*American Economic Review 2014, 104(4): 1149–1171*
*http://dx.doi.org/10.1257/aer.104.4.1149*

# Misallocation and Growth†

*By* Boyan Jovanovic*

*This paper models growth via on-the-job learning when firms and workers are heterogeneous. It is an overlapping generations model in which young agents match with the old. More efficient assignments lead to faster long-run growth, more inequality, and less turnover in the distribution of human capital. Constant-growth paths are characterized for general functional forms and then, for the Cobb-Douglas case, the transition dynamics are solved analytically when the skill of the young is log-normally distributed and the initial human capital of the old generation is also log-normal. Growth and inequality move together on the transition to the balanced growth path.* (*JEL* D83, J24, J31, J41)

This article employs an overlapping generations model to study the effect of mismatch in human capital on an economy's growth rate in a partnership production setting with skill complementarity. The main model consists of four components: a production function requiring one old agent and one young agent; an education function that determines the skill of a young agent as a function of his innate ability and the average skill of the old alive in the period; an on-the-job learning function which determines the skill of an agent when he is old; and an imperfect signal of young skill. This common, but imperfect information about young workers' abilities then guides who matches with whom, and the imperfection of this signal acts as an "assignment friction."

Abilities of the old and young are complements in production and in training, and a mismatch occurs when positive assortative matching is not in effect. The article solves for the growth rate or per capita output along the balanced path and studies the effects of changes in the signal-to-noise ratio regarding young skill on and off the balanced growth path (BGP).

The results are twofold. First, BGPs alone are compared. Better signals lead to more efficient assignments which, in turn, lead to faster long-run growth, to more income inequality, and to less turnover in the distribution of firm productivity. Second, transition dynamics are worked out analytically for a Cobb-Douglas economy with log-normal signals. Following an exogenous improvement in the quality

* Economics Department, New York University, 19 W. Fourth St., New York, NY 10012 (e-mail: bj2@nyu.edu).
The author has no financial or other material interests related to this research to disclose. He thanks Axel Anderson, Robert Lucas, and Gianluca Violante for comments, the National Science Foundation for support, and Gaston Navarro for the computations.
† Go to http://dx.doi.org/10.1257/aer.104.4.1149 to visit the article page for additional materials and author disclosure statement(s).

of signals, the efficiency of assignments rises, and growth and inequality move monotonically and together on the transition to their higher BGP values.

The signal-to-noise ratio is the parameter that the comparative dynamics focus on. A system that rewards ability instead of family background or social connections is a system in which the signal-to-noise ratio would be high. There are several historical episodes that this view helps explain. One of the changes associated with Japan's growth miracle is a shift in the hiring system away from one based on connections, and towards one based on merit. Europe differs from the United States in terms of income inequality and its education system; in Europe, the public university system probably yields a lower signal-to-noise ratio than in the United States, where a degree from a top university carries a lot more weight. The model thus offers an explanation for why the United States has outperformed Europe in the twentieth century despite being more unequal in its income distribution. The model also points to the first universities established in Italy, France, Spain, and England some 900 years ago as possible sources of better signals about ability that could have raised the quality of assignments and helped raise growth rates in the centuries that followed.

The model has three building blocks. The first is Prescott and Boyd (1987) who introduce lifetime participation constraints in an overlapping generations model; to this Chari and Hopenhayn (1991) add heterogeneous firms but keep the young generations homogeneous. The second is Becker (1973) who introduces the Becker-Brock theorem on monotone sorting that is central here. The third is MacDonald (1982) and Kremer (1993) who develop the idea of that the efficiency of assignments depends on the accuracy of public signals about agents' qualities.

Related papers are many and are discussed in Section V. The closest antecedents are Anderson and Smith (2010) and Anderson (2011) who analyze partnerships that produce both output and training, but they do not have growth. Whereas many papers consider the effect of frictions on the level of output, the present article takes a look at the effect of frictions on the growth rate. Moreover, it is one of the few to solve in closed form the general equilibrium dynamics of an economy in which the state is a distribution of capital over agents.

The policy implications are straightforward, yet different from those of the bulk of the papers on misallocation: equilibrium in the model is competitive and, in spite of an intergenerational externality in human capital accumulation, the equilibrium assignment maximizes the rate of growth of aggregate output subject to the learning friction. Equilibrium is constrained–Pareto efficient, and taxes and transfers cannot improve the allocation. If anything, policy should aim to reduce the friction.

Several recent papers focus on the allocational consequences of financing frictions—Buera and Shin (2009) and Midrigan and Xu (2014)—and a number of others focus on implicit taxes—Restuccia and Rogerson (2008) and Hsieh and Klenow (2009). Here, by contrast, the friction is the "labeling" friction of the MacDonald-Kremer type, which one could call a labor-market friction or human capital friction. In common with some search models, this model assumes that an agent can learn through experience, so that each old agent knows his ability. However, a young agent cannot rematch until he reaches old age, so that work *versus* search decision that Burdett and Coles (1997) and Shimer and Smith (2000) analyze in contexts where agents know their types but must search for one another randomly does not exist here.

*Plan of Paper*.—Section I introduces the model and characterizes equilibrium in general and then studies two extreme cases, namely, perfect sorting and random sorting. Section II solves the full dynamics for the case in which the production and training functions are Cobb-Douglas, and the distributions of human capital and of the signals log-normal. When the signals are better, the learning friction is smaller and the assignment is less "garbled." Section III draws out the implications of the friction for the dynamics, for growth and inequality, and for turnover in the TFP distribution. Section IV discusses the literature in more detail and outlines the policy implications and some limitations of the analysis. Section V concludes. The online Appendix contains proofs and computation details.

## I. Model

The model has infinitely many two-period overlapping generations. An agent's lifetime utility depends on consumption in youth, $c_y$, and consumption in old age, $c_o$:

$$c_y + \beta c_o.$$

Population is fixed, with a unit measure of young being born each period. An agent is endowed with a unit of labor in both periods of life.

*Firms*.—A firm uses the quality, $x$, of an old agent and the quality, $y$, of a young agent to produce output, $q$, and training, $x'$. Training is measured in units of skill that the young agent will have when he gets old. The production functions are

$$\text{output} \qquad q = f(x, y), \tag{1}$$

$$\text{training} \qquad x' = \phi(x, y), \tag{2}$$

where $f$ and $\phi$ are homogeneous of degree one, increasing, and have positive cross partials. Thus, a firm must have one old and one young worker. A justification for assuming this is that the team manager must be more experienced than the workers he trains.

*State of the System*.—The state of the system is the distribution function (CDF) of skill among the members of the old generation, denoted by $H(x)$. The law of motion of this aggregate state is a map from the set of distributions $H$ into itself. Notation is simpler, however, if equilibrium objects are indexed by $t$ rather than by $H$, and we shall stress the Markovian nature of the equilibrium only when we display the log-normal dynamics. Let $\bar{x}_t = \int x \, dH_t(x)$ be the mean of $H_t$.

*Innate Talent versus Working Skill*.—Let $\varepsilon$ denote the talent at birth of a young agent. Talent is distributed according to the C.D.F. $\Psi(\varepsilon)$, the same for each generation. A young agent's working skill $y$ depends on his innate talent $\varepsilon$, and, externally, on the average skill of the old, $\bar{x}$, as follows:[1]

$$y = b\bar{x}_t\varepsilon > 0. \tag{3}$$

The resulting C.D.F. of $y$ at date $t$ is denoted by

$$G_t(y) = \Psi\left(\frac{y}{b\bar{x}_t}\right). \tag{4}$$

To summarize, production is complementary in the two agents' skills. A young worker's skill depends on his innate ability and the average ability of older workers in the economy—an influence of the human capital of all the old combined for which the old receive no direct compensation. Additionally, old workers train the young ones, so the skill of a worker when old depends positively on the skill of his trainer. We now introduce an assignment friction in the form of common, but imperfect, information about young workers' abilities. There is a publicly observable signal about young workers' skills. Older workers' skills are known (or become known since output in their first period is observed).

*Public Signal of y.*—Let $s$ be a public signal of $y$, such as a "test score" having the C.D.F. $F(s|y)$ that has a density. This signal is all that anyone—including the young agent himself—knows about $y$. By contrast, an old agent's skill is publicly known, being revealed by the firm's previous-period output.[2] We shall assume that the posterior on $y$ given $s$,

$$d\tau_t(y|s) = \frac{dF(s|y)\, dG_t(y)}{\int dF(s|y')\, dG_t(y')}, \tag{5}$$

is stochastically increasing in $s$.[3] Since $F$ has a density, so does $\tau$. In that case,

$$\mathbf{E}_t(q|x, s) = \int f(x, y)\, d\tau_t(y|s) \quad \text{and} \quad \mathbf{E}_t(x'|x, s) = \int \phi(x, y)\, d\tau_t(y|s)$$

are increasing in $s$, and the cross partials $\partial^2 \mathbf{E}_t(q)/\partial x \partial s > 0$ and $\partial^2 \mathbf{E}_t(x')/\partial x \partial s > 0$, as will be clear in the log-normal case. The population of young workers by their observed type, $s$, then has the C.D.F.

[1] We may interpret $b$ along the lines of equation (10) of Becker and Tomes (1979) which posits a positive effect of "social capital" or "culture" on the human capital of the young, or Luttmer (2007) where an entering firm tries to imitate a randomly sampled incumbent.

Why is the external effect needed here? If $y$ was homogeneous, we could replace (1) and (2) by $q = Ax\hat{f}(y)$ and $x' = x\hat{\phi}(y)$, we could drop (3), and $x$ and $q$ would then grow at the constant rate $\hat{\phi}(y)$. But if there is more than one type $y$ and if $\hat{\phi}$ is increasing in $y$, firms $x$ that employ the high $y$ types would grow faster than others, and the coefficient of variation of $x$ would go to infinity. With (3) combined with diminishing returns to both $x$ and $y$, growth and inequality both settle down on the BGP.

[2] Given $x$, observing output $q = f(x, y)$ perfectly reveals $y$, and this information is then sufficient to reveal $x' = \phi(x, y)$. Thus if the date-zero $x$s are known, they are known in each subsequent period.

[3] In the log-normal application, this will certainly be true since $\tau$ will itself be log normal with a mean that is increasing in $s$ and a variance that is independent of $s$.

$$\Phi_t(s) = \int F(s|y)\, dG_t(y). \tag{6}$$

Thus $\Phi$ becomes the distribution of types of the young which are to be matched to the old that have distribution $H$.

*Market Structure*.—The old agent owns the firm and consumes its profits. The young worker takes over the firm when he gets old. There are no financial assets. Firms take as given the generation-$t$ workers' lifetime utility function denoted by $V_t(s)$. In other words, to attract a worker of type $s$ the firm must deliver a lifetime utility of at least $V_t(s)$. Moreover, the abilities $x$ of the old agents are publicly known at the time of contracting, which means that in calculating his own lifetime utility the worker will take into account not only the wage that he will receive, but also the training that firm of type $x$ is capable of providing.

*Positive Assignment*.—A positive assignment $s \equiv \alpha_t(x)$ solves the demand = supply equation $\Phi_t(s) = H_t(x)$ for all $x$, so that the market-clearing assignment is

$$\alpha_t(x) = \Phi_t^{-1}(H_t(x)). \tag{7}$$

*Lifetime Utility*.—Agents are risk neutral. With the assignment specified in (7), let $w_t(x)$ be the amount that firm $x$ chooses to pay the worker it has hired—that decision problem is defined below in (9)–(10). The object $w_t(\cdot)$ then uniquely determines agents' lifetime utilities. An agent born at date $t$ of type $s$ has a lifetime utility

$$V_t(s) = w_t(\alpha_t^{-1}(s)) + \beta \int \pi_{t+1}(\phi[\alpha_t^{-1}(s), y])\, d\tau_t(y|s), \tag{8}$$

where $\pi_{t+1}(x)$ is his old-age consumption, and the profit of the firm that he manages.

*Old-Age Consumption*.—A type-$x$ old agent chooses a type of young worker to hire, $s$, and the wage, $w$, to maximize his firm's profit

$$\pi_t(x) = \max_{w,\, s} \left\{ \int f(x, y)\, d\tau_t(y|s) - w \right\}, \tag{9}$$

subject to the participation constraint of the young

$$w + \beta \int \pi_{t+1}(\phi[x, y])\, d\tau_t(y|s) \geq V_t(s), \tag{10}$$

taking $V_t(s)$ as given.[4]

[4] In equilibrium the solution to the problem (9)–(10) for the pair $(w, s)$ is $(w_t(x), \alpha_t(x))$. These two functions can be used to define a wage as a function of the worker's type $s$. Call that function $W_t(s)$. One could have used $W$ instead of $w$ as an object of equilibrium; the two are equivalent.

*Law of Motion for the Aggregate State*.—In (4), since $\bar{x}$ depends on $t$, so does $G$ and, hence, so does $\tau$ in (5). An assignment $\alpha$ determines the following law of motion for $H$:

$$H_{t+1}(x') = \int_{\{(x,\, y)\,|\,\phi(x,\, y)<x'\}} d\tau_t(y\,|\,\alpha_t(x))\; dH_t(x). \tag{11}$$

## A. *Equilibrium*

Equilibrium consists of five functions of $(t, x)$ that satisfy the following conditions:

(i) $(\alpha, H)$ solve (7) and (11) subject to an initial condition $H_0$.

(ii) $(\pi, w, V)$ solve (8)–(10).

Since (7) and (11) do not include $(\pi, w, V)$, the solution algorithm first takes step (i) and solves for $(\alpha, H)$. Then step (ii) yields $(\pi, w, V)$. Step (i) is valid only if the positive assignment is incentive compatible; we shall check that ex post.

## B. *General Characterization*

The criterion (9) states that an old agent cares only about his current consumption and chooses $s$ so as to maximize it. In this sense his decision problem is myopic. But because the participation constraint (10) is forward looking, the old agent in fact chooses $s$ so as to maximize an expression resembling the discounted sum of his firm's expected profits. To whichever worker-type $s$ he chooses to hire, the old agent will pay a wage $w$ no higher than one that satisfies (10) with equality. Eliminating $w$ from the choice set in (9) reduces the old agent's problem to the unconstrained one of choosing just $s$:

$$\pi_t(x) = \max_s \left\{-V_t(s) + \int \left[f(x, y) + \beta\pi_{t+1}(\phi[x, y])\right] d\tau_t(y\,|\,s)\right\}. \tag{12}$$

Equation (12) is the expression for the equilibrium consumption of the old agent. Because (12) contains two unknown functions, $\pi$ and $V$, we can describe things more compactly in terms of the wage function. Proposition 1 provides a partial characterization of $w$ in the general case.[5]

PROPOSITION 1: *$w$ is differentiable and satisfies the functional equation*

$$w_t'(x) = U_t(x) + \beta \int w_{t+1}'(\phi[x, y])\,\phi_x(x, y)\; d\tau_t(y\,|\,\alpha_t[x]), \tag{13}$$

[5] To shorten the notation, we write $\tau_t(y|s = \alpha_t[x])$ as $\tau_t(y|\alpha_t[x])$ and $\tau_{t+1}(y'|s' = \alpha_{t+1}[x'])$ as $\tau_{t+1}(y'|\alpha_{t+1}[x'])$.

*in which*

$$(14)\ U_t(x) = \alpha_t'(x)\int f(x, y)\frac{\partial}{\partial s}d\tau_t(y\,|\,\alpha_t[x])$$
$$- \beta\int\phi_x(x, y)\int f_x(\phi[x,y],y')\,d\tau_{t+1}(y'\,|\,\alpha_{t+1}(\phi[x,y]))\,d\tau_t(y\,|\,\alpha_t[x])$$
$$- \beta\int\alpha_{t+1}'(\phi[x,y])\phi_x(x,y)\int f(\phi[x,y],y')\frac{\partial}{\partial s'}d\tau_{t+1}(y'\,|\,\alpha_{t+1}[\phi(x, y)])\,d\tau_t(y\,|\,\alpha_t[x])$$

*and in which $\alpha$ satisfies* (7). *A unique bounded solution for $w'$ exists if $U$ is bounded and if*

$$(15)\qquad \beta\sup_{x,y}|\phi_x(x, y)| < 1.$$

The term $U_t(x)$ is known, being a function of the primitives and of the assignment $\alpha$ which we already have in (7). Although $\alpha'(x) > 0$, only the first term on the RHS of (14) is positive; it reflects the marginal contribution of perceived quality in production. The second and third terms reflect the marginal contribution of quality to training. Both terms are negative because they reduce the pressure for high-$x$ firms to pay more today because such firms offer the worker more future income, especially when $\phi_x$ is high. For some parameter values the training effect may dominate, and high-$x$ firms will end up paying lower wages, i.e., we may have $w_t'(x) < 0$ for all $(t, x)$.

Two issues arise with this characterization that we need to note:

(i) *Multiple Equilibria May Exist*.—Proposition 1 makes no mention of a boundary condition for $w$, and unless one is specified we will have many solutions to the differential equation in (13). A natural condition is that when the lowest-$x$ firm produces virtually nothing, and when the only endowment of the young is labor, then the initial condition is that $w_t(0) = 0$ and equilibrium is then unique. This is the case in the log-normal example of Section III, in which the smallest $x$ and $y$ are both zero. More generally, when more than one equilibrium in this class does exist, it still implements the assignment (7).[6]

(ii) *Equilibrium Need Not Exist for All Parameter Values*.—There are two reasons why this is so. First, the contraction property may fail; since $\phi$ is linear homogeneous, the marginal product of $x$ is a function of the ratio $y/x$ and when $y$ has significant dispersion, (15) may fail, and so does the proof of existence. Second, for parameter values for which $U < 0$ for some $x$, a positive-sorting equilibrium exists only if the young *do* have an endowment or (equivalently) if their consumption can be negative. On this point see the

[6]This issue in the assignment game literature arises when the lowest payoff exceeds the outside option (zero). But it exists only if the measure of the two sides of the market are identical. Once there is a long side, the price is pinned down.

discussion of compensating differentials following the solutions for the wage given in (19) and (20).[7]

We shall specialize Proposition 1 in three ways. First, we study the case when the signal on $y$ is perfect. Second, we deal with the polar opposite case when the signal is uninformative. In each of these cases we solve only the BGP. Lastly, in Section III we shall study the full dynamics for a log-normal Cobb-Douglas economy.

## C. *The Case When the Signal is Perfect*

Suppose $s = y$ for all $y$. Let $\bar{\varepsilon} = \int \varepsilon d\Psi(\varepsilon)$, and let

$$\Gamma = \phi(1, b\bar{\varepsilon}). \tag{16}$$

Then when signals are perfect, $\Gamma$ is the long-run growth factor, and the distribution of $x$ is, at each date, a scaled-up version of the talent distribution $\Psi$. More precisely, we have

PROPOSITION 2: *Let* $H_0(x) = \Psi(\frac{x}{c})$, *where* $c > 0$. *Then for all* $t = 1, 2, \ldots,$

$$y = \alpha_t(x) = b\bar{\varepsilon}x, \tag{17}$$

*and*

$$H_t(x) = \Psi\left(\frac{x}{c\Gamma^t}\right), \tag{18}$$

*where* $\Gamma$ *is given in* (16). *The share of wages in output is constant and equal to*

$$\omega = \frac{1}{1 - \beta\phi_x(1, b\bar{\varepsilon})}\left(\frac{f_y(1, b\bar{\varepsilon})b\bar{\varepsilon}}{f(1, b\bar{\varepsilon})} - \beta\phi_x(1, b\bar{\varepsilon})\right) < 1, \tag{19}$$

*so that*

$$w_t(x) = \omega f(1, b\bar{\varepsilon})x, \qquad \text{and} \qquad \pi_t(x) = (1 - \omega)f(1, b\bar{\varepsilon})x. \tag{20}$$

Because we have the solution for $w$ and $\pi$, we can check that the solution to the firm's problem is incentive compatible. Since $w$ and $\pi$ are linear functions and since the assignment is linear, we find in (8) that $V(s)$ is linear and that, as a result, the firm's objective in (12) is strictly concave in $s$ provided that $\rho$ and/or $\theta$ are strictly positive and less than unity.

Note the following:

(i) *Compensating Differentials.*—As long as $\beta > 0$ and $\phi_x > 0$, wages include a compensating differential reflecting the training received. This is because

[7] An equilibrium may still exist in which the allocation is not positively assorted—see Anderson and Smith (2010) on this point.

there is not a separate market for workers and for training, and the wage reflects the net payment for the two combined. Moreover, if $f_y/f$ is small relative to $\phi_x$, $\omega$ can be negative, so that the worker in effect pays directly for the training. If the young have no endowment with which to finance this payment and if their consumption cannot be negative, then in this case equilibrium does not exist.

(ii) *The Experience Premium*.—The cross-sectional old-to-young wage ratio is $(1-\omega)/\omega$. The larger is the compensating differential, the larger is the experience premium. The premium is negative when $\omega > 1/2$. This can arise when the share of the young in production $f_y/f$ is large relative to $\phi_x$. When $x$ contributes little to output and little to training, the bulk of an agent's contribution to discounted production and, hence, lifetime income, occurs in one's youth.

(iii) *TFP and Wages*.—High-$x$ firms will have high labor productivity, but it also is likely that they will have high TFP as conventionally measured. Customarily one computes a Solow residual of (1) as shown in the Cobb-Douglas example in (21). If an old agent's $x$ is measured less accurately than the output that his firm produces, a high-$x$ firm will appear to the statistician to be more efficient relative to the measured quality and quantity of its inputs. Such a firm would then appear to have high TFP. When $\omega > 0$, high-$x$ firms pay higher wages. It is possible, however, for $\omega$ to be negative, in which case $w' < 0$ and high-$x$ firms pay the young workers less, as is the case among top economics departments, for example.

(iv) *Uniqueness*.—When the support of $\Psi$ includes zero, and when the parameters are such that $\omega > 0$, then a unique equilibrium exists—see the discussion under $(i)$ and $(ii)$ above.

**Cobb-Douglas example:** Let

$$f(x, y) = x^{1-\rho} y^{\rho} \quad \text{and} \quad \phi(x, y) = A x^{1-\theta} y^{\theta}. \tag{21}$$

If, in addition, $b\bar{\varepsilon} = 1$, (16) and (19) yield

$$\Gamma = A \quad \text{and} \quad \omega = \frac{\rho - \beta(1-\theta)A}{1 - \beta(1-\theta)A}. \tag{22}$$

The experience premium is negative when $2\rho > 1 + \beta(1-\theta)A$, whereas wages are negative if $\rho < \beta(1-\theta)A$. The Solow residual is $q - (1-\rho)\tilde{x} - \rho\tilde{y}$ where $(\tilde{x}, \tilde{y})$ are the qualities measured by the statistician.

### D. *The Case where the Signal is Uninformative*

As the signal-to-noise ratio in (5) goes to zero, the posterior over $y$ ceases to depend on $s$, i.e.,

$$\tau_t(y \mid s) \;\rightarrow\; G_t(y) \tag{23}$$

for all $s$, and workers become ex ante identical. Since the high-$x$ firms offer more training, they pay lower wages, so that in this case $w_t'(x)$ must be negative for all $x$. Now (11) simplifies to

$$H_{t+1}(x') \;=\; \int_{\{(x,y)\,|\,\phi(x,y)<x'\}} dG_t(y)\; dH_t(x). \tag{24}$$

The representation again simplifies, but now we cannot solve it explicitly except in the special case given in (26)–(27):

PROPOSITION 3: *If* (*23*) *holds*, *then* (*13*) *and* (*14*) *reduce to the single equation*

$$w_t'(x) \;=\; \beta \int \Big( w_{t+1}'(\phi[x,y]) - \int f_x(\phi[x,y], y')\; dG_{t+1}(y') \Big)\, \phi_x(x,y)\; dG_t(y). \tag{25}$$

*Moreover, in the Cobb-Douglas case* (*21*),

$$\lim_{\rho\to 1} \frac{1}{1-\rho}\, w_t'(x) \;=\; -\frac{B_t}{x}, \tag{26}$$

*where*

$$B_t \;=\; \beta(1-\theta)\Big(B_{t+1} + \int y\, dG_{t+1}(y)\Big) \;=\; \sum_{j=1}^{\infty} \beta^j (1-\theta)^j E(q_{t+j}). \tag{27}$$

So, for $\rho$ close enough to unity, in return for working with a firm that has an additional unit of $x$, the worker is willing to give up $(1-\rho)B_t/x$ in wages. In fact, $B_t$ satisfies a standard asset-pricing formula. This is because working for a better firm entitles the worker to a higher stream of rents that derive from the additional training that he will receive. The worker will use that training himself once, and will hand it down to his employee via training.

When assignment is purely random, only a fraction $1-\theta$ of the additional unit of $x$ is transmitted to the next period and, hence, it enters the discount factor. First, the marginal product of $x$ in training is proportional to $1-\theta$. Next, why should $w'$ be unit elastic? Because as $\rho \to 1$, the rate at which tomorrow's marginal product of $x'$, namely $(1-\rho)(y'/x')^{\rho}$, declines with $x'$ tends to unity as $\rho \to 1$. Thus, the wage reduction induced by a unit rise in $x$ is precisely the expected present value of the additional output obtainable from it. To get a nonzero wage elasticity, however, one must divide by $1-\rho$ because, as $\rho \to 1$, training (which the high-$x$ firm provides more of) ceases to contribute to output at all and wages cease to depend on $x$.

## II. Dynamics in the log-normal Case

We now specialize the model so that we can solve the dynamics analytically. When $\phi$ is Cobb-Douglas, it turns out that if $H_t$ is log normal, so is $H_{t+1}$, i.e., if $H_0$ is log normal, it forever remains log normal. Moreover, since a log-normal distribution has just two parameters, the model's dynamics can be described as a sequence in $R^2$ that satisfies an autonomous first-order difference equation. We shall describe the dynamics and focus on the effect of a reform that permanently raises $r$.

Let $f$ and $\phi$ be Cobb-Douglas as in (21) and that $\varepsilon$ and the signal are log-normal variates. Letting hats denote logs, let $\hat{\varepsilon} \sim N(\mu_{\hat{\varepsilon}}, \sigma_{\hat{\varepsilon}}^2)$, and let the log of the signal be

$$\hat{s} = \hat{y} + \eta,$$

where $\eta \sim N(0, \sigma_\eta^2)$. Denote the squared correlation coefficient (proxying the signal-to-noise ratio) by $r^2 = \sigma_{\hat{y}}^2/(\sigma_{\hat{y}}^2 + \sigma_\eta^2)$. Taking logs in (3), $\sigma_{\hat{y}}^2 = \sigma_{\hat{\varepsilon}}^2$, and therefore the variance of the log of $s$ is a constant: $\sigma_{\hat{s}}^2 = \sigma_{\hat{\varepsilon}}^2 + \sigma_\eta^2$, so that

$$r^2 = \frac{\sigma_{\hat{\varepsilon}}^2}{\sigma_{\hat{\varepsilon}}^2 + \sigma_\eta^2}. \tag{28}$$

Now if $H_0$ is log normal, the assignment will be log linear and at each date $H_t$ will be log normal, so that $\hat{x} \sim N(\mu_t, \sigma_t^2)$.[8] Let us take logs of the left-hand side of the assignment in (7) and obtain the deterministic relation

$$\hat{s} = \ln \alpha_t(x) = \mu_{\hat{s},t} + \frac{\sigma_{\hat{s}}}{\sigma_t}(\hat{x} - \mu_t). \tag{29}$$

The dynamics are summarized as follows:

PROPOSITION 4: *The law of motion for H in* (11) *is equivalent to the following difference equation for the pair* $(\mu, \sigma)$*:*

$$\mu_{t+1} = \mu_t + \hat{A} + \theta\left(\hat{b} + \mu_{\hat{\varepsilon}} + \frac{1}{2}\sigma_t^2\right), \tag{30}$$

$$\sigma_{t+1} = \sqrt{(1-\theta)^2\sigma_t^2 + 2\theta(1-\theta)r\sigma_{\hat{\varepsilon}}\sigma_t + \theta^2\sigma_{\hat{\varepsilon}}^2}. \tag{31}$$

*For* $\theta \in (0, 1]$, *the BGP is globally stable;* $\sigma_t$ *converges to*

$$\sigma(r) = \frac{\sigma_{\hat{\varepsilon}}}{2-\theta}\left(r(1-\theta) + \sqrt{1-(1-r^2)(1-\theta)^2}\right), \tag{32}$$

*and the growth factor* $\Gamma_t \equiv \int x\, dH_{t+1}(x)/\int x\, dH_t(x)$ *converges to*

$$\Gamma(r) = Ab^\theta \exp\left(\theta\mu_{\hat{\varepsilon}} + \frac{\theta}{2}[\sigma(r)]^2\right). \tag{33}$$

[8] We do not use the $\hat{x}$ subscript on $\mu_t$ and $\sigma_t^2$ so as to simplify notation.

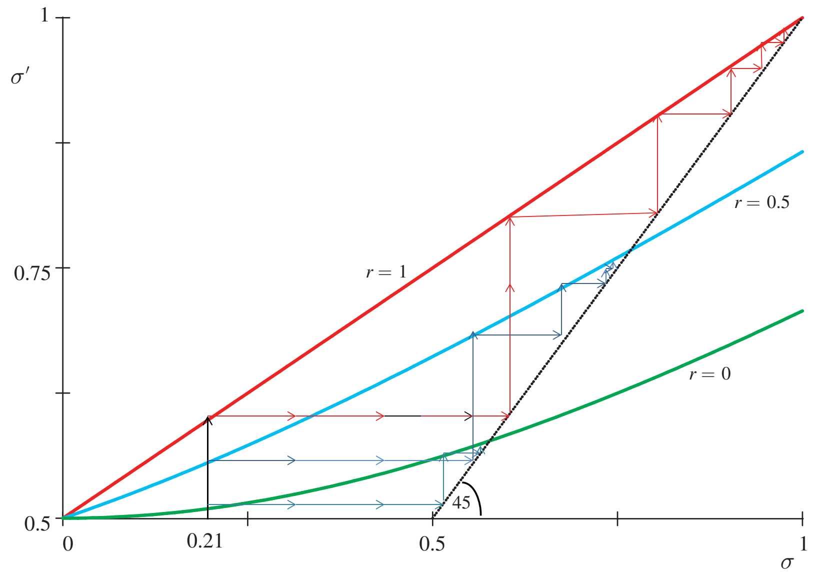


Figure 1. Transition Dynamics for $\sigma$ Based on (31) for $\theta = 1/2$, $\sigma_{\hat{\varepsilon}} = 1$, and $r \in \left\{0, \frac{1}{2}, 1\right\}$

We make two remarks about this proposition:

(i) The right-hand side of (31) is homogeneous of degree 1 in $(\sigma_t, \sigma_{\hat{\varepsilon}})$. Letting $u_t \equiv \sigma_t / \sigma_{\hat{\varepsilon}}$, (31) reads $u_{t+1} = \sqrt{(1-\theta)^2 u_t^2 + 2\theta(1-\theta) r u_t + \theta^2}$, so that from any starting value $\sigma_0$, the solution for $\sigma_t$ is proportional $\sigma_{\hat{\varepsilon}}$.

(ii) The proposition does not apply to the case $\theta = 0$. In that case $x' = Ax\sigma_t = \sigma_0$ for all $t$. That is, $\sigma_t$ remains at its initial value.

## A. *Mode of Convergence of* $\sigma$

If $\sigma_{\hat{\varepsilon}}^2 = 0$ and if we start with homogeneous old agents, the model is of an "Ak" type, with no transitional dynamics, and with no intragenerational inequality. Otherwise, the model has transitional dynamics the nature of which depends on whether the initial inequality, $\sigma_t$, among the old is above or below its steady-state value. The mode of convergence of $\sigma$ is the focus of Figure 1, which plots the right-hand side of (31) for $r \in \left\{0, \frac{1}{2}, 1\right\}$, along with the 45° line.

Since only the ratio $\sigma_t / \sigma_{\hat{\varepsilon}}$ is determined (see remark (i) above) we set $\sigma_{\hat{\varepsilon}} = 1$. The figure starts three economies off at the same initial level of inequality, namely $\sigma_0 = 0.21$. The intercept of all three curves is at $\theta$, which is set at 0.5.

Evidently, convergence of $(\sigma, \Gamma)$ to their BGP values given in (32)–(33) is slower for higher values of $r$. For fixed $r$, (30) shows that $\mu_{t+1} - \mu_t$ is increasing in $\sigma_t^2$ and, therefore, converges to its limit from the same direction and at the same rate as $\sigma_t^2$ converges to $[\sigma(r)]^2$. Thus, $(\sigma_t, \Gamma_t)$ converge monotonically either both from above or both from below. Monotonicity follows because the right-hand side of (31) is strictly increasing in $\sigma_t$. The $r = 1$ economy sees the fastest rise in inequality right away, the $r = 0$ economy the slowest. But $r = 0$ economy converges fastest to its BGP level, and the $r = 1$ economy the slowest. The last statement applies to convergence from above as well, say from a value $\sigma_0 > 1$, although this is not pictured in the figure. Concretely, let $r$ rise from $r = 0$ to $r = 0.5$, and let us trace the effect in Figure 1. The natural starting point is the steady-state value of $\sigma_t$ that we would read off the intersection of the $r = 0$ schedule with the 45° line. The new steady state is where the $r = 0.5$ line now crosses the 45° line. This would be a major reform, and it would take a couple of generations (i.e., a couple of steps in Figure 1) to get the economy about 90 percent of the way to its new steady-state level of inequality. Therefore, the experiment outlined in the previous subsection seems to imply that the response to the policy change from $r = 0$ to $r = 0.5$ should be about 90 percent complete in two generations. But if a unit of time is a generation's half life, convergence to the BGP is slow. In any case, Figure 1 shows that following a change in $r$, $\sigma$ converges monotonically to its long-run value, and therefore so does $\Delta\mu_t \equiv \mu_{t+1} - \mu_t$.

### B. *Dynamics after a Permanent Rise in r*

We now study an economy that starts off in a zero-growth steady state and that then experiences an unexpected "reform" that raises $r$ permanently. As we saw, (30) and (31) imply that $\Delta\mu_t$ and $\sigma_t$ both rise monotonically to their new BGP levels. Thus, as $\mu$ and $\sigma$ respond to the reform, we should see a positive comovement of $\mu_t$ and $\Delta\mu_t$ on the one hand, and $\sigma_t$ on the other. Let us now make this relationship precise. As a baseline value we take $r = 0.25$; the other parameters—chosen so that the economy has zero long-run growth—are listed in Table 1.

With these parameter values, equation (32) gives $\sigma = 0.38$. Thus the starting point is the pair $(\sigma, \Delta\mu) = (0.38, 0)$, and the initial fixed, $r$ then rises from 0.25 to 0.5, 0.75, and 1.0, respectively.[9]

Each curve in Figure 2 plots the response of $(\sigma, \Delta\mu)$ to a particular reform size, and the vector $(\sigma, \Delta\mu)$ moves North-East toward its new steady-state value which is $(0.012, 0.44)$, $(0.027, 0.51)$, and $(0.046, 0.57)$, respectively and which is marked off by dashed lines. Along each path we show a handful of dates; convergence to the steady state is mostly over after four generations. The paths are linear in light of (30) but the "step sizes" diminish rapidly. Since (31) contains $r$, when $r$ rises at date zero, $\sigma_1$ rises. But nothing happens to $\Delta\mu$ until the next period because (30), which reads $\Delta\mu_t = \hat{A} + \theta(\hat{b} + \mu_{\hat{\varepsilon}} + \frac{1}{2}\sigma_t^2)$, does not include $r$. Aggregate output and human capital do rise right away because their levels do depend on $r$. This will be clear when we discuss the impulse response of output growth next.

### C. *The Response of Aggregate Output*

While $\Delta\mu_t$ and $\sigma_t$ both rise monotonically to their new BGP levels, the growth of aggregate output responds nonmonotonically when the reform is large, though the response is monotonic for smaller reforms. The nonmonotonicity arises because (as

[9] To achieve this rise, $\sigma_\eta^2 = \sigma_{\hat{\varepsilon}}^2(1 - r^2)/r^2$ is lowered to 0.99, 0.26, and to zero, respectively.

Table 1—Baseline Parameter Values Used in the Computation

| $\theta$ | $\rho$ | $\beta$ | $b$ | $\mu_{\hat{\varepsilon}}$ | $\sigma^2_{\hat{\varepsilon}}$ | $\sigma^2_\eta$ | $A$ |
|---|---|---|---|---|---|---|---|
| 0.5 | 0.66 | 0.36 | 0.85 | 0 | 0.33 | 4.95 | 1.045 |

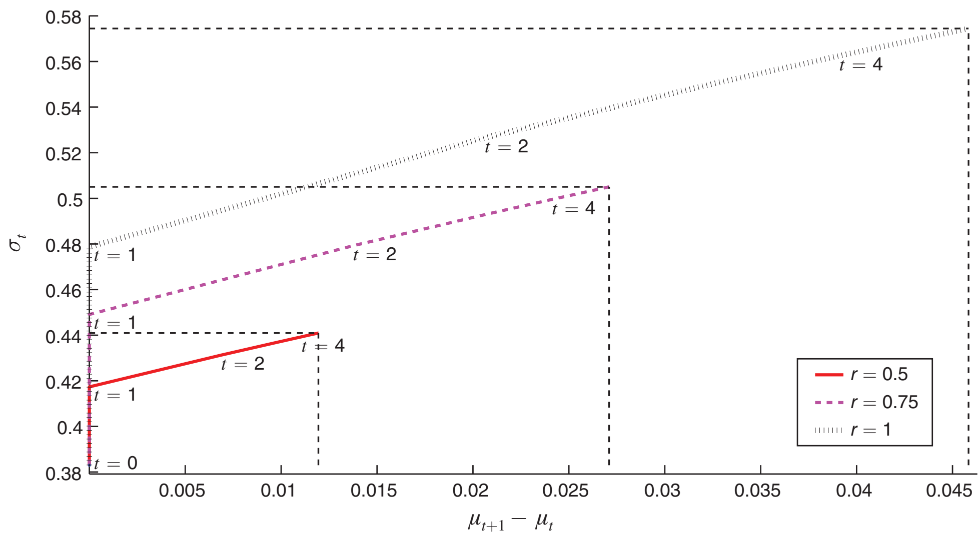


Figure 2. Transition Dynamics for $(\sigma, \Delta\mu)$ Following Three Reforms

the online Appendix shows) the level of aggregate output, $Y$, is an increasing function not only of $\mu$ but also of $\sigma$:

$$Y_t = C \exp\left\{D\mu + \frac{D^2}{2}\sigma_t^2\right\}, \tag{34}$$

where

$$C \equiv \exp\left\{\rho(1-r^2)\mu_{\hat{y}} + \frac{\rho^2}{2} r^2 \sigma_\eta^2\right\} \quad \text{and} \quad D \equiv 1 - \rho + \rho r^2 \sqrt{\frac{\sigma^2_{\hat{\varepsilon}} + \sigma^2_\eta}{\sigma_t^2}}. \tag{35}$$

As the growth of $\sigma$ slows down, this may dominate the rising growth in $\mu$ and lead to an eventual decline in output growth.

Figure 3 illustrates the response of output growth to the same three reforms that we analyzed in Figure 2. Taking again the baseline $r$ of 0.25, and the other parameters as listed in Table 1, this gives us a baseline path for $\Delta \ln Y_t$ along the horizontal line—i.e., the case of zero growth forever. Then each reform for $r$ results in a different impulse response. For a small rise in $r$ from 0.25 to $r = 0.5$, output growth responds monotonically: $\Delta \ln Y_t$ keeps on rising on its way to its new long-run value of 0.09. For larger rises in $r$, however, $\Delta \ln Y_t$ first overshoots its long-run value and then converges to it monotonically from above. These are the two top curves in Figure 3. Thus, although the growth of $\mu$ accelerates along the transition to a higher rest point, the growth of $\sigma$ decelerates, and this sometimes mitigates (but does not overturn) the initial positive response of the growth of $Y$ to the rise in $r$.

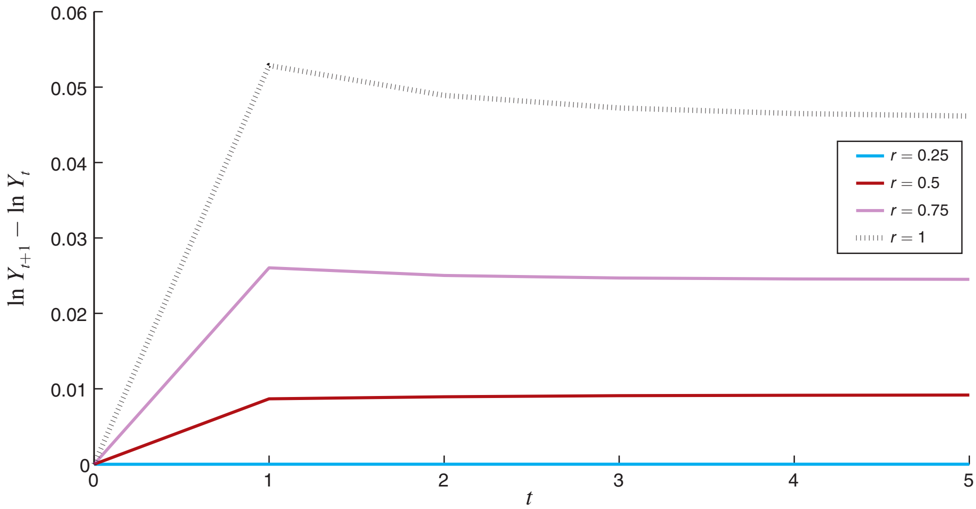


Figure 3. Response of Output Growth to Three Reforms

## III. Long-Run Effects of Garbling in the log-normal Case

We shall now discuss other long-run relationships with a focus on the effects of changes in $r$. The term "misallocation" refers to departures from the ideal assignment in (7) that is attainable under full information about $x$ and $y$. Since the ideal assignment is unknown when $r < 1$, agents generally fail to achieve it.

Let us refer to this departure from the ideal assignment as a "garbling." The smaller is $r$, the more garbling exists in the sense that $\alpha_t(x)$ (as given in (7)) is further from the ideal assignment. The following four subsections show that garbling (i) lowers long-run growth, (ii) lowers long-run inequality, (iii) lowers the returns to scale (suitably defined), and (iv) creates more turnover in the distribution of $x$. We saw (i) and (ii) in the simulations, and now we revert to analysis.

### A. *Garbling and Growth*

A garbling of assignments reduces long-run growth as long as the young are heterogeneous and as long as $\theta$ is neither zero nor one. The effect is easiest seen if one compares the worst and the best scenarios. The largest possible garbling reduction is one that raises $r$ from zero to one. Substituting for $\sigma(r)$ from (32) into (33) and evaluating at the two extreme values of $r$ yields the resulting growth-improvement ratio:

$$\frac{\Gamma(1)}{\Gamma(0)} = \exp\left(\frac{\theta(1-\theta)}{2-\theta}\,\sigma_{\hat{\varepsilon}}^2\right). \tag{36}$$

The effect of garbling on this ratio thus depends on $\theta$ and $\sigma_{\hat{\varepsilon}}^2$ alone.

*The Role of* $\sigma_{\hat{\varepsilon}}$.—The effect of $r$ on $\Gamma$ is proportional to $\sigma_{\hat{\varepsilon}}^2$. If the young are homogenous, $r$ does not matter.

*The Role of* $\theta$.—The effect of $r$ on $\Gamma$ is largest when $\theta$ is 0.59. The effect of $\theta$ itself has an inverted-U shape, with zero at the corners. When $\theta = 1$, only $y$ matters for growth, and $\ln y$ always has variance $\sigma_{\hat{\varepsilon}}^2$. In that case, $\Gamma = Ab \exp\left(\mu_{\hat{\varepsilon}} + \frac{1}{2}\sigma_{\hat{\varepsilon}}^2\right)$, and there are no transitional dynamics for $\Gamma$. The dynamics for $\sigma_t$ last just one period, i.e., $\sigma_t = \hat{\sigma}_\varepsilon$ is attained in a single generation and does not change thereafter. If $\theta = 0$, however, $x' = Ax$ and the equilibrium is qualitatively different: Proposition 4 is not valid for this case because (32) does not hold. Rather, (31) reads $\sigma_{t+1} = \sigma_t$, which means that $\sigma_t$ remains fixed at its initial value, $\sigma_0$, an extreme form of hysteresis in $\sigma_t$, whereas growth is constant at $\Gamma(r) = A$.

*A Rough Estimate of* $\Gamma(1)/\Gamma(0)$.—When $r = 1$ wages are proportional to $\varepsilon$ since the latter is then proportional to $x$—see (20). Since the computation revealed wages to be concave in $x$, and their maximal variation is when $r = 1$, a lower bound for $\sigma_{\hat{\varepsilon}}^2$ is the variance of log wages. Heathcote, Storesletten, and Violante (2005, Figure 1A) provide us with the estimate $\sigma_{\hat{\varepsilon}}^2 = 0.33$, an average of the variance of log wages and the variance of log earnings for a 33-year-old worker. Parameter estimates for the role of teacher and student ability in training functions are hard to find, so one may look at the achievement of children as a function of own and parents' skills. From Cunha, Heckman, and Schennach (2010, Table 1) we take $\theta$ as the share (in the production function for a child's cognitive skill formation) of the child's cognitive skills relative to sum of the shares of the child's and the parent's cognitive skills, then $\theta$ is 0.91. Plugging these parameter values into (36), the predicted effect of a unit change in $r$ on the growth rate over a half-life of a generation is 0.025, or 2.5 percent. That is, $\Gamma(1)/\Gamma(0) = 1.025$. If we take that half-life to be 22.5 years, the resulting effect is about 11 basis points per year. A small effect, but one that compounds for as long as the policy differential lasts. For instance, if we date the rise in $r$ to the first universities established in Italy, France, Spain, and England some 900 years ago, we would have an improvement factor of $\left(\Gamma(1)/\Gamma(0)\right)^{\frac{900}{22.5}} = 2.72$. The maximal effect of $r$ obtains if $\theta = 0.59$, in which case (continuing to assume that $\sigma_{\hat{\varepsilon}}^2 = 0.33$) the ratio $\Gamma(1)/\Gamma(0)$ would equal 1.058, and the 900-year improvement factor would be 9.54.

## B. *Garbling and Human Capital Inequality*

A garbling reduces lifetime inequality. It also reduces inequality in $x$ and, hence, inequality among the old. It also reduces the response of wages to $x$ and thereby reduces inequality among the young.

*Lifetime Inequality*.—Garbling reduces lifetime inequality and, in the limit as $r \to 0$, lifetime inequality also goes to zero because $V_t(s)$ in (8) ceases to depend on $s$ and becomes constant $V_t$; workers are the ex ante indistinguishable and must in equilibrium earn the same lifetime income. Workers opting for high-$x$ firms will earn less in youth and more in old age, but the present values are equal.

*Inequality Among the Old*.—Garbling reduces long-run inequality of $x$ among the old, but not to zero. The reduction is larger when $\theta$ is low. Using (32) we find that

$$\frac{\sigma(1)}{\sigma(0)} = \sqrt{\frac{2-\theta}{\theta}}. \tag{37}$$

Using once again Cunha, Heckman, and Schennach's (2010) estimate of $\theta = 0.91$, this ratio is just 1.09.

*Wage Inequality*.—Cases 1 and 2 both entail wage inequality, but when signals are perfect, the high-$x$ firms pay more, whereas when $r = 0$, the low-$x$ firms pay more. So presumably at intermediate levels of $r$, wages are relatively flat in $x$, and wage inequality is at a minimum. Based on this, we conclude that wage inequality is U-shaped as a function of $r$.

When $r < 1$, we cannot solve for $w_t(x)$ analytically. Simulations reveal that for lower values of $r$, $w$ becomes increasingly concave in $x$, but at some point this tendency reverses and then straightens and then become convex as $r \to 0$. Figure 4 plots $w(x)$ for various values of $r$. The other parameters as given in Table 1, but the parameter $A$ was adjusted to keep the economy's growth rate at zero. That is as $r$ varied, we set $A = b^{-\theta}\exp\left(-\frac{\theta}{2}[\sigma(r)]^2\right)$, where $\sigma(r)$ is the stationary value of $\sigma_t$ given in (32). In other words, as $r$ is lowered from unity towards zero, we compensate by raising $A$ so that $\Gamma(r)$ remains at unity. As $r \to 1$, we get the frictionless Cobb-Douglas case of (21) with the solution in (22). Given the parameter values in Table 1, $A \to 1$, and $w'(x) \to \omega = 0.59$. This is the constant slope of the top line in Figure 4. Anticipating the discussion in Section IVC, $w(x)$ has a curvature that rises as $r^2$ falls. The figure also shows that the sensitivity to $r$ is much larger when $r$ is close to zero.

The lowest curve is for an economy in which the signal-noise ratio is very close to zero. Confirming the result in Proposition 3, the slope is negative. The details of this computation are at the end of the online Appendix.

## C. *Garbling and the Returns to Scale*

The firm is always composed of two workers, and so by "returns to scale" we mean the response of $f$ or $\phi$ to a proportional increase in $(x, s)$ in the cross-section of firms at a point in time. The online Appendix shows that conditional on $s$, the log of $y$ is the variate

$$\hat{y} = (1 - r^2)\mu_{\hat{y}} + r^2\hat{s} + \sqrt{1-r^2}\,\sigma_{\hat{\varepsilon}}\zeta,$$

where $\zeta \sim N(0, 1)$. Thus in the Cobb-Douglas example (21), we find that the expected output of the pair $(x, s)$ is

$$E(x^{1-\rho}y^{\rho} \,|\, x, s) = B_1 x^{1-\rho} s^{\rho r^2},$$

where $B_1 = \exp(\rho(1-r^2)\mu)$ is a constant. Therefore, even though both $f$ is homogeneous of degree 1 in $(x, y)$, it is homogeneous of degree $1 - \rho(1 - r^2)$.

Now consider the reduced form after substituting out the assignment, and consider output as a function of $x$ alone. We now have returns that diminish. The Appendix

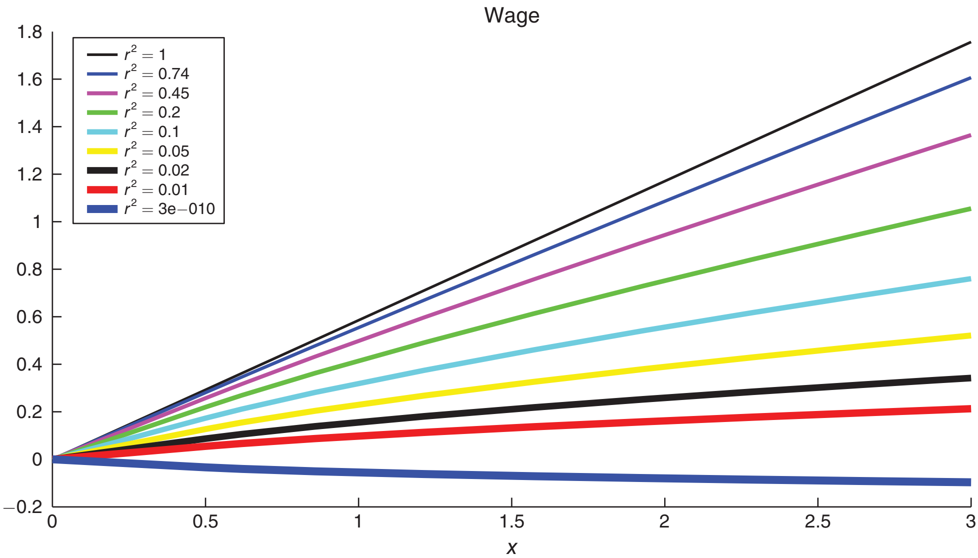


Figure 4. $w(x)$ for Various Values of $r^2$

shows that assignment in (29) leads to the following expression for the quality of the worker assigned to a firm of quality $x$:

$$\hat{y} = \mu_{\hat{y}} + r\frac{\sigma_{\hat{\varepsilon}}}{\sigma_t}(\hat{x} - \mu_t) + \sqrt{1 - r^2}\,\sigma_{\hat{\varepsilon}}\zeta. \tag{38}$$

Substituting this into $f$ and conditioning on $x$ alone,

$$E(x^{1-\rho}y^{\rho}|x) = B_2 x^{1-\rho\left(1 - r\frac{\sigma_{\hat{\varepsilon}}}{\sigma_t}\right)},$$

where $B_2$ is a constant.

Interestingly, returns to scale can exceed unity on the transition path to the BGP. If $r\sigma_{\hat{\varepsilon}}^2/\sigma_t > 0$, i.e., the quality of the young is highly dispersed relative to that of the old, in (29) we see that $d\hat{s}/d\hat{x}$ becomes large so output rises faster than $x$. This can occur only on a transition path from a $\sigma_0$ below its long-run level of $\sigma(r)$ which is given in (32), and which is at the intersection with the 45° line in Figure 1. At the steady state, however, the ratio $r\sigma_{\hat{\varepsilon}}^2/\sigma_t$ is never above unity and rises from zero to one as $r$ ranges from zero to one.

Similar remarks apply to the returns to $x$ in $\phi$, and to the coefficient of $\hat{x}_t$ in equation (39) below. Since $\phi$ governs the dynamics of $x$, that leads us into the next topic.

### D. *Garbling and Turnover in the TFP Distribution*

When $r < 1$, each firm's $x$ is stationary around a trend. Small firms will grow faster than larger firms, consistent with the data. When frictions are absent, however, as in Proposition 2, and $r = 1$, Gibrat's Law holds for output and TFP. When $r < 1$, unproductive, low-$x$ firms can end up with productive young people, and this will tend to raise their growth. The opposite happens to productive, high-$x$ firms. Thus, garbling induces mean reversion; it introduces turnover in the distribution of TFP across firms, i.e., a reversion to trend.

Let us see exactly how this works. The online Appendix derives the growth rate of firm $x$ to be

$$(39)\quad \hat{x}_{t+1} - \hat{x}_t = \underbrace{\hat{A} + \theta\left(\mu_{\hat{y}} - r\mu_t \frac{\sigma_{\hat{\varepsilon}}}{\sigma_t}\right)}_{\text{independent of } \hat{x}_t, \text{ and constant on BGP}} \underbrace{- \theta\left(1 - r\frac{\sigma_{\hat{\varepsilon}}}{\sigma_t}\right)}_{>0 \text{ on BGP}} \hat{x}_t + \underbrace{\theta\sqrt{1-r^2}\,\sigma_{\hat{\varepsilon}}\,\zeta}_{\text{zero-mean}}.$$

The larger are $\theta$ and $\sigma_{\hat{\varepsilon}}$ and the smaller is $r$, the larger the influence that the disturbance, $\zeta$, will have on firm growth, and the larger will be the turnover in the productivity distribution. Ábrahám and White (2006) find that the annual autocorrelation of the plant-specific productivity shock is only 0.4. On the other hand, output autocorrelations are much higher but still less than unity.

## IV. Discussion

This model suggests an explanation for secular and cross-sectional differences in growth caused by changes in the signal to noise ratio. This section discusses other questions that the model can address, extensions, and caveats.

### A. *Relation to Work on Misallocation*

Closest work to mine are the papers of Anderson and Smith (2010) and Anderson (2011) which analyze a model that has a version of (1) and (2) with infinitely lived agents. They analyze a partnership model as I do, and they allow for randomness in (2). An agent's history then acts as a public signal on the agent's ability, much like the signal $s$ in my model. They focus on life cycle issues; they also characterize matching patterns that are not perfectly assortative, and the resulting type, wage, and inequality dynamics in a world with stationary aggregates. Since I focus on growth and transitions, our papers are complements.

Related work focuses on search frictions (Burdett and Coles 1997; Shimer and Smith 2000; Lagos 2006; Gautier and Teulings 2011), while other work highlights implicit taxes (Restuccia and Rogerson 2008; Hsieh and Klenow 2009) or misallocation of capital due to finance (Buera and Shin 2009; Midrigan and Xu 2014; Moll forthcoming). Papers on growth highlight the misallocation of intermediate goods (Jones 2011), and papers that look at growth impacts and transition impacts (e.g., Buera and Shin) focus on capital market distortions. Most papers consider the effect of frictions on the level of output, and when they do deal with growth, they do not focus on the labor market and on human capital.

A vast literature also exists on market failures caused by private information. Earlier drafts of the article included a section in which either the young or the old have (perfect) private information about quality. As long as the private information is one sided, there is a separating equilibrium that supports the same positive assignment as in the full-information equilibrium assignment in (17). In other words, the friction does not change the allocations—matching remains positively assortative; the only difference is in the wages that support that outcome. For example, when

$y$ is privately known to the young, separation is achieved by a wage that decreases in $x$ by enough to dissuade the low-$y$ workers from taking the high-$x$ firms' offers. High-$y$ workers accept the lower wages because they are more than compensated for it by the additional training that they will receive. Conversely, when $x$ is privately known to the old, a separating equilibrium is still achieved, but this time with wages that increase in $x$. A high-$x$ firm cares more about the quality of the worker it will get and, hence, will pay more. Having found these separating equilibria in which the assignment was unchanged, I moved to a different friction, namely, the imperfect public signal analyzed here. Based on this discussion I would conjecture, however, that if $s$ were private information to the worker, a separating equilibrium would exist that would replicate the assignment in (7), at least for some parameter values. If that is so, the friction we have analyzed here can be interpreted as a private-information friction.

### B. *Role of Policy*

Since the equilibrium assignment maximizes the rate of growth of aggregate output subject to the learning friction, equilibrium is constrained Pareto efficient, and so taxes and transfers cannot improve the allocation. Rather, policy should strive to reduce the friction, at least if the benefits of using public funds for the purpose exceed the costs.

The signal $s$ is best thought of as a record of a person's schooling or as his score on a test he may have taken at the start of his career. A young agent knows how well he did at school but is typically not so sure how he will function in the workplace once he leaves the classrooms.[10] On this view, $r$ depends on the quality of schooling and on the accuracy of the test, and a good educational system should raise $r$. Investment in schooling varies among societies both in its scale and in its mix: Europe relies on public inputs, the United States more on private inputs.

If it relates to the provision of screening, policy should recognize that firms also screen internally and solve internal assignment problems as in Prescott and Visscher (1980). Instead of $F(s \mid y)$, the signal would be written as $F(s \mid y, z)$, where $z$ is a costly decision, and the policy question is what part of $z$ should be provided publicly, and what part privately. A discussion of education policy that takes an information-theoretic view is in Davies and MacDonald (1984).

Perhaps we can consider $r$ as an index of meritocracy as suggested in the introduction, which brings up the examples of Europe, Japan, and the United States. In a related vein, Caselli and Gennaioli (2013) argue that in some societies if the owner of a firm cannot hire the appropriate manager, he may bequeath the firm to his offspring who may be of inferior ability. This lowers $r$ in the sense of the present article.

Also, $r$ may proxy for physical barriers. Good roads, public transportation, and even a functioning rental market all should help agents to overcome geographical barriers and to better approximate the efficient assignment. Kremer (1993) argues

[10] Examples of forecast errors arise in the process of drafting of professional athletes in the National Football League, National Basketball Association, or Major League Baseball. Closer to home, the job market for new PhDs involves forecast errors: a star student may fail to get tenure, and someone who is denied tenure may end up with a Nobel prize.

that a market with many participants, a thick market in a city or in a densely populated economy, should have a high $r$.

### C. *Caveats*

There are two that we ought to note before closing: the complementarity assumption, and the assumption that a match must last an entire half-working life.

*Complementarity.*—Imperfect sorting on quality need not reflect a garbling friction. Instead, the assumption that $x$ and $y$ are complements may not hold always and everywhere. Evidence shows that sorting is positive, but not perfectly so. Lentz and Mortensen (2010), Eeckhout and Kircher (2011), and others discuss the problems in identifying the pattern of sorting between workers and firms, mainly in the context of static models of production. For training, evidence for complementarity between student and teacher quality is in Lockwood and McCaffrey (2009), and complementarity between child and parent quality is in Cunha, Heckman, and Schennach (2010). But when we fail to find highly positive sorting, other forces may be at work; some models imply a hierarchical internal structure in equilibrium so that workers of *different* abilities end up in the same firm, e.g., Garicano and Rossi-Hansberg (2004).

*No Recontracting.*—Even if the complementarity assumptions do hold, the response of the allocations to changes in $r$ may be smaller than described in (36) and (37). An $(x, s)$ match is treated as being permanent—the worker and his manager are stuck with each other for the remainder of the manager's life, and an actionable revelation of ability occurs only when the agent enters old age. If bad matches could be undone, the effects of $r$ on growth would decline—by how much would depend on the speed of learning the worker's general ability.

Lange (2007) estimates that the half-life of learning the worker's ability in an employment relationship is three years. If this information is about *general* ability, and if it becomes available to the market at large, then after some six years, an assignment approximating the $r = 1$ outcome would be attainable. The smaller is $r$, the larger is the average initial mismatch, and the higher would be the ex post desire to recontract. So, with gradual accumulation of signals on the job, the effects of $r$ would be smaller than described in (36) and (37). Now, half a working life is roughly between 20 and 25 years. On the other hand, match quality may depend on more than simply the interaction of the parties' general abilities; the presence of such effects would slow down the speed at which a worker's general ability could be ascertained on the basis of the output that a single match generates. In that case, three years is an underestimate of the speed of learning $y$. McCall (1990) studies a related issue in which an agent has a firm-specific as well as an industry-specific ability, neither of which he knows, and he can sort these out only by sampling jobs in more than one industry.

A way to bring this model closer in spirit to models of undirected search would be to endogenize the precision of the signal. Noise reduction could result from additional years in school, and this introduces a trade-off similar to that present in the search decision: a more precise signal improves the match, but obtaining one means spending less time producing and less time being trained.

## V. Conclusion

This article has studied the role of a learning friction in the labor market when the friction reduces output and training. Because training is a process that exhibits constant returns, there is long-run growth of human capital and output. For a class of economies, the article solves for the full dynamics, including the dynamics of the cross-section distribution of human capital.

The term "misallocation" refers to a departure from the ideal assignment in (7) that would be attainable under full information about $x$ and $y$. Unless a planner has such information, the market's failure to achieve the ideal assignment does *not* reflect an inefficiency. Equilibrium output and growth are both at a maximum conditional on the signal quality. Policy should, if anything, try to reduce the learning friction. If policy can raise signal quality cheaply enough, then it will be worth pursuing, although the model says that resulting rise in growth will be accompanied by a rise in inequality.

## REFERENCES


**Ábrahám, Árpád, and Kirk White.** 2006. "The Dynamics of Plant-Level Productivity in U.S. Manufacturing." Center for Economic Studies, US Bureau of the Census CES 06-20. ftp://ftp2.census.gov/ces/wp/2006/CES-WP-06-20.pdf (accessed January 19, 2014).

**Anderson, Axel, and Lones Smith.** 2010. "Dynamic Matching and Evolving Reputations." *Review of Economic Studies* 77 (1): 3–29.

**Anderson, Axel.** 2011. "A Dynamic Generalization of Becker's Assortative Matching Result." http://www9.georgetown.edu/faculty/aza/Dynamic_Sorting_August_2011.pdf (accessed January 19, 2014).

**Becker, Gary S.** 1973. "A Theory of Marriage: Part I." *Journal of Political Economy* 81 (4): 813–46.

**Becker, Gary S., and Nigel Tomes.** 1979. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy* 87 (6): 1153–89.

**Buera, Francisco J., and Yongseok Shin.** 2009. "Productivity Growth and Capital Flows: The Dynamics of Reforms." National Bureau of Economic Research Working Paper 15268.

**Burdett, Ken, and Melvyn G. Coles.** 1997. "Marriage and Class." *Quarterly Journal of Economics* 112 (1): 141–68.

**Caselli, Francesco, and Nicola Gennaioli.** 2013. "Dynastic Management." *Economic Inquiry* 51 (1): 971–96.

**Chari, V. V., and Hugo Hopenhayn.** 1991. "Vintage Human Capital, Growth, and the Diffusion of New Technology." *Journal of Political Economy* 99 (6): 1142–65.

**Cunha, Flavio, James J. Heckman, and Susanne M. Schennach.** 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78 (3): 883–931.

**Davies, James, and Glenn MacDonald.** 1984. *Information in the Labour Market: Implications for Education in Ontario*. Toronto: University of Toronto Press.

**Eeckhout, Jan, and Philipp Kircher.** 2011. "Identifying Sorting--In Theory." *Review of Economic Studies* 78 (3): 872–906.

**Garicano, Luis, and Esteban Rossi-Hansberg.** 2004. "Inequality and the Organization of Knowledge." *American Economic Review* 94 (2): 197–202.

**Gautier, Pieter A., and Coen N. Teulings.** 2011. "Sorting and the Output Loss due to Search Frictions." Institute for the Study of Labor (IZA) Discussion Paper 5477.

**Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante.** 2005. "Two Views of Inequality over the Life Cycle." *Journal of the European Economic Association* 3 (2–3): 765–75.

**Hsieh, Chang-Tai, and Peter J. Klenow.** 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics* 124 (4): 1403–48.

**Jones, Charles I.** 2011. "Intermediate Goods and Weak Links in the Theory of Economic Development." *American Economic Journal: Macroeconomics* 3 (2): 1–28.

**Kremer, Michael.** 1993. "The O-Ring Theory of Economic Development." *Quarterly Journal of Economics* 108 (3): 551–75.

**Lagos, Ricardo.** 2006. "A Model of TFP." *Review of Economic Studies* 73 (4): 983–1007.

**Lange, Fabian.** 2007. "The Speed of Employer Learning." *Journal of Labor Economics* 25 (1): 1–35.

**Lentz, Rasmus, and Dale T. Mortensen.** 2010. "Labor Market Models of Worker and Firm Heterogeneity." *Annual Review of Economics* 2 (1): 577–602.

**Lockwood, J. R., and Daniel F. McCaffrey.** 2009. "Exploring Student-Teacher Interactions in Longitudinal Achievement Data." *Education Finance and Policy* 4 (4): 439–67.

**Luttmer, Erzo G. J.** 2007. "Selection, Growth, and the Size Distribution of Firms." *Quarterly Journal of Economics* 122 (3): 1103–44.

**MacDonald, Glenn M.** 1982. "Information in Production." *Econometrica* 50 (5): 1143–62.

**McCall, Brian P.** 1990. "Occupational Matching: A Test of Sorts." *Journal of Political Economy* 98 (1): 45–69.

**Midrigan, Virgiliu, and Daniel Yi Xu.** 2014. "Finance and Misallocation: Evidence from Plant-level Data." *American Economic Review* 104 (2): 455–58.

**Moll, Benjamin.** Forthcoming. "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?" *American Economic Review*.

**Prescott, Edward C., and John H. Boyd.** 1987. "Dynamic Coalitions: Engines of Growth." *American Economic Review* 77 (2): 63–67.

**Prescott, Edward C., and Michael Visscher.** 1980. "Organization Capital." *Journal of Political Economy* 88 (3): 446–61.

**Restuccia, Diego, and Richard Rogerson.** 2008. "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments." *Review of Economic Dynamics* 11 (4): 707–20.

**Shimer, Robert, and Lones Smith.** 2000. "Assortative Matching and Search." *Econometrica* 68 (2): 343–69.

**This article has been cited by:**


1. Qiangmin Xi, Lin Mei. 2022. How did development zones affect China's land transfers? The scale, marketization, and resource allocation effect. *Land Use Policy* **119**, 106181. [Crossref]
2. Xinze Li, Luojia Wang, Kerui Du. 2022. How do environmental regulations influence resource misallocation in China? The role of investment flows. *Business Strategy and the Environment* **124**. . [Crossref]
3. Laiqun Jin, Xiuyan Liu, Sam Hak Kan Tang. 2022. High-technology zones, misallocation of resources among cities and aggregate productivity: evidence from China. *Applied Economics* **54**:24, 2778-2794. [Crossref]
4. Yian Chen. 2022. Misallocation of Talent and Innovation: evidence from China. *Applied Economics* **54**:14, 1598-1624. [Crossref]
5. Xu Dong, Yali Yang, Qinqin Zhuang, Weili Xie, Xiaomeng Zhao. 2022. Does Environmental Regulation Help Mitigate Factor Misallocation?—Theoretical Simulations Based on a Dynamic General Equilibrium Model and the Perspective of TFP. *International Journal of Environmental Research and Public Health* **19**:6, 3642. [Crossref]
6. Orlando Gomes. 2022. Human capital and growth in an OLG-life cycle model. *SN Business & Economics* **2**:1. . [Crossref]
7. Faisal Sohail. 2021. From employee to entrepreneur: Learning, employer size, and spinout dynamics. *Journal of Economic Dynamics and Control* **133**, 104270. [Crossref]
8. Xin Deng, Panpan Lian, Miao Zeng, Dingde Xu, Yanbin Qi. 2021. Does farmland abandonment harm agricultural productivity in hilly and mountainous areas? evidence from China. *Journal of Land Use Science* **16**:4, 433-449. [Crossref]
9. Daniel Erdsiek. 2021. Dynamics of overqualification: evidence from the early career of graduates. *Education Economics* **29**:3, 312-340. [Crossref]
10. Ling-Yun He, Xiao-Feng Qi. 2021. Resource Misallocation and Energy-Related Pollution. *International Journal of Environmental Research and Public Health* **18**:10, 5158. [Crossref]
11. Justin Yifu Lin, Wei Wang, Venite Zhaoyang Xu. 2021. Catch-up industrial policy and economic transition in China. *The World Economy* **44**:3, 602-632. [Crossref]
12. Ruipeng Tan, Mengmeng Xu, Chuanwang Sun. 2021. The impacts of energy reallocation on economic output and CO2 emissions in China. *Energy Economics* **94**, 105062. [Crossref]
13. Xu Dong, Yali Yang, Xiaomeng Zhao, Yingjie Feng, Chenguang Liu. 2021. Environmental Regulation, Resource Misallocation and Industrial Total Factor Productivity: A Spatial Empirical Study Based on China's Provincial Panel Data. *Sustainability* **13**:4, 2390. [Crossref]
14. Vishal K. Gupta. Purpose 43-72. [Crossref]
15. Stephan Heblich, Alex Trew. 2019. Banking and Industrialization. *Journal of the European Economic Association* **17**:6, 1753-1796. [Crossref]
16. Maria Minniti, Martin Andersson, Pontus Braunerhjelm, Frédéric Delmar, Annika Rickne, Karin Thorburn, Karl Wennberg, Mikael Stenkula. 2019. Boyan Jovanovic: recipient of the 2019 Global Award for Entrepreneurship Research. *Small Business Economics* **53**:3, 547-553. [Crossref]
17. Chao Feng, Miao Wang. 2019. The heterogeneity of China's pathways to economic growth, energy conservation and climate mitigation. *Journal of Cleaner Production* **228**, 594-605. [Crossref]
18. Ioana A. Igna, Francesco Venturini. 2019. The impact of educational mismatch on returns to R&D: evidence from manufacturing in OECD countries. *Economics of Innovation and New Technology* **28**:5, 435-464. [Crossref]

19. Yongjian Liu, Hong Geng. 2019. Regional Competition in China under the Price Distortion of Construction Land: A Study Based on a Two-regime Spatial Durbin Model. *China & World Economy* **27**:4, 104-126. [Crossref]
20. David de la Croix, Matthias Doepke, Joel Mokyr. 2018. Clans, Guilds, and Markets: Apprenticeship Institutions and Growth in the Preindustrial Economy*. *The Quarterly Journal of Economics* **133**:1, 1-70. [Crossref]
21. Yen-Po Chen, Ting-Wei Lai, Wen-Chieh Lee, Hao-Chung Li. 2017. Trade barrier and misallocations: The case of the photovoltaic manufacturing industry in China. *International Review of Economics & Finance* **52**, 352-367. [Crossref]
22. Hector Chade, Jan Eeckhout, Lones Smith. 2017. Sorting through Search and Matching Models in Economics. *Journal of Economic Literature* **55**:2, 493-544. [Abstract] [View PDF article] [PDF with links]
23. Hao-Chung Li, Wen-Chieh Lee, Bo-Ting Ko. 2017. What determines misallocation in innovation? A study of regional innovation in China. *Journal of Macroeconomics* **52**, 221-237. [Crossref]
24. Joel M. David, Hugo A. Hopenhayn, Venky Venkateswaran. 2016. Information, Misallocation, and Aggregate Productivity *. *The Quarterly Journal of Economics* **131**:2, 943-1005. [Crossref]
25. Pieter A. Gautier, Coen N. Teulings. 2015. SORTING AND THE OUTPUT LOSS DUE TO SEARCH FRICTIONS. *Journal of the European Economic Association* **13**:6, 1136-1166. [Crossref]
26. Axel Anderson. 2015. A dynamic generalization of Becker's assortative matching result. *Journal of Economic Theory* **159**, 290-310. [Crossref]
27. Burak Uras, Ping Wang. 2014. Techniques Choice, Misallocation and Total Factor Productivity. *SSRN Electronic Journal* . [Crossref]