A Spillover-Based Theory of Credentialism *

Chris Bidner†
University of British Columbia
Version: November 18, 2007

Abstract

This paper studies a large economy in which workers' productivities are influenced by the skills of their coworkers. Each worker's skill is the result of an ability-augmenting investment that is made prior to matching with a coworker. These skills are 'soft' (unquantifiable) characteristics of workers. In contrast, investments are 'hard' characteristics. This feature gives investments a credential quality, since a worker can use their investment attract desirable coworkers. The main result of the model is that, despite the positive externality, there is over-investment in equilibrium. Apart from providing a rationalization for credentialism that overcomes some recent criticisms of signaling models, the analysis provides insights into the relationship between spillovers and productivity, welfare, and inequality. The model offers a different interpretation of returns to education, and demonstrates how modeling spillovers in this way produces conclusions that are dramatically different from standard treatments. I extend the model to a simple dynamic setting and show that the qualitative conclusions of the static model are unaffected when utility is non-transferable, and that efficiency is restored as workers become patient when utility is transferable.

*Job Market Paper. Thanks to participants in the Development, Institutions, Environment, and Trade lunch at UBC, the 2007 Canadian Economics Association Meetings in Halifax, and the 2006 TARGET UBC-SFU student workshop. Thanks are especially due to Ken Jackson, Patrick Francois, and Mike Peters for valuable discussions. All errors are mine and any further comments are most welcome.

†PhD Candidate, Department of Economics, University of British Columbia, 997 - 1873 East Mall Vancouver, BC Canada, V6T 1Z1. email: cbidner@interchange.ubc.ca, web: grad.econ.ubc.ca/cbidner

1
1 INTRODUCTION

A society’s prosperity depends in large part upon how productive its members are and therefore on its capacity to provide individuals with suitable incentives to enhance their skills. This paper studies such incentives when ‘interactive’ aspects of production are central: agents’ productivities depend not only on their skills, but also on the skills of those who they work with. As I argue below, such an interactive setting is becoming an increasingly relevant feature of the modern workplace. The focus of the paper is the relationship between skill spillovers in the workplace and the phenomenon of credentialism. I define credentialism as the tendency for individuals to be motivated, at least in part, to engage in an activity (e.g. higher education) because of the credential it offers (e.g. a degree), as opposed to the intrinsic benefit associated with the activity (e.g. learning). Despite the prominent perception that this phenomenon exists, the standard explanation for it, at least among economists, suffers a number of problems. This paper offers a rationalization of credentialism that avoids such problems, and thereby restores a theoretical underpinning for the phenomenon. A solid understanding of the phenomenon is of great policy relevance since the social benefit of education is weakened in the presence of credential-motivated attainment.

Economists typically comprehend credentialism through models of signaling and/or screening, where the essential point is that education acts as a credential because it changes employer beliefs about a worker’s unobserved productivity. However, in order for signaling models to be convincing, it must be that employers take a relatively long time to learn the true productivity of their employees. A recent criticism of signaling models asserts that this fundamental requirement is, at best, dubious. As Gary Becker puts it:

The signaling interpretation of the benefits of going to college originated in the 1970’s and had a run of a couple of decades, but is seldom mentioned any longer. I believe it declined because economists began to realize that companies rather quickly discover the productivity of employees who went to college, whether a Harvard or a University of Phoenix. Before long, their pay adjusts to their productivity rather than to their education credentials.

Not only is this a compelling criticism on intuitive grounds, it also receives recent empirical support. If the criticism is accepted, then one clear possibility is that credentialism is an illusion. Indeed, in

---

1For instance, there is credentialism inherent in signaling models because workers are paid on the basis of their observed education credential, independent of their true productivity. There is no credentialism inherent in human capital models because workers are paid according to their productivity, regardless of any investment credentials.

2That there is a widely-held view that credentialism exists is evidenced by the large empirical literature devoted to evaluating the magnitude of signaling relative to the human capital explanation of the link between education and wages (see Weiss (1995), Ferrer and Riddell (2002), and Lange and Topel (2006) for surveys). The perception also exists outside Economics. Sociologists have discussed the phenomenon at length (e.g. see Labaree (1997), Brown (1995), Collins (1979) and Berg (1970)) and is also prominent in the popular press (e.g. Jacobs (2004)).


4Taken from a January 29, 2006, entry in the Becker-Posner Blog, which can be found at: http://www.becker-posner-blog.com/archives/2006/01/on_forprofit_co.html.

5Recent work by Lange (2007) and Lange and Topel (2006) estimate the speed of employer learning and come to
light of criticisms of this nature, the recent dramatic increase in educational attainment experienced in many countries has been widely interpreted as a natural response to the changing demands for skill in the modern workplace. The second possibility is simply that we need a different approach to understanding credentialism. The model developed here offers such an alternative. The basic ingredients of this explanation are based upon salient features of the modern workplace.

The first observation is intimately tied to the above criticism of signaling models: worker productivity is relatively easy to reward. Lemieux, MacLeod, and Parent (2007) document the growing incidence of explicit performance-based pay in the U.S. and argue that technological change, especially in regards to monitoring and reporting technology, has made this feasible. The practice is even more widespread than suggested since performance may be implicitly rewarded by continued employment (as in Academia). Even when objective measures of performance are unavailable, subjective performance evaluations often offer a suitable substitute. To make this feature as stark as possible, I assume that workers are simply paid their output. This assumption addresses the above criticism of signaling models since paying workers their realized output is equivalent to immediate employer learning.

The central feature of the model is that a worker’s productivity is influenced by who they work with. Modern work practices have greatly increased the scope for such spillovers to arise. From a long-term perspective, the fact that workers have moved off the farm and out of the workshop and into the factory and office building means that physical proximity to other workers has increased. Trends in Human Resource Management, such as those described in Ichniewski and Shaw (2003), stress the importance of “pay-for-performance plans like gain-sharing or profit-sharing, problem-solving teams, broadly defined jobs, cross-training for multiple jobs, employment security policies and labor-management communication procedures”. Many of these practices enhance the interconnectedness of employees. Finally, the introduction of computers in the workplace has changed the nature of tasks being performed. Autor, Levy, and Murnane (2003) and Levy and Murnane (2004) argue that computers can not easily perform non-routine tasks, and document the growing importance of tasks related to expert thinking (“solving problems for which there are no rule-based solutions”), and complex communication (“interacting with humans to acquire information, to explain it, or to persuade others of its implications for action”). A worker’s ability to perform such tasks depends on the skills of their coworkers - expert thinking is often a collaborative process and the efficacy of complex communication relies on the skills of both sender and receiver.

I assume, in the human capital tradition, that a worker’s skill is the result of a costly investment that augments their natural ability. In order to introduce spillovers, a worker’s productivity depends upon their skill as well as upon the skill of a coworker. If we define a firm to simply be a pair of the general conclusion that “employers learn quickly”. For instance, Lange (2007) concludes that signaling accounts for around 15% of the gain from an additional year of schooling under his preferred specification.

6For discussion and a theoretical treatment, see Baker, Gibbons and Murphy (1994), and for an empirical demonstration of the relevance of subjective performance evaluations, see Lazear (1999).

7For example, Gant, Ichniewski, and Shaw (2002) provide evidence that worker productivity is improved following the introduction of innovative work practices because of stronger social capital developed between workers. Drago and Garvey (1998) find that ‘helping effort’ is more readily expended by workers engaged in a large variety of tasks.
workers, then we can say that spillovers are ‘local’ in the sense that they occur within the boundaries of a firm. It is important to note however that there is no sense in which the ‘firm’ internalizes the skill externality. This is because of the assumption that workers are paid according to their productivity (and not their skill, as in Kremer (1993) for example).

In compiling a list of ‘new basic skills’ - skills that are valued in the modern workplace - Murnane and Levy (1996) write:

A surprise in the list of New Basic Skills is the importance of soft skills. These skills are called “soft” because they are not easily measured on standardized tests. In reality, there is nothing soft about them. Today more than ever, good firms expect employees to raise performance continually by learning from each other through written and oral communication and by group problem solving.

To capture this feature, skills are modeled as unquantifiable qualities such as creativity, capacity for working in teams, self-reliance, etc., for which there is no natural metric. The significance of this is that it is infeasible for workers to meet each other on the basis of skills. In contrast, I assume that investment is quantifiable (e.g. years of schooling, courses taken, grades, institution attended, etc.). This means that workers are able to use their investment in the ‘competition’ for desirable coworkers in the matching market, and thereby provides the rationale behind credential-minded educational attainment.

The model is clearly related to papers that model skill spillovers. Unlike Lucas (1988) and Moretti (2004), I do not model spillovers as a ‘global’ phenomenon in which some measure of the average skill in an area improves the productivity of everyone in that area. Instead, I model spillovers arising as a result of ‘local’ interactions between workers. I do not believe that this approach does any injustice to the motivation underlying those papers, yet I show that this approach leads to dramatically different conclusions. For example, policy prescriptions and most comparative statics are completely reversed.

The literature on learning in cities, including Jovanovic and Roy (1989), Glaeser (1999), and many others, takes this local aspect of spillovers more seriously. These papers typically model local interaction by assuming that agents meet in a random fashion and then enter into some kind of exchange. In contrast, this paper is interested in the consequences of the fact that individuals are able to take actions that influence who they interact with. It therefore differs from models of search and matching, such as Shimer and Smith (2000), Smith (2006), and Burdett and Coles (1997), which also assume random matching. Here, spillovers alter incentives to invest precisely because interaction is not random.

The feature that agents take actions mindful of the fact that it will affect their matching prospects

---

8There are no explicit firms in the model, but we can think of ‘workers meeting each other’ as the reduced-form of a process in which firms post vacancy notices requiring applicants possess particular quantifiable characteristics. By their ‘soft’ nature, firms could not meaningfully use skills for this purpose.

9The distinction between ‘global’ and ‘local’ does not arise in Lucas (1988) because he studies an economy with a representative agent. This assumption allows for a sharper focus on the growth dynamics.
is prominent in the literature on premarital investment, e.g. Peters (2004b), Peters and Siow (2002), Cole, Mailath, and Postelwaite (2001). In contrast to these papers, I allow the desirability of an agent as a match partner to depend on more than just their investment: in particular, an agent’s desirability will also depend on their type (ability). This is also the primary difference between this model and Kremer’s O-Ring Theory (Kremer (1993)).

If we think of a ‘firm’ as being the shell that houses groups of coworkers, then firms are ex-ante homogeneous in the model. However workers are not indifferent to which firm they work in because firms are differentiated by the skill of the coworkers they house. Hopkins (2005) analyzes a model in which firms house a single worker but firms themselves are heterogeneous. The main qualitative difference to the present model is that the heterogeneity of firms is exogenous in that model. In contrast, firms are endogenously differentiated in the present model since the skills of their workers are endogenously determined by the workers’ optimal investments. I assume that workers interact in pairs, making the analysis necessarily two-sided.  

The group of models that are most similar to the model presented here are those of two-sided signaling; including Hoppe, Moldovanu, and Sela (2005), Rege (2007), and Damiano and Li (2007). These papers share the general feature that agents are matched in a positive assortative manner on the basis of some unproductive investment made prior to matching. The model introduced in this paper assumes that this investment is productive. Although this distinction is largely trivial in standard signaling models based on Spence (1973), the distinction is not so trivial in the two-sided case considered here. The first reason for this is that the some key features of the environment, including the type of equilibrium employed, needs to be altered in order to produce separating equilibria. I take a competitive approach in that agents optimize taking as given a market return function that specifies the expected skill of coworker that will arise from the matching process given a particular investment.

The second important distinction is that when investment is a pure waste, there is clearly over-investment in any equilibrium in which any agent makes a positive investment. The issue is not so clear once investments are productive, since an agent’s investment confers a positive externality upon their partner. If partners were fixed, then the conclusion from this observation is that there is too little investment. But partners are not fixed, and in equilibrium more investment allows an agent to obtain a more desirable partner. I show that this latter aspect dominates and consequently that there can never be under-investment in equilibrium. In fact the lowest type invests efficiently, whereas all other types invest too much. Thus, not only does the model rationalize why individuals are concerned with the credentials offered by an education, it also justifies the conclusion that such a concern encourages excessive attainment. Interestingly, the existence of positive skill spillovers does not, on its own, justify the use of subsidies.

The third significant difference is that models with unproductive investment require that inter-

\footnote{A model in which heterogeneity is generated due to ‘investment’ behavior is presented in Cole, Mailath, and Postelwaite (1995). In that model the ‘investment’ is purely a waste and therefore more investment actually detracts from one’s quality, all else equal. In addition, that model is one-sided.}
action be complementary (at least when the cost of investment is independent of type). This generates a trade-off: investment-based matching involves wasted resources, but improves the efficiency of the matching pattern. Whilst this trade-off is interesting, it is a direct consequence of complementary interaction. By allowing for productive investment, I am able to relax the assumption of complementary interaction and thereby remove the necessity of the trade-off. However, a new trade-off emerges: investment-based matching generates too much investment in equilibrium whereas random matching generates too little investment.

This paper differs from all of the above studies in that it is concerned with the consequences of greater spillovers. Such an exercise is of little interest in models of two-sided unproductive investment because aggregate outcomes, such as productivity and inequality, are fixed by the exogenous distribution of types and are therefore insensitive to the degree of spillovers. By considering productive investments, I identify a novel channel through which technological change increases productivity: new technologies that embody spillovers (e.g., communication technologies, including computers) raise productivity because they provide incentives for workers to increase their skill-enhancing investments, not because the technologies enhance worker skills per se.

Spillovers between workers are argued to be a crucial force underlying credentialist behaviour. This is because investments are efficient in the absence of spillovers and the degree of over-investment monotonically increases in the degree of spillovers. The fact that credentialism relies on spillovers demonstrates how ‘education externalities’ and ‘credentialism’ are structurally related, and need not represent separate phenomena.

A second crucial element of the environment is that spillovers occur locally (within worker-coworker pairs) and that investment influences workers’ matching prospects. I show how more standard approaches to modeling spillovers produce dramatically different conclusions to those produced here. For instance, it matters that there is heterogeneity of types: the results produced are necessarily absent in models that use a representative agent. Further, it matters that the spillovers that a worker is exposed to is sensitive to their investment: the central comparative static results of the model are in direct contrast to those of models in which spillovers accrue to the population at large, or in which interaction is random.

The model suggests that care needs to be taken in interpreting observed returns to education. Unlike both signaling and human capital models, education raises a worker’s productivity even when education has no capacity to raise skills. The reason is that a higher education enhances matching prospects and raises productivity via spillovers. This implies that even if the researcher had individual-level data on ability, productivity and education, the resulting implied private effect

---

11This feature is desirable as it overcomes another common criticism of signaling - that costs are lower for higher types. This is not at all obvious if costs reflect opportunity costs since higher types would reasonably be forgoing better alternatives (Weiss (1995)).

12In this context, ‘interaction is complementary’ just means that marginal product of one’s own characteristic is increasing in the level of one’s partner’s characteristic (supermodularity). Equilibria will exist when interaction is only weakly complementary (the marginal product is independent of one’s partner), but these are ‘knife-edge’ equilibria in which all agents are indifferent to all equilibrium investments.

13Unproductive education raises a worker’s wage in signaling models, but not their actual productivity.
of education on productivity would over-state the true social return (even though education has a positive external effect on the coworker). The fact that spillovers are local means that the true social return will generally not be uncovered by including controls for region-level aggregate education as is typically done in the empirical literature.

Finally, I examine the interpretation that pre-match investments represent ‘signaling’ by considering some simple dynamic extensions in which workers can reject a proposed match after immediately observing their partner’s skill. The main conclusions from these exercises are that the analysis is qualitatively unaltered when utility is non-transferable. When utility is transferable, I show that changes in workers’ discount factors allows us to endogenize the degree of spillovers, and, in particular, equilibrium investment becomes efficient as the discount factor approaches unity.

The model is presented in Section 2. After laying out the fundamentals, the equilibrium concept is introduced. A graphical approach is taken to illustrate equilibria. The qualitative features of the unique separating equilibrium is then analyzed in Section 3. Some further results are produced within the context of an illustration provided in Section 4. Some extensions to the base model are introduced in Section 5 before conclusions are drawn in Section 6.

2 Model

2.1 Fundamentals

The economy is populated with a unit measure of workers indexed by \( i \in [0, 1] \). Each worker is endowed with an ability, \( \theta_i \in \Theta \equiv [\underline{\theta}, \overline{\theta}] \). The distribution of abilities is common knowledge and is given by \( F \), with an associated density of \( f \), where it is assumed that \( 0 < f(\theta) < \infty \) for \( \theta \in \Theta \).

Agents have the opportunity to make an investment, \( x \geq 0 \), at a cost of \( c(x) \). I assume that \( c : \mathbb{R}_+ \to \mathbb{R}_+ \) is differentiable, convex and strictly increasing. If agent \( i \) makes an investment of \( x \), then they produce a skill of

\[
    s_i = s(x, \theta_i) = \theta_i \cdot g(x),
\]

where \( g : \mathbb{R}_+ \to \mathbb{R}_+ \) is differentiable, concave and strictly increasing. For the problem to be well-defined, I assume that for each \( \theta \in \Theta \) there exists a unique, interior, finite solution to the problem:

\[
    \max_x s(x, \theta) - c(x) \text{ s.t. } x \geq 0.
\]

The fact that \( s \) is multiplicatively separable is not qualitatively important. The important assumption is that the marginal benefit of investment is higher for higher ability workers: \( s_{x\theta} > 0 \).

The productivity with which a worker produces is influenced by the skill of their coworker. If worker \( i \) has coworker \( j \), then worker \( i \)'s productivity is given by \( y_i = y(s_i, s_j) \), where \( y \) is strictly increasing in both of its arguments. In particular, I assume a simple additively separable specification:

\[
    y(s, s') = \phi_1 \cdot s + \phi_2 \cdot s'.
\]

Apart from being simple, this specification has a number of advantages. First, it allows for an easier comparison with similar papers (Peters and Siow (2002)). Second, it highlights the point that, unlike models of unproductive two-sided signaling, complementary interaction is not necessary for the results. In any case, the model easily accommodates complementarity.
interaction (see the Appendix for an elaboration). Since I am interested in spillovers, I introduce a spillover parameter, $\phi \equiv \phi_2/(\phi_1 + \phi_2)$. By normalizing $\phi_1 + \phi_2 = 1$, this allows me to express $y$ in terms of the spillover parameter alone:

$$y(s, s'; \phi) = (1 - \phi) \cdot s + \phi \cdot s'.$$

The benefit of introducing spillovers in this way is that the parameterization is ‘neutral’ in the following sense. If we were to fix $s_i$ for each worker and the assignment of workers to coworkers, then changes in $\phi$ have no effect on pair-level output (and therefore on aggregate output). Furthermore, if it happened to be the case that $s_i = s_i'$ (each worker is paired with an equally skilled coworker), then changes in $\phi$ have no effect on individual-level output. Thus $\phi$ is constructed in such a way that any effect of changes in $\phi$ arise only through changes in incentives to invest.

Once workers have made their investment, they enter a frictionless matching market in which participants are matched on the basis of their investment. If a worker of skill $s$ leaves the matching market with a partner of skill $s'$, the worker produces and consumes an output given by $y(s, s')$. The worker’s total payoff is

$$u(x, \theta, s') \equiv y(s(x, \theta), s') - c(x).$$

A matching market equilibrium is characterized by an assignment of workers to other workers, where the assignment is made on the basis of workers’ investments. Such an assignment is characterized by a rule, $m(x)$, which says that workers that have an investment of $x$ are to be matched with workers that have an investment of $m(x)$.

The assignment rule is required to be feasible. For all $A \subseteq \mathbb{R}$, define:

$$M(A) \equiv \{ x \mid m(x) = a \text{ for some } a \in A \}$$

An assignment rule is feasible if the measure of agents for which $x_i \in M(A)$ is no less than the measure of agents for which $x_i \in A$. In words, if we take some group of workers - all those that make investments in $A$ - and consider the group of workers that they are supposed to be matched with - workers that invest in $M(A)$ - then it must be that there are at least as many workers in the latter group as in the former group.

The assignment rule is also required to be stable. A matching rule is stable if we cannot find two non-paired workers such that both would prefer to be matched together (at least one strictly) to remaining with their assigned partner. To describe these preferences, note that workers do not care about their partner’s investment per se, but rather their skill. Worker $i$’s evaluation of worker

---

14In the Appendix I show that if the productivity function takes a CES form: $y(s, s') = \left[ (1 - \phi) \cdot s^\rho + \phi \cdot s'^\rho \right]^{1/\rho}$ where $\rho \in (-\infty, 1]$, then the associated separating equilibrium is independent of the substitution parameter, $\rho$. Since the additively separable specification is the special case in which $\rho = 1$, it follows that the equilibrium calculated below will also be the equilibrium for all other values of $\rho$ including Cobb-Douglas ($\rho = 0$) and, in the limit, Leontief ($\rho \to -\infty$).

15In general, this assignment rule could be stochastic. That is, a worker that invests $x$ is given a lottery over workers that they are to be matched with, $M(x)$. As will become clear, the non-stochastic version is sufficient.
j’s skill will of course in general depend on worker j’s investment. Given this, let \( \hat{\mu}(x') \) represent workers’ beliefs about the skill of a worker that has a visible investment of \( x' \). Given these beliefs, preferences over potential partners - who are differentiated by their investment \( x' \) - are described by the utility function: \( u(x, \theta, \hat{\mu}(x')) \).

Beliefs are required to be consistent with equilibrium behaviour. In particular, the following rational expectations condition places a restriction on beliefs:

**Definition 1 (Rational Expectations)** Beliefs satisfy rational expectations if:

\[
\hat{\mu}(x) = E \left[ s(x(\theta'), \theta') \mid x(\theta') = x \right],
\]

for all \( \theta \in \Theta \).

Note that this condition only restricts beliefs when evaluated at investment levels that arise in equilibrium.

Given beliefs and the matching rule, a worker’s total payoff from investing \( x \) is therefore given by \( u(x, \theta, \mu(x)) \), where \( \mu(x) \equiv \hat{\mu}(m(x)) \). The \( \mu \) function acts as a return function in the sense that it describes the coworker skill that can be expected given an investment of \( x \) and can be thought of as a Hedonic return (Peters and Siow (2002)). I impose the following ‘free-entry’ condition upon this function.

**Definition 2 (Free Entry)** Let \( \Delta(\Theta) \) be the set of probability distributions defined on \( \Theta \). A return function, \( \mu \), satisfies free entry, if for all \( x \in \mathbb{R}_+ \) there exists a \( \pi \in \Delta(\Theta) \) such that

\[
\mu(x) \geq \int_{\Theta} s(x, \theta) d\pi(\theta).
\]

The condition reflects free entry in the following sense. Consider a ‘market-maker’ that announces some investment level, \( x \), and invites any agent that has made this investment to participate in a specially organized sub-market (perhaps for a fee) in which participants are randomly matched. Agents know that their partner will have also made an investment of \( x \), but, ex-ante, do not know the ability of their eventual partner. If the equilibrium return function does not satisfy free entry, then there exists an off-equilibrium investment level such that if a market-maker were to establish a sub-market requiring this investment, then some workers would strictly prefer to enter the sub-market regardless of the beliefs they hold about their eventual partner’s ability. Such workers would be willing to pay a positive fee for this service and the market-maker makes a positive profit. Free entry of such market-makers makes such a situation impossible.\(^{16}\)

Finally I require that agents are investing optimally given the equilibrium return function, \( \mu \). That is, an agent of type \( \theta \) is investing optimally if \( x(\theta) \in \arg\max_x \ u(x, \theta, \mu(x)) \). Given these above conditions, an equilibrium can be defined as follows.

**Definition 3 (Equilibrium)** An equilibrium is an investment function, \( x(\theta) \), a matching rule, \( m(x) \), and beliefs, \( \hat{\mu}(x) \), such that:

\(^{16}\)It may be important to note that the condition is not designed in order to restrict off-equilibrium beliefs to reflect the reality of some underlying matching mechanism, but, rather, to ensure that the equilibrium is robust to the existence of (unmodeled) competitive market-makers.
1. Given \(x(\theta)\) and \(\tilde{\mu}(x)\), the matching rule is stable and feasible.

2. Given \(x(\theta)\), beliefs satisfy rational expectations.

3. Given \(\tilde{\mu}(x)\) and \(m(x)\), the equilibrium return function, \(\mu(x) \equiv \tilde{\mu}(m(x))\), satisfies free entry.

4. Given \(\mu(x)\), all worker are investing optimally.

The following Lemma allows us to simplify the definition of equilibrium.

**Lemma 1** There is positive assortative matching (perfect segregation) in equilibrium. That is, \(m(x) = x\).

**Proof** See Appendix.

That is, in equilibrium, workers are randomly matched with another worker that has made the same investment.\(^{17}\) Importantly, the positive assortative matching is an equilibrium outcome and not an inbuilt feature of the matching game (as it is in Peters (2004b) and Hoppe, Moldovanu, and Sela (2005), for example). Given perfect segregation, the following simpler definition of equilibrium will suffice.

**Definition 4 (Equilibrium’)** An equilibrium is an investment function, \(x(\theta)\) and a return function, \(\mu(x)\), such that:

1. **Rational Expectations:** \(\mu(x) = \mathbb{E}[s(x(\theta'), \theta') \mid x(\theta') = x] \text{ for all } \theta \in \Theta\).

2. **Optimality:** \(x(\theta) = \arg \max_x u(x, \theta, \mu(x)) \text{ for all } \theta \in \Theta\).

3. **Free Entry:** there exists a \(\pi \in \Delta(\Theta)\) such that \(\mu(x) \geq \int_{\Theta} s(x, \theta) d\pi(\theta) \text{ for all } x \in \mathbb{R}_+\).

**2.1.1 A Graphical Approach**

In order to depict equilibria, it is convenient to consider various relationships in investment/skill space. To begin, we can plot the indifference curves of the utility function \(u(x, \theta, s')\) for each type \(\theta \in \Theta\), where the vertical axis measures units of coworker skill, \(s'\). This is shown in Figure 1. The indifference curves are U-shaped because at low investment levels the net marginal benefit of increasing investment is positive, which requires that a lower coworker skill is required to maintain indifference. At higher investment levels the opposite is true.

Furthermore, the fact that investment acts to augment ability in a complementary manner implies a single-crossing property: for any proposed increase in investment, higher types need fewer units of coworker skill in order to remain indifferent.\(^{18}\)

---

\(^{17}\)This can be generalized to a model in which workers are exogenously designated a ‘class’ ex-ante and matching occurs across classes (as in the original marriage model). See the Appendix for details.

\(^{18}\)Both the U-shapedness and the single-crossing properties can be demonstrated by direct inspection of the indifference curve expression. If we define \(y^{-1}(s, z)\) as the inverse of \(y\) with respect to coworker skill (so that \(y(s, y^{-1}(s, z)) = z\)), then the equation of indifference curves are given by \(I(x, \theta, s'; u) = y^{-1}(s(x, \theta), u + c(x))\). The slope of the indifference curve is \(I_x(\cdot) = [c_s(x) - y_u(s(x, \theta), s') \cdot s_s(x, \theta)] / y_{s'}(s(x, \theta), s')\). Using the additively sep-
In the same space we can plot the return function $\mu(x)$, where workers conjecture that an investment of $x$ allows them to match with a coworker of skill $\mu(x)$. This is depicted in Figure 2. Given an arbitrary return function, the optimality condition requires that each worker chooses the value of $x$ that places them on the highest indifference curve, subject to the return function. Note that this condition, along with the single-crossing property, implies that investment must be a non-decreasing function of type in equilibrium (see the proof of Lemma 1).

Although we know that investment must be a non-decreasing function of ability, there will still be many possible equilibria. I will focus on separating equilibria - equilibria in which each type invests a unique amount. The primary reason for this focus is that it makes comparisons with the literature more transparent. Furthermore, pooling equilibria are often non-robust to standard refinements proposed in the signaling literature (e.g. see Cho and Kreps (1987)). I do not pursue such refinements here, but a discussion of pooling equilibria is presented in the Appendix.

### 2.1.2 Separating Equilibria

A separating equilibrium is an equilibrium in which each type of worker invests a different amount. The single-crossing property implies that investment is in fact a strictly increasing function of type (see the Proof of Lemma 1). The following Lemma adds further structure.

**Lemma 2** If $x(\theta)$ is the investment function in a separating equilibrium, then $x(\theta)$ is differentiable.

**Proof** See Appendix.

In deriving a separating equilibrium, note that the rational expectations condition becomes:

$$\mu(x(\theta)) = s(x(\theta), \theta), \quad \forall \theta \in \Theta.$$ 

That is, workers must believe that if they invest an amount that is an equilibrium investment for some type, say $x(\theta)$, then they will be matched with a coworker that is of skill $s(x(\theta), \theta)$. This condition is depicted in Figure 3.

An equivalent, and much more convenient way to express the rational expectations condition is to work with the inverse investment function, as follows. Let $X$ represent the set of investments that arise in equilibrium, and for all $x \in X$, let $\xi(x)$ be the type of worker that finds it optimal to invest $x$. The rational expectations condition then becomes:

$$\mu(x) = s(x, \xi(x)), \quad \forall x \in X.$$ 

This expression can be substituted into the objective function so that optimality requires:

$$x(\theta) = \arg \max_x \{ y(s(x, \theta), s(x, \xi(x))) - c(x) \}.$$ 

arable form gives $I(\cdot) = \frac{1}{\phi} \cdot [u + c(x) - (1 - \phi) \cdot \theta \cdot g(x)]$ and $I_x(\cdot) = [c_x(x) - (1 - \phi) \cdot s_x(x, \theta)]/\phi$. Notice how single-crossing does not rely on complementary interaction, but rather the fact that investment complements natural ability.
The differentiability of \( x(\cdot) \) implies the differentiability of \( \xi(\cdot) \) (whenever \( x'(\theta) > 0 \)), which in turn implies the differentiability of the objective function. The first-order condition for this problem, once re-arranged, gives us a differential equation that \( \xi(x) \) must satisfy:

\[
\xi'(x) = \frac{c_x(x) - \left[y_x(s(x, \xi), s(x, \xi)) + y_x'(s(x, \xi), s(x, \xi)) \cdot s_x(x, \xi)\right]}{y_x'(s(x, \xi), s(x, \xi)) \cdot s_x(x, \xi)} \equiv \Gamma(x, \xi),
\]

where I have used the fact that \( \theta = \xi(x(\theta)) \). Whenever \( y \) has the property that \( y_x(z, z) = (1 - \phi) \) and \( y_x(z, z) = \phi \) for all \( z \geq 0 \) (as in the additively separable case) we have:

\[
\Gamma(x, \xi) = \frac{c_x(x) - s_x(x, \xi)}{\phi \cdot s_x(x, \xi)}.
\]

This differential equation, once combined with an initial condition, \( \{x_0, \xi_0\} \), defines an initial values problem. The solution to the initial values problem represents an equilibrium if \( \xi(x) \) is a strictly increasing function for \( x \geq x_0 \) and the second-order condition is satisfied.

The solution to the initial values problem can be analyzed quite easily by means of the function’s direction field, as depicted in Figure 4. At each point in \((x, \xi)\) space, we know the slope of \( \xi \) is \( \Gamma \). By selecting points in the space, the properties of the solution emerge. In particular, we can plot the locus of points for which the slope of \( \xi \) is zero. This is simply given by \( \{(x, \xi) \mid c_x(x) = s_x(x, \xi)\} \).

This locus is represented by an increasing function in \((x, \xi)\) space since \( s_x \theta > 0 \). Note that \( \Gamma(x, \xi) > 0 \) for points ‘south-east’ of this locus and \( \Gamma(x, \xi) < 0 \) for points ‘north-west’ of this locus. Since \( \xi \) is required to be an increasing function in equilibrium, the initial condition, \( \{x_0, \xi_0\} \) must be such that \( \xi_0 = \bar{\theta} \) and \( x_0 \geq x^*(\bar{\theta}) \) where \( x^*(\bar{\theta}) \) satisfies \( s_x(x^*(\bar{\theta}), \bar{\theta}) = c_x(x^*(\bar{\theta})) \).

The free entry condition is the key to determining the initial condition. One implication of the free entry condition is that \( \mu(x) \geq s(x, \bar{\theta}) \) (since the most pessimistic beliefs place all the weight on \( \bar{\theta} \)). This implies that workers of the lowest type must invest at the point at which their indifference curve is tangent to their skill production function. If this were not the case, then their indifference curve would at some point pass below their skill production function, but this can not be optimal since the return function must lie above the indifference curve at this point (since free entry requires it to never lie below the skill production function of the lowest type). This tangency condition requires that \( x_0 \) be such that \( s_x(x_0, \bar{\theta}) = c_x(x_0) \). This solution to the initial values problem is verified to be an equilibrium once the second-order condition is checked. The separating equilibrium is depicted in Figure 5.

**Proposition 1** A separating equilibrium exists and it is unique. The inverse investment function is the solution to the initial values problem defined by (i) \( \xi'(x) = \Gamma(x, x) \) and (ii) \( \{\xi_0, x_0\} \), where \( \xi_0 = \bar{\theta} \) and \( x_0 \) satisfies \( c_x(x_0) = s_x(x_0, \bar{\theta}) \). The solution to the initial values problem is:

\[
\xi(x) = \frac{c_x(x_0)}{g_x(x_0)} \left[\frac{g(x_0)}{g(x)}\right]^\frac{1}{2} + \frac{1}{\phi} \int_{x_0}^{x} \frac{c_x(z)}{g(z)} \cdot \left[\frac{g(z)}{g(x)}\right]^\frac{1}{2} \, dz.
\]

\(^{19}\)In other words, if this condition did not hold, then there can not be any \( \pi \in \Delta(\Theta) \) such that \( \int_{\Theta} s(x, z) d\pi(z) \leq \mu(x) \) since \( \mu(x) < s(x, \bar{\theta}) \leq \int_{\Theta} s(x, z) d\pi(z) \).
In order to better understand the following analysis of this equilibrium, it is useful to be clear about the difference between the approach taken here and the non-cooperative approach taken in the literature (e.g. Hoppe, Moldovanu, and Sela (2005)).

The non-cooperative approach to modeling the economy is to assume that workers make an investment, then are matched in a positive assortative manner on the basis of this investment. I will follow Peters (2004b) in calling this type of game a ‘premarital investment game’. Given that all workers prefer workers that invest higher amounts as partners, the imposition of positive assortative matching seems quite reasonable.

In contrast, the approach taken here is to assume that workers do not have any particular second-stage matching game in mind when they invest, but rather, simply hold a conjecture about the quality of partner they can expect following any given first-stage investment level. Once workers enter the matching stage they are paired off in a manner that is stable and that does not contradict the initial conjecture. Positive assortative matching is an outcome of this process, rather than an assumption imposed at the outset.

Despite the fact that both approaches produce positive assortative matching, the nature and resulting policy implications of equilibrium, as we shall soon see, are quite different. Thus, it is important to clearly understand the costs and benefits of each approach in order to determine the cases in which one approach will be more useful than the other.

One shortcoming of the ‘competitive’ approach taken here has been highlighted in a series of recent papers by Michael Peters (Peters (2004a), (2004b), (2006), and (2007)). The argument is that the competitive (hedonic) equilibrium requires that some agents misunderstand the consequences of making off-equilibrium investments. In particular, agents of the lowest type do not recognize that they will not receive a lower quality partner if they were to cut their investment.

This is a very important criticism, and underlies his surprising conclusion that investment behaviour in the premarital investment game does not converge to the competitive outcome as the number of agents becomes infinite.

One response to this criticism is that it only has bite to the extent that we think the premarital investment game (i.e. imposing positive assortative matching) truly reflects the matching market that one has in mind. For example, another plausible underlying game could be that workers must use their investment to gain entry into a matching sub-market, where each agent can visit only one sub-market. This structure could represent a reduced form of a matching process in which firms post vacancy notices that require a particular investment level of their applicants and workers are unable to send applications to all firms for which they qualify. Alternatively, we can think of firms as hiring either two or three workers. This assumption is awkward in a marriage context, but much less so in a worker-coworker context. Both of these variants share the feature that if a worker of the lowest type were to cut their investment then they will be unmatched. This restores the plausibility of the conjectures held about off-equilibrium play. Unfortunately, it also allows for a multiplicity of

---

20This is because they are already being matched with the lowest quality partner in the economy.
equilibria to form since very harsh off-equilibrium beliefs are not ruled out. I have overcome this by imposing the ‘free entry’ condition. By introducing the notion of competitive market-makers, this condition rules out equilibria in which workers invest a very high amount out of an unreasonable fear of being unmatched if they were to cut their investment.

This then indirectly raises a counter-criticism of the premarital investment game - equilibria are not likely robust to the existence of competitive market-makers, especially when the number of agents gets large. That is, it would be profitable for an entrepreneur to establish a service in which she randomly matches together workers that make some minimum investment. In some cases, especially cases in which the number of agents is small, this is not an important criticism. My point is that there is no one-size-fits-all approach to modeling this type of environment, and that care needs to be taken in choosing an approach, because, as we are about to see, the conclusions are often significantly different.

The above criticism of the competitive approach is set in an environment of complete information, however, an important shortcoming of the non-cooperative approach is not fully apparent in such environments. In particular, although separating equilibria exist in non-cooperative models of two-sided pure signaling with a continuum of agents and a non-degenerate continuous distribution of types (as in Hoppe, Moldovanu, and Sela (2005)), such equilibria fail to exist when ‘signals’ are productive. The reason is that agents of the lowest type must be investing as if they were taking their partner as fixed, since they always get the worst partner in equilibrium and can therefore cut their investment without consequence. This investment is called their Nash investment. However, the Nash investment will not be optimal for them because they can invest a slightly higher amount and match with an agent that has also invested this slightly higher amount. Since the Nash investment represents an under-investment, this deviation is profitable (and even more so because the new partner is also of a higher type).

In contrast, separating equilibria always exist within the approach taken in this paper. One reconciliation for this contrast is that the non-cooperative approach requires that workers of the lowest type make their Nash investment, whereas the approach here requires that the lowest types invest at the point at which their indifference curve is tangential to their skill production function. It so happens that the these two investment levels coincide when investment is completely unproductive, but not otherwise. In other words, the two approaches are identical if and only if investment is unproductive. Given that non-cooperative separating equilibria only exist in this special case, it seems that the approach taken in this paper is a relatively fruitful avenue for future research on two-sided premarital investments without complete information.

This establishes that there can not be any pure strategy equilibria, but a generalization of the argument establishes that there will not be any mixed strategy equilibria either. The reason is that the Nash investment must be in the support of the mixed strategy for agents of the lowest type since the Nash investment is a best response to being matched with an agent of the lowest type. Rational expectations can be used to show that there can be at most one other investment in the support (basically, the indifference curve intersects the skill production function at most twice). This other investment in the support represents an over-investment. But this means that slightly higher types will also prefer to invest this amount, contradicting the separating nature of equilibrium.

21This establishes that there can not be any pure strategy equilibria, but a generalization of the argument establishes that there will not be any mixed strategy equilibria either. The reason is that the Nash investment must be in the support of the mixed strategy for agents of the lowest type since the Nash investment is a best response to being matched with an agent of the lowest type. Rational expectations can be used to show that there can be at most one other investment in the support (basically, the indifference curve intersects the skill production function at most twice). This other investment in the support represents an over-investment. But this means that slightly higher types will also prefer to invest this amount, contradicting the separating nature of equilibrium.
3 Analysis

This section analyzes some of the properties of the separating equilibrium derived in the previous section.

3.1 Efficiency

One central issue of interest is whether investment is ‘efficient’. Before this is addressed, we need to be clear about what is meant by ‘efficient’. There are two aspects. First, given equilibrium investments does the equilibrium matching arrangement maximize total welfare? Since investments are fixed, this requires that the matching arrangement maximizes total output. Second, given the equilibrium matching arrangement, do equilibrium investments maximize total welfare? Since the matching arrangement is fixed this allows us to check whether welfare is maximized on a type-by-type basis.

One advantage of studying a production function that is additively separable in own- and coworker skill is that the issue of whether the matching arrangement is efficient is easily addressed. This is because all matching arrangements produce the same total output. This is not the case in models that assume a supermodular production function, such as Hoppe, Moldovanu, and Sela (2005) and Rege (2007), where positive assortative matching is the unique output-maximising matching arrangement (see Becker (1973)). Therefore the equilibrium matching arrangement is efficient and we can focus attention on the efficiency of investments.

Consider a particular pair of coworkers, $i$ and $j$. The total welfare generated within this match is $W(x_i, x_j; \theta_i, \theta_j) = s(x_i, \theta_i) + s(x_j, \theta_j) - c(x_i) - c(x_j)$. The efficient investments are defined to be those that maximize $W(\cdot)$. Since i) there is perfect segregation in equilibrium ($\theta_i = \theta_j$), and ii) $[s(x, \cdot) - c(x)]$ is concave in $x$, there is a unique efficient investment associated with each type defined by:

$$x^*(\theta) = \arg \max_x 2 \cdot [s(x, \theta) - c(x)] = \arg \max_x s(x, \theta) - c(x). \quad (2)$$

---

22 The total output always equals $\int_{\Theta} \theta \cdot g(x(\theta)) dF(\theta)$ since each worker contributes $\theta \cdot g(x(\theta))$ to total output regardless of who they are matched with. Different matching arrangements will influence the distribution of output, but each possible arrangement is Pareto efficient.

23 I assume an additively separable production function for simplicity, comparability, and to make the point that supermodularity is not essential. However, the model can easily accommodate a supermodular production function. Since equilibrium matching is positive assortative, the conclusion that equilibrium matching is efficient remains intact in this case also.

24 One may argue that this is an inappropriate (i.e. excessively stringent) definition because it is based on a set-up in which utility is transferable across partners, whereas utility is non-transferable in the model. A more reasonable criteria would then be that the investments are Pareto efficient (within the match). Since there is perfect segregation in equilibrium, both members of the match invest the same amount. Therefore, in checking whether investments are Pareto efficient we need only compare them to the symmetric Pareto efficient investment level. This level coincides with the investment that maximizes $W(x, x; \theta, \theta)$. In this sense, nothing is lost by using this definition of efficiency in the analysis of equilibrium investments. The advantage of this definition is that it allows us to speak of the efficient investments for each member of the match without having to continually refer to the set of (Pareto) efficient investments.
There are good reasons to suspect that workers will invest too little in equilibrium due to the presence of positive externalities. Since workers do not take into account that their investment benefits their coworker, we could reasonably suspect that workers invest too little. This turns out to not be the case however, since workers are provided an added incentive to invest in the form of a better coworker. Peters and Siow (2002) and Cole, Mailath, and Postelwaite (2001) show how a concern for matching can induce agents to invest efficiently when they would otherwise under-invest. These papers make the assumption that all an agent cares about in their partner is their partner's investment. This is not the case here however, since workers care about their partner's skill, which depends on their partner's ability in addition to their partner's investment.

To explore this, note that the first-order condition associated with \( (2) \) is \( s_x(x, \theta) = c_x(x) \). If we let \( \xi^*(x) \) be the inverse efficient investment function, then we have \( s_x(x, \xi^*(x)) = c_x(x) \). But notice how this corresponds exactly the locus of points for which \( \Gamma(x, \xi) = 0 \) in Figure 4. Since the inverse equilibrium investment function, \( \xi \), has a slope of zero along this locus and a positive slope at all points 'to the right' of the locus it follows that \( \xi(x_0) = \xi^*(x_0) \) and \( \xi(x) < \xi^*(x) \) for all \( x > x_0 \). That is, the lowest type invests efficiently but all other types invest too much since the type that actually invests \( x \) is lower than the type that would efficiently invest \( x \). This is formalized in the following Proposition.

**Proposition 2** Workers of the lowest type invest efficiently in equilibrium, but all other types over-invest. Equilibrium investments approach efficient investments as spillovers disappear.

How can we reconcile the positive externality with over-investment? To begin, suppose that a firm or some other institution were able to internalize the externality by rewarding skill rather than productivity. Workers invest efficiently when they are paid the full marginal product of their skill: i.e. when they solve

\[
\max_x \ y(s(x, \theta), s') + y(s', s(x, \theta)) - c(x).
\]

That is, workers invest efficiently when they perceive the net marginal benefit to be:

\[
y_s(s(x, \theta), s') \cdot s_x(x, \theta) + y_{s'}(s(x, \theta)) \cdot s_x(x, \theta) - c_x(x). \tag{3}
\]

Now consider the situation in which workers are paid according to their productivity. The nature of the matching market is such that higher investment levels allows workers to work with higher skilled partners. Thus, there is an added benefit to investing that is introduced by matching. The worker actually solves

\[
\max_x \ y(s(x, \theta), \mu(x)) - c(x).
\]

That is, workers actually perceive a net marginal return of:

\[
y_s(s(x, \theta), \mu(x)) \cdot s_x(x, \theta) + y_{s'}(s(x, \theta), \mu(x)) \cdot \mu'(x) - c_x(x). \tag{4}
\]

Workers will therefore invest efficiently in equilibrium when \( \mu \) is such that (3) equals (4). When evaluated at the equilibrium investment level, (i) the rational expectations condition ensures that
\( s' = \mu(x) \), (ii) by definition, we have \( \theta = \xi(x) \), and (iii) perfect segregation implies \( s' = s(x, \theta) \).

Using these observations when equating (3) and (4) allows us to cancel expressions so that we are left with the requirement that:

\[
s_x(x, \xi(x)) = \mu'(x).
\]

The same three observations also imply that \( \mu(x) = s(x, \xi(x)) \) (the alternative expression of the rational expectations condition). Differentiating this with respect to \( x \), we have that 

\[
\mu'(x) = s_x(x, \xi(x)) + s_\theta(x, \xi(x)) \cdot \xi'(x).
\]

Therefore, \( \xi \) must be such that:

\[
s_x(x, \xi(x)) = s_x(x, \xi(x)) + s_\theta(x, \xi(x)) \cdot \xi'(x),
\]

which can only happen when \( \xi'(x) = 0 \). This means that workers would invest efficiently if and only they expected that their investment would influence the investment made by their match, but not the ability of their match. Of course, this does not work here because, in equilibrium, matching with a higher-investing coworker necessarily means matching with a higher-ability coworker.25 Thus, the incentive to invests are too great. The fact that workers invest efficiently in Peters and Siow (2002) can be seen as a consequence of the fact that the agents in their model do not care about the ability of their partner, only their investment.

This proposition has three main implications. The first is that spillovers are central to the overinvestment result in the sense that there is over-investment if and only if spillovers exist. In this light, credentialism and education externalities are not two separate phenomena that need to be weighted against each other in uncovering the social benefits of education. Rather, the analysis here suggests that the phenomena are structurally related: credentialism is a consequence of spillovers.

The second is from a policy perspective: despite the existence of positive spillovers, an investment subsidy could never be optimal (an analysis of optimal policy is contained in the Appendix). Thus, not only does the model provide an explanation for credentialism, it also confirms that credentialism implies excessive educational attainment.26

The third implication from this result is that heterogeneity matters. If agents were assumed to be homogeneous, then there would be no over-investment (since all workers are of the lowest type). Apart from this, heterogeneity forces us to think more carefully about modeling skill externalities because any ‘macro’ approach that uses a representative agent will necessarily overlook the mechanism offered here.

---

25Note that the above argument is general to the extent that I did not rely on any specific functional form for either \( y(s, s') \) or \( s(x, \theta) \).

26Whilst credentialism is casually associated with over-investment, it need not be the case. For example, Hopkins (2005) considers an economy in which firms are (exogenously and observably) differentiated by some quality measure and workers invest in education in order to compete for the ‘good jobs’. Under my definition, this economy exhibits credentialism since workers are assigned to firms on the basis of their investment. He shows that when worker investments are productive, low ability workers will tend to invest too little. Intuitively, these workers do not take into account the fact that their investment also benefits the firm. Such possibilities do not arise in the present model however since investment is two-sided.
3.2 Entry of Lower Types

Although I have exogenously fixed the lowest type, one could imagine a slightly more involved model in which there are two sectors - a non-skilled sector and a skilled sector. Workers in the non-skilled sector produce an output of $y_u$ regardless of their skill, whereas workers in the skilled sector produce an output that depends on their skill as well as the skill of their coworker, as described above. There will be some cut-off type such that lower types enter the non-skilled sector and higher types enter the skilled sector. This cut-off type can then be considered as endogenizing the value of $\theta$. Note that as $y_u$ increases $\theta$ also increases as more workers are drawn into the non-skilled sector. Therefore it may be of interest to understand how equilibrium investments are affected by $\theta$.

**Proposition 3** An increase in $\theta$ lowers the equilibrium investment of all types.

Intuitively, we can think of each worker’s investment as being sufficiently high so that marginally lower types are discouraged from imitating their behaviour. Since this is recursive in a sense, lower types have fewer workers ‘below them’ to discourage. All workers find fewer workers below them as the lowest type is increased, and therefore do not have to invest as much in order to discourage such types from imitating.

In terms of the two-sector model sketched briefly above, a decrease in the non-skilled sector’s wage raises productivity and incomes in the skilled sector purely because some low types move from the unskilled sector to the skilled sector, which results in all workers in the skilled sector increasing their equilibrium investment. Thus, if wages in the unskilled sector are relatively high, or if there are artificial barriers to entry into the skilled sector (perhaps based on gender or caste), then aggregate productivity remains relatively low since there are reduced incentives for workers in the skilled sector to invest in their skills.

3.3 The Effect of Spillovers

If technological progress tends to increase the scope for spillovers (see discussion in the Introduction) then it is of interest to examine the consequences of rising spillovers. Recall that the definition of the spillover parameter is $\phi_2/(\phi_1 + \phi_2)$, where $\phi_1 + \phi_2$ was normalized to unity. Thus ‘an increase in spillovers’ means a simultaneous increase in $\phi_2$ and decrease in $\phi_1$ (so that the sum remains at unity). With this construction, changes in $\phi$ are ‘neutral’ in the sense that individual productivities would not be affected in equilibrium if behaviour were not affected. Thus, changes in equilibrium variables arising from changing spillovers occur purely because of the fact that incentives to invest are changed.

**Proposition 4** Workers of the lowest type are unaffected by spillovers, however the equilibrium investment of all other workers is increasing in spillovers.

**Proof** See Appendix.

This result indicates that the spillover dimension of new technologies may be an important source of productivity growth in the development process. To be sure, it is not that spillovers enhance the
productivity of existing worker skills, but rather, that spillovers provide incentives for workers to improve their skills.

This theory is difficult to test against the theory that new technology simply makes existing skill more productive.\footnote{It is important to emphasize that there is nothing in the model that says that new technology can not also make existing skills more productive. That is, new technology will likely raise both \( \phi_1 \) and \( \phi_2 \). The point is that productivity will increase even if \( \phi_1 + \phi_2 \) does not change. In this light, it is unsurprising that the two theories are difficult to differentiate.} Both theories imply that new technologies raise productivity and involve a rising educational attainment. One possibility is examining the OLS return to education. This should increase if technology raises the productivity of existing skills, but will decrease if it is induced purely by greater spillovers.

Although greater spillovers raise productivity and incomes, the effect on welfare is not favorable.

**Corollary 1** The welfare of workers of the lowest type are unaffected by spillovers, however the welfare of all other workers is decreasing in spillovers.

This follows from the fact that i) there is over-investment and ii) spillovers increase investment. This result implies that if all differences in income across economies were attributable to differences in spillovers, then average income and welfare at the economy level would be negatively correlated in the cross section.

### 3.3.1 Returns to Education

In this section, I briefly explore some of the ways in a naive interpretation of observed returns to education will be misleading in the presence of spillovers. To begin, consider the true private returns from investment.

In equilibrium, a worker of type \( \theta_i \) evaluates the relationship between their investment and output as being given by:

\[
y(x_i, \theta_i) = y(s(x_i, \theta_i), \mu(x_i)) = y(s(x_i, \theta_i), s(x_i, \xi(x_i))),
\]

at least for investments that arise in equilibrium, \( x_i \in X \). The marginal return to education perceived by such a worker, when evaluated at their equilibrium investment, is:

\[
\frac{\partial}{\partial x} y(x_i, \theta_i) = (1 - \phi) \cdot s_x(x_i, \theta_i) + \phi \cdot [s_x(x_i, \theta_i) + s_\theta(x_i, \theta_i) \cdot \xi_x(x_i)] \tag{5}
\]

\[
= s_x(x_i, \theta_i) + \phi \cdot s_\theta(x_i, \theta_i) \cdot \xi_x(x_i), \tag{6}
\]

where I have used the fact that \( \xi(x_i) = \theta_i \) in equilibrium, as well as the fact that \( y_s(z, z) = (1 - \phi) \) and \( y_s'(z, z) = \phi \) for all \( z > 0 \).

The true private return will be overstated if the researcher were simply to fit a relationship between output and education due to what looks like a standard ‘ability bias’ problem. To be sure, investment and output are related in equilibrium by:

\[
\hat{y}(x_i) = y(s(x_i, \xi(x_i)), s(x_i, \xi(x_i)));
\]
implying a marginal return of
\[
\frac{\partial}{\partial x} \hat{y}(x_i) = s_x(x_i, \theta_i) + s_\theta(x_i, \theta_i) \cdot \xi_x(x_i),
\]
where I have use the same arguments as above. This implied return is never less than the true private return. Notice however that the presence of spillovers adds a subtlety: although the returns given by (7) are strictly greater than the true returns given by (6) for \( \phi \in [0, 1) \), the difference converges to zero as spillovers approach unity. This could never happen in a standard ‘ability bias’ problem, since investment is a strictly increasing function of ability. Spillovers are therefore introducing new scope for misinterpretation of observed returns. One feature that is perhaps puzzling at first glance is that the ability bias problem is actually reduced when spillovers - the source of imperfection in the model - are more pronounced. Indeed, the ability bias is completely overcome when spillovers are at their most severe.

To explore this further, it is useful to imagine the economy lying on a continuum that represents the extent to which a worker’s investment affects their income because of the fact that the investment augments their skills. The degree of spillovers determines the economy’s location on this continuum. At one end of the continuum, spillovers are zero. This corresponds to a pure ‘human capital’ economy in which investment raises skills, and these skills are rewarded directly. At this end, the entire benefit of investment arises because skills are enhanced. As spillovers are increased and we move along the continuum, the effect of investment on skills becomes less important relative to the effect of investment on attracting higher skilled coworkers. At the other end of the continuum, when spillovers are equal to one, none of the benefit of investment is directly due to skills being enhanced, but, rather, is entirely due to the fact that investment allows one to work with better coworkers. The significance of this is that an ability bias only arises when the effect of ability on income is mistakenly attributed to investment. But the effect of ability is relatively small when the effect of own skill is relatively small: that is, when spillovers are relatively high. Thus, a declining ability bias is consistent with rising spillovers. If the ability bias is ignored, then the resulting implied returns to education are exaggerated - but the extent of this exaggeration is reduced when spillovers are greater.

Suppose now that the ability bias problem is perfectly resolved (by assuming the researcher observes ability, or has some perfect instrument), and the researcher re-estimates the relationship between wages and investment while dealing with the ability bias. Such an exercise will reveal the true private returns to education (assuming the researcher uses the correct functional forms, etc). Such estimates are useful for individuals attempting to decide upon their optimal investment, but are not necessarily useful to a government that is evaluating the social benefit of education. The social return from a worker’s investment is given by:
\[
y^*(x_i) = y(s(x_i, \theta_i), s_j) + y(s_j, s(x_i, \theta_i)),
\]
where \( s_j \) is the skill of the worker \( i \)'s coworker. The marginal social return in equilibrium is
\[
\frac{\partial}{\partial x_i} y^*(x_i) = s_x(x_i, \theta_i),
\]
where I have used the fact that \( s_j = s(x_i, \theta_i) \) in equilibrium. It is important to note that evidence of a positive causal relationship between education and wages offers little, if any, insight into this social return. To see this most clearly, note that there is a positive causal relationship between education and wages in both a pure signaling model and a pure human capital model. This is a central reason why the debate between the human capital model and the signaling model has persisted for such a long time.

If we are preoccupied with distinguishing between the ‘human capital’ and ‘signaling’ models, then it may be of some value to possess data on worker-level output (as opposed to just wages). For various reasons, this is largely infeasible, but in principal it would shed light on the issue. The human capital explanation would posit that the effect of education on output is roughly the same as the effect on wages (since wages are tied to output directly). The signaling explanation would predict that education has less of an effect on output (possibly zero) than it does on wages. So, imagine that the researcher determined that the effect on output is roughly the same as the effect on wages. This provides evidence against signaling, and would likely produce the conclusion that there is no over-investment in education, given the evidence in support of the human capital model. This conclusion is misleading in the presence of spillovers. The reason is that wages are tied to output in this model also. In other words, such an exercise may be able to distinguish between signaling and the model presented here (where the human capital model is a special case), but it will not be able to tell us anything about the degree of spillovers (and therefore the social value of education).

The above observation leads to the following paradox: **education raises a worker’s productivity, even if education has no capacity to raise skills.** The reason is that more education allows workers to work with higher skilled coworkers, which raises their productivity via spillovers. In short, finding evidence in favor of a human capital model does not rule out over-investment in the presence of spillovers.

How then can we distinguish between a pure human capital model (\( \phi = 0 \)) and models with credentialism (\( \phi \in (0, 1] \))? One way is to note that when spillovers are positive, the mechanisms at play share the main qualitative features of signaling models - that education has a value as a credential. Therefore, the most compelling evidence in favor of signaling models (e.g. Tyler, Murnane, and Willett (2000), Bedard (2001), and Lang and Kropp (1986)) is also consistent with positive spillovers.\(^{28}\)

The ideal way in which to determine the extent of spillovers is to include measures of coworker skills. There are a number of practical problems in being able to achieve this, perhaps explaining why there are relatively few papers that attempt such an exercise. A more standard way to estimate spillovers is to include a measure of region-wide education in a wage regression. For example, Acemoglu and Angrist (1999) include state-level education aggregates in the wage equation, whilst Sand (2007) and Moretti (2004) uses city-level education aggregates. Such an approach will likely have

\(^{28}\)In fact, the model here perhaps offers a better interpretation of the results of Tyler, Murnane, and Willett (2000) because they find that the GED credential has a positive effect on wages, but this effect does not manifest itself for five years after attainment. This seems implausible in a signaling world, because the worker’s true productivity would almost surely be revealed during that time.
little power in detecting the type of spillovers introduced here for the simple reason that spillovers occur on a much more local level. Approaches like these will produce spurious estimates if regions differed in their spillover parameter, or if there is sorting across regions and a mis-specification in the wage equation. This issues are left to future research.

3.4 A Comparison with Global Spillovers

How would matters be different if we were to model skill spillovers in a more standard way? In the spirit of Lucas (1989) and Moretti (2004), suppose that each worker benefits from some aggregate of skill in the economy. One parsimonious way to address this issue is to suppose that a worker’s productivity is given by:

\[ y(s, \bar{s}) = (1 - \phi) \cdot s + \phi \cdot \bar{s}, \]

where \( \bar{s} \) is the average skill in the economy. As before, \( s = s(x, \theta) = \theta \cdot g(x) \). The additive separability is convenient because we know that investment with global spillovers, \( x^G(\theta) \), satisfies:

\[ (1 - \phi) \cdot s(x^G(\theta), \theta) = c_x(x^G(\theta)). \]

The key difference between this set-up and the model above is that a worker’s investment does not influence the skill spillover that they are exposed to. As such, the standard under-investment problem arises since no worker takes into account their positive impact on other workers. In fact, many of the central conclusions from the analysis are reversed. For instance, an increase in the spillover parameter lowers investment. Intuitively, inequality will tend to be relatively low under global spillovers since all workers are exposed to the same spillover level. As spillovers increase, this common component becomes relatively more important, which implies that spillovers tend to lower inequality when spillovers are global. The opposite is true in the model (see the following Section).

One way to view global spillovers is to note it’s qualitative equivalence to a setting in which matching on the basis of investment was infeasible (e.g. if investments were hidden). That is, \( \bar{s} \) can be interpreted as the expected skill that one will obtain from a coworker, given that matching is random. In this light, the comparison between equilibrium welfare and welfare with global spillovers is equivalent to the comparison of welfare between fully observable investments and fully hidden investments.

The trade-off here is different to that studied in the literature because interaction is not complementary. Here, visible investments lead to over-investment whereas hidden investments lead to under-investment. Welfare is the same across the cases in the absence of spillovers (\( \phi = 0 \)), since

\[ 29 \text{Models in which spillovers enter in this way are qualitatively the same (for our purposes) as models in which interaction is local but meetings are random, as in Glaeser (2001).} \]

\[ 30 \text{The trade-off considered in the literature (e.g. Rege (2007) and Hoppe, Moldovanu, and Sela (2005)) is that signaling is wasteful but, thanks to complementary interaction, is socially productive because it facilitates superior matching patterns. A trade-off of this nature is noted in Arrow (1973) and Stiglitz (1975). This trade-off can not be at work in the present model however since interaction is not strictly complementary. As noted previously, matching patterns are irrelevant for efficiency in this model.} \]
investment is efficient in both cases. When there are complete spillovers ($\phi = 1$), investment is zero in the hidden case. Average welfare is therefore also zero. In contrast, workers still separate in the visible case but each earns a utility equal to that obtained by the lowest type when they invest efficiently. The reason is that indifference curves are the same for all workers when spillovers are complete. Visible investment therefore tends to produce higher welfare when spillovers are very high. Intermediate cases are ambiguous, and depend in the extent to which investment augments ability. Along this dimension, hidden investments tend to produce greater welfare when investment is not very productive (e.g. hidden investments deliver the efficient investment level - zero - when investment is unproductive). The nature of this trade-off is analyzed in further detail in the following Section, where I use functional forms to explicitly derive equilibrium quantities. This exercise also shows how spillovers increase inequality in the model but decrease inequality when spillovers are global.

4 An Illustration

This section utilizes particular functional forms to demonstrate an equilibrium in closed form. This exercise illustrates some of the points made above and introduces some additional points of interest. Let the skill production function be given by $g(x) = x^\eta$ for some $\eta \in (0, 1)$, and let investment costs be given by $c(x) = c \cdot x$. Using (1), the inverse investment function turns out to be:

$$\xi(x) = \Lambda_0 \cdot \left[ \frac{x_0}{x} \right]^\eta \cdot x_0^{(1-\eta)} + \Lambda \cdot x^{1-\eta},$$

where $x_0 = \left[ (\eta/c) \cdot \tilde{\theta}^{(1/(1-\eta))} \right]$ and

$$\Lambda_0 = \frac{c}{\eta} \cdot \frac{\phi(1-\eta)}{\eta + \phi(1-\eta)},$$

$$\Lambda = \frac{c}{\eta + \phi(1-\eta)}.$$

Since $\xi$ starts off convex, then turns concave, the investment function starts off concave before turning convex.

To verify that there is over-investment, note that the inverse of the efficient investment function is $\xi^*(x) = (c/\eta) \cdot x^{1-\eta}$. There is over-investment since $\xi(x)/\xi^*(x) \leq 1$ (strict for $x > x_0$).$^{31}$ Thus, the type that efficiently invests $x$ is greater than the type that does invest $x$. Furthermore, the ratio approaches unity (investments become efficient) as spillovers go to zero.

To verify that a higher lowest type lowers investment, it is straightforward to see that $\xi$ is increasing in $x_0$ (and therefore in $\tilde{\theta}$). Also, the fact that $\xi$ is decreasing in $\phi$ demonstrates that increases in $\phi$ increase equilibrium investment.

The output produced by a worker that invests $x$, $g(x) \cdot \xi(x)$, is:

$$y(x) = \Lambda_0 \cdot \left[ \frac{x_0}{x} \right]^\eta \cdot x_0^{(1-\eta)} + \Lambda \cdot x,$$

$^{31}$The actual expression equals $\frac{\eta + \phi(1-\eta)}{\eta + \phi(1-\eta)} \cdot \left[ (x_0/x)^{(\eta/\phi)} \right]^{1+1-\eta}$. The result follows since the bracketed term is equal to one when $x = x_0$ and is strictly less than one for $x > x_0$. 

23
which is a convex function. If the first term is ignored, the OLS slope will appear to be \( \Lambda \). Interestingly, this slope is decreasing in the degree of spillovers. This result is consistent with the observation that Mincerian returns to education tend to be higher in poorer countries and in countries that have lower aggregate educational attainment (since economies with low spillovers have also have lower aggregate productivity and educational attainment).

The functional form used for \( g \) is convenient because the case in which investment is unproductive corresponds to the limiting case in which \( \eta \to 0 \). In the limit, the inverse investment function is \((c/\phi) \cdot x\), implying that the investment function is \( x(\theta) = (\phi/c) \cdot \theta \). Thus, the over-investment result does not rely on investment being productive.\(^{32}\) The case of unproductive investment highlights the way in which returns to education can be misinterpreted, even if ability were observed. This is because output is:

\[
E[y_i | \theta_i, x_i, \eta = 0] = (1 - \phi) \cdot \theta_i + \phi \cdot \xi(x_i)
\]

Therefore, an OLS regression of output on investment would reveal a positive correlation, even controlling for ability. Often, controlling for ability makes the researcher more comfortable in making claims of causality, and this is true here. However, education causes increased output at the individual level because of the matching market, not because education augments ability.

An analytic solution for the investment function is available if we let \( \theta = 0 \). Equilibrium investment is:

\[
x(\theta | \theta = 0) = [\Lambda^{-1} \cdot \theta]^{\frac{1}{1-\eta}},
\]

which makes equilibrium output:

\[
y(\theta | \theta = 0) = \Lambda^{-\frac{\eta}{1-\eta}} \cdot \theta^{\frac{1}{1-\eta}}.
\]

Equilibrium welfare is:

\[
u(\theta | \theta = 0) = \left[\Lambda^{-\frac{\eta}{1-\eta}} - c \cdot \Lambda^{-\frac{1}{1-\eta}}\right] \cdot \theta^{\frac{1}{1-\eta}},
\]

which is a decreasing function of \( \phi \).

Comparing the equilibrium outcome to that which would arise with global spillovers, it turns out that the equilibrium variables are the same replacing \( \Lambda \) with \( \Lambda^G \equiv c/[\eta \cdot (1 - \phi)] \). Average welfare is greater with visible investments when \( \eta > 1/2 \) and greater with hidden investments when \( \eta < 1/2 \) (this holds for all values of \( \phi \in (0, 1) \)). Welfare is the same across the two cases when \( \phi = 0 \) or \( \phi = 1 \). When \( \phi = 0 \) behaviour is the same across the cases. However, when \( \phi = 1 \) the

\(^{32}\)When \( \eta = 0 \), there is an equilibrium in which workers invest according to \( x(\theta) = (\phi/c) \cdot \theta \). Given this, an investment of \( x \) yields a partner of skill \( \mu(x) = (c/\phi) \cdot x \). Once inserted into the objective function, it becomes clear that payoffs are independent of investment (among equilibrium investments anyway). The investment function therefore constitutes an equilibrium. Uniqueness relies on investments being productive however, since there is another equilibrium surviving the refinement in which all workers invest zero.
behaviour is quite different since investment is zero with hidden investments but is very large with visible investments (so much so that all workers end up with a payoff of zero).

One benefit of a closed-form expression for the investment function is that it allows us to explicitly examine inequality. Although there are many available measures of inequality, a prominent measure is the percentile gap (e.g. the 90-10 and the 90-50 percentile differences are often documented for wage distributions). If \( y_p \) is the output (and in this case, wage) at the \( p \)th percentile, then the inequality measure is

\[
\Delta(p, p') \equiv y_{p'} - y_p,
\]

for \( 0 \leq p < p' \leq 1 \). Since equilibrium output is a strictly increasing function of type, \( y_p \) is simply the output produced by the type at the \( p \)th percentile. That is, the type \( \theta_p \) such that \( F(\theta_p) = p \). It then follows that the inequality measure is given by:

\[
\Delta(p, p') = y_{\theta_p'} - y_{\theta_p} = \Lambda^{-\frac{n}{1-\eta}} \left[ \theta_{\frac{1}{1-\eta}}^{\frac{1}{\eta} - \theta_p} \right].
\]

The first point to note is that inequality is decreasing in \( \Lambda \), and is therefore increasing in spillovers. Furthermore, if \( F \) is not too convex (e.g. a uniform distribution) then the inequality measure is increasing in \( p \), holding \( p' - p \) fixed.\(^{33}\) That is to say that inequality is greater at the top end of the distribution. This feature is not due solely to spillovers, since this holds even when all workers are investing efficiently - the relevant point is that the effect of spillovers on inequality will be most pronounced at the top of the distribution.

**Result 1** Rising spillovers induce an increase in inequality. For example, the 90/10 wage differential increases. If \( F \) is not too convex (e.g. a uniform distribution), then this increase is concentrated at the top end of the distribution. For example, the 90-50 wage differential increases more than the 50/10 wage differential.

Rising inequality within many OECD countries since the 1970’s - particularly in the 1980’s - has been well-documented (see Gottschalk and Smeeding (1997) for a survey), and Autor, Katz, and Kearny (2006) document the continued growth in the 90/50 (log) wage differential since the late 1980’s in the U.S.

In order to better understand potential sources of inequality, it is common to compare outcomes within and between ‘skill groups’. In practical terms, these skill groups often refer to different educational categories, such as ‘college educated’ and ‘high school educated’. Since educational attainment is termed ‘investment’ in the model, we can perform the exercise of comparing outcomes within and between ‘investment groups’. To this end, suppose that the researcher does not observe investment, but, rather, observes the region in which the investment falls. For instance, the

\[^{33}\text{Specifically, if } \psi(z) \equiv [F^{-1}(z)]^{1/(1-\eta)} \text{ is a convex function. For example, any distribution of the form } F(z) = z^q \text{ will do as long as } q \leq 1/(1-\eta) \text{ (e.g. a uniform distribution). All distributions with a non-increasing density will work (e.g. the class of exponential distributions). If we let } a \in (0, \bar{a}) \text{ and assume that } (\theta - a) \text{ is distributed log-normal (with ‘mean’ parameter } m \text{ and ‘standard deviation’ parameter } s) \text{, then the condition holds for sufficiently large } s.\]
researcher knows that a person has a college degree but does not know what major the degree is in, the person’s grades, the quality of the institution, and so on. This creates a situation in which workers of different skills (in the sense used in the model) are observationally equivalent.

To make some progress on this, assume that abilities are uniformly distributed on \([0, 1]\). The researcher observes whether or not a worker’s investment is greater than some cut-off, \(\hat{x}\). At the risk of confusing terminology, let us say that workers with investments less than \(\hat{x}\) are ‘un-skilled’, and workers with investments greater than \(\hat{x}\) are ‘skilled’.

Given the cut-off investment, the equilibrium investment function implies a cut-off ability, where a worker is ‘skilled’ in equilibrium if and only if their type is above this cut-off, is given by \(\hat{\theta} \equiv \hat{x}^{1-\eta} \cdot \Lambda\). Assuming that \(\hat{\theta} \leq 1\), the ‘supply of skilled workers’ is therefore given by:

\[
S(\hat{x}; \cdot) = 1 - \hat{\theta} = 1 - \hat{x}^{1-\eta} \cdot \Lambda.
\]

Since \(S\) is decreasing in \(\Lambda\), greater spillovers lead to a greater supply of skilled workers. The average ability of a skilled worker is \((1 + \hat{\theta})/2\), and the average ability of an unskilled worker is \(\hat{\theta}/2\). Greater spillovers therefore lowers both of these averages. The effect on average output is more complex because greater spillovers induce more investment. The average output (and wage) among the unskilled is:

\[
Y_U = \frac{1 - \eta}{2 - \eta} \cdot \hat{x} \cdot \Lambda,
\]

which is monotonically decreasing in spillovers.\(^{34}\) Intuitively, there is an upper limit to the investment made by an unskilled worker (by definition). As spillovers increase, this maximum investment is being made by lower ability workers. The average output (and wage) among the skilled is:

\[
Y_S = \frac{1 - \eta}{2 - \eta} \cdot \hat{x} \cdot \Lambda \cdot \left[ \frac{\frac{1}{\hat{x} \Lambda^{1-\eta}}} {1 - \hat{x}^{1-\eta} \Lambda} \right]
\]

This expression can be either increasing or decreasing in spillovers: greater spillovers induces more investment, but also induces entry of lower ability workers. Unlike the unskilled group, there is no upper limit on investment within the skilled group.

Notice how the large bracketed term will equal \(Y_S/Y_U\); a measure of the skill premium. This expression is decreasing in \(\Lambda\), implying that greater spillovers increase the apparent skill premium.

**Result 2** Rising spillovers induce a positive correlation between the skill premium, \(Y_S/Y_U\), and the supply of skilled workers, \(S\). Furthermore, the rise in the skill premium involves a decline in the average wage of unskilled workers.

\(^{34}\) The result would not hold for all distributions of ability since it could be the case that there are relatively few workers on the verge of becoming identified as skilled. As greater spillovers induce these workers to change categories, the reduction in average output among the unskilled due to the composition change will not fall much relative to the increase in average output owing to the fact that all workers invest more.
This result lends itself to a comment on skill-biased technical change. The above correlation has been widely documented for the US and has been interpreted as ‘skill-biased technical change’ (see Acemoglu (2002)). Furthermore, the rising skill premium is due to falling wages for the unskilled as opposed to rising wages of the skilled, as happens here. Although both models predict this correlation, the policy implications are dramatically different. Under skill-biased technical change, encouraging educational attainment is valuable because the production technology is becoming increasingly geared toward skilled workers. On the other hand, there is already over-investment in education in the above model and any further encouragement will lower welfare.

Another well-documented trend is that of residual inequality. This can be analyzed by examining the inequality within each skill group separately. The worker that has their ability at the \( p \)th percentile among the unskilled, \( \theta^U_p \), is that value that satisfies \( F(\theta^U_p)/\hat{\theta} = p \). Similarly, the worker that has their ability at the \( p \)th percentile among the skilled, \( \theta^S_p \), is that value that satisfies \( (F(\theta^S_p) - F(\hat{\theta}))/ (1 - F(\hat{\theta})) \). Utilizing the uniform distribution, we can write:

\[
\begin{align*}
\theta^U_p &= p \cdot \hat{\theta} = p \cdot \hat{x}^{1-\eta} \cdot \Lambda \\
\theta^S_p &= p + (1 - p) \cdot \hat{\theta} = p + (1 - p) \cdot \hat{x}^{1-\eta} \cdot \Lambda.
\end{align*}
\]

Inequality within skill group \( K \in \{U, S\} \) then becomes:

\[
\Delta^K(p', p) = \left[ \frac{\theta^K_{p'}}{\Lambda^\eta} \right]^{-\frac{1}{1-\eta}} - \left[ \frac{\theta^K_p}{\Lambda^\eta} \right]^{-\frac{1}{1-\eta}}.
\]

**Result 3** Spillovers increase inequality among the skilled, but lowers inequality among the unskilled.

Lemieux (2006) shows trends in residual inequality by education group and shows how higher education groups have greater residual inequality. Further, it seems that this inequality has grown faster for higher educational groups since the early 1970’s (and has actually declined for the very lowest education groups).

Finally, although there is evidence that inequality has risen in many economies, largely attributable to residual inequality of the most skilled groups, it must be that this inequality is a between-firm phenomenon if the above theory is to concord with the evidence. Some recent papers provide empirical support for this (Faggio, Salvanes, and Van Reenen (2007) and Dunne, Foster, Haltiwanger, and Troske (2004)).

### 5 Extensions

#### 5.1 A Dynamic Version

##### 5.1.1 Non-Transferable Utility

One might think that the over-investment result is due to the fact that, in a static model, workers are ‘locked-in’ once they are assigned a coworker. The point of this section is to provide a very basic extension, set in discrete time, in which workers are assigned coworkers, then observe the skill of
their assigned coworker, then have the opportunity to reject an undesirable coworker. I argue that the nature of equilibrium is unaffected by changes in the workers’ discount factor, and therefore that the above results do not hinge on the fact that the model is static.

Consider a generalization of the economy described above which operates in discrete time. All workers have a discount factor of $\beta \in [0, 1)$ so that a sequence of outputs $\{y_t\}_{t=0}^{\infty}$ produces a (gross) payoff of $Y = \sum_{t=0}^{\infty} \beta^t y_t$. In general $Y$ will depend on the worker’s type, $\theta$, and investment, $x$, so write this as $Y(x, \theta)$. The worker’s net payoff incorporates the cost associated with this investment, $C(x)$:

$$U(x, \theta) = Y(x, \theta) - C(x)$$

A new generation of workers are born each period. In the first period of existence workers choose their investment and are assigned a coworker in the labour market. If both workers within a match agree to the match, then the workers produce together for all future periods. If a worker of skill $s$ forms such a partnership with a coworker of skill $s'$ then the worker produces and consumes $y(s, s') = (1 - \phi) \cdot s + \phi \cdot s'$ each period and therefore obtains a gross payoff of $(1 - \beta)^{-1} \cdot y(s, s')$. If at least one of the workers does not agree to the match, then both workers re-enter the labour market the following period and are assigned new coworkers according to the equilibrium matching function. I assume that the worker’s only verifiable characteristic in the labour market is their investment (i.e. the output history is non-verifiable, as is the skill of their previous partners, etc). This assumption is clearly extreme, but it allows me to retain the feature that matches must be formed on the basis of investment.

Since we are primarily interested in the qualitative impact of changes in $\beta$ upon equilibrium investment, it is useful scale up investment costs by the factor $(1 - \beta)^{-1}$. The reason is that an increase in patience will mechanically increase incentives to invest simply because workers care more about the future payoffs that the investment allows for. Scaling up costs in this way purges the overall impact of patience of such mechanical effects. Thus, I assume that $C(x) = (1 - \beta)^{-1} \cdot c(x)$.

Given the analysis of the static model, I will look for a separating equilibrium. As in the static case, there is positive assortative matching on investment, and a worker that invests $x$ is matched with a coworker of skill $s(x, \xi(x))$, where $\xi(x)$ is again the inverse investment function. Now we can write a worker’s total payoff as $U(x, \theta) = (1 - \beta)^{-1} \cdot u(x, \theta, s(x, \xi(x)))$.

Even though the objective function is equivalent to the objective function in the static case, note that the constraints on the optimization problem are different. In the static case, we only required that $x \geq 0$. This is because matches would never be refused since there is only one period. This is not true in the dynamic case. If a worker of skill $s$ is matched with a coworker of skill $s'$, then the partnership is worth $(1 - \beta)^{-1} \cdot y(s, s')$ to the worker. In equilibrium the worker is supposed to match with a coworker of skill $s$, and can therefore guarantee a partnership worth $(1 - \beta)^{-1} \cdot y(s, s)$ in the following period if they are to reject the coworker of skill $s'$. Therefore, a worker of skill $s$ agrees to a match with a worker of skill $s'$ if and only if $y(s, s') \geq \beta \cdot y(s, s)$. If we let $A(s, \beta)$ be such that $y(s, A(s, \beta)) = \beta \cdot y(s, s)$, then $A(s, \beta)$ is the minimum skill that a worker of skill $s$ will accept in equilibrium given a patience of $\beta$. Notice that $A(s, \beta)$ has the following properties.
1. Workers always accept partners of equal skill: $A(s, \beta) < s$.

2. All workers are accepted for low enough patience: $\exists \beta \in (0, 1)$ such that $A(s, \beta) = 0$.

3. Workers become more selective as patience increases: $A(\beta) > 0$ and $\lim_{\beta \rightarrow 1} A(s, \beta) = s$.

The first property intuitively reflects the fact that workers prefer to accept their equilibrium match sooner rather than later. The second property tells us that the conclusions drawn from the static case (which is equivalent to $\beta = 0$) remain unaltered for sufficiently small levels of patience. The third property is key since it opens the possibility that patience can have a qualitative effect on equilibrium investments because the value of masquerading as a higher type is reduced as patience increases.

One intuition underlying over-investment in the static model is that workers are investing in order to distinguish themselves from lower types. Greater patience may reduce the pressure to do so because some lower types know that they will never be accepted by some higher types and, as a consequence, will never even attempt to masquerade as one. The question is whether this reduced ‘pressure’ translates into more efficient investment.

The minimum acceptance rule implies the following constraint on a worker’s optimization problem:

$$s(x, \theta) \geq A(s(x, \xi(x)), \beta).$$

That is, if a worker invests $x$ expecting to match with a partner of skill $s(x, \xi(x))$ then it must be the case that the worker’s skill is at least as great as the minimum required by their intended partner. In equilibrium, the constraint places an upper limit on how much a type $\theta$ worker can invest.\(^3\) The problem facing each worker is therefore

$$\max_x (1 - \beta)^{-1} \cdot u(x, \theta, \xi(x))$$

subject to (8).

Although the constraint has real implications for off-equilibrium behaviour - it places an upper limit on how much a worker would masquerade - it has no qualitative effect whatsoever on equilibrium investment behaviour. To see this, suppose that workers ignored the constraint. This produces an equilibrium investment identical to the one derived in the static case. Is the constraint satisfied? By the first property of $A(s, \beta)$, the answer is ‘yes’ for all possible values of $\beta$. In other words, if the constraint were binding for some type then this type must be matched with a worker of a strictly higher skill. But this then implies that the equilibrium is not separating since two workers of different abilities invest the same amount.

\(^3\)To see this, apply the function $y(s(x, \xi(x)), \cdot)$ to each side. This leads to $y(s(x, \xi), s(x, \xi)) \geq \beta y(s(x, \xi(x)), s(x, \xi(x)))$. Since $\tilde{y}(x, \theta) \equiv y(s(x, \xi(x)), s(x, \theta))$ cuts $\tilde{y}(x) \equiv y(s(x, \xi(x)), s(x, \xi(x)))$ at the equilibrium investment ‘from above’ (since $\theta > \xi(x)$ for $x < x(\theta)$), it also cuts $\beta \cdot \tilde{y}(x)$ from above. Therefore, the constraint requires that a type $\theta$ worker invests no more than $x^e(\theta, \beta)$ where this value satisfies $\tilde{y}(x^e(\theta, \beta), \theta) = \beta \cdot \tilde{y}(x^e(\theta, \beta))$. Furthermore, notice that $\beta < 1$ implies that $x^e(\theta, \beta) > x(\theta)$.
Result 4 When utility is non-transferable, equilibrium investment behaviour is not qualitatively affected by workers’ degree of patience, $\beta$.

The conclusion drawn from this exercise is that the over-investment that arises in the static case has nothing to do with the fact that workers are ‘locked in’ to their partnership. The same qualitative behaviour arises once we allow workers to reject partners. In contrast to the previous intuition, the ‘pressure’ leading a worker to over-invest is very much a local phenomenon. In other words, if a change in patience does not induce workers of type $\theta$ to change their investment, then it will not induce workers of slightly higher types to change their investment regardless of what all types lower than $\theta$ are doing.

One interpretation of this is that ‘signaling’ is not quite the right way to view the over-investment result. This claim is made in light of the facts that i) the hidden information - workers’ skills - are revealed once workers meet one another, and ii) the degree of patience can be thought of as the speed of learning this hidden information: the single period in which one’s partner’s skill is hidden becomes increasingly irrelevant as perfect patience is approached. Thus, the relevant feature is that matches are co-ordinated on the basis of investment and not skill. The ‘speed of learning’, once matched, is irrelevant.\footnote{This is a key difference between this model and standard signaling models. In those models the speed of learning is important. See the Appendix for a demonstration of how efficient investments can be supported in the analogous signaling environment.}

This result is special in at least two ways. First, it matters that types are not discrete. The efficient investments can always be supported when types are discrete for sufficiently high patience levels. Thus, a weaker version of the result holds: When utility is non-transferable and types are discrete, equilibrium investment behaviour is qualitatively unaffected when patience is sufficiently low. In the static model, if the efficient investments can not be supported it is because a type $\theta_k$ worker profits from raising their investment to the level that is supposed to made by a type $\theta_{k+1}$ worker. However, if patience is high enough, the type $\theta_{k+1}$ worker will find it optimal to reject any type $\theta_k$ workers that deviate by making the higher investment. This discourages the deviation and the efficient investments can be supported. Second, it matters that utility is non-transferable across partners. This case is now analyzed.

5.1.2 Transferable Utility: Bargaining

In the previous section the wage accruing to each worker was determined purely by the spillover parameter, and there was an implicit assumption that workers do not engage in side payments. An alternative approach is to assume that each worker produces an output equal to their skill and that workers within a match bargain over the total match output (the sum of the two individual outputs). This section studies this type of setting with Nash bargaining. Apart from wage determination, the economy is the same as just described.

Suppose that a worker $i$ is paired with a worker $j$ so that the pair bargain over the total match output, $s_i + s_j$. If $i$ gets a payment of $w_i$, then $j$ gets a payment equal to the remaining output,
\( s_i + s_j - w_i \). A wage of \( w \) produces a present value of \((1 - \beta)^{-1} \cdot w \). If \( k \in \{i, j\} \) has an outside option of \( q_k \) (to be determined), the wage to \( i \) is determined by (symmetric) Nash bargaining:

\[
\frac{w_i}{1 - \beta} - q_i = s_i + s_j - w_i - q_j.
\]

(9)

In equilibrium, worker \( j \) expects to be matched with a worker that also has a skill of \( s_j \). Given the symmetry, \( j \) gets paid a wage equal to \( s_j \) (since the match output, \( 2s_j \), is split evenly). Thus, the present value of future consumption for \( j \), once they meet their equilibrium partner, is \((1 - \beta)^{-1} \cdot s_j \).

Worker \( j \)’s outside option, \( q_j \), is the value associated with waiting one period in order to meet their equilibrium partner. That is, \( q_j = \beta \cdot (1 - \beta)^{-1} \cdot s_j \).

In equilibrium, the fact that worker \( i \) is optimizing means that if worker \( i \) is able to negotiate a wage of \( w_i \) this period, then this will also be the best that he can do in the following period. Thus, \( i \)’s outside option is the value associated with waiting one period in order to receive a constant wage stream of \( w_i \). That is, \( q_i = \beta \cdot (1 - \beta)^{-1} \cdot w_i \).

Once these outside options are substituted into (9), the resulting expression can be re-arranged to get:

\[
w_i = \left[ \frac{1}{2 - \beta} \right] \cdot s_i + \left[ \frac{1 - \beta}{2 - \beta} \right] \cdot s_j,
\]

(10)

Thus, if the pair agree to form a partnership, then the total output is divided according to \( \{w_i, s_i + s_j - w_i\} \). Since I have already incorporated the fact that workers are behaving optimally, the pair will always agree to form a partnership (i.e. if the partnership were rejected, then the rejected party is acting sub-optimally in bothering to show up in the first place). Again, positive assortative matching in equilibrium means that \( s_i = s_j \) in equilibrium - implying that \( w_i = w_j = s_i = s_j \). If we let \( \hat{\phi} = (1 - \beta)/(2 - \beta) \), then equation (10) indicates that the equilibrium will share the qualitative features of the equilibrium with non-transferable utility in which spillovers equal \( \hat{\phi} \). Thus, the transferable utility case can be thought of as one way to endogenize the spillover parameter. Note that \( \hat{\phi} \in (0, 1/2] \), where higher patience lowers \( \hat{\phi} \). Unlike the non-transferable case, changes in patience have a qualitative effect on equilibrium investments since patience determines relative bargaining power. Whilst there is always over-investment, investments approach the efficient investment as patience approaches perfect patience (\( \beta \to 1 \)). This is because worker \( i \) perceives that he will be able to appropriate the full marginal return on his investment.

Result 5 When utility is transferable, equilibrium investment behaviour is qualitatively affected by workers’ degree of patience, \( \beta \). Equilibrium outcomes in the dynamic model share the qualitative features of a static model in which \( \phi = (1 - \beta)/(2 - \beta) \). In particular, investment becomes efficient as patience becomes perfect (\( \beta \to 1 \)).

6 Conclusions

The model developed here reconciles two common, seemingly opposing, views of education: first, the externality view that education entails positive benefits to others, and second, the credentialist view
that education is, at least in part, wasteful because educational attainment is motivated more by the fact that it offers one a credential than by the fact that it makes one more productive. The analysis shows how the two views are not necessarily separate phenomena that need to be weighed against each other, but rather that credentialism actually relies on the existence of spillovers, implying that there is a more structural connection between the viewpoints than has previously been recognized.

The model allows us to draw out some implications of a greater degree of interaction in the workplace. Despite being parameterized in a ‘neutral’ manner, higher spillovers are shown to raise productivity and inequality, but lower welfare. This is suggestive of a novel way in which to interpret the impact of spillover-conducive technologies: productivity and incomes are raised because there are greater incentives for individuals to invest in their skill level, not because the technology enhances existing skills per se. Far from being welfare-improving, these rising productivity and income levels are associated with lower utility levels since greater spillovers exacerbate a type of ‘rat-race’, especially among the relatively high ability workers.

The conclusions reached in this paper must be taken within the context of the model’s limitations. Specifically, I make no attempt at arguing that spillovers of a more global nature are unimportant and effective education policy clearly needs to take these types of spillovers into account. My goal is simply to point out that making a seemingly small, but realistic and relevant, departure in the modeling of spillovers forces us to dramatically re-evaluate the role of education in a modern economy.
7 Appendix

7.1 Proofs

Proof of Lemma 1.

**Proof.** For any given \( \mu, \) optimality implies that investment is a non-decreasing function of ability. To see this, consider two investments, \( x \) and \( x' > x, \) and two types \( \theta \) and \( \theta' > \theta. \) If \( \theta \) prefers \( x' \) to \( x, \) then so too will \( \theta'. \) This is because if

\[
(1 - \phi) \cdot [s(x', \theta) - s(x, \theta)] + \phi \cdot [\mu(x') - \mu(x)] - [c(x') - c(x)] \geq 0,
\]

then

\[
(1 - \phi) \cdot [s(x', \theta') - s(x, \theta')] + \phi \cdot [\mu(x') - \mu(x)] - [c(x') - c(x)] \geq 0,
\]

since \( s(x', \theta') - s(x, \theta') > s(x', \theta) - s(x, \theta) \) (by virtue of \( x' > x \) and the property that \( s_{x\theta} > 0). \)

Given this, rational expectations implies that beliefs are non-decreasing in investment. That is, all workers prefer to be matched with those that invest more. Suppose that matching was not positive assortative. Then there is some \( x \) such that \( m(x) = x'. \) If \( x' < x \) then workers that invest \( x \) prefer to match among themselves than according to \( m. \) By feasibility, at least some workers that invest \( x' \) will match with workers that invest \( x. \) If \( x' > x \) then workers that invest \( x' \) would prefer to match among themselves than according to \( m. \) Therefore, the only stable matching is positive assortative (which is also feasible). ■

Proof of Lemma 2.

**Proof.** In equilibrium, an investment of \( x(\theta') \) produces a productivity of \( y(s(x(\theta'), \theta), s(x(\theta'), \theta')). \) By revealed preference arguments:

\[
y(s(x(\theta), \theta), s(x(\theta), \theta)) - c(x(\theta)) \geq y(s(x(\theta'), \theta), s(x(\theta'), \theta')) - c(x(\theta')),
\]

and

\[
y(s(x(\theta'), \theta'), s(x(\theta'), \theta')) - c(x(\theta')) \geq y(s(x(\theta), \theta'), s(x(\theta), \theta)) - c(x(\theta)).
\]

Together, these imply:

\[
y(s(x(\theta), \theta'), s(x(\theta), \theta)) - y(s(x(\theta'), \theta), s(x(\theta'), \theta')) \leq c(x(\theta)) - c(x(\theta')) \leq y(s(x(\theta), \theta), s(x(\theta), \theta)) - y(s(x(\theta'), \theta'), s(x(\theta'), \theta')).
\]

Define the function \( \Lambda(q, r) \equiv y(s(x(q), r), s(x(q), q)), \) and note that since \( y \) and \( s \) are differentiable, \( \Lambda_r \) is well-defined. The above conditions can be written as

\[
- [\Lambda(\theta', \theta') - \Lambda(\theta, \theta')] \leq c(x(\theta)) - c(x(\theta')) \leq \Lambda(\theta, \theta) - \Lambda(\theta', \theta).
\]

By adding \( \Lambda(\theta', \theta) - \Lambda(\theta, \theta') \) to each term, we get:

\[
- [\Lambda(\theta', \theta') - \Lambda(\theta', \theta)] \leq c(x(\theta)) - c(x(\theta')) + \Lambda(\theta', \theta) - \Lambda(\theta, \theta') \leq \Lambda(\theta, \theta) - \Lambda(\theta, \theta').
\]
Dividing each term by $\theta - \theta'$ gives:
\[
\frac{\Lambda(\theta', \theta') - \Lambda(\theta', \theta)}{\theta'-\theta} \leq \frac{c(x(\theta')) - c(x(\theta)) + \Lambda(\theta', \theta) - \Lambda(\theta, \theta')}{\theta - \theta'} \leq \frac{\Lambda(\theta, \theta) - \Lambda(\theta', \theta')}{\theta - \theta'}. 
\]
Taking the limit as $\theta' \rightarrow \theta$ gives:
\[
\Lambda_r(\theta, \theta) \leq \lim_{\theta' \rightarrow \theta} \frac{c(x(\theta')) - c(x(\theta)) + \Lambda(\theta', \theta) - \Lambda(\theta, \theta')}{\theta - \theta'} \leq \Lambda_r(\theta, \theta).
\]
Since it is ‘sandwiched’, the middle term must also equal $\Lambda_r(\theta, \theta) = (1 - \varphi) \cdot s_{\theta}(x(\theta), \theta)$. Adding and subtracting $\Lambda(\theta, \theta)/\theta - \theta'$ to the middle term gives:
\[
\lim_{\theta' \rightarrow \theta} \frac{c(x(\theta)) - c(x(\theta'))}{\theta - \theta'} - \frac{\Lambda(\theta, \theta) - \Lambda(\theta', \theta)}{\theta - \theta'} + \frac{\Lambda(\theta, \theta) - \Lambda(\theta, \theta')}{\theta - \theta'},
\]
which is the definition of
\[
\frac{\partial}{\partial \theta} c(x(\theta)) = \Lambda_q(\theta, \theta) + \Lambda_r(\theta, \theta).
\]

For these derivatives to exist, $x$ must be differentiable.

As a side note, the differential equation that the investment function must satisfy can also be derived by noting that the ‘sandwiching’ of the middle term means that $\frac{\partial}{\partial \theta} c(x(\theta)) = \Lambda_q(\theta, \theta)$. Adding and subtracting $\Lambda(\theta, \theta)/\theta - \theta'$ to the middle term gives:
\[
\frac{\partial}{\partial \theta} c(x(\theta)) = \Lambda_q(\theta, \theta) + \Lambda_r(\theta, \theta).
\]

Once this is expanded out, a differential equation of the form $x_\theta(\theta) = \hat{\Gamma}(x, \theta)$ arises. Furthermore, $\hat{\Gamma}(x, \theta) = \Gamma(\theta, x)^{-1}$, where the latter function is the differential equation governing the inverse investment function derived in the text. This verifies that the two approaches lead to the same conclusion.

**Proof of Proposition 1.**

**Proof.** I have showed in the text that all separating equilibria must solve the initial values problem. Furthermore, all solutions to the initial values problem are equilibria if the solution is a strictly increasing function and the second-order condition is satisfied. Thus, first I show that the initial values problem has a unique solution. Second, since I have showed that this solution is strictly increasing in the text, I verify that the second-order condition is satisfied.

That a unique solution exists to the initial values problem follows from the observation that $\Gamma(x, \xi)$ is a well-defined first-order linear differential equation for $x_0 > 0$ since $s(x, \theta) = \theta \cdot g(x)$. To see this, let $a(x) = -\frac{g'(x)}{g(x)}$ and $b(x) = \frac{g(x)}{\frac{g'(x)}{g(x)}}$, so that we can write:
\[
\xi'(x) = a(x) \cdot \xi(x) + b(x).
\]

The solution to this first-order linear differential equation is:
\[
\xi(x) = \left[ K + \int^x b(z) \cdot \exp \left( - \int^z a(t) dt \right) dz \right] \cdot \exp \left( \int^x a(z) dz \right),
\]

34
where $K$ is a constant that adjusts so that the initial condition is satisfied and the notation $\int^x f(z)dz$ represents the indefinite integral of $f(x)$. Note that $\int^x a(z)dz = -(1/\phi) \ln(g(x))$, so that $\exp(\int^x a(z)dz) = g(x)^{-1/\phi}$, which lets us write:

$$
\xi(x) = \left[ K + \int^x b(z) \cdot g(z)^{1/\phi} \, dz \right] \cdot g(x)^{-\frac{1}{\phi}}
$$

For any given $\{\xi_0, x_0\}$, we know that:

$$
K = \xi_0 \cdot g(x_0)^{1/\phi} - \frac{1}{\phi} \int^{x_0} c'(z) \cdot g(z)^{1-\frac{1}{\phi}} \, dz,
$$

so that:

$$
\xi(x) = \left[ \xi_0 \cdot g(x_0)^{1/\phi} + \frac{1}{\phi} \int^{x_0} c'(z) \cdot g(z)^{1-\frac{1}{\phi}} \, dz \right] \cdot g(x)^{-\frac{1}{\phi}}
$$

Finally, the initial condition is that $x_0$ satisfies $\xi_0 = \theta = c_x(x_0)/g_x(x_0)$, which gives the result in the text once the substitution is made.

The first term in the expression goes to zero as $x_0$ goes to zero, implying that the expression also represents an inverse investment function when $x_0 = 0$. Geometric arguments using the direction field can be used to show that this is the unique solution. Suppose there was some other function, $\hat{\xi}$, such that $\hat{\xi}(x) = \Gamma(x, \hat{\xi})$ and $\hat{\xi}(x_0) = c_x(x_0)/g_x(x_0) = 0$. For some $x > 0$, if $\hat{\xi}(x) < \xi(x)$, then there will exist some $x' \in (0, x]$ such that $\hat{\xi}(z) < 0$ for $z \in (0, x')$, which is impossible since $\hat{\xi}$ represents a type, and all types are non-negative. Similarly, if $\hat{\xi}(x) > \xi(x)$ then there exists a $x'$ such that $\hat{\xi}_s(z) < 0$ for all $z \in (0, x')$. This is not allowed either since the inverse investment function must be strictly increasing. Thus, there is a unique solution to the initial values problem for all $x_0 \geq 0$.

Either the geometric arguments used in the text, or an inspection of the actual solution reveals that the solution is strictly increasing for all $x \geq x_0$. This solution constitutes an equilibrium once the second-order condition is verified. This done in two steps. First I show that equilibrium investments constitute a local maximum, then second show that this is a global maximum because indifference curves touch the return function only once.

The curvature of the objective function at the equilibrium investment is:

$$
\frac{\partial^2}{\partial x^2} u(x, \theta, \mu(x)) = \frac{\partial}{\partial x} \left\{ (1 - \phi) \cdot s_x(x, \theta) + \phi \cdot [s_x(x, \xi) + s_x(x, \xi) \cdot \xi_x] - c_x \right\} - c_x
$$

$$
= \frac{\partial}{\partial x} \left\{ (1 - \phi) \cdot s_x(x, \theta) + \phi \cdot s_x(x, \xi) - s_x(x, \xi) \right\}
$$

$$
= (1 - \phi) \cdot \frac{\partial}{\partial x} \left\{ s_x(x, \theta) - s_x(x, \xi) \right\}
$$

$$
= (1 - \phi) \cdot [s_{xx}(x, \theta) - s_{xx}(x, \xi) - s_{xx}(x, \xi) \cdot \xi_x]
$$

$$
= -(1 - \phi) \cdot s_{x\theta}(x, \theta) \cdot \xi_x,
$$
which is strictly negative for \( x > x_0 \) since \( s_x(\theta, x, \theta) > 0 \). The investment is therefore a local maximum for all types \( \theta > \theta_0 \). By setting the return function suitably for investments below \( x(\theta) \) we can ensure that the investment of the lowest type is also a local maximum.

Next, I claim that each worker’s indifference curve touches the return function only once. This is a consequence of single-crossing: suppose to the contrary that there exists some \( \theta' \neq \theta \), such that type \( \theta' \)’s indifference curve coincides with \( \mu(x') \), where \( x' = x(\theta') \) for some \( \theta' \neq \theta \). Now, if \( \theta' > \theta \), then the indifference curve of a worker of type \( \theta' \) is flatter than the indifference curve of a type \( \theta \) worker at the point \((x', \mu(x'))\). But optimality requires that the slope of the former indifference curve equal the slope of \( \mu \) at this point. Thus, \( \mu \) has a smaller slope than the slope of the indifference curve of a type \( \theta \) worker. This implies that \( \mu \) crosses the indifference curve of a type \( \theta \) worker ‘from above’. But since \( \mu \) is continuous (on the set of equilibrium investments) and is below the indifference curve of a type \( \theta \) worker for investments in the neighborhood of \( x(\theta) \), any point at which the indifference curve coincides with \( \mu \) must have the property that \( \mu \) crosses the indifference curve ‘from below’, which is a contradiction. A similar argument applies if \( \theta' < \theta \): single-crossing implies that the slope of indifference curve of type \( \theta \) workers is flatter than the slope of the indifference curve of a type \( \theta' \) worker. This implies that \( \mu \) crosses the indifference curve of a type \( \theta \) worker ‘from above’ (going right to left), whereas we know that \( \mu \) must cross ‘from below’ (again, going right to left). Again, a contradiction. From this we can conclude that the proposed investments represent solutions to the maximization problem (this was only shown for deviations to other equilibrium investments, but we need not worry about off-equilibrium investments because the return function can always be chosen suitably - equal to \( s(x, \theta) \) for example). ■

**Proof of Proposition 2.**

**Proof.** By construction, the lowest ability workers invest efficiently. So focusing on higher types we have already used geometric arguments to show that \( \xi'(x) > 0 \) for all \( x > x_0 \). Notice that since \( s_x(x, \xi(x)) = c_x(x) \), we can write:

\[
\xi'(x) = \frac{s_x(x, \xi'(x)) - s_x(x, \xi(x))}{\phi \cdot s_\theta(x, \xi(x))}.
\]

Since \( \xi'(x) > 0 \) for \( x > x_0 \), the right side must also be positive for \( x > x_0 \). This implies that \( \xi'(x) > \xi(x) \) since \( s_x(x, \theta) \) is strictly increasing in \( \theta \). The inequality implies over-investment.

For the second part, let \( \xi(x; \phi) \) be the inverse investment function when spillovers are \( \phi \). We know that \( \xi(x; \phi) \) must always be finite for any \( \phi \in (0, 1] \) when evaluated at any equilibrium investment (since \( \xi(x) \) must be no greater than \( \theta_0 \)). But since \( \xi(x; \phi) = \xi_0 + \int_{x_0}^{x} \xi'(z; \phi) \cdot dz \), it must be that \( \xi'(x; \phi) \) is finite for all \( \phi \in (0, 1] \) when evaluated at almost all equilibrium investments. This then requires that

\[
\frac{1}{\phi} \frac{s_x(x, \xi'(x)) - s_x(x, \xi(x; \phi))}{s_\theta(x, \xi(x; \phi))}
\]

(11)
is finite for all \( \phi \in (0, 1] \) for almost all \( x \). But the first term goes to infinity as \( \phi \) goes to zero, which means that the second term must go to zero. The denominator is always positive and finite (and in fact, independent of \( \phi \) since \( s_\theta(\cdot) = g(x) \) is independent of \( \xi \)), so therefore
\[s_x(x, \xi^*(x)) - s_x(x, \xi(x; \phi))\] goes to zero as \(\phi\) goes to zero. Since \(s_x(x, \theta)\) is strictly increasing in \(\theta\), this implies that \([\xi^*(x) - \xi(x; \phi)]\) approaches zero when evaluated at almost all equilibrium investments. The continuity and non-decreasing nature of both \(\xi\) and \(\xi^*\) ensures that this converges occurs for all equilibrium investments. But this says that equilibrium investments approaches the efficient investments in the limit as spillovers disappear (investments are always efficient when spillovers are zero).

**Proof of Proposition 3.**

**Proof.** The proof follows immediately from noticing that the inverse investment function, \(\xi(x)\), given by (1), is an increasing function of \(x_0\). The derivative of \(\xi(x)\) with respect to \(x_0\), once simplified, equals \(g(x_0)^{1/\phi} \cdot \partial / \partial x_0 \{c_x(x_0)/g_x(x_0)\}\), which is positive.

**Proof of Proposition 4.**

**Proof.** There is a simple geometric proof. Consider the direction field and note that a higher value of \(\phi\) lowers the slope of \(\xi\) at every \((\xi, x)\) pair. Thus, \(\xi(x; \phi')\) must cross \(\xi(x; \phi)\) 'from above'. Since \(\xi(x_0; \phi') = \xi(x_0; \phi)\), it follows that \(\xi(x; \phi')\) must lie below \(\xi(x; \phi)\) for all \(x > x_0\).

### 7.2 Complementary Interaction

Let worker interaction take the following CES form:

\[
y(s, s') = [(1 - \phi) \cdot s^\rho + \phi \cdot s'^\rho]^{\frac{1}{\rho}},
\]

for \(\rho \in (\infty, 1]\). This form captures complementarity because:

\[
y_{ss'}(s, s') = (1 - \phi) \cdot \phi \cdot (1 - \rho) \cdot (s s')^{\rho - 1} \left[(1 - \phi) \cdot s^\rho + \phi \cdot s'^\rho\right]^{\frac{1}{\rho} - 2},
\]

which is non-negative (strictly positive for \(\rho < 1\)). Furthermore, notice that

\[
y_s(s, s')|_{s'=s} = (1 - \phi) \cdot s^{\rho - 1} \left[(1 - \phi) \cdot s^\rho + \phi \cdot s'^\rho\right]^{\frac{1}{\rho} - 1} = (1 - \phi),
\]

and

\[
y_{s'}(s, s')|_{s'=s} = \phi \cdot s'^{\rho - 1} \left[(1 - \phi) \cdot s^\rho + \phi \cdot s'^\rho\right]^{\frac{1}{\rho} - 1} = \phi.
\]

These expressions can be substituted into the general expression for \(\Gamma(\xi, x)\) to get:

\[
\Gamma(\xi, x) = \frac{c_x(x) - s_x(x, \xi)}{\phi \cdot s_\rho(x, \xi)}.
\]

(12)

This is exactly the same expression as that derived in the additively separable case (which is not surprising given that this case is the special case in which \(\rho = 1\)). The initial condition is determined by the efficient investment for the lowest types. This is the value of \(x_0\) that satisfies:

\[
[y_s(\cdot) + y_{s'}(\cdot)] \cdot s_x(x, \theta) = c_x(x_0).
\]
Applying the expressions derived for \( y_s(\cdot) \) and \( y_x(\cdot) \) gives us that \( x_0 \) satisfies \( s_x(x_0, \theta) = c_x(x_0) \), which is identical to the additively separable case.

I conclude that the equilibrium investment function is the same for all values of \( \rho \in (−\infty, 1] \) since the initial values problem is identical. Thus, the equilibrium analyzed in body of the paper \((\rho = 1)\) also applies to any other function in the CES class, such as Cobb-Douglas \((\rho = 0)\) and, in the limit, Leontief \((\rho \to −\infty)\).

The main benefit of assuming \( \rho = 1 \) in the body of the paper (rather than allowing for a general CES function) is that it simplifies the discussion since \( \mu(x) \) can be interpreted as the ‘expected skill of a coworker that invests \( x \).’

### 7.3 Policy

Given the over-investment in equilibrium, it will never be the case that an investment subsidy is optimal despite the positive spillover associated with investment. In order to characterize the optimal policy, suppose a planner is able to charge an investment tax of \( \tau(x) \). Incorporating this tax into the model is straightforward since we can derive the equilibrium under the assumption that investment costs are \( \tilde{c}(x) = c(x) + \tau(x) \). That is, the inverse investment function becomes:

\[
\xi(x) = \frac{\theta}{g(x)} \left[ \frac{g(x_0)}{g(x)} \right]^{\frac{1}{\phi}} + \frac{1}{\phi} \int_{x_0}^x \frac{c_x(z) + \tau_x(z)}{g(z)} \cdot \left[ \frac{g(z)}{g(x)} \right]^{\frac{1}{\phi}} dz.
\]

The tax schedule is chosen so that the induced inverse investment function coincides with the inverse efficient investment schedule: \( \xi(x) = \xi^*(x) \). We know that \( \xi^*(x) = c_x(x)/g_x(x) \), so the optimal tax schedule must satisfy:

\[
g(x)^{\frac{1}{\phi}} \frac{c_x(x)}{g_x(x)} - \frac{\theta}{g(x_0)} g(x)^{\frac{1}{\phi}} = \frac{1}{\phi} \int_{x_0}^x [c_x(z) + \tau_x(z)] \cdot g(z)^{\frac{1}{\phi}} dz.
\]

If we define the left side to be \( L(x) \), then we can differentiate both sides to get that:

\[
\tau_x(x) = \frac{\phi \cdot L'(x)}{g(x)^{\frac{1}{\phi}}} - c_x(x),
\]

where \( L'(x) = (1/\phi) \cdot g(x)^{(1-\phi)/\phi} \cdot c_x(x) + g(x)^{1/\phi} \cdot [\partial/\partial x \{ c_x(x)/g_x(x) \}] \). It follows then that:

\[
\tau_x(x) = \phi \cdot g(x) \cdot \left[ \frac{\partial}{\partial x} \left\{ \frac{c_x(x)}{g_x(x)} \right\} \right] = \phi \cdot g(x) \cdot \xi_x^*(x) > 0.
\] (13)

Since the lowest types invest efficiently, we also have that \( \tau(x_0) = 0 \). The solution to the initial values problem defined by (13) and the initial condition constitutes the optimal taxation policy. It is immediate that higher spillovers increase the optimal tax at each investment level. Furthermore, if the investment function were altered by a technology term, \( A^g \), such that \( g(x) = A^g \cdot \tilde{g}(x) \), then the optimal tax rate is independent of \( A^g \). If investment costs are similarly augmented by \( A^c \), then the optimal tax is also scaled up by a factor of \( A^c \). The last two results come from the fact that \( \xi(x)/\xi^*(x) \) is independent of both \( A^g \) and \( A^c \). Therefore, changing \( A^g \) has no effect on relative inefficiency and therefore on optimal tax policy. If both components of total investment costs - the
private costs and the tax - are raised in the same proportion, then optimal policy is unaffected. Therefore, raising the private cost alone must be accompanied by a proportional raise in the total tax cost.

The optimal tax policy for the illustration presented in Section 4 can be derived by using (13) along with the initial condition. In this case the optimal policy is a (piecewise) linear tax given by:

\[
\tau^*(x) = \begin{cases} 
0 & \text{if } x \in [0, x_0) \\
\phi \cdot c \cdot \frac{1-n}{\eta} \cdot [x - x_0] & \text{if } x \geq x_0.
\end{cases}
\]

The optimal tax rate is non-zero if and only if spillovers exist. Furthermore, the optimal marginal tax rate is increasing in marginal investment costs and approaches infinity as investment becomes unproductive (as \(\eta \to 0\)).

### 7.4 A Model With Classes

The model presented above can be generalized to a setting in which each agent belongs to one of two classes, where a match requires one agent from each class. The classic example is marriage, where the classes are males and females. Let there be a unit measure of both males and females. Each male \(i\) is endowed with a type, \(\theta_i \in \Theta\). The distribution of male types is \(G_M(\cdot)\), where \(G'_M(z) \in (0, \infty)\) for all \(z \in \Theta\). Each female \(j\) is endowed with a type \(\tilde{\theta}_j \in \tilde{\Theta}\). The distribution of female types is \(G_F(\cdot)\), where \(G'_F(z) \in (0, \infty)\) for all \(z \in \tilde{\Theta}\).

I am going to restrict attention to fully revealing equilibria. Since investment is increasing in type, equilibrium will involve a type \(\theta\) male being matched with a type \(h(\theta)\) female, where \(G_M(\theta) = G_F(h(\theta))\). If males perceive a return function of \(\mu(x)\), and females perceive a return function of \(\tilde{\mu}(x)\), then rational expectations requires:

\[
\mu(x(\theta)) = s(\hat{x}(h(\theta)), h(\theta)) \\
\tilde{\mu}(\hat{x}(h(\theta)))) = s(x(\theta), \theta),
\]

where \(x(\cdot)\) and \(\hat{x}(\cdot)\) are the equilibrium investment functions for males and females respectively. Given these return functions, the investment functions must be optimal:

\[
x(\theta) = \arg\max_{x \geq 0} u(x, \theta, \mu(x)), \; \forall \theta \in \Theta \\
x(h(\theta)) = \arg\max_{x \geq 0} u(x, h(\theta), \tilde{\mu}(x)), \; \forall \theta \in \Theta.
\]

If we define \(\hat{x}(\theta) \equiv \hat{x}(h(\theta))\), then the the same procedure as described above can be employed, and we end up with a system of first-order differential equations:

\[
x'(\theta) = \Gamma(x, \hat{x}, \theta) \\
\hat{x}'(\theta) = \hat{\Gamma}(x, \hat{x}, \theta).
\]

This system, along with the initial conditions, \([x(\theta), \hat{x}(\theta)]\), form an initial values problem. The solution to this problem will constitute an equilibrium when both functions in the solution are strictly increasing, and the second-order condition is satisfied.
Analysis of this problem, especially the geometric arguments, is considerably more complicated than that of the case in which there is a single class because of the extra dimension.\textsuperscript{37} Phase diagrams are infeasible since the system of differential equations is non-autonomous, and direction fields need to be given a three-dimensional treatment. Note however that assuming unproductive investment makes the problem tractable since $\Gamma$ is no longer a function of $\hat{x}$ and $\hat{\Gamma}$ is no longer a function of $x$. Thus, the two equations can be treated separately according to the procedure used in the single-class case.

One way to make some progress is to use the method of undetermined coefficients. That is, employ a particular structure and guess that the solution will take a particular parameterized form, then use the optimality and rational expectations conditions to solve for the parameters. This is of course a much less general approach to the problem since particular functional forms are used, however the exercise is useful because it allows for some closed-form solutions.

Let $G_M$ be the uniform distribution on $[0,1]$, and let $G_F$ be the uniform distribution on $[0,F]$. This implies that $h(z) = z \cdot F$ for $z \in [0,1]$. Let $g(x) = x$ and let $c(x) = (c/2) \cdot x^2$. I will guess that the solution is linear: $x(\theta) = \beta \cdot \theta$ and $\hat{x}(\theta) = \hat{\beta} \cdot \theta$.

Given $\mu$ and $\hat{\mu}$, optimality requires that:

\begin{align*}
(1 - \phi) \cdot \theta + \phi \cdot \mu'(x) &= c \cdot x, \\
(1 - \phi) \cdot F\theta + \phi \cdot \hat{\mu}'(x) &= c \cdot x,
\end{align*}

and rational expectations requires:

\begin{align*}
\mu(x(\theta)) &= \hat{x}(\theta) \cdot F\theta \\
\hat{\mu}(\hat{x}(\theta)) &= x(\theta) \cdot \theta.
\end{align*}

Taking the derivative of each of these with respect to $\theta$, and substituting the conjectured form gives:

\begin{align*}
\mu'(x(\theta)) &= \frac{\hat{\beta}}{\beta} \cdot 2F\theta \\
\hat{\mu}'(\hat{x}(\theta)) &= \frac{\beta}{\hat{\beta}} \cdot 2\theta.
\end{align*}

Substituting these into the first-order conditions, and again applying the conjectured forms, gives:

\begin{align*}
(1 - \phi) + \phi \cdot \frac{\hat{\beta}}{\beta} \cdot 2F &= c \cdot \beta, \\
(1 - \phi) \cdot F + \phi \cdot \frac{\beta}{\hat{\beta}} \cdot 2 &= c \cdot \hat{\beta}F.
\end{align*}

An equilibrium of the form conjectured exists if we can find a $\beta > 0$ and a $\hat{\beta} > 0$ such that the above two conditions are satisfied. These conditions can be re-written as follows:

\begin{align*}
\hat{\beta} &= \frac{1}{2F\phi} \left[ c \cdot \beta^2 - (1 - \phi) \cdot \beta \right] \\
\beta &= \frac{F}{2\phi} \left[ c \cdot \hat{\beta}^2 - (1 - \phi) \cdot \hat{\beta} \right].
\end{align*}

\textsuperscript{37}The two-class problem is equivalent to the single-class problem when the distribution of types is the same for both classes ($h$ is the identity function).
Depicting these relationships in $(\beta, \hat{\beta})$ space easily reveals that there are unique positive values of $\beta$ and $\hat{\beta}$ that satisfy these relationships.

Are the equilibrium investments efficient? To begin, we can calculate the set of Pareto efficient investment pairs by equating the slope of indifference curves in $(x, \hat{x})$ space. This leads us to the conclusion that if the female invests $\hat{x} = \hat{\beta} \cdot \theta$ (some multiple of the male’s type), then the corresponding Pareto efficient investment from the male is $x = \kappa^* (\hat{\beta}) \cdot \theta$, where:

$$
\kappa^* (\hat{\beta}) \equiv \frac{1 - \phi}{c} + \left[ \frac{\phi}{c} \right]^2 \cdot \frac{1}{\hat{\beta} - \frac{1 - \phi}{c}}.
$$

From here we can ask the question of whether there exists an $F$ such that the resulting equilibrium investments are Pareto efficient. As it turns out, such a value does not exist. To show this, suppose that the female invested $\hat{\beta} \cdot \theta$. Then, by substituting out for $F$, it can be shown that the male will invest $\beta = \kappa (\hat{\beta}) \cdot \theta$ in equilibrium, where:

$$
\beta = \kappa (\hat{\beta}) \equiv \frac{1 - \phi}{c} + 2 \left[ \frac{\phi}{c} \right]^2 \cdot \frac{1}{\hat{\beta} - \frac{1 - \phi}{c}}.
$$

Since $\kappa (\hat{\beta}) > \kappa^* (\hat{\beta})$, it follows that equilibrium investments are always inefficiently great. In other words, there will always exist lower investments from both the male and female that results in a Pareto-improvement.

Furthermore, since higher values of $F$ correspond to lower values of $\hat{\beta}$ (which correspond to higher values of $\beta$), an increase in $F$ leads to more investment for males and less investment for females. Intuitively, an increase in $F$ increases the dispersion of female types, which provides incentives for males to compete harder at each investment level.

One interesting matter of non-existence arises when the two-sides are explicitly considered. The reason is that when the two sides become too ‘asymmetric’ - in this case a very high or very low $F$ - the less differentiated side must invest a great deal in equilibrium. This investment can in fact be so great that they would prefer making their Nash investment, even if it meant leaving them without a partner. This is under the assumption that having no partner is equivalent to having a partner of zero skill. There is no reason why this need be the case, but it does not seem unreasonable. When $F$ is too extreme, all agents on one side of the market have a profitable deviation in investing their Nash level and remaining without a partner. However, if all workers did this, then agents would still be fully separated (since Nash investment is a strictly increasing function of type), which implies that matching would still be positive assortative. Thus, there would be a profitable deviation to be had from raising investment a little in order to pretend to be a worker of higher ability. Thus, there may be no symmetric equilibrium in pure strategies. Having said this, there always exists an equilibrium when the two sides are sufficiently similar (i.e. when $F$ is sufficiently close to unity). The possibility of non-existence never arises in the single-class model because this can be thought of as a model in which the two sides have the same distribution.

\[38\]In fact, the argument goes through for any utility level above negative infinity that one wishes to assign to the state of not being matched.
7.5 Signaling With Patience

Consider the following dynamic signaling model. A generation of workers are born each period, and each is endowed with some ability, \( \theta \in \Theta \). For comparability, take \( \Theta = [\theta, \theta] \) and assume that \( \theta \) is distributed on this interval with positive density. Workers are able to invest in a signal, \( x \), at some positive increasing cost. The investment is a pure signal - a worker’s output simply equals their productivity.

In their first period of existence workers make their investment and enter the labour market. Firms do not observe ability but do observe investment, and offer workers a wage schedule, \( w(x) \). The wage contract is enforceable, implying that if a firm hires a worker that ends up producing less than their wage, their only recourse is to fire the worker. Workers select which firm they wish to work at (free entry ensures that each worker gets a job), get paid \( w(x) \) and produce output. At the end of the period the firm decides whether they wish to fire the worker. If the worker is fired, then they re-enter the labour market the following period. Finding a new job takes one period. To maintain comparability, I assume that the labour market does not observe the worker’s work history (only their investment).

Workers each have a discount factor of \( \beta \in [0, 1) \), implying that the present value of lifetime earnings given a wage stream of \( \{ w_t \}_{t=0}^{\infty} \) is \( V \equiv \sum_{t=0}^{\infty} \beta^t \cdot w_t \). The standard static model is a special case in which \( \beta = 0 \).

The efficient outcome is to have all workers investing zero, and this can not be supported by a separating equilibrium at \( \beta = 0 \). Intuitively, if all workers invested zero then all workers appear the same. Any worker can claim to be the highest type worker, and get paid \( \theta \). The firm offering this wage would fire the worker, but since the future is completely discounted this does not concern the worker. What about if \( \beta > 0 \)?

There is always a non-degenerate (and continuous) distribution of types such that it is possible to support a ‘separating’ and efficient equilibrium when \( \beta > 0 \). To see this, firms offer a continuum of wage contracts of the form \( w(x) = w' \) for each \( w' \in \Theta \). Since the wages on offer do not depend on investment, it follows that all workers will invest zero (efficiently). In essence, workers choose their own wage mindful of the fact that if they choose a wage for which they are under-qualified, then the firm will fire them and they will have to spend a period looking for a new job. If a worker chooses a wage for which they are qualified, \( w \), their payoff is \((1 - \beta)^{-1} \cdot w \). If they instead choose a wage for which they are not qualified, \( w' \), then their payoff is \( (1 + \beta) \cdot (1 - \beta)^{-1} \cdot w' \). The most attractive deviation possible is that made by the lowest type worker claiming to be the highest type worker.

---

39 An alternative assumption would be that the wage is re-negotiated, but this introduces unnecessary complications. For instance, if the firm must fire the worker before offering them the renegotiated wage (which involves the worker waiting one period), and if the worker has all the bargaining power, then the conclusions (conditions required for efficiency) are exactly the same. Giving bargaining power to the firm only makes a deviation less attractive, in which case the efficient outcome is easier to support, strengthening the argument.

40 To see this, note that the worker alternates between employment and job search each period. While employed they get \( w' \) and while unemployed get zero. The payoff stream is therefore: \( V = w' + \beta^2 \cdot w' + \beta^4 \cdot w' + ... = (1 - \beta^2)^{-1} w' = [(1 + \beta) \cdot (1 - \beta)]^{-1} \cdot w' \).
If this deviation is unprofitable, then all deviations are unprofitable. This deviation is unprofitable if $(1 + \beta) \cdot \bar{\theta} > \bar{\theta}$. Therefore efficient investments can be supported in a separating equilibrium with a continuum of types if the highest type is not too great relative to the lowest type. Note that if $\beta = 0$ (the static case) such an equilibrium can only be supported if the distribution is degenerate.

As patience increases, so too does the maximum allowable value of the highest type.

The conclusion to be taken from this is that different levels of patience can qualitatively change the nature of separating equilibria in standard signaling models with a continuum of types. This highlights a key difference between the matching model and signaling models.

### 7.6 Pooling Equilibria

This section examines equilibria in which at least some workers of different types invest the same amount. The simplest version is a complete pooling equilibrium in which all workers invest some $x_p$, as depicted in Figure 6. The rational expectations condition simply requires that $\mu$ passes through the point $(x_p, \bar{s}(x_p))$, where $\bar{s}(x) = E[\theta] \cdot g(x)$ is the expected skill of a worker randomly drawn from the population, given that all workers invest $x$.

There will in general be equilibria in which any finite number of investments are made (as in signaling models). The fact that equilibrium investments are non-decreasing implies that these equilibria can be calculated according to the following.

**Proposition 5** For any finite integer $N$, fix any partition of the type space, $\{[\theta, \theta_1), [\theta_1, \theta_2), ..., [\theta_{N-1}, \theta]\}$. There exist investments, $\{x_1, ..., x_N\}$ such that a pooling equilibrium exists in which all members in the $n$th partition invest $x_n$.

The key to demonstrating this is to use the following recursive method. Take some $x_1$, and assume that all $\theta \in [\theta, \theta_1)$ invest $x_1$. The return function must pass through the point $(x_1, \bar{s}_1)$, where $\bar{s}_1 = g(x_1) \cdot E[\theta \mid \theta \in [\theta, \theta_1)]$. From here we can draw in the indifference curve of workers of type $\theta_1$ that passes through this point. Now consider workers with types in $[\theta_1, \theta_2)$. We can depict the curve of $\bar{s}_2(x) = g(x) \cdot E[\theta \mid \theta \in [\theta_1, \theta_2)]$. This curve will cut the indifference curve of the type $\theta_1$ workers at one investment level. Simply let $x_2$ be this investment level, then repeat the same procedure until the final partition is reached. One implication of this is that $x_1 < x_2 < ... < x_N$.

Apart from these equilibria, there will be hybrid equilibria in which some types pool and some types separate.
References


Figure 1: Indifference Curves and Single-Crossing ($\theta' > \theta$)

Figure 2: Optimal investment for a type $\theta$ worker given $\mu(x)$
Figure 3: The rational expectations condition in a separating equilibrium

Figure 4: Direction Field
Figure 5: Separating Equilibrium

Figure 6: An equilibrium with complete pooling at $x_p$