INSTITUTO DE ECONOMÍA

DOCUMENTO de TRABAJO

466

2015

The Causal Impact of Human Capital on R&D and Productivity: Evidence from the United States

Ignacio Loeser, Verónica Mies, Matías Tapia

www.economia.puc.cl • ISSN (edición impresa) 0716-7334 • ISSN (edición electrónica) 0717-7593

The Causal Impact of Human Capital on R&D and Productivity: Evidence from the United States*

Ignacio Loeser † (P. Universidad Católica de Chile)

Verónica Mies‡ (P. Universidad Católica de Chile)

Matías Tapia § (P. Universidad Católica de Chile)

October 2015

Abstract

We use census micro data aggregated at the state level data for US cohorts born between 1915 and 1939 to test the impact of secondary and tertiary schooling in the US at the state-cohort level on R&D and TFP growth across industries in 1970. We instrument our measures of schooling by using the variation in compulsory schooling laws and differences in mobilization rates in WWII, which we relate to the education benefits provided by the GI Bill Act (1944). This novel instrument provides a clean source of variation in the costs of attending college. Two-stage least squared regressions find no effect of the share of population with secondary schooling on outcomes such as n R&D per worker or TFP growth. On the other hand, the share of population with tertiary education has a significant effect on both R&D per worker or TFP growth.
Keywords: Human Capital, R&D, TFP, secondary and tertiary education .

*We acknowledge the support from VRI Incio, Pontificia Universidad Catlica de Chile and from CONICYT/Programa de Investigacin Asociativa (Proyecto SOC 1102). We would also like to thank Juan Pablo Ossa for outstanding research assistance.

†mailto:jiloeser@uc.cl.

‡mailto:vmies@uc.cl. Address: Av. Vicuna Mackenna 4860, Macul, Santiago, Chile.

§mailto:mtapiag@uc.cl. Address: Av. Vicuna Mackenna 4860, Macul, Santiago, Chile.

1 Introduction

Theoretical models have long stressed the direct effect of human capital accumulation in growth and long-run income differences across countries (Lucas, 1988; Romer, 1986). Besides its direct effect on output as a productive factor, schooling can also impact output through changes in productivity if more educated workers lead firms to adopt more advanced technologies (Nelson and Phelps, 1966; Benhabib and Spiegel, 2002; Acemoglu and Zilibotti, 2001).

On the empirical side, there appears to be a strong empirical cross-country correlation between schooling and growth (Temple, 1999; De la Fuente and Domenech, 2001, 2006), as well as between schooling and TFP (Hall and Jones, 1999). These results seem to suggest that convergence in (quality-adjusted) measures of human capital should lead to convergence in income measures. However, over the last 50 years there has been strong convergence in average schooling years across countries while per capita income has not converged.

Moreover, cross-country studies typically have difficulties with endogeneity, so assessing causality remains troublesome. Individual country studies, which can potentially use stronger identification strategies, might be able to provide a clearer picture. However, when looking at data on US cities and states, Acemoglu and Angrist (2000) and Cicconi and Peri (2006) do not find any significant TFP effects of an increase in average schooling.

These apparent contradictions suggest two key issues that lie at the core of the empirical evaluation on the impact of human capital on income and technology. The first one, what is human capital, and how do we measure it? The second one, how do we deal with endogeneity issues?

Regarding the first question, both theory and data suggest that the composition of human capital matters. Aggregate measures of human capital (i.e. average years of schooling) can be misleading if different types of human capital are imperfect substitutes (Caselli and Coleman, 2002, 2006). In that sense, the composition of human capital can be crucial for determining aggregate income levels if workers with different educational levels are not perfect substitutes in production. In particular, the choice of production technology at the firm level might be determined by the relative abundance of different types of human capital. Advanced technologies require highly skilled workers, (evidence of skill biased technological change: Katz & Murphy, 1992; Autor, Katz, -& Krueger, 1998; -& Kearney, 2008), and thus will only be chosen by firms if the relative cost of highly skilled workers makes those technologies profitable. Thus, aggregate measures of human capital that are able to indicate compositional effects, such as the share of workers with tertiary/schooling, can complement the picture provided by a one-dimensional measure such as average years of schooling.

While the theoretical channel and the need to use separate measures might be clear, endogeneity concerns imply that obtaining empirical estimates of the impact of education on macro/micro outcomes is complicated. In equilibrium, technology and human capital are jointly determined by firms and households. More educated workers can drive firms to adopt advanced technologies, and more firms with advanced technologies can increase the return on schooling and the number of educated workers. Thus, reverse causality (Bils and Klenow, 2000) and omitted variables become a concern when trying to estimate the impact of human capital on firm decisions and outcomes. As a consequence, OLS estimates of the effects of schooling are potentially biased. Estimating the causal effect of human capital on technology and TFP implies finding a source of exogenous variation in human capital. If, as discussed above, one wants to distinguish between different types of human capital, more than one source of exogenous variation is required. In particular, separating the effect of secondary and tertiary schooling, as this paper does, requires at least two instruments.

At the individual country level, changes in education policy that have an impact on schooling, but are arguably exogenous to future technology, are a natural candidate to become an instrument for human capital. As an example, Acemoglu and Angrist (2000) exploit compulsory schooling laws (CLS) across states/time as an exogenous shifter in the number of years of high school per workers The same instrument has been widely used in subsequent studies (Ciccone and Peri, 2011; Iranzo and Peri, 2009; Milligan et al, 2004; Moretti and Lochner, 2004). While this is indeed a strong instrument for measures of human capital such as the number of schooling years or enrollment in secondary schooling, it appears as a less likely candidate for tertiary schooling. Moreover, and as discussed before, estimating the separate effect of secondary and tertiary schooling still requires an additional instrument. Finding this second source of exogenous variation has been somehow troublesome in the literature, and no consensus on a strong instrument has emerged as was the case with compulsory schooling laws. For example, Iranzo and Peri (2009) instrument tertiary education in the US across states by push-driven immigration and the share of state population living close to land grant colleges, while Aghion, Boustan, Hoxby and Vandenbussche (2009) rely on spending shocks in state level education produced by filling of vacancies in the education committee

This paper contributes to the empirical literature on the impact of human capital on technology adoption and the production structure of the economy by using census micro data aggregated at the state level data for US cohorts born between 1915 and 1939. We test the impact of secondary and tertiary schooling in the US at the state-cohort level on R&D and TFP growth across industries in 1970. While we follow the literature in using the variation in the timing of compulsory schooling laws across states to instrument secondary schooling, we propose a novel instrument for tertiary enrollment. In particular, we exploit, as in Acemoglu, Autor and Lyle (2004), the differences across states and cohorts in World War II mobilization rates. While Acemoglu, Autor, and Lyle (2004) used this variation as an exogenous shift in female labor supply, we exploit the fact that WWII veterans were benefited by the GI Bill Act (1944), which granted them free college education once they were discharged from service. This provides a clean source of variation in the costs of attending college, which allows us to exploit differences in college enrollment across states and cohorts.

Our results suggest that, consistent with the initial discussion, different types of human capital are associated to different effects on the productive structure of the economy. Two-stage least squared regressions find no effect of the share of population with secondary schooling over outcomes such as R&D per worker or TFP growth. On the other hand, the share of population with tertiary education has a significant effect on both R&D per worker or TFP growth. In particular, a 1% increase in the share of workers with tertiary education increases R&D per worker by 1.8 percentage points, and annual TFP growth by 1% for 17 years.

Section 2 presents a simple model that illustrates our underlying hypotheses. Section 3 discusses the data and the estimation strategy. Section 4 presents the empirical results. Section 5 presents some concluding remarks.

2 A simple analytical framework

We develop a stylized setup to illustrate how exogenous changes in schooling at different levels can have heterogeneous effects on R&D and TFP. This model is a static version of more general models that study education and TFP choices in a general equilibrium framework. Here we take education decisions as given, as we want to illustrate the mechanisms through which the economy's production structure can be affected by changes in human capital endowments. For simplicity, we show results on the extensive margin of technology adoption, but a similar model could be built to include an intensive margin for technology investments.

Endowments–. There are three types of workers endowed with different types of human capital. A measure λ of workers are college-educated workers, each one endowed with advanced human capital a_h. A measure γ of workers have high school as their highest degree, which endows them with general human capital, a_l. Finally, a measure $(1-\lambda-\gamma)$ of workers is unskilled and possesses basic human capital, which we normalize to 1, with 1, with $a_h > a_l > 1$.

Firms are heterogeneous and endowed with a productivity z drawn from a probability measure ψ_t on $[0, \infty)$.

Technologies–. The economy's final good can be produced with three distinct technologies. There is an innovative technology with productivity ζ that requires *R&D* investment which corresponds to the pay of a fixed cost F_h. This technology can only be operated with high-skilled workers

$$y = \zeta z^{\eta} h^{1-\eta},\ \eta \in (0,1)\,, \tag{1}$$

where h denotes units of high-skilled human capital. There is a second technology, the modern technology, which provides a productivity η, with $\zeta > \xi$ and requires low-skilled workers. This technology also involves investment $R\&D$, in particular, the payment of a fixed cost F_l with $F_h > F_l$.

$$y = \xi z^{\nu} l^{1-\nu},\ \nu \in (0,1)\,, \tag{2}$$

where l denotes units of low-skilled human capital. Finally, there is a basic CRS technology that uses basic (unskilled) human capital u. This technology does not involve a fixed cost and produces with a basic productivity A, with $\zeta > \xi > A$.

$$y = Au, \tag{3}$$

where u denotes units of unskilled human capital. There is imperfect substitution among the different types of human capital. Workers may work in sectors that require less skills than they possess. However, when working in those sectors, they only supply the minimum level of skills required by each technology.

Given their productivity and wages, firms choose which technology to operate and the quantity of labor to hire to maximize profits. The maximization problem for each sector is presented equations (4) to (6). Wages for high, low and basic human capital are denoted by w^h, w^l, w^u, respectively.

$$\text{innovative sector} \quad : \quad \pi_h\left(z\right) = \max_h \left(\zeta z^{\eta} h^{1-\eta} - w^h h\right) \tag{4}$$

$$\text{modern sector} \quad : \quad \pi_l\left(z\right) = \max_l \left(\xi z^{\nu} l^{1-\nu} - w^l l\right) \tag{5}$$

$$\text{basic sector} \quad : \quad \pi_u\left(z\right) = \max_l \left(Au - w^u u\right. \tag{6}$$

Optimal levels for each type of human capital yield the following expressions for profits in each sector[1]:

$$\text{innovative sector} \quad : \quad \pi_h^*\left(z;w^h\right) = z\zeta^{\frac{1-\eta}{\eta}}\eta\left(\frac{1-\eta}{w^h}\right)^{\frac{1-\eta}{\eta}} - F_h \tag{7}$$

$$\text{modern sector} \quad : \quad \pi_l^*\left(z;w^l\right) = z\xi^{\frac{1-\nu}{\nu}}\nu\left(\frac{1-\nu}{w^l}\right)^{\frac{1-\nu}{\nu}} - F_l \tag{8}$$

$$\text{basic sector} \quad : \quad \pi = 0. \tag{9}$$

Equilibrium Conditions–. Next, we define a competitive equilibrium in which all three sectors are open.

Definition 1 *Given a distribution of productivities $\phi(z)$, production functions presented in equations (1), (2), (3), fixed costs F_h and F_l, and supplies of workers $H^s = (1-\lambda)h$, $L^s = (1-\gamma)l$, U^s=$(1-\lambda-\gamma)u$, a competitive equilibrium in which all sectors are operating consists of*

i. Wages w^h, w^l, w^b paid to workers with high, low, basic human capital, respectively.

ii. A double cutoff productivities $(z*_h, z_l^*)$ such that firms with productivity $z > z_h^*$ operate innovating firms; firms with productivity $z_l^* < z < z_h^*$ operate modern firms and firms with productivity $z < z_l^*$ operate basic firms.

iii. Firms that maximize profits.

iv. Workers who do not wish to change occupation which implies $w^h \geq w^l \geq A$.

v. Markets that clear, i.e., $H^s = H^d$, $L^s = L^d$, and $U^s = U^d$ (subscripts s and d denote supply and demand, respectively).

[1] First order conditions are: $\zeta\,(1-\eta)\,(z/h)^{\eta} = w^h$, $\xi\,(1-\nu)\,(z/h)^{\nu} = w^h$, u is $w^u = A$.

Given wages $\left\{w^h, w^l, w^u\right\}$, there are two cutoffs productivities z_h^* and z_l^* that determine each firm's type.

$$
\begin{aligned}
z_h^*(\zeta,\xi) &= \frac{F_h - F_l}{\zeta^{\frac{1-\eta}{\eta}}\eta\left(\frac{1-\eta}{w^h}\right)^{\frac{1-\eta}{\eta}} - \xi^{\frac{1-\nu}{\nu}}\nu\left(\frac{1-\nu}{w^l}\right)^{\frac{1-\nu}{\nu}}} && (10)\\
z_l^*(\xi) &= \frac{F_l}{\xi^{\frac{1-\nu}{\nu}}\nu\left(\frac{1-\nu}{w^l}\right)^{\frac{1-\nu}{\nu}}} && (11)
\end{aligned}
$$

The conditions under which the three sectors operate in equilibrium are the following: i) at the reservation wage of workers with high human capital (i.e., $w_R^h = w^l$) profits in the innovative sector are nonnegative (equation 12; i) at the reservation wage of workers with low human capital (i.e., $w_R^l = A$) profits in the modern sector are nonnegative (equation 13. Conditions i) and ii) ensure that $w^h \geq w^l \geq A$ hold in equilibrium (check!) and iii) evaluated at the low productivity threshold z_l^*, the modern sector has larger profits than the innovating sector and profits' difference between the latter and the former sector increases with the level of the firm's productivity z (equation 14). Finally, as $z^l > 0$, the mass of firms in the basic sector will be positive.

$$
\begin{aligned}
z\zeta^{\frac{1-\eta}{\eta}}\eta\left(\frac{1-\eta}{A}\right)^{\frac{1-\eta}{\eta}} - F_h &\geq 0 && (12)\\
z\xi^{\frac{1-\nu}{\nu}}\nu\left(\frac{1-\nu}{A}\right)^{\frac{1-\nu}{\nu}} - F_l &\geq 0 && (13)
\end{aligned}
$$

$$
\frac{F_h}{F_l} \geq \frac{\zeta^{\frac{1-\eta}{\eta}}\eta\left(\frac{1-\eta}{w^h}\right)^{\frac{1-\eta}{\eta}}}{\xi^{\frac{1-\nu}{\nu}}\nu\left(\frac{1-\nu}{w^l}\right)^{\frac{1-\nu}{\nu}}} \geq 1 \qquad (14)
$$

Given the productivities cutoffs, the maximization problem presented in equations (4) to (6) yields the following demands for each type of labor: $H_t^d(\zeta) = \int_{z_h^*}^{\infty} z\left(\zeta\frac{1-\eta}{w^h}\right)^{\frac{1}{\eta}} \psi_t\,(dz)$ and $L_t^d(\zeta) = \int_{z_l^*(\zeta,\xi)}^{z_h^*} z\left(\xi\frac{1-\nu}{w^l}\right)^{\frac{1}{\nu}} \psi_t\,(dz)$. Wages for each type of workers derived from the market clearing conditions are the following:

$$w_h = \frac{\zeta(1-\nu)[\int_{z^*}^{Z} z\psi_t\,(dz)]^{\nu}}{H^s} \tag{15}$$

$$w_l = \frac{(1-\eta)[\int_{z^*}^{Z} z\psi_t\,(dz)]^{\eta}}{L^s} \text{and} \tag{16}$$

$$w_u = A \tag{17}$$

R&D and TFP–. R&D per sector corresponds to the sum of R&D performed by every firm in every sector and depends on the mass of firms in each sector: R&D in the high sector equals $R\&D_h = F_h \int_{z_h^*}^{\infty} \psi_t(dz)$ and R&D in the low sector corresponds to $R\&D_l = F_l \int_{z_l^*}^{z_h^*} \psi_t(dz)$. Aggregate R&D corresponds to total R&D. TFP assigned to each sector corresponds to the productivity of each sector weighted by the fraction of firms in that sector and aggregate TFP corresponds to sum of each of the previous measure. That is, TFP assigned to the high sector is $TFP_h = \zeta \int_{z_h^*}^{\infty} \psi_t(dz)/\Psi$; TFP in the low sector is $TFP_l = \xi \int_{z_l^*}^{z_h^*} \psi_t(dz)/\Psi$; and TFP in the basic sector is $TFP_b = A \int_0^{z_l^*} \psi_t(dz)/\Psi$, where $\Psi = \int_0^{\infty} \psi_t(dz)$ corresponds to the total mass of firms.

Comparative statics–. Now, we analyze the effects on R&D and TFP per sector of an exogenous increase in the number of workers with college (an increase of λ) maintaining the number of workers with high school unchanged. As high-skilled labor becomes more abundant, w^h falls (equation 15), the innovating sector becomes relatively more profitable and the cutoff productivity $z_h^\star$ falls (equation 10). Modern firms switch to the innovating sector which produces an increase in $R\&D_h$ and

TFP_h. As total demand for low-skilled workers falls, w^l and, consequently, $z_l^\star$ fall, too, and basic firms switch to the modern sector. In all, $R\&D$ and TFP in this sector decrease. At the aggregate level, R&D and TFP increase as in both margins (basic/modern and modern/innovative), firms switch to more technological sectors.

Next, we analyze the effects of exogenously increasing the number of workers with high school (an increase of γ) maintaining the number of workers with college unchanged. In contrast to the previous case, now low-skilled workers are more abundant producing a fall of w^l (equation 16). The modern sector becomes more profitable and expands in both margins as z_l^* falls and z_h^* raises. As a consequence the other two sectors shrink. Accordingly, $R\&D_h$ and TFP_h decrease and $R\&D_l$ and TFP_l increase. In this case, the total effect on both variables is ambiguous as some firms switch to a more advanced sector (basic to modern), but others, to a more backward one (innovative to modern).

3 Data and Empirical Strategy

We analyze the impact of different levels of schooling on technological outcomes, in the spirit of our theoretical framework, by building a state of birth/year of birth cohort panel for the US and running regressions of the form:

$$Outcome_{ij} = \alpha + \beta_1 Tertiary_{ij} + \beta_2 Secondary_{ij} + \theta X_{ij} + \mu_i + \delta_j + \epsilon_{ij} \qquad (18)$$

for cohort j born in state i. Outcomes variables are measures of the technological characteristics of jobs, such as R&D per worker and TFP growth in the sectors and occupations in which agents in a given state/cohort are working in a given point of time. Regressors include state and cohort fixed effects and two different measures of education, enrollment in secondary and tertiary schooling. ‘X” accounts for additional controls.

Our main source of information is IPUMS,from which we extract a 1% sample of 1970 census data for males born in the US between 1915 and 1939. This gives us a sample of 287,336 individuals, born in 49 different states (we exclude Alaska and Hawaii) and belonging to 25 different birth cohorts. For each individual, we extract information on his birth state, highest educational degree attained, war veteran status, labor status in 1970, and, for working men, their occupation and working sector. Notice that, in 1970, our oldest individuals are 55 years old, while the youngest are 31, so they are all working age and one can reasonably assume that they have completed their education. We collapse these individuals observations by birth place (state) and year of birth (cohort). This allows us to build a balanced panel with 1225 observations. In the panel, individual level dummies that measured a specific characteristic (such as whether an individual was a WWII veteran or was workimg in a particular sector) now measure the share of individuals with a given characteristic for a state/cohort.

Our choice of sample, and the way we construct our variables, is directly related to our identification strategy. Our interest lies on identifying the causal impact of two separate measures of human capital, the amount of workers with secondary schooling and the amount of workers with tertiary schooling, on economy’s the production structure. Secondary (Tertiary) schooling is measured as the share of individuals in state/cohort whose highest degree is some level of secondary (tertiary) schooling.To address endogeneity concerns, we then require two separate instruments.

For secondary schooling, we follow Acemoglu and Angrist (2000), and use the state-level laws in the number of years of compulsory school attendance. We can argue that the implementation of these laws was not related to technological and labor market outcomes in 1970, but rather a reflection of wide political discussions during the early XXth century (the "high school movement", Goldin, 2001). We exploit the state-time variation in the legislation that implies that individuals at different state/cohorts were probably facing different regulations. The census data only provides us information on the individual's birth state, and on his residence state in 1970. Thus, we do not have information on the residence state at the time when his secondary schooling attendance decisions were made. However, we argue that, given the young age at which people typically decide whether to finish or drop out from secondary schooling (14-18 years old), state of birth provides a more reasonably approximation to his state of residence at that time than his state of residence several decades in the future. Moreover, using state of residence in 1970 implies that our results could be contaminated by migration decisions, a concern that is not relevant for the birth state as individuals cannot choose where they are born.

Our instrument for tertiary schooling is the paper's main empirical contribution, as we think it can provide a cleaner source of exogenous variation that other instruments that have been used in the past. Our initial inspiration comes from the WWII mobilization rate instrument used by Acemoglu, Author and Lyle (2004), which we combine with the fact that WWII veterans were elegible for free tertiary education upon their return from the war."The Servicemen's Readjustment Act of 1944", better known as G.I. Bill was a public policy that provided financial help to WWII veterans (1940-47) to start a business or attend college. Thus, this policy significantly reduced the costs of tertiary education for WWII veterans. Several papers (Bound & Turner, 2002; Stanley, 2014) show empircally that, in fact, was associated to an important increase in tertiary enrollment for the cohorts that received the benefits. We argue that this policy is a valid instrument as the G.I. Bill was first discussed in the US Congress on January 10th of 1944. By then, the recruitment process was mandatory, so arguably no WWII veteran enlisted to obtain benefits, a concern for the validity for other iterations of the GI Bill in future conflicts. Moreover, the goal of the GI Bill was to provide support for veterans and, as with the compulsory schooling laws, it seems hard to argue that its implementation was related to production technology outcomes in 1970.

We build this instrument by proxying access to the G.I. Bill educational benefits by the WWII mobilization rate in a given state/cohort interacted by the eligibility of each cohort to be a beneficiary. We argue that mobilization rates were not related to individual characteristics that could be related to college attendance, such as intellectual ability or discipline. As in Acemoglu, Author and Lyle (2004), we include variables that may have affected mobilization rates by state, such as racial composition (blacks were relatively underrepresented) and productive structure (mobilization rates were smaller in states with a larger share of agriculture), and our results are not changed. Notice that the mechanism underlying this second instrument differs from the one on compulsory schooling laws. While the first instrument operates by increasing secondary enrollment directly forcing agents who would have otherwise chosen not to do to stay in school, the GI Bill instrument represents a reduction in education costs, which can change the decisions of agents who would have otherwise not attended college, particularly those that were financially constrained and those for which the net returns of tertiary schooling, with tuition, were marginally negative.

All outcome variables are measured in 1970. The outcome variables measure the characteristics of different jobs and occupations, and the intensity in which members of a given state/cohort work in different types of jobs. To construct the variables, we first use industry level data for 1970 to construct industry level data, and combine that data from the census data on the sector and occupation in which individuals are working.

Average R&D per worker uses data from the National Science Foundation to construct a measure of R&D spending per employee by industry, at the national level, in 1970. The same data is used to construct a measure of R&D spending per sales. Similarly, data from KLEMS is used to construct measures of average annual TFP growth across industries. Finally, the R&D spending data is used to define innovative/ non-innovative occupations, related to how different types of occupations are used across different industries.

For each of these variables, the national level data is combined with the state/cohort data. For instance, each worker is assigned the R&D spending per employee of the sector in which he is working in 1970. By collapsing the data by state and cohort, we get a state/cohort measure of R/D spending, measured as the average of R/D spending across industries, weighted by the share of employment the industry has in that particular state/cohort.

Notice that our outcome variables can not be interpreted as realizations taking place in a given state in 1970, but rather as the labor outcomes in 1970 of different cohorts born in a given state. This is, our outcome is not R&D per worker in California in 1970 from workers born in 1925, but the R&D per worker in 1970 associated to workers born in California in 1925.This means that our empirical exercise implicitly defines the US an integrated labor market, with state/cohort observations providing us variation in terms of human capital composition. Under this interpretation, endogeneity on residence decisions in 1970 is not an issue in our exercise.

Descriptive statistics for all variables are presented in Tables 1 to 6.

4 Estimation results

As we discussed in detail in the previous sections the sources of variations of our measures of secondary and tertiary education could be correlated with R&D and TFP and OLS estimates may be biased. In order to address this problem, we adopt an instrumental variable strategy. As we want to evaluate the impact of secondary and tertiary education, we need to use at least two instruments, namely the Compulsory Schooling Laws (CSL) and the G.I. Bill Benefits. All regressions control for state of birth and cohort effects. Standard errors are clustered at the state level.

Tables 7 and 8 reports the estimates of the two instruments on Secondary Education Coverage and Tertiary Education Coverage per cohort and state, respectively. In both tables, the second column shows estimations controlling for the share of farmers and share of non-whites. Table 7 shows that Compulsory Schooling Laws have a positive and significant effect on Secondary Education Coverage. According to the estimates, one more year of compulsory schooling increases Second Education Coverage by 2.5 percentage points. The second instrument, in contrast, shows no significant effect on it. These results suggest that Compulsory Schooling Laws did affect the decision of going to high school but did not generally affect the decision of attending College. A caveat; one might have expected a positive effect of CLS on Tertiary Education Coverage considering that a larger share of people with secondary education translates to a larger share of people that can access tertiary education. However, to enter college, students have to have completed secondary education and not just have obtained some secondary education which is the measure that we are using here. In this sense, the result suggests that a higher compulsory schooling attendance predicts that more people attain some secondary education, but not necessarily that these people will attain a high school Diploma.

We find exactly the opposite effect when we analyze the explanatory power of both instruments on Tertiary Education Coverage. Table 8 reports that Compulsory Schooling Laws have no significant impact on this variable, while the second instrument, G.I. Bill Benefits, has a positive and significant impact on it. G.I. Bill Benefits reduced the costs of attaining education and thus affected the margin of going/not going to College. While the benefits of the law were not confined to finance tertiary education, find a significant effect only for tertiary education is expected since the G.I Bill beneficiaries were veterans over 18 years old and, therefore, were likely to affect people who already had attained some secondary education.

4.1 Two stage estimates

In this section we discuss regressions results for the specifications described in equation (18). We will analyze the effects of the different types of human capital coverage on two outcome variables: R&D per worker and TFP. The definition and construction of all data used here is presented in section *Data and Definitions*.

Before discussing our instrumental variables results, we present the results for ordinary least squares regressions (OLS) of Secondary and Tertiary Education on our measures of R&D per worker and TFP, both per state and cohort. The first column of Table 9 shows the effects of the different types of educational measures on R&D. We find that Secondary Education has a significant and negative impact on

R&D per worker and tertiary education has a significant and positive effect on this variable. These results are in line with the ones predicted by our theoretical model; however, as we discussed in detail in the previous sections, estimates may be biased. Columns 2 and 3 show the IV estimation results for R&D per worker. We control for Race and for "Working on a Farm" in the last column. Tertiary Education Coverage has a significant and positive effect on R&D per worker. According to the point estimate reported in column 3, an increase in one percentage point in Tertiary Education Coverage increases R&D per worker in 1.8 percentage points. Comparing it with the OLS coefficient, the IV estimate is more than two-times larger. Secondary Education Coverage, in contrast, shows no significant impact on R&D per worker while the OLS estimate shows a negative and significant impact on this measure.

Both IV results are in line with our theoretical model. In the case of Secondary Education Coverage, the lack of significance of the point estimate may reflect the two opposing effects of this variable on R&D already discussed in the analytical framework.

The result for TFP is consistent with the result obtained for R&D. Table 10, column 3, shows that for Secondary Education Coverage, the impact is again not significant while the impact of tertiary education is positive and significant. In the latter case, a one percent increase in Tertiary Education Coverage predicts a one percent increase in the level of TFP in a time span of 13 years.

One could argue that R&D per worker does not necessarily measure different R&D efforts if sectors produce with different labor intensities. As a robust exercise, we scale R&D expenditures by the level of sales instead of the number of workers. The results remain unchanged: Secondary Education Coverage has no significant impact on this new measure while Tertiary Education Coverage shows a positive and significant impact.

A different way to evaluate the impact of different types of human capital on R&D related measures consists of analyzing whether workers with higher levels of human capital work in activities associated with higher levels of R&D. To this end, we measure the impact of the two measures of education on the type of occupation in which the employee works. The results are again consistent with the previous ones. Column 1 of Table 10 reports the results. Secondary education has no significant impact on the type of activities performed by employees while a higher tertiary education coverage is associated with activities that are more R&D intensive.

In this analysis, tertiary education is instrumented by the G.I. Bill Benefits. One could argue that the effects found from tertiary education on R&D and TFP are explained because the veterans acquired special skills that are important for R&D and TFP and not because they have gone to College after the war. A simple way to explore this hypothesis is to analyze whether there are significant differences between people with College who are veterans versus those who did not go to war. Table 11 reports some descriptive statistics for the various measures of R&D and TFP for each group. As shown there, there are no significant differences between groups.

Robustness checks. In the previous section, we defined that a worker had secondary/tertiary education if she had spent some time in high-school/college, regardless of whether she had been in the system a couple of months or had completed the corresponding level. However, one could argue that the effects of each type of human capital depends not only on whether the person had access to a certain level of education but also on the time spent studying at that level. In this section, we repeat the previous exercises, but defining secondary/tertiary education as the average time spent in each education level. In contrast to the previous variables, these new definitions measure *education intensity.*

Table 12 presents the first stage estimates for Average Years of Secondary Schooling (column 2) and Average Years of Tertiary Education (column 3). Comparing them with the previous results, we find that in the case of Average Years of Secondary Schooling both instruments, Compulsory Schooling Laws, and G.I. Bill Benefits have a positive and significant effect on it. We interpret the change of significance of G.I. Bill Benefits as follows. It is likely that a percentage of people who went to war had not obtained a high school diploma. Returning from the war, these veterans used G.I. Bill Benefits to complete secondary education. It is important to note that if they got some secondary education before going to war, completing it afterwards does not affect the measure of education used in the previous section (Secondary Education Coverage), but does affect average years of schooling spent at high school. This would explain why under this new measure of education, the impact of G.I. Bill Benefits became significant. In the case of Average Years of Tertiary Education, results remain unchanged and only G.I. Bill Benefits has a significant impact on it.

Table 13 reports the estimation results for the second stage. The dependent variable is R&D per worker. The results are consistent with the previous measures of education. Average Years of Secondary Schooling has no significant impact on R&D per worker while Average Years Tertiary Education has a positive and significant impact on R&D per worker. Again, the IV point estimate for this latter variable is more than two times higher than the OLS point estimate (column 1). This suggests that the OLS biases these new measures of education through the same channels than it did under the previous definition. Results are robust to the other outcomes measures, namely R&D per sales and Occupation Type. Columns 1 and 2 of Table 14 report the results. Finally, the impacts on TFP are similar under both measures of education. Average Years of Secondary Schooling has no impact on TFP and Average Years of Tertiary Education has a positive and significant impact on it.

5 Conclusions

This paper contributes to the empirical literature on the impact of human capital on technology adoption and the production structure of the economy by using census micro data aggregated at the state level data for US cohorts born between 1915 and 1939. We test the impact of secondary and tertiary schooling in the US at the state-cohort level on R&D and TFP growth across industries in 1970. While we follow the literature in using the variation in the timing of compulsory schooling laws across states to instrument secondary schooling, we propose a novel instrument for tertiary enrollment. In particular, we exploit the differences across states and cohorts in World War II mobilization rates and, particularly, the fact that WWII veterans were benefited by the GI Bill Act (1944), which granted them free college education once they were discharged from service. This provides a clean source of variation in the costs of attending college, which allows us to exploit differences in college enrollment across states and cohorts.

Our results suggest that different types of human capital are associated to different effects on the productive structure of the economy. Two-stage least squared regressions find no effect of the share of population with secondary schooling over outcomes such as R&D per worker or TFP growth. On the other hand, the share of population with tertiary education has a significant effect on both R&D per worker or TFP growth. In particular, a 1% increase in the share of workers with tertiary education increases R&D per worker by 1.8 percentage points, and annual TFP growth by 1% for 17 years.

References

[1] Acemoglu, Daron, and Joshua Angrist, How Large Are the Social Returns to Education? Evidence from Compulsory Schooling Laws (pp. 959), in Ben Bernanke and Kenneth Rogoff (Eds.), NBER Macroeconomic Annual (2000).

[2] Acemoglu, Daron, David H. Autor and David Lyle. Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury. Journal of Political Economy (2004). Vol. 112(3). 497-547.

[3] Acemoglu, Daron, and Fabrizio Zilibotti, Productivity Differences, Quarterly Journal of Economics 116 (May 2001), 563-606.

[4] Aghion Philippe, Leah Boustan, Caroline Hoxby, and Jerome Vandenbussche J. The Causal Impact of Education on Economic Growth: Evidence from U.S. Manuscript. (2009).

[5] Autor, David, Lawrence Katz, and Alan Krueger, Computing Inequalit Have Computers Changed the Labor Market? Quarterly Journal of Economics 113:4 (November 1998), 1169-1213.

[6] Autor, David, Lawrence Katz, and Melissa Kearney, Trends in U.S. Wage Inequality. Revising the Revisionists, Review of Economics and Statistics 90:2 (May 2008), 300-323.

[7] Benhabib, Jess and Mark Spiegel, The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data, Journal of Monetary Ecomics 34: (1994), 142-173. Journal

[8] Bils, Mark, and Peter J. Klenow "Does Schooling Cause Growth?" American Economic Review 90(5): (May 2000) 1160-1183.

[9] Bound, John, and Sarah E. Turner. "Going to War and Going to College: Did the World War II and the G.I. Bill Increase Educational Attainment for Returning Veterans?" Journal of Labor Economics (2002). Vol 20(4): 784-815.

[10] Caselli, Francesco, and John Coleman, The U.S. Technology Frontier, American Economic Review PP 92:2 (May 2002), 148-152.

[11] Caselli, Francesco, and John Coleman, The World Technology Frontier, American Economic Review 96:3 (June 2006), 499-522.

[12] Ciccone, Antonio, and Giovanni Peri, Identifying Human Capital Externalities: Theory with Applications, Review of Economic Studies 73 (April 2006), 381-412.

[13] Ciccone, Antonio, and Giovanni Peri, Schooling Supply and the Structure of Production: Evidence from US States 1950-1990. Barcelona GSE Working Paper Series Working Paper 595 (November 2011).

[14] De la Fuente, Angel, and Rafael Domenech, Schooling Data, Technical , and the Neoclassical Model, American Economic Review PP 90:5 (May 2001), 323-327.

[15] De la Fuente, Angel, and Rafael Domenech, Human Capital in Growth Regressions: How Much Difference Does Data Quality Make? Journal of the European Economic Association 4:1 (March 2006), 136.

[16] Goldin, Claudia. The human-capital century and American leadership: Virtues of the past. The Journal of Economic History (2001). Vol 61(2): 263-292.

[17] Hall, Robert E., and Charles I. Jones, Why Do Some Countries Produce So Much More Output per Worker than Others? Quarterly Journal of Economics 114:1 (February 1999), 83116.

[18] Iranzo, Susana, and Giovanni Peri, Schooling Externalities, Technology and Productivity: Theory and Evidence from U.S. States, Review of Economics and Statistics91(2): (May 2009), 420-431.

[19] Katz, Lawrence F., and Kevin M. Murphy, Changes in Relative Wages 19631987: Supply and Demand Factors, Quarterly Journal of Economics 107:1 (February 1992), 35-78.

[20] Lucas, Robert, Jr. On the mechanics of economic development. Journal of Monetary Economics 22: (1988), 3-42.

[21] Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos, Does EducationImprove Citizenship? Evidence from the U.S. and the U.K., Journal of Public Economics 88:9 (August 2004), 1667-1695.

[22] Moretti, Enrico, and Lance Lochner, The Effect of Education on Criminal Activity: Evidence from Prison Inmates, Arrests and SelfReport, American Economic Review 94:1 (March 2004), 155-189.

[23] Nelson, Richard and Edmund Phelps. "Investment in Humans, Technological Diffusion, and Economic Growth," American Economic Review. Papers and Proceedings (1966). Vol. 56: 69-75.

[24] Romer, Paul M. "Increasing Returns and Long-run Growth," Journal of Political Economy 94(5): (October 1986), 1002-37.

[25] Temple, Jonathan, The New Growth Evidence, Journal of Economic Literature 37:1 (March 1999), 112-15.

Tables

Table 1: Education Summary statistics
Divided by Age

	Main Variables		Alternative Variables	
Born Between	Secondary Education	Tertiary Education	Av. Years of Sec.	Av. Years of Ter.
1915 and 1919	.6478208	.2303981	3.289096	.719899
	(.1061528)	(.0943142)	(.6712542)	(.304942)
1920 and 1924	.627985	.2779322	3.533455	.879339
	(.0943674)	(.1096548)	(.6252588)	(.343875)
1925 and 1929	.6274554	.2995683	3.615228	.978012
	(.0816643)	(.0955858)	(.5445332)	(.338956)
1930 and 1934	.6106361	.3311547	3.842793	1.109317
	(.0831992)	(.1053622)	(.5076084)	(.390670)
1935 and 1939	.6088792	.3595022	4.048984	1.147925
	(.0982925)	(.1111888)	(.4211492)	(.392974)
All the Sample	.6245553	.2997111	3.665911	.9668982
	(.0941127)	(.1124238)	(.6177948)	(.388089)

Note: Standard errors in parenthesis.

Table 2: Education Summary statistics
Divided by Region

	Main Variables		Alternative Variables	
Region	Secondary Education	Tertiary Education	Av. Years of Sec.	Av. Years of Ter.
NE region	.6516965	.2981499	3.735242	.9867831
	(.0895368)	(.1045539)	(.4561561)	(.3769322)
MW region	.6720552	.3133483	3.936669	1.012854
	(.0583976)	(.0708933)	(.3463321)	(.2513227)
S region	.5880387	.2129355	2.999895	.6690638
	(.066772)	(.0677023	(.5038629)	(.2133631)
W region	.5838172	.3892322	4.075713	1.24525
	(.1204234)	(.124293	(.4971682)	(.4430465)
Total	.6245553	.2997111	3.665911	.9668982
	(.0941127)	(.1124238)	(.6177948)	(.3880891)

Note: Standard errors in parenthesis.

Table 3: Instruments Summary statistics
Divided by Age

Born Between	Compulsory Schooling Laws	G.I. Bill Benefits
1915 and 1919	8.69	0.00
	(1.23)	
1920 and 1924	8.77	.77
	(1.15)	(.08)
1925 and 1929	8.84	.56
	(1.06)	(.30)
1930 and 1934	9.17	0.00
	(.90)	
1935 and 1939	9.208	0.00
	(.85)	
All the Sample	8.93	.27
	(1.07)	(.36)

Note: Standard errors in parenthesis.

Table 4: Instruments Summary statistics
Divided by Region

Region	Compulsory Schooling Laws	G.I. Bill Benefits
NE region	8.81	.28
	(.63)	(.38)
MW region	9.24	.26
	(.99)	(.35)
SR region	8.59	.24
	(1.23)	(.33)
W region	9.14	.27
	(1.21)	.(37)
Total	8.93	.26
	(1.06)	(.36)

Note: Standard errors in parenthesis.

Table 5: Technology Summary statistics
Divided by Age

Born Between	Average R& D per worker	R& D by Sales	Occ. Type	TFP Variation
1915 and 1919	2.01632	.070634	.2243643	.0180957
	(.3605704)	(.0131009)	(.0632533)	(.0017551)
1920 and 1924	2.09595	.0740295	.2307955	.0181849
	(.4377093)	(.0165306)	(.059414)	(.0021364)
1925 and 1929	2.05558	.0728159	.2323364	.0179529
	(.3714016)	(.0138524)	(.060286)	(.001946)
1930 and 1934	2.123333	.0752994	.2451789	.0181607
	(.3904951)	(.0148303)	(.0664801)	(.0020329)
1935 and 1939	2.114231	.0750066	.2454296	.0182284
	(.3884444)	(.0139437)	(.0669672)	(.0018244)
All the Sample	2.081082	.0735571	.2356209	.0181245
	(.3920147)	(.0145752)	(.0638007)	(.001943)

Note: Standard errors in parenthesis.

Table 6: Technology Summary statistics
Divided by Region

Region	Average R& D per worker	R& D by Sales	Occ. Type	TFP Variation
NE region	2.214328	.0789095	.2531334	.0191421
	(.4498603)	(.0167125)	(.0704222)	(.0021582)
MW region	2.015542	.0713491	.2286115	.018474
	(.2156565)	(.0082209)	(.0347031)	(.0015455)
SR region	1.889254	.0660709	.2043213	.0172055
	(.217847)	(.0078603)	(.0367744)	(.001002)
W region	2.221816	.0784877	.2595615	.0176266
	(.5010311)	(.0184853)	(.0851544)	(.0022251)
Total	2.081082	.0735571	.2356209	.0181245
	(.3920147)	(.0145752)	(.0638007)	(.001943)

Note: Standard errors in parenthesis.

Table 7: First Stage on Secondary Education

Dependent Variable: Secondary Education Coverage (%)

	(1)	(2)
Compulsory Schooling	0.0249***	0.0247***
	(0.0083)	(0.0080)
G.I. Bill Benefits	-0.0192	-0.0157
	(0.0541)	(0.0583)
Farm Control		0.0440
		(0.0636)
Race Control		0.0388
		(0.0928)

Note: Standard errors (in parentheses) account for clustering on states. Number of observations: 1225, 49 states and 25 cohorts. Regressions with time and state fixed effects.

Table 8: First Stage on Tertiary Education

Dependent Variable: Tertiary Education Coverage (%)

	(1)	(2)
Compulsory Schooling	-0.0052	-0.0056
	(0.0035)	(0.0037)
G.I. Bill Benefits	0.1800***	0.1880***
	(0.0666)	(0.0672)
Farm Control		0.0867**
		(0.0399)
Race Control		0.0314
		(0.0779)

Note: Standard errors (in parentheses) account for clustering on states. Number of observations: 1225, 49 states and 25 cohorts. Regressions with time and state fixed effects.

Table 9: OLS and IV regressions.
Dependent Variable: Average R& D per worker (USD)

	(1) OLS	(2) IV	(3) IV
Secondary Education	-0.427**	-0.561	-0.591
	(0.212)	(0.617)	(0.625)
Tertiary Education	1.564***	3.743***	3.650***
	(0.306)	(0.890)	(0.831)
Farm Control			-.278
			(0.183)
Race Control			-0.867**
			(0.411)

Note: Standard errors (in parentheses) account for clustering on states. Number of observations: 1225, 49 states and 25 cohorts. Regressions with time and state fixed effects. 2SLS Regressions.

Table 10: IV regressions.
Different Dependent Variable

	(1) IV	(2) IV	(3) IV
Dependent Variable	R&D Sales	Occupation Type	TFP Variation
Secondary Education	-0.0170	-0.0821	-0.000554
	(0.0218)	(0.132)	(0.00516)
Tertiary Education	0.142***	0.650***	0.0202***
	(0.0312)	(0.220)	(0.00743)
Farm Control	-0.00622	-0.0561	-0.00121
	(0.00462)	(0.0503)	(0.00104)
Race Control	(0.0154)	(0.0425)	(0.00185)

Note: Standard errors (in parentheses) account for clustering on states. Number of observations: 1225, 49 states and 25 cohorts. Regressions with time and state fixed effects. 2SLS Regressions.

Table 11: Summary Statistics: Outcome Variables

Variable	Individuals who went to college and are WWII veterans	Individuals who went to college and did not go to war
Average R& D per worker	3.14	3.13
	(2.35)	(2.37)
R&D by Sales	0.114	0.113
	(0.092)	(0.091)
TFP Variation	0.021	0.02
	(0.012)	(0.011)
Occupation Type	0.432	0.405
	(0.495)	.(0.491)
Number of Observations	21015	3762

Note: Standard errors in parenthesis.

Table 12: First Stage, Education measured as average years
Dependent Variable: Average Years of Secondary and Average Years of Tertiary

	(1) Secondary (years)	(2) Tertiary (years)
Compulsory Schooling	0.0710***	-0.0248
	(0.0147)	(0.0149)
G.I. Bill Benefits	1.054***	0.504*
	(0.213)	(0.256)
Farm Control	0.499***	0.0348
	(0.128)	(0.146)
Race Control	0.600	-0.130
	(0.488)	(0.322)
R-squared	0.607	0.362

Note: Standard errors (in parentheses) account for clustering on states. Number of observations: 1225, 49 states and 25 cohorts. Regressions with time and state fixed effects.

Table 13: OLS and IV regressions. Different measure of education.
Dependent Variable: Average R& D per worker (USD)

	(1) OLS	(2) IV	(3) IV
Average Years of Secondary	0.146**	0.00109	-0.00526
	(0.0595)	(0.179)	(0.182)
Average Years of Tertiary	0.535***	1.369***	1.389***
	(0.0772)	(0.420)	(0.411)
Farm Control			-0.0333
			(0.175)
Race Control			-0.592
			(0.439)

Note: Standard errors (in parentheses) account for clustering on states. Number of observations: 1225, 49 states and 25 cohorts. Regressions with time and state fixed effects.

Table 14: IV regressions. Different measure of education.
Different Dependent Variable

	(1)	(2)	(3)
	IV	IV	IV
Dependent Variable	R&D Sales	Occupation Type	TFP Variation
Average Years of Secondary	0.000973	0.00357	0.000501
	(0.00633)	(0.0407)	(0.00137)
Average Years of Tertiary	0.0515***	0.237***	0.00650**
	(0.0156)	(0.0750)	(0.00326)
Farm Control	0.00309	-0.0134	4.14e-05
	(0.00568)	(0.0487)	(0.00106)
Race Control	-0.0180	0.0127	3.32e-05
	(0.0166)	(0.0710)	(0.00197)

Note: Standard errors (in parentheses) account for clustering on states. Number of observations: 1225, 49 states and 25 cohorts. Regressions with time and state fixed effects. 2SLS Regressions.