This is a post-peer-review, pre-copyedit version of an article published in Empirical Economics. The final authenticated version is available online at: http://dx.doi.org/10.1007/s00181-020-01837-z

Noname manuscript No.
(will be inserted by the editor)

Human Capital vs. Signaling is Empirically Unresolvable

Nick Huntington-Klein

Received: date / Accepted: date

Abstract Economists offer two main explanations for the causal labor market returns to education. The first is human capital accumulation: education improves ability. The second is signaling: education allows high-ability students to distinguish themselves. A major point of interest is the relative contributions of these effects. I demonstrate the theoretical and empirical conditions necessary to identify the relative contribution of the two models. Then, I review the existing literature to evaluate whether the feasible set of empirical estimates is capable of meeting those conditions and so informing theory. Empirical evidence is capable of rejecting pure human capital and signaling models, and usually does so. I argue that, for the general question of relative contribution, necessary identification conditions are not met, and partial identification bounds are wide. Two models with different non-zero contributions of human capital and signaling cannot be empirically distinguished, limiting the usefulness of human capital vs. signaling as a framing for understanding the return to education and for policy.

Keywords Human capital · signaling · identification · education · JEL I23 · I26 · J24

I thank Oded Gurantz, Tyler Ransom, Jacob Vigdor, anonymous reviewers and the associate editor, and seminar attendees at the University of California Santa Barbara for comments and suggestions.

Nick Huntington-Klein
800 N. State College Blvd., Fullerton, CA, 92831.
Tel.: +1657-278-3918
ORCID 0000-0002-7352-3991
E-mail: nhuntington-klein@fullerton.edu

1 Introduction

A central question in the economics of education is the effect of education on earnings. A large part of the empirical work in the economics of education concerns estimates of the causal effect of education on earnings and the implications of that effect. Much of the theoretical work in the economics of education concerns explanations of why that causal effect exists and is positive.

The two prevailing explanations for the return are human capital and signaling. Human capital theory (Schultz, 1963; Becker, 1964) suggests that education has a positive causal effect on student ability, which in a competitive labor market translates into higher earnings. Those with education earn more because they learn.

Under signaling theory (Spence, 1973), education does not improve student ability. Instead, education is used to identify workers who already had high levels of ability.[1]

Human capital and signaling are not mutually exclusive. There are multiple empirical studies, many of them discussed in later sections, that convincingly show both that human capital explains a non-zero portion of the returns to education, and that signaling explains a non-zero portion of the returns to education.

However, showing that both effects are non-zero does not provide information on which of the explanations should be given primacy, or to what degree each should be given weight. The question of the relative importance of each theory is a part of the way that the question is presented publicly in the economics of education, both colloquially and in encyclopedia entries (Page, 2010; Gunderson and Oreopoulos, 2010), popular-press books (Caplan, 2018), and in textbooks (Mankiw 2014 Chapter 19; Lovenheim and Turner 2018 Section 5.5). In the academic literature, some studies calculate human capital and signaling shares in the context of their own data (Fang, 2006; Lange, 2007; Kaymak, 2012; Bingley et al., 2015; Eble and Hu, 2016; Aryal et al., 2019). Other studies review the literature to make a case that one explanation should be preferred, but they do not agree on which explanation it is (Layard and Psacharopoulos, 1974; Weiss, 1995; Lange and Topel, 2006; Caplan, 2018). The literature on human capital and signaling presents the question about the primacy of the human capital and signaling models as solvable, but unsettled because of the well-acknowledged fact that human capital and signaling effects are very difficult to distinguish from each other empirically.[2]

In this paper, I review the literature on the returns to education and make the case that the question of relative importance of human capital and signaling cannot be meaningfully addressed by evidence, and that

[1] Throughout the paper I use the term signaling to also refer to the screening hypothesis (Arrow, 1973; Stiglitz, 1975; Wolpin, 1977), which is similar but differs in timing and some implications (Stiglitz and Weiss, 1994).

[2] This is contrary to the quote from Lang and Kropp (1986), "In fact, many members of the profession maintain (at least privately) that these hypotheses cannot be tested against each other and that the debate must therefore be relegated to the realm of ideology." Their claim is stronger still, in that it implies that human capital and signaling effects cannot be identified at all, as opposed to being unable to identify the relative importance of either model.

the conception of these models within the field should shift. Evidence has been used to convincingly reject a model of education returns in which either human capital or signaling play *no* part. However, I claim that any model of education returns within those bounds, from a model that is almost entirely human capital to a model that is almost entirely signaling, is empirically indistinguishable from another model that assigns different weight to the two explanations. Identification is only partial, and the bounds of partial identification are too wide to be useful.

To make this case I provide two main contributions. First, I link theory to empirical observation by laying out the necessary identifying conditions for the question of relative importance. I do this by presenting human capital and signaling as both existing in empirical form as part of a returns-to-education model with mediating variables. Both explanations imply that education should improve earnings, and the distinction between them can be understood as emphasizing different mediating variables that explain *why* education improves earnings. In this framework, I show in Sections 2 and 3 the conditions necessary to identify the human capital or signaling shares of the return, or to narrow partial-identification bounds to a useful range. This process could be applied in future work to other questions to determine the conditions for identification of competing theoretical propositions.

The second contribution, in Section 4, is to thoroughly examine the empirical literature on human capital and signaling specifically, and the returns to education more broadly, to make the case that the necessary conditions for identification cannot be realistically met, and that partial-identification bounds are wide. This argument relies on the empirical content of the human capital and signaling models, and on multiple examples of debates where evidence considered to be in favor of one model or the other is actually indeterminate. These indeterminate cases cover nearly all of the empirical content of the two models that the literature has uncovered.

The conditions for identification tend to fail for three reasons: (Section 4.1) There are too few observable mediating variables that can be assigned to *only one* of human capital or signaling, (Section 4.2) both theories place heavy emphasis on unobservable mediating variables, which prevents falsification, and (Section 4.3) situations in which these concerns can be overcome, like in some quasiexperimental studies, generally cannot be used to address the question of relative importance.

In effect, the argument is: the task of estimating the human capital and signaling shares of the return to education requires that researchers estimate how a non-experimentally derived causal effect is mediated. This is in itself a difficult, although not impossible, statistical feat, one that is accomplished in the case of many academic questions concerning two competing theories. The specific case of human capital vs. signaling is different in that the mediating variables of interest are unmeasurable, both theories are too loosely defined in empirical terms for nearly any proxies to be believable, and, when identifiable, the local average treatment effect provides only a little information in answering the question.

I argue that while human capital and signaling are useful theoretical tools, and can be productively used to generate testable hypotheses, the actual testing of these hypotheses cannot usefully inform the theory. This prevents the theory from being practically applicable in prediction or policy. Human capital and signaling is then a subpar approach to understanding education returns in the real world. I suggest in Section 5 two alternatives: an atheoretical approach to understanding the returns to education, and a theoretical framework that places at its center the concepts of the private and external returns to education. Section 6 concludes.

2 A Mediating-Variables Model of the Returns to Education

In this section I provide a general model of the returns to education. Figure 1 shows a directed acyclic graph that describes the structural relationship between education and various outcomes of interest (Pearl, 2009; Morgan and Winship, 2014). The use of a diagram model allows me to focus on the relationship between theoretical structure and inference, encoding the same statistical assumptions as would be expressed in potential outcomes notation while highlighting economic theory rather than statistical assumptions. Appendix A demonstrates the same model in potential outcomes notation.

Variation in education is driven by both endogenous selection pressures (family background, ability) and exogenous selection pressures (compulsory education policy changes, experimental assignment). Education can be defined in the model at any given margin, such as "high school degree vs. bachelor's degree," "years of education," or even something that simply changes the nature of education rather than the amount, such as "was placed with a great teacher rather than an average one" or "was exposed to advanced pedagogical methods vs. business as usual."

The outcomes of interest can be measured at the individual level, such as earnings at a certain age or over a lifetime (Card, 1999), but also unemployment, occupation held, measured productivity, or a particular age-earnings profile. Non-labor outcomes like marital status, health, or happiness (Oreopoulos and Salvanes, 2011; Heckman et al., 2018) or committing crime (Machin et al., 2011) can also be considered. Individual outcomes then build to affect aggregate outcomes such as productivity, economic growth, and inequality (Goldin and Katz, 2009) or the market conditions and wages for labor markets of more-educated and less-educated workers (Bedard, 2001; Moretti, 2004).

Education does not affect outcomes of interest directly, but rather influences a set of mediating variables $x_1, ..., x_J$ that affect the outcomes of interest. The mediating variables are defined broadly enough so as to intercept any direct effect that education might have on the outcomes of interest. These include things like cognitive skills (Ritchie and Tucker-Drob, 2018), non-cognitive and social skills (West et al., 2016), job-specific skills (Van Der Velden and Bijlsma, 2016; Brunello and Rocco, 2017), exposure to peers of certain qualities (Sacerdote, 2001), cultural socialization (Rivera, 2016), knowledge of one's extant abilities (Stinebrickner and Stinebrickner, 2014; Arcidiacono et al., 2016), knowledge of the labor market (Botelho and Pinto, 2004), potential-employer beliefs about one's skills (Arcidiacono et al., 2010), or having a degree (Jaeger and Page, 1996; Belman and Heywood, 1997). Some of these mediating variables may have their own sources of exogenous variation z_j.[3]

These mediating variables are key to identifying the different explanations of the returns to education. With the exception of the selection explanation of educational premia (in which education is simply correlated with outcomes because both are determined by endogenous selection pressures), explanations of the returns to education assume that education has an effect on *something*, and then that something affects our outcomes of interest.

Each mediating variable x_j has a corresponding effect β_j on the Private Outcome. Each effect β_j can be further broken up based on theoretical understanding into four parts: a human capital portion κ_j, a signaling portion σ_j, and two "other" portions: ω_j, for non-signaling and non-human-capital portions of the returns to education, and ϵ_j for the portion of the effect of the mediator that is not part of the return to education. In total, $\beta_j = \kappa_j + \sigma_j + \omega_j + \epsilon_j \; \forall \; j$.[4]

Under this framework, the total causal effect of education is defined as $B \equiv \sum_j(\kappa_j + \sigma_j + \omega_j)$, the total signaling effect is defined as $\Sigma \equiv \sum_j \sigma_j$, and the total human capital effect is defined as $K \equiv \sum_j \kappa_j$. For the purposes of this paper I will assume that the causal effect of education B has been identified and so ignore ϵ_j for most discussion.

If a particular mediating variable x_j is strictly an example of human capital and not signaling, then $\sigma_j = \omega_j = 0$. Alternately, if it is strictly signaling and not human capital, then $\kappa_j = \omega_j = 0$. If it is a mixture of both, then $\kappa_j \neq 0$ and $\sigma_j \neq 0$.

The human capital model assumes that education improves individual and aggregate outcomes because it improves the broadly defined job-relevant skills of the student, and these skills are rewarded in the labor

[3] The model is general but is still by necessity a simplification, and there are several obvious variations. Depending on what is considered as an outcome, some outcomes may be considered mediating variables sometimes: for example, education may affect the occupation held, which affects earnings, but also individual productivity and thus returns through job match (Van Der Velden and Bijlsma, 2016) and aggregate productivity through production complementarities (Kremer, 1993). Mediating variables may also affect each other in some way, such as how having a degree can impact a potential employer's beliefs about a employee's skills. While not pictured, these complexities are generally understood and incorporated into the discussion of identification.

[4] Standard analysis of causal diagrams would instead specify the effect of education via a mediator as the product of the effect of education on the mediator, and the effect of the mediator on the outcome. My approach is equivalent to the standard if we assume that only one of κ_j, σ_j, or ω_j is nonzero $\forall \; j$, i.e. that all measured mediators are specific enough to have clear theoretical interpretation. This variation on notation is for the purpose of allowing for unclear theoretical interpretation of observed mediators, which Section 4.1 will show to be necessary.

market. In a pure human capital model, the set of mediating variables $x_1, ..., x_J$ that fully describe the effect of education would all have $\sigma_j = \omega_j = 0$. Signaling can similarly be defined using mediating variables; in a pure signaling model, $\kappa_j = \omega_j = 0$ for each of the mediators $x_1, ..., x_J$ that are sufficient to fully describe the effect of education. Pure signaling or human capital models can be rejected by showing that $\kappa_j \neq 0$ or $\sigma_j \neq 0$ for some j, respectively.

Other explanations similarly fit the mediating-variables setting with $\omega_j \neq 0$. If students use education to discover their own abilities (Arcidiacono et al., 2016, e.g.), then "beliefs about one's own abilities" fits into $x_1, ..., x_J$. If exposure to certain kinds of other students improves skills and socialization, or offers networking opportunities, then "exposure to students with quality X" is a part of $x_1, ..., x_J$. For this paper, I will focus on signaling and human capital and so assume that $\omega_j = 0 \; \forall \; j$ for most discussion.

Figure 1 addresses only the effect of education on earnings and productivity, why that effect exists, and how such estimates can be derived from observed data. It does not directly address other testable hypotheses that the human capital or signaling models might provide, for example the testable signaling-model assumption that the effort costs of education are negatively related to ability. These other hypotheses may offer ways to test for the presence of human capital and signaling effects, but they would not inform the proportion of the return that could be explained by either model, which is the effect of interest.

The relative importance of signaling and human capital are defined as K/B and Σ/B, respectively, and so answering the question of relative importance requires that one of these be identified. Under the assumption that $\omega_j = 0 \; \forall \; j$ it would be sufficient to identify two of K, Σ, and B, and use $K + \Sigma = B$ to identify K/B and Σ/B. The benefit of this mediating-variables model is that it outlines what must actually be done to identify the relative importance of these differing explanations of the education premium:

1. For some given explanation of the education premium, for example signaling, translate signaling from a theoretical proposition into an empirical one. The same steps would follow for human capital, swapping the place of the two explanations.
   - Theoretically determine a subset of mediating variables $\chi_\sigma \subseteq \{x_1, ..., x_J\}$ for which $\kappa_j = 0$ that are indicative only of signaling, a set χ_κ for which $\sigma_j = 0$ and so are indicative only of human capital, and a complement set χ^C of mediators that are partially human capital and partially signaling for which $\sigma_j \neq 0$ and $\kappa_j \neq 0$. $\chi_\sigma \cup \chi_\kappa \cup \chi^C = \{x_1, ..., x_J\}$.
2. Estimate the part of the effect of education on the outcome of interest that occurs because of signaling or because of human capital.[5]

[5] Each of these steps is done either by controlling fully for endogenous selection pressures, or by utilizing exogenous selection pressures such as instrumental variables (Pearl, 2009; Morgan and Winship, 2014).

- Proceed by adding: Estimate the effect of education on χ_σ, and then, separately, the effects of χ_σ on outcomes to get a signaling effect $\Sigma = \sum_{j|x_j \in \chi_\sigma} \beta_j$.[6]
- For each element j of χ^C, either identify and estimate σ_j and add it to Σ to get an estimate of $\Sigma = \sum_{j|x_j \in \chi_\sigma} \beta_j + \sum_{j|x_j \in \chi^C} \sigma_j$, or estimate the total effects of χ^C on outcomes, add them to Σ, and then identify, estimate, and subtract the κ_js to get an estimate of $\Sigma = \sum_{j|x_j \in \chi_\sigma} \beta_j + \sum_{j|x_j \in \chi^C} (\beta_j - \kappa_j)$.

or

- Proceed by subtracting: Estimate the effect of education on the outcome of interest while either controlling for all elements of χ_κ or while using a margin of education or source of exogenous variation that should not cause any element of χ_κ to get a total signaling effect $\Sigma = B - \sum_{j|x_j \in \chi_\kappa} \beta_j$.
- For each element j of χ^C, either identify and estimate κ_j and subtract it from Σ to get an estimate of $\Sigma = B - \sum_{j|x_j \in \chi_\kappa} \beta_j - \sum_{j|x_j \in \chi^C} \kappa_j$, or instead estimate Σ while controlling for both χ_κ and χ^C, and then identify, estimate, and add back the σ_js in χ^C to get an estimate of $\Sigma = B - \sum_{j|x_j \in \chi_\kappa} \beta_j - \sum_{j|x_j \in \chi^C} (\beta_j - \kappa_j)$.

3. Under the assumption that the steps have been followed, the estimated $\hat{\Sigma}$ identifies the signaling effect Σ.
4. Estimate either the total effect of education B, or, using these same steps, the human capital effect K. Compare this to the signaling effect to produce an estimate of the relative importance of signaling, defined as Σ/B and human capital, defined as $K/B = (B - \Sigma)/B$.[7]

Modeling the human capital/signaling divide in a mediating-variables setting, as in the previous section, already implies that distinguishing the two will be difficult. Mediation analysis is difficult even in randomized settings (Green et al., 2010), and in observational settings requires conditional ignorability both between treatment and mediator, and between mediator and outcome (Imai et al., 2010), making following step 2 difficult.

Despite these statistical difficulties, the above approach can, and has, fruitfully led to useful information about how the effects of education are mediated. However, there are limitations to translating this empirical knowledge back into a theoretical understanding of education. In the next section I consider the feasibility of actually following these steps, and what theoretical parameters can be identified using them.

[6] This is, in effect, the "front-door method" in the case of multiple mediators, as in Bellemare and Bloem (2019).
[7] These calculations assume that $\omega_j = 0 \ \forall \ j$ and must be modified otherwise, depending on whether the researcher wants to estimate the importance of signaling relative to the total return, or relative to human capital.

3 Partial Identification

In this section we will assume that an empirical study or set of empirical studies has followed the identification steps in the previous section, and examine what is identified in imperfect real-world settings. The steps identify a point estimate of the human capital or signaling share only if assignment of mediators to χ_σ, χ_κ, and χ^C is done perfectly, and each of the relevant effects can be point-identified. This is a high bar, and partial identification would still be valuable and inform policy. We can consider what the partial identification bounds would be under different realistic scenarios about which component parameters can be identified.

For simplicity I will discuss the signaling share Σ/B, although the same argument holds for the human capital share. Further, I will assume that B can be identified, and so focus on partial identification bounds for Σ itself, which then implies the bounds on Σ/B.

The partial identification bounds for Σ can be derived based on the subset of mediators $\chi^*_\kappa, \chi^*_\sigma, \chi^{C*}$ for which individual κ_j and σ_j effects are identified,[8] and the subsets $\chi^X_\kappa, \chi^X_\sigma, \chi^{CX}$ for which individual κ_j and σ_j effects are not identified. x_j may be in $\chi^X_\kappa, \chi^X_\sigma$, or χ^{CX} because the research design is incapable of identifying β_j, because x_j has not been conceived of as a mediator, because β_j can be identified but theory cannot determine whether x_j is a member of χ_σ, χ_κ, or χ^C, or because β_j is identified and $x_j \in \chi^C$, but neither σ_j nor κ_j are individually identified.

First consider cases where Σ is estimated by adding, following the first method in step 2 in the previous section by taking the effects that go through χ_σ and adding either $\sum_{j|x_j\in\chi^C}\sigma_j$ (possibly by adding $\sum_{j|x_j\in\chi^C}(\beta_j - \kappa_j)$, equivalent under the $\omega_j = 0 \;\forall\; j$ assumption). In these cases estimation identifies a lower bound $\underline{\Sigma}$:

$$\Sigma \geq \sum_{j|x_j\in\chi^*_\sigma} \beta_j + \sum_{j|x_j\in\chi^{C*}} \sigma_j \equiv \underline{\Sigma} \tag{1}$$

This approach does not identify an upper bound unless it achieves point identification by identifying σ_j for every element of χ_σ and χ^C (i.e. χ^X_σ and χ^{CX} are empty).

Second, consider cases where Σ is estimated by subtracting, following the second method in step 2. Similarly, this identifies an upper bound $\overline{\Sigma}$:

$$\Sigma \leq B - \sum_{j|x_j\in\chi^*_\kappa} \beta_j - \sum_{j|x_j\in\chi^{C*}} \kappa_j \equiv \overline{\Sigma} \tag{2}$$

[8] Throughout this section I refer to the identification of individual κ_j or σ_j parameters; it is not necessary in all cases that these individual parameters are identified as long as they are included in some aggregate estimate. For example, if χ^C is empty, then by blocking χ_κ while estimating the overall effect of education, the aggregate Σ can be identified without identifying the individual σ_js. This does not change the argument, and so I ignore the possibility of identifying these parameters in aggregate for simplicity of explanation.

Since both approaches are feasible, either both estimated in the same data set or aggregated across multiple studies, we can see these two terms as the full set of partial identification bounds for $\varSigma$, and by extension these bounds divided by B are the bounds for $\varSigma/B$, $[\underline{\varSigma}/B, \overline{\varSigma}/B] \subseteq [0,1]$.[9]

The conditions for some level of partial identification are weak: $\underline{\varSigma}/B > 0$ if χ_σ^* or χ^{C*} are not empty, and $\overline{\varSigma}/B < 1$ if χ_κ^* or χ^{C*} are not empty. In other words, there needs to only be one identified σ_j or κ_j to shift the bounds from 0 and 1, respectively.

However, narrowing identification bounds to a meaningful degree requires that the elements of $\chi_\sigma^*, \chi_\kappa^*$, and χ^{C*} make up a meaningful portion of the overall effect of education B.

In the following section, I make the case that meaningful narrowing of these bounds cannot occur. This argument proceeds in Sections 4.1 and 4.2 by showing that, in real-world settings, elements of $\chi_\sigma^*, \chi_\kappa^*$, and χ^{C*} make up a small portion of B. This happens for two main reasons:

First, because of the theoretical flexibility of the human capital and signaling models, there are few elements of χ_σ or χ_κ that are “pure” examples of either theory, and therefore few elements of χ_σ^* and χ_κ^*. Nearly all mediators, even those traditionally considered to be elements of χ_σ or χ_κ, are actually elements of χ^C.

Second, while I show that χ^C has many elements, I argue that χ^{C*} is usually empty. For a given mediating effect $x_j \in \chi^C$, an attempt to identify σ_j or κ_j separately runs into the same identification problem as in the attempt to identify $\varSigma$ or K. The issue recurs.

In Section 4.2 I consider the possibility that bounds could be meaningfully narrowed if the underlying latent variable “ability” could be measured, as this would define χ_κ comprehensively and be a member of χ_κ^*, but ability is too abstractly defined to be measured.

I acknowledge that there are certain contexts in which there may be many elements of χ_σ^* or χ_κ^*. This can occur because a certain source of, usually quasiexperimental, variation sets $\sigma_j = 0$ or $\kappa_j = 0$ for some set of variables that would normally be in χ^C. I discuss in section 4.3 why this does not solve the identification problem in the intended way.

4 Difficulties in Explaining the Returns to Education

The previous section outlined how explanations of the returns to education can be identified. In general, “a signaling effect” or “a human capital effect” can be found by selecting a set of mediating variables that can be convincingly labeled as clear empirical examples of signaling (χ_σ) or human capital (χ_κ), and isolating only the part of the effect of education that works through these variables.

[9] This assumes that the human capital and signaling effects are both nonnegative.

However, in order to relate results produced by this approach back to a theoretical explanation of the returns to education, we must be able to firmly establish which explanation these mediating variables are examples of. If the mediating variable x_j can be plausibly considered an example of both signaling and human capital (χ^C), then a mediation analysis using x_j must further identify the signaling and human capital shares of x_j itself (σ_j, κ_j) to inform the overall share of signaling and human capital.

There are three reasons why empirical evidence can have only a limited effect on our understanding of human capital or signaling, which will be addressed in the following subsections. Section 4.1 shows that too few mediating variables can be plausibly assigned as examples of exactly one explanation, in other words that χ_σ and χ_κ are small. Section 4.2 shows that the human capital and signaling models are both flexible enough in regards to the definitions of "ability" and "beliefs" that falsification is nearly impossible; alternate ways of defining mediators such that χ_σ and χ_κ are large do not work. Section 4.3 shows that human capital and signaling effects are too heterogeneous to be able to accumulate results across contexts where the issues from Sections 4.1 and 4.2 are avoided; the contexts in which χ_σ and χ_κ are truly large do not solve the problem.

4.1 Multiple Explanations

The process of identifying effects aligned with different theoretical explanations in a mediating-variables framework requires that different mediating variables can be claimed by a given explanation.

As shown in Section 3, the ability to narrow the partial identification bounds of the signaling or human capital share of education is weakened by the presence of mediators x_j that cannot be theoretically assigned to one explanation or the other (in χ^C rather than χ_κ or χ_σ). Elements of χ^C widen the partial identification bounds, unless σ_j and κ_j can themselves be individually identified, which is itself an identification problem with similar features to the attempt to identify overall human capital or signaling shares, as will be discussed further at the end of the section.

In this section I consider different areas of empirical research on human capital and signaling. The theoretical goal here is to show that areas of research that have claimed to find elements of χ_κ or χ_σ have in fact found elements of χ^C with indeterminate theoretical interpretation. Since these areas of research highlight mediators of the returns to education that appear to make up a large portion of the return, and these have so far been the most promising attempts to find elements of χ_κ or χ_σ, showing that these are elements of χ^C means that the partial identification bounds on human capital and signaling shares are barely informative.

I use Arteaga (2018) as an initial illustration.[10] In this study, the author looks at a top economics and business program in Colombia that reduced its coursework requirements. Graduates lacked a certain set of knowledge they would have otherwise had, but the population of students graduating from the program did not change immediately. So, the margin of education being examined (pre- and post-change in requirements) should affect earnings solely through the mediating variable x_1 "exposed to the set of knowledge taught in courses no longer required."

Large observed effects of education on early career earnings through the mediator of choice, β_1, are taken as evidence of human capital. In other words, it is assumed that $\sigma_1 = 0$ and $x_1 \in \chi_\kappa$, and so a non-zero β_1 captures a pure human capital effect.

However, as described in the paper, top employers in the region commonly gave applicants written exams including questions about the knowledge no longer covered by coursework. β_1 could be nonzero either because the knowledge acquired actually makes the students better workers ($\kappa_1 > 0$) or, framed in an alternative way to what is given in the paper, because employers found that the knowledge had been in the past a good signal of desirable employee qualities ($\sigma_1 > 0$, $x_1 \in \chi^C$),[11] and the results were more a consequence of employers either relying out-of-equilibrium on an outdated signal, or finding that hiring based on the signal is still the best screening approach despite being weaker than it once was. Distinguishing the two explanations requires the researcher to know whether the material learned is actually productive, which is a high bar.

One could make a reasonable case that the effects in Arteaga (2018) are *better*-suited to human capital than to signaling. But this establishes only that $\kappa_1 > \sigma_1$, not that $\sigma_1 = 0$, and so the human capital share of the effect, defined as κ_1/β_1, is unidentified. We do not know how far from 0 it can be bounded.

This same argument applies to any effect of education on outcomes that operates through skills that accumulate through education but are also visible to employers before hiring. Treating these effects as human-capital affiliated makes sense, but there are heavy requirements on the data to establish that signaling has no part to play in these results. And so, in evaluating the overall model of the returns to education, the entire portion of that return that operates through observable learned skills is in χ^C rather than χ_κ.

These sorts of interpretation issues apply to many of the observed phenomena that are used to inform our understanding of human capital and signaling. I will consider three here: employer learning, sheepskin effects, and the effect of education on aggregate productivity.

[10] I use this same paper as an example many times in this paper. I pick Arteaga (2018) specifically for this because it applies to many arguments, and because the work itself is of high quality and its empirical results are believable, so any issues I point out can be attributed to the flexibility of the human capital and signaling models rather than flaws in the paper.

[11] To demonstrate the precise argument being made here, consider a student who learns Shakespeare in college, and then makes a Shakespeare reference during a job interview, impressing the interviewer and getting the job. This knowledge of Shakespeare is a skill acquired in education, and improved their earnings, even though it may have no effect on productivity.

Employer learning refers to the ability of employers to learn employee productivity better through observation after hiring them than they could learn before hiring on the basis of imprecise signals. As the employer learns the employee's productivity, earnings should increasingly reflect actual productivity (Jovanovic, 1979; Farber and Gibbons, 1996; Altonji and Pierret, 2001).

The employer learning model is shown in the mediating-variables framework in Figure 2. Education impacts early earnings through employee ability and employer beliefs, which each have clear human capital and signaling interpretations. Since ability and beliefs are not directly observed, the approach of the employer learning model is to compare estimates of the direct early returns to education, $\kappa_{A1} + \sigma_B$, to estimates of the direct late returns to education, κ_{A2}.

This bounds Σ by subtraction, estimating B while removing the influence of χ_κ, using the assumption that elements of χ_κ affect early and late returns similarly. If the returns to education fade quickly and we assume $\kappa_{A1} \approx \kappa_{A2}$, σ_B must have been a large part of the early returns and so this is generally taken to be evidence in favor of signaling. If the returns to education persist or grow, σ_B must be small, and this is taken to be evidence against signaling.

The empirical literature on employer learning typically finds that the returns to education persist or grow over time, and that employer learning is too quick for unobserved ability to go unrewarded for long, a result given a human capital interpretation (Altonji and Pierret, 1997; Bauer and Haisken-Denew, 2001; Lange and Topel, 2006). Lange (2007) uses these results to place an upper bound on the contribution of signaling at no more than 45%, under the assumptions most generous to signaling. Aryal et al. (2019) derive a signaling share of 30% by contrasting short-term instrumented returns in cases where the instrument is visible to employers against cases where it is invisible.

Using evidence on employer learning to inform theory about human capital and signaling assumes both that rapid learning could not erase human capital effects and that there is no way for signals to affect late earnings. The presence of elements of χ^C or χ_σ that affect late earnings widen the bounds on Σ when estimating by subtracting the effects of χ_κ.

Habermalz (2011) revisits the original interpretation of the Lange (2007) model and reconciles the findings with the signaling model. Up to a point, faster employer learning actually improves the value of an educational signal. If employer learning is slow, then the out-of-equilibrium educated but low-ability workers who are the focus of the Lange (2007) analysis are difficult to detect, which reduces the quality of the signal. Fast learning reduces the benefits of college for low-ability workers, sharpening the separating equilibrium and increasing the value of college as a signal.

Arcidiacono et al. (2010) argue that higher levels of education allow underlying ability to be communicated to employers much more accurately than for lower levels, partially because there are many more signals that can be sent (club participation, test scores).[12] Education should have a moderating effect on σ_B itself, complicating the comparison of early to late returns. If part of educational signaling is in providing a highly refined and accurate signal, rather than broadly separating an educated person from a less-educated person, then we could observe short employer learning periods even if a large part of the return to education is signaling.

Standard interpretations of employer learning results assume that mistaken employer beliefs at the time of hiring cannot affect late earnings because employers will learn the true underlying ability. Aryal et al. (2019) make this explicit by describing a model in which the causal effect of education on late earnings is exactly equivalent to the impact of education on productivity. However, part of productivity in the workforce is firm-, industry-, or task-specific human capital that is acquired on the job rather than during education. In these cases, being assigned to a high-earning job early because of employer misperception allows the employee to gain specific human capital. Through experience, the misassigned employee improves their actual productivity beyond someone initially more skilled who simply failed to send the relevant signal. A similar phenomenon arises if sorting within firms into coworker groups occurs on the basis of education credentials and there are coworker spillovers (Bidner, 2014). Further, the education signal may lead to later favor in promotion by education level, holding ability constant (Waldman, 2016). This effect is present in Figure 2 by the fact that there is an arrow from Early Earnings to Late Earnings: it is possible for Employer Beliefs to affect Late Earnings via Early Earnings.

Under the arguments associated with Habermalz (2011) and Arcidiacono et al. (2010), information about employer learning does not identify the human capital or signaling share at all, since quick learning can in fact be evidence against a human capital interpretation. If neither of those arguments apply, then under the specific human capital argument, Employer Beliefs becomes an element of χ^C instead of χ_σ when looking at late earnings as an outcome variable, widening partial identification bounds by the size of the beliefs effect. What is identified by the employer learning literature is the combination of the human capital effect and the effect of training in specific human capital.

A second empirical observation is often taken as a major source of evidence in the human capital vs. signaling debate: sheepskin effects. The returns to years of education are much higher for years in which a degree is earned (Hungerford and Solon, 1987; Jaeger and Page, 1996; Belman and Heywood, 1997; Flores-Lagunes and Light, 2010). This is an empirical regularity that is observed globally, and I refer to it here as a sheepskin effect. In the mediating-variables framework, the relevant treatment variable is "years of

[12] The Arcidiacono et al. (2010) empirical result that there is no employer learning for college graduates is disputed (Light and McGee, 2015). However, the use of the Arcidiacono study here does not rely on their empirical result.

education," the mediating variable of interest x_2 is "holding a degree," and the sheepskin effect argument assigns this mediating effect to signaling such that $\kappa_2 = 0$ and $x_2 \in \chi_\sigma$.[13]

Before considering the argument that sheepskin effects may not be entirely signaling, it is worth noting the contradiction between the employer learning and sheepskin literature. Both literatures frame, and sometimes explicitly refer to, their mediating effect of interest not as just *a* signaling effect, but as *the* signaling effect Σ/B, as a point estimate or bounds. These literatures cannot both be right. The generous 45% or preferred 10% maximum signaling share found in Lange (2007) in the employer learning literature is mutually exclusive with the conservative 50% or preferred 80% minimum signaling share suggested by Caplan (2018) from a review of the sheepskin effect literature. This tension can be resolved if the estimates are statistically noisy enough that a true identified share can produce these inconsistent results due to noise, or if the partial identification bounds are not as narrow as claimed in either case.

There are several explanations of observed sheepskin effects that do not rely on signaling. The first is that sheepskin effects simply reflect selection into graduation on the basis of prior observables. However, sheepskin effects tend to persist after adjusting for selection pressures and prior observables (e.g. Frazis, 1993). The second is that sheepskin effects reflect selection into graduation on the basis of factors that could not be known ahead of time: students learn of their own return to education through the process of education and drop out, ensuring that those with the lowest returns are seen terminating their education at non-degree years. This explanation is explored theoretically in Chiswick (1973), Hungerford and Solon (1987), and Lange and Topel (2006), and the underlying mechanism can be seen in studies of ability revelation in college (Stinebrickner and Stinebrickner, 2014; Arcidiacono et al., 2016). In both of these explanations, endogenous selection pressures are not properly adjusted for in sheepskin effect estimation. Under the second explanation, the holding a degree mediator x_2 is in χ_σ only for students who enroll in a program knowing for certain whether they will graduate or not, which may be a small and unidentifiable group. $x_2 \in \chi^C$ otherwise.

The third is that the years of education and earnings may be negatively correlated among graduates but positively correlated among dropouts because more-skilled students graduate more quickly but drop out more slowly, leading to high observed returns in degree-granting years (Flores-Lagunes and Light, 2010). This implies that measured years of education is a poor proxy for amount of education received, and so looking at the impact of years of education mediated by having a degree does not identify even β_2.

[13] I refer here specifically to estimates that compare the returns to education between degree-granting and non-degree-granting years. Several of the arguments presented here that interpret sheepskin effects in human capital terms do not apply to natural experiments that estimate the return to holding a degree in other ways, like Tyler et al. (2000). These will be addressed in Section 4.3.

The fourth is that the original argument that sheepskin effects cannot reflect human capital may be partially incorrect, and at least some small part of the sheepskin effect can be explained by students learning more valuable skills in the final year than in earlier years, or becoming more capable of absorbing what is taught ($\kappa_2 > 0$, even if it is small, and $x_2 \in \chi^C$). This may be plausible in any context, like college, where curriculum becomes more specialized in later years.

In addition to these well-known arguments as to why sheepskin effects may not be entirely signaling, I present another. I take it as given that part of the return to education is that it provides a credential certifying student ability. However, this credential certifies not just prior ability but also skills learned in the process of education. There should be an arrow in the diagram from the mediator "human capital accumulated in school" to the mediator "has a degree." Since the effect of education on earnings mediated by having a degree incorporates the effect of human capital accumulation, $\kappa_2 > 0$ and may even be large, and $x_2 \in \chi^C$. The fact that sheepskin effects persist after controlling for initial student ability measures lends plausibility to the idea that part of what is credentialed is learned in school. This argument is similar to Graetz (2017), who points out the distinction between the "information content" of a degree and the signaling content. Under this argument, sheepskin effects identify $\beta_2 > \sigma_2$. What is identified is a combination of the signaling effect and the effect of anything learned in school that contributed to graduation.

To illustrate, consider a mass of students of identical prior ability 1 facing four years of education, the fourth of which earns a degree. Identical prior ability ensures there can be no signaling of prior ability or selection on prior ability. Every year t, each student i sees their ability increase through learning by $a_{it} \sim U[0, 1]$, and the enrolled students in the bottom decile drop out because the university gives them failing grades. Wages are equal to average ability within education group after graduation. The observed returns will produce a sheepskin effect derived purely from human capital accumulation differences, with 28%, 31%, and 115% returns for the second, third, and degree year relative to the year before.[14]

I discuss one last area in which empirical results are commonly used to make inference about the relative importance of human capital and signaling: the impact of education on aggregate productivity and growth. This is also referred to (with some variation in concept) as the external or social benefits of education, or education spillover effects. "Aggregate worker skill" is the mediating variable of interest x_3. Human capital implies that part of the impact of worker skill on output is because aggregate worker skill is driven by aggregate education ($\kappa_3 > 0, x_3 \in \chi_\kappa$). In this case, the simplest signaling model does not assume that $\sigma_3 > 0$ or that $x_3 \in \chi_\sigma \cup \chi^C$. Rather, it assumes that $\kappa_3 = \sigma_3 = 0$ and $\beta_3 = 0$.[15] A result that higher levels

[14] These figures come from a simulation using 1,000,000 students, with log ability regressed on schooling level indicators to estimate returns.

[15] $\beta_3 = 0$ and $\beta_3 > 0$ here can be substituted with $\beta_3 = \epsilon_3$ and $\beta_3 > \epsilon_3$.

of education causally improve productivity or lead to economic growth ($\beta_3 > 0$) is generally taken as an example of human capital.

Unlike with employer learning and sheepskin effects, the empirical effect of interest is less settled here. Evidence on national growth generally ranges from the null to the optimistic (Topel, 1999; Lange and Topel, 2006; Goldin and Katz, 2009; Valero and Van Reenen, 2019; Caicedo, 2019). Studies using within-country regional variation often find positive external benefits of education (Acemoglu and Angrist, 2000; Moretti, 2004), but these results too are inconsistent (Ciccone and Peri, 2006). There are studies of the underlying necessary mechanism finding improved productivity within firms on the basis of education (Battu et al., 2003; Crook et al., 2011; Bentsen et al., 2019).

Regardless of what the impact of education on aggregate productivity levels actually is, a positive finding would be taken by standard interpretation as an example of human capital accumulation, bounding K from below by β_3. However, this assumes that the signaling function of education is nonproductive, which is untrue in any version of the signaling model in which the return to skill varies between occupations, for example shown theoretically in Hopkins (2012) or empirically in van der Meer (2011) and Van Der Velden and Bijlsma (2016). Here, "talent allocation" x_4 is the mediating variable of interest, and the ability of the signaling model to sort workers to the right jobs unambiguously improves productivity ($\sigma_4 > 0$). What is identified by estimates of the effect of education on growth is $\kappa_3 + \sigma_4$, and so K is not bounded without further ability to separate the two. If worker ability is complementary with the skill of other workers (as in Kremer, 1993), the impact of this sorting will be heightened. Signaling may also be productive because it allows firms to predict worker ability and thus equalize marginal products across different forms of production (Wolpin, 1977).

These three effects: employer learning, sheepskin effects, and external returns, have made up the backbone of the literature separating human capital and signaling. In addition to returns that operate through employer-visible skills, as previously discussed, estimates of these effects in the literature appear to make up a large portion in total of the observed return to education.

When applying these empirical results to the theories of human capital and signaling, these studies explicitly or implicitly assume that their mediators of interest belong to χ_κ or χ_σ, or that κ_j or σ_j can be separately identified. The ability to narrow the partial identification bounds of Σ/B or K/B relies on this interpretation. I have made the case in this section that these mediators actually belong to χ^C, and that the individual κ_j or σ_j parameters for $x_j \in \chi^C$ are not being identified. Because this identification does not occur for x_j mediators that appear to account for a large portion of B, the partial identification bounds cannot be narrowed far beyond $[0, 1]$.

Bounds could be improved by theoretical assumptions about the primacy of each explanation within given mediators. For example, even if “having a degree” is in χ^C rather than χ_σ, economists may agree that it is *more* supportive of signaling than of human capital. An assumption that $\sigma_2 > \kappa_2$ would narrow the bounds by at least half the amount that assuming $x_2 \in \chi_\sigma$ would. But it is not clear on what basis these claims can be made, since theory rarely specifies a magnitude for relative share, and it is unlikely that they could come from empirical results.

While these empirical effects still may intuitively rest more with one explanation than the other, the important point is that they cannot be clearly assigned to being *entirely* one explanation or the other from theory alone. There is no clear way to break them down further such that κ_j and σ_j can be separated; any attempt would face the same problems as trying to break down the return to education as a whole into a human capital share and a signaling share. If a set of mediators explaining large portions of the return have the same problem, which I have argued in this section is the case, then the mediation evidence necessary to narrow the partial identification bounds to a useful degree does not exist.

This finding about wide partial identification bounds occurs under some theoretical assumptions made thus far that favor narrow bounds. So far, my analysis, and most of these studies themselves, have assumed that explanations of the causal return to education other than human capital or signaling, $\mathit{\Omega}$, are zero. Relaxing this assumption makes the theoretical assignment of mediated effects much more difficult, especially when the goal is to estimate bounds by subtraction - what is really identified in these cases is $K+\mathit{\Omega}$ or $\mathit{\Sigma}+\mathit{\Omega}$ rather than K or $\mathit{\Sigma}$. Another issue is that many of the empirical predictions derived from the theoretical models presented in this section assume a strong relationship between productivity and earnings as in a perfectly competitive labor market. Monopsony, discrimination, or other frictions strain this relationship in sometimes-unpredictable ways, making the theoretical interpretation of empirical observations more difficult and the partial identification bounds wider.

4.2 Ability and Beliefs

Section 4.1 details the problems associated with assigning different observed mediators to human capital or signaling in order to identify the contribution of each, as in Figure 1. Perhaps it does not need to be so difficult. The empirical model that may be in the mind of some researchers as they consider the relative contributions of human capital and signaling may not be the complex Figure 1 but instead the simpler Figure 3 in which the assignment of each mediating effect is clear.

Under Figure 3, human capital and signaling effects can be cleanly defined by simply selecting appropriate proxies for Employee Ability or Employer Beliefs. χ^C is empty. The human capital share, for example, is given in Figure 3 as $\kappa_A/(\kappa_A + \sigma_B)$ or κ_A/B. This is effectively the same approach that is taken in Section 2, but adds the identifying assumption that Employee Ability can be fully proxied by observable measures of ability such as test scores, or that Employer Beliefs can be fully proxied by some measure of employer beliefs, or fully controlled by examining a situation where it would be difficult for employers to see variation in education.

This approach fails because both theories resist the use of proxies.

The problem grows from the fact that "ability" is broadly defined, both in the human capital and signaling models. In these models, ability is not limited to intelligence, but rather is any quality that makes someone a more productive employee. Ability is necessarily multidimensional, varies across occupations, and includes features that researchers do not have access to or are effectively unmeasurable.

The abstract nature of ability makes both human capital and signaling exceedingly difficult to falsify using proxies of ability. If measures of ability learned in school do mediate the returns to education, this is taken as evidence of human capital. But if this ability can be observed by employers, as in Arteaga (2018), it can be argued that these learned skills increase wages because they are signals and do not contribute to productivity. If measures of ability learned fail to mediate the returns to education, this is taken as evidence of signaling.

But in any case it can be argued that the wrong sort of ability has been measured, as long as the list of skills that employers actually value is not known or includes unmeasurable characteristics. The partial-identification bounds on κ_A relies on the strength of the proxies. The lower end is bounded by the effects of measurable proxies that can be said to be unambiguous examples of productive employee ability, which may be a small share of κ_A if employee-observable abilities do not count, and most available ability measures have uncertain effects on actual productivity. The higher end is bounded by the presence of examples of productive employee ability that are unmeasured, which may be a large share of κ_A if productive ability is truly high-dimensional and contains many intangibles.

We can consider the implications for human capital and signaling theories under hypothetical empirical findings that should be strong evidence that each provides a small share of the education return. Finding that education has little impact on ability, for example, should be strong evidence against the human capital model, since if education only weakly affects ability, then κ_A is necessarily small or zero. Similarly, finding that education significantly contributes to the development of job-relevant skill should minimize the potential impact of signaling (κ_A is large), as would findings that employer beliefs are not affected by education (σ_B is small or zero).

First, we consider the implications of findings that education has little impact on ability. Arum and Roksa (2011), for example, argue that recent cohorts of college students retain relatively little of the knowledge they are taught in class. Further, the psychology of learning literature finds limited student ability to learn things far removed from what is taught in class (Barnett and Ceci, 2002; Ambrose et al., 2010; Sala et al., 2019). Let us take these empirical results as given, and consider the implications on theoretical understanding. Given this evidence, Caplan (2018) argues that it is effectively impossible that skills are heavily improved in college, and so the human capital model is likely to apply little. These approaches claim to identify a small κ_A and thus point-identifying a small human capital share.

However, even if there is little evidence that education moves measured skills,[16] the human capital model is flexible enough to accommodate.

Unless learning is zero, understanding whether learning is "large" or "small" requires actual measurement of the outcome of interest, not just measurement of skill. This places heavy data demands on this particular argument against human capital, and has been pointed out by several responses to Arum and Roksa (Pascarella et al., 2011; Haswell, 2012). Skills of labor-market interest could be close enough to what is directly taught in class that transfer across closely-linked domains occurs. Education could directly teach other skills - learning to turn something in on time, for example, does not appear in the "learning objectives" part of a syllabus and would not be included in a follow-up test of learning, but it does appear on the syllabus and is a skill practiced in school. This implies that these studies do not bound κ_A because this requires research in which actual productivity is the dependent variable.

The argument that estimates of the effect of education on skills measure the wrong skills can be made regardless of how many abilities education may be shown not to affect. This frames human capital theory as being so flexible as to be unfalsifiable through the measurement of ability.

There is evidence to back up the conditions that lead to unfalsifiability. Heckman et al. (2013) provide one example of this, in which the authors find that the Perry Preschool program had effects on student personality despite fading or null effects on achievement tests. Chetty et al. (2011), Carneiro and Ginja (2014), and Baker et al. (2015) provide similar evidence in other contexts. Chetty et al. (2014) find that the assignment of different teachers affects adult labor market performance, even though the effect of a given teacher on cognitive skills is generally recognized to decay much sooner than adulthood. This literature does not mean that a null finding of the effect of education on intermediating skills is non-informative, but it limits the extent to which theoretical inference can be drawn from empirical results. These results imply large and meaningful non-measured effects, pushing the upper bound of κ_A/B closer to 1.

One potential approach to restoring falsifiability via measured skill to the human capital model would be to find a set of variables representing measurable ability that fully mediate the effects of education,

[16] There is reason to doubt that the effects are actually zero - there is no shortage of studies that find effects of various educational interventions on test scores. The well-established ability to affect test scores at the margin implies a general effect of education on measurable ability, although an argument could be made that the effect is small with some definition of small.

which would point-identify relative contribution. Hanushek (2016) finds that cognitive scores fully mediate the relationship between education and national growth. However, if this finding were to hold on individual data, it would contradict the standing evidence on individual returns via other measures.

The signaling model, like the human capital model, relies on a broad measure of ability, which can make the model flexible in the same way. Arteaga (2018) is an example of education clearly improving some measure of ability that mediates the returns to education. We can take this result for granted and consider the implications for the signaling model. As previously argued in regards to Arteaga (2018), very strict conditions must be placed on the visibility of that skill in order to ensure that the phenomenon cannot be explained using signaling, which pushes the lower bound on κ_A/B closer to 0. The necessity of these strict conditions follows from the flexible definition of ability. If the measured skills learned in these classes are a full description of "ability", then a human capital interpretation of the results can be taken regardless of visibility to employers.

The broad measure of ability makes the signaling model flexible in another way. A common critique of the signaling model is that, if education is largely about signaling, then employers should be able to find far less expensive ways than education of identifying high-quality workers. Most employers have yet to find a way to do this. The standard response to this critique is that education does not just signal easy-to-measure things like intelligence, but a host of wider, inherently unobservable characteristics (Lang and Kropp, 1986), such as conscientiousness. This response mirrors the human capital-supporting argument that, if education does not improve measured skills, it may still improve other, unmeasurable skills. The upper bound on σ_B/B is pushed towards 1. Like the argument in favor of human capital, this defense of signaling frames it as flexible enough to avoid falsification on the basis of measured ability because any set of measurements for ability is incomplete. For this reason, the aforementioned Hanushek (2016) result, if it held at the individual level, would pose a challenge to the signaling model as well as to the human capital model.

There is a potential alternate avenue for tests of signaling. While employer belief cannot be proxied by the use of measured skill variables, it may be feasible to measure holistic employer belief using subjective report data. Researchers can ask employers how they evaluate employee skill and interpret the presence of education, as in Rivera (2011, 2016). Measurements of this kind are common in many fields in both quantitative and qualitative form, and have seen development in economics as well, mostly in quantitative form (see for example Manski, 2004).

If employer beliefs can be accurately measured and used in a model of the overall returns to education, this offers a realistic potential path forward for point-identifying the signaling share of the return. However, in order to answer questions about relative share, the subjective data would need to pass several difficult barriers beyond common issues related to self-reported perception data. Responses would need to be quantifiable in order to provide information about how much the partial-identification bounds are narrowed. Employer reports would also need to be structured so as to isolate the signaling, human capital, and selection parts of their beliefs. This would entail separating out the part of an employer's perception of education that is non-causal. Then, of the causal part, the study would need to determine how the effect is mediated by actual skills learned in education. As discussed in the sheepskin effect part of Section 4.1, there are several human capital-consistent reasons why an employer might change their perception of a candidate upon learning their education level, and employers may not be able to distinguish and discount these reasons in self-report. Rivera (2016), for example, finds extensive evidence about how elite employers rely on education signals about acculturation, but cannot distinguish what portion of the perception comes from acculturation that is learned in college. Because of the potential for close experimental control in a subjective-expectations study, an analysis making these distinctions may be possible, but has not been performed yet, and would face a number of difficult hurdles.

4.3 Quasiexperiments and Heterogeneity

In the previous two sections I made the case that it is extremely difficult to cleanly identify the extent to which signaling or human capital explain the return to education. These problems can be overcome in serendipitous circumstances. For example, a natural experiment may push students across a particular margin of education in a way that is invisible to employers (as in Pischke, 2007; Goodman, 2019), or change what employers believe about ability without changing the actual ability (as in Tyler et al., 2000). By identifying the entirety of K or Σ without needing to specify mediators, it still may be possible to construct a general model of education returns. This is how theoretical inference advances for many empirical problems.

In this section I argue that a general model of the returns to education, in which the relative contributions of human capital and signaling are properly estimated, is unlikely to come from an accumulation of quasiexperimental evidence from different contexts. The structure of these studies requires that they be aggregated together in order to answer a question about relative contribution. However, the effects, even if plausibly estimated within any given study, are simply too heterogeneous to be aggregated with confidence. As a result, the local average treatment effects these studies uncover are useful and of interest, but do not meaningfully inform the debate about the relative importance of human capital and signaling. These quasi-experimental studies often, correctly, interpret their own results as providing evidence of the existence of human capital or signaling effects, rather than the relative importance of those effects.

To make use of the set of plausible quasiexperimental studies that identify either human capital or signaling effects, it is necessary to consider exactly what is identified in these studies. Consider a study that quasiexperimentally identifies a human capital effect. In accordance with the steps in Section 2, such a study uses a source of variation that identifies the effect of education while excluding a signaling interpretation. This either identifies κ_j for some mediators j, or, if the total effect of education is estimated while blocking a signaling explanation, the total human capital effect K is identified for the given educational margin. The ability to identify total effects without needing to measure and specify mediators is what separates these studies from others, and suggests a possibility for point-estimation of the relative share.

An important distinction here is that the ability to block a signaling explanation to identify K does not mean that $\Sigma = 0$. Arteaga (2018) uses a policy change in required curriculum not visible to employers, arguing that this isolates the human capital effect (κ_1) of the mediating variable x_1 "exposed to the set of knowledge taught in courses no longer required." However, even if κ_1 is identified in this study, that does not mean that $\sigma_1 = 0$ in the true model. Taking as given that employers could not observe the *change* in curriculum, they could observe the curriculum in the first place, and could take it as a signal.

In order for Arteaga (2018) or similar studies like Pischke (2007) or Goodman (2019) to identify the relative importance of signaling and human capital, it is not enough to identify κ_1, or even the full K. As per the steps in Section 2, K must be compared to Σ or to the full return B to get relative importance. Similarly, studies identifying Σ would also need to also identify K or the full return. The nature of these studies is that, in nearly all cases, the exogenous variation that allows either K or Σ to be identified precludes the estimation of the remaining Σ or K, or the overall effect. If variation in education (or a mediator) reflects only K or Σ, then that variation is incapable of providing information about the alternative explanation at the same time. In the Arteaga (2018) context, if there had been variation in x_1 relevant only to signaling so that σ_1 could be identified, it would have interfered with, and possibly invalidated, the identification of κ_1, which relied on no signaling-related changes occurring at the same time.

Outside of cases where estimates show $K = 0$ or $\Sigma = 0$ or where one study manages to estimate both K and Σ (or one of them and B), then, quasiexperimental studies are incapable of point-identifying the human capital or signaling share, and in fact cannot even provide bounds on it, since for a given K, for example, Σ could be anywhere from 0 ($K/B = 1$) to ∞ ($K/B \approx 0$). To narrow these bounds, it is necessary to either assume a maximum plausible Σ or B, or to compare across multiple cleanly-identified studies to identify the relative importance of signaling and human capital, using the K from one and the Σ or full return B from another. However, this sort of comparison requires a consistency across settings and margins that is not present in the total, human capital, and signaling returns to education. These effects are too heterogeneous to be able to make these cross-study comparisons.

The return to education itself, like many causal effects in the social sciences, can be expected to be heterogeneous. Evidence on the return backs this up; the return differs across the margin of education studied

(Jaeger and Page, 1996), across countries (Trostel et al., 2002), across demographics (Cunha and Heckman, 2007; Henderson et al., 2011), and across labor market conditions (Altonji et al., 2016). The literature on the returns to education has long had to confront the difficulties of attempting to make generalizable statements about the returns to education when the best evidence that addresses endogeneity is necessarily context-specific or produces a local average treatment effect (Card, 1999).

In addition to heterogeneity in the overall return, the human capital and signaling shares of the return should similarly be heterogeneous, perhaps even moreso than the return itself.

There is, of course, variation in human capital effects that can be measured in a fairly straightforward way. Different students see different amounts of improvement in their measured ability following schooling on the basis of personal characteristics such as race (Fryer and Levitt, 2004) or different qualities of the education they are exposed to such as the teacher or school assigned (Deming, 2014; Chetty et al., 2014).

Ability improvements as a result of education can be measured more directly when students enter jobs where productivity can be measured. Chingos and Peterson (2011) find that a master's degree in education is uncorrelated with effectiveness as a teacher, even though it guarantees higher pay under many teacher payment agreements. Hussey (2012) finds that the personal returns to an MBA are not reflected in a causal productivity improvement. Both results imply that the any selection-corrected returns to these degrees are signaling or some other non-human-capital explanation. These results relating to particular educational margins differ from the broader literature, implying heterogeneity in human capital effects across degree types.

Signaling effects are heterogeneous as well. Theoretically, signaling effects should be more heterogeneous than human capital effects. Human capital effects should vary with the individual ability to acquire skills from education and the quality and format of that education, all of which are likely to follow well-behaved distributions. Signaling effects, on the other hand, should vary with the skill level of other people sending the same signal and also with the sum total of all other information the employer has about the employee. Employer beliefs are likely to vary sharply on the basis of what information is available, and so the signaling effect should vary sharply too.

Consistent with the theoretical prediction, studies of signaling effects that examine heterogeneity in the effect tend to find it. Bedard (2001) finds gender differences, and Tyler et al. (2000) find racial differences. Clark and Martorell (2014) is an exception, finding that high school degree signaling effects were uniformly zero across all groups studied. More broadly, sheepskin effects vary significantly across time, geography, and demographics (Belman and Heywood, 1991; Gibson, 2000; Belman and Heywood, 1997; Bitzan, 2009; Bol and Van De Werfhorst, 2011). While I have argued that sheepskin effects are not fully signaling, it is unlikely that major variation in sheepskin effects can be fully explained by heterogeneity in the human capital portion of the sheepskin effect.

This generalizability issue applies to most causal effects in social science, including the overall effect of education, but is particularly damaging to the attempt to identify the relative importance of signaling and human capital. Unlike with other questions in social science, the problem is not in whether a particular estimate is generalizable. Rather, two estimates from different settings must be assumed to generalize in order narrow the partial identification bounds at all. Additionally, because of the theoretical inference issues outlined in previous sections, the contexts in which effects can be cleanly estimated are narrow, and so generalizability is difficult to check through replication.

One potential way around this heterogeneity problem is to estimate the human capital and signaling shares is to model the problem structurally, so that total K and Σ can both be estimated in a single context without needing quasiexperiments, point-identifying the relative share. Fang (2006) uses a simplified structural model in which the human capital and signaling shares are identified on the basis of the model and assumptions about the ability distribution. This approach may offer the most hope for plausible generalizability. However, a structural approach necessarily relies on selecting a particular structure by which human capital and signaling operate. In effect, a structural approach addresses the problem from Section 4.1 that both models are flexible enough to explain wide ranges of behavior by disallowing that flexibility. Creating variants of the human capital and signaling models rigid enough that they can actually be pinned down may be preferable to declaring the unresolvability of the issue, as this paper does. But the resulting versions of human capital and signaling will not match the flexible theoretical versions, and so standard theoretical implications of the theories may not hold.

5 Related Questions and Paths Forward

Human capital and signaling underlie economists' understanding of why there are positive causal returns to education. They are useful concepts for developing a theoretical understanding of the returns to education, and they can be used to generate hypotheses that can be tested empirically. However, for the reasons outlined throughout Section 4, it is exceedingly difficult, and perhaps not possible, to use these results to generate an understanding of how educational returns function overall by estimating the relative importance of the two theories.

In other words, these two theories, both designed to explain why there are labor market returns to education, have limited application in explaining what is important in driving the labor market returns to education we see in the world. I suggest that an ability to explain the reasons for the returns to education, and the importance of those reasons, is a valuable goal. There are ways to divide and explain the returns to education that may be more fruitful and more amenable to empirical observation than human capital and signaling. I suggest two approaches here. Fortunately, both are already underway within the economics of education.

One approach is to be generally atheoretical. Understanding the variables that mediate the returns to education is a valuable goal in itself, and has important policy implications. There is already an extensive list of studies, many of them cited in this paper, that examine variables that mediate the returns to education without attempting to infer anything about human capital or signaling. The atheoretical approach focuses on measurable mediators and considers those to be of direct interest rather than as proxies of deep theoretical constructs.

The atheoretical approach has obvious downsides. Without building towards a general theory of the returns to education, results are difficult to generalize. An atheoretical approach ignores the unmeasurable latent variables like "ability" or "beliefs" that, while unmeasurable and flexibly defined, remain important. However, due to the amorphousness of these terms, as discussed in Section 4.2, a focus on trying to make the leap between measured variable and theoretical construct may act to inhibit our understanding rather than expand it. Empirical work that does not attempt to make this leap is valuable on its own. For example, Cardoso et al. (2018) acknowledges the place of ability in the model of the returns to education, but focuses its implications on decompositions of the returns to education according to measurable factors. Hanushek (2016) is another example.

Despite the value of atheoretical work, having a general theoretical framework is useful and allows for predictive and policy analysis outside the bounds of what has already been observed. But in the absence of a theoretical model that can make use of empirical findings, those empirical findings can only be used in an atheoretical context. There may be theoretical framings other than human capital and signaling that are more amenable to be informed by empirical data and are as relevant to policy, or moreso. The second path forward is to use such a framing.

The use of the human capital vs. signaling framework in the context of policy prescription has often focused on the question of how much education subsidy is justified. The standard policy argument assumes that, if education is mostly human capital, then external returns will be large and positive, and subsidy is justified. If education is mostly signaling, then external returns will be small, and education may have undesirable effects on income distribution (Stiglitz, 1975), so subsidy is unjustified or less justified.

However, this line of argument is misdirected. First, because the question of whether education is "mostly" signaling or human capital is empirically unresolvable, and knowing that both effects exist in some unknown proportion does not allow for strong policy prescription. Second, even if the relative contributions of human capital and signaling could be estimated, these policy prescriptions do not actually follow. As discussed in Section 4.1, education as signaling can improve productivity by leading to a more efficient allocation of skill to jobs or by concentrating talent geographically. Conversely, education that builds human capital can reduce productivity if the skills attained allow graduates to enter industries that rest on rent-seeking or negative externalities.

I argue that the exact same policy question of interest can be answered more directly and accurately using a theoretical framing that is already in use and is more amenable to being informed by empirical data: the identification of private and external returns. The primary policy application of human capital vs. signaling terms effectively uses human capital and signaling as stand-ins for private and external returns anyway (Lange and Topel, 2006; Caplan, 2018; Aryal et al., 2019). It makes more sense to simply study the question of actual policy interest, which conveniently is also more amenable to empirical analysis.

A focus on private vs. external returns would not replace exactly the human capital vs. signaling debate (except perhaps in policy relevance), as it answers a different question. This would not be a different means to resolve the same debate, but rather a shift in focus to a question that is more capable of incorporating empirical results and producing useful policy prescriptions.

Private and external returns can be estimated in a model like Figure 1, focusing on the overall effect of education on private and external outcomes, or more directly by ignoring the mediating variables altogether as in Figure 4. This approach does not need to concern itself with assigning each mediating variable x_j to one explanation or another, avoiding effectively all of the issues raised by Sections 4.1 and 4.2, and additionally being more robust to the existence of noncompetitive labor markets where the link between ability and earnings is unclear. Private and external returns do not require as much emphasis on relative importance. For most policy discussions, the *share* of the return that is private or external is not as important as the absolute levels of private or external returns - subsidy is justified by the presence of external returns, not the share. Many of the difficulties outlined in this paper relate specifically to attempts to estimate a share, which requires the estimation of two competing causal effects, and runs into many of the problems outlined in Section 4.3.

As a means of organizing empirical data on the returns to education into theory, the private vs. external distinction may prove far more useful than human capital vs. signaling and offers a clear path forward. Unlike the human capital and signaling framework, where relatively few studies inform either theory, all existing research on the overall return to education applies to our understanding of the private return, and work on the effect of education on economic growth, productivity, and the wages of others applies to our understanding of the external return.

The estimation of private vs. external returns has its own set of difficulties, as do most empirical questions of interest. Endogeneity problems are difficult to resolve when the outcome is aggregate and so is determined by many interlocking factors. A measurement problem specific to the private vs. external returns framework is that the researcher must define the external group that should be affected, and determine how to separate out the returns an individual receives.

This problem with measurement and theoretical assignment is real, but is simpler than the similar problem faced by signaling vs. human capital. While the underlying construct of "private vs. external" may be difficult to define empirically, its theoretical meaning is clearly understood, which is often not the case with signaling vs. human capital mediators. Difficulties with identifying private and external effects, or narrowing partial identification bounds, is a statistical or definitional problem rather than a theoretical one. That definitional problem is of a narrower focus than the problems with human capital and signaling, where there are many different mediators for which the effect must be partitioned theoretically. One additional way in which private vs. external returns are more easily identified is that they are an exhaustive description of the causal returns, and studies of these effects do not need to assume away or grapple with alternative explanations Ω.

The theoretical explanation of why external returns exist, of course, incorporates the discussion of human capital and signaling and so empirically would face a similar kind of empirical unresolvability. Still, all of these difficulties with private vs. external returns are of a very different sort than in human capital and signaling. With human capital vs. signaling, the question of interest cannot be answered empirically. With private vs. external returns, the question of interest is at a level where it can be answered empirically, and the answers can be used to form a useful model of educational returns, even if the mechanisms that model describes are not as deep as with human capital vs. signaling.[17]

6 Conclusion

The current theoretical view of the returns to education is that these returns can be explained using human capital accumulation and signaling. In many contexts, empirical evidence convincingly supports evidence that both effects exist. In this paper I make the case that empirical evidence can do little more to inform theory in this case. Human capital and signaling are both theoretically flexible enough that most observed behavior can be explained by either. The mediating effects that we use to empirically formalize human capital and signaling are rarely pure examples of either model. The flexible conception of ability prevents it from being proxied accurately. Finally, the circumstances in which other problems can be overcome in order to generate cleanly identified estimates of one model's effects cannot be used to provide an overall picture of

[17] For any theoretical framing, there is always a deeper level on which mechanisms are not explained. Under a pure human capital model, for example, why does education improve skills? This is an unaddressed, deeper mechanism.

the relative importance of the two theories. These issues mean that the relative importance of human capital and signaling is only partially identified, and that the bounds on partial identification are too wide to be useful.

This paper examines the empirical and theoretical literature on human capital and signaling to establish the unresolvability of the question of relative contribution. Of course, while relative contribution is often presented by economists as being an important topic for future research, it is also generally recognized that the question is, if not unresolvable, at least extremely difficult. What, then, is the point of this paper?

What this paper aims to do is not only to establish that the relative contribution of human capital and signaling is unresolvable, but to detail precisely why that is the case. Section 2 provides a set of requirements for establishing whether or not a question of the form "why, of these competing nonzero explanations, does X have a causal impact on Y?" is theoretically identified, and the remainder of the paper pinpoints the precise reasons why human capital vs. signaling fails to meet these requirements. This also highlights that these theories, like many economic theories, provide testable predictions about the presence of certain empirical relationships, but are not precisely specified enough to predict the magnitude of those relationships, which would allow a given empirical observation to cleanly inform theory.

This structured approach to identification helps specify what parts of theory, exactly, can be informed by empirical analysis. The theoretical claims "there are no signaling effects in education" and "there are no human capital effects in education" can be tested empirically, because they only require one mediating variable for each theory that can be unambiguously interpreted, or one instance with plausible quasiexperimental variation in only one theory's effects. The claim "X% of the effect of education on earnings is due to human capital accumulation" fails, because, as outlined by the steps in Section 2, identification requires far more from the data and theory in terms of interpretation of mediators, measurement of unmeasurable mediators, and generalizability.

The general approach this paper takes to the possibility of identification can be applied to other theoretical claims. This could be a fruitful path for distinguishing difficult empirical problems, which most empirical problems are, from effectively impossible ones, and outlining the exact conditions necessary for general identification. This same approach can show that not all difficult-to-study theoretical claims in economics are unresolvable, even ones that can be framed in a mediating-variables framework. Using another example from the economics of education, the difficulties of empirically estimating peer effects are well known. However, the claim "X% of the impact of attending a more selective school is because of exposure to higher-intelligence peers" can be tested empirically using the steps from Section 2, because the exposure to peers is observable and has a clearly interpretable connection with the stated theory. "Intelligence" is a theoretical concept like "ability", except that it has believable measurable proxies.

I claim in Section 5 that another empirical question that is difficult but answerable is how much of the returns to education are private, as opposed to external. This framing offers a means of organizing empirical results into a theoretical framework that accepts empirical answers. Understanding the returns to education as being separable into private and external returns has in the past been seen as a restatement of the human capital vs. signaling debate. But the potential for signaling to produce productivity improvements and external returns means that the analogy is flawed, and implies that the private/external distinction is actually more useful for making policy prescriptions and making sense of empirical data.

Human capital and signaling remain useful theoretical concepts, and the underlying explanation of education returns should naturally include both human capital and signaling. But any empirical tests of human capital or signaling-derived theories should rarely be understood as having theoretical implications for the models they are derived from. The primary theories about educational returns that generate testable hypotheses are not then updated and improved by the results of those tests.

Following a debate on human capital and signaling that has lasted for nearly fifty years without approaching resolution, the field would do well to reorient its attempts to explain the returns to education. Theoretical advancement in the economics of education would be improved by focusing on theoretical framings that are more responsive to the wealth of empirical results the field is capable of generating.

References

Acemoglu D, Angrist JD (2000) How Large are Human-Capital Externalities? Evidence from Compulsory Schooling Laws. NBER Macroeconomics Annual 15(2000):9–74, DOI 10.2307/3585383

Altonji JG, Pierret CR (1997) Employer Learning and the Signaling value of Education. In: Ohashi I, Tachibanaki T (eds) Industrial Relations, Incentives, and Employment, Macmillan, London, England

Altonji JG, Pierret CR (2001) Employer Learning and Statistical Discrimination. The Quarterly Journal of Economics 116(1):313–350, DOI 10.1162/003355301556329

Altonji JG, Kahn LB, Speer JD (2016) Cashier or Consultant? Entry Labor Market Conditions, Field of Study, and Career Success. Journal of Labor Economics 34(S1):S361–S401, DOI 10.1086/682938

Ambrose SA, Bridges MW, DiPietro M, Lovett MC, Norman MK (2010) How Learning Works: Seven Research-based Principles for Smart Teaching. John Wiley & Sons, New York City, NY

Arcidiacono P, Bayer P, Hizmo A (2010) Beyond Signaling and Human Capital: Education and the Revelation of Ability. American Economic Journal: Applied Economics 2(4):76–104, DOI 10.1257/app.2.4.76

Arcidiacono P, Aucejo E, Maurel A, Ransom T (2016) College Attrition and the Dynamics of Information Revelation, NBER Working Paper No. 22325

Arrow KJ (1973) Higher Education as a Filter. Journal of Public Economics 2:193–216, DOI 10.1007/s003550050148

Arteaga C (2018) The Effect of Human Capital on Earnings: Evidence from a Reform at Colombia's Top University. Journal of Public Economics 157:212–225, DOI 10.1016/j.jpubeco.2017.10.007

Arum R, Roksa J (2011) Academically Adrift: Limited Learning on College Campuses. University of Chicago Press, Chicago, IL

Aryal G, Bhuller M, Lange F (2019) Signaling and Employer Learning with Instruments, DOI 10.3386/w25885, NBER Working Paper No. 25885

Baker M, Gruber J, Milligan K (2015) Non-Cognitive Deficits and Young Adult Outcomes: The Long-Run Impacts of a Universal Child Care Program, NBER Working Paper No. 21571

Barnett SM, Ceci SJ (2002) When and Where Do We Apply What We Learn? A Taxonomy of Far Transfer. Psychological Bulletin 128(4):612–637, DOI 10.1037//0033-2909.128.4.612

Battu H, Belfield CR, Sloane PJ (2003) Human Capital Spill-Overs Within the Workplace. Oxford Bulletin of Economics and Statistics 65(5):575–594, DOI 10.1111/j.1468-0084.2003.00062.x

Bauer TK, Haisken-Denew JP (2001) Employer Learning and the Returns to Schooling. Labour Economics 8:161–180, DOI 10.1016/S0927-5371(01)00026-4

Becker GS (1964) Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education. National Bureau of Economic Research

Bedard K (2001) Human Capital versus Signaling Models: University Access and High School Dropouts. Journal of Political Economy 109(4):749–775, DOI 10.1086/322089

Bellemare MF, Bloem JR (2019) The Paper of How: Estimating Treatment Effects Using the Front-Door Criterion, Unpublished

Belman D, Heywood JS (1991) Sheepskin Effects in the Returns to Education: An Examination of Women and Minorities. The Review of Economics and Statistics 73(4):720–724, DOI 10.2307/2109413

Belman D, Heywood JS (1997) Sheepskin Effects by Cohort: Implications of Job Matching in a Signaling Model. Oxford Economic Papers 49(4):623–637, DOI 10.1093/oxfordjournals.oep.a028628

Bentsen KH, Munch JR, Schaur G (2019) Education spillovers Within the Workplace. Economics Letters 175:57–59, DOI 10.1016/j.econlet.2018.11.025

Bidner C (2014) A Spillover-based Theory of Credentialism. Canadian Journal of Economics/Revue canadienne d'économique 47(4):1387–1425, DOI 10.1111/caje.12110

Bingley P, Christensen K, Markwardt K (2015) Signaling and Productivity in the Private Financial Returns to Schooling, Unpublished

Bitzan JD (2009) Do Sheepskin Effects Help Explain Racial Earnings Differences? Economics of Education Review 28(6):759–766, DOI 10.1016/j.econedurev.2008.10.003

Bol T, Van De Werfhorst HG (2011) Signals and Closure by Degrees: The Education Effect Across 15 European Countries. Research in Social Stratification and Mobility 29(1):119–132, DOI 10.1016/j.rssm.2010.12.002

Botelho A, Pinto LC (2004) Students' Expectations of the Economic Returns to College Education: Results of a Controlled Experiment. Economics of Education Review 23(6):645–653, DOI 10.1016/j.econedurev.2004.03.005

Brunello G, Rocco L (2017) The Labour Market Effects of Academic and Vocational Education over the Life Cycle: Evidence from Two British Cohorts. Journal of Human Capital 11(1):106–166, DOI 10.1086/690234

Caicedo FV (2019) The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America. The Quarterly Journal of Economics 134(1):507–556, DOI 10.1093/qje/qjy024

Caplan B (2018) The Case Against Education: Why the Education System is a Waste of Time and Money. Princeton University Press, Princeton, NJ

Card D (1999) The Causal Effect of Education on Earnings. In: Ashenfelter OE, Card D (eds) Handbook of Labor Economics, vol 3, North-Holland, New York, chap 30, pp 1801 – 1863

Cardoso AR, Guimarães P, Portugal P, Reis H (2018) The Returns to Schooling Unveiled, IZA Discussion Paper No. 11419

Carneiro P, Ginja R (2014) Long Term Health and Behavioral Impacts of Compensatory Pre-School: Evidence from Head Start. American Economic Journal: Economic Policy 6(4):135–173, DOI 10.1257/pol.6.4.135

Chetty R, Friedman JN, Hilger N, Saez E, Schanzenbach DW, Yagan D (2011) How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star. The Quarterly Journal of Economics 126(4):1593–1660, DOI 10.3386/w16381

Chetty R, Friedman JN, Rockoff JE (2014) Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. American Economic Review 104(9):2633–2679, DOI 10.3386/w19423

Chingos MM, Peterson PE (2011) It's Easier to Pick a Good Teacher than to Train One: Familiar and New Results on the Correlates of Teacher Effectiveness. Economics of Education Review 30(3):449–465, DOI 10.1016/j.econedurev.2010.12.010

Chiswick BR (1973) Racial Discrimination in the Labor Market: A Test of Alternative Hypotheses. Journal of Political Economy 81(6):1330–1352, DOI 10.1086/260130

Ciccone A, Peri G (2006) Identifying Human-Capital Externalities: Theory with Applications. The Review of Economic Studies 73(2):381–412, DOI 10.1111/j.1467-937X.2006.00380.x

Clark D, Martorell P (2014) The Signaling Value of a High School Diploma. Journal of Political Economy 122(2):282–318, DOI 10.1086/675238

Crook TR, Todd SY, Combs JG, Woehr DJ, Ketchen DJ (2011) Does Human Capital Matter? A Meta-Analysis of the Relationship Between Human Capital and Firm Performance. Journal of Applied Psychology 96(3):443–456, DOI 10.1037/a0022147

Cunha F, Heckman JJ (2007) Identifying and Estimating the Distributions of Ex Post and Ex Ante Returns to Schooling. Labour Economics 14(6):870–893, DOI 10.1016/j.labeco.2007.06.002

Deming DJ (2014) Using School Choice Lotteries to Test Measures of School Effectiveness. American Economic Review 104(5):406–411, DOI 10.1257/aer.104.5.406

Eble A, Hu F (2016) The Importance of Educational Credentials: Schooling Decisions and Returns in Modern China, Unpublished

Fang H (2006) Disentangling the College Wage Premium: Estimating a Model with Endogenous Education Choices. International Economic Review 47(4):1151–1185, DOI 10.1111/j.1468-2354.2006.00409.x

Farber HS, Gibbons R (1996) Learning and Wage Dynamics. Quarterly Journal of Economics 111(4):1007–1047, DOI 10.2307/2946706

Flores-Lagunes A, Light A (2010) Interpreting Degree Effects in the Returns to Education. Journal of Human Resources 45(2):439–467

Frazis H (1993) Selection Bias and the Degree Effect. Journal of Human Resources 28(3):538–554, DOI 10.2307/146159

Fryer RG, Levitt SD (2004) Understanding the Black-White Test Score Gap in the First Two Years of School. The Review of Economics and Statistics 86(2):447–464, DOI 10.1162/003465304323031049

Gibson J (2000) Sheepskin Effects and the Returns to Education in New Zealand: Do they Differ by Ethnic Groups? New Zealand Economic Papers 34(2):201–220, DOI 10.1080/00779950009544323

Goldin CD, Katz LF (2009) The Race Between Education and Technology. Harvard University Press, Cambridge, MA

Goodman J (2019) The Labor of Division: Returns to Compulsory High School Math Coursework. Journal of Labor Economics 37(4):1141–1182, DOI 10.1086/703135

Graetz G (2017) Human Capital, Signaling, and Employer Learning: What Insights Do We Gain from Regression Discontinuity Designs?, IZA Working Paper No. 11125

Green DP, Ha SE, Bullock JG (2010) Enough Already about "Black Box" Experiments: Studying Mediation Is More Difficult than Most Scholars Suppose. The Annals of the American Academy of Political and Social Science 628(1):200–208

Gunderson S, Oreopoulos P (2010) Returns to Education in Developed Countries. In: Baker E, Peterson P, McGaw B (eds) International Encyclopedia of Education, 3rd edn, Elsevier, Oxford

Habermalz S (2011) The Speed of Employer Learning and Job Market Signalling Revisited. Applied Economics Letters 18(7):607–610, DOI 10.1080/13504851.2010.482514

Hanushek EA (2016) Will More Higher Education Improve Economic Growth? Oxford Review of Economic Policy 32(4):538–552, DOI 10.1093/oxrep/grw025

Haswell RH (2012) Methodologically Adrift. College Composition and Communication 63(3):487–491

Heckman JJ, Pinto R, Savelyev P (2013) Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes. American Economic Review 103(6):1–35, DOI 10.1257/aer.103.6.2052

Heckman JJ, Humphries JE, Veramendi G (2018) Returns to Education: The Causal Effects of Education on Earnings, Health and Smoking. Journal of Political Economy 126(S1):S197–S246, DOI 10.3386/w22291

Henderson DJ, Polachek SW, Wang L (2011) Heterogeneity in Schooling Rates of Return. Economics of Education Review 30(6):1202–1214, DOI 10.1016/j.econedurev.2011.05.002

Hopkins E (2012) Job Market Signaling of Relative Position, or Becker Married to Spence. Journal of the European Economic Association 10(2):290–322, DOI 10.1111/j.1542-4774.2010.01047.x

Hungerford T, Solon G (1987) Sheepskin Effects in the Returns to Education. The Review of Economics and Statistics 69(1):175–177, DOI 10.2307/1937919

Hussey A (2012) Human Capital Augmentation Versus the Signaling Value of MBA Education. Economics of Education Review 31(4):442–451, DOI 10.1016/j.econedurev.2011.12.004

Imai K, Keele L, Tingley D (2010) A General Approach to Causal Mediation Analysis. Psychological Methods 15(4):309–334, DOI 10.1037/a0020761

Jaeger DA, Page ME (1996) Degrees Matter: New Evidence on Sheepskin Effects in the Returns to Education. The Review of Economics and Statistics 78(4):733–740, DOI 10.2307/2109960

Jovanovic B (1979) Job Matching and the Theory of Turnover. Journal of Political Economy 87(5):972–990, DOI 10.1086/260808

Kaymak B (2012) Quantifying the Signaling Role of Education. In: Society of Labor Economists, p 40

Kremer M (1993) The O-Ring Theory of Economic Development. The Quarterly Journal of Economics 108(3):551–575, DOI 10.2307/2118400

Lang K, Kropp D (1986) Human Capital Versus Sorting: The Effects of Compulsory Attendance Laws. The Quarterly Journal of Economics 101(3):609–624, DOI 10.2307/1885699

Lange F (2007) The Speed of Employer Learning. Journal of Labor Economics 25(1):1–35, DOI 10.1086/508730

Lange F, Topel R (2006) The Social Value of Education and Human Capital. In: Hanushek E, Welch F (eds) Handbook of the Economics of Education, vol 1, North-Holland, pp 459–509

Layard R, Psacharopoulos G (1974) The Screening Hypothesis and the Returns to Education. Journal of Political Economy 82(5):985, DOI 10.1086/260251

Light A, McGee A (2015) Does Employer Learning Vary by Schooling Attainment? The Answer Depends on How Career Start Dates are Defined. Labour Economics 32:57–66, DOI 10.1016/j.labeco.2014.12.002

Lovenheim M, Turner S (2018) Economics of Education, 1st edn. Worth Publishers, New York

Machin S, Marie O, Vujic S (2011) The Crime Reducing Effect of Education. The Economic Journal 121(2011):463–484, DOI 10.1111/j.1468-0297.2011.02430.x.

Mankiw NG (2014) Principles of Microeconomics, seventh edn. Cengage Learning, Mason, OH

Manski CF (2004) Measuring Expectations. Econometrica 72(5):1329–1376, DOI 10.1111/j.1468-0262.2004.00537.x

Moretti E (2004) Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-sectional Data. Journal of Econometrics 121(1-2):175–212, DOI 10.1016/j.jeconom.2003.10.015

Morgan SN, Winship C (2014) Counterfactuals and Causal Inference. Cambridge University Press

Oreopoulos P, Salvanes KG (2011) Priceless: The Nonpecuniary Benefits of Schooling. Journal of Economic Perspectives 25(1):159–184, DOI 10.1257/jep.25.1.159

Page M (2010) Signaling in the Labor Market. In: Baker E, Peterson P, McGaw B (eds) International Encyclopedia of Education, 3rd edn, Elsevier, Oxford

Pascarella ET, Blaich C, Martin GL, Hanson JM (2011) How Robust are the Findings of Academically Adrift? Change: The Magazine of Higher Learning 43(3):20–24, DOI 10.3200/CHNG.40.4.50-54

Pearl J (2009) Causality, 2nd edn. Cambridge University Press, Cambridge, MA

Pischke JS (2007) The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Years. The Economic Journal 117(523):1216–1242, DOI 10.1111/j.1468-0297.2007.02080.x

Ritchie SJ, Tucker-Drob EM (2018) How Much Does Education Improve Intelligence? A Meta-Analysis. Psychological Science pp 1358–1369, DOI 10.1177/0956797618774253

Rivera LA (2011) Ivies, Extracurriculars, and Exclusion: Elite Employers' Use of Educational Credentials. Research in Social Stratification and Mobility 29(1):71–90, DOI 10.1016/j.rssm.2010.12.001

Rivera LA (2016) Pedigree: How Elite Students Get Jobs. Princeton University Press, Princeton, NJ

Sacerdote B (2001) Peer Effects with Random Assignment: Results for Dartmouth Roommates. The Quarterly Journal of Economics 116(2):681–704, DOI 10.1162/00335530151144131

Sala G, Aksayli ND, Tatlidil KS, Tatsumi T, Gondo Y, Gobet F (2019) Near and Far Transfer in Cognitive Training: A Second-Order Meta-Analysis. Collabra: Psychology 5(1):18, DOI 10.1525/collabra.203

Schultz TW (1963) The Economic Value of Education. Columbia University Press, New York, NY

Spence M (1973) Job Market Signaling. The Quarterly Journal of Economics 87(3):355, DOI 10.2307/1882010

Stiglitz JE (1975) The Theory of Screening, Education, and the Distribution of Income. American Economic Review 65(3):283–300, DOI 10.2307/1804834

Stiglitz JE, Weiss A (1994) Sorting Out the Differences Between Signaling and Screening Models. In: Bacharach MO, Dempster MA, Enos JL (eds) Mathematical Models in Economics, Oxford University Press, Oxford, England

Stinebrickner R, Stinebrickner TR (2014) A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout. Review of Economic Studies 81(1):426–472, DOI 10.3386/w19165

Topel R (1999) Labor Markets and Economic Growth. In: Ashenfelter O, Card D (eds) Handbook of Labor Economics, vol 3, 1st edn, Elsevier, chap 44, pp 2943–2984

Trostel P, Walker I, Woolley P (2002) Estimates of the Economic Return to Schooling for 28 Countries. Labour Economics 9:1–16, DOI 10.1016/S0927-5371(01)00052-5

Tyler JH, Murnane RJ, Willett JB (2000) Estimating the Labor Market Signaling Value of the GED. The Quarterly Journal of Economics 115(2):431–468, DOI 10.1162/003355300554818

Valero A, Van Reenen J (2019) The Economic Impact of Universities: Evidence from Across the Globe. Economics of Education Review 68:53–67, DOI 10.1016/j.econedurev.2018.09.001

van der Meer PH (2011) Educational Credentials and External Effects: A Test for the Netherlands. Research in Social Stratification and Mobility 29(1):107–118, DOI 10.1016/j.rssm.2010.12.003

Van Der Velden R, Bijlsma I (2016) College Wage Premiums and Skills: A Cross-country Analysis. Oxford Review of Economic Policy 32(4):497–513, DOI 10.1093/oxrep/grw027

Waldman M (2016) The Dual avenues of Labor Market Signaling. Labour Economics 41:120–134, DOI 10.1016/j.labeco.2016.05.001

Weiss A (1995) Human Capital vs. Signalling Explanations of Wages. Journal of Economic Perspectives 9(4):133–154, DOI 10.1257/jep.9.4.133

West MR, Kraft MA, Finn AS, Martin RE, Duckworth AL, Gabrieli CFO, Gabrieli JDE (2016) Promise and Paradox: Measuring Students' Non-Cognitive Skills and the Impact of Schooling. Educational Evaluation and Policy Analysis 38(1):148–170, DOI 10.3102/0162373715597298

Wolpin KI (1977) Education and Screening. American Economic Review 67(5):949–958

7 Figures

Fig. 1 General Causal Model of the Effect of Education on the Labor Market

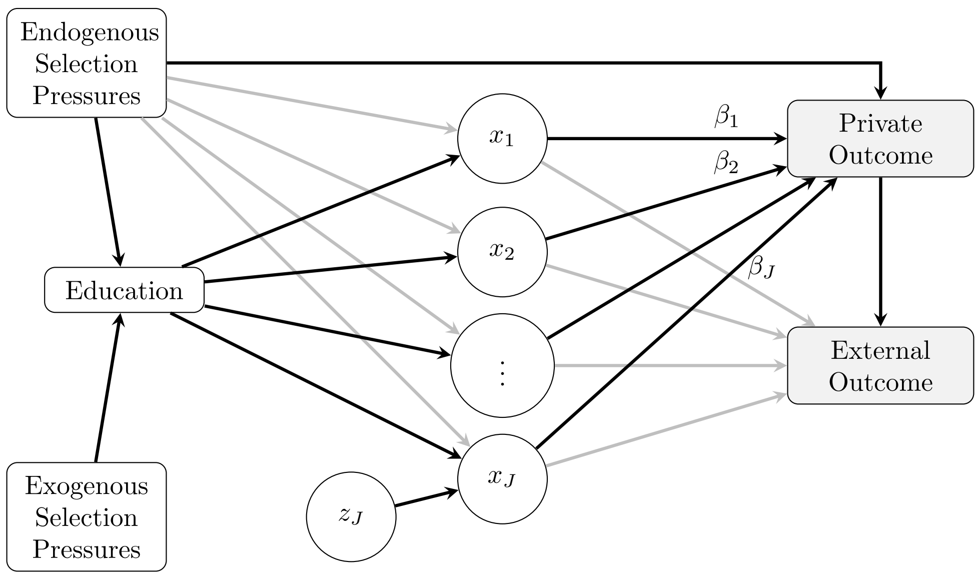


Fig. 2 Basic Employer Learning Model

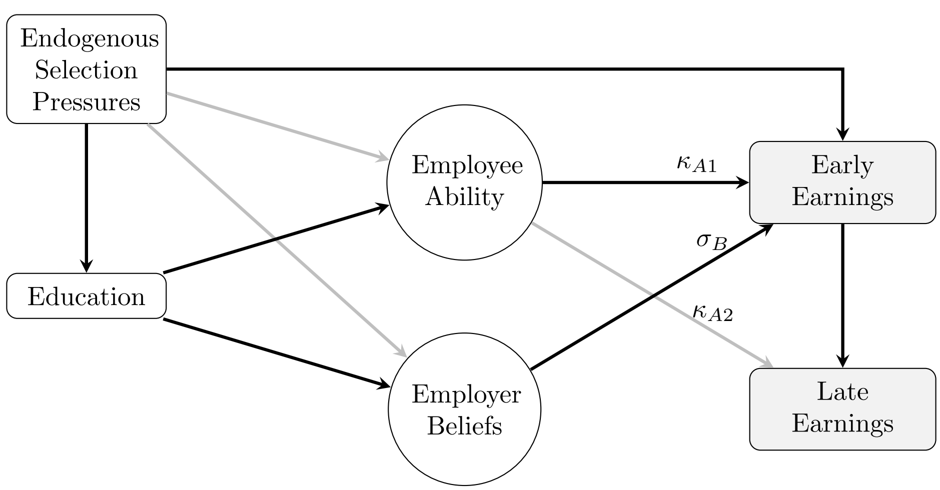


Fig. 3 Simple Theoretical Model of the Returns to Education

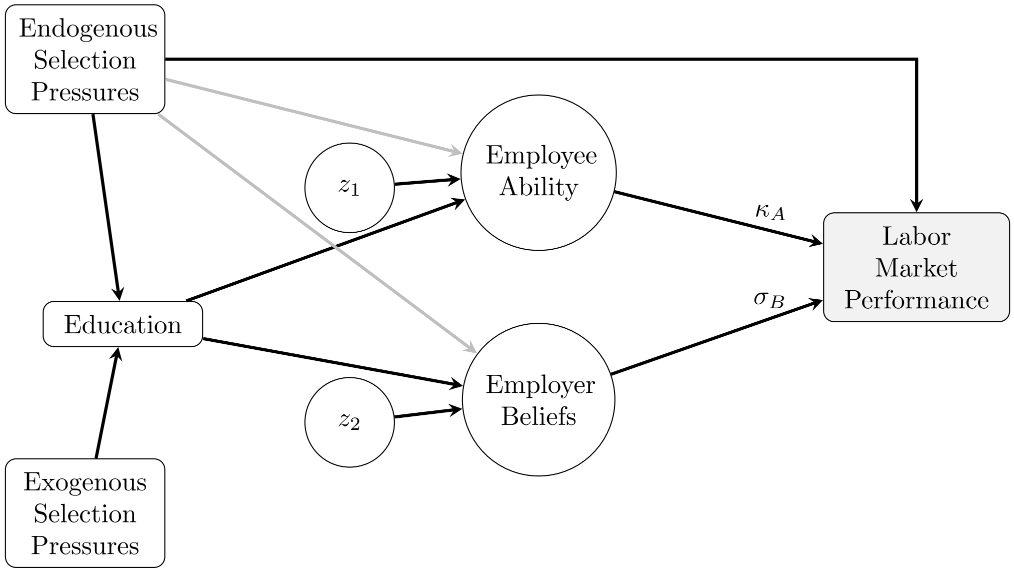
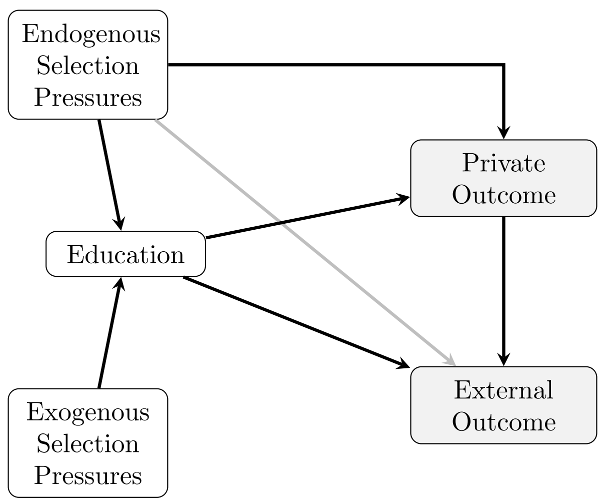


Fig. 4 The Returns to Education Without Mediating Variables



A The Effect of Education on the Labor Market in Potential Outcomes Notation

This appendix draws from Imai et al. (2010), albeit with the use of superscript counterfactual notation rather than function notation.

We are interested in the causal effect of education Ed_i on some outcome Y_i for individual i. For simplicity, assume that the margin of interest compares $Ed_i = 1$ against $Ed_i = 0$. Let Y_i^C be the possibly-counterfactual outcome for individual i under the condition C.

The average causal effect of education is

$$E(Y_i^{Ed_i=1} - Y_i^{Ed_i=0}) \equiv B \equiv \Sigma + K + \Omega \tag{3}$$

where Σ, K, and Ω are the parts of the education return attributable to signaling, human capital, and other explanations, respectively.

In observed data, either Y_i^1 or Y_i^0 is missing, such that the raw relationship between Y_i and Ed_i does not identify the effect of education:

$$E(Y_i|Ed_i=1) - E(Y_i|Ed_i=0) \neq E(Y_i^{Ed_i=1} - Y_i^{Ed_i=0}) \tag{4}$$

However, for the purposes of this paper, we will assume that the causal effect of education can be identified. For simplicity, assume that we do this by conditioning on a set of controls W_i such that

$$E(Y_i|Ed_i=1,W_i) - E(Y_i|Ed_i=0,W_i) = E(Y_i^{Ed_i=1} - Y_i^{Ed_i=0}) \tag{5}$$

The effect of Ed_i on Y_i is fully mediated by a set of mediating variables $x_1, x_2, ..., x_J$. For simplicity of notation, assume that each of these mediators is binary. We have

$$E(Y_i^{Ed_i=1} - Y_i^{Ed_i=0}|x_1, x_2, ..., x_J) = 0 \tag{6}$$

$$E(Y_i^{x_{ij}=1} - Y_i^{x_{ij}=0}) \equiv \sigma_j + \kappa_j + \omega_j + \epsilon_j \ \forall \ j \in \{1, ..., J\} \tag{7}$$

where σ_j, κ_j, ω_j, and ϵ_j are the parts of the effect of x_j on Y that are attributable to signaling, human capital, other educational explanations, and other non-educational explanations. $\Sigma = \sum_j \sigma_j$, $K = \sum_j \kappa_j$, and $\Omega = \sum_j \omega_j$. Isolating the part of each mediator that is driven by education excludes non-educational explanations such that

$$E\left(Y_i^{Ed_i=1, x_{ij}^{Ed_i=1}} - Y_i^{Ed_i=0, x_{ij}^{Ed_i=0}}\right) = \sigma_j + \kappa_j + \omega_j \ \forall \ j \in \{1, ..., J\} \tag{8}$$

and for most points of discussion assume also that $\omega_j = 0 \ \forall \ j$.

Divide the set $x_1, x_2, ..., x_J$ into the subset χ_σ for which $\kappa_j = 0 \ \forall \ x_j \in \chi_\sigma$, χ_κ for which $\sigma_j = 0 \ \forall \ x_j \in \chi_\kappa$, and χ^C for which $\kappa_j \neq 0$ and $\sigma_j \neq 0 \ \forall \ x_j \in \chi^C$. For an element of χ^C, we could identify σ_j or κ_j separately if there were a way to vary the mediator while holding the other explanation constant, or by varying only the part of the mediator associated with one explanation, expressed as

$$E\left(Y_i^{Ed_i=1, x_{ij}^{Ed_i=1}|\kappa_j} - Y_i^{Ed_i=0, x_{ij}^{Ed_i=0}|\kappa_j}\right) = \sigma_j \tag{9}$$

The share of the educational return that is due to, for example, signaling, is defined as $\Sigma/B = (B-K)/B = \Sigma/(\Sigma+K)$ under the assumption that $\Omega = 0$. Identifying this share requires following the steps given in Section 2, which includes the tasks of identifying at least some of:

- The total effect of education B, which is $E\left(Y_i^{Ed_i=1} - Y_i^{Ed_i=0}\right)$
- The part of the effect of education that goes through χ_σ, which is $\sum_{j|x_j\in\chi_\sigma} E\left(Y_i^{Ed_i=1, x_{ij}^{Ed_i=1}} - Y_i^{Ed_i=0, x_{ij}^{Ed_i=0}}\right)$
- The part of the effect of education that goes through χ_κ, which is $\sum_{j|x_j\in\chi_\kappa} E\left(Y_i^{Ed_i=1, x_{ij}^{Ed_i=1}} - Y_i^{Ed_i=0, x_{ij}^{Ed_i=0}}\right)$
- The parts of the effect of education that go through χ^C that are signaling-related, which is $\sum_{j|x_j\in\chi^C} E\left(Y_i^{Ed_i=1, x_{ij}^{Ed_i=1}|\kappa_j} - Y_i^{Ed_i=0, x_{ij}^{Ed_i=0}|\kappa_j}\right)$

- The parts of the effect of education that go through χ^C that are human capital-related, which is $\sum_{j|x_j \in \chi^C} E\left(Y_i^{Ed_i=1,x_{ij}^{Ed_i=1}|\sigma_j} - Y_i^{Ed_i=0,x_{ij}^{Ed_i=0}|\sigma_j}\right)$

with the particular elements that must be identified varying depending on which explanation is being identified and what strategy is being taken for following the steps in Section 2.