Misallocation and Growth

Boyan Jovanovic*

July 8, 2012

Abstract

This paper models growth via on-the-job learning in which learning is a by-product of production, and when both firms and workers are heterogeneous. Two-period OLG. The young must work with the old, and as a result an assignment problem arises. We can calculate the transition dynamics analytically when the skill of the young is log-normally distributed and the initial human capital of the old generation is also log-normal. Assignment frictions affect long-run growth. Rates of convergence to the BGP are slow. When types are private information, a separating equilibrium emerges in which the allocations stay the same as in the full-information case, but not the prices.

1 Introduction

This paper derives an economy’s growth rate as a function of the quality of the sorting mechanism. Thus we focus not on investment rates, as in most growth models, but on the efficiency of sorting. The latter depends, in turn, on how well the quality of workers is measured.


This paper is about labor-market misallocation and growth. Growth arises because of on-the-job training and because of external effects flowing from the old to the young via some process such as the quality of elementary schooling or upbringing by parents. Both mechanisms are in Lucas (1988) in a representative agent model.

*Economics Dept. NYU. I thank many colleagues for comments, especially A. Anderson, F. Cunha, E. Hanushek, A. Monge-Naranjo, G. Navarro, and G. Violante.
Boyd & Prescott (1987) is similar because it has two-period lived overlapping generations, because it has no vintage capital and because growth is endogenous. In my model firms are heterogeneous, and the skills that a young agent can learn differ from firm to firm as in Chari & Hopenhayn (1991), Jovanovic and Nyarko (1995) and Monge-Naranjo (2011) and wages include a compensating differential. Even within these models one can study misallocation of labor across firms or technologies and it would have consequences for growth in the Boyd-Prescott model, and for the transitional dynamics in the Chari-Hopenhayn model. My model adds heterogeneity to the abilities of each young generation so that a Becker-Brock type of assignment problem arises (Becker 1973). Firms are perfect competitors in both the product and factor markets.

When agents in a generation are homogeneous, the model is of an “Ak” type, with no transitional dynamics. With heterogeneity among the talent of the young, transitional dynamics arise, and in the long run, the inequality among the old is determined by the (exogenous) inequality at birth in innate skills. We then add a friction to allocation in the factor market. The friction is expressed in the noise in a public signal of a worker’s ability; the noise impedes the market’s ability to reach a positive assignment, an idea that is in MacDonald (1982) and Kremer (1993). The friction reduces both growth and inequality in the long run. Moreover, when the talent of the young and the signal noise are log-normally distributed, we can solve the dynamics in closed form if the initial distribution of human capital among the old (which also is the state of the system) is itself log normal.

How much can such a friction explain? Based on numbers from Heathcote et al. (2005), the development gap explained by the model is not likely to exceed 10 percent of the leader in one generation, but the effects cumulate as long as the frictions differential persists. The question is answered also for the transitional episode that may itself be triggered by a change in the accuracy with which skills are measured. At the micro level we find that output and inequality rise when the garbling frictions diminish and assignments more accurately reflect workers’ talents.

We also can solve for the full dynamics for a fairly wide class of . We show global stability, and monotone convergence. In fact, following a reform in which the friction of assignment is somehow permanently reduced, growth and inequality both rise monotonically to their new BGP levels. Thus such a reform induces a positive relation between inequality and growth along the transition path. Whether


2In MacDonald (1982), a continuum of imperfectly labeled types is assigned to two tasks. Kremer generalizes to a continuum types and tasks.

3Ljungqvist and Sargent (1998) and Violante (2002) also study the evolution of inequality in response to a change of regime.
this explains the rise in U.S. wage inequality since the 1970s that Heathcote et al. (2010) document is not clear because be have not seen an accompanying rise in the rate of growth.

The paper also studies the effect about private information about quality. Several sets of assumptions are entertained. It turns out that as long as the private information is one sided, the friction does not change allocations – matching remains positively assortative. The difference is in the wages that support these outcomes. The biggest difference in wages arises when the private is about the quality of the young. In this case, to support the positive sorting assignment, wages must be negatively related to the quality of the firm.

Evidence

A. Substitutability in production and training: Some evidence on this assumption follows.

(i) Skilled-Unskilled.—Griliches (1989), Krusell et al. (2000). Katz and Murphy (1992) elasticity of substitution is 1.4. But much of this evidence assumes aggregation of efficiency units and is not relevant here.

(ii) Firm-worker.—Lentz and Mortensen (2010) survey the evidence and the issues. Interpreting the wage data required modeling assumptions, and the same is true here. E.g., if a worker’s quality is his private information, wages are then negatively correlated with firm quality. Lazear, Shaw, and Stanton (2011) find a complementarity between the qualities of bosses and their workers. Andersson et al. (2006) find strong relationships between a firm’s existing talent, newly hired talent, and overall quality of software firms. Gabaix and Landier (2008) find that better CEOs match with larger firms. Gavilan (2012) documents some recent rise in segregation of worker quality in plants in the U.S. presumably because qualities became more complementary.

(iii) Teacher-student.—Evidence on production functions for learning appears to be scant. Hanushek (2011) and the empirical literature on economics of education usually assumes a cross-partial derivative that is zero, which allows the use of linear fixed-effect models. An exception is Lockwood and McCaffrey (2009) who find weak positive complementarity between quality of teachers and students. Behrman, Todd and Wolpin (in progress) find complementarity between teacher and student’s efforts in generating student achievement.


C. Gibrat’s Law.—Evans (1987), Luttmer (2010), Sutton (1997) provide evidence that firm TFP and output show a reversion to a common trend. In the model such reversion is caused by assignment frictions.

Plan of paper.—Section 2 deals with homogeneous generations; its purpose is to clarify the growth mechanism. Section 3 adds heterogeneity in each young generation and defines an equilibrium. Section 4 solves for the balanced growth path prices and allocations. Section 5 deals with several ways in which private information may affect the outcome. Section 6 introduces a different, “mislabeled” friction that necessarily leads to a garbled assignment and that features globally stable transition dynamics, and it quantifies the long-run effects of that friction on allocations. Section 7 discusses some other models that feature ability signals, and Section 8 concludes.

2 Homogeneous skill endowments

The model is a 2-period overlapping generations with lifetime utility depending linearly on consumption in youth, $c_y$, and consumption in old age, $c_o$:

$$c_y + \beta c_o$$

An agent is endowed with a unit of labor in both periods of life. Population is fixed, with a unit measure of young being born each period. Let $x_t$ denote the skill of each old agent and $y_t$ the skill of each young agent at date $t$.

Firms.—A firm must have one old & one young worker. The idea is that a manager must be more experienced than the workers he oversees. The firm produces a consumption good and training using the following production functions:

$$\text{output} \quad q_t = f(x_t, y_t) \quad (1)$$

$$\text{training} \quad x_{t+1} = \phi(x_t, y_t) \quad (2)$$

where $f$ and $\phi$ are homogeneous of degree one.

---

4“Oversight” may be a misnomer because incentive problems in this two-member team are assumed to be absent. If, say, the young worker must to exert costly effort, Alchian and Demsetz (1972) argue that efficiency dictates that a principal should monitor the behavior of the agent, and that the monitor should be the residual income recipient. The old member of the firm is naturally viewed as having the monitor role.

5A version of (1) and (2) is analyzed by Anderson and Smith (2010) and Anderson (2011); they allow for randomness in (2), but they specialize in other ways: They analyze a partnership model with no defined roles and symmetric production and transitions, and their dynamic model degenerates into a static model whenever perfect sorting obtains, since in that case no types change.
External effect.—Let $\varepsilon$ denote as an innate talent parameter for a young agent. In this section $\varepsilon$ is the same for each young agent in each generation. There is an externality operating on the innate skill $\varepsilon$, transforming the latter into productive skill $y$ as follows:

$$y = b\varepsilon \bar{x} > 0.$$  
(3)

The term $\bar{x}_t$ is the skill of the each old agent, and it exerts a positive spillover effect on the marketable skill of the young. Since the old agents are all the same,

$$\bar{x}_t = x_t.$$ 

Substituting from (3) into (1) we find that output is

$$q_t = f(x_t, b\varepsilon x_t)$$  
(4)

This external effect is not needed for growth to arise in the homogeneous case, because we could simply assume an Ak technology in (2) so that $x_{t+1} = Ax_t$, whence the growth factor would become exogenous and equal to $A$. We shall return to this in the next section where we introduce heterogeneity among the young.6

LR growth.—The firm’s capital grows as follows:

$$\frac{x_{t+1}}{x_t} = \phi \left( 1, \frac{1}{x_t} b\varepsilon x_t \right) = \phi (1, b\varepsilon) \equiv \Gamma.$$  
(5)

Output then also grows at the same rate:

$$\frac{q_{t+1}}{q_t} = \frac{f(x_{t+1}, b\varepsilon x_t)}{f(x_t, b\varepsilon x_t)} = \frac{x_{t+1} f(1, b\varepsilon)}{x_t f(1, b\varepsilon)} = \Gamma.$$ 

To sum up: A CRS model (1)-(2) with a fixed factor y is converted into an Ak model via the external effect (3) making $y$ a de-facto reproducible factor. The homogeneous-generation model has no transitional dynamics. Note that neither $\phi$ nor the spillover $b\bar{x}$ can generate long-run growth on its own; both are needed for $\Gamma > 1$ as $x \to \infty$.

Example.—Let $\phi$ be Cobb-Douglas:

$$\phi (x, y) = Ax^{1-\theta}y^\theta$$  
(6)

Their focus has thus been characterizing matching patterns that are not perfectly assortative, and the resulting type, wage, and inequality dynamics. In my model, by contrast, the old and young have defined roles in production, different distributions over human capital, and interesting dynamics obtain even with assortative matching. So these two papers on the one hand and mine on the other have simplified along different dimensions toward a different sort of theory, and studying each of these cases in isolation appears to simplify the task.

6 Instead of (3), we could have put the external effect directly into $f$ and $\phi$ as a labor-augmenting (i.e., $\varepsilon$-augmenting) variable, provided that it was of the same form in both places.
Then (5) implies that
\[ \Gamma = A (b \varepsilon)^\theta. \] (7)
The model would be of the Ak type if \( \theta = 0 \); The external effect, \( b \), would then not matter for growth, since it is \( b^\theta \) that matters.

2.1 Policy and multiple equilibria

In spite of the endogeneity of growth, the model is effectively a simple two-period OG model with pure exchange of two indivisible commodities: The young agent exchanges his indivisible unit of time for the old agent’s indivisible firm and the capital that resides in that firm. We have called this human capital, but it could be labeled firm-specific physical capital. Since there are equal numbers on both sides of the market, the market clearing wage is indeterminate. And a tax on wages and profits would have no allocative effects. The indeterminacy of wages goes away if we add population growth; this transfers all rents to the old generation but keep the solution much the same, as Boyd and Prescott’s (1977) related model would show.

When the young are born heterogeneous, however, the mathematics stays simple as long as population growth is zero or declining. We shall then be able to work out transitional dynamics, and long-run growth will be related to that in (5), but will also depend on the frictions in assignment.

3 Heterogeneous skill endowments and no frictions

We continue to maintain the production function \( f \) in (1) and the training function \( \phi \) in (2). But the inputs \((x, y)\) in both will now be heterogeneous. The state of the system will be the C.D.F. of \( x \) denoting the distribution of the skills among the members of the current old generation. Denoted by \( H(x) \), this distribution will also summarize a direct influence that the skills of the old will exert on the skills of the young. This influence will be external, perhaps operating through the schooling mechanism. The C.D.F. of skills among the young will be denoted by \( G \). To summarize,

\[
\begin{align*}
x &= \text{skill of an old agent } x \sim H(x), \\
y &= \text{skill of a young agent } y \sim G^H(y),
\end{align*}
\]

where \( H \) represents external effect.

Why the external effect in \( G \)?—Why do we assume the external effect in eq. (3)? This is because the young are heterogeneous, and because the assignment of young to old is one-to-one. In particular, we certainly do not need the external effect (3) if
the young are homogeneous, as indeed they were in section 2. If \( y \) were the same for all the young, we could replace (1) and (2) by

\[
q = Ax\hat{f}(y) \quad \text{and} \quad x' = x\hat{\phi}(y),
\]

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 \( y \) so that the equilibrium assignment of \( y \) to \( x \) is positive at each date, then (8) would yield a permanently higher growth for firms \( x \) that employ the high \( y \) type, and the coefficient of variation of \( x \) would go to infinity, contrary to what the data seem to show. Income inequality does not appear to be exploding. To prevent that from happening, we introduce diminishing returns to both \( x \) and \( \hat{\phi} \), but then generate growth in \( y \) via a linear external effect of \( x_H \), and this allows both growth and inequality to settle down on the BGP.

3.1 Recursive equilibrium

In all that follows, the functions \( f \) and \( \phi \) will be assumed to possess positive cross partials and that a positive assignment of \( y \) to \( x \) obtains at each date. We shall describe a Markovian equilibrium in which the valuation, \( \pi(x) \), of human capital \( x \) by an old agent is defined recursively. Wages will then be stated in terms of \( \pi(.) \). Carrying the state “\( H \)” as an argument will go on for a few equations, and we shall drop it after equilibrium is stated. Although firms and workers are heterogeneous, assignment will be frictionless, the prices of each type of worker will be taken as given by the firm as in complete-markets theory.

**Assignment.**—The assignment of the young to the old will be written as \( y = \alpha(x) \).

For a positive-sorting assignment \( \alpha \) to clear the market requires that

\[
H(x) = G^H(y)
\]

for all \( x \), so that the assignment is

\[
\alpha^H(x) = (G^H)^{-1}(H[x]).
\]

**Aggregate law of motion \( \chi \).**—This function maps the set of C.D.F.s on the positive line into itself. Letting a prime \( \omega^n \) denote the next-period’s value of \( H \),

\[
H' = \chi(H),
\]

Let \( x' = \xi^H(x) \) denote the evolution of productivity in firm \( x \) when the aggregate state is \( H \). That is,

\[
\xi^H(x) = \phi(x, \alpha^H(x))
\]

Adding up we require that for all \( x' \), the number of tomorrow’s old with skills \( \leq x' \) be the same as the number of today’s old \( x \) that, when matched with a young person
of type \( \alpha(x) \), will generate training \( \leq x' \). I.e., for all \( x' \), \( H' \) must satisfy \( H'(x') = H' \left[ (\xi^H)^{-1}(x') \right] \). Stated compactly, the aggregate law of motion is

\[
\chi(H) = H \cdot (\xi^H)^{-1}.
\]  

(13)

It represents an equilibrium evolution if we can show that the positive assignment – which has so far been assumed to hold at each date – can be decentralized. Note, however, that (10) and (13) do not refer to prices which means that the calculation of the equilibrium is itself recursive – the quantities can be calculated first, and then the prices.

Wages and profits.—Denote by \( w^H(x) \) the wage paid by firm \( x \) in state \( H \). The profits of firm \( x \) and the consumption of an old agent of type \( x \) then are

\[
\pi^H(x) = f(x, \alpha^H(x)) - w^H(x)
\]  

(14)

Equilibrium lifetime utility of a young agent of type \( \chi \) in state \( H \) is:

\[
V^H(\chi) = w^H \left( \left( \alpha^H \right)^{-1}(\chi) \right) + \beta \pi^H \left( \phi \left( \left( \alpha^H \right)^{-1}(\chi), \chi \right) \right)
\]  

(15)

It depends on wages received in youth, and on the training carried into old age.\(^7\)

Firm \( x \)’s decision problem.—Firm \( x \) chooses a type to match with, \( y \), and a wage to pay him, \( w \), so as to maximize its consumption and subject to providing agent \( y \) with his equilibrium lifetime utility

\[
\pi^H(x) = \max_{w,y} \{ f(x,y) - w \}
\]  

(16)

s.t.

\[
w + \beta \pi^H \left( \phi[x,y] \right) \geq V^H(y).
\]  

(17)

Now, (17) must bind at an optimum, and it can be used to eliminate \( w \) from (16) to yield a Bellman equation for \( \pi^H \):

\[
\pi^H(x) = \max_y \left\{ f(x,y) - V(y) + \beta \pi^H \left( \phi[x,y] \right) \right\}.
\]  

(18)

This equation takes \( \chi \) as parametrically given, and \( \chi \) is given in (13) with no reference to \( \pi, w, \) or \( V \). This is because quantities can be calculated with no reference to prices.

Definition of equilibrium.—It consists of functions \( \alpha \) and \( \chi \) satisfying (10) and (13), and \( \pi, w, \) and \( V \) satisfying (15), (16), and (18) with

\[
\alpha^H(x) = \arg \max_y \{ \text{RHS of (18)} \}.
\]  

(19)

\(^7\)Note the updating of \( H \) to \( \psi(H) \) on the RHS of (15).
4 The BGP under full information

To have a balanced-growth path (BGP) on which long-run growth in \( x \) coexist with a CRS form for \( \phi \) in (2), we shall need the spillover that is captured by the effect of \( H \) in \( G^H (y) \) to roughly raise the talent of the young to a level proportional to the talent of the old.

4.0.1 Spillover mechanism

Let \( \bar{x}_H = \int x dH (x) \). A young agent with raw talent \( \varepsilon \) will then have market skill.

\[
y = b \bar{x}_H \varepsilon. \tag{20}
\]

In other words, except for the substitution of \( \bar{x}_H \) for \( \bar{x}_t \), (20) is the same as (3) in how a given agent’s \( \varepsilon \) translates into his \( y \). But now the young differ in their raw talents: \( \varepsilon \) is assumed to be distributed according to the C.D.F. \( \hat{G} (\varepsilon) \), assumed to be identical over generations. Together with (20), this implies that \( y \) is distributed according to the C.D.F.

\[
G^H (y) = \hat{G} \left( \frac{y}{b \bar{x}_H} \right). \tag{21}
\]

Now we explain how even under heterogeneity, the mechanism in (20) is likely to enable a balanced-growth path (BGP) to exist. the role that For any C.D.F. \( H (x) \) and constant \( \Gamma > 0 \), let the resulting distribution of \( \tilde{x} = \Gamma x \) be denoted by

\[
H^\Gamma (\tilde{x}) = H \left( \frac{\tilde{x}}{\Gamma} \right). \tag{22}
\]

The assignment that enables \( H \) to evolve into \( H^\Gamma \) must be homogeneous of degree one in \( (x, \Gamma) \):

**Proposition 1** If (20) holds,

\[
\alpha^{H^\Gamma} (x) = \Gamma \alpha^H \left( \frac{x}{\Gamma} \right) \tag{23}
\]

**Proof.** \( \alpha^{H^\Gamma} (x) \) solves for \( y \) the equation (9), which reads

\[
H^\Gamma (x) = G^{H^\Gamma} (y). \tag{24}
\]

But according to (21), \( G^{H^\Gamma} (y) = \hat{G} \left( \frac{y}{b \bar{x}_{H^\Gamma}} \right) = \hat{G} \left( \frac{y}{b \bar{x}_H} \right) = G^H \left( \frac{y}{\Gamma} \right) \), the second equality follows because \( \bar{x}_{H^\Gamma} = \Gamma \bar{x}_H \). Substituting into (24) and using (22), \( \alpha^{H^\Gamma} (x) \) solves for \( y \) the equation

\[
H \left( \frac{x}{\Gamma} \right) = G^H \left( \frac{y}{\Gamma} \right) \Rightarrow \alpha^{H^\Gamma} (x) = \Gamma (G^H)^{-1} \left[ H \left( \frac{x}{\Gamma} \right) \right] = \Gamma \alpha^H \left( \frac{x}{\Gamma} \right)
\]
Lemma 1 shows that if all agents’ \( q \)'s are multiplied by \( \Gamma \), each old agent will be able to match a young agent with quality \( \Gamma q \), where \( q \) was that agent’s previous assignment. In other words, regardless of \( H \), if every \( x \) were to double, the quality of the equilibrium assignment of \( q \)'s would double too.

The result in (23) does not invoke the linear homogeneity of \( \phi \); it relies only on (20) and on the assumption that each \( x \) has grown by the same proportion. But the result makes it clear that since \( f \) and \( \phi \) indeed are both linearly homogeneous, the economy is capable of growing at a constant rate once a particular \( H \) is attained. If it reaches such an \( H \), the economy will then be on its BGP.

4.0.2 \((\alpha, \Gamma, H)\) on the BGP

Recall that we can solve for the quantities first, and the prices later. So, we first solve for the triple \((\alpha, \Gamma, H)\) on the BGP.

**Proposition 2** On the BGP,

\[ \alpha(x) = b\bar{x}x, \quad \text{and} \]
\[ H(x) = \hat{G}\left(\frac{\varepsilon}{C}\right), \]

where \( C > 0 \) is any real number. If, at \( t = 0 \), \( H_0 = \hat{G}\left(\frac{\varepsilon}{C_0}\right) \), then

\[ H_t = \hat{G}\left(\frac{\varepsilon}{C_0\Gamma^t}\right) \]

where

\[ \Gamma = \phi(1, b\bar{x}) \]

is the growth factor.

**Proof.** Suppose the economy is at \( H \). If each \( x \) grows at some common factor \( \Gamma \), in terms of (12), this means that for all \( x \) in the support of \( H \), \( \Gamma x = \phi(x, \alpha^H(x)) \) or, dividing both sides by \( x \), that

\[ \Gamma = \frac{1}{x}\phi(x, \alpha^H(x)) = \phi\left(1, \frac{\alpha^H(x)}{x}\right) \]

where the second equality follows by the linear homogeneity of \( \phi \). This means that

\[ \alpha^H(x) = \alpha^*x, \]

where \( \alpha^* > 0 \) is a constant satisfying \( \phi(1, \alpha^*) = \Gamma \). Evaluating the RHS of (29) at the means, \( \bar{y} = \alpha^*\bar{x} = b\bar{x}\bar{e} \) (using (20)), which, together with (29) implies (25). Dividing \( \alpha^*\bar{x} = b\bar{x}\bar{e} \) by \( \bar{x} \) implies

\[ \alpha^* = b\bar{e}, \]
which proves (28). Next, (20) and (29) imply $\alpha^* x = b \bar{\varepsilon}$, and this leads to

$$\frac{x}{\bar{x}} = \frac{b}{\alpha^* \varepsilon} = \frac{\varepsilon}{\bar{\varepsilon}}$$

where the second equality follows from (30). Therefore $x$ is a scaled version of $\varepsilon$ which proves (26). Finally, with a constant growth factor $\Gamma$, (27) follows $\blacksquare$

Remarks on Proposition 2:

1. All firms’ TFPs grow at the same factor $\Gamma$ given in (28). Notice that (28) and (5) become the same if the $\varepsilon$’s are identical.

2. Growth depends positively on the talent of the young and on the externality parameter $b$, as well as on the parameters of $\phi$. E.g.,

$$\phi(x, y) = A x^{1-\theta} y^\theta \Rightarrow \Gamma = A (b \bar{\varepsilon})^\theta$$

(31)

Again, note the correspondence to (7)

3. Inequality depends solely on the inequality the young’s talents $\varepsilon$, and (29) implies that the inequality in $\ln x$ and $\ln \varepsilon$ are the same

4. The assignment in (25) is a special case of the assignment in (23). The restriction in (25) follows from the linear homogeneity of $\phi$ that Proposition 2 invokes.

4.0.3 $(w, \pi, V)$ on the BGP

We now solve for the prices that make the BGP allocations incentive compatible.

**Proposition 3** On the BGP wages and profits are

$$w(x) = \omega f(1, b \bar{\varepsilon}) x$$
$$\pi(x) = (1 - \omega) f(1, b \bar{\varepsilon}) x$$

(32) (33)

and lifetime utility is

$$V(y) = \left( \omega f \left( \frac{1}{b \bar{\varepsilon}}, 1 \right) + \beta (1 - \omega) \frac{\Gamma}{b \bar{\varepsilon}} \right) y,$$

(34)

where the constant share of output paid to the young is

$$\omega = \frac{f_y(1, b \bar{\varepsilon})}{f(1, b \bar{\varepsilon})} + \beta \phi_y(1, b \bar{\varepsilon}) - \frac{\bar{\varepsilon} \beta \Gamma}{b \bar{\varepsilon}^2 + \beta \phi_y(1, b \bar{\varepsilon}) - \frac{\bar{\varepsilon} \beta \Gamma}{b \bar{\varepsilon}}},$$

(35)
**Proof.** We proceed by construction. That is, we substitute the claimed prices into the Bellman equation (18), and show that they solve this equation as well as its FOC and that the solution for \( y \) of the FOC is then given by (25). This is the unique maximizing choice of \( y \) for the firm because its objective will be shown to be strictly concave in \( y \). With the substitution from (32)-(34), (18) reads

\[
(1 - \omega) \bar{f} x = \max_y \left\{ f(x, y) - V(y) + \beta (1 - \omega) \bar{f} \phi(x, y) \right\}
\]

where \( \bar{f} = f(1, b\bar{e}) \). The FOC is

\[
0 = f_y \left( 1, \frac{y}{x} \right) - V_y + \beta (1 - \omega) \bar{f} \phi_y \left( 1, \frac{y}{x} \right)
= f_y (1, b\bar{e}) - \omega f \left( \frac{1}{b\bar{e}}, 1 \right) - \beta (1 - \omega) \bar{f} \phi \left( \frac{1}{b\bar{e}}, 1 \right) + \beta (1 - \omega) \bar{f} \phi_y (1, b\bar{e})
\]

where (37) after substituting for the assignment in (25). For any \( \omega \in [0, 1] \), the SOC holds in that the RHS of (36) is concave in \( y \) because \( V \) is linear in \( y \), while and \( f \) and \( \phi \) are strictly concave in \( y \). Concavity is true as long as \( \omega \leq 1 \). Evaluating the RHS of (36) at \( y = b\bar{e} x \), and noting that \( f \left( \frac{1}{b\bar{e}}, 1 \right) y = f \left( \frac{1}{b\bar{e}}, 1 \right) b\bar{e} x = f (1, b\bar{e}) x \) and that \( \phi \left( \frac{1}{b\bar{e}}, 1 \right) y = \phi (1, b\bar{e}) x = \Gamma x \), it reads

\[
(1 - \omega) \bar{f} x = f \bar{x} - V(b\bar{e} x) + \beta (1 - \omega) \bar{f} \Gamma x = \bar{f} x - \omega \bar{f} x - \beta (1 - \omega) \bar{f} \Gamma x + \beta (1 - \omega) \bar{f} \Gamma x = (1 - \omega) \bar{f} x,
\]
i.e., (18) evidently holds for all \((x, \omega)\). Therefore we can just solve the FOC (37) for the one unknown, \( \omega \). Noting that \( f \left( \frac{1}{b\bar{e}}, 1 \right) = \frac{1}{b\bar{e}} \Gamma \) and that \( f \left( \frac{1}{b\bar{e}}, 1 \right) = \frac{1}{b\bar{e}} \bar{f} \), we get

\[
\omega = \frac{f_y (1, b\bar{e}) + \beta \bar{f} \phi_y (1, b\bar{e}) - \frac{1}{b\bar{e}} \beta \bar{f} \Gamma}{f \left( \frac{1}{b\bar{e}}, 1 \right) + \beta \bar{f} \phi_y (1, b\bar{e}) - \frac{1}{b\bar{e}} \beta \bar{f} \Gamma},
\]
i.e., (35). 

**Cobb-Douglas example:** Let \( b\bar{e} = 1 \), and let \( f \) and \( \phi \) both be Cobb-Douglas:

\[
\begin{align*}
  f(x, y) &= x^{1-\rho} y^\rho \\
  \phi(x, y) &= A x^{1-\theta} y^\theta
\end{align*}
\]

\[
\Gamma = A \\
\begin{align*}
  f_y (1, 1) &= \rho \\
  \phi_y (1, 1) &= \theta
\end{align*}
\]

Substituting into (35), we obtain

\[
\omega = \frac{\rho - \beta (1 - \theta) A}{1 - \beta (1 - \theta) A}
\]

Remarks on Proposition 3 and the example.
1. *Compensating differentials.*—As long as \( \beta > 0 \), wages include a compensating differential, and if this differential is large, wages can turn into tuition payments. This can happen if \( x \) matters a lot relative to \( y \) in both production and training. In the Cobb-Douglas example (38) leading to (39), \( \omega < 0 \) iff, \( \rho < \beta A (1 - \theta) \), which reflects the intuition just mentioned. The compensating differential component of \( \omega \) in (39) predictably goes to zero if \( \beta \to 0 \) in which case the young do not value the training that the firm provides, or if \( \theta \to 1 \), in which high-\( x \) firms are no better at providing training than low-\( x \) firms.

2. When \( \omega < 0 \), the young must pay tuition, and for this they would need an endowment, or a second asset like a bond on which the young could go short, or banks that would extend credit. The equilibrium rate would be \( \beta^{-1} - 1 \). But both instruments (credit, short sales) require collateral in the real world, and without an endowment, the only collateral the young could offer would be having employment at the firm, and human capital is not a legally valid form of collateral. This problem arises here only when \( \omega < 0 \), but in Sec. 5.1 which treats the case in which \( y \) is private information, it may always arise for a subset of the agents if \( \varepsilon \) is dispersed enough.

3. Note that even \( \omega < 0 \), the objective in (36) remains concave.

4. *The experience premium.*—Over an agent’s lifetime wages grow faster than output. Wages in youth are a fraction \( \omega \) of the output of the firm in which the young agent works. When the worker gets old, his firm’s output has grown by the factor \( A \), and he then receives a fraction \( 1 - \omega \) of that larger output. Therefore his wage growth and experience premium are

\[
\text{total wage growth} = \frac{1 - \omega}{\omega} A = \frac{1 - \rho}{\rho - (1 - \theta) \beta A} A, \tag{40}
\]

\[
\text{experience premium} = \frac{1 - \omega}{\omega} A - A = \frac{1 - 2\rho + (1 - \theta) \beta A}{\rho - (1 - \theta) \beta A} A. \tag{41}
\]

The parameters affect the experience premium as one would expected. Impatience (a lower \( \beta \)) lowers the premium because workers are not willing to sacrifice current wages as much for the prospect of higher income later. The premium rises with \( A \), which measures the efficiency of the training mechanism. A rise in \( \theta \) diminishes the role of the old in training, and with that, the compensating differential for the training that the young receive goes down. Finally, a rise in \( \rho \) raises the marginal product of the young which raises their current contribution relative to the future product of any training that the young receive.
5 The BGP under private information

So far we assumed that \( \xi \) and \( \eta \) were public information. But either or both may be private information. It turns out that a separating equilibrium still exists and that it supports the same assignment and the same BGP, the wages support it differ from (32). The main difference arises when \( y \) is private information of the young, in which case better firms pay lower wages. By contrast, when the quality of firms is private information, wages are increasing in the quality of firms, but the compensating differential for training disappears. In either case, allocations are unchanged by the information friction because a separating equilibrium overcomes them. Allocations change only if the quality of agents on both sides of the market is private information, we get a constant wage and a random assignment.

5.1 Private information about \( \eta \)

A separating equilibrium with positive sorting still survives if \( \eta \) is known privately, but now it requires that high-\( \xi \) firms pay lower wages regardless of the parameters. Wages are still paid up front and not contingent on the output produced. As long as \( \theta > 0 \), the worker’s quality affects how much training is received, and this also depends on \( \xi \), and this is an effective screening device because the high-\( y \) types would pay more to work with a high \( \xi \) type.\(^8\)

Assume that \( \xi \) (and hence whatever \( \xi' \) that worker \( \eta \) ends up with) is public information. In this case wages must be decreasing in \( x \) regardless of the parameters, otherwise every worker would seek the highest-\( x \) job since that job would offer highest wages in youth and in old age. Write the wage as

\[
w(x) = w_0 \bar{x} - w(x). \tag{42}
\]

Consider the decision problem of a worker of type \( \eta \). He seeks to maximize his lifetime value over the choice of the type of firm \( x \). Then, assuming that he will face the same linear assignment (30) in old age,

\[
\hat{V}(\eta) = \max_x \{w_0 \bar{x} - wx + \beta (-w_0 \bar{x} \Gamma + [w + f(1, \beta \bar{\varepsilon})] \phi (x, y))\} \tag{43}
\]

where (28) still holds so that \( \Gamma = \phi (1, \beta \bar{\varepsilon}) \). The problem is strictly concave in \( x \). The FOC is

\[
w = \beta (w + f(1, \beta \bar{\varepsilon})) \phi_x (x, y)
= \beta (w + f(1, \beta \bar{\varepsilon})) \phi_x (1, \beta \bar{\varepsilon})
\]

and therefore at the uniquely maximizing assignment,

\[
w = \frac{\beta f(1, \beta \bar{\varepsilon}) \phi_x (1, \beta \bar{\varepsilon})}{1 - \beta \phi_x (1, \beta \bar{\varepsilon})}. \tag{44}
\]

\(^8\)The intuition goes back to at least Salop and Salop (1976) where the unobservable type was the propensity to quit a job, and where a stable worker prefers a backloaded contract.
Proposition 4 When \( y \) is private information, \((\alpha, H, \Gamma)\) still satisfy (25)-(28), but \( w(x) \) is given in (42) in which the slope coefficient is given in (44) and in which \( w_0 \) is arbitrary.

Remarks:

1. The constant \( w_0 \) affects only the distribution of income between the old and the young, as well as the lifetime value of the young in (43).

2. \( w'(x) = -w \) is the compensating differential per unit of \( x \). This means that if \( x \) (and ultimately \( \varepsilon \)) is dispersed enough, some workers will receive negative wages.

Cobb-Douglas example, again: Just as in (38), let \( b\varepsilon = 1 \), and let \( f \) and \( \phi \) both be Cobb-Douglas: \( f(x, y) = x^{1-\rho}y^\rho \), \( \phi(x, y) = Ax^{1-\theta}y^\theta \), so that \( \Gamma = A \), \( f(1, 1) = 1 \) and \( \phi_x(1, 1) = (1 - \theta)A \). Substitute these values into (44) to get the compensating differential per unit of \( x \):

\[
w = \frac{\beta (1 - \theta) A}{1 - \beta (1 - \theta) A}
\]

(45)

Note that in this case, \( w \) is identical to the compensating differential component of \( \omega \) in (39). The differential again goes to zero if \( \beta \to 0 \) because then the young do not value training, or if \( \theta \to 1 \), in which case high-\( x \) firms provide the same training as low-\( x \) firms.

5.2 Private information about \( x \)

We now show that if \( x \) is private information, a BGP with a separating equilibrium still exists, but the wage must now be higher for higher \( y \) types. The hypothetical BGP wage paid to a young type \( y \) is \( w^H(y) \equiv \Gamma'w(y\Gamma^{-1}) \). Since worker \( y \) matches with \( b\varepsilon x \), and since \( x \) grows by the factor \( \Gamma \), under balanced growth, the worker’s \( y \) would become \( (b\varepsilon)^{-1} y\Gamma \) in the next period, and next period the output of his firm would equal

\[
f[(b\varepsilon)^{-1} y\Gamma, y\Gamma] = f[(b\varepsilon)^{-1}, 1] y\Gamma
\]

and the wage that he would pay to the young worker that he would then hire would be

\[
\Gamma w \left( \frac{y\Gamma}{\Gamma} \right) = \Gamma w(y)
\]

Therefore, in such an equilibrium, his lifetime value would be

\[
V(y) = w(y) + \beta \Gamma \left[ f[(b\varepsilon)^{-1}, 1] y - w(y) \right]
\]

This lifetime value, however, will not play the same role in the decision of the firm that the value in (15) played when \( x \) was known.
Consider now the problem of the old, i.e., the private information counterpart to the problem posed in (16)-(18). We simplify to the BGP version. The key difference, however, is that a worker of type $y$ assumes that the firm hiring him is of type $(b \varepsilon)^{-1} y$ because that’s what the equilibrium assignment is. Thus a worker that is paid $w(y)$ assumes he is being hired by firm $(b \varepsilon)^{-1} y$, and that as a result he will get earn his equilibrium lifetime value $V(y)$. The problem becomes one without a constraint of the type expressed by (15). Rather, firm $x$ solves the unconstrained, strictly concave problem:
\[
\max_y \{ f(x, y) - w(y) \} \tag{46}
\]
with the FOC (evaluated at the equilibrium assignment) equal to
\[
w'(y) = f_y ((b \varepsilon)^{-1} y, y) = f_y ((b \varepsilon)^{-1}, 1).
\]
which means that the equilibrium wage is
\[
w(y) = C + f_y ((b \varepsilon)^{-1}, 1) y \tag{47}
\]
where $C$ is a constant. We summarize this case as follows:

**Proposition 5** When $x$ is private information, $(\alpha, H, \Gamma)$ still satisfy (25)-(28), but $w(x)$ is given in (47) in which $C$ is arbitrary

Remarks:

1. Wages reflect only the properties of $f$. Since $x$ is not observed by the young, incentive compatibility for $x$ does not involve the properties of $\phi$.

2. The constant $C$ affects only the distribution of income between the old and the young.

3. The expression in (47) is increasing in $y$ and since sorting is positive, high-$x$ firms pay higher wages. Compared to the other two BGP wages in (32) and (42), this case displays the steepest relation between $x$ and $w$, because $w(y)$ in (47) contains no compensating differential for training. Separation requires that this should be so.

**Cobb-Douglas example once again:** Again, let $b \varepsilon = 1$, and let $f(x, y) = x^{1-\rho} y^{\rho}$, so that $f_y (1, 1) = \rho$. Then
\[
w(y) = C + \rho y \tag{48}
\]
5.3 Private information about both $x$ and $y$

It would seem that the only possible outcome is that all firms will pay the same wage, and that the assignment should be random—there is no mechanism to force the positive assignment. Under a random allocation, economy $H$ generates $Q = \int f(x, b x^H \varepsilon) dG(\varepsilon) dH(x)$ and $H'(x_0) = \Pr \{ \phi(x, b x^H \varepsilon) \leq x_0 \mid H \}$. The allocations (but not the wage payments) will correspond to the case in the next section which $r = 0$ which also has a purely random assignment.

6 Assignment frictions

Assignment frictions are now assumed to originate in imperfect information about a young worker’s ability. We shall assume no private information, only imperfect public signals. Let $s =$ publicly observed signal of $y$, a “test score”, distributed as

\[
\Pr (\hat{s} \leq s \mid y) = F(s \mid y).
\]

We may think of $s$ as the CV of a new job applicant—his grade point average, recommendations, etc. Of course this is not an exact measure of his future performance on the job.

Old agents’ productivities are publicly known, perhaps because the firm’s output is observed. Then a young worker’s type is $s$ and wages and assignment depend on $s$. In other words, we continue to assume a frictionless assignment conditional on $s$;

We match $x$ to $s$ using the distributions $\Phi$ and $H$ as shown below:

\[
\Phi(s) = \int F(s \mid y) dG(y) = \text{signal distribution}
\]

\[
s = \Phi^{-1}(H(x)) = \alpha(x) = \text{assignment}
\]

\[
\pi(y \mid s) = \text{posterior}
\]

\[
\pi(y \mid \alpha^{-1}(x)) = \text{distribution of } y \text{ given } x
\]

\[
\Pr (\bar{x} \leq x' \mid x) = \Pr (Ax^{1-\theta}y^{\theta} \leq x' \mid x) = \pi \left( \frac{x'}{Ax^{1-\theta}} \right)^{1/\theta} \mid \alpha(x)
\]

because $x' = Ax^{1-\theta}y^{\theta}$.

Log normal $\hat{G}, H_0,$ and $F.$—The rest of this section assumes log normality of $\hat{G}(\varepsilon)$ and of the initial conditions for $H(x)$; the noise in the signal will also be assumed to be log normal. These assumptions will allow us to characterize the dynamics off the BGP and on. Let

\[
\hat{s} = \hat{y} + \eta
\]

---

9If the quality $x$ of the manager is known, observing output $q = f(x, y)$ perfectly reveals $y$, and this information can then be used to infer $x' = \phi(x, y)$. 

17
where hats denote logs, and where \( E(\eta) = 0 \). Denote the squared correlation coefficient (proxying the signal-to-noise ratio) by

\[
r^2 = \frac{\sigma_y^2}{\sigma_y^2 + \sigma_\eta^2} = \frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma_\eta^2}.
\]

**Labeling errors vs. other informational frictions.**—An \((x, s)\) match is attained frictionlessly and yet it is permanent – the worker and his manager are stuck with each other for the duration of the manager’s life. Whenever \( r^2 < 1 \), there will ex post arise a mismatch between the worker and his boss, because generally, \( s \neq y \). The smaller is \( r \), the larger the average mismatch, and the higher the ex-post desire to recontract. Since the latter is ruled out, the recontracting cost is infinite. If a third period of life were added, for example, and if the output of each match were publicly observed, a reshuffling would certainly arise. It is tempting to say that reshuffling does not occur because output is private information, but then one could not assume (as this section does) that \( x \) is publicly observed at the stage to which (49) pertains.\(^{10}\)

**BGP growth with perfect signals.**—When \( \sigma_\eta^2 = 0 \) so that signals are perfect – the case we analyzed in the previous section. With log normality so that \( \bar{\varepsilon} \sim N(\mu_\varepsilon, \sigma_\varepsilon^2) \), we have \( \varepsilon = \exp(\mu_\varepsilon + \frac{1}{2} \sigma_\varepsilon^2) \). Assuming still that \( y = b\bar{\varepsilon}, \) (31) then yields the BGP growth factor of

\[
\Gamma = A(b\varepsilon)^\theta = Ab^\theta \exp\left(\theta \mu_\varepsilon + \frac{\theta}{2} \sigma_\varepsilon^2\right) \tag{50}
\]
or, if we take logs and denote them by hats “\( \hat{} \)”,

\[
\hat{\Gamma} = \hat{A} + \theta \left(\ln b + \mu_\varepsilon + \frac{\theta}{2} \sigma_\varepsilon^2\right) \tag{51}
\]

6.1 **The full dynamics**

In the time domain, \( H \) evolves as follows:

\[
H_{t+1} (x') = \int \pi \left( \left( \frac{x'}{Ax_1^{1-\theta}} \right)^{1/\theta} \mid \Phi_t^{-1} (H_t (x)) \right) dH_t (x).
\]

But we can derive the first-order law of motion for \( H \) if \( H_0 \) is log-normal, because then \( H \) remains log-normal for ever, and is fully described by its first two moments. To ease notation, we shall write \((\mu, \sigma)\) for \((\mu_\varepsilon, \sigma_\varepsilon)\) so that (11) becomes

\[
\begin{align*}
\mu' &= \chi_1 (\mu, \sigma) \tag{52} \\
\sigma' &= \chi_2 (\mu, \sigma) \tag{53}
\end{align*}
\]

\(^{10}\)A related set of outcomes when there are search frictions and when search two-sided and directed, as in Chernomykin et al. (2012). Multiple equilibria then arise, entailing different degrees of direction, resembling what one would obtain in this model as \( r^2 \) varies.
I.e., instead of mapping distributions on the line into itself, we shall (without changing the notation) treat $\chi$ as the map $\chi : (R \times R_+) \rightarrow (R \times R_+)$. 

*Deriving $\mu' = \chi_1 (\mu, \sigma)$.*—The assignment in (49) yields the deterministic relation
\[
\dot{s} = \dot{\lambda} (x) = \mu_s + \frac{\sigma_s}{\sigma} (\dot{x} - \mu),
\]
but the relation between $y$ and $s$ and, hence, between $y$ and $x$ is now stochastic: Letting $\zeta \sim N (0, 1)$,
\[
\hat{y} = (1 - r^2) \mu_y + r^2 \dot{s} + \sqrt{1 - r^2} \sigma_y \zeta = (1 - r^2) \mu_y + r^2 \left( \mu_x + \frac{\sigma_x}{\sigma} (\dot{x} - \mu) \right) + \sqrt{1 - r^2} \sigma_y \zeta
\]
\[
= \mu_y + r \frac{\sigma_y}{\sigma} (\dot{x} - \mu) + \sqrt{1 - r^2} \sigma_y \zeta
\]
(55)
because $\mu_y = \mu_s$, $\sigma_y = \sigma_x$, and because
\[
r^2 \frac{\sigma_y}{\sigma} = r \frac{\sigma_x}{\sigma}.
\]
(56)
Since $\phi (x, y) = Ax^{1-\theta} y^\theta$, $x$ evolves as follows:
\[
\dot{x}' = \dot{A} + (1 - \theta) \dot{x} + \theta \hat{y} = \dot{A} + (1 - \theta) \dot{x} + \theta \left( \mu_y + r \frac{\sigma_y}{\sigma} (\dot{x} - \mu) + \sqrt{1 - r^2} \sigma_y \zeta \right)
\]
(57)
Integrating both sides over $H = N (\mu, \sigma^2)$, and noting that that $A \beta \exp (\theta \mu_x + \frac{\theta}{2} \sigma_x^2)$, we have
\[
\mu' = \dot{A} + \theta \mu_y + (1 - \theta) \mu = \dot{A} + \theta \left( \ln b + \mu_x + \mu + \frac{1}{2} \sigma^2 \right) + (1 - \theta) \mu_x
\]
\[
= \mu + \dot{A} + \theta \left( \ln b + \mu_x + \frac{1}{2} \sigma^2 \right)
\]
(58)
Thus the evolution of $\mu$ depends on $\sigma$, but not on $r$. Moreover, (58) makes it clear that the absolute change in $\mu$ is a function of $\sigma$ alone:
\[
\mu' - \mu = \dot{A} + \theta \left( \ln b + \mu_x + \frac{1}{2} \sigma^2 \right) \\
\equiv \chi_3 (\sigma)
\]
(59)
\[
p (\hat{y} | \dot{s}) = \frac{p (\dot{s} | \hat{y}) p (\hat{y})}{p (\dot{s})} \text{ has variance } \frac{\sigma_y^2 \sigma_x^2}{\sigma_y^2 + \sigma_x^2} = (1 - r^2) \sigma_y^2
\]
Therefore this is the variance of $\hat{y}$ conditional on $s$.

Because $r^2 \frac{\sigma_y}{\sigma} = r^2 \sqrt{\frac{\sigma_y^2 + \sigma_x^2}{\sigma_y^2}} = r^2 \frac{\sigma_x}{\sigma} \sqrt{\frac{\sigma_y^2 + \sigma_x^2}{\sigma_y^2}} = r^2 \frac{\sigma_x}{\sigma} = r \frac{\sigma_x}{\sigma}$.
Deriving $\sigma' = \chi_2(\sigma)$.—The evolution of $\sigma$ does not depend on $\mu$. Taking the variance of both sides of (57) we get the difference equation for $\sigma^2$:

$$ (\sigma^2)' = (1 - \theta + \theta r \frac{\sigma_\xi}{\sigma})^2 \sigma^2 + (1 - r^2) \theta^2 \sigma_\xi^2 $$

$$ = (1 - \theta)^2 \sigma^2 + 2\theta (1 - \theta) r \sigma_\xi \sigma + \theta^2 \sigma_\xi^2 $$

The sequence $\sigma_t$ satisfies (60), which is quadratic, and does not contain $\mu$. Therefore

$$ \sigma' = \chi_2(\sigma) = \sqrt{(1 - \theta)^2 \sigma^2 + 2\theta (1 - \theta) r \sigma_\xi \sigma + \theta^2 \sigma_\xi^2} $$

(60)

Figure 1 plots $\chi_2(\sigma)$ for $\theta = 0.5$, and $r \in \{0, 0.5, 1\}$, along with the $45^0$ line. The figure starts three economies with the same level of inequality, namely $\sigma/\sigma_\xi = 0.21$. The $r = 1$ economy sees the fastest rise in inequality right away, the $r = 0$ economy the slowest.

Effect of a permanent rise in $r$.—Suppose there is an unexpected “reform” that raises $r$ from, say, $r = 0$ to $r = 0.5$. We can trace its effect in Fig. 1. The natural starting point is the steady-state value of $\sigma$ that we would read off the intersection of the $r = 0$ schedule with the $45^0$ line. The new steady state is where the $r = 0.5$ line now crosses the $45^0$ line. This would be a major reform, and it would take a couple of
generations (i.e., a couple of steps in the Fig. 1 to get the economy about 90 percent of the way to its new steady state level of inequality.

**Speed of convergence to the BGP.**—Eq. (59) implies that \( \mu' - \mu \) converges to its BGP value as fast as \( \sigma \) converges to its steady-state level \( \sigma (r) \). Therefore, the experiment outlined in the previous paragraphs seem 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.

**BGP inequality.**—That BGP value of \( \sigma \) solves the equation \( \sigma = \chi_2 (\sigma) \) which is a quadratic. The unique positive solution, call it \( \sigma (r) \) is

\[
\sigma (r) = \frac{\sigma_\varepsilon}{2 - \theta} \left( r (1 - \theta) + \sqrt{1 - (1 - r^2) (1 - \theta)^2} \right). \tag{61}
\]

We state a few properties of \( \sigma (r) \) next:

**Proposition 6 (BGP inequality)** The solution (61) is (i) proportional to \( \sigma_\varepsilon \), (ii) increasing and convex in \( r \), and (iii) equals \( \sigma_\varepsilon \) when \( r = 1 \). Moreover,

\[
\frac{\partial \sigma}{\partial r} \geq \left. \frac{\partial \sigma}{\partial r} \right|_{r=0} = \frac{1 - \theta}{2 - \theta} \sigma_\varepsilon \quad \text{and} \quad \tag{62}
\]

\[
\frac{\partial \sigma}{\partial r} \leq \left. \frac{\partial \sigma}{\partial r} \right|_{r=1} = (1 - \theta) \sigma_\varepsilon. \tag{63}
\]

**Proof.** Claims (i) and (iii) are immediate. For (62), (63), and (ii), we note that

\[
\frac{\partial \sigma}{\partial r} = \frac{\sigma_\varepsilon}{2 - \theta} \left( 1 - \theta + r (1 - \theta)^2 \left[ 1 - (1 - r^2) (1 - \theta)^2 \right]^{-1/2} \right),
\]

and that

\[
\frac{\partial^2 \sigma}{\partial r^2} \propto \left[ 1 - (1 - r^2) (1 - \theta)^2 \right]^{-1/2} - r^2 (1 - \theta)^2 \left[ 1 - (1 - r^2) (1 - \theta)^2 \right]^{-3/2} \propto 1 - (1 - r^2) (1 - \theta)^2 - r^2 (1 - \theta)^2 = 2 \theta (1 - \theta) \geq 0.
\]

Thus (62) gives us the lower bound and (63) the upper bound of how “garbling” reduces BGP inequality. We can observe the following:

1. The effects are proportional to \( \sigma_\varepsilon \). Realizing this, it then helps to plot the ratio \( \sigma (r) / \sigma_\varepsilon \) on the parameters, and this is done in Figure 2 which plots \( \sigma / \sigma_\varepsilon \) as a function of \( r \), for three alternative values of \( \theta \).

2. The biggest effect on the inequality of the old occurs when \( \theta \) is low, which means that the effect of the old on training, \( 1 - \theta \) is high. A talented young worker foregoes that amount of training when he is mislabelled and assigned to an average employer. The same conclusion does not pertain to the rate of growth on the BGP: Garbling hurts most when \( \theta \) is in the middle range, so that we get a sort of an inverted-U curve in terms of \( r \), reminiscent of the Laffer curve, except that here we are talking about growth losses and not tax revenues.
Figure 2: Plots of $\sigma(r)/\sigma_\varepsilon$ for $\theta = 0.1$ (blue), 0.5 (red), 0.9 (green)

The next result states that the situation depicted by Figure 1 in which the fixed point is globally stable, in fact obtains for all parameter values.

**Proposition 7** The fixed point (61) of the map $\chi_2(\sigma)$ is unique and globally stable.

**Proof.** Eq. (60) shows that $\chi_2(.)$ is homogeneous of degree 1 in $(\sigma, \sigma_\varepsilon)$. Dividing both sides by $\sigma_\varepsilon$, squaring both sides of the resulting expression, and letting

$$u \equiv \frac{\sigma^2}{\sigma_\varepsilon^2},$$

we have,

$$u' = \theta^2 + 2\theta (1 - \theta) r u^{1/2} + (1 - \theta)^2 u \equiv \tilde{\chi}(u)$$

For all $\theta < 1$, $\tilde{\chi}$ is strictly increasing and concave in $u$, with $\tilde{\chi}(0) = \theta^2$, and $\tilde{\chi}(1) = \theta^2 + 2\theta (1 - \theta) r + (1 - \theta)^2 \leq \theta^2 + 2\theta (1 - \theta) + (1 - \theta)^2 = (\theta + [1 - \theta])^2 = 1$, and the inequality is strict for $r < 1$. Thus $\tilde{\chi} : [0, 1] \rightarrow [0, 1]$, and being continuous, concave and ending below unity, $\tilde{\chi}$ has exactly one fixed point at which it crosses the 45° line from above. ■

The long-run effects of garbling are summarized in the proposition below
Corollary 8 The BGP is globally stable, the growth rate converges to
\[
\hat{\Gamma} = \hat{A} + \theta \left( \ln b + \mu_{\hat{\varepsilon}} + \frac{1}{2} \sigma^2 \right),
\]
and the growth factor to
\[
\Gamma = A (b \mu_{\hat{\varepsilon}})^{\theta} \exp \left( \frac{1}{2} \sigma^2 \right),
\]
where \(\sigma = \sigma (r)\) is given in (61)

Proof. From (58),
\[
\hat{\Gamma} (\sigma) \equiv \chi_1 (\mu, \sigma) - \mu = \hat{A} + \theta \left( \ln b + \mu_{\hat{\varepsilon}} + \frac{1}{2} \sigma^2 \right)
\]
and so if \(\sigma\) converges monotonically to \(\sigma (r)\), so does \(\hat{\Gamma}\) to \(\hat{\Gamma} (\sigma [r])\). ■

By part (iii) of Proposition 6, \(\sigma (1) = \sigma_{\hat{\varepsilon}}\), from which it follows that (50) and (51) are special cases of (66) and (65) when \(r = 1\).

The speed of convergence of \(H\).—Eq. (65) or eq. (66) depict a comparative BGP relation. The speed of convergence of \(\sigma\) and the discussion in the context of Fig. 1 does not apply to \(H\), the same cannot be said of the speed of convergence to the new BGP; the reason is that \(r\) affects \(\Gamma\) only through the parameter \(\sigma\).

6.2 Garbling and growth

Substituting from (61) into (65) we have \(\hat{\Gamma}\) as a function of the parameters alone
\[
\hat{\Gamma} = \hat{A} + \theta \left( \ln b + \mu_{\hat{\varepsilon}} + \frac{1}{2} \left( \frac{\sigma_{\hat{\varepsilon}}}{2 - \theta} \left( r (1 - \theta) + \sqrt{1 - (1 - r^2) (1 - \theta)^2} \right) \right)^2 \right)
\]
and differentiating,
\[
\frac{\partial \hat{\Gamma}}{\partial r} = \frac{\theta}{2 - \theta} \sigma_{\hat{\varepsilon}} \left( r (1 - \theta) + \sqrt{1 - (1 - r^2) (1 - \theta)^2} \right) \left( 1 - \theta + r (1 - \theta) (1 - (1 - r^2) (1 - \theta)^2)^{-1/2} \right)
\]
and so we have
\[
\left. \frac{\partial \hat{\Gamma}}{\partial r} \right|_{r=0} = \frac{\theta}{2 - \theta} \sigma_{\hat{\varepsilon}} (1 - (1 - \theta)^2)^{1/2} (1 - \theta) (1 - (1 - \theta)^2)^{-1/2} = \frac{\theta (1 - \theta)}{2 - \theta} \sigma_{\hat{\varepsilon}},
\]
and
\[
\left. \frac{\partial \hat{\Gamma}}{\partial r} \right|_{r=1} = \frac{\theta}{2 - \theta} \sigma_{\hat{\varepsilon}} (2 - \theta) (1 - \theta + 1 - \theta) = 2 \theta (1 - \theta) \sigma_{\hat{\varepsilon}},
\]
This equation (68) illustrates the inverted-U relation with respect to $\theta$. Changes in $r$ do not matter for BGP growth when $\theta$ is close to zero or close to unity. Maximal effects occur when $\theta$ is close to $1/2$, and evaluated at $\theta = 1/2$, these bounds are

$$\frac{1}{2 - \theta} \frac{\partial \Gamma}{\partial \theta} (1 - \theta) \sigma_{\hat{\xi}} \leq \hat{r} \leq 2 \theta (1 - \theta) \sigma_{\hat{\xi}}$$

\text{(68)}

\textbf{Effect of $r$ on the growth rate.}—One estimate of $\sigma_{\hat{\xi}}$ would use wages. In this case, Heathcote \textit{et al.} (2005 Fig. 3 reported in the Appendix) suggest $\hat{\sigma}_{\hat{\xi}} = 0.33$. Another estimate of $\sigma_{\hat{\xi}}$ would use TFP or, rather, fixed effects in plant TFP.\footnote{If we do not correct for labor quality at all, then $\xi$ is TFP. If we fully correct for quality of all workers, then there is no difference in TFPs among firms.} In this case Table 1 Abraham and White (2006) suggest $\hat{\sigma}_{\hat{\xi}} = 0.50$. These are for a generation, i.e., for the duration of one half of a working life, say 23 years, the range is roughly between one half a percent and one percent per year. This is still very large. However, the derivative with respect to $r$ measures the effect of an increase from $r = 0$ (no correlation) to $r = 1$ (perfect correlation), which presumably is so large as to be infeasible. Thus the derivative should be multiplied by the fraction of the rise in $r$ that is considered feasible.

We mentioned that garbling hurts BGP growth the most when $\theta$ is in the middle range, and that we get an effect that, in terms of $\theta$ has an inverted-U shape. Let us illustrate by plotting (67) as a function of $\theta$ and $r$ when we set $A = b = 1$, and $\sigma_{\hat{\xi}} = 0.42$, the mid-point between the Heathcote \textit{et al.}, and the Abraham-White estimates mentioned above. What is interesting in Figure is not the level of growth—this is an Ak model and the average rate of growth is easily changed by changing the parameters $A$ or $b$. Rather, the interest is in the effect of $r$, and this effect is proportional to $\sigma_{\hat{\xi}}$ and \textit{de facto} to $\theta (1 - \theta)$, as (68) shows. As expected, the effect is strongest when $\theta = 0.5$, which is when the slope of the line is steepest. From the middle line, the effect is about 3.8 percent per generation which we may take to be 22 years in our OG model. So the annual effect on (balanced) growth can be at most $\frac{3.8}{22} = 0.17$, i.e. 17 basis points per year. Off the BGP the effect can be larger, as we show in Section 6.4.1

\textbf{6.3 TFP growth of firms}

Imperfect signals reduce the economy’s growth rate and the TFP growth rate of the average firm, and they also introduce turnover in the distribution of TFP across firms. By injecting a stochastic element into assignments and, hence, into firm growth, which means that there will be turnover in the distribution of firm TFPs and revenues, i.e., a reversion to trend. 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. Let us see why that is. Subtracting \(\hat{x}\) from both sides of (57), and using (61), we obtain the growth rate of firm \(x\):

\[
\hat{x}' - \hat{x} = A + \theta \left( \mu_{\hat{y}} - r \mu_{\hat{x}} \frac{\sigma_{\hat{x}}}{\sigma_{\hat{y}}} \right) - \theta \left( 1 - r \frac{\sigma_{\hat{x}}}{\sigma_{\hat{y}}} \right) \hat{x} + \theta \sqrt{1 - r^2 \sigma_{\hat{x}} \zeta} \tag{69}
\]

Remarks on (69):

1. Figure 2 shows that \(\frac{\sigma_{\hat{x}}}{\sigma_{\hat{y}}} > 1\) for all \(r < 1\), and on the BGP small firms grow faster than large firms. The figure also shows that \(r/\sqrt{u}\) is monotonically increasing in \(r^{14}\).

2. On the transition path, if the initial \(\sigma_{\hat{x}}\) is small relative to \(\sigma_{\hat{y}}\) so that \(\sqrt{u}\) is low, the relation between size and growth will be positive for some periods until \(\sqrt{u}\) reaches its BGP level in (61).

3. The larger are \(\theta\) and \(\sigma_{\hat{x}}\) 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.

\(^{14}\)Indeed, if one knew \(\theta\), one could infer \(r\) from the estimates in Evans (1987) or Hall (1987).
In the log-normal family, the state of the system, \( H \), is described by \((\mu_x, \sigma_x)\). Having described the evolution of \( \sigma_x \), we now discuss the evolution of \( \mu_x \). First, since 
\[ x = \exp\left(\mu_x + \frac{1}{2} \sigma_x^2\right) \]
so that
\[ \mu_y = \ln b + \mu_x + \mu_x + \frac{1}{2} \left(\sigma_x^2 + \sigma_z^2\right) \]
Substituting into (57) we get
\[ \hat{x}' = \hat{A} + \theta \left( \ln b + \mu_x + \mu_x + \frac{1}{2} \left(\sigma_x^2 + \sigma_z^2\right) - r \frac{\sigma_x^2}{\sigma} \mu_x \right) + (1 - \theta) \hat{x} + \theta r \frac{\sigma_x^2}{\sigma} \hat{x} + \theta \sqrt{1 - r^2} \sigma_x \zeta \]

The economy evolves deterministically due to the law of large numbers, a firm’s growth stochastic. The relation between growth and TFP depends on the position of the economy.

### 6.4 Garbling and inequality

When \( r \) declines, inequality shrinks, because positive assignment provides leverage to a high \( x \).\(^{15}\) In the models of Restucchia and Rogerson (2008) and Hsieh and Klenow (2009), the same result obtains in that an efficient firm will, when scale is variable, choose to operate at a higher scale and therefore raise its profits by more than if it would had its scale remained unchanged, and this force is attenuated when there is misallocation. As a theory of income inequality, this theory provides a rationale for the upward-sloping portion of the Kuznets curve, but not a downward portion. The evidence in favor of a rising inequality with development curve is mixed. Barro (2000) reports evidence that a negative effect of inequality on growth shows up for poor countries but that the relationship for rich countries is positive. The model implies a positive long-run relationship between growth and income inequality or perhaps TFP inequality. If the model can apply to industries, (51) could explain the evidence in Dwyer (1998, eq. 2) that dispersion in productivity is larger in industries with more rapid productivity growth. In the model, TFP varies only because labor quality does The estimates of Hsieh and Klenow (2009) calculate dispersion of TFP after correcting for labor quality by multiplying the labor input by the wage; thus they correct for the quality of labor by multiplying by the wage, and my model implies that the variance of TFP thus constructed should be zero.

In the transition, inequality can rise or fall, depending on where it is relative to its steady state value. The most reasonable time-series experiment is a sudden rise in \( r \) due, perhaps, to a policy reform that allows quality of the young to be better labelled and more appropriately rewarded. In that case, the economy would find its inequality and its development rising along the transition to the new steady state. In

\(^{15}\)A similar inequality result is also in models of marital sorting (Kremer 1997, Fernandez and Rogerson 2001), and evidence seems to support it.
In this sense, the model can generate a rising inequality in the time-series sense. Of course, we can then calculate the entire transitional dynamics based on the reform. That will need to be done numerically. For now, we can calculate analytically the impact effect of the reform.

### 6.4.1 Impact effect of a reform on output

Since gambling has growth effects, the long-run level effects of a rise in \( r \) are infinite. But the transitional effects are finite, and we can analytically calculate the following

1. Start two economies off in the same state: i.e., let them both have the same \( H \).
2. Raise \( r \) in one of the economies so that the labeling of the young generation improves
3. Calculate the ratio of outputs of the two economies and call that the “output gap” on impact

Log output is

\[
\hat{q} = (1 - \rho) \hat{x} + \rho \hat{y}.
\]

**Distribution of output conditional on \( x \).**—From (55), log output of firm \( x \) is distributed as \( N \left( \mu_{\hat{q}(x)}, \sigma_{\hat{q}(x)}^2 \right) \) where

\[
\begin{align*}
\mu_{\hat{q}(x)} & = (1 - \rho) \hat{x} + \rho \left( \mu_y + r \frac{\sigma_{\hat{z}}}{\sigma_{\hat{x}}} (\hat{x} - \mu_{\hat{x}}) \right) \\
\sigma_{\hat{q}(x)}^2 & = (1 - r^2) \rho^2 \sigma_{\hat{x}}^2
\end{align*}
\]

Output is unconditionally distributed as \( N \left( \mu_{\hat{q}}, \sigma_{\hat{q}}^2 \right) \), where

\[
\begin{align*}
\mu_{\hat{q}} & = (1 - \rho) \mu_{\hat{x}} + \rho \mu_{\hat{y}} \\
\sigma_{\hat{q}}^2 & = \left( 1 - \rho + \rho r \frac{\sigma_{\hat{z}}}{\sigma_{\hat{x}}} \right)^2 \sigma_{\hat{x}}^2 + (1 - r^2) \rho^2 \sigma_{\hat{z}}^2
\end{align*}
\]

In light of (56), we have

\[
\sigma_{\hat{q}}^2 = \left( 1 - \rho + \rho r \frac{\sigma_{\hat{z}}}{\sigma_{\hat{x}}} \right)^2 \sigma_{\hat{x}}^2 + \rho^2 (1 - r^2) \sigma_{\hat{z}}^2
\]

\[
= (1 - \rho)^2 \sigma_{\hat{x}}^2 + \rho^2 r^2 \sigma_{\hat{z}}^2 + 2 \rho (1 - \rho) r \sigma_{\hat{z}} \sigma_{\hat{x}} + \rho^2 \left( 1 - r^2 \right) \sigma_{\hat{z}}^2
\]

\[
= (1 - \rho)^2 \sigma_{\hat{x}}^2 + 2 \rho (1 - \rho) r \sigma_{\hat{z}} \sigma_{\hat{x}} + \rho^2 \sigma_{\hat{z}}^2
\]

---

\(^{16}\) Inequality may be bad for growth for other reasons, reasons not included in the model. Easterly (2004) finds some evidence for a causal negative effect of inequality on development.
Thus interchanging \((1 - \rho, \sigma_\bar{\xi})\) and \((\rho, \sigma_\bar{\xi})\) leaves the value of \(\sigma_\bar{\xi}^2\) unchanged.

Compare two economies starting with the same \(x \sim N (\mu_\bar{q}, \sigma_\bar{q}^2)\) but with a different \(r\), say \(r_1\) and \(r_2\). Then the ratio of the two economies’ aggregate outputs is

\[
\frac{Q_1}{Q_2} = \exp (\rho (1 - \rho) \sigma_\bar{\xi} \sigma_{\bar{\xi}} [r_1 - r_2])
\]

Aggregate output is

\[
Q^H \equiv \int f (x, y) d\tau (y \mid \alpha (x)) dH (x) = \int f (x, y) d\tau (y \mid \alpha (x)) dH (x)
\]

Consider the long-run gap between two economies with different values for \(r\). In the long run, (27) holds, and it states that \(\sigma_\bar{\xi} = r \sigma_{\bar{\xi}}\). Substituting into (70), the steady-state output ratio is

\[
\lim_{t \to \infty} \frac{Q_{1,t}}{Q_{2,t}} = \exp (\rho (1 - \rho) \sigma_{\bar{\xi}}^2 [r_1^2 - r_2^2]).
\]

Figure 4 plots this as a function of \(r^2\) for various values of the parameter combination \(\rho (1 - \rho) \sigma_{\bar{\xi}}^2\), assuming that one of the economies has perfect assignment:

**Development gap relative to leader.**—Suppose the leader has \(r = 1\), and that a follower starts with \(r = 0\). Then (27) and Remark 3 below it imply that in the leader’s economy, \(\sigma_x = \sigma_{\bar{\xi}}\). Now, according to (32) we may use the standard deviation of log wages to estimate \(\sigma_{\bar{\xi}}\). Depending on how TFP is measured, we may we can infer \(\sigma\) from \(\sigma_{\bar{\xi}}\). We once again use the two measures we used in Section 5.2, summarized in the following table:

<table>
<thead>
<tr>
<th>ineq. measure</th>
<th>(\hat{\sigma}_{\bar{\xi}})</th>
<th>max. impact output gap = (1 - e^{-\sigma_{\bar{\xi}}^2/4})</th>
</tr>
</thead>
<tbody>
<tr>
<td>wages(^{17})</td>
<td>0.33</td>
<td>0.08</td>
</tr>
<tr>
<td>TFP fixed effects(^{18})</td>
<td>0.50</td>
<td>0.12</td>
</tr>
</tbody>
</table>

We shall use these numbers to estimate the impact effect of a change in \(r\) from \(r\) to unity, which is

\[
1 - \exp \left( - \left(1 - r^2\right) \rho (1 - \rho) \sigma_{\bar{\xi}}^2 \right). \tag{72}
\]

Since \(\rho (1 - \rho) \leq 1/4\), the largest gap that the model can explain is \(1 - \exp \left(- \frac{1}{4} \sigma_{\bar{\xi}}^2\right)\)

\[
0.05 \leq 1 - \exp \left(- \frac{1}{4} \sigma_{\bar{\xi}}^2\right) \approx 0.08 \leq 0.11.
\]

which is based on the inequalities \(0.2 \leq \sigma_{\bar{\xi}}^2 \approx 0.33 \leq 0.47\) taken from Heathcote et al. (2005) where in Fig. 1A they report the variance of log earnings at various ages.

\(^{17}\)Source: Heathcote et al. (2004) Fig. 3

\(^{18}\)Source: Table 1 Abraham and White (2006) We take the estimate of the standard deviation of log(\(\bar{\xi}\)) from the unweighted whole sample.
The top line in Fig. 4 assumes that $1/2$ and that $\sigma^2_\epsilon = 0.33$, the latter being the most reasonable estimate of this parameter based on the data in the paper by Heathcote et al.. The bottom line is the slightly larger effect based on the estimates of Abraham and White (2006). This is a transitional effect of roughly 10 percent per generation, or about $10 \times 0.45 = 4.5$ basis points per year, slightly higher than the BGP estimate of 17 basis points per year.

The exercise in Figure 4 assumes that the economies being compared are closed and that $H$ refers to the distribution of $x$ in a given economy. An extension would regard the spillover in (20) as including an inflow of knowledge from abroad which would seem reasonable. Such spillovers would probably reduce the model’s ability to explain world inequality. At some level the model would be suited for explaining within-country as well as cross-country inequality as in documented in Sala-i-Martin (2006).

### 6.5 Wages and profits on the transition path

The results of this section so far concern only the evolution of $H(\cdot)$. It remains to solve for wages and profits so that we can talk about income inequality. We continue to assume that the initial $H_0$ is log-normal – thus, only a subset of the full dynamics will be analyzed in which every member of the sequence $(H_t)_{t=0}^\infty$ is log normal. The
personal distribution and the share of labor will change along the transition path. It turns out that as long as \( \rho < 1 \), wages no longer are a linear function of output, even if we were to let the slope coefficients depend on the economy’s state \( H_t \) as summarized in the log-normal case by \((\mu_t, \sigma_t)\). To keep track of dates, we shall introduce time subscripts.

Since the young does not know his type, his equilibrium lifetime utility is now an expectation over possible \( y \) realizations; they affect not his wage because that is paid up front, but his old-age profits. His lifetime utility given his revealed quality \( s \) and given the date-\( t \) assignment \( s = \alpha_t(x) \) is

\[
V_t(s) = w_t\left(\alpha_t^{-1}(s)\right) + \beta \int \pi_{t+1}\left(\phi\left[\alpha_t^{-1}(s), y\right]\right) d\tau(y \mid s) \tag{73}
\]

**Firm \( x \)'s decision problem.**

\[
\pi_t(x) = \max_{w,*} \left\{ \int f(x,y) d\tau(y \mid s) - w \right\} \tag{74}
\]

s.t.

\[
w + \beta \int \pi_{t+1}\left(\phi(x,y)\right) d\tau(y \mid s) \geq V_t(s) \tag{75}
\]

leading to the consolidated problem

\[
\pi_t(x) = \max_s \left\{ -V_t(s) + \int [f(x,y) + \beta \pi_{t+1}(\phi(x,y))] d\tau(y \mid s) \right\} \tag{76}
\]

### 6.5.1 Existence of equilibrium wages and profits

Let us prove that there exist equilibrium prices \( w_t(x) \) that support the transitional dynamics. We shall define this equilibrium as follows. First, (74) implies that \( \pi_t(x) = \int f(x,y) d\tau(y \mid s) - w_t(\alpha_t[x]) \), so that \((\alpha_t, \pi)\) imply \( w \).

**Equilibrium.**—It consists of three functions \( \alpha_t(x), \pi_t(x), \) and \( V_t(s) \) satisfying (73), (76), and (54) which we need to re-write more explicitly:

\[
\hat{s} = \hat{\alpha}_t(x) = \mu_{\hat{s},t} + \frac{\sigma_{\hat{s}}}{\sigma_t}(\hat{x} - \mu_t), \tag{77}
\]

where, by (59) and (60), \((\mu_t, \sigma_t)\) satisfy the difference equations

\[
\mu_{t+1} = \mu_t + \hat{A} + \theta \left(\ln b + \mu_{\hat{z}} + \frac{1}{2}\sigma_{\hat{z}}^2\right), \quad \text{and} \tag{78}
\]

\[
\sigma_{t+1} = \sqrt{(1 - \theta)^2 \sigma_{\hat{z}}^2 + 2\theta (1 - \theta) r \sigma_{\hat{z}} \sigma_t + \theta^2 \sigma_t^2} \tag{79}
\]

with \((\mu_0, \sigma_0)\) given and where, by (20), since \( \mu_{\hat{s},t} = E_t(\ln y) \), since \( \ln y = \ln b + \ln \bar{x} + \ln \varepsilon \), since \( E_t(\ln \bar{x}) = \mu_t + \frac{1}{2}\sigma_t^2 \), and since \( \mu_{\varepsilon} = 0 \),

\[
\mu_{\hat{s},t} = \ln b + \mu_t + \frac{1}{2}\sigma_t^2 \tag{80}
\]
Algorithm. — Substituting into (77)

\[ \hat{\alpha}_t (x) = \ln b + \left( 1 - \frac{\sigma_s}{\sigma_t} \right) \mu_t + \frac{1}{2} \sigma_t^2 + \frac{\sigma_s}{\sigma_t} \hat{x}, \]

Now, Assuming that \( w \) is differentiable, the FOC reads

\[ V_t' (\alpha [x]) = \frac{\partial}{\partial s} \int [f (x, y) + \beta \pi_{t+1} (\phi [x, y])] d\tau (y \mid s) \bigg|_{s=\alpha(x)} \quad (81) \]

At the equilibrium assignment, (76) holds as an accounting identity, and the wage is to be solved using (81). Now (73) implies that

\[ \pi_0 (x) = \frac{\partial}{\partial x} \left( \frac{\partial}{\partial s} \int \pi_{t+1} (\phi [x, y]) d\tau (y \mid s) \bigg|_{s=\alpha(x)} \right) \frac{d\alpha_t^{-1}}{ds} \bigg|_{s=\alpha(x)} \]

Now \( \frac{d\alpha_t^{-1}}{ds} \bigg|_{s=\alpha(x)} = \frac{1}{\alpha_t (x)} \). Substituting this into (82) and combining the result with (81), we obtain the optimality condition

\[ w_t' (x) = \frac{\partial}{\partial s} \int f (x, y) d\tau (y \mid s) \bigg|_{s=\alpha(x)} \alpha_t' (x) - \frac{\partial}{\partial x} \beta \int \pi_{t+1} (\phi [x, y]) d\tau (y \mid s) \bigg|_{s=\alpha(x)} \]

Finally using the accounting relation

\[ \pi_t (x) = \int f (x, y) d\tau (y \mid \alpha_t (x)) - w_t (x), \]

updated to \( t + 1 \), we eliminate \( \pi_{t+1} \) from (83) so that it becomes

\[ w_t' (x) = \frac{\partial}{\partial s} \int f (x, y) d\tau (y \mid s) \bigg|_{s=\alpha(x)} \alpha_t' (x) - \frac{\partial}{\partial x} \beta \int \left[ \int f (\phi [x, y], y') d\tau (y' \mid \alpha_{t+1} (\phi [x, y])) - w_{t+1} (\phi [x, y]) \right] d\tau (y \mid s) \bigg|_{s=\alpha(x)} \]

This is the Bellman equation to be solved for \( w_t' (x) \). It can be written more explicitly when the derivative w.r.t. \( x \) is calculated:

\[ w_t' (x) = \alpha_t' (x) \frac{\partial}{\partial s} \int f (x, y) d\tau - \frac{\partial}{\partial x} \beta \int \left( \int f (\phi [x, y], y') d\tau \mid y' \mid \alpha_{t+1} (\phi [x, y]) \right) d\tau + \beta \int w_{t+1}' (\phi [x, y]) \frac{\partial \phi (x, y)}{\partial x} d\tau \quad (84) \]

where we use the shorthand notation \( d\tau = d\tau (y \mid s) \bigg|_{s=\alpha_t (x)} \).
Next, let us make (84) more explicit. We shall show that the RHS of (84) is the sum of the terms given in (89), (91), and (92). First, from (54),

$$\alpha_t (x) = e^{\mu_t - \frac{\sigma_t}{\sigma_t} \mu_t^t x^t} \quad \text{and} \quad \alpha_t^{-1} (s) = e^{\mu_t - \frac{\sigma_t}{\sigma_t} \mu_t^t s^t},$$

so that

$$\alpha_t' (x) = \frac{\sigma_t s}{\sigma_t x}.$$  

(86)

We maintain the Cobb-Douglas forms $f = x^{1-\rho} y^\rho$ and $\phi = Ax^{1-\theta} y^\theta$. From the first line of (55),

$$E (y^\rho | s) = e^{\rho (1-r^2) \mu_y + \frac{\sigma_y^2}{2} (1-r^2) s^2 r^2 \rho^2}.$$  

(87)

Assignment is given by (54) which implies (55). Then the first integral in (84) is

$$\int f (x, y) d\tau (y | s) = x^{1-\rho} E_t (y^\rho | s) = e^{\rho (1-r^2) \mu_y + \frac{\sigma_y^2}{2} (1-r^2) x^t x^{-\rho} s^2 r^2 \rho^2} = \hat{D} x^{1-\rho} s^2 r^2 \rho$$

where

$$\hat{D} = e^{\rho (1-r^2) \mu_y + \frac{\sigma_y^2}{2} (1-r^2) s^2 r^2 \rho}.$$  

(88)

Since $\frac{\sigma_t}{\sigma_t} \int f (x, y) d\tau = r^2 \rho \hat{D} x^{1-\rho} s^2 r^2 \rho - 1$, the first term on the RHS of (84) is

$$\alpha_t' (x) \frac{\partial}{\partial s} \int f (x, y) d\tau = \frac{\sigma_t s}{\sigma_t} r^2 \rho \hat{D} x^{1-\rho} s^2 r^2 \rho - 1$$

$$= \frac{\sigma_t s}{\sigma_t} r^2 \rho \hat{D} x^{1-\rho} s^2 r^2 \rho = \frac{\sigma_t s}{\sigma_t} r^2 \rho \hat{D} x^{1-\rho} \left( e^{\mu_t - \frac{\sigma_t}{\sigma_t^t} \mu_t^t x^t} \right)^{\rho^2}$$

$$= \frac{\sigma_t s}{\sigma_t} r^2 \rho \hat{D} e^{\rho^2 (\mu_t - \frac{\sigma_t}{\sigma_t^t} \mu_t^t) x^t (r^2 \sigma_t^2 - 1)}$$

(89)

where the third equality follows from a substitution for $s$ from (85).

Now the second term on the RHS of (84) is $\int f (\phi [x, y], y') d\tau [y' | \alpha_{t+1} (\phi [x, y])] d\tau (y | s)$. The term

$$E_t (q_{t+1} | x, s)$$

$$= E_t \{ x_{t+1} \} E_t (y_{t+1}^t | x_{t+1}) | x, s \}$$

$$= E_t \{ \phi (x, y_t)^{1-\rho} E_t (y_{t+1}^t | \phi [x, y_t]) | x, s \}$$

$$= B_{t+1} E_t \{ \phi (x, y_t) \}^{\alpha_{t+1}} | x, s \} = A^{a_{t+1}} B_{t+1} x^{(1-\theta) a_{t+1}} E (\gamma_{t+1}^a | s)$$

$$= A^{a_{t+1}} B_{t+1} e^{\theta a_{t+1} [1-r^2] \mu_y + \frac{\sigma_y^2}{2} (1-r^2) s^2 x^{(1-\theta) a_{t+1}} s^2 r^2 a_{t+1}} = C_{t+1} x^{(1-\theta) a_{t+1}} s^2 r^2 a_{t+1}$$

where the last line follows if one replaces $\rho$ by $\theta a_{t+1}$ in (87), where

$$a_t = 1 - \rho + \rho r^2 \frac{\sigma_t}{\sigma_t} \quad \text{and} \quad B_t = \exp \left\{ \rho \mu_y + \rho r^2 \frac{\sigma_t}{\sigma_t} (x - \mu_t) + r^2 (1-r^2) \sigma_t^2 \right\}$$

(90)
Therefore the second term on the RHS of (84) is

$$\frac{\partial}{\partial x} \beta \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\phi, y') d\tau (y') d\tau (y) = (1 - \theta) a_{t+1} C_{t+1} (1^{(1-\theta)a_t+1 \theta^2 a_t+1} (1^{(1-\theta)a_t+1 \theta^2 a_t+1}

$$

where

$$C_{t+1} = (1 - \theta) a_{t+1} C_{t+1} \left( e^{\mu_t - \frac{\sigma_t^2}{2}} \right)^{\theta^2 a_t+1} \left( 1^{(1-\theta)(1-\theta)^2 a_t+1 \theta^2 a_t+1} \right)$$

Lastly, the third term on the RHS of (84) is

$$\beta (1 - \theta) A x^{\theta} \int_{-\infty}^{\infty} w'_{t+1} \left( A x^{1-\theta} y^{\theta} \right) y^{\theta} d\tau (y | \alpha_t [x]),$$

where, from the third line of (55), we know that ln $y$ is normally distributed with mean $\mu_{\hat{y},t} + r \frac{\sigma_t}{\sigma_t} (\hat{x} - \mu_t)$ and variance $(1 - r^2) \sigma_t^2$, and therefore

$$d\tau (y | \alpha_t [x]) = \frac{1}{y} \frac{1}{\sqrt{2\pi (1 - r^2) \sigma_t^2}} \exp \left( \frac{(\ln y - \mu_{\hat{y},t} - r \frac{\sigma_t}{\sigma_t} [\hat{x} - \mu_t])^2}{2 (1 - r^2) \sigma_t^2} \right),$$

and so the third term on the RHS of (84) is

$$\beta (1 - \theta) A x^{\theta} \int_{-\infty}^{\infty} w'_{t+1} \left( A x^{1-\theta} y^{\theta} \right) y^{\theta-1} \exp \left( \frac{(\ln y - \mu_{\hat{y},t} - r \frac{\sigma_t}{\sigma_t} [\ln x - \mu_t])^2}{2 (1 - r^2) \sigma_t^2} \right) dy$$

(92)

Now, the last term shows we do not have a contraction of functions $w'$ in a normed space if we use the sup norm $\|w'\| \equiv \sup_{x,t} w'_t (x).$ \footnote{Denote by $T$ the operator on the RHS of (84), we see that since

$$\int_{-\infty}^{\infty} y^\theta d\tau (y | \alpha_t [x]) = \exp \left( \theta \left[ \mu_{\hat{y},t} + r \frac{\sigma_t}{\sigma_t} (\hat{x} - \mu_t) \right] + \frac{1}{2} \theta^2 (1 - r^2) \sigma_t^2 \right),$$

and since

$$\beta (1 - \theta) A x^{\theta} \int_{-\infty}^{\infty} y^\theta d\tau = \beta (1 - \theta) A \exp \left( \theta \left[ \mu_{\hat{y},t} + r \frac{\sigma_t}{\sigma_t} (\hat{x} - \mu_t) \right] + \frac{1}{2} \theta^2 (1 - r^2) \sigma_t^2 \right) x^{\theta (r \frac{\sigma_t}{\sigma_t} - 1)},$$

we see that

$$\|T w'\| \leq C \|w'\| \sup_{x,t} x^{\theta (r \frac{\sigma_t}{\sigma_t} - 1)} = +\infty,$$

for any constant $C > 0$, we pick because, since $\sigma_t$ is not constant, there will exist an $x, t$ pairs that lead to arbitrarily large $x^{\theta (r \frac{\sigma_t}{\sigma_t} - 1)}.$} We need a complete metric space
with a weaker mode of convergence.

7 What does $r$ depend on?

In reality, $r$ probably depends on private and public spending as well as the individual’s receptivity to such spending and his understanding of what it takes to translate his innate skills into an exam grade. The mix of the spending and its total amount probably vary among societies: Europe relies on public inputs, the U.S. on private inputs, etc. A discussion of policy that takes an information-theoretic view is Davies and MacDonald (1984). One would like to see a list of policies that may affect $r$ and maybe see if countries that use those policies do indeed grow faster. Instead, we shall just list a few models that feature learning about a worker’s ability and perhaps these will suggest the sorts of policies that.

When $y$ measures absolute advantage.—In the model, $y$ is general ability, and it is exogenous. But better signals lead to a better distribution of $x’$. In a related model, Kremer (1993, Sec. II) argues that signal garbling lowers investment in $y$. Since we found that $r$ raises output and growth, signal accuracy is a good thing here as well. But there are models that say the opposite: In Holmström (1999), a better signal about a worker’s quality can lower welfare by reducing that worker’s reputational concerns and causing an underprovision of working effort. Even in Spence (1974), removing information from the system leads to higher output, though not a Pareto improvement.

When $y$ measures comparative advantage.—Prescott and Visscher (1980) explain how, after hiring the worker, a firm would test him so as to learn about his comparative advantage; the signal helps the firm better allocate the worker to the task that he can do better. When he is being tested, the worker cannot perform other tasks. MacDonald (1982) also assumes that signals accumulate over time about advantage that is both comparative and absolute, but that learning does not require a sacrifice of working time.\footnote{Hanushek (2011), Lockwood and McCaffrey (2009), Cunha, Heckman and Schennach (2010), Del Boca, Flinn and Wiswall (2012), and Behrman, Todd and Wolpin (in progress) have, or will have, evidence on the form of $\phi(x, y)$, but perhaps less about $r$.}

8 Conclusion

We have derived a growth effect of misallocation, in a model where growth is endogenous and in which the BGP is stable from any log-normal starting state. The BGP growth effect was estimated to be roughly at most 17 basis points per year, but in the transition in can be more than twice that value.

Aggregate output and aggregate growth are both maximized by a positive assignment, and quality of the assignment depends on how accurate the signal is. The
accuracy of the signal is exogenous to the model, and a social planner could improve on the competitive allocation only if he could improve signals. Thus, factors are not misallocated in light of the information available, but only when compared to some allocations that would be attainable the information was better. Education policy can affect signal quality, and thinking more seriously about what determines signal accuracy is a natural next step.

References


38


[65] Spence, M. “Competitive Responses to Signals.” *JET* 1974


Figure 5: Figure 1A from Heathcote, Storesletten and Violante (2005)