961, we can calculate $\overline{CSL}_t$ for the first seven model periods directly, and assume $\overline{CSL}_t = \overline{CSL}_7$ for all $t \geq 7$.[10]

4.2 Estimated Parameters

Each state is allowed to have its own paths of TFP A_t. The state-specific TFP A_t is allowed to grow in the first 15 periods with a state-specific growth rate g, and is assumed to stay constant at its value in the 15th period after that. This gives us two parameters for each state: the initial level A_1 and growth rate g of TFP in the first 15 periods. The distribution of innate ability z and initial stock of human capital h_0 is also state-specific and follows a joint log normal distribution given by

$$\begin{pmatrix} \log z \\ \log h_0 \end{pmatrix} \sim Normal \left[\begin{pmatrix} \mu_z \\ \mu_h \end{pmatrix}, \begin{pmatrix} \sigma_z^2 & \\ \rho_{z,h}\sigma_z\sigma_h & \sigma_h^2 \end{pmatrix} \right]$$

[10]Due to data limitations, $\overline{CSL}_1$ is calculated from $\{RS_t\}_{t=1905}^{1908}$, while $\overline{CSL}_7$ is calculated from $\{RS_t\}_{t=1959}^{1961}$.

Besides the state-specific parameters, we also need to estimate five parameters common to all states, including the externality parameter θ, the parameters governing human capital production $(\alpha_1, \alpha_2, \alpha_3)$, and the ratio of learning ability at work relative to the learning ability in school c. Table 8 lists the set of parameters to be estimated.

4.3 Moments and Identification

Two sets of moments are used to estimate the model parameters. The first set of moments are calculated using data from all states and include

1. OLS and IV estimates of external returns to schooling for all workers. From Table 2, the OLS and IV (with RS) estimates are 0.069 and 0.061 respectively.

2. OLS and IV estimates of external returns to schooling by education. We include the estimates for workers with 9-11 years of schooling and workers with 12 or more years of schooling from Table 6. It is important to note that, although we refer to the statistics as either OLS or IV estimate, indirect inference does not require either of them to be causal or unbiased. All statistics used in indirect inference are just different ways of summarizing the data. The only requirement is that these statistics are calculated in the same way using simulated data from the model.

3. OLS and IV estimates of the private return to schooling. The OLS estimate is 0.072. In a wage equation with region-by-YOB dummies where individual schooling is instrumented with RS, Stephens and Yang (2014) find the IV estimate of the private return to schooling is about -0.02 and not statistically significant. Applying the same strategy to our data produces an insignificant estimate of -0.1. We use -0.1 as the baseline but results are similar when -0.02 is used. Again, we are only using the IV estimate as one way to summarize the data. It does not mean that we believe a negative private return to schooling.

4. The share of educational expenditure in GDP. As the model ignores elementary education, we match the expenditure on secondary and tertiary education as a share of GDP, which is about 4% in 1970 according to the Digest of Education Statistics.[11]

The second set of moments are state-specific and include

1. Average wage of workers aged 22-61 in 1960, 1970 and 1980.

2. The distribution of schooling in 1960, 1970 and 1980. Three moments of the distribution in each year are used: average years of schooling, the fraction with at most 8 years of schooling, and the fraction with at most 12 years of schooling. All moments are calculated using workers aged 22-61.

3. The Mincerian return to schooling in 1960, 1970 and 1980, calculated from workers between 42 and 51 years old in each census.

4. Wage growth of workers between 22 and 31 years old in 1960. We first calculate their average wage in 1960, and then calculate the average wage of workers between 42 and 51 years old in 1980. With these two numbers, we can calculate wage growth assuming that the two numbers are for the same cohort of workers.

All moments are calculated using census data. Individual schooling in the model must be greater than or equal to $\overline{CSL_t}$ set by the government, while in the data there are workers with schooling below the relevant requirement. To make sure the model and data moments are comparable, in the data we set $s_i^* = \max\{s_i, RS\}$ before calculating relevant moments based on schooling. All data moments related to schooling are calculated using s_i^*.

Although the parameters are identified jointly from all moments, some moments are especially important for the identification of particular parameters. For example, IV estimates of external returns to schooling are critical in identifying θ. Other things equal, a larger θ leads to a larger γ_1. The share of educational expenditure is particularly useful in identifying α_2,

[11] http://nces.ed.gov/pubs2012/2012001.pdf.

the elasticity of human capital production in school with respect to the goods input. Wage growth between 1960 and 1980 for the cohort of workers between 22 and 31 years old in 1960 is particularly important for identifying α_3, the human capital production technology during the working phase. Conditional on other parameters, estimates of the private return to schooling are used to identify α_1. The state-specific distribution of schooling and the Mincerian returns to schooling are mainly used to identify the state-specific distribution of ability and initial human capital $F(z, h_0)$, while the state-specific average wages in 1960-1980 are mainly used to identify the initial level A_1 and growth rate g of TFP.[12]

4.4 Estimation Results

Given a set of parameters β, we solve for transition dynamics state by state and use the simulated data to calculate the statistics $\Psi(\beta)$. The set of parameters that minimizes the criterion function (4.1) will be our estimated parameters. Appendix A provides more details on model computation and estimation.

Table 8 reports the estimated parameters. Tables 9 and 10 report the relevant data and model moments. For brevity, state-specific parameters and moments are not reported here but their averages across states are reported instead. The estimated return to scale for the schooling technology is 0.663 (0.529+0.134), and it is 0.574 for the training technology. Both are within the range of estimates reported in Browning, Hansen, and Heckman (1999). The effective learning ability at work is estimated to be about 94.7% of the learning ability in school. The average decadal growth rate of TFP is estimated to be 10.5%, implying an annual rate of 1%. This is substantially smaller than traditional estimates, due to the presence of human capital externalities. The two initial conditions z and h_0 are estimated to be positively correlated, with a correlation coefficient of 0.153. The elasticity of firm productivity with respect to average human capital θ, the structural measure of human capital externalities, is estimated to be 0.121, suggesting that a 1% increase in average human capital raises the

[12] Results are similar when the average wages are replaced with the state-specific per capita income in 1960-1980 from the Bureau of Economic Analysis. `http://www.bea.gov/iTable/index_regional.cfm`.

productivity of a typical firm by about 0.121%.

The model moments calculated with the estimated parameters are generally comparable to the corresponding data moments. The OLS and IV estimates of external returns to schooling from all workers are only slightly larger in the model than in the data (0.072 vs 0.069 and 0.063 vs 0.061), and we are able to match the OLS and IV estimates of external returns by schooling almost exactly. The OLS and IV estimates of the private return to schooling are slightly smaller in the model than in the data (0.068 vs 0.072 and -0.106 vs -0.1), so is the share of educational expenditure in GDP (0.038 vs 0.4). However, the differences in all three cases are very small. The model underpredicts the growth of schooling and the increase in Mincerian returns over time as well as the cohort-specific wage growth. One explanation may be that we are abstracting from the evolution of human capital production technologies over time. Overall, the match between model and data is reasonably good.

To provide some evidence on the effect of θ and how it is identified, we re-estimate the model by setting $\theta = 0$ and use the simulated data to calculate moments reported in the last column of Tables 9 and 10. With $\theta = 0$, the model is able to match the OLS estimates of external returns as well as other moments reasonably well. However, the model can no longer match the IV estimates of external returns to schooling either for all workers or by schooling, all estimates are now very close to zero. This suggests that θ is identified from the IV estimates of external returns to schooling γ_1 and a positive θ is required to generate the positive estimates of γ_1.

In summary, we estimate a positive human capital externality: a 1% increase in average human capital raises the productivity of a typical firm by about 0.121%. This positive externality is required to account for the positive effect of average schooling on individual wages estimated earlier in the paper.

5 Heterogeneous Human Capital

Our baseline model assumes that human capital is homogeneous and subscribes to an efficiency units view of human capital. Assuming that different types of human capital are imperfect substitutes, Ciccone and Peri (2006) argue that the Mincerian approach to human capital externalities (as in equation (3.1)) generates an estimator subject to an upward bias. In this case, a positive estimate of γ_1 from equation (3.1) may be consistent with a production function without human capital externalities $\theta = 0$. Earlier in this paper, we show that the positive external returns for workers with 9-11 years of schooling suggest that our estimates are robust to this critique. To provide further evidence, in this section, we extend our baseline model to include two types of human capital (high skilled and low skilled) that are imperfect substitutes. The goal is to demonstrate that the presence of complementarities between the two types of human capital by itself is insufficient to generate a strongly positive estimate of γ_1, and a positive θ is required to match the empirical estimate of γ_1 from section 3.

Assume there is no externality. For simplicity, we focus on steady states and ignore the time subscript t. We also assume that all individuals start their lives with the same amount of initial human capital h_0, and they differ from each other only in terms of innate learning ability z, the distribution of which is given by $F(z)$.

Let w_H be the price of high skilled human capital h_H, and w_L be the price of low skilled human capital h_L. Given these two prices, an individual chooses the level of schooling and, in turn, whether to be a high skilled worker or a low skilled worker. In order to be a high skilled worker, an individual has to pay a fixed cost χ and years of schooling must exceed some threshold $\bar{S}$. Let $V(z; w_H)$ be the PDV of lifetime income for an individual with ability z who chooses to be a high skilled worker, and $V(z; w_L)$ be the corresponding value if the individual chooses to be a low skilled worker. Define the cut-off value of ability $\hat{z}$ by

$$V(\hat{z}; w_H) - \chi = V(\hat{z}; w_L)$$

In equilibrium, individuals with $z \geq \hat{z}$ will become high skilled, and other individuals will be low skilled.

Let $h_a(z)$ and $n_a(z)$ be the stock of human capital and time spent on training for an individual with ability z at age a, respectively. The total amount of human capital used in the production of the final good will be given by

$$\begin{aligned} h_H &= \int_{\hat{z}} \sum_{a=0}^{R} \left[h_a(z)\left(1 - n_a(z)\right)\right] dF(z) \\ h_L &= \int^{\hat{z}} \sum_{a=0}^{R} \left[h_a(z)\left(1 - n_a(z)\right)\right] dF(z) \end{aligned}$$

The two types of human capital are imperfect substitutes. In particular, the effective amount of human capital used in production is given by

$$h = \left[\kappa h_H^{\rho} + (1-\kappa) h_L^{\rho}\right]^{\frac{1}{\rho}}$$

with $\kappa \in (0,1)$ and $\rho \leq 1$. In the limit, when $\rho = 1$, the two types of human capital will be perfect substitutes, and the model is reduced to our baseline model when $\kappa = 0.5$. Otherwise, for $\rho < 1$, the two types of human capital will be complements.

Assume each human capital is paid by their marginal product, we have

$$\frac{w_H}{w_L} = \frac{\kappa}{1-\kappa}\left(\frac{h_H}{h_L}\right)^{\rho-1}$$

For $\rho < 1$, the relative wage $\frac{w_H}{w_L}$ is decreasing in the relative supply of human capital $\frac{h_H}{h_L}$. Otherwise, for $\rho = 1$, the relative wage is independent of the relative supply.

Because the typical value for required years of schooling RS is 8 and the largest value of RS in the data is 11,[13] the direct effect of CSL is on h_L if $\bar{S} > 11$. In particular, h_L will increase with RS if some low skilled workers are forced to acquire more human capital

[13]Specifically, according to the data from Stephens and Yang (2014), $\Pr(RS \leq 7) = 0.25$, $\Pr(RS = 8) = 0.43$, $\Pr(RS = 9) = 0.24$, $\Pr(RS = 10) = 0.06$, $\Pr(RS = 11) = 0.01$.

through schooling. With complementarity $\rho < 1$, an increase in h_L will raise $\frac{w_H}{w_L}$, which may induce more individuals to become high skilled. That is, the cut-off value of ability $\hat{z}$ is decreasing in RS. Because of this, a tighter CSL reduces the average ability of both types of workers through selection, and this is the extra effect induced by complementarity. Without complementarity ($\rho = 1$), because the relative wage is independent of the relative supply of human capital, an increase in h_L due to a higher RS has no effect on individual decisions on whether to be a high skilled or low skilled worker.

When the negative effect of RS on $\hat{z}$ is strong enough, the model with complementarity but no externality predicts a negative effect of average schooling on earnings $\gamma_1 < 0$, and a larger θ is required to match the empirical estimates of γ_1. This is confirmed though numerical simulations reported in table 11. For each value of ρ, we calibrate the model to obtain an estimate of θ that matches the empirical estimates of γ_1.[14] We normalize the estimate of θ in the case of $\rho = 1$ to be one. Table 11 shows that the estimate of θ is decreasing in ρ. This is expected because a smaller ρ implies a larger elasticity of the relative wage $\frac{w_H}{w_L}$ with respect to the relative supply $\frac{h_H}{h_L}$, which, through the effect of relative wage on the cut-off value of ability $\hat{z}$, leads to a larger negative effect of RS on the average ability of workers within a given type. Katz and Murphy (1992) estimate that the elasticity of substitution between college and high school workers is about 1.41. According to Table 11, an elasticity of 1.41 means that the magnitude of human capital externalities θ is 5-7% larger than our baseline estimate.

The results in Table 11 suggest that complementarity by itself is insufficient in generating a large positive external return γ_1. With complementarity, we need an even larger externality θ to match the empirical estimates of γ_1. The key difference between our model and that of Ciccone and Peri (2006) is that we allow individual schooling as well as the decision to become a high skilled worker to be endogenous. This endogenous response affects the composition of workers with the same level of schooling or the same type of human capital in states with

[14] Appendix B describes how this is done in detail.

different CSLs. The composition effect is large enough to keep the model from generating a positive estimate of external returns in the absence of human capital externalities.

6 Conclusion

This paper estimates human capital externalities formulated in Lucas (1988). We incorporate externalities into an overlapping generations model of human capital accumulation with Compulsory Schooling Laws (CSL). The model implies that human capital externalities can be estimated from the effect of CSL affecting the schooling decision of one generation on the wage of other generations. Using an instrumental variable strategy deduced from the model, we find that one more year of average schooling at the U.S. state level raises individual wage by about 6-8%. We also find the effect is larger when the exogenous variation in average schooling is driven by the tightening of CSL for older cohorts, and it is larger for less-educated workers. Our estimates are robust to different empirical specifications.

Taking the reduced form estimates into account, we estimate the model by indirect inference. We solve for the transitional dynamics state by state, use the simulated data to calculate moments on earnings and schooling and match them with actual data. We find that the elasticity of a typical firm's productivity with respect to the average human capital of an economy is 0.121. We show that a positive human capital externality is required to match the estimated effect of state average schooling in a model featuring complementarity across two types of human capital.

Our findings are complementary to Gennaioli et al. (2013) who find overwhelming evidence that human capital fosters development through entrepreneurial education and human capital externalities. This paper also complements Choi (2011) and Malley and Woitek (2017), both of which find significant learning externalities along the lines of Tamura (1991). An interesting direction for future research is to estimate the externalities in Lucas (1988) and Tamura (1991) jointly.

References

Acemoglu, Daron and Joshua Angrist. 2001. "How Large are Human-Capital Externalities? Evidence from Compulsory-Schooling Laws." In *NBER Macroeconomics Annual 2000, Volume 15*, NBER Chapters. National Bureau of Economic Research, Inc, 9–74.

Angrist, Joshua D. and Alan B. Krueger. 1991. "Does Compulsory School Attendance Affect Schooling and Earnings?" *The Quarterly Journal of Economics* 106 (4):979–1014.

Angrist, Joshua D. and Jorn-Steffen Pischke. 2014. *Mastering 'Metrics: The Path from Cause to Effect*. Princeton University Press.

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy* 75:352.

Browning, Martin, Lars Peter Hansen, and James J. Heckman. 1999. "Micro data and general equilibrium models." In *Handbook of Macroeconomics*, *Handbook of Macroeconomics*, vol. 1, edited by J. B. Taylor and M. Woodford, chap. 8. Elsevier, 543–633.

Choi, Seung Mo. 2011. "How Large Are Learning Externalities?" *International Economic Review* 52 (4):1077–1103.

Ciccone, Antonio and Giovanni Peri. 2006. "Identifying Human-Capital Externalities: Theory with Applications." *Review of Economic Studies* 73 (2):381–412.

Davies, Jim. 2003. "Empirical Evidence on Human Capital Externalities." University of Western Ontario, Economic Policy Research Institute Working Papers 20035, University of Western Ontario, Economic Policy Research Institute.

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer. 2013. "Human Capital and Regional Development." *The Quarterly Journal of Economics* 128 (1):105–164.

Goldin, Claudia and Lawrence Katz. 2008. *The Race Between Education and Technology.* Cambridge, MA: Belknap Press.

Gourieroux, Christian, Alain Monfort, and Eric Renault. 1993. "Indirect inference." *Journal of applied econometrics* 8 (S1):S85–S118.

Imbens, Guido W and Joshua D Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62 (2):467–475.

Katz, Lawrence F and Kevin M Murphy. 1992. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *The Quarterly Journal of Economics* 107 (1):35–78.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman. forthcominga. "Life-Cycle Human Capital Accumulation Across Countries: Lessons From U.S. Immigrants." *Journal of Human Capital* .

———. forthcomingb. "Life-Cycle Wage Growth Across Countries." *Journal of Political Economy* .

Lochner, Lance. 2011. "Nonproduction Benefits of Education: Crime, Health, and Good Citizenship." In *Handbook of the Economics of Education*, vol. 4, edited by Erik Hanushek, Stephen Machin, and Ludger Woessmann, chap. 2. Elsevier, 183–282.

Lucas, Robert Jr. 1988. "On the mechanics of economic development." *Journal of Monetary Economics* 22 (1):3–42.

Malley, James and Ulrich Woitek. 2017. "Estimated Human Capital Externalities in an Endogenous Growth Framework." Working paper, University of Glasgow.

Moretti, Enrico. 2004a. "Estimating the social return to higher education: evidence from longitudinal and repeated cross-sectional data." *Journal of Econometrics* 121 (1-2):175–212.

———. 2004b. "Human capital externalities in cities." In *Handbook of Regional and Urban Economics*, vol. 4, edited by J. V. Henderson and J. F. Thisse, chap. 51. Elsevier, 2243–2291.

———. 2004c. "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions." *American Economic Review* 94 (3):656–690.

Rauch, James E. 1993. "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities." *Journal of Urban Economics* 34 (3):380–400.

Stephens, Melvin and Dou-Yan Yang. 2014. "Compulsory Education and the Benefits of Schooling." *American Economic Review* 104 (6):1777–92.

Tamura, Robert. 1991. "Income Convergence in an Endogenous Growth Model." *Journal of Political Economy* 99 (3):522–40.

Appendix A: Baseline Model Estimation

This appendix describes how we calculate the model and estimate the parameters. We assume each state was in its own initial steady state in the first period with state-specific A_1 and $\overline{CSL}_1$, and reaches its new steady state in period $T = 150$ with state-specific A_{15} and $\overline{CSL}_7$. To estimate the model parameters, we proceed as follows,

1. Start with an initial guess of model parameters.

2. Solve for the transitional path for each state as follows

   (a) Start with a guess of the path of average human capital $\{\overline{H}_{t,guess}\}$, which, combined with the path of TFP $\{A_t\}$, allows us to solve for the path of human capital prices $\{w_t\}$ using equation (2.3).

   (b) With $\{w_t\}$, we can solve the problem in equation (2.1) for each worker in each period.

   (c) Given the solution to each worker's problem in each period, we can calculate the implied average human capital $\bar{H}_t$ in each period using equation (2.4).

   (d) Calculate the distance between the two paths $\{\bar{H}_{t,guess}\}$ and $\{\bar{H}_t\}$. In practice, the distance is defined to be $d = \max_t \left\{ \frac{\left|\bar{H}_t - \bar{H}_{t,guess}\right|}{\overline{H}_{t,guess}} \right\}$.

   (e) If d is smaller than some tolerance value, move on to the next step. Otherwise, set $\bar{H}_{t,guess} = \bar{H}_t$ for each t and redo (a)-(d).

3. Given the solution to each worker's problem in each period for each state, we can calculate the relevant moments with the data simulated from the model.

4. Calculate the distance between the model moments and the corresponding data moments.

5. Redo steps 1-3 with another guess of model parameters. Keep iterating until the distance between model moments and data moments is minimized. The set of parameters that minimizes this distance will be our estimated parameters.

Appendix B: Calibration of the Model with Heterogeneous Human Capital

This appendix documents how we calibrate the model with heterogeneous human capital and use it to generate the results reported in Table 11.

For simplicity, we ignore transitional dynamics and consider steady states only. Normalize TFP $A = 1$ and initial human capital $h_0 = 1$. We assume that a worker become skilled with at least one year of college education.

Assume $\log z \sim Normal\left(\mu_z, \sigma_z^2\right)$. Relative to the baseline case, the model with heterogeneous human capital introduces five additional parameters $(\rho, \kappa, \chi, \mu_z, \sigma_z)$. Given a value of ρ, we can calibrate θ and $(\kappa, \chi, \mu_z, \sigma_z)$. To do so, we treat the 49 states as a single economy and use the 1960 census to calculate four data moments:[15] average schooling (mainly used to identify mean ability μ_z), standard deviation of schooling (mainly used to identify ability dispersion σ_z), fraction of workers with at least one year of college education (mainly used to identify the fixed cost of becoming a skilled worker χ), and the college wage premium (mainly used to identify the relative efficiency κ of skilled human capital) estimated from a regression of log wage on years of schooling and an indicator for college education. We use the empirical estimate of γ_1 as an additional moment to identify θ.

For each value of ρ, we

1. Start with an initial guess of $(\theta, \kappa, \chi, \mu_z, \sigma_z)$.

2. Set CSL to a value RS_1 measuring the average required years of schooling for respon-

[15] Results from 1970 and 1980 censuses are similar.

dents in 1960 census.

3. Discretize the ability distribution $\log z \sim Normal\left(\mu_z, \sigma_z^2\right)$.

4. Solve the lifecycle problem for each worker with a given ability.

5. Use data from step 4 to calculate the model counterparts of the four moments calculated from the census.

6. Set CSL to a different value $RS_2 \neq RS_1$.

7. Solve the lifecycle problem for each worker under the new CSL.

8. Combine data from steps 4 and 7 to estimate the equation $\log Y_{i,j} = b_0 + b_1 \bar{S}_j + b_2 s_{i,j} + \xi_{i,j}$, where i is an index for worker, $j = \{1, 2\}$ is an index for the two cases with different CSLs, $Y_{i,j}$ and $s_{i,j}$ are the wage and schooling of worker i in case j, $\bar{S}_j$ is the average schooling in case j, and $\xi_{i,j}$ is the error term. b_1 is the model counterpart of γ_1.

9. Calculate the distance between the five model moments and the corresponding data moments.

10. Redo steps 1-9 with another guess of $(\theta, \kappa, \chi, \mu_z, \sigma_z)$. Keep iterating until the distance between model moments and data moments is minimized. The set of parameters that minimizes this distance will be our estimates.

Note that the solutions to workers' problem in steps 4 and 7 require solving a fixed points for (h_H, h_L). That is, we need to start with an initial guess of the equilibrium stock of high and low skilled human capital (h_H, h_L), use them to solve for the human capital prices (w_H, w_L), solve the workers' problem under these prices, use the solutions to calculate the implied (h_H, h_L) and make sure that they match the initial guesses.

Table 1: Effects of CSLs on State Average Schooling

	CL		CA		RS	
Age	coef.	s.e.	coef.	s.e.	coef.	s.e.
21	0.065	0.147	0.062	0.047	-0.242	0.065
22	0.066	0.133	0.031	0.046	0.420	0.120
23	-0.470	0.142	-0.039	0.042	-0.146	0.140
24	0.084	0.123	-0.002	0.043	0.241	0.132
25	0.226	0.082	-0.089	0.031	-0.370	0.149
26	-0.502	0.137	-0.301	0.051	0.319	0.107
27	-0.049	0.128	0.051	0.043	-0.243	0.108
28	0.727	0.162	0.232	0.050	-0.328	0.085
29	-0.121	0.096	0.111	0.033	0.205	0.103
30	-0.112	0.068	-0.064	0.021	0.146	0.104
31	-0.058	0.127	0.000	0.057	-0.087	0.087
32	-0.084	0.128	-0.075	0.068	-0.085	0.146
33	-0.208	0.144	0.089	0.056	0.181	0.093
34	0.341	0.186	0.047	0.053	0.055	0.113
35	0.099	0.068	-0.134	0.025	0.268	0.106
36	0.134	0.102	-0.166	0.057	-0.292	0.108
37	-0.335	0.182	0.271	0.046	-0.378	0.114
38	0.350	0.170	-0.082	0.046	0.261	0.094
39	-0.078	0.067	0.180	0.036	0.192	0.094
40	-0.098	0.077	-0.138	0.037	0.024	0.143
41	-0.587	0.154	-0.029	0.057	-0.429	0.142
42	-0.098	0.143	0.126	0.061	0.321	0.117
43	0.376	0.144	0.015	0.048	-0.234	0.126
44	0.344	0.193	-0.036	0.057	0.762	0.141
45	0.142	0.102	-0.153	0.036	-0.343	0.132
46	0.176	0.136	0.042	0.048	-0.281	0.114
47	-0.029	0.113	0.093	0.047	-0.149	0.103
48	-0.032	0.120	-0.117	0.039	0.307	0.081
49	-0.151	0.077	0.190	0.041	0.040	0.085
50	0.091	0.044	0.030	0.009	0.096	0.093
51	-0.027	0.112	-0.083	0.047	-0.535	0.103
52	0.026	0.140	0.166	0.047	0.199	0.079
53	-0.161	0.153	-0.090	0.049	0.343	0.074
54	0.050	0.113	-0.041	0.045	0.065	0.051
55	-0.022	0.115	0.019	0.040	-0.162	0.040
56	-0.200	0.093	0.012	0.038		
57	0.258	0.075	-0.068	0.039		
58	0.025	0.011	0.046	0.010		
F-CSL	26.720		12.805		22.743	

Table 2: Baseline Estimates of External Returns to Schooling

	OLS	CL	CA	RS
Panel A: Individual schooling is exogenous				
Private Return	0.072***	0.072***	0.072***	0.072***
to Schooling	(0.001)	(0.001)	(0.001)	(0.001)
External Return	0.069***	0.069***	0.081***	0.061**
to Schooling	(0.019)	(0.022)	(0.023)	(0.024)
Panel B: Individual schooling is endogenous				
Private Return		0.061***	0.087***	0.069***
to Schooling		(0.018)	(0.015)	(0.022)
External Return		0.073***	0.076***	0.061**
to Schooling		(0.024)	(0.026)	(0.027)
F-CSL		26.725	12.805	22.753
F-QOB		5.568	5.612	5.583

Notes. All regressions use data on US-born white men in their 40s with positive earnings reported in decennial censuses 1960-1980. The dependent variable is log weekly wage. The key explanatory variables are individual schooling and the average schooling of all workers (not restricted to US-born white men in their 40s with positive earnings) in the worker's state of residence in the relevant census. The coefficients in front of these two variables are referred to as the private return to schooling (for individual schooling) and the external return to schooling (for average schooling), and their estimates are reported in the table. Other control variables included in all columns are dummies for state of birth, state of residence (SOR), year of birth, year of census and interactions between year of birth dummies and dummies for the four census regions. The first column reports OLS estimates. The second to fourth columns report 2SLS estimates where average schooling is instrumented by child labor laws CL, (corrected) compulsory attendance laws CA and a composite measure of required schooling RS respectively. See the main text for the construction of these measures of compulsory schooling laws CSLs. Individual schooling is treated as exogenous in panel A but endogenous in panel B. F-CSL reports the (partial) F-statistics of joint significance of CSL measures in the first-stage regression of average schooling on QOB dummies, relevant CSL measures and all explanatory variables used in the main wage equation other than individual schooling and average schooling. Similarly, F-QOB reports the (partial) F-statistics of joint significance of QOB dummies in the first-stage regression of individual schooling on the same set of variables. All standard errors reported in the parentheses are clustered to the SOR-by-census level. *** denotes 1%, ** denotes 5%, and * denotes 10% significance.

Table 3: Estimates of External Returns from Alternative Specifications

	CL	CA	RS
Panel A: Control for own CSL			
Private Return	0.072***	0.072***	0.072***
to Schooling	(0.001)	(0.001)	(0.001)
External Return	0.072***	0.083***	0.061**
to Schooling	(0.022)	(0.024)	(0.024)
F-CSL	27.341	12.620	22.849
Panel B: Excluding $\{\overline{CSL}_{jt}^{a}\}_{a=40}^{50}$			
Private Return	0.072***	0.072***	0.072***
to Schooling	(0.001)	(0.001)	(0.001)
External Return	0.077***	0.084***	0.066**
to Schooling	(0.024)	(0.025)	(0.030)
F-CSL	14.780	9.725	11.804

Notes. All regressions use data on US-born white men in their 40s with positive earnings reported in decennial censuses 1960-1980. The dependent variable is log weekly wage. The key explanatory variables are individual schooling and the average schooling of all workers (not restricted to US-born white men in their 40s with positive earnings) in the worker's state of residence in the relevant census. The coefficients in front of these two variables are referred to as the private return to schooling (for individual schooling) and the external return to schooling (for average schooling), and their estimates are reported in the table. Other control variables included in all columns are dummies for state of birth, state of residence SOR, year of birth, year of census and interactions between year of birth dummies and dummies for the four census regions. The three columns report 2SLS estimates where average schooling is instrumented by child labor laws CL, (corrected) compulsory attendance laws CA and a composite measure of required schooling RS respectively. See the main text for the construction of these measures of compulsory schooling laws CSLs at the individual level and their age-by-SOR averages, the latter of which are used as instruments for average schooling. Panel A includes individual level CSL as an additional control in both the first and the second stage, while panel B excludes the CSLs effective for workers in their 40s from the first stage. F-CSL reports the partial F-statistics of joint significance of CSL measures for average schooling in the first stage. All standard errors reported in the parentheses are clustered to the SOR-by-census level. *** denotes 1%, ** denotes 5%, and * denotes 10% significance.

Table 4: Estimates of External Returns with SOR Instruments

	CL	CA	RS
Private Return	0.072***	0.072***	0.072***
to Schooling	(0.001)	(0.001)	(0.001)
External Return	0.063**	0.107***	0.065***
to Schooling	(0.026)	(0.030)	(0.023)
F-CSL	18.112	16.680	49.005

Notes. All regressions use data on US-born white men in their 40s with positive earnings reported in decennial censuses 1960-1980. The dependent variable is log weekly wage. The key explanatory variables are individual schooling and the average schooling of all workers (not restricted to US-born white men in their 40s with positive earnings) in the worker's state of residence in the relevant census. The coefficients in front of these two variables are referred to as the private return to schooling (for individual schooling) and the external return to schooling (for average schooling), and their estimates are reported in the table. Other control variables included in all columns are dummies for state of birth, state of residence SOR, year of birth, year of census and interactions between year of birth dummies and dummies for the four census regions. The three columns report 2SLS estimates where average schooling is instrumented by child labor laws CL, (corrected) compulsory attendance laws CA and a composite measure of required schooling RS respectively. The difference from Tables 2 and 3 is that here the three measures of compulsory schooling laws CSLs are calculated by assuming that individuals were educated in their current SOR instead of their state of birth. F-CSL reports the partial F-statistics of joint significance of CSL measures for average schooling in the first stage. All standard errors reported in the parentheses are clustered to the SOR-by-census level. *** denotes 1%, ** denotes 5%, and * denotes 10% significance.

Table 5: Estimates of External Returns by CSL Cohorts

CSL Age Cohort	CL	CA	RS
40-58	0.091***	0.118***	0.041
	(0.029)	(0.027)	(0.031)
F-CSL	8.412	4.412	6.182
21-39	0.008	0.072***	0.022
	(0.033)	(0.026)	(0.034)
F-CSL	5.118	8.031	5.807

Notes. All regressions use data on US-born white men in their 40s with positive earnings reported in decennial censuses 1960-1980. The dependent variable is log weekly wage. The key explanatory variable is the average schooling of all workers (not restricted to US-born white men in their 40s with positive earnings) in the worker's state of residence in the relevant census. The coefficient in front of the average schooling is referred to as the external return to schooling and its estimates are reported in the table. Other control variables included in all regressions are individual schooling and dummies for state of birth, state of residence SOR, year of birth, year of census and interactions between year of birth dummies and dummies for four census regions. The three columns report 2SLS estimates where average schooling is instrumented by child labor laws CL, (corrected) compulsory attendance laws CA and a composite measure of required schooling RS respectively. See the main text for the construction of these measures of compulsory schooling laws CSLs. The upper (lower) panel reports estimates where the instruments are CSLs affecting schooling decisions of workers aged 40-58 (21-39). F-CSL reports the partial F-statistics of joint significance of CSL measures for average schooling in the first stage. All standard errors reported in the parentheses are clustered to the SOR-by-census level. *** denotes 1%, ** denotes 5%, and * denotes 10% significance.

Table 6: Estimates of External Returns by Years of Schooling

Years of Schooling	OLS	CL	CA	RS
0-8	0.114***	0.119***	0.130***	0.087**
	(0.029)	(0.033)	(0.034)	(0.035)
F-CSL		38.697	17.421	26.051
9-11	0.048*	0.037	0.080***	0.025
	(0.025)	(0.029)	(0.029)	(0.030)
F-CSL		32.932	12.701	25.349
12+	-0.007	-0.006	-0.001	-0.010
	(0.017)	(0.021)	(0.021)	(0.022)
F-CSL		23.309	12.664	22.788

Notes. The main data uses information on US-born white men in their 40s with positive earnings reported in decennial censuses 1960-1980. The top panel uses the sample of workers with at most 8 years of schooling, the middle panel uses workers with 9-11 years of schooling, and the bottom panel uses workers with at least 12 years of schooling. The dependent variable is log weekly wage in all regressions. The key explanatory variable is the average schooling of all workers (not restricted to US-born white men in their 40s with positive earnings) in the worker's state of residence in the relevant census. The coefficient in front of this variable is referred to as the external return to schooling and its estimates are reported in the table. Other control variables included in all columns are dummies for state of birth, state of residence SOR, year of birth, year of census and interactions between year of birth dummies and dummies for the four census regions. The first column reports OLS estimates. The second to fourth columns report 2SLS estimates where average schooling is instrumented by child labor laws CL, (corrected) compulsory attendance laws CA and a composite measure of required schooling RS respectively. See the main text for the construction of these measures of compulsory schooling laws CSLs. F-CSL reports the partial F-statistics of joint significance of CSL measures for average schooling in the first stage. All standard errors reported in the parentheses are clustered to the SOR-by-census level. *** denotes 1%, ** denotes 5%, and * denotes 10% significance.

Table 7: Estimates of External Returns from Workers Between 30 and 39 Years Old

	OLS	CL	CA	RS
Panel A: Baseline Estimates				
Private Return	0.063***	0.063***	0.063***	0.063***
to Schooling	(0.001)	(0.001)	(0.001)	(0.001)
External Return	0.098***	0.104***	0.108***	0.104***
to Schooling	(0.015)	(0.016)	(0.017)	(0.015)
F-CSL		28.637	14.272	26.037
Panel B: External Returns by CSL Age Cohorts				
40-58		0.111***	0.138***	0.080***
		(0.024)	(0.025)	(0.021)
F-CSL		9.272	4.604	6.990
21-39		0.105***	0.107***	0.105***
		(0.027)	(0.024)	(0.021)
F-CSL		5.744	8.447	6.518
Panel C: External Returns by Years of Schooling				
0-8	0.164***	0.181***	0.171***	0.165***
	(0.029)	(0.035)	(0.033)	(0.032)
F-CSL		44.643	21.275	36.577
9-11	0.044**	0.051*	0.049*	0.040
	(0.021)	(0.028)	(0.025)	(0.024)
F-CSL		36.787	15.449	30.137
12+	0.047**	0.037**	0.059***	0.043**
	(0.017)	(0.016)	(0.019)	(0.019)
F-CSL		25.918	13.787	25.090

Notes. The main data uses information on US-born white men in their 30s with positive earnings reported in decennial censuses 1960-1980. The specifications used in Panels A, B and C are the same as those in Panel A of Table 2, Table 5 and Table 6, respectively. *** denotes 1%, ** denotes 5%, and * denotes 10% significance.

Table 8: Estimated Parameters

Parameter	Notation	Value
Human capital externality	θ	0.121***
		(0.038)
Human capital production		
	α_1	0.529***
		(0.001)
	α_2	0.134
		–
	α_3	0.574***
		(0.010)
Ability at work relative to ability in school	c	0.947***
		(0.049)
Average of state-specific parameters		
Initial TFP	A_1	0.974
Decadal Growth rate of TFP	g	0.105
Distribution of ability and initial human capital		
	μ_z	0.518
	μ_h	4.423
	σ_z	0.543
	σ_h	0.656
	$\rho_{z,h}$	0.153

Notes. This table reports the estimated model parameters. Standard errors are in the parentheses. *** denotes 1%, ** denotes 5%, and * denotes 10% significance.

Table 9: Key Moments: Data vs Model

	Data	Model	$\theta = 0$
External returns to schooling			
All workers			
OLS	0.069	0.072	0.079
IV	0.061	0.063	-0.007
Workers with 9 to 11 years of schooling			
OLS	0.048	0.046	0.046
IV	0.025	0.025	-0.004
Workers with 12 years of schooling or more			
OLS	-0.007	-0.007	-0.010
IV	-0.010	-0.010	0.000
Other moments involving all states			
Private return to schooling: OLS	0.072	0.068	0.070
Private return to schooling: IV	-0.100	-0.106	-0.110
Educ exp/GDP in 1970	0.040	0.038	0.038

Notes. This table reports the data and model moments calculated using information from all states.

Table 10: State-Specific Moments: Data vs Model

	Data	Model	$\theta = 0$
Normalized average wage			
1960	1.183	1.172	1.181
1970	1.487	1.354	1.365
1980	1.451	1.458	1.471
Average years of schooling			
1960	10.741	10.972	10.546
1970	11.627	11.289	10.817
1980	12.528	11.569	11.106
Fraction of workers $S \leq 8$			
1960	0.312	0.279	0.331
1970	0.173	0.194	0.291
1980	0.076	0.089	0.128
Fraction of workers with $S \leq 12$			
1960	0.796	0.774	0.830
1970	0.721	0.764	0.790
1980	0.599	0.651	0.770
Mincerian returns to schooling			
1960	0.079	0.071	0.070
1970	0.090	0.075	0.072
1980	0.094	0.076	0.074
Cohort-specific wage growth			
	104.7%	98.7%	94.3%

Notes. This table reports the averages of state-specific moments. All moments are averaged across states.

Table 11: Complementarity and Externality

ρ	Elasticity of Substitution $\frac{1}{1-\rho}$	Normalized θ
1	∞	1
0.5	2	1.047
0	1	1.072
-0.5	0.667	1.098
-1	0.5	1.107

Notes. This table reports the effect of complementarity, measured by the elasticity of substitution between skilled and unskilled human capital, on estimates of human capital externalities θ. Each row reports an estimate of θ for a different value of ρ. The estimated θ for $\rho = 1$ is normalized to be one. See section 5 for details.