

# Signaling and Employer Learning with Instruments<sup>\*</sup>

Gaurab Aryal<sup>†</sup>

Manudeep Bhuller<sup>‡</sup>

Fabian Lange<sup>§</sup>

January 26, 2019

## Abstract

We show how to use instruments to disentangle the social and the private returns to education when schooling increases productivity and is used to signal productivity. Within the employer learning framework (Farber and Gibbons (1996) and the Altonji and Pierret (2001)), and absent production externalities, IV estimates of the returns to education among experienced workers identify the social return of education. What IV estimates identifies among less experienced workers depends on assumptions about how informed employers are about the instruments. If employers observe the instrument and are aware of their correlation with ability and schooling, then instrumental variables identify the social returns to education. When employers are uninformed, they identify the private returns to education, at each experience level. As experience accumulates, these estimates (of private returns) converge to the social returns to education. We use an IV based on local variation in compulsory schooling laws across multiple cohorts in Norway to implement these ideas. Our findings indicate social returns to education at 6.4%. This compares with a private internal rate of return aggregating the returns over the life-cycle of 8%. These estimates indicate that 80% of the private return to education can be attributed to education raising the productivity of workers. The remaining 20% can be attributed to the signaling value of education.

**Keywords:** signaling, human capital, employer learning, instruments

**JEL code:** J24, J31, D83

---

<sup>\*</sup>We thank Emily Nix and participants at the 2018 Cowles Conference in Honor of Joseph Altonji and seminar participants at University of Bergen, Ohio State University, Norges Bank and Goethe University Frankfurt for helpful comments.

<sup>†</sup>University of Virginia, Department of Economics. E-mail: aryalg@virginia.edu

<sup>‡</sup>University of Oslo, Department of Economics; Statistics Norway. E-mail: manudeep.bhuller@econ.uio.no

<sup>§</sup>McGill University, Department of Economics. E-mail: fabolange@gmail.com

# 1 Introduction

Two competing models rationalize positive returns to education. According to the human capital model (Becker, 1962) education increases skills valued by employers. The model of job-market signaling (Spence, 1973) instead posits that education signals innate differences in productivity, among workers, to potential employers. Both models, however, imply a positive relationship between education and earnings.

From an individual worker’s perspective, it is irrelevant whether education increases earnings because of job-market signaling or because of human capital acquisition. However, signaling is socially inefficient as workers expend valuable resources to signal their productivity –resources that otherwise could be spent on other productive activities. Abstracting from other spillovers of education, if we define *social returns to education* to be solely the effect of education on productivity, signaling creates a wedge between private and social returns by raising the private returns above the social returns, which in turn can lead to over-investment in education.<sup>1</sup>

Education policy, thus requires empirical guidance on the relative contribution of signaling and human capital acquisition to the returns to education. It is, however, difficult to empirically separate these two effects, for at least two reasons. First, both models postulate that observable earnings reflecting payments for latent skills –skills that the signaling model specifically assumes is privy only to the employers and is thus unobserved by employers, and

---

<sup>1</sup>We thus abstract from spillover effects of education that arise because education might induce productive externalities beyond the employer-employee relationship and because education might entail various non-production and non-pecuniary benefits of education such as reducing crime or improving public health; see Lange and Topel (2006), Lochner (2011), and Oreopoulos and Salvanes (2011) among others. The literature suggests that such education spillovers tend to be positive, which implies that the social returns to education can actually outweigh the individual increases in worker productivity, and can provide a rational for large public investments in education (OECD, 2018). Likewise, signaling might improve the match between a worker and her tasks (Pastorino, 2018). Our study abstracts from these various other avenues through which education might generate positive social returns and instead considers only the productive and signaling effects of education on the own work activity individuals engage in.

thus presumably also by researchers. And, second, both models impose the same behavioral restrictions: workers choose schooling to maximize the present value of earnings and employers maximize profits by employing workers as long as wages do not exceed productivity. The models thus do not provide additional behavioral restrictions that can be used to identify the model using data on wages and education only. These difficulties in separately identifying signaling from human capital acquisition have long been recognized in the literature; see Lange and Topel (2006) and references therein.<sup>2</sup> These difficulties ultimately result in widely divergent estimates of the contribution of signaling to the returns to education.

One way forward is to make the a priori reasonable assumption that employers learn over time. In two influential papers, Farber and Gibbons (1996) and Altonji and Pierret (2001) (henceforth FG and AP, respectively) propose a tractable model of employer learning (henceforth, EL), and use it to test if employers use schooling to infer unobserved ability. The key identifying assumption underlying their approach is the availability of a correlate of unobserved ability to researcher that is not available to employers. FG propose the AFQT score available in the NLSY1979 to be such a correlate. They show that the increasing association between the AFQT score and earnings, over the life-cycle, is consistent with the assumption that employers learn about workers abilities over time. AP then show that the partial correlation (after controlling for the AFQT score) between earnings and schooling declines over the life-cycle, consistent with statistical discrimination on the basis of schooling.

Following their lead, Lange (2007) provides evidence that employers learn fast, and also shows that the contribution of signaling cannot be point identified in the EL model, *even* when a correlate of ability such as the AFQT score is available. However, using the first-order condition for schooling decision Lange (2007) identifies bounds for the contribution of signaling and estimates it to be less than 25% of the private returns to schooling over the life-cycle.

We contribute to this literature by estimating a model of EL that builds on FG, AP,

---

<sup>2</sup>Some other papers in this area include Tyler et al. (2000); Bedard (2001); Fang (2006); Hopkins (2012); Clark and Martorell (2014); Feng and Graetz (2017) and Arteaga (2018).

and Lange (2007) using a conventional instrumental variable for schooling. IV estimates over the life-cycle allow us to point-identify the effect of education on productivity and the signaling returns to education. This allows us to answer the question of how much do the private and social returns to education differ from each other due to signaling. Unlike Lange (2007), however, we do not have to rely on the assumption that we can measure the major components to the costs of schooling that enter the optimality (first-order) condition for schooling that he exploits. In fact, we can dispense with using this optimality condition entirely, thus we do not have to impose strong behavioral assumptions behind the optimality condition, which in turn means that our estimates are robust. Equally important is the fact that our result does not rely on the assumption that we have access to a correlate of ability that is not observed by employers.<sup>3</sup>

Our approach rests on carefully considering the following question: “What do conventional IV estimates of returns to education identify within the EL framework?” Our first result is that, within an EL framework, any conventional IV estimate of returns to education in earnings measured at sufficiently high levels of experience is a consistent estimate of the causal effect schooling on productivity.<sup>4</sup> This result implies that, as long as the structure of the EL model is maintained, and we have access to an IV and a long panel dataset with earnings over the workers’ careers, we can identify the productivity effect of schooling without a hidden correlate of ability or other behavioral assumptions. This interpretation of ‘long-run’ IV estimates follows directly from the (limit) result in the EL model that earnings eventually converge to productivity. Instrumental variables that identify the causal effect of schooling on earnings thus also identify the causal effect of schooling on productivity.

Our second result is that what ‘short-run’ IV estimates identify depends on what employers know about the IV at hand. To that end, we distinguish between *hidden* and *transparent*

---

<sup>3</sup>Should we have access to such a correlate, then we can obtain over-identifying restrictions that can possibly be used to relax the assumptions of the learning model.

<sup>4</sup>Specifically, with heterogeneous returns and an IV that satisfies the assumptions of conditional independence, relevance and monotonicity, any ‘long-run’ IV estimate of returns to education is a consistent estimate of the effects of schooling on productivity for a subpopulation of individuals, i.e., the compliers, who are induced to take more education due to the instrumental variation (Imbens and Angrist, 1994).

instruments. We refer to an instrument variable as *hidden* if it is unobserved by the employers (e.g., eligibility rules for student aid or schooling laws that depend on unobserved personal characteristics) and as *transparent* if it is observed and priced into the wage equation (e.g., labor market conditions). We show that if the IV is *hidden* then the ‘short-run’ IV estimates identify *private returns to education* at each stage in the life-cycle. In comparison, if the IV is *transparent* then the ‘short-run’ IV estimates identify *productivity effect of education*.

For our third result, we propose a way to identify the speed of employer learning using IV estimates and link it to the parameters identified by the hidden instrument. In the EL framework, employers use past performances to update their belief about workers’ skills, so as work experience accumulates the private returns converge to the productivity effect of education, i.e., the social returns to education. We show that the estimates of the speed of learning and the ‘short-run’ and ‘long-run’ IV estimates of returns to education that are based on a hidden instrument are sufficient to *point-identify* the contributions of human capital and signaling in lifetime earnings.

To implement these ideas we use a unique dataset consisting of the population of Norwegian males born between 1950 and 1980, with earnings and employment histories between 1967 and 2014. We also observe an ability correlate – a cognitive test administered by the Norwegian military and taken by male conscripts around the age of 18, and a hidden instrument based on local variation in compulsory schooling laws across many birth cohorts.

First, we compare our OLS estimates to those in previous studies when we use an ability correlate similar to the AFQT score. The patterns uncovered in the Norwegian data are strikingly similar to those found by FG, AP, Lange (2007), and Arcidiacono et al. (2010) for the NSLY. In Norway, the estimated return to one standard deviation increase in the ability score increases from near zero in the first few years in the labor market before it stabilizes to about 7% after around 15 years of experience. The experience pattern in the NLSY with respect to the AFQT is similar, except that the return to a standard deviation in the AFQT score converges to approximately 14%. Controlling for the interaction between the ability

score and experience, we find that the coefficients on years of schooling decline rapidly from about 10% to 3% within the first 20 years. Likewise, in NLSY these coefficients decline from about 9% to around 6%.<sup>5</sup>

Second, we examine how the IV returns to schooling vary over the life-cycle, and interpret this through the lens of our EL model. The returns to schooling start high at 14% in the first year following graduation, and then decline, rapidly at first and then slowly, until they stabilize to 6%-7%, after approximately 20 years of experience. These findings are consistent with the learning hypothesis and the assumption that our instrument is a hidden instrument. We use these estimates to calculate the rate at which employers learn about workers' ability, and like in Lange (2007) our estimates also suggest employer learning is fast.

Finally, using our estimates of the speed of employer learning and experience-specific IV estimates of returns to education, we quantify the contribution of signaling. Our analysis reveal a productivity effect of education of 6.4% and a private internal rate of return in lifetime earnings, discounted to the time of schooling choice, of 8%. These estimates suggest that 80% of the total private returns to schooling over the life-cycle represents a productivity-enhancing effect of schooling while only the remaining 20% can be attributed to the signaling value of schooling. Thus, we argue that there is indeed a modest role of signaling in explaining the positive returns to education estimated in our data.

The rest of our paper proceeds as follows. Section 2 describes the model of employer learning as developed by FG, AP, and Lange (2007) and defines the private and the social returns to education within this structure. Section 3 then discusses identification of the private and the social returns to education in the model of employer learning using instrumental variables. Section 4 presents the data and our empirical setting. Section 5 presents our results. We discuss further extensions in Section 6 and conclude in Section 7.

---

<sup>5</sup>Similar to Arcidiacono et al. (2010), we also find that the association between ability test score and log-earnings increases with experience only for those with a high school degree or less. For those with more a college degree, the returns to the ability score remain constant at around 6-7% from a standard deviation increase in the ability score across all years of experience, consistent with a college degree revealing ability.

## 2 Model of Employer Learning

In this section we present the model of employer learning in a perfectly competitive labor market first proposed by FG and AP. Under perfect competition, workers are paid their expected output, where the expectation is taken with respect to the employers' belief about worker's productivity. Let worker  $i$ 's productivity be

$$\chi_{it} = \exp(\beta_{ws}S_i + \beta_{wq}Q_i + A_i + H(t) + \varepsilon_{it}) \equiv \exp(\psi_{it}), \quad (1)$$

where  $S$  is the years of schooling,  $Q$  is a correlate of ability observed by the employers, and  $A$  is unobserved (to employers and researchers) ability that is possibly correlated with the employer-observed correlates  $(S, Q)$ . The function  $H(t)$  captures how log-productivity varies with experience  $(t)$ . While  $H(t)$  can be a fully flexible function of  $t$ , we assume that it does not depend on schooling or ability. Finally,  $\varepsilon_t$  represents time-varying noise in the production process that is independent of all other variables.

To model employer learning we follow Lange (2007) and assume that  $\varepsilon_{it} \stackrel{i.i.d}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2)$  and  $(S_i, Q_i, A_i) \stackrel{i.i.d}{\sim} \mathcal{N}(\boldsymbol{\mu}, \Sigma)$  across workers and across time. Let  $\sigma_0^2 = \text{Var}(A_i | S_i, Q_i)$  be the conditional variance of  $A_i$  given  $(S_i, Q_i)$ . Besides knowing  $(S_i, Q_i)$ , every period employers also observe total output  $(\chi_{it})$ . Observing  $\chi_{it}$  is equivalent to observing a signal  $\xi_{it} \equiv A_i + \varepsilon_{it}$  about  $i$ 's productivity and employers use this signal to update their beliefs about the skills of workers as they age. If we let  $\mathcal{E}_{it}$  to be employers' information about  $i$  in period  $t$ , then  $\mathcal{E}_{it} := (S_i, Q_i, \xi_i^t)$  with  $\xi_i^t \equiv \{\xi_{i\tau}\}_{\tau < t}$  as the collection of all signals from all previous periods.<sup>6</sup>

It will be useful to introduce some additional notation. Let

$$A_i = \phi_{A|S}S_i + \phi_{A|Q}Q_i + \varepsilon_{A|S,Q}, \quad (2)$$

---

<sup>6</sup>We consider a model with symmetric information about workers' ability and past output across all employers. An important literature (Kahn, 2013; Schönberg, 2007; Waldman, 1984) analyzes how asymmetric information across the current employer and the other firms affects the labor market.

where  $\varepsilon_{A|S,Q} \equiv A_i - \mathbb{E}[A_i|S, Q]$ . The assumption  $(S_i, Q_i, A_i) \stackrel{i.i.d}{\sim} \mathcal{N}(\boldsymbol{\mu}, \Sigma)$  implies that the conditional expectation of  $A_i$  given the information at  $t = 0$ , i.e.,  $\mathbb{E}[A_i|\mathcal{E}_{i0}] = \mathbb{E}[A_i|S, Q]$ , is linear in  $S$  and  $Q$ . The wage in period  $t$  is equal to the expected productivity conditional on  $\mathcal{E}_{it}$ , so  $W_{it} = \mathbb{E}[\chi_{it}|\mathcal{E}_{it}] = \mathbb{E}[\chi_{it}|S_i, Q_i, \xi_i^t]$ . Taking the expectation of the log of (1) and using log normality of  $\exp(A_i + \varepsilon_{it})$  with conditional variance  $v_t := \text{Var}(A_i + \varepsilon_{it}|\mathcal{E}_{it})$ , we get

$$\ln W_{it} = \beta_{ws}S_i + \beta_{wq}Q_i + \tilde{H}(t) + \mathbb{E}[A_i|\mathcal{E}_{it}], \quad (3)$$

where  $\tilde{H}(t) \equiv H(t) + \frac{1}{2}v_t$  collects the terms that vary only with  $t$  but not across the realizations of  $\xi_i^t$ . For notational ease we set  $\tilde{H}(t) = 0$ .<sup>7</sup>

Under these assumptions we can represent the process by which employers update their expectations using the Kalman filter, which admits a particularly simple form

$$\mathbb{E}[A_i|\mathcal{E}_{it}] = \theta_t \mathbb{E}[A_i|S, Q] + (1 - \theta_t) \bar{\xi}_i^t, \quad (4)$$

where  $\bar{\xi}_i^t = \frac{1}{t} \{\sum_{\tau < t} \xi_{it}^t\}$  is the average of signals received in the market up to period  $t$  and the weight on the initial signal  $\theta_t = \frac{1-K_1}{1+(t-1)K_1}$  with  $K_1 = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_\varepsilon^2} \in [0, 1]$ . Equation (4) shows that the conditional expectation of ability at time  $t$  is equal to the weighted average of the expectation at  $t = 0$ , before any additional information about productivity has been received, and the average of all additional signals received up to  $t$ .

The weight  $(\theta_t)$  declines with experience  $(t)$  because observed realizations of productivity become more important in forming expectations compared to the correlates  $(S_i, Q_i)$  that are available at  $t = 0$ . The rate at which this weight decreases depends on the parameter  $K_1$  that Lange (2007) refers to as the “speed of learning.” It governs how quickly information about individual productivity accumulates in the market, which in turn depends on the information value of the signals. In particular,  $K_1$  is closer to 1 and the market learns faster about  $A$

---

<sup>7</sup>This notational simplification is without loss of generality because we can residualize log earnings on experience  $\tilde{H}(t)$ .



if the signal-to-noise ratio is high. Similarly, the more uncertain employers are about the priors relative to new information that arrives (i.e,  $\sigma_\varepsilon/\sigma_0^2$  is small), the larger is  $K_1$ .

## Social and Private Returns to Education

If education has a signaling value, then the private returns to education should exceed the social returns to education, conceived here as the effect of education on productivity. The coefficient  $\beta_{ws}$  in Equation (1) is not the causal effect of education on productivity, but only the “partial” causal effect of schooling *holding the employers-observed ability correlate  $Q$  and the unobserved ability  $A$  constant*. Schooling, however, might also causally affect  $(Q, A)$ , so the “total” causal effect of schooling on productivity also includes any (indirect) effect on productivity mediated through  $(Q, A)$ . We refer to this total causal effect as the *social returns to education*.

Next, we formalize these two measures of returns, and for that we need to introduce new notation and a simplifying assumption. For a random variable  $Y$ , let  $\delta^{Y|S}$  denote the causal effect of  $S$  on  $Y$  and let  $\tilde{Y}$  denote variation in  $Y$  that is not caused by schooling  $S$  but which may correlate with  $S$ . Let there be a linear causal relationship between  $S$  and  $(Q, A)$  so that

$$\begin{aligned} Q_i &= \delta^{Q|S} S_i + \tilde{Q}_i; \\ A_i &= \delta^{A|S} S_i + \tilde{A}_i. \end{aligned} \tag{5}$$

Substituting  $(S, Q, A)$  from the above equations into (1), we obtain

$$\begin{aligned} \psi_{it} &= \underbrace{(\beta_{ws} + \beta_{wq}\delta^{Q|S} + \delta^{A|S})}_{:=\delta^{\psi|S}} S_i + \underbrace{\beta_{wq}\tilde{Q}_i + \tilde{A}_i + \varepsilon_{it}}_{:=u_{it}}, \\ &= \delta^{\psi|S} \times S_i + u_{it}. \end{aligned} \tag{6}$$

The first term  $(\delta^{\psi|S})$  in (6) is the total causal effect of schooling on productivity, which is

also the social returns to education.<sup>8</sup> It consists of a direct causal effect and indirect effects on other ability components  $(Q, A)$ . In particular, an extra year of schooling increases  $Q$  by  $\delta^{Q|S}$  units, which in turn raises productivity by  $\beta_{wq}$ . An example of a  $Q$  could be students learning foreign languages that are mentioned in their résumé and can easily be verified. Schooling also raises ability  $(A)$  by  $\delta^{A|S}$  units that are of value in the labor market, but are unobserved by the employer.

Consider now the private returns to education, defined as the effect an additional year of schooling has on expected log earnings at time  $t$  when evaluated at the start  $t = 0$ . Schooling can affect expected log earnings at  $t$ : (i) directly, in that employers use schooling to form expectations about productivity; (ii) indirectly, because schooling causes changes in credentials  $Q$  observed by the firm (see Equation (5)); and (iii) because schooling affects productivity, which firms ultimately discover by observing workers output over time.

Substituting (2) and (4) in (3), and using  $\bar{\xi}_i^t = \frac{1}{t} \sum_{\tau < t} (A_i + \varepsilon_{i\tau}) = A_i + \bar{\varepsilon}_i^t$ , we get:

$$\ln W_{it} = (\beta_{ws} + \theta_t \phi_{A|S}) S_i + (\beta_{wq} + \theta_t \phi_{A|Q}) Q_i + (1 - \theta_t) (A_i + \bar{\varepsilon}_i^t). \quad (7)$$

Furthermore, using the causal relationships from (5) in the above equation, we obtain

$$\begin{aligned} \ln W_{it} &= (\beta_{ws} + \theta_t \phi_{A|S}) S_i + (\beta_{wq} + \theta_t \phi_{A|Q}) (\delta^{Q|S} S_i + \tilde{Q}_i) + (1 - \theta_t) (\delta^{A|S} S_i + \tilde{A}_i + \bar{\varepsilon}_i^t) \\ &= \underbrace{(\beta_{ws} + \beta_{wq} \delta^{Q|S} + \delta^{A|S} + \theta_t (\phi_{A|S} + \phi_{A|Q} \delta^{Q|S} - \delta^{A|S}))}_{:= \delta_t^{W|S}} S_i \\ &\quad + \underbrace{(\beta_{wq} + \theta_t \phi_{A|Q}) \tilde{Q}_i + (1 - \theta_t) (\tilde{A}_i + \bar{\varepsilon}_i^t)}_{:= \tilde{u}_{it}}. \end{aligned} \quad (8)$$

We thus arrive at the following wage equation:

$$\ln W_{it} = \delta_t^{W|S} S_i + \tilde{u}_{it}. \quad (9)$$

---

<sup>8</sup>As noted earlier (footnote 1), we abstract from other spillover effects of education.

Note that  $\delta_t^{W|S}$  in Equation (9) is the private return to education and given by the coefficient in front of S in Equation (8). Comparing this term and  $\delta_t^{W|S}$  in Equation (6) reveals that

$$\underbrace{\delta_t^{W|S}}_{\text{private returns}} = \underbrace{\delta^{\psi|S}}_{\text{social returns}} + \underbrace{\theta_t}_{\text{weight}} \underbrace{(\phi_{A|S} + \phi_{A|Q}\delta^{Q|S} - \delta^{A|S})}_{\text{adjustment term}}. \quad (10)$$

We can see that the private return  $\delta_t^{W|S}$  deviates from the social return  $\delta^{\psi|S}$  if the effect of schooling on expected  $A$ , based on information available to firms ( $\phi_{A|S} + \phi_{A|Q}\delta^{Q|S}$ ), differs from the causal effect of schooling on unobserved ability  $A$  ( $\delta^{A|S}$ ). The signaling literature assumes that this “adjustment term” is non-negative, so that  $\delta_t^{W|S} \geq \delta^{\psi|S}$  for all  $t$ . It also follows here that since  $\lim_{t \rightarrow \infty} \theta_t = 0$ , we must have  $\lim_{t \rightarrow \infty} \delta_t^{W|S} = \delta^{\psi|S}$ . Thus, as experience accumulates the private returns converge to the social returns to education.

In this Section, we have introduced the model of employer learning and used this model to define the social and the private returns to education. We now consider how one might identify these returns.

### 3 Identification

We now study how to identify the private and social returns to education in the above model. We begin by considering the simple case when only education, experience, and earnings are observed. In addition, we consider two cases with additional information: (i) a hidden correlate of ability is available (as in FG, AP and Lange (2007)); and (ii) an instrument for schooling is available, which might be observed by employers or not.

#### 3.1 Bias in the OLS

To begin, consider the coefficients obtained when regressing log earnings on years of education after controlling for a flexible specification in experience and fully interacting years of schooling with experience dummies. The resulting coefficient estimates for years of schooling

are consistent for

$$\text{plim}(b_{OLS,t}) = \delta^{\psi|S} + \frac{\text{cov}(\beta_{wq}\tilde{Q} + \tilde{A}, S)}{\text{var}(S)}, \quad (11)$$

Inspecting equation (11) reveals that the coefficient estimates do not vary with experience so that this specification collapses to the famous Mincer equations with log earnings profiles that are parallel in experience across years of schooling. More to the point, equation (11) illustrates the standard omitted variable bias in estimating the causal effect of schooling on earnings that arises when omitting ability correlated but not *caused* by schooling. This bias does not change over the life-cycle since the ability components  $(\tilde{Q}, \tilde{A})$  and schooling are constant in  $t$ .

### 3.2 Exploiting a Hidden Correlate of Ability

Now suppose that the researcher has access to a correlate of ability,  $Z_i$  that is not observed by employers.<sup>9</sup> Furthermore, suppose  $A_i = \beta_{Az}Z_i + \eta_i$ , then using it in (1) gives

$$\chi_{it} = \exp(\beta_{ws}S_i + \beta_{wq}Q_i + \beta_{Az}Z_i + \eta_i + H(t) + \varepsilon_{it}) = \exp(\psi_{it}). \quad (12)$$

Here,  $Q_i$  is observed by employers but not by researchers,  $Z_i$  is observed by researchers but not by employers, and  $\eta_i$  represents the productivity component that is neither observed by the researchers nor the employers. Following Lange (2007) we can show that

$$\mathbb{E}[\ln W_{it}|S, Z, t] = \theta_t \mathbb{E}[\ln W_{i0}|S_i, Z_i] + (1 - \theta_t) \mathbb{E}[\ln W_{i\infty}|S_i, Z_i], \quad (13)$$

where  $W_{i0}$  is the wage received in period  $t = 0$  and  $W_{i\infty}$  is the wage that would be received at  $t \rightarrow \infty$ , when enough information has been revealed so that productivity is observed in the market.

---

<sup>9</sup>Much of the literature on employer learning, including FG, AP, and Lange (2007) used the AFQT score reported by the NLSY for this purpose.

The conditional expectation in (13) is linear so we can estimate the weight ( $\theta_t$ ) and, moreover, the speed of learning ( $K_1$ ) by projecting log wages on  $(S, Z)$  across different experience levels  $t$ . The regression coefficients on  $(S, Z)$  converge from their  $t = 0$  value to the regression coefficients in  $\mathbb{E}[\ln W_{i\infty}|S_i, Z_i]$ , at a rate that depends on  $K_1$ .

The projection coefficients obtained from estimating of (13) across experience do however not identify the causal effect of  $S$  or  $Z$  on productivity. These coefficients are biased, even when  $t \rightarrow \infty$ , because  $(S, Z)$  correlate with the omitted variables  $(Q, \eta)$ . That is, schooling and  $Z$  correlate with information on individual productivity that is not observed by the researcher but is potentially observed by employers.

Thus, as shown in Lange (2007), while we can identify the Speed of Learning  $K_1$  assuming that we have a hidden correlate of ability, we can't identify the signaling value of schooling or its social return directly using this assumption. Rather, we require additional, strong assumptions on the schooling decision to identify or bound the contribution of signaling to the returns to education.

### 3.3 Instrumental Variable

Next, we consider the case when we have access to an instrumental variable. For the sake of argument, we consider a binary instrument  $D_i \in \{0, 1\}$ . Assume that it satisfies the standard assumptions to qualify as a valid instrumental variable:

**Assumption 1.** *Instrumental Variables*

1. (Conditional Independence):  $u_{it} \perp D_i | S_i$ , where  $u_{it}$  is defined in (6).
2. (First Stage):  $\mathbb{E}[S_i | D_i = 0] \neq \mathbb{E}[S_i | D_i = 1]$ .
3. (Monotonicity):  $S_i(D_i = 1) \geq S_i(D_i = 0)$  for all  $i$ .

Under Assumption 1, from Imbens and Angrist (1994) we know that in period  $t$

$$\text{plim } \hat{b}_{IV,t} = \frac{\mathbb{E}[\ln W_{it} | D_i = 1, t] - \mathbb{E}[\ln W_{it} | D_i = 0, t]}{\mathbb{E}[S_i | D_i = 1, t] - \mathbb{E}[S_i | D_i = 0, t]}. \quad (14)$$

Further, using (i) the invariance of  $S$  with respect to  $t$  and (ii)  $\lim_{t \rightarrow \infty} \ln W_{it} = \psi_i$ , we get

$$\text{plim} \left( \lim_{t \rightarrow \infty} \hat{b}_{IV,t} \right) = \text{plim} \hat{b}_{IV,t \rightarrow \infty} = \frac{\mathbb{E}[\psi_i | D_i = 1] - \mathbb{E}[\psi_i | D_i = 0]}{\mathbb{E}[S_i | D_i = 1] - \mathbb{E}[S_i | D_i = 0]} = \delta^{\psi|S},$$

where the second equality follows from the fact that the part of  $\psi_i$  that is not caused by  $S_i$  is orthogonal to  $D_i$ , i.e., for  $\psi_i = \delta^{\psi|S} S_i + \tilde{\psi}_i$ , Assumption 1 implies that  $\tilde{\psi}_i \perp D_i | S_i$ . Therefore, as  $t \rightarrow \infty$ , the instrument identifies the causal effect of schooling on productivity.<sup>10</sup>

Note that the identification argument is valid *irrespective* of what employers know about the instrument. This is because employers ultimately base compensation on signals observed over the career of the worker only. Thus, we can identify the social returns to education from the instrument directly by focusing on the returns to education in the long run. For  $t < \infty$ , however, what IV identifies under Assumption 1 depends on whether the employers know about  $D$ . To show how the information of employers affects the interpretation of the IV estimates we distinguish between *hidden* and *transparent* instruments next.

## Hidden Instrument

We begin with the case when employers do not observe the instrument  $D$ . We refer to such instruments as *hidden* instruments.

**Assumption 2.** (*Hidden Instrument*)  $D_i \notin \mathcal{E}_{it}$  which implies  $\ln W_{it} \perp D_i | (S_i, Q_i, \xi_i^t)$  for all  $i$ .

Note that Assumption 2 is conceptually different from Assumption 1-1. Assumption 1-1 asserts that the instrument is conditionally independent of determinants of productivity not caused by schooling. Assumption 2 captures the idea that given information  $\mathcal{E}_{it}$  available to the employer about  $i$ , this worker's wage does not depend on the instrument. When the instrument is hidden and Assumption 2 holds, we get  $\ln W_{it} = \mathbb{E}[\psi_i | \mathcal{E}_{it}] = \mathbb{E}[\psi_i | \mathcal{E}_{it}, D_i]$ .

In many settings, assumption 2 is a natural assumption to impose. The clearest examples

---

<sup>10</sup>With heterogeneous returns, this result says that IV estimates of returns to education at sufficiently high levels of experience provide consistent estimates of the effects of schooling on productivity for a subpopulation of individuals, i.e., the compliers, who are induced to take more education due to the instrumental variation.

relate to field experiments providing subsidies or information that induces higher school enrollment. In these cases, whether a student is in the control or treatment group is generally not known to potential employers. Examples from quasi-experimental settings in the empirical literature may include, but are not limited to, (i) the interaction of draft lottery number and year of birth in Angrist and Krueger (1992); (ii) the interaction of a policy intervention, family background and season of birth in Pons and Gonzalo (2002); (iii) parents' education and number of siblings in Taber (2001); and (iv) the elimination of student aid programs interacted with an indicator for a deceased father in Dynarski (2003). Many studies also exploit interactions of birth year and location of birth with locally implemented policy reforms (Duflo (2001), Meghir and Palme (1999)) and we do the same in our empirical application.

So, what do IV estimates of returns to education based on hidden instruments identify? Denote by  $\Delta_D$  the difference across  $D = 1$  and  $D = 0$ . Then the numerator in  $\text{plim} \left( \hat{b}_{IV,t} \right)$  for a binary, hidden instrument  $D_i$  (equation 14) is:

$$\Delta_D \mathbb{E} [\ln W_{it} | D_i, t] = \Delta_D \mathbb{E} [\beta_{ws} S_i + \beta_{wQ} Q_i + \mathbb{E} [A_i | S_i, Q_i, \xi_i^t] | D_i, t],$$

where  $\ln W_{it}$  does not directly depend on  $D_i$  because it is not used by the employers in the wage setting. The instrument  $D_i$  affects  $\ln W_{it}$  only indirectly by affecting  $(S_i, Q_i, \xi_i^t)$  that make up the information  $\mathcal{E}_{it}$  used by employers to infer productivity. From Assumption 1 we get  $\mathbb{E} [Q_i | D_i, S_i] = \delta^{Q|S} S_i$ , and using that with Equations (2) and (5) we get

$$\begin{aligned} \mathbb{E} [\ln W_{it} | D_i, t] &= (\beta_{ws} + \beta_{wQ} \delta^{Q|S}) \mathbb{E} [S_i | D_i] + \mathbb{E} [\mathbb{E} [A_i | S_i, Q_i, \xi_i^t] | D_i] \\ &= (\beta_{ws} + \beta_{wQ} \delta^{Q|S}) \mathbb{E} [S_i | D_i] + \mathbb{E} [\theta_t \mathbb{E} [A_i | S_i, Q_i] + (1 - \theta_t) \bar{\xi}_i^t | D_i] \\ &= (\beta_{ws} + \beta_{wQ} \delta^{Q|S}) \mathbb{E} [S_i | D_i] + \theta_t \mathbb{E} [\phi_{A|S} S_i + \phi_{A|Q} Q_i | D_i] + (1 - \theta_t) \mathbb{E} [A_i | D_i] \\ &= (\beta_{ws} + \beta_{wQ} \delta^{Q|S}) \mathbb{E} [S_i | D_i] + \theta_t \mathbb{E} \left[ \phi_{A|S} S_i + \phi_{A|Q} \left( \delta^{Q|S} S_i + \tilde{Q}_i \right) | D_i \right] \\ &\quad + (1 - \theta_t) \mathbb{E} [\delta^{A|S} S_i + \tilde{A}_i | D_i]. \end{aligned}$$

Simplifying this equation further, we get

$$\begin{aligned}\mathbb{E}[\ln W_{it}|D_i, t] &= ((\beta_{ws} + \beta_{wQ}\delta^{Q|S}) + \theta_t(\phi_{A|S} + \phi_{A|Q}\delta^{Q|S}) + (1 - \theta_t)\delta^{A|S}) \mathbb{E}[S_i|D_i], \\ &= (\delta^{\psi|S} + \theta_t(\phi_{A|S} + \phi_{A|Q}\delta^{Q|S} - \delta^{A|S})) \mathbb{E}[S_i|D_i],\end{aligned}$$

and

$$\text{plim } \hat{b}_{IV,t} = \frac{\Delta \mathbb{E}[\ln W_{it}|D_i]}{\Delta \mathbb{E}[S_i|D_i]} = \delta^{\psi|S} + \theta_t(\phi_{A|S} + \phi_{A|Q}\delta^{Q|S} - \delta^{A|S}). \quad (15)$$

Comparing this expression with the private returns defined in (10) we observe that the hidden IV identifies private returns to education at  $t$ . In addition, we obtain  $\text{plim}(\hat{b}_{IV,t})$  across  $t$  and can thus use the convergence of  $b_{IV,t}$  from  $b_{IV,t=0}$  to  $b_{IV,t \rightarrow \infty}$  over the life-cycle to identify the speed of learning  $K_1$ .

In summary, hidden instruments identify the private returns to schooling at each  $t$  as well as the parameter  $K_1$  governing the learning process.

### Transparent Instrument

We define an instrument as transparent if it is observed by the employers and as a consequence enters the wage setting directly. When the instrument  $D_i$  is transparent, it is included in the information set of the employers. To accommodate that the new information we have to expand the information set. Let  $\tilde{\mathcal{E}}$  denote the new information set.

**Assumption 3.** (*Transparent Instrument*) For all  $i$ ,  $\tilde{\mathcal{E}}_{it} = \mathcal{E}_{it} \cup \{D_i\}$  and  $\ln W_{it} \not\perp D_i$  and  $\ln W_{it} \perp D_i | (S_i, Q_i, \xi_i^t)$ .

This is to say that when employers determine wages they take into account the fact that as  $D_i$  varies exogenously, the years of schooling  $S_i$  changes too. So while unlike with the hidden instrument,  $\ln W_{it} = \mathbb{E}[\psi_i|\tilde{\mathcal{E}}_{it}] = \mathbb{E}[\psi_i|\mathcal{E}_{it}, D_i] \neq \mathbb{E}[\psi_i|\mathcal{E}_{it}]$ ,  $D_i$  still satisfies the exclusion restriction because  $D_i$  does not directly affect the earning. Examples of such



instruments employed in the empirical literature may include (i) tuitions at two and four years state colleges in Kane and Rouse (1995); (ii) a dummy for being a male aged 19-22 from Ontario in Lemieux and Card (2001); (iii) local labor market conditions in Cameron and Heckman (1998); Cameron and Taber (2004) and Carneiro et al. (2011); (iv) change in minimum school-leaving age in the United Kingdom from 14 to 15 in Oreopoulos (2006); and, perhaps also, (v) the distance to the college in Card (1993), Kane and Rouse (1995), Kling (2001) and Cameron and Taber (2004).

In order to characterize the learning process with a transparent instrument, we need to make an additional assumption about the conditional distribution of productivity  $\psi_i$  given  $(S_i, D_i)$ . To that end, we assume that  $\psi_i|S_i, D_i$  is normally distributed and with a homoscedastic variance, i.e., the variance of  $\psi_i$  is unaffected by  $D$ . In this environment, employers (on average) price the labor correctly given  $D_i$ , thus:

$$\begin{aligned}\mathbb{E}[\ln W_{it}|S_i, D_i] &= \mathbb{E}\left[\delta^{\psi|S} S_i + \tilde{\psi}|S_i, D_i\right] = \delta^{\psi|S} \times S_i \\ \mathbb{E}[\ln W_{it}|D_i] &= \delta^{\psi|S} \mathbb{E}[S_i|D_i].\end{aligned}$$

The Wald estimator for a binary transparent instrument gives  $\text{plim } \hat{b}_{IV,t} = \delta^{\psi|S}$ . Therefore, when firms are informed about the instrument, the IV estimate of returns to education is a consistent estimate of the productivity effect of education on earnings, i.e., we can use a transparent instrument to identify the the social returns to education at all  $t$ .

## 4 Data and Empirical Setting

In this section, we first provide information about our data sources, sample construction and describe the key variables that we utilize in our analysis. We describe the Norwegian compulsory schooling reform that provides a source of exogenous variation in education to construct IV estimates of the returns to education in log-earnings at each year of experience. We discuss the empirical specifications motivated by discussion in Section 3.

## 4.1 Data Sources and Sample Construction

Our empirical analysis uses several registry databases maintained by Statistics Norway. These databases allow us to construct a rich longitudinal dataset containing records for *every* Norwegian male from 1967 to 2014. The variables available to us include individual demographic information (e.g., cohort of birth and childhood municipality of residence) and socio-economic data (e.g., years of schooling and annual earnings). Importantly, the dataset also includes unique personal identifiers which allow us to follow individual earnings across time and record individuals' labor market experiences. The personal identifiers also allow us to merge information from Statistics Norway's registry databases with data from the Norwegian Armed Forces that provide us with ability test scores for male conscripts.

The Norwegian earnings data have several advantages over those available in most other countries. First, there is no attrition from the original sample beyond natural attrition due to either death or out-migration. Second, our earnings data pertain to all individuals, and not only to jobs covered by social security. Third, we can construct long earnings histories that allow us estimate the returns to education at each year of labor market experience.

We restrict our sample to Norwegian males born between 1950 and 1980, including several cohorts with earnings observed over a wide range of labor market experience.<sup>11</sup> We restrict the sample to males because the ability test scores are not available for females. We further exclude immigrants as well as Norwegian males with missing information on years of schooling, childhood municipality of residence, ability test score or indicator for exposure to the compulsory schooling reform. Applying these restrictions we retain a sample consisting of 732,163 individuals, which is 80.6% of all Norwegian males born between 1950 and 1980.

Our primary outcome variable is the natural logarithm of pre-tax annual labor-earnings.<sup>12</sup> To avoid variation in earnings across labor market experience due to the intensity of part-time

---

<sup>11</sup>In our annual income panel data from 1967 to 2014, we can observe the oldest cohort (1950) between ages 17 and 64 and the youngest cohort (1980) up to age 34.

<sup>12</sup>Note that this income measure thus does not include income from self-employment, capital income or unconditional cash transfers such as social economic assistance, housing assistance, child allowance, etc.

work, we focus solely on full-time workers defined as having annual labor earnings (adjusted for wage inflation) above the substantial gainful activity (henceforth, SGA) threshold in the Norwegian Social Security System, which in 2015 levels amounted to 10,650 USD.<sup>13</sup> Restricting the sample to males with full-time work for three consecutive years, we retain 718,234 individuals (i.e., most males are recorded having a three-year full-time work spell at least once) and a panel data set comprising of 14,758,689 person-year observations (i.e., on average an individual is observed with full-time work for 20.5 years), which we label as the estimation sample utilized in this analysis. It should be noted however that this sample is unbalanced; we have earnings for 579,386 individuals in the first year of potential experience and for 191,100 individuals in the 30th year of potential experience.<sup>14</sup>

## 4.2 Measures of Schooling and Ability

The first key regressor of interest is the number of years of schooling corresponding to the highest level of completed education. This variable is taken from Statistics Norway’s Education Register and is based on the educational attainment reports submitted by educational establishments directly to Statistics Norway, thereby minimizing measurement error due to misreporting. Our second regressor of interest is the ability test score received from the Norwegian Armed Forces. In Norway, military service was compulsory for all able males in the birth cohorts we study. Before entering the service, each male conscript’s medical and psychological suitability were assessed. The majority of eligible Norwegian males took this examination around their 18<sup>th</sup> birthday. The ability measure is a composite unweighted mean from three speeded tests—arithmetics, word similarities, and figures.<sup>15</sup>

Figure 1 plots the average ability and the conditional probability density function of ability for each year of schooling between 7 and 21 years. This figure illustrates two striking

---

<sup>13</sup>The earnings data is top-coded only at the very high earnings levels, and less than 3% of observations have right-censored earnings in any given year.

<sup>14</sup>As a robustness exercise, we perform an analysis also for a more balanced sample of cohorts 1950-1960.

<sup>15</sup>The arithmetic test mirrors the test in the Wechsler Adult Intelligence Scale (WAIS), the word test is similar to the vocabulary test in WAIS, and the figures test is comparable to the Raven Progressive Matrix test. See Sundent et al. (2004) and Thrane (1977) for details.

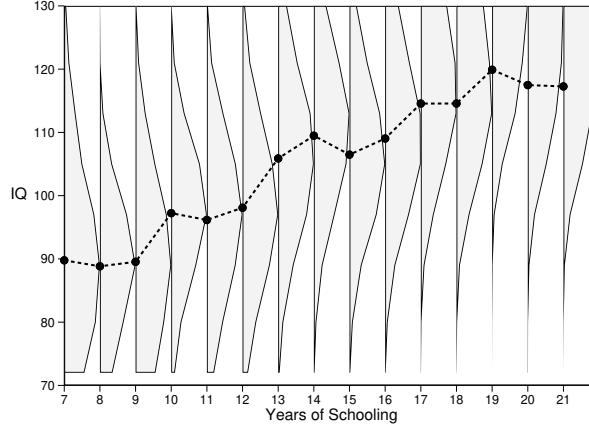


Figure 1: The Conditional Probability Density of Ability Test Scores on Years of Schooling.

*Note:* The sample consists of Norwegian males born 1950-1980 observed in earnings data over years 1967-2014 with years of potential experience between 0 and 30 years with annual earnings above 1 SGA threshold ( $N=14,758,689$ ). The ability test score (IQ) along the y-axis is here standardized to have a mean of 100 and a standard deviation of 15. The black dotted line plots the average ability for each year of schooling, while the shaded areas plot the conditional probability density of IQ.

patterns of our data that are worth noting. First, the measures of ability and schooling are strongly (positively) correlated, with a correlation of almost 0.5. Second, spikes in the average ability tend to occur around the entry years of high school (10 years), bachelor's degree (14 years), master's degree (17 years), and PhD (19 years). The latter pattern is perhaps surprising given that there are virtually no pecuniary cost of schooling (such as tuition or fees) in Norway. One possibility is that students with low non-pecuniary costs may find it easier to enroll in high school. This pattern is also consistent with there being selective entry requirements at various levels of higher education in Norway.<sup>16</sup>

Arguably, Norway is an interesting setting to assess employer learning and the signaling value of education for several reasons. First, the strong correlation between schooling and ability test scores in our data suggests that schooling has the potential to function as a predictor of ability, satisfying a necessary condition for schooling to have a signaling value. Second, employers can not request test scores from the military conscription boards tests nor is it common to (voluntarily) disclose these in job applications. We therefore believe

<sup>16</sup>As documented in Kirkeboen et al. (2016), there are limited slots available in higher education and access to fields with excess demand is based almost exclusively on the basis of merit through a centralized admissions process. Students with a higher GPAs from high school can thus more easily select into fields with excess demand. The same students tend to have achieved a higher ability test score in military conscriptions.

that it is reasonable to assume that the ability test scores from military conscriptions are unobserved by employers, such that a researcher can infer about the process of employer learning from the correlation between these test scores and earnings across experience. We do allow for the possibility that other correlates applicants' ability test scores are revealed in the application process and the model allows for the existence of such correlates (as captured by  $Q_i$  in Section 2). Finally, most cohorts in our sample would have entered the labor market before the arrival of online recruitment tools in the early 2000s, which could have altered the way in which employers tended to screen workers.

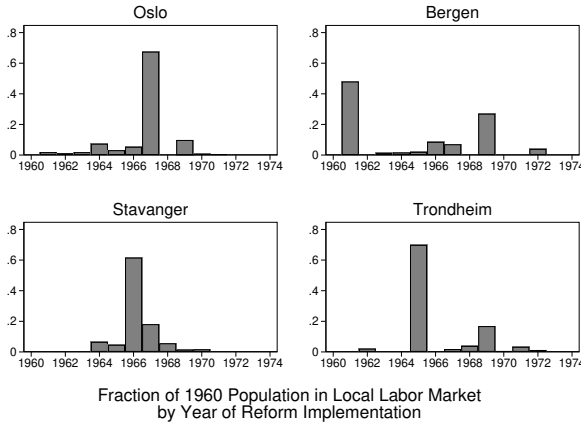
### 4.3 The Compulsory Schooling Reform

Between 1960 and 1975 Norway increased compulsory schooling from 7 to 9 years. This increase was implemented in different years in the different municipalities (the lowest level of local administration). Thus, for more than a decade, Norwegian schools were divided into two separate systems, where the length of compulsory schooling depended on the year in which an individual was born and the municipality of residence at age 14 (which we refer to as the childhood municipality). We use the timing differences across municipalities induced by the staggered implementation of the reform as our source of exogenous variation in educational attainment. For further details about the reform see Black et al. (2005).<sup>17</sup>

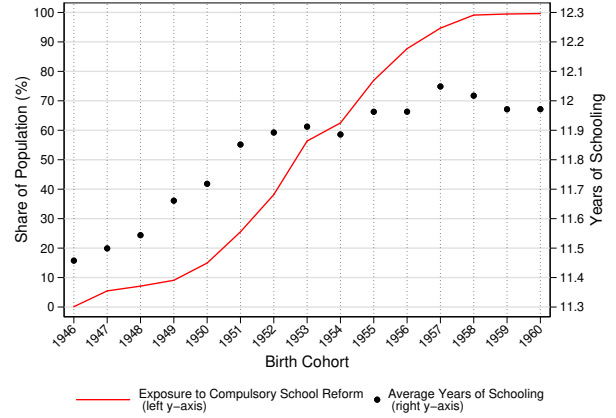
Historical records provide information about the year in which the reform was implemented for 672 out of the 732 municipalities that existed in 1960. For the remaining 60 municipalities this information is unknown; see Monstad et al. (2008). As shown in Figure 2, there is a considerable variation in the timing of compulsory schooling reform within local labor markets (panel (a) of the figure) and in the overall fraction of each birth cohort that was exposed to the reform (panel (b) of the figure). In particular, nobody born before 1946 was subjected to 9 years of compulsory schooling law, whereas all individuals born after 1960 were affected by the new law.

---

<sup>17</sup>This compulsory schooling reform in Norway has also been used previously, albeit in different contexts, by Monstad et al. (2008); Aakvik et al. (2010); Machin et al. (2012), and Bhuller et al. (2017).



(a) Timing Within Local Labor Markets



(b) Exposure Across Birth Cohorts

Figure 2: Compulsory School Reform Across Birth Cohorts and Local Labor Markets.

*Note:* Plot (a) shows the fraction of 1960 population in the four biggest local labor markets (concentrated around the four major cities) by the year of reform implementation. Using the 1960 classification of municipalities, there were 40 municipalities in the Oslo region, 27 municipalities in the Trondheim region, and 25 municipalities each in the Bergen and Stavanger regions. The variation in the timing of reform within local labor markets (LLMs) is due to variation in the timing of reform across municipalities within LLMs. The red line in plot (b) shows the cohort-specific share of population exposed to the compulsory school reform, while the black dots indicate the average years of schooling for Norwegian male cohorts born 1946-1960.

In particular, Figure 2-(a) shows that there is considerable variation within the four largest local labor markets (concentrated across the four biggest cities in Norway). For instance, the municipality of Oslo city, which accounted for 2/3 of the population in the Oslo labor market region in 1960, implemented the reform in 1967, whereas the timing of reform varied between 1961 and 1971 across the remaining 1/3 of population living in one of other 39 municipalities in this labor market.<sup>18</sup> In our analysis, we utilize both the overall variation in timing of reform across birth cohorts and municipalities, and the variation in timing of reform across birth cohorts and municipalities within local labor markets.

As discussed in Section 3, for the purpose of identifying both the private and the social returns to education, the instrument needs to satisfy not just the standard IV assumptions (Assumption 1), but also the hidden instrument assumption (Assumption 2). As implication of the latter assumption is that employers must be uninformed about the interaction between a worker's birth cohort and the timing of compulsory school reform in the worker's

<sup>18</sup>We utilize a classification of Norway in 160 local labor markets constructed by Gundersen and Juvkam (2013); on average each market comprises of 5 municipalities.

municipality of birth. There are two reasons why we think this assumption is reasonable in our setting.

First, in contrast to the compulsory schooling laws legislated elsewhere in a centralized manner (e.g., the U.S. states), the timing of the Norwegian compulsory school reform was decentralized and decided at the municipal level. This is consistent with our data, e.g., Figure 2-(a) shows substantial variation in the timing of reform also within local labor markets. Within the local labor markets, there are high rates of commuting and mobility. This means that to know whether or not an individual application was treated, a potential employer would have to not only know the exact date of implementation for each municipality but would also have to determine the municipality of each worker (or job applicant) when he was at age 14. While it might be easier to discern the place of residence and birth year, from the CV, say, determining the childhood municipality would be difficult and expensive if not impossible. Second, even if employers had information on each applicant's birth year and childhood municipality, the task of retrieving information on exposure to compulsory school reform for each applicant would still be onerous and costly. The information on the timing of compulsory reform was until recently not readily available in online public databases.<sup>19</sup> Therefore, it is reasonable to assume that for cohorts 1946-1960, graduating in an era long before internet, this information would not have been easily traceable for employers seeking to utilize such information in their wage setting decisions.

Even though we cannot directly test the hidden instrument assumption, we perform a robustness analysis in Section 5 in which we restrict our estimation sample to workers growing up in municipalities other than the municipality with the largest population in each local labor market. This will allow us to determine whether employers in a local labor market form their expectations based on the timing of reform in the largest municipality in that region. By relying on the sample of workers in each local labor market who did not grow up in the the largest population municipality, we think that our design makes it even more

---

<sup>19</sup>Previously, Black et al. (2005) and Monstad et al. (2008) have tracked various historical documents and databases to construct information on the timing of reform for 672 out of 732 municipalities.

plausible that the variation induced by our instrument is indeed not known to employers.

## 4.4 Empirical Specifications

Our first empirical specification projects log-earnings on schooling and control variables at each experience level:

$$y_{it} = \alpha_t^{(1)} + \beta_{s,t}^{(1)} s_i + \tau_t^{(1)} x_{it} + \omega_{i,t}^{(1)}, \quad (16)$$

where  $y_{it}$  represents log-earnings,  $s_i$  represents schooling,  $x_{it}$  represents all control variables and  $\left\{ \alpha_t^{(1)}, \beta_{s,t}^{(1)}, \tau_t^{(1)} \right\}_{t=0, \dots, T}$  are parameters of interest. The parameter  $\alpha_t^{(1)}$  varies freely with experience  $t$  and thus absorbs the experience profile  $H(t)$ .

Our primary focus is on  $\beta_{s,t}^{(1)}$ , which we estimate using the IV approach described below. Under Assumption 2 we can use these estimates to obtain the social returns to education for  $\beta_{s,\infty} = \beta_{s,T} = \lim_{t \rightarrow T} \beta_{s,t}^{(1)}$ , whereas the estimates for any experience  $t < T$  is the estimate of the private returns at that ( $t$ ) experience level. In addition, we can estimate the speed of learning by using the rate of convergence of  $\beta_{s,t}^{(1)}$  to its limit  $\beta_{s,\infty}$ .

We then obtain estimates of the speed of learning  $K_1$  obtained using our IV estimates with estimates obtained using the ability test score in our data as a hidden correlate. This latter approach requires projecting log earnings on schooling, ability ( $z$ ), and controls ( $x$ ) at different experience levels ( $t$ ):

$$y_{it} = \alpha_t^{(2)} + \beta_{s,t}^{(2)} s_i + \beta_{z,t}^{(2)} z_i + \tau_t^{(2)} x_{it} + \omega_{i,t}^{(2)}. \quad (17)$$

Under the assumptions that schooling does not independently enter  $H(t)$  and that  $z$  is unobserved in the market, we can use the OLS estimates of  $\left\{ \beta_{z,t}^{(2)}, \beta_{s,t}^{(2)} \right\}$  to obtain two independent estimates of the speed of learning. See Lange (2007) for further details.

Our IV model consists of the second-stage equation (16) and the first-stage equation:

$$s_i = \mu + \lambda d_i + \rho x_i + \kappa_i, \quad (18)$$



Table 1: First-Stage Estimates on Years of Schooling

Outcome Variable:	Years of Schooling
Instrument: <i>Exposure to Compulsory Schooling Reform</i>	0.266*** (0.026)
F-stat (instrument)	107.92
Sample Mean Years of Schooling	12.44
Standard Deviation Years of Schooling	2.58
Number of Observations	718,468

*Note:* The estimation sample consists of Norwegian males born 1950-1980 observed any time in earnings data over years 1967-2014 with years of potential experience between 0 and 30 years and annual earnings above 1 SGA threshold. All estimations include fixed effects for birth cohort and childhood municipality. Standard errors are clustered at the local labor market region. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

where the binary instrument  $d_i$  is equal to 1 if the individual was exposed to the reformed schooling law, and 0 otherwise. The reform exposure indicator is coded based on whether the reform had been implemented in  $i$ 's childhood municipality of residence by the time he had turned 14 and thus required to complete at least 9 years of schooling.<sup>20</sup> Here  $x_i$  includes a full set of dummy indicators for childhood municipality and birth cohorts.

We estimate the system of equations (18) and (16) by 2SLS, separately for each year of experience ( $t$ ).<sup>21</sup> We use the childhood municipality indicators to control unobservable determinants of earnings or schooling that are fixed at the municipality level, and use the birth cohort indicators to control the aggregate changes in schooling and earnings across cohorts (e.g., due to technical change). Notably, this specification allows  $H(t)$  to vary in an arbitrary manner across birth cohorts and childhood municipality for a given  $t$ .

Table 1 displays the first-stage estimates of years of schooling on our instrumental variable for the full estimation sample consisting of 718,468 individuals. The results indicate that the

<sup>20</sup>The school starting age in Norway at that time was 7 years and the pre-reform system required 7 years of compulsory schooling, which meant that the critical age at which a pupil would be required to take two additional years of schooling is 14 years. Cohorts aged 14 or below at the time of school reform would be required to take the two additional years, while all cohorts aged above 14 at the time the new law went into effect would not.

<sup>21</sup>Unlike in the second stage equation, there are no experience subscript on  $s$  and  $d$  in the first stage equation because these variables are time-invariant for a given individual. However, with a unbalanced panel, the first stage estimates of  $b$  will vary with experience because of changing cohort composition. In practice, the  $b$  estimates are fairly stable across the experience range that we consider in our analysis.

exposure to compulsory schooling reform increased completed years of schooling by 0.266, which is more than 10 % increase of a standard deviation. Notably, this instrument is strong with a partial F-statistic above 100, implying that weak instrument bias is not a concern for our analysis.<sup>22</sup>

A complication arises when we compare the OLS and IV estimates because earnings are not necessarily log-linear in schooling. It is well known that when the true model is non-linear, OLS and IV estimates of the linear specification identify different weighted averages of the true marginal effects of schooling on log earnings. See for example Angrist and Imbens (1995); Angrist and Krueger (1999); Heckman et al. (2006); Løken et al. (2012) and Mogstad and Wiswall (2016). It is, however, possible to make the OLS estimates comparable to the IV estimates by re-weighting the data. We call these estimates IV-weighted OLS and denote them by  $\beta_{s,t}^{(3)}$ :

$$\begin{aligned} y_{i,t} &= \alpha_t^{(3)} + \sum_{s=8}^{21} \gamma_{s,t}^{(3)} \times \mathbb{1}(S_i \geq s) + \beta_{z,t}^{(3)} \times z_{it} + \tau_t^{(3)} \times x_{it} + \omega_{it}^{(3)}; \\ \beta_{s,t}^{(3)} &= \sum_{s=8}^{21} \gamma_{s,t}^{(3)} \times \omega_s; \quad \omega_s = \frac{\text{cov}(\mathbb{1}(S_i \geq s), d)}{\text{cov}(s, d)}, \end{aligned} \tag{19}$$

where  $\mathbb{1}(S_i \geq s)$  is an indicator for at least  $s$  years of schooling, and  $d \in \{0, 1\}$  is the binary instrument equal to 1 if the compulsory schooling law has been implemented in individual  $i$ 's childhood municipality by the time the individual turned 14 years old. This specification thus takes the OLS estimates of the non-linear relationship between schooling and log-earnings and combines it in a way to mimic the variation exploited by the IV estimator. Intuitively, the IV estimates emphasize the marginal effects of schooling for those most affected by the instrument. Comparisons between the OLS and the IV estimates in the presence of non-linearities can therefore be misleading simply because the OLS and the IV estimates weight different parts of the support of schooling differently.

---

<sup>22</sup>Bhuller et al. (2017) show that the first-stage analysis is robust to including, pre-reform, linear and quadratic trends in schooling as additional controls, and that the timing of reform is uncorrelated with baseline municipality characteristics.

## 5 Results

In this section we present our estimation results. We begin by presenting the IV estimates and then compare those estimate with the findings relying on the standardized IQ score. Figure 3-(a) displays the IV estimates based on Equation (16) for two samples. These coefficients represent private returns to schooling. In Figure 3-(b) we further restrict the sample to workers growing up in municipalities other than the municipality representing the largest population in each local labor market.

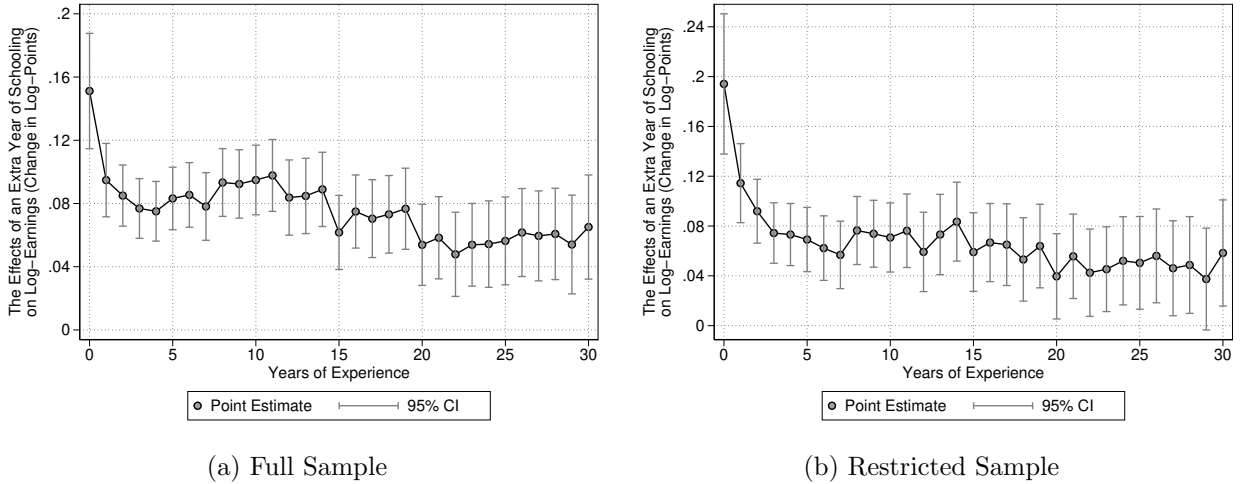


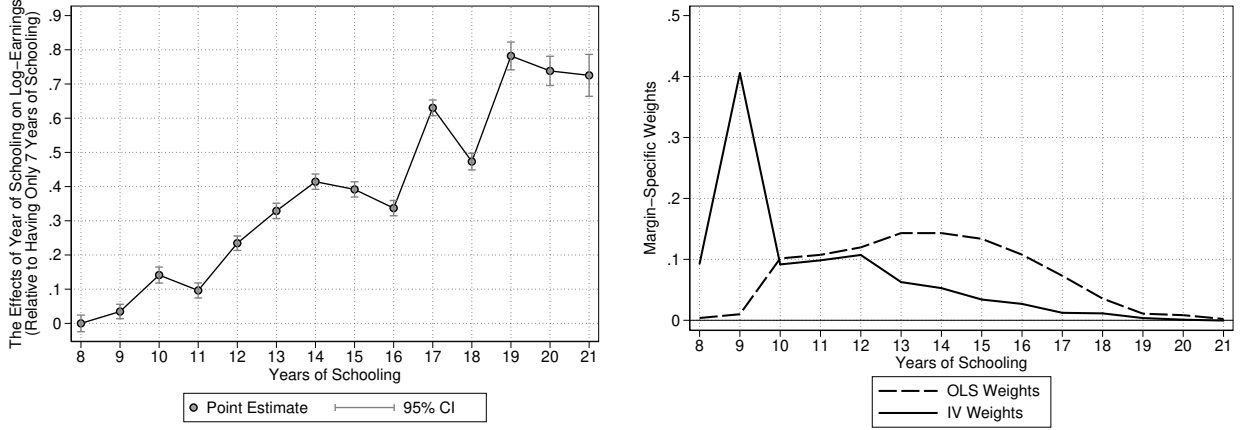
Figure 3: IV Estimates of the Returns to Schooling By Year of Experience.

*Note:* The full estimation sample consists of Norwegian males born 1950-1980 observed in earnings data over years 1967-2014 with years of potential experience between 0 and 30 years and annual earnings above 1 SGA threshold ( $N=14,758,689$ ). The restricted estimation sample further drops those growing up in the municipality with largest population size in each the 160 labor market regions in Norway ( $N=8,704,844$ ).

Both panels exhibit a relatively steep decline in the return to schooling early on, stabilizing after a few years. These estimates are consistent with employers learning about workers' ability and that employers are not fully pricing in that the variation in schooling is induced by the local variation in compulsory attendance laws across cohorts.

Next, we compare these IV estimates with the OLS estimates of Equation (17) for each year of experience  $t$  using the standardized IQ score at our measure of cognitive ability and controlling for municipal and cohort fixed effects. As argued above, the relationship between earnings and schooling is not log-linear. To compare OLS with the IV estimates we use the

specification in Equation (19). Figure 4 shows in panel (a) the estimates  $\gamma_{s,t}$  for experience  $t = 20$  and in panel (b) the weights  $\omega_s$  used to obtain  $\beta_{s,t}^{(3)} = \sum_s \gamma_{s,t} \omega_s$ .



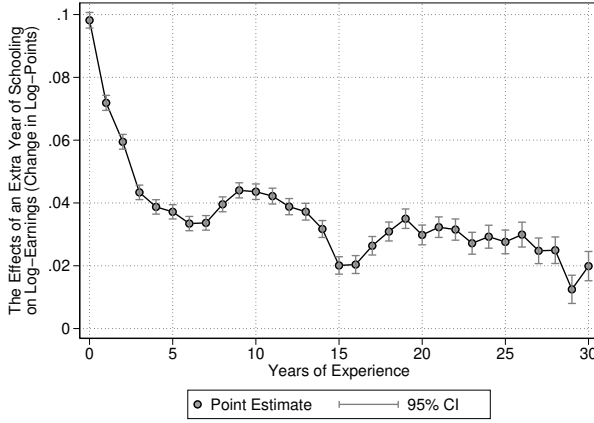
(a) Non-Linear Returns to Schooling

(b) Margin-Specific OLS and IV Weights

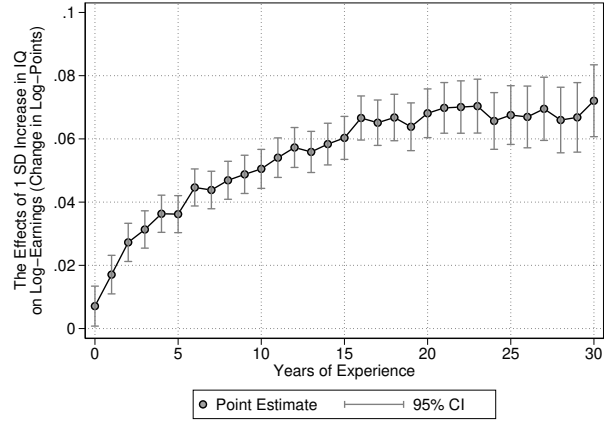
Figure 4: Non-Linear Returns to Schooling and Margin-Specific OLS and IV Weights.

*Note:* Panel (a) plots OLS estimates of returns to schooling from a specification with dummies for each year of schooling, controlling for cohort fixed effects, flexible time trends and childhood municipality fixed effects. The estimation sample consists of Norwegian males born 1950-1980 observed in earnings data over years 1967-2014 with years of potential experience between 10 and 20 years and annual earnings above 1 SGA threshold. The estimates show the returns to each year of schooling relative to having only 7 years of compulsory schooling. Panel (b) plots the margin-specific OLS and IV weights at each year of schooling.

Figure 5 displays the estimates of  $\beta_{s,t}^{(3)}$  and  $\beta_{z,t}^{(3)}$ . Despite employing different variation, the experience patterns in the returns to schooling (Figures 3 and 5) are consistent with each other. Both  $\beta_{s,t}^{(1)}$  and  $\beta_{s,t}^{(3)}$  decline rapidly within the first few years and then stabilize over the remainder of the life-cycle. Furthermore, the estimates from  $\beta_{z,t}^{(3)}$  increase over the life-cycle, rapidly at first, and then stabilize.



(a) The Returns to Schooling



(b) The Returns to Ability

Figure 5: IV-Weighted OLS Estimates of the Returns to Schooling and Ability.

*Note:* The estimation sample consists of Norwegian males born 1950-1980 observed in earnings data over years 1967-2014 with years of potential experience between 0 and 30 years and annual earnings above 1 SGA threshold (N=14,758,689).

At this point, we note that the patterns in the returns to schooling and ability over the life-cycle in our data are surprisingly similarly to those found in the NLSY using the AFQT score. We believe that this empirical finding is noteworthy regardless of its interpretation - it suggests that similar underlying economic phenomena are at work. To be specific, the rapid decline in the returns to schooling and increase in the returns to the ability score early in the life-cycle and the convergence to long-run values after a few years are present in both data (see Figure 1 in Lange (2007)). Further, we find similar patterns in the variation of the interaction between the AFQT and experience across education distribution as reported by Arcidiacono et al. (2010). These authors report that the increase in the returns to the AFQT score with experience is absent for those with more than a high school degree. In Figure 6 we show the returns estimated using three groups differentiated by their education. Figure 6-(a) shows the returns to ability over the life-cycle among those with 7-9 years of schooling, those with 9 years of schooling up to receiving a high school degree (panel b of the figure), and those with more than a high school degree (panel c of the figure).

We find that the interaction between experience and the cognitive test score is driven

entirely by those with less than a high school degree, consistent with Arcidiacono et al. (2010). Like Arcidiacono et al. (2010), we conjecture that employers are better informed about differences in cognitive ability among those with post-secondary education, possibly because they observe the grades, choice of field of study, reference letters, and because individuals get internships that reveals information about their ability.

Returning to the IV and OLS estimates from Figures 3 and 5, we find through visual inspection that these estimates are consistent with each other. This variation in the coefficients on schooling and the cognitive test score is driven by employer learning about ability differences across individuals with relatively little education.

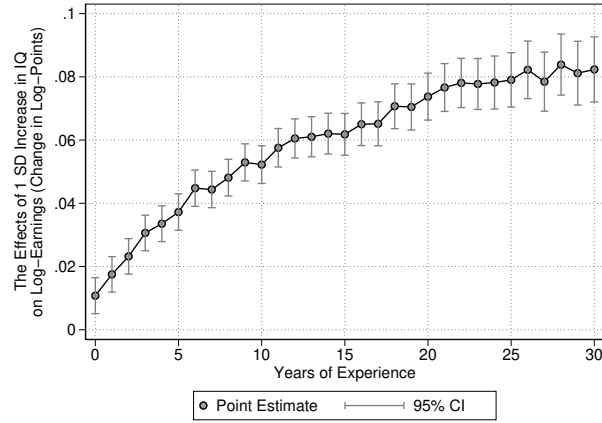
Next, we estimate how rapidly this learning process unfolds using three different sets of estimates: either of the estimates  $\{\beta_{s,t}^{(3)}, \beta_{z,t}^{(3)} : t = 1, \dots, T\}$  from the OLS specification or the IV estimate of  $\{\beta_{s,t}^{(1)} : t = 1, \dots, T\}$ . From Equation (13), we know that the coefficients  $\{\beta_{s,t}^{(3)}, \beta_{z,t}^{(3)} : t = 1, \dots, T\}$  are given by

$$\begin{aligned}\beta_{s,t}^{(3)} &= \theta_t b_{s,0}^{(3)} + (1 - \theta_t) b_{s,\infty}^{(3)}, \\ \beta_{z,t}^{(3)} &= \theta_t b_{z,0}^{(3)} + (1 - \theta_t) b_{z,\infty}^{(3)},\end{aligned}$$

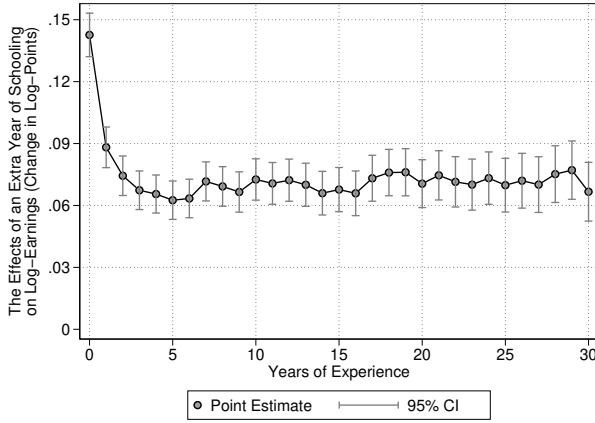
where  $\{b_{s,0}^{(3)}, b_{z,0}^{(3)}\}$  are the projection coefficients of log-earnings on schooling and ability at the onset of the individuals career and  $\{b_{s,\infty}^{(3)}, b_{z,\infty}^{(3)}\}$  are the projection coefficients that would be observed once productivity of individuals was fully revealed in the market. The parameters  $\{b_{s,0}^{(3)}, b_{z,0}^{(3)}, b_{s,\infty}^{(3)}, b_{z,\infty}^{(3)}\}$  themselves do not have an easy interpretation in term of the social or private returns to education. The IV estimates of  $\beta_{s,t}^{(1)}$  can also be expressed as

$$\beta_{s,t}^{(1)} = \theta_t b_{s,0}^{(1)} + (1 - \theta_t) b_{s,\infty}^{(1)},$$

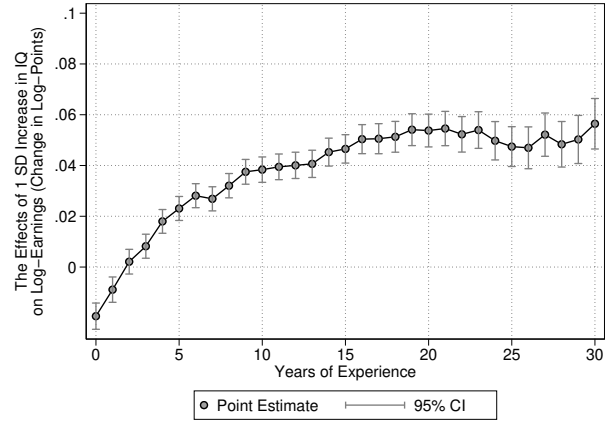
where the weights  $\{\theta_t\}$  are the same as before. The parameters  $\{b_{s,0}^{(1)}, b_{s,\infty}^{(1)}\}$ , however, have



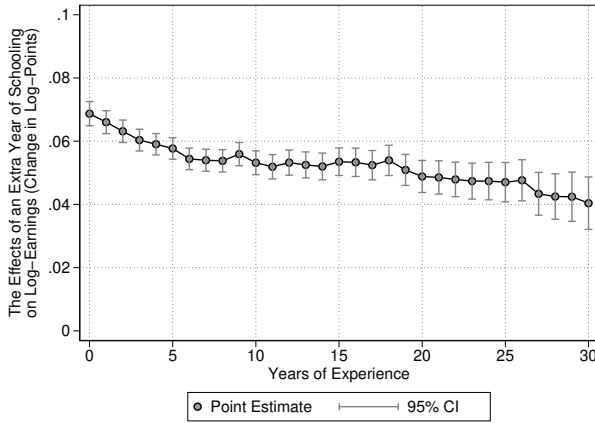
(a) The Returns to Ability: Compulsory School



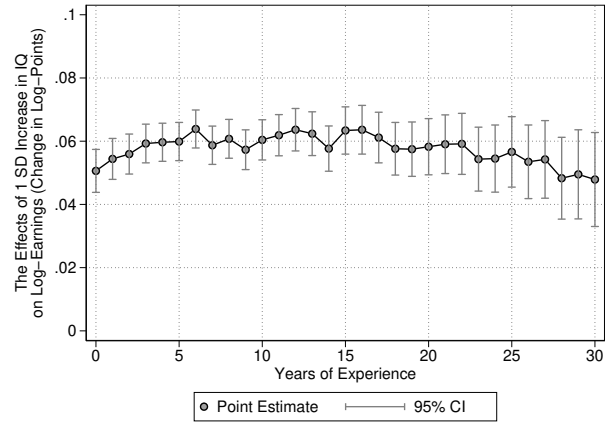
(b) The Returns to Schooling: High School



(c) The Returns to Ability: High School



(d) The Returns to Schooling: College/University



(e) The Returns to Ability: College/University

Figure 6: OLS Returns to Schooling and Ability and Ability by Level of Education.

*Note:* The estimation sample consists of Norwegian males born 1950-1980 observed in earnings data over years 1967-2014 with years of potential experience between 0 and 30 years and annual earnings above 1 SGA threshold ( $N=14,758,689$ ).

Table 2: Estimates of the Speed of Employer Learning, Initial Value and Limit Value.

	IV-Weighted OLS Estimates				IV Estimates	
	Two values of $K_1$		One value of $K_1$		Full Sample	Restricted Sample
	(1) Years of Schooling	(2) Ability Test Score	(3) Years of Schooling	(4) Ability Test Score	(5) Years of Schooling	(6) Years of Schooling
Speed of Learning $K_1$	0.377*** (0.046)	0.116*** (0.023)	0.207*** (0.029)		0.464*** (0.126)	0.543*** (0.058)
Initial Value $b_0$	0.098*** (0.005)	0.008** (0.004)	0.084*** (0.004)	-0.001 (0.004)	0.146*** (0.013)	0.193*** (0.010)
Limit Value $b_\infty$	0.023*** (0.002)	0.087*** (0.005)	0.017*** (0.002)	0.077*** (0.003)	0.064*** (0.004)	0.051*** (0.003)

*Note:* The full estimation sample consists of Norwegian males born 1950-1980 observed in earnings data over years 1967-2014 with years of potential experience between 0 and 30 years and annual earnings above 1 SGA threshold (N=14,758,689). The estimated experience-specific coefficients of schooling and IQ on log-earnings presented in Figure 5 are used to construct the IV-weighted OLS estimates of speed of learning, initial values of limit values in columns (1)-(4), while the estimates plotted in Figure 3(a) are used to construct the corresponding IV estimates of speed of learning, initial value of limit value in column (5). The IV estimates in column (6) are based the estimates plotted in Figure 3(b) for a restricted estimation sample in which the municipality with largest population size in each the 160 labor market regions in Norway is dropped (N=8,704,844).  
\*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

an economic interpretations. The parameter  $b_{s,0}^{(1)}$  represents the private return to education at  $t = 0$  and the parameter  $b_{s,\infty}^{(1)}$  the social returns. As shown above, the IV estimates of  $\beta_{s,t}^{(1)}$  represent the private returns at time  $t$ .

Table 2 displays the estimates of the parameters that enter these weighted averages for all three sets of estimates obtained by fitting the parameter estimates in Figures 3 and 5 using non-linear least squares. We observe that the estimates of  $K_1$  using schooling coefficients are similar across the OLS and the IV specifications, but they are different from the estimate of  $K_1$  that is obtained from using the standardize IQ score. Although the IV estimate of  $K_1$  is imprecisely measured, its 95% confidence interval does not overlap with the estimate using the IQ, suggesting that a simple learning model can not explain the full dynamics in the education and the IQ coefficients over life-cycle.

Given our assumptions, we can construct estimates of the social and private returns to



education based on the results reported in Table 2, column 3. As shown above, the limit  $b_\infty$  estimates the social returns to education, while the returns for fewer years of experience represent the private returns to education. Figure 7 displays the private and social returns based on the estimates from Table 2. The scatter plot displays the estimate obtained from the full sample (as in Figure 3) and the horizontal solid line represents the social returns.

To construct this figure we use the values  $b_0 = 0.146$  and  $b_\infty = 0.064$  from column 3 of Table 2. The latter is an estimate of the social returns to education and the former represents an estimate of the private returns to education prior to any learning. Given the large spread in estimates for  $K_1$ , we show the two estimates that bracket the value of  $K_1$  in Table 2. The value  $K_1 = 0.464$  which implies very rapid learning and the value  $K_1 = 0.116$  which implies much slower learning about individual ability. As expected from our model (see Equation 10), we can see that the speed of learning ( $K_1$ ) substantially affects the wedge between private returns and the social returns to education for medium and long experience levels.

Given these estimates, we can calculate the internal rate of return for an additional year of schooling. This equates the present discounted value of earnings over the career assuming a career length of 40 years. The internal rate of return associated with  $K_1 = 0.464$  and the lower broken curve in Figure 7 is 8.0% which is 1.6 percentage points higher than the social returns to schooling. The internal rate of return associated with slower learning ( $K_1 = 0.116$ ), and the upper broken line in Figure 7 is 10.7% which exceeds the social returns by 4.3 percentage points. Finally, using the IV estimates, by year of experience (the scatter) as our experience specific of the private returns to education, we obtain an internal rate of return of 8.1 %.<sup>23</sup>

We can also compare the estimate of the social returns to education of 6.4% with the Mincer returns to education - that is the coefficient on schooling in a non-interacted specification controlling for a flexible experience profile. This comparison provides an indication

---

<sup>23</sup>After experience 31, when increasing rates of individuals start retiring, we use 6.4%.

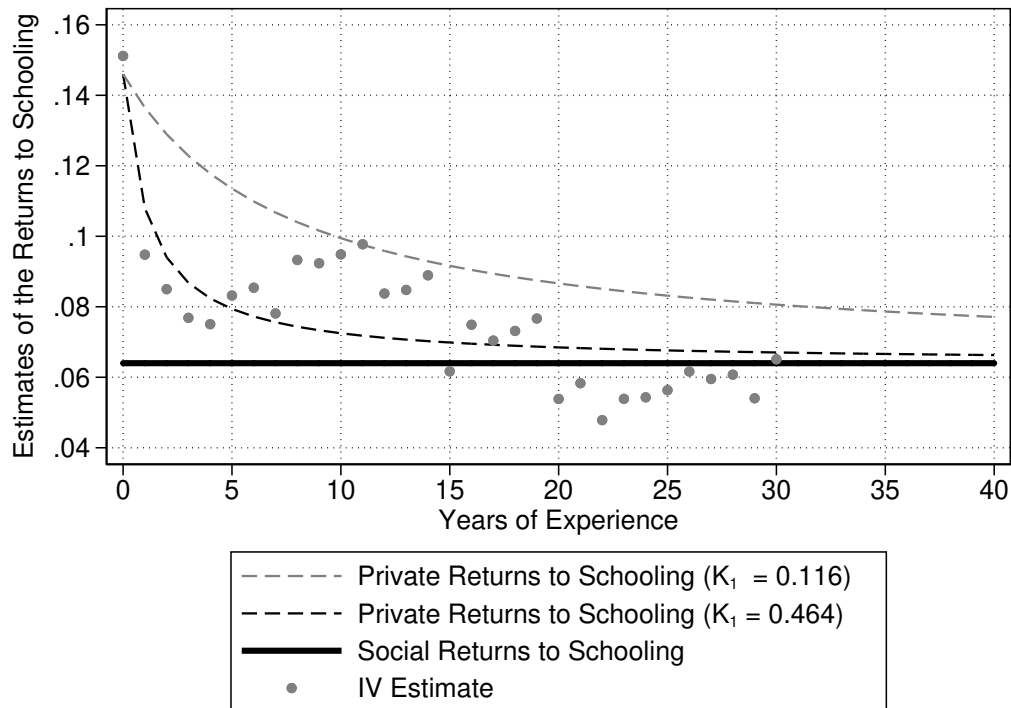


Figure 7: The Private and social returns to Schooling.

how much the social returns differ from the observed average differences in earnings in the population and is of interest since the Mincer coefficient is a very commonly used indicator of the value of education. We find that the Mincer coefficient is 6.8%, exceeding the social returns by about half a percentage point.

## 6 Extensions

## 7 Conclusion

In this paper we develop a model of employer learning with signaling that is based on FG, AP and Lange (2007), and use it to estimate the returns to education when an IV for schooling is available. We show that using the wages and schooling of workers with long work experience we can identify the causal effects of schooling on returns. On the other hand, for those with

shorter work experience what IV estimates identify depends on the information employers have about the IV at hand. In particular, we show that if the IV is hidden from the employers then the IV estimates the private returns to education, and if the IV is transparent to the employers then the IV estimates the social returns to education. In contrast to the short run returns, the long run returns are robust with respect to employers' information about the IV. To the best of our knowledge, these interpretations of IV estimates are novel.

Using data from Norway we flexibly estimate that the causal effect of schooling on productivity is 6.4%, which is also the social returns to education. We estimate the private returns to education to be 8%, where the difference is attributable to the role of signaling in education choice. In particular, we estimate that 80% of the total returns to education accrues to human-capital and the 20% to signaling. Our estimates also suggest that employers learn the unobserved skill very quickly.

There are many avenues for future research on related topics. First, like other studies of employer learning, we model labor market as a competitive market. This limits our ability to examine the role of labor market frictions, in particular asymmetric learning between current employer and potential employers. Relatedly a second interesting extension could be to allow for the possibility that the workers have career concerns (Holmström, 1999). This will allow us to study how workers' concern for future jobs affects effort choice on the job, which in turn will affect the rate of employer learning and the returns-to-job-experience. If the effort cost is inversely proportional to the ability, like in signaling (Spence, 1973), then the returns to experience will depend on both experience and schooling. Future research should seek to explore these questions in depth.

## References

- Aakvik, Arild, Kjell G. Salvanes, and Kjell Vaage. 2010. Measuring heterogeneity in the returns to education in norway using educational reforms. *European Economic Review* 54, no. 4:483–500.
- Altonji, Joseph G. and Charles R. Pierret. 2001. Employer learning and statisical discrimination. *The Quarterly Journal of Economics* 116, no. 1:313–350.
- Angrist, Joshua D. and Guido W. Imbens. 1995. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association* 90:431–442.
- Angrist, Joshua D. and Alan B. Krueger. 1992. Estimating the payoff to schooling using the vietnam-era draft lottery. *NBER working paper 4067* .
- . 1999. Empirical strategies in labor economics. In *Handbook of labor economics*, eds. Orley Ashenfelter and David Card. Amsterdam: North-Holland, 1277–1366.
- Arcidiacono, Peter, Patrick Bayer, and Aurel Hizmo. 2010. Beyond signaling and human capital: Education and the revelation of ability. *American Economic Journal: Applied Economics* 2, no. 4:76–104.
- Arteaga, Carolina. 2018. The effect of human capital on earnings: Evidence from a reform in columbia’s top university. *Journal of Public Economics* 157:212–225.
- Becker, Gary S. 1962. Investment in human capital: A theoretical analysis. *Journal of Political Economy* 70, no. 5:9–49.
- Bedard, Kelly. 2001. Human capital versus signaling models: University access and high school dropouts. *Journal of Political Economy* 109, no. 4:749–775.
- Bhuller, Manudeep, Magne Mogstad, and Kjell G. Salvanes. 2017. Life-cycle earnings, education premiums, and internal rate of return. *Journal of Labor Economics* 35, no. 4:993–1030.

- Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2005. Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *American Economic Review* 95, no. 1:437–449.
- Cameron, Stephen V. and James J. Heckman. 1998. Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of american makes. *Journal of Political Economy* 106, no. 2:262–333.
- Cameron, Stephen V. and Christopher R. Taber. 2004. Estimation of educational borrowing constraints using returns to schooling. *Journal of Political Economy* 112, no. 1:132–182.
- Card, David. 1993. Using geographic variation in college proximity to estimate the return to schooling. *NBER working paper 4483* .
- Carneiro, Pedro, James J. Heckman, and Edward J. Vytlacil. 2011. Estimating marginal returns to education. *American Economic Review* 101, no. 6:2754–2781.
- Clark, Damon and Paco Martorell. 2014. The signaling value of a high school diploma. *Journal of Political Economy* 122, no. 2:282–318.
- Duflo, Esther. 2001. Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American Economic Review* 91, no. 4:795–813.
- Dynarski, Susan M. 2003. Does aid matter? measuring the effect of student aid on college attendance and completion. *American Economic Review* 93, no. 1:279–288.
- Fang, Hanming. 2006. Disentangling the college wage premium: Estimating a model with endogenous education choices. *International Economic Review* 47, no. 4:1151–1185.
- Farber, Henry S. and Robert Gibbons. 1996. Learning and wage dynamics. *The Quarterly Journal of Economics* 111, no. 4:1007–1047.

- Feng, Andy and Georg Graetz. 2017. A question of degree: The effects of degree class on labor market outcomes. *Economics of Education Review* 61:140–161.
- Gundersen, Frantz and Dag Juvkam. 2013. Inndelinger i senterstruktur, sentralitet og barregioner. Tech. rep., NIBR - rapport 2013:1.
- Heckman, James J., Sergio Urzua, and Edward J. Vytlacil. 2006. Understanding instrumental variables in models with essential heterogeneity. *Review of Economic Studies* 88, no. 3:389–432.
- Holmström, Bengt. 1999. Managerial incentive problems: A dynamic perspective. *Review of Economic Studies* 66:169–182.
- Hopkins, Edward. 2012. Job market signaling of relative position, or becker married to spence. *Journal of the European Economic Association* 10, no. 2:290–322.
- Imbens, Guido W. and Joshua D. Angrist. 1994. Identification and estimation of local average treatment effects. *Econometrica* 62, no. 2:465–475.
- Kahn, Lisa B. 2013. Asymmetric information between employers. *American Economic Journal: Applied Economics* 5, no. 4:165–205.
- Kane, Thomas J. and Cecilia Elena Rouse. 1995. Labor-market returns to two- and four-year college. *American Economic Review* 85, no. 3:600–614.
- Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad. 2016. Field of study, earnings, and self-selection. *Quarterly Journal of Economics* 131, no. 3:1057–1111.
- Kling, Jeffrey R. 2001. Interpreting instrumental variables estimates of the returns to schooling. *Journal of Business and Economic Statistics* 19, no. 3:358–364.
- Lange, Fabian. 2007. The speed of employer learning. *Journal of Labor Economics* 25, no. 1:1–35.

- Lange, Fabian and Robert Topel. 2006. The social value of education and human capital. *Handbook of the Economics of Education* 1:459–509.
- Lemieux, Thomas and David Card. 2001. Education, earnings, and the ‘canadian g.i. bill’. *Canadian Journal of Economics* 34, no. 2:313–344.
- Lochner, Lance. 2011. Non-production benefits of education: Crime, healthm, and good citizenship. *NBER Working Paper* .
- Løken, Katrine V., Magne Mogstad, and Matthew Wiswall. 2012. What linear estimators miss: The effects of family income on child outcomes. *American Economic Journal: Applied Economics* 4, no. 2:1–35.
- Machin, Stephen, Kjell G. Salvanes, and Panu Pelkonen. 2012. Education and mobility. *Journal of the European Economic Association* 10, no. 2:417–450.
- Meghir, Costas and Mårten Palme. 1999. Assessing the effect of schooling on earnings using a social experiment. *Working Paper* .
- Mogstad, Magne and Matthew Wiswall. 2016. Testing the quantity–quality model of fertility: Estimation using unrestricted family size models. *Quantitative Economics* 7, no. 1:157–192.
- Monstad, Karin, Carol Propper, and Kjell G. Salvanes. 2008. Education and fertility: Evidence from a natural experiment. *The Scandinavian Journal of Economics* 110, no. 4:827–852.
- Oreopoulos, P. 2006. Estimating average and local average treatment effects of education when compulsory schooling laws really matter. *American Economic Review* 96:152–175.
- Oreopoulos, Philip and Kjell G. Salvanes. 2011. Priceless: The nonpecuniary benefits of schooling. *Journal of Economic Perspectives* 25, no. 1:159–184.

- Pastorino, Elena. 2018. The role of job assignment, learning, and human capital acquisition for careers in firms. *Mimeo* .
- Pons, Empar and Maria Teresa Gonzalo. 2002. Returns to schooling in Spain: How reliable are IV estimates? *Labour* 16, no. 4:747–770.
- Schönberg, Uta. 2007. Testing for asymmetric employer learning. *Journal of Labor Economics* 25, no. 4:651–691.
- Spence, Michael. 1973. Job market signaling. *The Quarterly Journal of Economics* 87, no. 3:355–374.
- Sundent, Jon Martin, Dag G. Barlaug, and Tore M. Torjussen. 2004. The end of the Flynn effect? A study of secular trends in mean intelligence test scores of Norwegian conscripts during half a century. *Intelligence* 32:349–362.
- Taber, Christopher R. 2001. The rising college premium in the eighties: Return to college or return to unobserved ability? *Review of Economic Studies* 68, no. 3:665–691.
- Thrane, Vidkun C. 1977. Evneprøving av utskrivningspliktige i Norge 1950-53. Tech. rep., Arbeidsrapport nr. 26, INAS.
- Tyler, John H., Richard J. Murnane, and John B. Willett. 2000. Estimating the labor market signaling value of the GED. *Quarterly Journal of Economics* 115, no. 2:431–468.
- Waldman, Michael. 1984. Job assignments, signalling, and efficiency. *The RAND Journal of Economics* 15, no. 2:255–267.