Estimating Aggregate Human Capital Externalities

Junjie Guo, Nicolas Roys and Ananth Seshadri*

October 2016

Abstract

Using measures of Compulsory Schooling Laws as instruments for state average schooling, we find that one more year of average schooling leads to a 6-8% increase in individual wages. The effect is statistically significant and robust to different specifications. We also find the effect to be larger for less-educated workers. The key difference from previous strategies is that, the exogenous variation in average schooling induced by our instruments comes mainly from workers not used directly in the wage regression. We construct a model of human capital accumulation where the average human capital of an economy affects the productivity of a typical firm in the economy. We estimate the elasticity of a firm's productivity with respect to the average human capital of the economy to be 0.121.

JEL: E23; J24; O47

*Guo: Department of Economics, University of Wisconsin-Madison, jguo27@wisc.edu; Roys: Federal Reserve Bank of St. Louis, nicolas.roys@stls.frb.org; 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 (CSLs) 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. Following Lucas (1988), we model human capital externalities as the effect of the average human capital of all workers on the productivity of a typical firm in an economy. Through its impact on productivity, average human capital also affects individual wages. Empirically, using average schooling as a measure for average human capital, human capital externalities are estimated as the impact of average schooling of all workers on the wage of a typical worker in an economy.

[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.

In the first part of the paper we use data on US-born white men aged 40-49 from the decennial censuses 1960-1980, the same data used in Acemoglu and Angrist (2001). Each state is taken to be a local economy, and human capital externalities are estimated as the impact of the change in state average schooling on the change in individual wages over time. We begin by replicating Acemoglu and Angrist (2001). For each individual used in the wage regression, their instruments for state average schooling are measures of CSLs affecting the schooling decision of that particular individual. They find no significant impact of average schooling on individual wages. Different from Acemoglu and Angrist (2001), our instruments for state average schooling faced by a worker are measures of CSLs affecting the schooling decisions of all workers aged 21-58 in the same state at the time of the census.[2] As an example, CSLs affecting schooling decisions of workers at age 30 in Wisconsin in 1960 are part of our instruments for average schooling in Wisconsin in 1960.[3] These CSLs, however, are not used in Acemoglu and Angrist (2001) because workers at age 30 are not included in the wage regression.

We believe our instruments are better for two reasons. First, with more instruments affecting a larger group of workers, our first-stage results are stronger and more robust to the inclusion of either state or region-specific trends. As shown in Angrist and Pischke (2014) and Stephens and Yang (2014), the impact of CSLs on individual schooling is greatly reduced with the inclusion of either state or region-specific trends. In terms of average schooling, we show that the F-statistics of the joint significance of the instruments in Acemoglu and Angrist (2001) are below 5 with either a state-specific trend as in Angrist and Pischke (2014) or a region-specific trend as in Stephens and Yang (2014). With our instruments, the same F-statistics are above 12 in all specifications. This is because our instruments affect the schooling decisions of a larger group of workers. In the above example, the schooling of workers at age 30 in Wisconsin in 1960, and in turn the average schooling of all workers in Wisconsin in 1960, is affected by our instruments but not those in Acemoglu and Angrist

[2] As explained later in the paper, the age restriction is due to data limitations.
[3] These are the CSLs around 1940 during the workers' schooling ages.

(2001). The additional variation in average schooling induced by our instruments leads to a stronger first stage. Second, our instruments are likely to be more "exogenous". While only workers aged 40-49 are used in the second-stage wage regression, our instruments are calculated from workers representing a much larger age range. Hence, most of the instruments have no direct impact on the workers used in the wage regression. In the above example, CSLs affecting schooling decisions of workers at age 30 in Wisconsin in 1960 should have no direct effect on wages of workers aged 40-49 in Wisconsin in 1960 other than the indirect effect through average schooling. In Acemoglu and Angrist (2001), by construction, the instruments for average schooling are correlated with the schooling decision of the worker in the wage regression. We believe it is problematic because absent CSLs, some individuals would not have attained their observed schooling level. If schooling is correlated with ability and part of the population was forced by CSLs to acquire the observed level of schooling, observationally identical individuals (who have the same schooling) will differ in ability, with workers facing stricter (higher) CSLs having a lower average ability. This could lead to a downward bias in estimates of human capital externalities.

Using this strategy, we find that state average schooling has a significant effect on individual wages. In particular, conditional on workers' own schooling and other observable characteristics, we estimate that one additional year of state average schooling increases individual wages by about 6-8%. The effect is larger for less-educated workers. And it is also larger when the exogenous variation in average schooling is due to older workers compared to the variation due to younger workers. The results are robust to the inclusion of state-specific trends as in Angrist and Pischke (2014) and to the inclusion of region-specific trends as in Stephens and Yang (2014). Our results are also robust to interstate migration.

In the second part of the paper, we propose a model of human capital acquisition to measure human capital externalities using the empirical estimates from the first part. Human capital externalities are formulated as the impact of average human capital on the productivity of a typical firm. Through its impact on firm productivity, average human capital also affects the price of human capital. Given the (expected) path of human capital prices, individuals make decisions on human capital accumulation both in school and at work to maximize their lifetime income. Individual decisions are subject to the CSLs set by the government. An increase in CSLs forces individuals at the margin to stay in school for a longer period of time and acquire more human capital, which, in turn, may affect the average human capital of the economy. With a human capital externality, the change in average human capital will affect the price of human capital and the optimal decisions and wages of workers not directly affected by the CSLs.

We estimate the model by treating each state as a small open economy. Each state is allowed to have its own paths of CSLs and TFP, where the former is taken from data and the latter is estimated to match the path of per capita income over time. We solve for transitional dynamics state by state to generate the joint distribution of schooling and earnings. Human capital production technologies in school and at work are estimated to match the schooling distribution and wages by schooling and age. To identify the human capital externality, we run regressions using simulated data and match the impact of average schooling on individual wages estimated in the first part of the paper. We find an elasticity of firm productivity with respect to average human capital to be about 0.121. We show using simulations that the model cannot match the empirical estimates of the effect of average schooling on individual wages in the absence of human capital externalities. Thus, the IV estimates from the first part of the paper help identify the important externality elasticity. 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.

Our baseline model assumes that human capital is homogenous 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 CSLs 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 CSLs raises the amount of low skilled human capital through more schooling. This, due to complementarity between the two types of human capital, raises the relative price of high skilled human capital. Low skilled workers at the margin will choose to become high skilled workers, and this selection effect reduces the average ability of low skilled workers and lead to a negative correlation between CSLs and individual wages conditional on schooling if there is no human capital externality. The positive effect of CSLs 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 model has other implications that are consistent with the available evidence. With significant human capital externalities, workers in rich countries accumulate more human capital over the life cycle than those in poor countries. This prediction is consistent with the findings in Lagakos et al. (2016, forthcoming).

The rest of the paper proceeds as follows. Section 2 introduces two measures of human capital externalities: one is reduced form and the other is structural. Section 3 presents empirical estimates of the impact of average schooling on individual wages. Section 4 introduces the model. Section 5 discusses the estimation strategy and presents the relevant results. Section 6 extends our baseline model to the case of heterogeneous human capital and demonstrates that our conclusion on human capital externalities holds with complementarity between different types of human capital. Section 7 concludes.

2 Measuring Human Capital Externalities

In an influential paper that sought to understand the mechanics of economic development and simultaneously account for cross-country income differences, Lucas (1988) proposes a simple way to capture external effects of human capital. The production function of a typical firm in an economy is given by

$$y = Ak^{\alpha}h^{1-\alpha}\bar{H}^{\theta} \tag{2.1}$$

where y is output, A is productivity, k and h are the quantities of physical and human capital rented by the firm, and $\bar{H}$ is the average human capital of the economy. θ is a measure of human capital externalities. The ultimate goal of this paper is to obtain an estimate of θ.

Theoretically, θ can be estimated from data on wages. Specifically, profit maximization by the firm implies

$$\begin{aligned} r &= \alpha A k^{\alpha-1}h^{1-\alpha}\bar{H}^{\theta} \\ w &= (1-\alpha)\,Ak^{\alpha}h^{-\alpha}\bar{H}^{\theta} \end{aligned}$$

where r and w are the rental rates of physical and human capital respectively. Assuming the rental rate of physical capital is given by the world interest rate r, we have

$$w = (1-\alpha)\left(\frac{\alpha}{r}\right)^{\frac{\alpha}{1-\alpha}} A^{\frac{1}{1-\alpha}}\overline{H}^{\frac{\theta}{1-\alpha}}$$

The log wage income $Y_i \equiv \log(wh_i)$ of a worker with h_i units of human capital is given by

$$Y_i = \beta_0 + \frac{1}{1-\alpha}\log A + \frac{\theta}{1-\alpha}\log \bar{H} + \log h_i$$

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. In the model we write down later, we are careful to distinguish between human capital and schooling. Empirically, we can estimate the following equation

$$Y_i = \gamma_0 + \gamma_1\bar{S} + \gamma_2 s_i + \epsilon_i \tag{2.2}$$

where s_i is the schooling of worker i, $\bar{S}$ is the average years of schooling of all workers in the economy, and ϵ_i is the error term that includes productivity A.

In the above formulation, γ_1 is a reduced form measure of human capital externalities. Intuitively, γ_1 should be related to the structural measure of human capital externalities θ. In what follows, we will first use instrumental variables (IV) to obtain a consistent estimate of γ_1, and then use a model to recover θ by forcing the model-generated data to match the empirical IV estimate of γ_1.

3 Empirical Analysis

We first present empirical estimates of the effect of state average schooling on individual wages in this section. Because our approach is based on the work of Acemoglu and Angrist (2001), we start by replicating their results and show that their empirical strategy is not very robust to the inclusion of state-specific or region-specific trends. We then present our empirical strategy and results. Contrary to Acemoglu and Angrist (2001), 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 Replicating Acemoglu and Angrist (2001)

Consider the following specification used in Acemoglu and Angrist (2001)

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

where Y_{ijt} is the log weekly wage of individual i in state j in census year t, u_{jt} is a state-year error component, and ϵ_i is an individual error term. The vector X_i includes state-of-birth (SOB) and year-of-birth (YOB) dummies, and δ_j and δ_t are state-of-residence (SOR) and census-year effects. The coefficient in front of individual schooling s_i, γ_2, will be referred to as the private return to schooling,[4] and the coefficient in front of average schooling $\overline{S}_{jt}$, γ_1, will be referred to as the external return to schooling. The goal is to obtain a consistent estimate of γ_1.

Equation (3.1) is essentially an extension of equation (2.2) derived from the production function (2.1), and it will be estimated using US decennial censuses 1960-1980,[5] which provide information on quarter of birth (QOB), SOB, SOR, highest grade completed and measures of earnings and labor supply for each respondent. Following Acemoglu and Angrist (2001), the schooling variable for an individual is defined to be the highest grade completed and capped at 17, and the weekly wage for an individual is calculated by dividing annual wage and salary income by weeks worked in the previous year. Average schooling for each state in each year is calculated to be the weighted average years of schooling of all US-born persons aged 16-64, where the weight is the product of the IPUMS weighting variable SLWT and weeks worked in the previous year. The main analysis in the paper is limited to US-born white men in their 40s with positive weekly wages reported in the censuses. All estimates are weighted by SLWT.[6]

As noted in Acemoglu and Angrist (2001), OLS estimates of γ_1 are likely to be biased because state average schooling $\overline{S}_{jt}$ is likely to be affected by other characteristics of a state and thus correlated with the error component u_{jt}.[7] To solve this problem, Acemoglu and Angrist (2001) propose an instrumental variables strategy where measures of Compulsory Schooling Laws (CSLs) are used as instruments for state average schooling. In particular, two measures of CSLs are constructed for each state in each year from 1914 to 1978, one for

[4]In Acemoglu and Angrist (2001), the private return to schooling is assumed to be random and takes the form $\gamma_{2i} = \gamma_2 \eta_i$. For simplicity, we assume γ_2 is constant across individuals.

[5]Acemoglu and Angrist (2001) also use 1950 and 1990 censuses in some cases. Because using these two datasets involves more data imputation, we ignore them for simplicity.

[6]For comparability, 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.

[7]For example, in equation (2.2), average schooling $\overline{S}$ will be correlated with the error term if it is correlated with productivity A.

compulsory attendance laws (CA) and the other for child labor laws (CL), as follows

$$CA = \max\{req\_sch, drop\_age - enroll\_age\}$$
$$CL = \max\{work\_sch, work\_age - enroll\_age\}$$

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.

With two measures of CSLs for each state in each year, we can assign these measures to each individual according to his SOB and YOB. In particular, the CSLs effective for an individual is taken to be the CSLs in effect in his SOB in the year when he was 14 years old. As the majority of individuals used in the regression have CL in the 6-9 range, while CA is concentrated in the 8-12 range with almost no one in the 11 category. Acemoglu and Angrist (2001) approximate the distributions of CL and CA using four dummies for each variable. For CL , the dummies are $CL6$ (for $CL \leq 6$), $CL7$ (for $CL = 7$), $CL8$ (for $CL = 8$) and $CL9$ (for $CL \geq 9$). For CA, the dummies are $CA8$ (for $CA \leq 8$), $CA9$ (for $CA = 9$), $CA10$ (for $CA = 10$) and $CA11$ (for $CA \geq 11$). These dummies are the instruments for state average schooling used in Acemoglu and Angrist (2001).

Table 1 presents our replication of the relevant results in Acemoglu and Angrist (2001). Estimates of γ_1 (external return to schooling) and γ_2 (private return to schooling) are reported for five different models of equation (3.1). Column 1 reports the OLS estimates, column 2 reports the two-stage least squares (2SLS) estimates where average schooling $\overline{S}_{jt}$ is instrumented by $CL7$, $CL8$ and $CL9$, column 3 reports the 2SLS estimates where $\overline{S}_{jt}$ is instrumented by $CA9$, $CA10$ and $CA11$, column 4 reports the 2SLS estimates where individual schooling s_i and average schooling $\overline{S}_{jt}$ are instrumented by QOB dummies and $CL7$, $CL8$ and $CL9$. Finally, column 5 reports the 2SLS estimates where s_i and $\overline{S}_{jt}$ are instrumented by QOB dummies and $CA9$, $CA10$ and $CA11$. The corresponding estimates in Acemoglu and Angrist (2001) are reported in column 1 (panel b) of Table 2 and columns 5, 6, 1 and 2 of Table 7.

We are able to replicate the coefficients in Acemoglu and Angrist (2001) almost exactly.[8] The standard errors in Table 1 are larger than the corresponding values in Acemoglu and Angrist (2001) because, instead of using the formula in Moulton (1986) to correct for state-year clustering, we account for both intragroup correlation of the error term and heteroskedasticity by clustering all standard errors to state-year level throughout this paper. By using the larger standard errors we are more conservative and make sure that the significant effects uncovered below are not due to specific choices of smaller standard errors.

From Table 1 we can see that, although the OLS estimate of the external return to schooling is about 7%, statistically significant and almost as large as the private return to schooling, the IV estimates of external returns to schooling are much smaller and not significant in any of the specifications. Based on these results, Acemoglu and Angrist (2001) conclude that there is no evidence in support of sizable human capital externalities.

3.2 Validity of Empirical Strategy in Acemoglu and Angrist (2001)

There are two potential threats to the empirical strategy in Acemoglu and Angrist (2001), and we discuss these threats in turn.

3.2.1 Parallel Trends

As pointed out in Angrist and Pischke (2014),[9] "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

[8] The only discrepancy is the OLS estimate of the private return to schooling. We get 0.072 while it is 0.073 in Acemoglu and Angrist (2001).

[9] Pages 226-227.

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". Following this argument, Angrist and Pischke (2014) report the estimates of the effect of CL on individual years of schooling using models with and without state-specific trends and show that, once a state-specific trend is included in the model, CL no longer matters for individual schooling in their sample. Based on this, Angrist and Pischke (2014) conclude that the strategy in Acemoglu and Angrist (2001) is "a failed research design".

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 the "first stage estimates of the impact of schooling laws on educational attainment are weakened by allowing regional differences in year of birth effects", and 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.

The variable of interest in Angrist and Pischke (2014) and Stephens and Yang (2014) is individual schooling s_i, while in this paper it is average schooling $\overline{S}_{jt}$. We examine the impact of CSLs on average schooling under different models of time trends. Specifically, we estimate three variants of equation (3.2), where CSL is a vector of CSL measures. In the first case, no control for time trends is included, as in Acemoglu and Angrist (2001). In the second case, a separate linear YOB trend for each SOB is included. Specifically, following Angrist and Pischke (2014), we assume $Trend = \varphi_{SOB_i} YOB_i$. In the third case, we follow Stephens and Yang (2014) by using the interactions between region dummies and YOB dummies as a way to control for different time trends across regions. That is, we assume $Trend = \delta_{region(SOB_i),YOB_i}$, where $region\,(SOB_i)$ is the region to which SOB_i belongs.

$$\overline{S}_{jt} = X_i'f + \delta_j + \delta_t + \kappa s_i + CSL'\Gamma + Trend + \varepsilon_{jt} \tag{3.2}$$

Table 2 reports the results for both CL and CA measures. Following Acemoglu and Angrist (2001), for CL, we set $CSL = (CL7, CL8, CL9)$, and we set $CSL = (CA9, CA10, CA11)$ in the case of CA. It is clear that once a trend term is included, the impact of CSLs on average schooling is reduced significantly. For instance, the coefficient on $CL7$ is reduced from 0.084 to 0.014 once a state-specific trend is included, and the coefficient on $CA9$ is reduced from 0.128 to 0.029 once a region-specific trend is included. The last but one line named F-CSL reports the F-statistics of the joint significance of the relevant CSL measures. While the F-statistic is above 8 for both CL and CA in the absence of a trend term, it is below 5 once a trend term is included. While one can debate the best way to model time trends, the small F-statistics for both CSL measures under both models of trends indicate that the empirical strategy in Acemoglu and Angrist (2001) may be contaminated and the resulting IV estimates of external returns to schooling may not be conclusive. Additional estimates using different strategies may be informative.

3.2.2 Ability Bias

Another potential issue with the strategy in Acemoglu and Angrist (2001) is that it may introduce a downward bias to the estimate of γ_1. To see this, assume individual schooling s_i is determined by ability z_i and CSLs CSL_i. We expect individuals with higher ability to choose more schooling and, other things equal, individuals in states with tighter CSLs should acquire more schooling. That is, $\frac{\partial s_i}{\partial z_i} \geq 0$ and $\frac{\partial s_i}{\partial CSL_i} \geq 0$. Consider two individuals with the same amount of schooling $s_1 = s_2$ but born and educated in two states with different CSLs $CSL_1 > CSL_2$. For these two individuals, we expect $z_1 \leq z_2$. That is, for individuals with the same years of schooling, those born and educated in a state with tighter CSLs should on average have lower ability because some of them were forced to acquire the observed years of schooling while some others did so by choice. This introduces a negative selection bias which implies that, in the absence of human capital externalities, individuals from a state with tighter CSLs would earn less than those with the same amount of schooling but from a state in which the CSLs are not as tight. That is, if $\gamma_1 = 0$, the strategy in Acemoglu and Angrist (2001) may generate a negative estimate of γ_1 if the average schooling is higher in the state with tighter CSLs.[10] If this is the case, the insignificant estimates of γ_1 in Acemoglu and Angrist (2001) are not conclusive evidence against a positive impact of average schooling on individual wages. A zero reduced form IV estimate of external returns to schooling in Acemoglu and Angrist (2001) could be perfectly consistent with a positive structural estimate of human capital externalities.

3.3 Empirical Strategy

We present our empirical strategy in this subsection. Before moving on to the details of our strategy, we first introduce two alternative measures of CSLs. We will show that our strategy and results are not affected by particular choices of CSL measures.

As noted by Goldin and Katz (2008), the first term in the max function for CA introduced earlier is an exception which allows the child to leave school before the dropout age, the correct calculation of CA would use a min function. Following this, we create a corrected measure of CA, CCA, defined as

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

[10] Using QOB as an instrument for individual schooling does not solve this problem. The reason is that the population whose schooling could potentially be affected by QOB depends on CSLs that vary across states. For example, assume the required years of schooling in state 1 is 6, and it's 8 in state 2. The population of workers whose schooling decisions are affected by QOB in state 1 will be those considering whether to finish the 6th grade. While in state 2, it will be those considering whether to finish the 8th grade. In particular, QOB has no effect on whether to finish the 6th grade in state 2. As a result, the population of workers with 6 years of schooling in state 1 will be different from the population of workers with 6 years of schooling in state 2. This leads to a selection bias that affects the estimates of external returns to schooling.

and CCA is used instead of CA hereafter.

Our last measure of CSLs 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.[11]

With all measures of CSLs introduced, we are now ready to present our empirical strategy. Below we use CL as an example to show how our instruments are constructed. Instruments based on CCA and RS are constructed in the same way. Given the individual level CL that potentially affects a worker, we can calculate $\overline{CL}_{jt}^{a}$, the average CL among all workers at age a in state j in census year t. Naturally, $\overline{CL}_{jt}^{a}$ is a measure of the strength of CSLs affecting schooling decisions of all workers at age a in state j in census year t, and it should potentially be correlated with $\overline{S}_{jt}^{a}$, the average years of schooling for this group of workers. Because $\overline{S}_{jt}^{a}$ is part of $\overline{S}_{jt}$, $\overline{CL}_{jt}^{a}$ is a potential determinant of $\overline{S}_{jt}$. Data availability allows us to calculate $\overline{CL}_{jt}^{a}$ for a between 21 and 58,[12] and $\{\overline{CL}_{jt}^{a}\}_{a=21}^{58}$ will be our instruments for average schooling $\overline{S}_{jt}$.

At first glance, it may seem strange to average individual level CL to state level $\overline{CL}_{jt}^{a}$

[11]Interested readers can find more details in Stephens and Yang (2014). In practice, we download the measure constructed by the authors directly from `http://www.aeaweb.org/articles.php?doi=10.1257/aer.104.6.1777`. The measure is available for all individuals born in mainland US between 1905 and 1961.

[12]$\overline{CCA}_{jt}^{a}$is also available for a between 21 and 58. $\overline{RS}_{jt}^{a}$, however, is only available for a between 21 and 55.

because some variation will be lost when averaging. However, the benefit of doing so is that, now we can use as instruments the CSLs for all workers, including not only those in their 40s who are used in the wage regression but also individuals in their 20s, 30s or 50s who are used in calculating average schooling $\overline{S}_{jt}$ but not included directly in the wage regression. The inclusion of CSLs for workers not used directly in the wage regression brings at least two benefits: (1) it could potentially introduce more variation in average schooling as long as some workers not used directly in the wage regression were affected by CSLs, and this will lead to a stronger first-stage estimate of the impact of CSLs on average schooling. For example, as workers at age 30 are not used directly in the wage regression, CSLs effective for them are not utilized in Acemoglu and Angrist (2001). By including $\overline{CL}_{jt}^{30}$ as an instrument, we can exploit the variation in average schooling $\overline{S}_{jt}$ induced by $\overline{CL}_{jt}^{30}$. (2) This additional variation may be more exogenous because it comes from other workers and has little direct impact on individuals used in the wage regression other than through average schooling. For example, $\overline{CL}_{jt}^{30}$ should have little direct effect on wages of workers in their 40s used in the wage regression, and the variation in average schooling $\overline{S}_{jt}$ induced by $\overline{CL}_{jt}^{30}$ is unlikely to be correlated with the error terms in the wage equation (3.1).

To be clear, for each individual used in the wage regression, the instruments for average schooling $\overline{S}_{jt}$ in Acemoglu and Angrist (2001) are measures of CSLs that could have affected the schooling decision of that individual (own CSLs), while our instruments consist of measures of CSLs that could have affected the schooling decisions of all workers, including women and men in their 20s, 30s and 50s who are not included directly in the wage regression, in the same state in the same census year. As long as schooling decisions of some workers in the same state in the same census are affected by CSLs, our instruments will be correlated with average schooling $\overline{S}_{jt}$.

To investigate the impact of our instruments on average schooling, we estimate equation (3.2) by setting the term $CSL'\Gamma$ to be $\sum_{a=21}^{58} \overline{CL}_{jt}^{a}\Gamma_a$, $\sum_{a=21}^{58} \overline{CCA}_{jt}^{a}\Gamma_a$ and $\sum_{a=21}^{55} \overline{RS}_{jt}^{a}\Gamma_a$, respectively. Tables 3 and 4 report estimates of Γ_a from equation (3.2) with state-specific and region-specific time trends, respectively. In both tables, we can see that not all estimates of Γ_a are positive. This is likely due to the correlation of the CSL measures over time. For example, if there is no migration, we should have $\overline{CL}^{a}_{j1960} = \overline{CL}^{a+10}_{j1970} = \overline{CL}^{a+20}_{j1980}$ for all j and a because the three measures are calculated from the same group of workers. As a change in one CSL measure (for example, $\overline{CL}^{30}_{j1960}$) is associated with a change in other CSL measures (for example, $\overline{CL}^{40}_{j1970}$ and $\overline{CL}^{50}_{j1980}$), we can no longer interpret the estimates of Γ_a as the marginal effect of $\overline{CL}^{a}_{jt}$ (or $\overline{CCA}^{a}_{jt}$ and $\overline{RS}^{a}_{jt}$) holding other variables constant.

To circumvent the complication in interpreting the estimates of Γ_a at each age, we focus on the F-statistics of the joint significance of the relevant CSL measures at all ages reported in the last row of tables 3 and 4. All 6 F-Statistics are above 12, and four of them are above 20. The large F-statistics indicate that our instruments are significantly related to average schooling $\overline{S}_{jt}$ for all three CSL measures under both models of trends. As long as CSLs are exogenous to the error terms in the wage equation (3.1) conditional on other controls including the trend term, the instruments will be valid and can be used to estimate external returns to schooling.

3.3.1 Alternative Specifications

As argued above in section 3.2.2, using own CSLs as an instrument for average schooling may introduce a downward bias to the estimated impact of average schooling on individual wages. Because our instruments are averaged to state-year-age level, the contribution of own CSLs to the instruments is very small, and almost all of the variation in average schooling induced by our instruments come from individuals other than the particular worker in the wage regression.

To further reduce the contribution of own CSLs to our instruments and use only the variation in average schooling due to other workers to identify external returns to schooling, we implement two alternative specifications. In the first case, we include own CSLs as an additional control variable into equation (3.1). With this additional control, the identification of γ_1 comes from workers affected by the same CSLs, and they face different average schooling purely due to different schooling choices of other workers in the same state in the same census. In the second case, we use $\{\overline{CSL}_{jt}^a\}_{a=21}^{39}$ and $\{\overline{CSL}_{jt}^a\}_{a=51}^{58}$ as instruments instead of $\{\overline{CSL}_{jt}^a\}_{a=21}^{58}$, where $\overline{CSL}_{jt}^a$ is either $\overline{CL}_{jt}^a$, $\overline{CCA}_{jt}^a$ or $\overline{RS}_{jt}^a$. Because the CSLs effective for workers in their 40s who are used in the wage regression are excluded, the variation in average schooling is again purely due to other workers.

As the effect of own CSLs on average schooling is either controlled for directly or removed completely, we expect the estimates of external returns to schooling from these two alternative specifications to be larger than our baseline specification, but the difference should be very small as the contribution of own CSLs to our baseline instruments is minimal. Results presented below are consistent with this prediction.

3.4 Estimates of External Returns to Schooling

Given the validity of our instruments, we report in this subsection the 2SLS estimates of external returns to schooling.[13] For brevity, we only report estimates from the model where equation (3.1) is 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.

3.4.1 Baseline Estimates

Table 5 reports the baseline results. The first column of Panel A reports OLS estimates. The estimated external return is about 7% and statistically significant as in our replication of Acemoglu and Angrist (2001) in the first column of Table 1. Controlling for trends reduces the estimated external return only slightly from 7.3% to 6.9%. The rest of the columns in panel A report 2SLS estimates where average schooling is instrumented by each of the three sets of instruments. Although estimates of external returns vary with the particular

[13]The estimates from using limited information maximum likelihood (LIML), which are available from the authors, are very similar to the 2SLS estimates shown below.

instruments used, they fall in a narrow range of 6-8% and are comparable to the OLS estimate, and all estimates are statistically significant. Finally, in panel B, we treat both individual schooling and average schooling as endogenous and instrument them with QOB dummies and relevant CSL measures. 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 estimates of the private return to schooling. For instance, the estimated external returns when CL measures are used as instruments is 6.9% when individual schooling is treated as exogenous, and it increases to 7.3% when individual schooling is treated as endogenous. The estimated external returns when RS measures are used as instruments is 6.1% no matter whether individual schooling is taken as exogenous or endogenous. Given this and the fact that F-statistics of joint significance of QOB dummies are only around 5,[14] in what follows we only report results treating individual schooling as exogenous. However, it is important to note that all results reported below are similar when individual schooling is treated as endogenous and QOB dummies are used as additional instruments.

Table 6 presents estimates from two alternative specifications discussed in section 3.3.1 to further reduce the negative selection bias induced by own CSLs. Panel A reports estimates of equation (3.1) augmented to include own CSLs as an additional control variable, and panel B reports estimates where $\{\overline{CSL}_{jt}^{a}\}_{a=40}^{50}$ are excluded from the set of instruments. Consistent with the idea that own CSLs introduce a negative bias, the estimated external returns in Table 6 are larger than our baseline estimates reported in panel A of Table 5. The differences between them, however, are very small. For instance, when average schooling is instrumented with CL measures, the estimated external returns is 6.9% in our baseline specification, and it increases to 7.2% and 7.7% in the two alternative specifications. Overall, controlling for own CSLs leads to an increase in estimates by about 0 to 0.8 percentage point

[14]This is the partial F-statistic of the joint significance of QOB dummies in a regression of individual schooling on QOB dummies, relevant CSL measures and all explanatory variables in the main wage regression other than individual schooling and average schooling.

or 0-11.6%, which is small and consistent with the fact that the contribution of own CSLs to our instruments used in the baseline specification is minimal. We proceed with our baseline specification.

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 SOB CSLs, invalidating SOB CSLs as an instrument for average schooling. SOR CSLs, defined as the CSLs effective for a worker in his SOR in the year when he was 14 years old, 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. To do so, we 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. Using these average SOR CSLs as instruments for average schooling, we redo our analysis and the results are reported in Table 7.

Clearly, estimates of external returns to schooling in Table 7 are very close to corresponding estimates in Table 5, 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.

3.4.2 Estimates from Different CSL Cohorts

As mentioned earlier, average schooling of a state is defined as the weighted average years of schooling of all workers aged 16-64 in that state. If average schooling in state 1 is larger than average schooling in state 2, it could either be that young workers in state 1 have more schooling on average than young workers in state 2, or older workers in state 1 have on average more schooling than older workers in state 2, or both. The effect of average schooling due to young workers may be different from the effect of average schooling due to older workers. Our approach allows us to explore this possibility. Instead of using average CSLs of all ages $\{\overline{CSL}^a_{jt}\}_{a=21}^{58}$ as instruments for average schooling at the same time, we break the instruments into two groups. The first group $\{\overline{CSL}^a_{jt}\}_{a=21}^{39}$ includes CSLs affecting schooling decisions of workers between 21 and 39 years old, and the second group $\{\overline{CSL}^a_{jt}\}_{a=40}^{58}$ includes CSLs affecting schooling decisions of workers at age 40 and beyond. When $\{\overline{CSL}^a_{jt}\}_{a=21}^{39}$ are used as instruments for average schooling, the variation is more likely to come from young workers, while the variation in average schooling induced by $\{\overline{CSL}^a_{jt}\}_{a=40}^{58}$ is more likely to result from older workers. We use these two groups of CSLs separately as instruments for average schooling to investigate whether the effect of average schooling due to older workers is different from the effect from young workers.

Table 8 reports estimates of external returns to schooling from this exercise. The top panel reports the estimates where average schooling is instrumented by $\{\overline{CSL}^a_{jt}\}_{a=40}^{58}$, and the bottom panel reports the estimates where the instruments are $\{\overline{CSL}^a_{jt}\}_{a=21}^{39}$. With fewer measures of CSLs, the instruments are not very significant in the first-stage regressions. This can be seen from the relatively small F-statistics of joint significance of CSL measures, mostly of which are 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 $\{\overline{CL}^a_{jt}\}_{a=40}^{58}$, 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 $\{\overline{CL}^a_{jt}\}_{a=21}^{39}$. Larger estimated effects from older workers are also observed with the other two CSL measures. Being able to use the variation in average schooling induced by CSLs effective for a much larger group of individuals is the main reason why our instruments are more significantly related to average schooling than those in Acemoglu and Angrist (2001), and the large estimates of external returns uncovered in this paper seem to come from the variation in average schooling due to older cohorts.

One potential explanation for the larger impact induced by older cohorts is human capital accumulation. According to the production function in equation (2.1), an increase in average human capital $\bar{H}$ due to more schooling acquired by workers raises the price of human capital if there is a positive human capital externality $\theta > 0$. If young workers anticipate a higher future price of human capital due to more schooling from older cohorts, they will take advantage of this by accumulating more human capital while young and earning a higher wage later on because of not only the higher price but also a larger stock of human capital. On the other hand, if an old worker realizes that the price of human capital is higher because of more schooling from younger workers, faced with a shorter horizon, he is not going to respond much in terms of human capital accumulation, and the increase in wage will be relatively smaller.

3.4.3 Estimates by Years of Schooling

Until now we have been assuming that external returns to schooling are homogeneous across the population. Next we divide individuals in our sample into three groups according to their completed years of schooling s_i, and estimate a separate external return for each group using equation (3.1). Workers used in the wage regression (US-born white men in their 40s with positive earnings reported) are grouped into 3 bins based on years of schooling: 0-8, 9-11 and 12+. Table 9 reports estimated external returns to schooling for each group under different models.

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), CCA 0.035 (0.025) and RS 0.009 (0.027). One potential explanation is diminishing returns to human capital accumulation. Because better-educated workers have a larger stock of human capital, the return to additional human capital accumulation is smaller. When the price of human capital increases due to a higher average schooling, all workers will have an incentive to accumulate more human capital, but the increase in human capital in percentage terms will be smaller for better-educated workers.

3.4.4 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 10 reports the results. Panel A reports the baseline estimates where the specification is the same as the one used for Panel A of Table 5. 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 8. Finally, Panel C reports the estimated external returns by years of schooling, where the specifications are the same as those in Table 9.

We can see from panel A that both the OLS and the 2SLS 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 8, the 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 age cohorts are smaller in panel B than the differences in Table 8.[15] 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 statistical significant. While in Table 8, only three of them are significant. However, as in Table 8, the F-statistics in panel B are

[15] 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 8 is 0.083 (0.091-0.008).

relatively small, suggesting that the results should be interpreted with caution.

Finally, from panel C we can see that the estimated external returns are larger for less-educated workers, consistent with previous estimates from workers in their 40s reported in Table 9. However, different from Table 9, all but one estimate in panel C are statistically significant at conventional significance levels.

3.5 Discussion

In summary, in contrast to Acemoglu and Angrist (2001), we find a large and statistically significant impact of average schooling on individual wages. The impact is larger for less-educated workers. The evidence suggests that the impact is larger when the variation in average schooling is due to older workers. We also find a large impact of average schooling on younger workers.

The positive estimates for workers with 9-12 years of schooling are evidence against the concern that the estimated external returns arise from supply-driven movements along a downward sloping relative demand curve for less educated workers (Ciccone and Peri (2006)). As shown in Acemoglu and Angrist (2001), CSLs primarily shift the distribution of schooling in middle- and high-school grades, and have no effect on the proportion of the population attending college.[16] In other words, CSLs reduce the fraction of workers with 0-8 years of schooling and raise the fraction of workers with 9-12 years of schooling without affecting the fraction of workers with more than 12 years of schooling. Without external returns to schooling, an increase in the fraction of workers with 9-12 years of schooling would reduce their wages in the case of a downward sloping demand curve. This is not what we observe in the data.

[16] 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 CCA are similar.

4 Model

We present in this section a model of schooling and human capital accumulation. The model is based on the classic work of Ben-Porath (1967), where an individual makes decisions on schooling and on-the-job training to maximize the present discounted value of lifetime income. We introduce CSLs into the baseline model of Ben-Porath (1967). Our model is able to account for the effects of average schooling on individual earnings estimated in the previous section. More importantly, by relating the reduced form parameter γ_1 in equation (3.1) to the structural parameter θ in the production function (2.1), we are able to recover a direct measure of human capital externalities.

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.

4.1 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. (z, h_0) is assumed to be independently and identically distributed (i.i.d) both across workers in a given generation and across generations.

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. Time endowment is normalized to be one in each period. Section 5 discusses the mapping between schooling $n_0 \in [0, 1]$ determined at age 0 in the model and years of schooling observed in the data. Each individual is characterized by (z, h_0, t_b) and solves the dynamic programming problem:

$$V\left(z, h_0, t_b\right) = \max_{n_a \in [0,1], x_0 \geq 0} \left\{ \left[w_{t_b} h_0 \left(1 - n_0\right) - x_0\right] + \sum_{a=1}^{R} \frac{w_{t_b+a} h_a \left(1 - n_a\right)}{\left(1 + r\right)^a} \right\} \tag{4.1}$$

subject to

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

where w_t is the 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 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.

When making schooling decisions n_0, an individual faces constraints from Compulsory Schooling Laws (CSLs) set by the government requiring the schooling time for all individuals born at time t_b to reach or exceed some threshold $\overline{CSL}_{t_b}$. As a result, individual decisions must satisfy $n_0 \geq \overline{CSL}_{t_b}$.

4.2 Firms

Firms produce the final goods using human capital h and physical capital k. The production function is given by equation (2.1) and repeated below for completeness

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

where y is output, A is productivity, and t is a time index. Human capital externalities are captured by θ, which is the key parameter of interest.

Markets are competitive. Taking interest rate r, wage rate w_t and productivity A_t as given, firms maximize profits

$$\pi_t = \max_{k_t, h_t} \left\{ A_t k_t^{\alpha} h_t^{1-\alpha} \bar{H}_t^{\theta} - r k_t - w_t h_t \right\} \tag{4.2}$$

by choosing k_t and h_t, where the price of final goods is normalized to be 1. From this we can get

$$w_t = (1-\alpha) \left(\frac{\alpha}{r} \right)^{\frac{\alpha}{1-\alpha}} A_t^{\frac{1}{1-\alpha}} \bar{H}_t^{\frac{\theta}{1-\alpha}} \tag{4.3}$$

4.3 Government

The government has only one role, making decisions in each period t on CSLs summarized by $\overline{CSL}_t$. For our purpose, no particular assumption on the objective of the government is necessary because we are not trying to work out the optimal choice of $\overline{CSL}_t$. The only assumption is that choices of $\overline{CSL}_t$ are exogenous to both workers and firms. Workers' responses to exogenous variations in $\overline{CSL}_t$ are critical for the identification of human capital externalities θ.

4.4 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 (4.1).

2. For each firm, k_t^* and h_t^* are optimal solutions to the problem given by equation (4.2).

3. The demand for human capital equals the supply of human capital.

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

$$\bar{H}_t = \frac{\sum_{a=0}^{R} \int h_a^* (1 - n_a^*) \, dF(z, h_0)}{\sum_{a=0}^{R} \int (1 - n_a^*) \, dF(z, h_0)} \tag{4.4}$$

where $F(z, h_0)$ is the cumulative distribution function of (z, h_0). Condition 3 requires workers and firms to have perfect foresight over the path on $\bar{H}_t$.

4.5 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 leads to an increase in both average schooling and average human capital of the economy $\bar{H}$ when these individuals enter the labor force. With a positive human capital externality $\theta > 0$, a larger average human capital raises the rental rate w of human capital through equation (4.3). The higher rental rate w has a direct effect on earnings of all workers, including those whose schooling decisions are not affected by the increase in $\overline{CSL}_t$. This explains the positive correlation between average schooling and individual wages estimated in the previous section. In the absence of human capital externalities $\theta = 0$, the increase in average schooling and average human capital due to a tighter $\overline{CSL}_t$ has no effect on the rental rate w of human capital and earnings of those not affected directly by the increase in $\overline{CSL}_t$, in which case we should get a zero estimate of γ_1.

With a positive human capital externality $\theta > 0$ where we get a positive effect of $\overline{CSL}_t$ on the rental rate w of human capital, there is also an indirect effect on earnings that works through human capital accumulation. Specifically, in anticipation of the higher rental rate w when the cohort born at time t enters the labor force, other cohorts will adjust their choices, x and n, to take advantage of this. The response will be larger for cohorts born after time t than those born before time t, because the latter cohort has a shorter time horizon to reap the benefit of increased investments. This implies a larger impact of average schooling on individual wages when the increase in average schooling arises from older cohorts. The magnitude of this indirect effect also varies across individuals within the same cohort depending on their ability z and stock of human capital h_a. Individuals with a larger z will benefit more from this indirect effect by investing more resources in human capital production. In the case of diminishing returns, $\alpha_3 < 1$, workers who have already accumulated a large stock of human capital h_a will not adjust their investments as much, resulting in a smaller benefit from this indirect effect.

5 Estimation

This section proposes the strategy used to estimate our model using indirect inference. Indirect inference works by selection of a set of statistics of interest which the model is asked to reproduce.[17] These statistics $\hat{\Psi}$ come primarily from the first part of the paper and are listed 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 derived by searching over the parameter space to find the parameter vector that minimizes the criterion function,

$$\hat{\beta} = \arg\min_{\beta} \left(\hat{\Psi} - \Psi(\beta)\right)' W \left(\hat{\Psi} - \Psi(\beta)\right) \tag{5.1}$$

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

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

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

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 that is 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.

5.1 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 US 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, CA and CCA) only account for the 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 $t \geq 7$.[18]

5.2 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

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

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 11 lists the set of parameters to be estimated.

5.3 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

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

1. OLS and 2SLS estimates of external returns to schooling for all workers. From Table 5, the OLS and 2SLS (with RS) estimates are 0.069 and 0.061 respectively.

2. OLS and 2SLS 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 9.

3. OLS and 2SLS 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 2SLS 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.

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.[19]

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 the average wage of workers between 42 and 51 years old in

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

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, 2SLS 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, 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.[20]

5.4 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 (5.1) will be our estimated parameters. Appendix A provides more details

[20] 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.

on model computation and estimation.

Table 11 reports the estimated parameters, and Tables 12 and 13 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 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 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 2SLS 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 2SLS estimates of external returns by schooling almost exactly. The OLS and 2SLS 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 simulated data to calculate moments reported in the last column of tables 12 and 13. 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 2SLS 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 2SLS estimates of external returns to schooling γ_1 and a positive θ is required to generate the positive estimates of γ_1.

In summary, estimation of the model results in 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.

6 Heterogeneous Human Capital

Our baseline model assumes that human capital is homogenous 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 (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 9-12 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 let $V(z; w_L)$ 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 $h_a(z)$ be the stock of human capital for an individual with ability z at age a, and let $n_a(z)$ be the amount of training time for this individual at age a. The total amount of human capital used in the production of the final good will be given by

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

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,[21] the direct effect of CSLs 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 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 with 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 14. For each value of ρ, we calibrate the model to obtain an estimate of θ that matches the empirical estimate of γ_1.[22] We normalizes the estimate of θ when $\rho = 1$ to be one. From table 14 we can see 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

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

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

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 14, 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 14 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 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.

7 Conclusion

We estimate two measures of human capital externalities in this paper. Empirically, we build on the work of Acemoglu and Angrist (2001) and estimate the impact of state average schooling on individual wages. Even though we also use measures of CSLs as instruments for average schooling, our empirical strategy is different from Acemoglu and Angrist (2001) in two ways. First, we use mainly the CSLs affecting workers not used in the wage regression as instruments and argue that the variation in average schooling induced by our instruments are more exogenous. Second, following the work of Angrist and Pischke (2014) and Stephens and Yang (2014), both of whom argue for the importance of controlling for the different trends across either states or regions, we include a trend term into the empirical specification and show that the instruments in Acemoglu and Angrist (2001) become insignificant with the inclusion of the trend term. Our instruments, however, are robust to different models of the trends. We find that one more year of average schooling leads to a 6-8% increase in individual wages. The effect is statistically significant and robust to different specifications. The effect is also found to be larger for less-educated workers and it is larger when the exogenous variation in average schooling comes from older workers.

Having established that the Instrumental Variable estimates of aggregate human capital externalities are sizable, we proceed to write down a model of human capital accumulation. The key ingredients of the model are a Ben-Porath style human capital production technology which is the workhorse of modern labor economics and an aggregate spillover that affects the productivity of the representative firm. We argue that the IV estimate we uncover in the first part of the paper can help identify the structural externality parameter. Our second measure of human capital externalities concerns the impact of average human capital in an economy on the productivity of a typical firm in the economy. To estimate this measure, we construct an overlapping generations model of human capital accumulation both in school and at work. Despite the model's simplicity, computing the transitional dynamics of the model with the changes in CSLs is a fairly involved exercise. The model is estimated using state level data in the 20th century and in particular, the causal effect of CSLs on individual wages identified in the first part of the paper. Estimation of the model reveals that a 1% increase in average human capital raises a firm's productivity by around 0.121%.

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 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.

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." Tech. rep.

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.

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. 2016. "Life-Cycle Human Capital Accumulation Across Countries: Lessons From U.S. Immigrants." NBER Working Papers 21914, National Bureau of Economic Research, Inc.

———. forthcoming. "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.

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.

Moulton, Brent R. 1986. "Random group effects and the precision of regression estimates." *Journal of Econometrics* 32 (3):385–397.

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.

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 rental rates $\{w_t\}$ using equation (4.3).

   (b) With $\{w_t\}$, we can solve the problem in equation (4.1) 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 (4.4).

   (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 14.

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:[23] 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 CSLs to a value RS_1 measuring the average required years of schooling for respon-

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

dents in 1960 census.

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

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 CSLs to a different value $RS_2 \neq RS_1$.

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

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 of the 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: Replication of Acemoglu and Angrist (2001)

	OLS	CL	CA	QOB+CL	QOB+CA
Private Return	0.072***	0.073***	0.073***	0.074***	0.074***
to Schooling	(0.001)	(0.001)	(0.001)	(0.016)	(0.016)
External Return	0.073***	0.002	0.018	0.003	0.017
to Schooling	(0.016)	(0.073)	(0.067)	(0.078)	(0.067)

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 in all columns. 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 and year of census. The first column reports the OLS estimates. The second and third column report the estimates where average schooling is instrumented by child labor laws CL and compulsory attendance laws CA respectively. The last two columns further instruments individual schooling with the worker's quarter of birth QOB. All standard errors reported in the parentheses are clustered to the SOR-by-census level. *** denotes 1%, ** denotes 5%, and * denotes 10% significance.

Table 2: Validity of Instruments in Acemoglu and Angrist (2001)

	CL			CA		
CL7	0.084**	0.014	0.030			
	(0.036)	(0.014)	(0.024)			
CL8	0.107***	0.027**	0.020			
	(0.029)	(0.012)	(0.018)			
CL9	0.226***	0.030**	0.090***			
	(0.050)	(0.015)	(0.030)			
CA9				0.128***	0.012	0.029
				(0.028)	(0.011)	(0.017)
CA10				0.122***	0.004	0.088***
				(0.039)	(0.014)	(0.025)
CA11				0.143***	0.021	0.020
				(0.031)	(0.019)	(0.019)
F-CSL	8.335	1.987	4.909	8.265	0.649	4.164
Trend	No	State	Region	No	State	Region

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 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. CL7, CL8 and CL9 are dummies indicating whether the worker was required by the child labor laws CL to stay in school for at least 7, 8 and 9 years, respectively. CA9, CA10 and CA11 are similar measures of compulsory attendance laws CA. F-CSL reports the F-statistic of the joint significance of compulsory schooling laws CSLs (CL7, CL8 and CL9 in the first three columns and CA9, CA10 and CA11 in the last three columns). Other variables included in all columns are individual schooling and dummies for state of birth, state of residence (SOR), year of birth and year of census. The first and fourth columns report the estimates with neither state nor region-specific trend. The second and fifth columns report the estimates with a linear year of birth trend for each state of birth as in Angrist and Pischke (2014). The third and sixth columns report the estimates with interactions between year of birth dummies and dummies for the four census regions as in Stephens and Yang (2014). 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: First-Stage Estimates: State-Specific Time Trends

	CL		CCA		RS	
Age	coef.	s.e.	coef.	s.e.	coef.	s.e.
21	0.048	0.142	0.086	0.044	-0.213	0.056
22	0.090	0.127	-0.000	0.041	0.476	0.102
23	-0.496	0.123	-0.063	0.041	-0.205	0.120
24	0.065	0.105	0.010	0.040	0.247	0.116
25	0.268	0.072	-0.062	0.030	-0.441	0.124
26	-0.490	0.116	-0.284	0.048	0.390	0.096
27	-0.061	0.118	0.057	0.041	-0.282	0.093
28	0.729	0.137	0.213	0.043	-0.309	0.079
29	-0.125	0.092	0.097	0.030	0.168	0.102
30	-0.104	0.062	-0.059	0.018	0.226	0.092
31	-0.043	0.112	0.051	0.052	-0.075	0.082
32	-0.126	0.119	-0.125	0.062	-0.114	0.128
33	-0.114	0.118	0.069	0.051	0.166	0.083
34	0.221	0.173	0.045	0.048	-0.014	0.106
35	0.147	0.060	-0.119	0.025	0.297	0.090
36	0.094	0.089	-0.159	0.057	-0.280	0.097
37	-0.311	0.165	0.255	0.043	-0.370	0.101
38	0.362	0.144	-0.052	0.045	0.207	0.090
39	-0.089	0.058	0.178	0.034	0.252	0.093
40	-0.098	0.074	-0.159	0.036	-0.032	0.128
41	-0.497	0.126	-0.005	0.055	-0.325	0.130
42	-0.193	0.130	0.108	0.061	0.315	0.106
43	0.394	0.129	0.001	0.046	-0.255	0.109
44	0.327	0.170	-0.042	0.055	0.739	0.122
45	0.153	0.100	-0.156	0.034	-0.406	0.101
46	0.226	0.129	0.055	0.045	-0.193	0.102
47	-0.100	0.100	0.106	0.045	-0.182	0.097
48	-0.000	0.105	-0.106	0.038	0.290	0.076
49	-0.192	0.067	0.197	0.039	0.046	0.076
50	0.092	0.041	0.030	0.009	0.101	0.080
51	-0.039	0.098	-0.098	0.044	-0.461	0.091
52	0.015	0.131	0.166	0.044	0.164	0.079
53	-0.132	0.127	-0.082	0.044	0.316	0.061
54	0.040	0.097	-0.037	0.043	0.087	0.045
55	0.019	0.100	0.015	0.039	-0.173	0.034
56	-0.266	0.072	0.015	0.036		
57	0.280	0.061	-0.061	0.037		
58	0.023	0.009	0.045	0.009		
F-CSL	28.901		13.896		30.028	

Table 4: First-Stage Estimates: Region-Specific Time Trends

	CL		CCA		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 5: Baseline Estimates of External Returns to Schooling

	OLS	CL	CCA	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 CCA 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 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 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 6: Estimates of External Returns from Alternative Specifications

	CL	CCA	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 $\left\{\overline{CSL}_{jt}^{a}\right\}_{a=40}^{50}$			
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 CCA 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 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 with SOR Instruments

	CL	CCA	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 CCA and a composite measure of required schooling RS respectively. The difference from tables 5 and 6 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 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 8: Estimates of External Returns by CSL Cohorts

CSL Age Cohort	CL	CCA	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 CCA 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 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 9: Estimates of External Returns by Years of Schooling

Years of Schooling	OLS	CL	CCA	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 CCA 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 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 10: Estimates of External Returns from Workers Between 30 and 39 Years Old

	OLS	CL	CCA	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 5, Table 8 and Table 9, respectively. *** denotes 1%, ** denotes 5%, and * denotes 10% significance.

Table 11: 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 12: Key Moments: Data vs Model

	Data	Model	$\theta = 0$
External returns to schooling			
All workers			
OLS	0.069	0.072	0.079
2SLS	0.061	0.063	-0.007
Workers with 9 to 11 years of schooling			
OLS	0.048	0.046	0.046
2SLS	0.025	0.025	-0.004
Workers with 12 years of schooling or more			
OLS	-0.007	-0.007	-0.010
2SLS	-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: 2SLS	-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 13: 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 14: 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 6 for details.