HONG KONG INSTITUTE FOR MONETARY RESEARCH

THE WEAK LINK THEORY OF ECONOMIC DEVELOPMENT

Charles I. Jones

HKIMR Working Paper No.4/2007

February 2007

The Weak Link Theory of Economic Development

Charles I. Jones*
University of California, Berkeley
Hong Kong Institute for Monetary Research

February 2007

Abstract

Per capita income in the richest countries of the world exceeds that in the poorest countries by more than a factor of 50. What explains these enormous differences? This paper returns to an old idea in development economics and proposes that complementarity and linkages are at the heart of the explanation. Just as a chain is only as strong as its weakest link, problems at any point in a production chain can reduce output substantially if inputs enter production in a complementary fashion. This paper builds a model with complementary inputs and links across sectors and shows that it can easily generate 50-fold aggregate income differences from plausible distributions of productivity in the underlying sectors.

* I would like to thank Daron Acemoglu, Pol Antras, Bill Easterly, Luis Garicano, Chang Hsieh, Pete Klenow, Alwyn Young and seminar participants at Berkeley, the Chicago GSB, Toulouse, USC, and the World Bank for helpful comments. I am grateful to the Hong Kong Institute for Monetary Research for hosting me during the early stages of this research and to the Toulouse Network for Information Technology for financial support.

1. Introduction

By the end of the 20th century, per capita income in the United States was more than 50 times higher than per capita income in Ethiopia and the Democratic Republic of the Congo (Zaire). Dispersion across the 95th-5th percentiles of countries was more than a factor of 32. What explains these profound differences in incomes across countries?[1]

This paper develops a model in which complementarity and linkages are at the heart of the explanation. High productivity in a firm requires a high level of performance along a large number of dimensions. Textile producers require raw materials, knitting machines, a healthy and trained labor force, knowledge of how to produce, security, business licenses, transportation networks, electricity, etc. Macroeconomics often works with production functions that exhibit substantial substitutability between inputs, but at the level of the production process itself, it is not clear that such a high degree of substitutability is warranted. Without electricity or production knowledge or raw materials or security or business licenses, production of textiles – or any other good for that matter – is likely to be severely hindered.

Linkages between activities are also likely to be important. Low productivity in transportation reduces agricultural productivity. Irregular electricity supplies hinder manufacturing. Lack of clean water leads to poor health among students and teachers, leading to inadequate training and low output elsewhere in the economy. Bureaucratic bottlenecks in trade may limit imports of replacement parts and have widespread effects. This notion that linkages affect development dates back to Hirschman (1958).

The metaphor that works best to describe this paper is the old adage, "A chain is only as strong as its weakest link". Complementarity and linkages in the economy mean that problems at any point in the production chain can sharply reduce overall output. The strength of a typical link need not differ by a large amount between rich and poor countries. Instead, what differs is the strength of the weakest links.

In any production process, there are ten things that can go wrong that will sharply reduce the value of production. In rich countries, there are enough substitution possibilities that these things do not often go wrong. In poor countries, on the other hand, any one of several problems can doom a project.

The contribution of this paper is to build a model in which these ideas can be made precise. We show that complementarity and linkages amplify small differences across economies. With plausible average differences in productivity across countries, we are able to explain 50-fold differences in per capita income.

The spirit of this paper is close to the O-ring theory put forward by Kremer (1993), but the papers differ substantially in crucial ways. These differences will be discussed in detail below.

---

1 Recent work on this topic includes Romer (1994), Klenow and Rodriguez-Clare (1997), Prescott (1998), Hall and Jones (1999), Parente and Prescott (1999), Acemoglu, Johnson and Robinson (2001), Acemoglu and Johnson (2005), Klenow and Rodriguez-Clare (2005), and Manuelli and Seshadri (2005).

The approach taken in this paper can also be compared with the recent literature on political economy and institutions. This paper is more about mechanics: can we develop a plausible mechanism for getting a big multiplier, so that relatively modest distortions lead to large income differences? The modern institutions approach builds up from political economy. This is useful in explaining why the allocations in poor countries are inferior – for example, why investment rates in physical and human capital are so low – but the institutions approach ultimately still requires a large multiplier to explain income differences. As just one example, even if a political economy model explains observed differences in investment rates across countries, the model cannot explain 50-fold income differences if it is embedded in a neoclassical framework. The political economy approach explains why resources are misallocated; the approach here explains why misallocations lead to large income differences. Clearly, both steps are needed to understand development.

2. The Role of Complementarity

Standard models of production emphasize the substitutability of different in-puts. While substitution will play an important role in the model that follows, so will complementarity. Since this is less familiar, we begin by focusing our attention on complementary inputs.

For this purpose, it is helpful to begin with a simple example. Suppose you'd like to set up a factory in China to make socks. The overall success of this project requires success along a surprisingly large number of different dimensions. These different activities are complementary, so that inefficiencies on any one dimension can sharply reduce overall output.

As one example, the managers of the firm require knowledge of exactly how to manufacture socks. This kind of knowledge plays a central role in the endogenous growth literature following Romer (1990).

Second, the firm needs the basic inputs of production. These include cotton, silk, and polyester; the sock-knitting machines that spin these threads into socks; a competent, healthy, and motivated workforce; a factory building; electricity and other utilities; a means of transporting raw materials and finished goods throughout the factory, etc.

Apart from the physical production of socks, other activities are required to turn raw materials into revenue. The entire production process must be kept secure from theft or expropriation. The sock manufacturer must match with buyers, perhaps in foreign markets, and must find a way to deliver the socks to these buyers. Legal requirements must also be met, both domestically and in foreign markets. Firms must acquire the necessary licenses and regulatory approval for production and trade.

The point of this somewhat tedious enumeration is that production – even of something as simple as a pair of socks – involves a large number of necessary activities. If any of these activities are performed inefficiently, overall output can be reduced considerably. Without a reliable supply of electricity, the sock-making machines cannot be utilized efficiently. If workers are not adequately trained or are unhealthy because of contaminated water supplies, productivity will suffer. If export licenses are not in order, the socks may sit in a warehouse rather than being sold. If property is not secure, the socks may be stolen before they can reach the market.

2.1 Modeling Complementary Inputs

A natural way to model the complementarity of these activities is with a CES production function:

$$Y = \left(\int_0^1 z_i^\rho di \right)^{1/\rho}. \tag{1}$$

We use z_i to denote a firm's performance along the i^{th} dimension, and we assume there are a continuum of activities indexed on the unit interval that are necessary for production. In terms of our sock example, z_a could be the quality of the instructions the firm has for making socks. z_b could be number of sock-making machines, z_c might represent the extent to which the relevant licenses have been obtained, etc.

The elasticity of substitution among these activities is $1/(1-\rho)$, so the degree of complementarity is a parameter. With $\rho = 0$, the elasticity of substitution is one and the production function is Cobb-Douglas. But if $\rho < 0$, inputs are even more complementary and the elasticity of substitution is less than one.

With $\rho < 0$, all inputs are necessary. That is, if any of the z_i are zero, output is also driven to zero. More generally, complementarity puts extra "weight" on the activities in which the firm is least successful. This is easy to see in the limiting case where $\rho \to -\infty$; in this case, the CES function converges to the minimum function, so output is equal to the smallest of the z_i.

This intuition can be pushed further by noting that the CES combination in equation (1) is called the *power mean* of the underlying z_i in statistics. The power mean is just a generalized mean. For example, if ρ = 1, *Y* is the arithmetic mean of the z_i . If $\rho = 0$, output is the geometric mean (Cobb-Douglas). If $\rho = -1$, output is the harmonic mean, and if $\rho \to -\infty$, output is the minimum of the z_i. From a standard result in statistics, these means are a decreasing function of ρ. Economically, a stronger degree of complementarity puts more weight on the weakest links and reduces output.

The essence of the story pursued here is this. On average, rich countries like the United States are only a little bit better – maybe by a factor of two – than the poorest countries in their underlying productivity at performing the key activities of production. Because of complementarity, however, it is not the average that matters. Instead, a chain is only as strong as its weakest link. Poor countries are poor because very low productivity at one or more essential activities reduces overall output. In a sense that will be made more precise below, poor countries have a thicker lower tail in the distribution of productivities, and complementarity among activities inflates these differences in the lower tail.

2.2 Comparing to Kremer's O-Ring Approach

It is useful to compare this approach to the O-ring theory of income differences put forward by Kremer (1993). Superficially, the theories are similar, and the general story Kremer tells is helpful in understanding the current paper: the space shuttle Challenger and its seven-member crew are destroyed because of the failure of a single, inexpensive rubber seal.

This paper differs crucially, however, in terms of how the general idea gets implemented. In particular, Kremer's modeling approach assumes a large degree of increasing returns, which is difficult to justify.

To see this, recall that Kremer assumes there are *N* different tasks that must be completed for production to succeed. Suppose workers have a probability of success q for any task, and assume these probabilities are independent. Expected output is then given by $Q = q^N$. Suppose the richest countries are flawless in production, so $q^{rich} = 1$, while the poorest countries are successful in each task 50 percent of the time, so $q^{poor} = 1/2$. The ratio of incomes between rich and poor countries is therefore on the order of 2^N . If there are five different tasks in production, it is quite easy to explain a 32-fold difference in incomes across countries.[2]

A problem with this approach is that the O-ring logic implies complementarity, but it does not imply the huge degree of increasing returns assumed in Kremer's $Q = q^N$ formulation. For example, an alternative production function that is also perfectly consistent with the O-ring story is $Q = q_1^{1/N} q_2^{1/N} \cdot ... \cdot q_N^{1/N}$ — that is, a Cobb-Douglas combination of tasks with constant returns. Notice that the O-ring complementarity applies here as well: if any q_i is zero, then $Q = 0$ and the entire project fails. With symmetry so that $q_i = q$, this approach leads to $Q = q$, so that a 2-fold difference in success on each task only translates into a 2-fold difference in incomes across countries.

So while the O-ring story is quite appealing, Kremer's formulation relies on an arbitrary and exceedingly strong degree of increasing returns to get big income differences. The approach taken here is to drop the large increasing returns inherent in Kremer's formulation and to emphasize complementarity instead.

3. Setting up the Model

We now apply this basic discussion of complementarity to construct a theory of economic development.

3.1 The Economic Environment

A single final good in this economy is produced using a continuum of activities that enter in a complementary fashion, as discussed above:[3]

$$Y = \zeta \cdot \left(\int_0^1 Y_i^\rho di \right)^{1/\rho}, \quad \rho < 0. \tag{2}$$

[2] Of course, this also suggests a problem with Kremer's approach: how many different tasks are involved in production in modern economies? If it is 5, then the model predicts a 32-fold difference in income. But if it is 20 instead, then incomes should differ by a factor of 2^{20}, or more than a million.

[3] Becker and Murphy (1992) consider a production function that combines a continuum of tasks in a Leontief way to produce output. They use this setup to study the division of labor and argue that it is limited by problems in coordinating the efforts of specialized workers.

In this expression, Y_i denotes the activity inputs, and ζ is a constant that we will use to simplify some expressions later.[4]

Activities are themselves produced using a relatively standard Cobb-Douglas production function:

$$Y_i = A_i \left(K_i^\alpha H_i^{1-\alpha} \right)^{1-\sigma} X_i^\sigma, \tag{3}$$

where α and σ are both between zero and one. K_i and H_i are the amounts of physical capital and human capital used to produce activity i, and A_i is an exogenously-given productivity level. The novel term in this production specification is X_i , which denotes the quantity of intermediate goods used to produce activity i.

Before discussing the role of X_i , it is convenient to specify the three resource constraints that face this economy:

$$\int_0^1 K_i di = K, \tag{4}$$

$$\int_0^1 H_i di = H, \tag{5}$$

and

$$C + \int_0^1 X_i di = Y. \tag{6}$$

The first two constraints are straightforward. We assume the economy is endowed with an exogenous amount of physical capital, K, and human capital, H, that can be used in production. Later on, we will endogenize K and H in standard ways, but it is convenient to take them as exogenous for now.

The last resource constraint says that final output can be used for consumption, C, or for the X_i intermediate goods. One unit of the final good can be used as one unit of the intermediate input in any activity.[5]

One can think of this as follows. Consider the production of the i^{th} activity Y_i, which we might take to be transportation services. Transportation is produced using physical capital, human capital, and some intermediate goods from other sectors (such as fuel). The share of intermediate goods in the production of the i^{th} activity is σ. To keep the model simple and tractable, we assume that the same bundle of intermediate goods are used in each activity, and that these intermediate goods are just units of final output.

4 In particular, we assume $\zeta = \sigma^{-\sigma}$, where σ will be defined below.

5 An issue of timing arises here. To keep the model simple and because we are concerned with the long run, we make the assumption that intermediate goods are produced and used simultaneously. We could, of course, incorporate a lag so that today's final good is used as tomorrow's intermediate input; the steady state of this setup would then deliver the result we have here.

The parameter σ, then, measures the importance of *linkages* in our economy. If $\sigma = 0$, the productivity of physical and human capital in each activity depends only on A_i and is independent of the rest of the economy. To the extent that $\sigma > 0$, low productivity in one activity feeds back into the others. Transportation services may be unproductive in a poor country because of inadequate fuel supplies or repair services. Low productivity in the telecommunications sector reduces productivity throughout the economy.

3.2 Substitution and Complementarity

Now is also a good time to note the roles played by substitution and complementarity in the model. The underlying activities, Y_i, are produced using Cobb-Douglas production functions. There is an exogenous component A_i to how good a country is at some activity, but defects in productivity can partly be offset by allocating high quantities of physical and human capital and intermediate goods to an activity.

To see this, suppose we let $\rho \to -\infty$ so that final output is equal to the minimum over the activities. In this case, the optimal allocation would place large amounts of K_i, H_i, and X_i in the activities with the lowest productivity, in an effort to offset this drawback. The ability to substitute inputs like capital and labor to offset low productivity helps to mitigate the extent to which complementarity reduces output. Weak links can be strengthened. This feature plays an important role in the results that follow.

3.3 Specifying Exogenous Productivity

At the moment, we have a continuum of exogenous productivity levels, A_i. In principle, we could refrain from specifying anything else about these productivity levels; indeed, some of the subsequent results will take this form.

However, for the purpose of quantifying the predictions of the model, it is helpful to parameterize this continuum more parsimoniously. This should be viewed as a convenient simplifying device rather than as something fundamental in the model. In particular, we assume the A_i are distributed independently according to a Weibull distribution. That is,

$$\Pr\left[A_i \leq a\right] \equiv F(a) = 1 - e^{-(a/\beta)^{\theta}}. \tag{7}$$

The mean of this distribution is $\beta\Gamma(1 + 1/\theta)$, where $\Gamma(\cdot)$ is Euler's factorial function (which will be discussed in more detail below). The Weibull distribution is chosen because it is very flexible and yet can be transformed and integrated up in nice ways.

When we turn to the quantitative analysis of the model, we will assume the parameters of this distribution – β and θ – are allowed to differ across countries. Figure 1 shows an example.

As a rough rule of thumb, one can think of β as determining the mean and θ as determining the thickness of the lower tail. For example, if $\theta = 1$, the Weibull distribution is an exponential distribution, and therefore has lots of mass in the lower tail. For $\theta > 1$, the Weibull looks sort of like a log-normal distribution.

In Figure 1, the "rich" country has $\beta = 1.93$ and $\theta = 5$, while the "poor" country has $\beta = 1$ and $\theta = 2$. The average value of productivity in the rich country works out to be twice that in the poor country, showing the role of β. The poor country has a thicker lower tail, as reflected in the θ parameters. Our two countries have different underlying productivities, but on average they are not that different. However, it is not the average that matters. Because of complementarity in production, bad draws from the distribution get magnified.

4. Allocating Resources and Solving

Taking the stocks of physical and human capital as given, we consider two alternative ways of allocating resources. The first is the optimal allocation of resources. In this case, weak links in the production chain get strengthened by substituting capital and intermediate goods for inferior productivity. Of course, in the poorest countries of the world, this strengthening may itself be imperfect. This leads us to consider an alternative allocation where resources are not allocated to strengthen weak links. These two allocations are defined in turn.

The optimal allocation of resources is the choice of $K_i,\ H_i$, and X_i that maximizes consumption:

DEFINITION 4.1 The *optimal allocation of resources* in this economy consists of values for the six endogenous variables $Y, C, \{Y_i, K_i, H_i, X_i\}$ that solve

$$\max_{\{X_i, K_i, H_i\}} C = Y - \int_0^1 X_i di$$

subject to

$$Y = \zeta \cdot \left(\int_0^1 Y_i^\rho di \right)^{1/\rho}$$

$$Y_i = A_i \left(K_i^\alpha H_i^{1-\alpha} \right)^{1-\sigma} X_i^\sigma$$

$$\int_0^1 K_i di = K$$

$$\int_0^1 H_i di = H$$

where the productivity levels A_i are given exogenously.[6]

What if weak links are not reinforced properly, i.e. inputs are not allocated optimally? Of course there are lots of ways that inefficiencies in allocation could arise. We will consider a simple benchmark case where intermediate goods, physical capital, and human capital are allocated equally across activities: weak links are not reinforced, but neither do highly productive sectors get a disproportionate share of resources. This case should give some idea of how misallocating resources can affect productivity.

6 In terms of counting equations and unknowns, notice that we get three sets of first order conditions from the maximization, and then we have three main equations determining consumption, output, and the activities. The last two resource constraint equations do not really count as they give a single restriction, but there are a continuum of capital allocations to be chosen.

DEFINITION 4.2 The *symmetric misallocation of resources* in this economy has $K_i = K$, $H_i = H$, $X_i = X$, and $X = \bar{s}Y$, where $0 < \bar{s} < 1$. Moreover, we assume $\bar{s} = \sigma$, which turns out to be the optimal share of output to use as intermediate goods. Y and Y_i are then determined from the production functions in (2) and (3).

We now turn to solving the model using these alternative schemes for allocating resources.

4.1 Solving for the Optimal Allocation

The solution for the optimal allocation is relatively straightforward, and we solve in two steps. First we obtain the optimal allocation of the intermediate goods X_i , and then we solve for the optimal allocations of physical and human capital. We report the solution in a series of propositions, not because the results are especially deep, but because this helps organize the algebra in a useful way, both for presentation and for readers who wish to solve the model themselves. (Outlines of the proofs are in Appendix B.)

PROPOSITION 4.1 *(The Optimal Allocation of X_i.) Let $a_i \equiv A_i(K_i^\alpha H_i^{1-\alpha})^{1-\sigma}$. If the intermediate goods X_i are allocated optimally, then output is given by*

$$Y = \left(\int_0^1 a_i^\lambda di\right)^{\frac{1}{\lambda}\cdot\frac{1}{1-\sigma}}, \tag{8}$$

where $\lambda \equiv \frac{\rho}{1-\rho\sigma}$.

This first proposition is really just a midpoint into our solution. However, it reveals two useful insights that will be reinforced later. First, notice the similarity of the output equation to the CES function we began our discussion of complementarity with, equation (1). Output is a CES combination of the input-adjusted productivities, a_i . Moreover, the curvature parameter in this CES function is no longer ρ but rather $\lambda \equiv \frac{\rho}{1-\rho\sigma}$. The optimal allocation of intermediate goods can partially offset low productivities, and this shows up as an increase in the effective elasticity of substitution.

The second insight is also important. In particular, the entire CES combination gets raised to the power $\frac{1}{1-\sigma} > 1$. That is, the presence of intermediate inputs in our setup generates a multiplier. This is very similar to the multiplier that emerges because of capital accumulation in a standard neoclassical growth model, where the term $\frac{1}{1-\alpha}$ appears frequently. We will discuss this multiplier in more detail once we derive the final solution, in the next proposition.

PROPOSITION 4.2 *(The Optimal Allocation of K_i and H_i.) When physical capital and human capital are also allocated optimally across activities, total production of the final good is given by*

$$Y = Q^{\frac{1}{1-\sigma}} K^\alpha H^{1-\alpha}, \tag{9}$$

where

$$Q \equiv \left(\int_0^1 A_i^{\frac{\rho}{1-\rho}} di\right)^{\frac{1-\rho}{\rho}}. \tag{10}$$

According to this proposition, our model delivers an elegant expression for aggregate output. Y is the familiar Cobb-Douglas combination of aggregate physical and human capital that exhibits constant returns to scale. Total factor productivity is a CES combination of the productivities of the individual activities. This combination features the same two properties documented in the previous proposition, which we now discuss in more detail.

First, the elasticity of substitution that matters for total factor productivity depends on the curvature parameter $\frac{\rho}{1-\rho}$ rather than just ρ itself. Notice that if the domain of ρ is $[0, -\infty)$, the domain of $\frac{\rho}{1-\rho}$ is $[0, -1)$, which means there is less complementarity in determining Q than there was in the original CES combination of activities. The reason is that the optimal allocation strengthens weak leaks by allocating more resources to activities with low productivity. Mathematically, this raises the effective elasticity of substitution that matters for output.

The second property of this solution worth noting is the multiplier associated with intermediate goods. Total factor productivity is equal to the CES combination of underlying productivities raised to the power $\frac{1}{1-\sigma}$. A simple example should make the reason for this transparent. Suppose $Y = aX^{\sigma}$ and $X = sY$; output depends in part on intermediate goods, and the intermediate goods are themselves produced using output. Solving these two equations gives $Y = a^{1/1-\sigma}s^{\sigma/1-\sigma}$, which is a simplified version of what is going on in our model.

The economic intuition for this multiplier is also straightforward. Low productivity in electric power generation reduces productivity in transportation services. But this reduces the ease with which the electricity industry can obtain new power-generating equipment and therefore further reduces output in electric power generation. Linkages between sectors within the economy generate an additional multiplier through which productivity problems get amplified.

4.2 Solving with Misallocation

When resources are (mis)allocated according to the symmetric misallocation defined above, the solution is even more straightforward. Notice that $Y_i = A_i m$, where $m = (K^{\alpha}H^{1-\alpha})^{1-\sigma}X^{\sigma}$ is constant across activities. Therefore final output just depends on the CES combination of the A_i with curvature parameter ρ, as stated in the following proposition:

PROPOSITION 4.3 *(The Symmetric Misallocation.) Under the symmetric misallocation of resources, total production of the final good is given by*

$$Y = Q_m^{\frac{1}{1-\sigma}} K^{\alpha} H^{1-\alpha}, \tag{11}$$

where

$$Q_m \equiv \left(\int_0^1 A_i^{\rho} di \right)^{\frac{1}{\rho}}. \tag{12}$$

It is useful to compare this result with the previous proposition. The aggregate production function takes the same form. The only difference is that the curvature parameter determining the productivity aggregate is the original ρ rather than the adjusted $\frac{\rho}{1-\rho}$. Weak links are not reinforced by allocating resources to unproductive activities, so the original complementarity among activities is not mitigated. This comparison can be illustrated with an example. Suppose $\rho \to -\infty$. In this case, the symmetric misallocation depends on the smallest of the A_i, but the optimal allocation depends on the harmonic mean of the productivities, since $\frac{\rho}{1-\rho} \to -1$. Weak links are even more crucial when they are not strengthened by the appropriate allocation of resources.

4.3 Evaluating TFP

The expressions for Q and Q_m above are nice, but it is not immediately obvious how to use them to judge TFP differences across countries. Our assumption that the A_i productivities are drawn from a Weibull distribution allows us to solve for Q as a function of the parameters of the distribution, leading to a more parsimonious expression. We do this now. Since this argument is less familiar than the algebra needed to understand the previous propositions, we go through the reasoning in more detail.

Let η represent the absolute value of the curvature parameter in determining the productivity aggregate Q or Q_m . For the optimal allocation of resources, $\eta = -\frac{\rho}{1-\rho}$, so that $\eta \in [0, 1)$ is a positive curvature parameter. For the symmetric misallocation, $\eta = -\rho \in [0, \infty)$. Also, define $z_i \equiv A_i^{-\eta}$. With a slight abuse of notation, the productivity aggregate can be written as

$$Q = \left(\int_0^1 A_i^{-\eta} di \right)^{-\frac{1}{\eta}} = \left(\int_0^1 z_i di \right)^{-\frac{1}{\eta}} . \tag{13}$$

Applying the law of large numbers to our model, Q can be viewed as the mean of the z_i across our continuum of sectors, raised to the power $-1/\eta$. To compute this mean, notice that

$$\begin{aligned}
\Pr[z_i \le z] &= \Pr[A_i^{-\eta} \le z] \\
&= \Pr[A_i \ge z^{-1/\eta}] \\
&= 1 - F(z^{-1/\eta}) \\
&= e^{-\left(\frac{1}{\beta} \cdot z^{-1/\eta}\right)^{\theta}} \\
&= e^{-(\beta^{\eta} z)^{-\theta/\eta}}.
\end{aligned}$$

This last expression is the cumulative distribution function for a Frechet random variable, which has a mean given by $\beta^{-\eta}\Gamma(1-\eta/\theta)$. This leads to the following proposition:

Proposition 4.4 *(The Solution for Q) If the underlying productivities A_i are distributed according to a Weibull distribution, as in equation (7), then the aggregate productivity term Q is given by*

$$Q^* = \beta \left(\Gamma(1 - \frac{\eta}{\theta}) \right)^{-1/\eta} \tag{14}$$

where $\Gamma(\cdot)$ is Euler's factorial function.

4.4 Quantifying the Mechanism

All of the ingredients we need to understand enormous differences in incomes across countries are now in place. We will conduct a full quantitative analysis of the model in a later section after we have endogenized physical and human capital. However, we pause now to show the complementarity and linkage mechanisms at work.

To begin, it is helpful to get more familiar with the $\Gamma(\cdot)$ function. The actual definition of the gamma function is not especially helpful for our purposes, but its properties are.[7] Some useful properties of this function are

$$\Gamma(1) = 1,$$

and

$$\Gamma(n+1) = n\Gamma(n), \tag{15}$$

so that $\Gamma(n+1) = n!$ if n is a positive integer. This is why the gamma function is sometimes referred to as Euler's extension of the factorial.

For our purposes, we are more concerned with the behavior of the factorial function for n between zero and one. To see what happens here, it is helpful to rewrite (15) as

$$\Gamma(n) = \frac{\Gamma(n+1)}{n}.$$

Because $\Gamma(1) = 1$, this expression tells us that $\Gamma(n)$ diverges to $+\infty$ as n falls to zero. With these properties in mind, Figure 2 shows the gamma function.

Now recall the solution for Q in equation (14): $Q = \beta\left(\Gamma(1 - \frac{\eta}{\theta})\right)^{-1/\eta}$. For our problem to yield an interior solution, we require the term inside the gamma function to be positive, which is equivalent to $\eta < \theta$; we will see below what happens if this condition is not met. Recall that $0 \leq \eta < 1$ for the optimal allocation of resources and $\theta > 0$, so there is plenty of room in the parameter space for this to occur. For the symmetric misallocation, we have only that $\eta > 0$, so there is more room for a corner solution.

We think of rich and poor countries as having different distributions of underlying productivity. In particular, a rich country may be expected to have a higher value of θ than a poor country, let's suppose, corresponding to a thinner lower tail.

---

7 For completeness, the gamma function is defined as

$$\Gamma(n) \equiv \int_0^\infty x^{n-1}e^{-x}dx$$

for $n > 0$.

If the poor country has a sufficiently low value of θ — that is, lots of mass at low productivity levels — then η/θ gets close to one, and $\Gamma(x) \to \Gamma(0) = +\infty$. Because Q depends on the inverse of the gamma function, this means that Q gets arbitrarily close to zero. This is the mathematics that allows us to explain large income differences.

What is the economics? A lower value of θ corresponds to a thicker lower tail, and a higher value of η corresponds to more complementarity in production. In this framework, the poorest countries of the world are poor because they have a number of weak links that play an important role because of complementarity. With the misallocation of resources, complementarity is even stronger because weak links are not reinforced. The parameter η is larger and this makes it more likely that the gamma function blows up.

A simple numerical example shows how this works. Suppose $\rho = -1$ and consider the optimal allocation of resources. In this case, $\eta \equiv -\frac{\rho}{1-\rho} = 1/2$. The basic elasticity of substitution between activities is $\frac{1}{1-\rho} = 1/2$, midway between Cobb-Douglas and Leontief. Allocating capital, labor, and intermediate inputs efficiently raises this elasticity of substitution to $\frac{1}{1+\eta} = 2/3$. Let's assume that a rich country is such that $Q_{rich} = 1$, and assume $\beta = 1$ for the poor country, so differences are driven by θ. Finally, let's take a share of intermediate goods in production of $\sigma = 1/2$. Table 1 shows the implied TFP differences in this example.

Before looking closely at the table, let us stipulate that to explain income differences of a factor of 32, one needs TFP differences of about 4 in a framework like this. This is the kind of number one gets from Klenow and Rodriguez-Clare (1997) or Hall and Jones (1999), and more details will be provided later.

Table 1 shows that enormous TFP differences can be obtained if the lower tail of the underlying productivity distribution is sufficiently thick, if activities are sufficiently complementary, and if linkages between sectors are sufficiently strong. This conclusion is especially true if resources are misallocated so that weak links are not strengthened. Notice that for the first two rows — corresponding to $\theta = 0.75$ and $\theta = 1$, the condition $\eta < \theta$ is violated. In this case, the lower tail of the distribution is so thick that the mean of z_i (approximately the inverse of A_i) does not exist, driving output in the poor country to zero.

5. Endogenizing K and H

The remainder of the paper proceeds in two steps. In this section, we enrich the model slightly by endogenizing a country's stocks of physical and human capital. The former gives us another multiplier in a familiar fashion, while the latter gives us another factor of 2. Both of these are useful in explaining large income differences across countries. The last main section of the paper will then turn to a full calibration exercise.

5.1 Endogenizing Physical Capital

We endogenize physical capital in a standard fashion. In particular, we assume that capital can be rented from the rest of the world at a constant and exogenous real rate of return, $\bar{r}$. This rate of return includes both the real interest rate and whatever country-specific distortions there are in the capital market. This parameter will therefore vary across countries.

The optimal allocation then hires capital until the marginal product of capital falls to equal this real rate of return (which includes depreciation). Given our Cobb-Douglas expression for output in equation (9), this condition is

$$\alpha \frac{Y^*}{K^*} = \bar{r}. \tag{16}$$

This equation implicitly determines the capital stock in a country.

5.2 Endogenizing Human Capital (Schooling)

We turn now to the human capital of the labor force, modeled as schooling. This is useful for two reasons. First, it allows us to present a very simple, tractable model of human capital that can be embedded in any theory of development. Second, it allows us to make additional quantitative predictions about the role of human capital in development. The specification below is closest to that in Mincer (1958). Richer models of human capital include Ben-Porath (1967), Bils and Klenow (2000), and Manuelli and Seshadri (2005). The approach here is purposefully stripped-down, trading generality and realism for simplicity and tractability.

Aggregate human capital H is labor in efficiency units: $H = hL$, where h is human capital per worker and L is the number of workers. Assume the (constant) population in a country is distributed exponentially by age and faces a constant death rate $\delta > 0$: the density is $f(a) = \delta e^{-\delta a}$. A person attending school for S years obtains human capital $h(S)$, a smooth increasing function. The representative individual's problem is to choose S to maximize the expected present discounted value of income:

$$\max_S \int_S^\infty w_t h(S) e^{-(\bar{r}+\delta)t} dt, \tag{17}$$

where the base wage w_t is assumed to grow exponentially at rate $\bar{g}$.

Solving this maximization problem leads to the Mincerian return equation:

$$\frac{h'(S^*)}{h(S^*)} = \tilde{r} \equiv \bar{r} - \bar{g} + \delta. \tag{18}$$

The left side of this equation is the standard Mincerian return: the percentage increase in the wage if schooling increases by a year. The first order condition says that the optimal choice of schooling equates the Mincerian return to the effective discount rate. In this case, the effective discount rate is the interest rate, adjusted for wage growth and the probability of death. The original Mincer (1958) specification pinned down the Mincerian return by the interest rate. The generalization here shows the additional role played by economic growth and limited horizons. Rather than being an exogenous parameter, as in the simple version of Bils and Klenow (2000) used by Hall and Jones (1999) and others, the Mincerian return in this specification is related to fundamental economic variables.

More progress can be made by assuming a functional form for $h(S)$. Consider the constant elasticity form $h(S) = S^{\phi}$. In this case, the Mincerian return is $h'(S)/h(S) = \phi/S$, so the Mincerian return falls as schooling rises. The first-order condition in equation (18) then implies the optimal choice for schooling is

$$S^* = \frac{\phi}{\bar{r} - \bar{g} + \delta}, \tag{19}$$

and the human capital of the labor force in efficiency units is

$$h^* = \left(\frac{\phi}{\bar{r} - \bar{g} + \delta}\right)^{\phi}. \tag{20}$$

We assume ϕ is the same across countries, so differences in schooling can be explained in this simple framework by differences in the effective discount rate. A higher interest rate, slower growth, and a higher death rate all translate into lower educational attainment.

People in this world go to school for the first S^* years of their lives and then work for the remainder of their lives. Anyone working has S^* years of schooling and therefore supplies h^* efficiency units of labor for production.

5.3 Solving the Extended Model

The Cobb-Douglas expression for output in equation (9) can be combined with the solutions for K^* and h^* in equations (16) and (20) to yield the following solution of the model:

PROPOSITION 5.1 *(The Solution for Y/L) In this weak link theory of economic development, output per worker when resources are allocated optimally is given by*

$$\begin{aligned} y^* \equiv \frac{Y^*}{L^*} &= Q^{*\frac{1}{1-\sigma}\frac{1}{1-\alpha}} \left(\frac{K^*}{Y^*}\right)^{\frac{\alpha}{1-\alpha}} h^* \\ &= Q^{*\frac{1}{1-\sigma}\frac{1}{1-\alpha}} \left(\frac{\alpha}{\bar{r}}\right)^{\frac{\alpha}{1-\alpha}} \left(\frac{\phi}{\bar{r} - \bar{g} + \delta}\right)^{\phi}. \end{aligned} \tag{21}$$

(22)

If resources are given by the symmetric misallocation, output per worker takes the same form, with Q_m^* replacing Q^*.

The wealth of nations is explained by two sets of parameters. Differences in $\bar{r}$ and the human capital parameters reflect the standard neoclassical forces. Now, however, we also have differences in TFP arising from complementary activities. The parameters β and θ reflect the differences in underlying productivities across countries.

The form of this solution should be familiar. Output per worker is determined by productivity, the cost of physical capital, and the factors that influence the accumulation of human capital. For the usual reasons, there is a $\frac{1}{1-\alpha}$ multiplier (exponent) associated with capital accumulation: anything that increases output leads to additional capital accumulation, which further increases output, etc. The sum of this geometric series is $\frac{1}{1-\alpha}$.

6. Quantitative Analysis

We now explore the model's quantitative predictions: can it help us to understand 50-fold differences in incomes across countries?

6.1 Calibration

There are five country-specific parameters in equation (21) that need to be calibrated: the Weibull distribution parameters β and θ, the interest rate $\bar{r}$, the growth rate $\bar{g}$, and the death rate δ. There are also four parameters that are assumed to be common across countries: the capital exponent α, the share of intermediate goods in production σ, the complementarity parameter ρ, and the schooling elasticity ϕ. Our benchmark values for all of these parameters are reported in Table 2.

Values for the parameters of the Weibull distribution have already been discussed briefly; these are the values used in the example shown back in Figure 1. The parameter choices imply that the mean of the distribution of managerial knowledge in the rich country is only twice that in the poor country, so in some average sense the countries do not look so different.

This factor of 2 difference in means largely pins down the β parameters. Pinning down the θ parameters that govern the thickness of the lower tail is harder. In the future, I plan to use some evidence on firm-level productivity from Hsieh and Klenow (2006) to shed some light on these parameters. For the moment, I choose the parameters so that (I hope) they are not too extreme. A section on robustness will attempt to exhaust the set of plausible values.[8]

The next set of parameters are related to schooling. We assume the interest rate for discounting future wages is 6% in the rich country and 12% in the poor country. Such values are well within the range of plausibility; see, for example, Caselli and Feyrer (2005). The parameter $\bar{r}$ plays two roles in the model, as the domestic cost of capital and the interest rate for discounting future wages. In theory, these interest rates could be determined by different forces. For example, the cost of capital could be higher because of capital taxation, while the (after tax) interest rate for discounting wages could be higher because of borrowing constraints. The 2-fold difference assumed here seems perfectly reasonable given the distortions to capital markets in Kenya or Ethiopia versus the United States.

---

8 The values of $\theta^r = 5$ and $\theta^p = 2$ are consistent with Hsieh and Klenow's evidence on the distribution of TFP *within* 4-digit sectors in China and India, and with Syverson's (2004) evidence for the United States. Of course, for the present purposes, it would be better to have data on the distribution of TFP *across* sectors, and that is what I will obtain in the near future.

We take a growth rate of 2% per year for the rich country and a growth rate of zero for the poor country. Many of the poorest countries of the world have exhibited essentially zero growth for the last 40 years.

For the death rate, we assume δ = 1% per year in the rich country and 2% per year in the poor country. With this constant probability of death, life expectancy is 50 years in the poor country and 100 years in the rich country.

These parameter values imply a Mincerian return to schooling of 5% in the rich country and 14% in the poor country. We also take ϕ = 0.6. Together with the other parameter values, this implies people in the rich country get 12 years of schooling, while people in the poor country get 4.3 years of schooling. These numbers are not a perfect match of the data (one might want a slighly smaller gap in the Mincerian returns and a slightly larger gap in the years of schooling, as documented by Bils and Klenow 2000), but they are certainly in the right ballpark, which is a nice accomplishment for the simple schooling framework used here.

The remaining parameters are common across countries. We pick $\alpha = 1/3$ to match the empirical evidence on capital shares; see Gollin (2002). For the share of intermediate goods in total output, Rotemberg and Woodford (1995) report an estimate of 0.5 for the U.S. economy since 1960. We tentatively take this value for σ, although it would be nice to have an estimate from a developing country as well — is the share lower there? Notice that this implies a substantial multiplier that works through intermediate goods: $\frac{1}{1-\sigma} = 2$.

The complementarity parameter is another parameter that is quite important but difficult to calibrate. Recall that we want ρ to be negative in the complementarity story. We take $\rho = -1$, which corresponds to an elasticity of substitution of 1/2, midway between Cobb-Douglas and Leontief. Once inputs are allocated optimally across sectors to reinforce weak links, this delivers a value for $\eta \equiv -\frac{\rho}{1-\rho} = 1/2$ and therefore an effective elasticity of substitution of $\frac{1}{1+\eta} = 2/3$. Obviously it is desirable to obtain better evidence on the extent of complementarity of activities in production. But given the stories we told to motivate this paper, this value of ρ does not seem extreme.

6.2 Results

To emphasize how this model explains differences in incomes between rich and poor countries, we evaluate the solution for output per worker in equation (21) for two countries and compute the ratio. Let the superscript r denote a rich country and the superscript p denote a poor country. Then income ratios are given by

$$\frac{y^{r*}}{y^{p*}} = \left[\underbrace{\left(\frac{\beta^r}{\beta^p} \cdot \left(\frac{\Gamma(1-\eta/\theta^p)}{\Gamma(1-\eta/\theta^r)} \right)^{1/\eta} \right)^{\frac{1}{1-\sigma}}}_{\text{TFP}} \underbrace{\left(\frac{\bar{r}^p}{\bar{r}^r} \right)^{\alpha}}_{\text{K/Y}} \underbrace{\left(\frac{\bar{r}^p - \bar{g}^p + \delta^p}{\bar{r}^r - \bar{g}^r + \delta^r} \right)^{\phi(1-\alpha)}}_{\text{h}} \right]^{\frac{1}{1-\alpha}}$$

Using the baseline parmeters from Table 2, the terms in this equation can quantified as follows. First, for the optimal allocation of resources:

$$\begin{aligned}\frac{y^{r*}}{y^{p*}} &\approx (6.44 \times 1.26 \times 1.51)^{1.5} \approx (12.3)^{1.5} \\ &\approx 16.4 \times 1.41 \times 1.85 \\ &\approx 42.9\end{aligned}$$

And next for the symmetric misallocation:

$$\begin{aligned}\frac{y^{r*}}{y^{p*}} &\approx (8.64 \times 1.26 \times 1.51)^{1.5} \approx (16.4)^{1.5} \\ &\approx 25.4 \times 1.41 \times 1.85 \\ &\approx 66.6\end{aligned}$$

The standard neoclassical terms for physical capital and schooling imply a difference in incomes of a factor of 1.41 X 1.85 = 2.6. This is smaller than the 4-fold difference between the 5 richest and 5 poorest countries documented by Hall and Jones (1999). With larger differences in $\bar{r}$, we could increase the difference in the model, but to be conservative, we keep these values.

Here, of course, we also have a theory of TFP differences, and the story goes as follows. Rich and poor countries on average are not that different (a factor of two, recall) in average productivity at various activities. However, these small differences get amplified in two distinct ways. First, activities enter production in a complementary fashion, so that problems in one area reduce the value of overall output. Second, intermediate goods provide linkages between activities. Low productivity in one activity leads to low productivity in the others.

The TFP differences in our calibration can be decomposed as follows. The basic factor of two is reflected in $\beta^r/\beta^p = 1.93$. This term is multiplied by the ratio of the gamma functions, reflecting complementarity. This ratio is 1.3 for the optimal allocation and 1.5 for the symmetric misallocation. The ratio of the Q productivity aggregates is then 1.93 X 1.3 = 2.5 for the optimal allocation and 1.93 X 1.5 = 2.9 for the symmetric misallocation. Because the intermediate goods share in production is 1/2, these numbers get squared in order to yield the basic TFP differences: 6.44 and 8.64. Capital accumulation provides further amplification, raising each of these numbers to the 3/2 power to yield 16.4 and 25.4.

The overall income difference predicted by this simple calibration is then the product of this TFP factor with the roughly 3-fold neoclassical effect. The model predicts differences between rich and poor countries of about 43 times for the optimal allocation and 67 times for the symmetric misallocation. These numbers can be compared to a 95th/5th percentile ratio for GDP per capita of 32.1 for the year 1999. The mechanism at work in this paper, then, seems to be perfectly capable of explaining the large income differences observed in the data.

6.3 Robustness

There are a number of parameter values in this quantitative exercise whose values we do not know especially well. In the future, I hope to gather more information about these parameters — for example, by using data from Hsieh and Klenow (in progress) on the distribution of total factor productivity at the firm level in rich and poor countries to help me calibrate the parameters of the Weibull distributions.

In the meantime, this section shows the robustness of the results to changes in the parameter values. In particular, we consider changing the complementarity parameter ρ, the Weibull distribution parameters θ_p and θ_r, and the share of intermediate goods in the economy, σ.

The results of these robustness checks are shown in Table 3. The first scenario simply repeats the baseline results, for comparison.

The second and third scenarios explore changes in the degree of complementarity in the economy. The baseline value for ρ is -1; we consider -1/2 and -2 as alternatives. Large income differences are clearly preserved by this change, and the differences explode to infinity at ρ = -2 for the case of misallocation.

The next four scenarios consider variations in the θ parameters, which govern the thickness of the lower tail of the Weibull distribution. Once again, large income differences are easily preserved for the range of values considered.

Finally, the last two rows show what happens when the share of intermediate goods in production is reduced. This parameter is more important, because of the nature of geometric sums. The case of σ = 0 is reported to show the extra power provided by the linkages channel in this paper.

It is worth noting that changes in these parameter values could interact in significant ways. For example, even if $\sigma = 1/4$ is eventually found to be the most appropriate share for intermediate goods, income differences of more than a factor of 22 may be observed if θ^p is less than 2.

6.4 Discussion

The model possesses two key features that seem desirable in any theory designed to explain the large differences in incomes across countries. First, relatively small and plausible differences in underlying parameters can yield large differences in incomes. That is, the model generates a large multiplier.

Second, improvements in underlying productivity along any single dimension have relatively small effects on output. If a chain has a number of weak links, fixing one or two of them will not change the overall strength of the chain.

This is important: if there were a single magic bullet for solving the world's development problems, one would expect that policy experimentation across countries would hit on it, at least eventually. The magic bullet would become well-known and the world's development problems would be solved. For example, this is a potential problem in the Manuelli and Seshadri (2005) paper: small subsidies to the production of output or small improvements in a single (exogenous) productivity level have enormous long-run effects on per capita income in their model.

Here, while it is true that small average differences in underlying productivity across countries lead to large differences in income, the development problem may be quite hard to solve in practice. Obtaining the instruction manual for how to produce socks is not especially useful if the import of knitting equipment is restricted, if cotton and polyester threads are not available, if property rights are not secure, and if the market to which these socks will be sold is unknown.

Multinational firms may help to solve these problems. For example, they may bring with them knowledge of how to produce, access to transportation and foreign markets, and the appropriate capital equipment. And yet domestic weak links may still be a problem. A lack of contract enforcement may make intermediate inputs and other activities hard to obtain. Weak property rights may lead to expropriation. Inadequate energy supplies may reduce productivity. Indeed many of the examples we know of where multinationals produce successfully in poor countries effectively give the multinational control on as many dimensions as possible: consider the maquiladoras of Mexico and the special economic zones in China and India.

The development problem is hard because there are 10 things that can go wrong in any production process. In the poorest countries of the world, productivity is low at many different stages, and complementarity means that reforms targetted at one or two problems have only modest effects. Linkages across sectors and the misallocation of resources provide additional amplification of inefficiencies.

7. Conclusion

This working paper provides the basics of a theory of economic development. I am continuing to develop these ideas. In particular, in the near future I plan to incorporate data on the distribution of productivity across firms and industries, in an effort to justify the calibration of the θ parameters. Also, the notion that there is a multiplier associated with intermediate goods will be explored more carefully in the coming months.

References

Acemoglu, Daron and Simon Johnson (2005), “Unbundling Institutions,” *Journal of Political Economy, forthcoming.*

__,__, and James A. Robinson (2001), “The Colonial Origins of Camparative Development: An Empirical Investigation,” *American Economic Review*, *91*(5):1369–401.

Aghion, Philippe, Peter Howitt and David Mayer-Foulkes (2005), “The Effect of Financial Development on Convergence: Theory and Evidence,” *Quarterly Journal of Economics,* 120(1):173–222.

Becker, Gary S. and Kevin M. Murphy (1992), “The Division of Labor, Coordination Costs, and Knowledge, ” *Quarterly Journal of Economics*, 107(4):1137–60.

Ben-Porath, Yoram (1967), “The Production of Human Capital and the Life Cycle of Earnings,” *Journal of Political Economy*, 75(4): 352–65.

Bils, Mark and Peter Klenow (2000), “Does Schooling Cause Growth?” *American Economic Review*, 90(5): 1160–83.

Caselli, Francesco and James Feyrer (2005), “The Marginal Product of Capital,” NBER Working Paper No.11551, Cambridge MA: National Bureau of Economic Research.

Gollin, Douglas (2002), “Getting Income Shares Right,” *Journal of Political Economy*, 110(2): 458–74.

Hall, Robert E. and Charles I. Jones (1999), “Why Do Some Countries Produce So Much More Output per Worker than Others?” *Quarterly Journal of Economics*, 114(1): 83–116.

Heston, Alan, Robert Summers and Bettina Aten (2002), “Penn World Table Version 6.1,” Center for International Comparisons at the University of Pennsylvania (CICUP).

Hirschman, Albert O. (1958), *The Strategy of Economic Development*, New Haven, CT: Yale University Press.

Hsieh, Chang-Tai and Peter J. Klenow (2006), “Misallocation and Manufacturing TFP in China and India,” University of California at Berkeley working paper.

Klenow, Peter J. and Andres Rodriguez-Clare (1997), “The Neoclassical Revival in Growth Economics: Has It Gone Too Far?” in Ben S. Bernanke and Julio J. Rotemberg, eds., *NBER Macroeconomics Annual 1997*, Cambridge, MA: MIT Press.

__and__ (2005), “Extenalities and Growth,” in Philippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, Amsterdam: Elsevier.

Kremer, Michael (1993), “The O-Ring Theory of Economic Development,” *Quarterly Journal of Economics,* 108(4): 551-76.

Maddison, Angus (2001), *The World Economy: A Millennial Perspective*, Paris: Organization for Economic Cooperation and Development.

Manuelli, Rodolfo and Ananth Seshadri (2005), “Human Capital and the Wealth of Nations,” University of Wisconsin working paper.

Mincer, Jacob (1958), “Investment in Human Capital and Personal Income Distribution,” *Journal of Political Economy*, 66(4): 281–302.

Parente, Stephen L. and Edward C. Prescott (1999), “Monopoly Rights: A Barrier to Riches,” *American Economic Review*, 89(5): 1216–33.

Prescott, Edward C. (1998), “Needed: A Theory of Total Factor Productivity,” *International Economic Review*, 39(3): 525–51.

Pritchett, Lant (1997), “Divergence: Big Time,” *Journal of Economic Perspectives*, 11(3): 3–17.

Romer, Paul M. (1990), “Endogenous Technological Change,” *Journal of Political Economy*, 98(5): S71–S102.

__ (1994), “New Goods, Old Theory, and the Welfare Costs of Trade Restrictions,” *Journal of Development Economics*, 43(1): 5–38.

Rotemberg, Julio J. and Michael Woodford (1995), “Dynamic General Equilibrium Models with Imperfectly Competitive Product Markets,” in Thomas J. Cooley, ed., *Frontiers of Business Cycle Research*, Princeton University Press.

Table 1. TFP Differences: A Numerical Example

	— Optimal Allocation —		— Symmetric Misallocation —	
θ	$\frac{Q^{rich}}{Q^{poor}}$	$\left(\frac{Q^{rich}}{Q^{poor}}\right)^{\frac{1}{1-\sigma}}$	$\frac{Q_m^{rich}}{Q_m^{poor}}$	$\left(\frac{Q_m^{rich}}{Q_m^{poor}}\right)^{\frac{1}{1-\sigma}}$
0.75	7.2	51.5	∞	∞
1.0	3.1	9.9	∞	∞
1.5	1.8	3.4	2.7	7.2
2.0	1.5	2.3	1.8	3.1
4.0	1.2	1.4	1.2	1.5

Note: This example assumes $\rho = -1$ so that $\eta = 1/2$ for the optimal allocation and $\eta = 1$ for the symmetric misallocation. Q for the rich country is taken to be 1.0, and Q for the poor country is calculated according to equation (14), with $\beta = 1$.

Table 2. Baseline Parameter Values

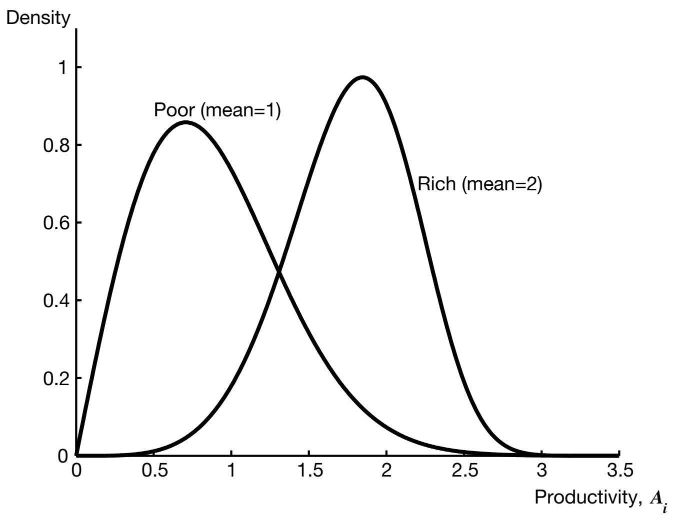
Parameter	Rich Country		Poor Country	Comments
β	1.93		1	Weibull location parameter
θ	5		2	Weibull curvature parameter
$\bar{r}$	.06		.12	Interest rate
$\bar{g}$	.02		0	Growth rate
δ	.01		.02	Death rate
α		1/3		Capital share
σ		1/2		Share of intermediate goods
ρ		-1		EofS=1/2 (midway)
ϕ		0.6		To match Mincerian returns

Table 3. Output per Worker Ratios: Robustness Results

Scenario	Change from Baseline	Optimal Allocation	Symmetric Misallocation
1	Baseline simulation	42.9	66.6
2	Less complementarity: $\rho = -1/2$	38.4	42.9
3	More complementarity: $\rho = -2$	48.7	∞
4	Thicker tail in poor country: $\theta^p = 1.5$	82.6	243.0
5	Thinner tail in poor country: $\theta^p = 3$	26.8	30.4
6	Thicker tail in rich country: $\theta^r = 4$	39.7	59.4
7	Thinner tail in rich country: $\theta^r = 8$	46.8	75.3
8	Reduced intermediate share: $\sigma = 1/4$	16.9	22.7
9	Zero intermediate share: $\sigma = 0$	10.6	13.2

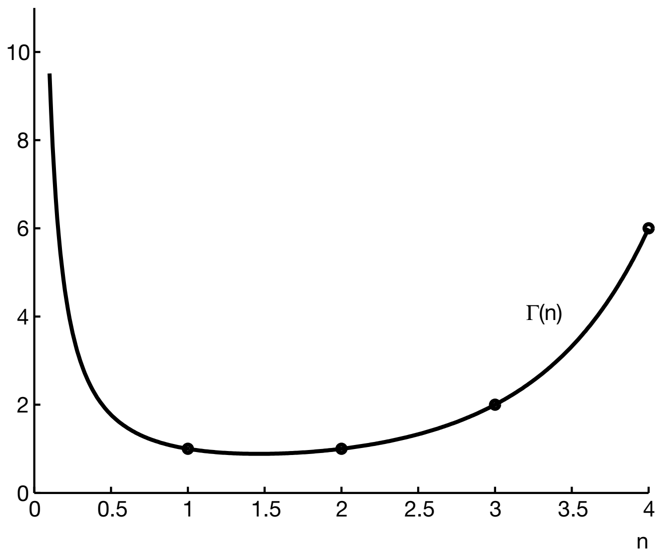
Note: The table reports income ratios for rich and poor countries. The baseline run uses the parameter values reported in Table 2, in particular: $\rho = -1$, $\theta^p = 2$, $\theta^r = 5$, and $\sigma = 1/2$. Other scenarios change one parameter at a time. The parameter β^r is changed when necessary to keep average underlying productivity twice as high in the rich country.

Figure 1. The Weibull Distribution of A_i



Note: The Weibull density for the "rich" country has $\beta = 1.93$ and $\theta = 5$, while the density for the poor country has $\beta = 1$ and $\theta = 2$.

Figure 2. The Gamma Function



Note: This figure plots Euler's factorial function $\Gamma(n)$. If n is a positive integer, $\Gamma(n+1) = n!$. The lowest value of n shown in the plot is 0.1, and $\Gamma(0.1) \approx 9.5$.

Appendix A. Facts about Income Differences

This appendix documents the large and growing income differences across the countries of the world.

At least since Pritchett (1997), it has been known that divergence characterizes the evolution of the world income distribution over the very long run. Up until two or three centuries ago, people everywhere were relatively poor, so that the ratio of incomes in the richest to poorest countries were probably on the order of two or three. In the last two hundred years, incomes have diverged, with the poorest countries remaining fairly close to the general subsistence-like level that characterized much of world history while the richest countries have grown rapidly. The ratio of per capita GDP in 2000 between the United States and Ethiopia, for example, is more than a factor of 50.

That said, the facts of the last half century are less well appreciated. Is the last half century characterized by convergence, divergence, or a relatively stable world income distribution? Much of the early work in the empirical growth literature emphasized the lack of convergence (but not divergence) in the world as a whole and the presence of convergence amount a group of relatively rich countries. More recently, increased appreciation has been given to the fact that the divergence that characterized much of history has continued in the 2nd half of the twentieth century (Maddison 2001, Aghion, Howitt and Mayer-Foulkes 2005). This continued divergence is shown graphically in Figure 1.

Between 1960 and 1999, the ratio of per capita GDP in the fifth richest country in the world to the fifth poorest country increased from 21 to 32. Similarly, the standard deviation across countries of the log of per capita GDP rose from 0.91 to 1.18. To put this increase in perspective, if the data were distributed normally with these standard deviations, the ratio of the 95th percentile to the 5th percentile would have increased over this period from 38 to 112.

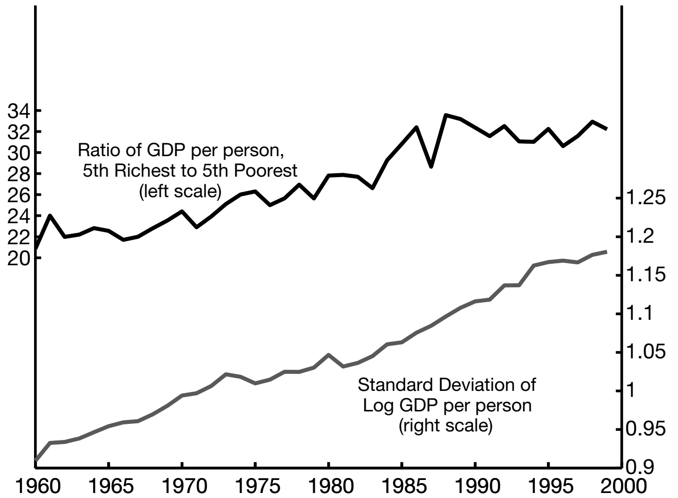
In fact, the data are not normally distributed, and the true 95th-5th ratio increased from 20.3 to 32.1 between 1960 and 1999, as shown in Table 1. More generally, what we see in this table is that the divergence occurred throughout the distribution of per capita income. For example, the 95th-50th ratio increases from 4.6 to 5.9 and the 80th-50th ratio rises from 2.2 to 3.8. The only place in the distribution where we do not see this divergence is at the very top. For example, the 95th-80th ratio remains relatively steady (even declining slightly) at just below a factor of two.

Table 1. Ratios of Per Capita GDP at Various Percentiles

	1960	1970	1980	1990	1999	Factor Increase
Max/Min	39.3	62.1	50.4	54.5	87.4	2.23
95/5	20.3	24.4	27.3	31.8	32.1	1.58
90/10	11.8	14.8	16.7	22.2	27.1	2.29
80/20	5.2	7.9	9.2	10.7	12.5	2.39
95/50	4.6	5.2	5.2	6.1	5.9	1.29
90/50	3.5	4.2	4.7	5.6	5.5	1.56
80/50	2.2	3.1	3.3	3.8	3.8	1.72
95/80	2.1	1.7	1.6	1.6	1.6	0.75
50/5	4.4	4.7	5.3	5.2	5.4	1.23
50/10	3.3	3.5	3.5	3.9	4.9	1.47
50/20	2.4	2.6	2.8	2.8	3.3	1.39

Note: Using the 104 countries that have continuous data for 1960 to 1999, the table reports the ratio of per capita GDP from various percentiles. For example, the 3rd row reports the 90th percentile to the 10th percentile in each year. The last column of the table shows the ratio of the 1999 column to the 1960 column. Underlying data from Penn World Tables 6.1.

Figure 1. Divergence in the Last Half Century



Note: Computed using Penn World Tables, Mark 6.1 of Heston, Summers and Aten (2002) using the 104 countries with continuous data from 1960 to 1999.

Appendix B. Proofs of the Propositions

Proposition 4.1: The Optimal Allocation of X_i

Proof. Notice that $Y_i = a_i X_i^\sigma$, so the optimal allocation of X_i solves

$$\max_{\{X_i\}} C = \zeta \left(\int_0^1 a_i^\rho X_i^{\rho\sigma} di \right)^{1/\rho} - \int_0^1 X_i di.$$

Solving this problem and substituting the solution back into the production function in equation (2) gives the result. (This is where the judicious definition of ζ comes in handy.) ▮

Proposition 4.2: The Optimal Allocation of K_i and H_i

Proof. Using the expression for output derived in the previous proposition, equation (8), the optimal allocations of K_i and H_i solve

$$\max_{\{K_i, H_i\}} \int_0^1 a_i^\lambda di$$

subject to the resource constraints in equations (4) and (5), where $a_i \equiv A_i (K_i^\alpha H_i^{1-\alpha})^{1-\sigma}$. The result follows from substituting the solution back into the expression for output given in equation (8). ▮

Proposition 4.3: The Symmetric Misallocation

Proof. Follows directly from the fact that $Y_i = A_i m$, where $m = (K^\alpha H^{1-\alpha})^{1-\sigma} X^\sigma$ is constant across activities. ▮

Proposition 4.4: The Solution for Q

Proof. Given in the text. ▮

Proposition 5.1: The Solution for Y/L

Proof. Given in the text. ▮