The Value of Information in Growth and Development

Charles I. Jones*

Department of Economics, U.C. Berkeley and NBER
E-mail: chad@econ.berkeley.edu
http://elsa.berkeley.edu/˜chad

September 12, 2005 – Version 0.30

EXTREMELY PRELIMINARY AND INCOMPLETE

This paper explores the role of information in the theory of economic growth and development. They way it is used here, "information" refers to every feature of an economy, including not only the economic environment, but also the institutions like markets and government policies that affect the allocation of resources. The first half of the paper considers a model where consumers are faced with an enormous range of possible goods they can purchase, so many that they do not know exactly how much utility they would get from consuming each good. Improvements in this information possessed by consumers increases welfare by making the allocation of resources more efficient, and these effects may be large in some cases. The second half extends these ideas to the production side of the economy and argues that a model where different kinds of information are crucial to successful production can lead to an information-based theory of TFP. In the theory outlined here, large differences in incomes can be explained by small differences in the ease with which people in a given country can acquire four different kinds of complementary knowledge.

Key Words:

JEL Classification: O40, E10

*This research is part of the Toulouse Network on Information Technology. I also thank the Hong Kong Institute for Monetary Research for hosting me during the early stages of this research.

1. INTRODUCTION

Modern growth theory views the discovery of new knowledge (new ideas) as the key driving force underlying economic growth. In this paper, I suggest that something closely related — information — plays an important role in the economic growth of advanced countries and in explaining the 50-fold differences between rich and poor countries that we observe in the data. By information, I mean all features of the underlying economy, including every possible action an individual could take and the economic consequences of those actions. New ideas in the Romer (1990) sense are information — the instructions for combining existing objects into a new good that yields increased utility are to be found in the economic environment; before the discovery of the new idea they are not known, but they are still there to be discovered. But information is broader than ideas and includes features we would not call ideas. For example, the utility that a particular person would get from reading a particular book is information, as is a specific method for greasing the bureaucratic wheels in order to obtain import licenses for the best textile machinery. Both pieces of information are part of the underlying fabric of an economy, even if they are not known by anyone in that economy.

Consider a more detailed version of the book example: suppose you'd like to pick up a book to read on your next business trip. According to Amazon.com, there are in excess of 2 million possible books available for purchase. The overwhelming majority of these books would be a poor match for any given person. Moreover, exactly how good the match quality is between a particular book and a particular person is information that is known only imperfectly, if at all, by a given person. The five books from among these millions that would produce the most utility may be entirely unknown to an individual. The names of these books and the utility levels they would produce are information. They are properties of the underlying economic environment that happen to be unknown.

Matching buyers and sellers is one of the fundamental purposes of markets. When markets work well, everyone whose marginal benefit for a good exceeds the marginal cost of production is matched with an additional unit of the good. One of the crucial conditions for this matching to work smoothly is that consumers must know about the existence of the good and they must know their own marginal benefit.

It is conventional in economics to assume that economic agents possess this information, but in this paper I would like to explore the possibilities that arise when important pieces of information are unknown. This exploration takes place in two main steps in what follows. First, I build a model of the book example. I show that improvements in information — like those that might occur because of advances in information technology — can potentially lead to large increases in welfare in an economy because of the improved quality of matches between people and the goods they consume. While books represent an obvious example of this improvement, the phenomenon itself is much more general. Other examples where matching up heterogeneous people with heterogeneous goods is the essence of the market include food, restaurants, wines, vacations, and movies. But items that make up a more significant share of expenditures also crucially possess a matching component, like cars and houses. As emphasized in a literature in macroeconomics, jobs may also be viewed as matches between workers and firms, and improvements in information technology may improve matching in the labor market, perhaps with a substantial effect on welfare.

In the second half of the paper, I apply this idea to economic development. Consider a potential entrepreneur in Kenya who starts a textile factory to make shirts. There is a large amount of complementary information that the entrepreneur needs in order for the business to succeed. First, the entrepreneur needs to know how to manufacture shirts. What materials are needed, what steps are involved, what machines are required? How does one make a shirt? Second, the entrepreneur needs to find a source for buying the materials and the textile equipment and needs to know how to hire and train a workforce. Third, the entrepreneur needs to know how to sell the shirts. This is the same matching problem considered in the first part of the paper, but this time from the producer side. One critical component determining the success or failure of the business is the ability of the entrepreneur to match with an intermediary who provides a substantial market for the goods once they are produced. Finally, the entrepreneur also needs to know how to manage the bureaucratic hurdles that may be encountered along the way.

There are at least four distinct kinds of information that entrepreneurs need in order to start a successful business. Failure along any one of these dimensions will reduce productivity sharply. The information is highly complementary, and we will model this aspect of production using an O-ring production function like that proposed by Kremer (1993). We will see that this view of information provides a theory that can explain the large productivity differences across countries with relatively small differences in the extent to which entrepreneurs can obtain the information necessary for production. Such a theory may also help us to understand the continued divergence of per capita incomes across countries during the last half century.

This (very rough) paper outlines some thoughts I have about the value of information to economic growth and development. At this point, I have some suggestive models, but much work remains to be done. The paper is divided into two main parts which are somewhat separate. The first is the analysis of information in the matching of consumers with goods. I try to give some indication of how important improvements in information technology may be to economic growth along this particular dimension.

After working on this first model, I started thinking about the application to development. In the end, this may well end up being the more important application.

2. PART I: THE ROLE OF INFORMATION IN MATCHING CONSUMERS AND GOODS

The simple model we construct here to illustrate the basic ideas in the paper involves matching heterogeneous consumers with heterogeneous goods. In some ways, the setup can be viewed as a simplified version of the discrete choice models of McFadden (1973) and Berry, Levinsohn and Pakes (1995), with a layer of matching and information added on top.

2.1. Setup

Suppose there are a continuum of individuals of measure one, and a continuum of goods indexed by i on the interval $[0, N]$. Each person has a random value for each different good, and these values are drawn randomly, both across goods and across people, from the distribution $F(v)$. The goods come in discrete bundles, and utility is such that the individual consumes either 1 unit of the good or none at all. The quality v_i of good i to a typical consumer may be known or unknown. Let this also be random across goods and people, such that each person knows her valuation for some fraction q of the goods, leaving the quality of fraction $1 - q$ unknown. All individuals are the same from the standpoint of the quality distribution they face and the fraction of goods for which they know the quality, allowing us to speak of a representative individual. However, individuals differ according to which goods have which quality and according to which goods are known.

As an example of this setup, consider books. There are millions of books out there to be read, and different books are valued differently by different people. For a given person, each book is either read once or not at all. The value to the person of reading a given book can be viewed as a draw from a distribution. For some fraction of the books, this quality is known in advance, e.g. because of experience with the author or based on reviews, etc. But for the vast majority of books, the quality is unknown.

An individual's utility function is

$$U = \int_0^N v_i x_i di + c, \tag{1}$$

where $x_i \in \{0, 1\}$ is the consumption of good i and c is the consumption of a generic good that we take as a numeraire. The budget constraint for an individual is

$$\int_0^N p_i x_i di + c \leq y, \tag{2}$$

where p_i is the price of good i and y is an exogenous endowment of income, assumed for simplicity to be the same across people.

On the production side, one unit of any good can be produced at constant cost w. We assume markets are perfectly competitive, so that all goods sell for a price equal to this marginal cost. We also assume that this cost of production is larger than the average valuation an individual places on goods, i.e. $w > Ev_i$. This implies that individuals will not purchase any good whose quality is unknown to them. In this sense, not knowing the quality of a good is equivalent to not knowing that the good exists at all, since in neither case will it be purchased.

2.2. Equilibrium

The representative individual chooses which goods to consume according to the following problem:

$$\max_{\{x_i\}, i \in Q} \int_Q v_i x_i di + y - \int_Q p_i x_i di, \tag{3}$$

where Q denotes the set of goods for which quality is known to the individual. Since $p_i = w$, this problem can be written

$$\max_{\{x_i\}, i \in Q} \int_Q (v_i - w) x_i di + y, \tag{4}$$

and the solution is straightforward:

$$x_i = 1 \iff v_i \geq w \text{ and } i \in Q. \tag{5}$$

That is, the individual consumes those goods whose values are known and whose values exceed the price, which is equal to marginal cost w.

This solution to the individual's problem allows us to characterize the overall market for good i. Let $G(v) \equiv 1 - F(v)$ denote the complement of the distribution, F. Because of the symmetry of the model, the distribution of values across goods for a single individual is the same as the distribution of values across individuals for a single good. Therefore, $G(p)$ describes the measure of the population whose valuation for the good exceeds the price p. Because only the fraction q of the population knows their valuation, total demand for a good, X_i, is given by

$$X_i = qG(p_i). \tag{6}$$

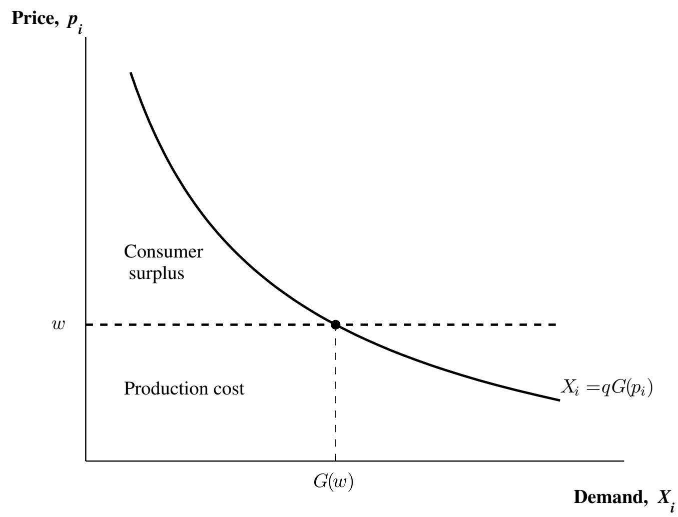
The market for good i can be characterized as in Figure 1.[1]

As it plays a role in our subsequent analysis, it is also useful to derive the expression for consumer surplus in this market, denoted CS. Consumer surplus can be written in any of the following three equivalent fashions:

$$\begin{aligned} CS &= q \cdot \int_w^\infty G(p) dp \\ &= q \cdot \int_w^\infty (v - w) f(v) dv \\ &= q \cdot \Pr[v \geq w] \left(E(v | v \geq w) - w \right). \end{aligned} \tag{7}$$

[1] This approach of using a statistical distribution for individual tastes in order to generate a market demand curve is explored in more detail in Garber, Jones and Romer (2005) in the context of the value of medical innovations.

FIGURE 1. The Market for Good i



Note: Strictly speaking, this is the demand curve that applies for $v_i \geq Ev_i$.

The first line states that consumer surplus is the area under the demand curve above the price. The second writes this as the integral over all of the valuations v that exceed the price w, multiplied by the density of valuations $f(v)$. Finally, the last line notes that the consumer surplus is the product of the fraction of valuations that exceed w (the measure of people purchasing the good) multiplied by the average surplus $v - w$ given that the good is purchased. All of these measures of consumer surplus are shifted in by the fraction q of people who know their valuations.

2.3. Welfare

Now that we know the individual's demand for goods, we can compute welfare for our representative person. This is given by

$$U^* = qN \cdot \int_w^\infty (v - w) f(v) dv + y \tag{8}$$

$$= qN \cdot \Pr[v \geq w] \left(E(v|v \geq w) - w\right) + y \quad (9)$$

$$= N \cdot CS + y. \quad (10)$$

That is, the welfare of the representative individual is equal to her base income, y, plus the consumer surplus she gets in each of the N quality markets; notice that our CS measure is already downweighted by the probability q, which is why this term does not appear in the final expression for welfare.

This second line of this welfare expression can be interpreted readily. Overall welfare is equal to base utility y plus the product of three terms. The first is qN, which reflects the measure of goods whose quality is known to the consumer. Next, $\Pr[v \geq w]$ represents the fraction of these goods for which the individual's valuation exceeds marginal cost. Finally, the last term in the product is the average utility gain an individual receives from a good that is purchased.

2.4. The Welfare Gain from Better Information

We are now ready to consider how improvements in information raise welfare. Suppose that the fraction of products that a typical consumer knows about increases from q to $q + \Delta q$. How valuable is this increase in information?

We can answer this question by computing the equivalent variation. That is, we can solve for the change in income that would deliver an increase in utility that is equivalent to that generated by the improvement in information. Let Δy denote this change in income. Then,

$$(1 + \frac{\Delta q}{q}) \cdot N \cdot CS + y = N \cdot CS + y + \Delta y.$$

Solving for Δy and expressing this as a percentage of initial income gives

$$\frac{\Delta y}{y} = \frac{\Delta q}{q} \cdot \frac{N \cdot CS}{y}.$$

Finally, an expression that is somewhat easier to implement empirically can be found by multiplying and dividing by $qw\,\Pr[v \geq w]$, yielding

$$\begin{aligned} \frac{\Delta y}{y} &= \Delta q \cdot \frac{wN\,\Pr[v \geq w]}{y} \cdot \frac{CS}{qwG(w)} \\ &= \Delta q \cdot \bar{\theta} \cdot \frac{CS}{\text{Production cost}}. \end{aligned} \tag{11}$$

This last expression says that the welfare gain from increasing the fraction of goods about which the consumer knows her valuation by proportion Δq is the product of three terms. The first is Δq itself. The second is $\bar{\theta} \equiv \frac{wN\,\Pr[v \geq w]}{y}$; this is equal to the fraction of income that would be spent on these "quality" goods if consumers knew their valuation for every good. Finally, the last term is the typical ratio of consumer surplus in a quality market to the total production cost in this market, $wqG(w)$. That is, this last term measures the average proportional improvement that comes from buying a quality good rather than just consuming the generic consumption good. As shown back in Figure 1, this last term is just the ratio of the consumer surplus triangle to the production cost rectangle.

Intuitively, the welfare gain from an increase in information depends on the size of the informational gain, the importance of quality goods in the budget, and the average consumer surplus in a quality market.

3. SOME QUANTITATIVE EXAMPLES

While the model is of course extremely stylized, it is helpful to plug in some numbers to get a rough sense for how large the gains associated with improvements in information can be. The equivalent variation in income shown in equation (11) is the product of three terms that have relatively straightforward economic counterparts.

First, we can consider a modest increase in information, say $\Delta q = .10$, so that the fraction of goods a consumer has knowledge of rises by 10 percentage points. Next, we need to say something about the fraction of income that would be spent on "quality" goods if consumers had perfect information. To the extent that we think information is important in the economy, we might expect this number to be close to one. To be conservative, we will consider values of 1/4, 1/2, and 1 in our exercise.

Finally, we need to calibrate the last term in equation (11), the ratio of consumer surplus to production cost (or total revenue) in a typical quality market. We have two complementary ways of assigning values to this term. The first looks at empirical evidence on the ratio of consumer surplus to revenue. The second parameterizes the distribution $F(v)$ and computes this ratio.

3.1. Using Evidence on Consumer Surplus Measures

The first approach to measuring this last term is to exploit the relatively large and growing literature on consumer surplus measurement, associated with Hausman (1981) and Hausman (1997), among many others. Hausman (1997) measures the consumer surplus associated with the introduction of Apple Cinnamon Cheerios and finds it to be on the order of $40 to $80 million. zzz Nevo (2001), etc. implies total revenues in the typical cereal market are about this large as well, suggesting a ratio of consumer surplus to production cost of about one.

Petrin (2002) provides estimates that can be used to infer the ratio of consumer surplus to revenue for minivans. He estimates a large average consumer surplus for a typical purchaser of $1247 in 1984 (in 1982-1984 prices). However, because the purchase price of a large ticket item like a minivan is so high — averaging about $8700 — this consumer surplus is only about 1/7th the magnitude of total revenue. Goolsbee and Petrin (2004) consider the consumer gains from the introduction of direct broadcast satellite television. They find the welfare gain to be about $150 per subscriber, which again works out to just under 1/3 of their estimate of the overall cost of satellite service ($30 per month, plus $50 per year in equipment). At the other end of the spectrum, Brynjolfsson (1996) estimates the consumer surplus associated with spending on information technology to be about 3 times the level of spending on information technology.

In an exercise that is especially relevant to the present paper, Brynjolfsson, Hu and Smith (2003) estimate the consumer surplus associated with the increase in the variety of books available to the typical consumer because of the emergence of online booksellers. Several of the findings in this paper are worth noting. First, the authors note that Amazon.com sells approximately 2.3 million unique titles, approximately 10 to 100 times more than the conventional brick-and-mortar bookstore. Brynjolfsson et al. the consider the additional variety offered by online booksellers as everything other than the top 100,000 ranked books; so they are evaluating the expansion of variety associated with books that are relatively obscure. Nevertheless, they find that sales of these "obscure" books account for just under 40% of Amazon's revenues and that the increase in consumer surplus associated with this expanded product variety was between $731 million and $1.03 billion in the year 2000 (about three dollars per person). Other data in BHS suggest that revenue from these titles to the online booksellers totaled about $580 million, suggesting a ratio of consumer surplus to revenue of between 1.25 and 1.75.

With these papers as background, we consider values for the ratio of consumer surplus to revenue of 1/2, 1, and 2.

Table 1 computes the equivalent variation in income that yields the same increase in welfare as raising q by 10 percentage points using these parameter values. The effect of increasing q by 10 percentage points is equivalent to an increase in income that ranges from a low of 1.25% to a high of 20%.

TABLE 1.
Equivalent Variation in Income (Percent) from Increasing q by .10

Expenditure Share, $\bar{\theta}$	Ratio of CS to Revenue 1/2	1	2
1/4	1.25	2.5	5
1/2	2.5	5	10
1	5	10	20

Note: The table shows $\Delta y/y \times 100$, i.e. the equivalent variation in income, expressed as a percent, that produces the same gain as increasing q by 10 percentage points. The calculation is based on the formula in equation (11).

Another way to think about these numbers is to consider the value of the multiplier on Δq. These multipliers range from a low of 1/8 to a high of 2. Therefore, if we thought that improvements in information technology led q to rise by a percentage point every year, this would be equivalent to an increase in income by an amount ranging from 1/8th of a percent to 2 percent each year. Since the overall growth rate of real per capita GDP in the United States during the last decade was 2.1 percent per year, these numbers are relatively large.[2]

3.2. Making Distributional Assumptions

An alternative approach to calibrating the ratio of consumer surplus to revenue relies more closely on the model. In particular, from Figure 1 and equation (7), we know that consumer surplus is given by the integral under the demand curve and therefore depends on the integral of the (complement of) the distribution of values in the population. Similarly, revenue in a market is $qwG(w)$, which can also be computed if we make a distributional assumption.

[2]This growth rate is for real per capita GDP between 1994 and 2004, from Table B-31 of the *Economic Report of the President, 2005*.

Suppose the distribution of values is Pareto, i.e. $F(v) = 1 - (v/\beta)^{-\alpha}$, where $\alpha > 1$ is the curvature parameter and $v \geq \beta > 0$ is the domain. In this case, the demand function is

$$X_i(p_i) = qG(p_i) = q\beta^{\alpha} p_i^{-\alpha}. \tag{12}$$

That is, the Pareto distribution for underlying values generates a constant elasticity demand function where the elasticity of demand is the Pareto parameter itself.[3]

Another useful result that can be gleaned from this distributional approach is related to a fact we have not really emphasized up until now. With millions of goods available, the typical individual only consumes a tiny fraction of the available goods. What is needed for the model to match this fact? Recall that the fraction of people consuming a particular good is equal to $q \Pr[v \geq w] = qG(w)$. Here, $G(w) = (w/\beta)^{-\alpha}$, so we have two parameters to play with. The parameter β is the smallest valuation realized in our distribution, so the ratio of w/β reflects the ratio of marginal cost to the smallest valuation. In general, to get $G(w)$ to be small, so that only a small fraction of people consume any given good, we need w/β to be large. As an example, if $\alpha = 2$ (motivated below), then $w/\beta = 100$ means that only 1 in 10,000 people consume a particular good.[4]

It is straightforward to show that the consumer surplus in this case is given by

$$CS = q \int_w^{\infty} G(p)dp = \frac{qw}{\alpha - 1} \left(\frac{w}{\beta}\right)^{-\alpha}. \tag{13}$$

[3] Jones (2005) shows that if ideas for production techniques are drawn from a Pareto distribution, the overall production function may take on a Cobb-Douglas shape.

[4] Of course, one thing the model completely misses is that some goods are inherently better than others, so that qualities are correlated across people to some extent. It seems like the model could readily be extended in this direction.

Similarly, total production cost is $qwG(w)$, so that the ratio of consumer surplus to cost is

$$\frac{CS}{Cost} = \frac{1}{\alpha - 1}. \tag{14}$$

There are two ways of moving forward, which give slightly different results. These correspond to thinking about α as the curvature parameter of the Pareto distribution, on the one hand, and thinking about α as the elasticity of demand, on the other hand.

As a matter of economics, we requires $\alpha > 1$ so that the elasticity of demand is larger than one in magnitude. Low values of α imply that the tail of the Pareto distribution is very thick. For example, by making α arbitrarily close to one, we can make the ratio of consumer surplus to revenue arbitrarily large. On a related point, for α between 1 and 2, the variance of the distribution is infinite (i.e. it does not exist), another symptom of the very thick tail of the distribution. A value of $\alpha = 2$ corresponds to a ratio of consumer surplus to profits equal to unity. In this view, it is easy to get large values for this consumer surplus ratio.

Interestingly, estimates of the Pareto parameter for the distribution of book sales are reported by Brynjolfsson et al. (2003) and Chevalier and Goolsbee (2003). These papers use data on the rank of books in best-seller lists and sales to regress rank on sales and recover a Pareto exponent. These estimates are related to the distribution of book sales rather than to the underlying heterogeneity in an individual's valuation for a book of given quality, but the estimates are suggestive, nonetheless. The estimates in Brynjolfsson et al. imply a Pareto curvature parameter of $\alpha = 1.15$, which implies a ratio of consumer surplus to revenue of 6.67. Chevalier and Goolsbee report a range of estimates for the Pareto coefficient, from a low of 0.9 to a high of 1.49, so that $1/(\alpha - 1)$ would range from infinite down to 2. They choose a value of 1.2 as their benchmark estimate, which corresponds to a ratio of consumer surplus to revenue of 5. Estimates of Pareto parameters from data on book sales, then, generally supports a quite high ratio of consumer surplus to profits.

On the other hand, it is also possible to get small values, and these seem more plausible when one thinks of α as a demand elasticity. Given that we are thinking of different books, cars, and houses as different goods, one might expect the elasticity of demand to be relatively high. With $\alpha = 5$, for example, the ratio of consumer surplus to revenue falls to 1/4. Notice also that if one were to add monopolistic competition to the model, the monopoly markup would equal $\alpha/(\alpha - 1)$. A value of $\alpha = 2$ corresponds to a (proportional) markup of 2, while a value of $\alpha = 5$ corresponds to a markup of 20% over marginal cost. As just one example, this latter value for an elasticity and markup corresponds roughly to the numbers that Berry et al. (1995) and Petrin (2002) find for cars and minivans.

My view of the distributional approach to measuring the ratio of consumer surplus to revenue is that it generally supports the range of numbers examined in Table 1. If the intriguing estimates of the Pareto parameter from the book sales data are representative, then the welfare gains may be even larger than is suggested in the table. On the other hand, large elasticities of demand and small markups — especially for large ticket items like cars and minivans — point toward smaller estimates.

4. PART II: INFORMATION AS TFP

The rest of the paper now shifts gears considerably. Up until now, we have considered a single, relatively narrow application of information to economic growth that occurs at the stage where goods are sold to final consumers. What motivates the remainder of the paper is the notion that information is important — even crucial — at *every* stage of production. Small gaps in information are then able to multiply up to explain large differences in incomes across countries. This question of what explains the 50-fold difference in incomes between, say, the United States and Ethiopia is one of the most important unanswered questions in macroeconomics. Perhaps thinking carefully about information can take us toward an answer.[5]

4.1. The Economic Environment

A team of $\bar{\ell}$ managers uses capital K and labor L in order to produce an output good Y, according to a relatively standard production function

$$Y = Q(K^{\alpha}L^{1-\alpha})^{\nu} \tag{15}$$

where $\alpha < 1$. As described below, Q captures the production knowledge embodied in the managers. The knowledge itself is nonrivalrous, justifying the increasing returns to Q, K, and L taken together. However, this knowledge must be embodied in the managers of the firms as rivalrous human capital in order for production to take place. Because managers are themselves rivalrous there are constant returns to scale in K, L, and management taken together, and therefore decreasing returns to K and L alone. With a given stock of knowledge embodied in a given management team, doubling the amount of capital and production workers leads to a less than doubling of output. Reflecting this fact, $\nu < 1$ is the span-of-control parameter of the management team (Lucas 1978).

4.1.1. Production Knowledge

The managers invest in learning knowledge about how to produce. There are n different kinds of knowledge that are necessary for production, and q_i is the amount of knowledge of type i that the managerial team possesses. Anticipating the discussion that follows, q_1 is the knowledge of how to

[5] Recent work on this topic includes Hall and Jones (1999), Klenow and Rodriguez-Clare (1997), Prescott (1997), Parente and Prescott (1999), Acemoglu, Johnson and Robinson (2001), Acemoglu and Johnson (2005).

produce a particular good, q_2 is knowledge related to obtaining and managing the inputs to production, q_3 represents knowledge of finding a market for the good and delivering the good to that market, and q_4 is knowledge related to managing the bureaucratic procedures associated with creating and running a firm.

There may be other kinds of knowledge beyond these examples; the model allows for n types. However, $n = 4$ distinct types will suffice, both to understand the nature of the production function and for the quantitative analysis that comes later.

Overall production knowledge, Q is given by

$$Q = \Pi_{i=1}^{n} q_i. \tag{16}$$

Different types of knowledge combine in a multiplicative fashion to yield overall production knowledge, as in the "O-ring" production function of Kremer (1993). Notice that doubling the amount of knowledge of each type increases overall production knowledge (and therefore TFP) by a factor of 2^n, so there are strong increasing returns built into this function. This key assumption requires some justification.

In thinking about how different kinds of knowledge combine in production, it is helpful to consider specific examples. For instance, consider the production of socks in one of the fastest-growing sock-producing regions of the world: Yiwu, China. In order to produce socks, a management team must first have the knowledge of exactly how to produce socks: what is the instruction set — the recipe — according to which one manufactures socks. This kind of knowledge plays a central role in the endogenous growth literature building on Romer (1990).

A second kind of information is related to managing the production process itself. The management team must know how and where to buy the cotton and polyester that is used in production. It must know what kind of sock-making machines to buy, where to purchase them, and how to keep them in good repair. And it must know how to hire, train, and manage a competent workforce. This second kind of knowledge fits in well with traditional notions of TFP. For example, one can imagine a production function such as $F(b_1K, b_2, L, b_3\text{Materials})$. If this production function is Cobb-Douglas with constant returns in capital, labor, and materials, then the b's can be factored out as a single TFP term. If $b_1 = b_2 = b_3 \equiv b$, then b would multiply the production function the way q_2 does in our formulation.

A third type of information needed by the management team is knowledge of the market for socks. The management team needs to know how to find the highest-value buyers of the socks and how to deliver the stocks to these buyers, potentially in many countries throughout the world. This is a traditional matching problem of the kind studied by Diamond (1982), Mortensen (1982), and Pissarides (1985). It is also similar to the matching problem described earlier in this paper: better knowledge allows the firm to find the customers who value the product most highly. The extent to which q_3 is less than one then reflects the extent to which the firm fails to match with the highest-value consumers of socks.

A final kind of knowledge needed for successful production is knowledge of the local laws, regulations, and customs. The management team must know how to acquire the necessary licenses and regulatory approval. A team that can do this rapidly and easily will be more productive. If the product is being sold in one or more foreign countries, the management team may also require this knowledge about the other countries. To the extent that bribes have to be paid or resources have to be diverted to tasks other than producing the output good, q_4 will be lower.

Now that we understand the different kinds knowledge that may be used in production, the next question is how these types of knowledge combine to produce output. One simple formulation popular in the growth literature is a perfect-substitutes approach. For example, in Romer (1990) and Aghion and Howitt (1992), what matters is the total number of ideas the economy has discovered. The total stock of knowledge is simply the sum of the individual ideas, and one idea can be thought of as a perfect substitute for another, at least insofar as measuring the total stock of knowledge relevant for production.

This same kind of formulation applies here for a particular type of knowledge. As we will see below, individuals can learn a specific type of knowledge, and the amount they learn is cumulative. However, as Kremer (1993) assumed for skills of different workers, we assume that different *types* of knowledge are complementary in production. One can think of each type of knowledge as being a necessary stage in production, and a loss in efficiency at any one stage reduces the overall value of production by the same percentage.

As one example, it is useful to consider two alternative production structures for Q. The first is the one we have, which is the O-ring structure. An alternative structure is a geometric mean: the O-ring structure raised to the power $1/n$, which would exhibit constant returns in the different types of knowledge. Now use these two alternatives to think about the production function for socks. Is it the geometric mean of the different kinds of knowledge — how to make socks, what kind of inputs to purchase and how to manage them, how to sell the socks, and how to manage the bureaucratic infrastructure — or the overall product that matters? For this intuition, consider the case where the q_i are all less than one. In the geometric mean interpretation, a management team could be 100% effective in knowing how to make socks and how to find the market, but only 50% effective in buying good quality inputs and managing the bureaucracy. Overall productivity would then be something like 75% (actually, the number turns out to be 87%).

In contrast, in the O-ring formulation, the q_i can be thought of as the fraction of the value of the product that is retained through the application of each type of knowledge. In this case, the 50% loss in buying the right inputs and the 50% loss in managing the bureaucracy means that the firm's overall efficiency is such that only 25% of the maximum value of production gets realized. We follow Kremer (1993) in his O-ring structure and use this latter interpretation. Each of the q_i should be thought of as being measured on a scale so that if a single q_i is reduced by 10%, the entire value of production is reduced by 10%.

4.1.2. Acquiring Production Knowledge

How does the management team acquire these n types of knowledge? At some level, this too is a matching problem. Learning to produce socks or automobiles or medical stents involves some combination of learning existing production techniques from existing producers and generating new and improved production techniques. The knowledge of what kind of materials to use in production, where to buy these materials, which sock-making machines to purchase, etc., are also typically learned from other producers. More general management skills such as how to motivate workers and how to manage bureaucracies may be learned in other industries.

The matching nature that surely characterizes some of this knowledge acquisition should be interesting to pursue further, but we have not yet done this. It plays a role in the literature on economic geography and surely helps us to understand the geographic concentration of firms of a similar type.

For now, though, we model this knowledge acquisition using a simple reduced-form matching function. This can be thought of as the knowledge-matching version of the Diamond-Mortenson-Pissarides matching function for the labor market. In particular, we assume that if the management team exerts effort ℓ_i, knowledge of type i is learned according to

$$q_i = A_i \ell_i. \tag{17}$$

In this equation, A_i is an exogenous parameter that governs how efficiently managerial effort gets turned into production knowledge. The assumption is that in some countries, it may be harder to learn certain kinds of production knowledge. For example, if there are no sock producers in a country, learning the knowledge related to sock production may be more difficult. International trade and FDI may make learning certain kinds of knowledge easier and would surely facilitate learning about other markets for particular products and how to transport goods to those markets. In the end, we will show that relatively small differences in A_i across countries, because of the complementary nature of production knowledge, can lead to large differences in per capita income. In the future, we plan to think more carefully about how these differences in A_i may arise.

The static nature of the knowledge-matching function is for simplicity only. One could imagine a richer dynamic setting such as $\dot{q}_i = A_i \ell_i - \delta q_i$, where δ reflects the depreciation of knowledge. Notice that the steady-state of the richer dynamic model takes the form of our static equation, so not much is lost in our formulation.

4.1.3. Closing the Model

The model is closed with several more equations. First is the resource constraint on managerial effort:

$$\sum_{i=1}^{n} \ell_i = \bar{\ell}. \tag{18}$$

A firm is a collection of $\bar{\ell}$ managers, a number that is exogenously given. Each manager has one unit of labor that can be used in acquiring knowledge.

With respect to capital, we assume there is a world capital market from which the firm can obtain capital at a constant exogenous rental price of $\bar{r}$. To simplify, we assume all capital is owned abroad.

Finally, the economy consists of $\bar{L}$ people, any of whom can be managers or workers. Let M denote the endogenous number of firms in an economy. Then the resource constraint for labor is

$$M(\bar{\ell} + L) = \bar{L}. \tag{19}$$

4.2. The Optimal Allocation of Resources

A natural allocation of resources to focus on in this economy is the allocation that maximizes consumption per person. Since all agents are identical, per capita consumption is given by $c \equiv M(Y - \bar{r}K)/\bar{L}$. This motivates the following definition:

DEFINITION 4.1. The optimal allocation of resources in this economy consists of values for $Y, Q, K, L, \{q_i\}, \{\ell_i\}, M, c$ that solve[6]

$$\max_{\{\ell_i\}, M, K, L} c \equiv M(Y - \bar{r}K)/\bar{L}$$

subject to

$$Y = Q(K^{\alpha} L^{1-\alpha})^{\nu}$$

$$Q = \Pi_{i=1}^{n} q_i$$

$$q_i = A_i \ell_i, \;\; i = 1, ..., n$$

$$\sum_{i=1}^{n} \ell_i = \bar{\ell}$$

$$M(\bar{\ell} + L) = \bar{L}.$$

[6]To get the number of equations and unknowns to match up, notice that there are Lagrange multipliers associated with the last two constraints of this maximization problem.

Solving for the optimal allocation of resources is straightforward in this model. Managerial effort in acquiring each type of knowledge enters symmetrically, yielding

$$\ell_i^* = \frac{\bar{\ell}}{n}. \tag{20}$$

Capital is hired until the marginal product of capital equals the rental price:

$$\alpha\nu \frac{Y^*}{K^*} = \bar{r}. \tag{21}$$

Since any worker can be used as a manager, they must both have the same marginal product of labor. This pins down the size of the firm as

$$L^* = \frac{\nu}{1-\nu}(1-\alpha)\bar{\ell}. \tag{22}$$

The number of firms is then pinned down by the resource constraint as

$$M^* = \frac{\bar{L}}{\bar{\ell} + L^*}. \tag{23}$$

Total managerial knowledge in a firm is

$$Q^* = \left(\frac{\bar{\ell}}{n}\right)^n \Pi_{i=1}^n A_i, \tag{24}$$

and our welfare measure, consumption per person c, is given by

$$c^* = \Psi \cdot Q^{*\frac{1}{1-\alpha\nu}} \left(\frac{\alpha\nu}{\bar{r}}\right)^{\alpha\nu} \cdot \bar{\ell}^{-\frac{1-\nu}{1-\alpha\nu}}, \tag{25}$$

where Ψ is a constant.[7] Per capita consumption is proportional to managerial knowledge, Q, raised to a power that depends on the degree of diminishing returns to capital accumulation. As in the standard neoclassical growth model, increases in Q lead to additional capital deepening, which produces a multiplier effect on output and consumption.

[7]It is given by $\Psi \equiv (1-\nu)\left(\frac{\nu}{1-\nu}(1-\alpha)\right)^{\frac{(1-\alpha)\nu}{1-\alpha\nu}}$.

At this stage it is convenient to introduce an assumption that we will sometimes use to simplify the model further. In particular, we can make a symmetry assumption that $A_i = \bar{A}$ for $i = 1, ..., n$. Under this symmetry assumption, $Q^* = (\bar{A}\bar{\ell}/n)^n$, and we get the result that per capita consumption is proportional to $\bar{A}$ in the following way:

$$c^* \propto (\bar{A}\bar{\ell})^{\frac{n}{1-\alpha\nu}}. \tag{26}$$

In the data, per capita consumption varies across the richest and poorest countries by something like a factor of 64 (the exact numbers will be discussed a bit later). How can the model account for such a difference?

The answer turns out to be surprisingly easily. For our back-of-the-envelope calibration, consider the following choices of parameter values. Following Gollin (2002), let's assume a capital share of 1/3, common across countries, so that $\alpha\nu = 1/3$. Assume $n = 4$, corresponding to the four kinds of knowledge discussed extensively in the previous section.

With these parameter choices, $n/(1 - \alpha\nu) = 4/(1 - 1/3) = 6$. To explain a 64-fold difference in c^* between the richest and poorest countries of the world then requires us to explain a 2-fold difference across countries in $\bar{A}\bar{\ell}$, since $2^6 = 64$, and this seems perfectly plausible. For example, it could be that the richest countries of the world are more than twice as productive as the poorest countries in producing each kind of knowledge. Or it could be that $\bar{A}$ and $\bar{\ell}$ each vary by a factor of just over 1.4 between the richest and poorest countries. The richest countries may have a slight edge in producing knowledge and may have a contractual structure that allows them to support firms with 40% larger managerial teams.[8]

[8]In discussing the results with respect to $\bar{\ell}$, we should note that we are ignoring the effect that comes from the diminishing returns associated with ν. In particular, there is an additional term in equation (25) relative to our simplification in (26) that depends on $\bar{\ell}^{-\frac{1-\nu}{1-\alpha\nu}}$. In the calibration, then, the elasticity of consumption with respect to $\bar{\ell}$ really is $(n - (1 - \nu))/(1 - \alpha\nu)$. However, conventional values of ν are 0.95 or so, so this effect is quite small quantitatively.

This is the main point of this section: because different kinds of knowledge are complementary, small differences across countries in the productivity of managers in acquiring knowledge can multiply up to yield large differences in productivity and income.

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 productivity parameters can yield large differences in incomes. Second, improvements in the productivity of knowledge acquisition along any single dimension have relatively small effects on output, again because of the complementary nature of knowledge. 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 the magic bullet relatively easily and this magic bullet would become well-known. 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 effects on per capita income in their model. But we see little evidence in the data for such large multipliers.

Here, while it is true that small differences in A_i across countries lead to large differences in income, there are a number (four in this case) of distinctly different productivities that must be improved in order to capture these large gains. Obtaining the instruction manual for how to produce socks is not especially useful if it is extremely difficult to buy the cotton and polyester inputs that are needed to make socks, if the bureaucratic hurdles that must be overcome are substantial, and if the market to which these socks will be sold is unknown.

5. DIVERGENCE

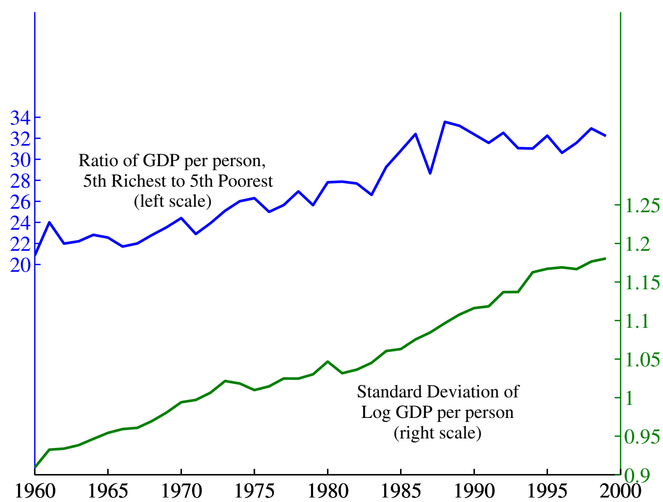
5.1. Facts

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 2.

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.

FIGURE 2. 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.

TABLE 2.
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.

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 2. 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.

5.2. Extending the Model to Account for Divergence

How do we understand this continued divergence of per capita income across countries? An extension of the model in the previous section may provide some insight. The way the model works can be summarized as follows. There are $\bar{N}_t$ intermediate goods in the world, each of which is produced using the production function in the previous section. These intermediate goods are combined as in Romer (1990) to produce a final output good. At any point in time, a country may know how to produce and use only a subset of these $\bar{N}$ goods. Over time, the number of intermediate goods in the world grows, and the number that a country knows how to produce may also grow. But these growth rates need not be the same. Divergence can be understood as the advanced countries of the world inventing, producing, and using more and more goods of higher and higher quality. If the poorest countries of the world are not learning to use these new goods as rapidly as the rich countries, divergence can occur.

To make this story more precise, suppose there are a variety of intermediate goods indexed by j on the interval $[0, N_t]$, where t denotes time. The production of any single good j occurs as described in the previous section. These goods are then used to produce a single final-output good according to

$$Y_t = \left(\int_0^{N_t} (Y_{jt} j^{\beta})^{\rho} dj \right)^{\frac{1}{\rho}} \tag{27}$$

where $0 < \rho < 1$ is the curvature parameter of the technology that combines these goods, with elasticity of substitution $1/(1-\rho) > 1$. Goods are ordered so that "newer" intermediate goods (i.e. goods with a higher index) are also better goods, with $\beta > 0$ governing how quality rises with the index.

From the previous section, we can write the production function for the jth good as

$$Y_j = \Phi Q_j^{\frac{1}{1-\alpha\nu}} L_j, \tag{28}$$

where L_j is the total amount of labor (managers and workers) devoted to the production of good j and where Φ is a constant.[9]

[9]Its value is $\Phi = ...$

Now it is time for another symmetry assumption to simplify the presentation. In particular, assume that $Q_j = Q$ so that production knowledge for each of the intermediate goods is the same.

In this case, the efficient allocation of labor across goods — that is, the allocation that maximizes output — is given by

$$\frac{L_j}{\bar{L}} = \frac{(1+\gamma)j^{\gamma}}{N^{1+\gamma}}, \tag{29}$$

where $\gamma \equiv \frac{\beta\rho}{1-\rho}$ and $\bar{L}$ is the total amount of labor used in production. Because intermediate goods enter production with an elasticity of substitution greater than one, goods that are more productive are used more intensively.

Substituting this allocation back into the two production functions in equations (27) and (28) yields, after some tedious algebra,

$$y \equiv \frac{Y}{\bar{L}} = \Phi^{1/\rho} N^{\mu+\beta} Q^{\frac{1}{1-\alpha\nu}}, \tag{30}$$

where $\mu \equiv 1/\rho - 1$.

Just as in the previous section, output per worker depends on production knowledge with an elasticity that is inflated because of capital deepening. So all of the results in the previous section continue to hold. Now, however, there is an additional term. Output per worker is proportional to the number of intermediate goods the economy knows how to use, N, raised to the power $\mu + \beta$.

This elasticity is the sum of two terms. The first, μ, reflects the standard gains from variety, originally emphasized in Romer (1990). Quantitatively, however, there are reasons to think these gains may be relatively small. For example, in variety models like this one, the elasticity of substitution between varieties is $1/(1-\rho)$, and the implied gross markup of price over marginal cost if a variety were produced under monopolistic competition would then be $1/\rho$.[10] The net markup is then $1/\rho - 1$, which is exactly the

[10]For an example with this notation, see Jones (forthcoming).

expression that we have defined to equal μ. That is, in a standard variety model, the gains from variety depend on the parameter combination that also determines the net markup. Empirically, these markups seem to be quite low, on the order of 10 or 20 percent (Basu and Fernald 1997), suggesting a high elasticity of substitution between varieties of 11 or 6. It is so easy to substitute across different varieties that new varieties are not especially valuable. If all varieties were created equal, then, differences in the range of varieties used in production by different countries would have a very small effect on productivity. As an example, if we thought the poorest countries of the world could only use 1/2 of the intermediate goods used by a rich country, this would reduce output by a factor of something like $(1/2)^{.1} = .93$, a tiny fraction of the 50-fold difference between the richest and poorest countries.

This is the reason we augmented our model to include a quality effect. In the final goods production function, equation (27), newer varieties are also better and more productive. The parameter β governs how much better they are, and what we see in the solution to this model is that β also enters the elasticity of output with respect to varieties. This gives differences in N an extra kick and may help to explain some of the differences in incomes across countries.

Notice that in terms of growth rates, $g_y = (\mu + \beta)g_N$ (under the maintained assumption that Q is constant over time). In the long run, countries grow because of an increase in the range of goods they can use in production. Part of this is an expanded variety effect (associated with μ) but part is also because newer varieties are more productive (associated with β).

At this point, I do not have a good way of calibrating g_N and β. However, it does not strike me as crazy to guess that perhaps $g_y = g_N$ so that $\mu + \beta = 1$ in a country like the United States. That is, maybe it is the case that per capita income and the number of new goods are growing at roughly the same rate. In this case, notice that the quality effects associated with β are much more important than the variety effects associated with μ. For example, if $\mu = .1$, then $\beta = .9$, and quality effects are 9 times more important than variety effects.

In terms of levels accounting, this assumption implies that if the poorest countries of the world use only the lower half of the intermediate goods in production, output per worker is reduced by a factor of 2 from the N term. By itself, of course, this is not huge, and this explains why the Q theory in the previous section is still needed. Nevertheless, because we get a factor of 2 from differences in the quality of goods used in rich versus poor countries, the necessary differences in $\bar{A}$ across countries are that much smaller.

5.3. How does the model explain divergence?

Let $\bar{N}_t$ denote the measure of intermediate goods invented in the world, and let's assume $\bar{N}_t$ grows at at rate $g_{\bar{N}}$.[11] Management teams in different countries must learn the n types of knowledge associated with producing each intermediate good. It may be that $A_{ij} = 0$ for some i for intermediate goods beyond some value N in a country, so that the country cannot use more than N intermediate goods. In this case, *a country's growth and development is limited by the extent to which it can use and produce intermediate goods.* Countries can potentially grow at different rates depending on how the A_{ij} get "switched on," i.e. depending on how countries learn to work with new intermediate goods.

In this model, per capita income in a country will grow at rate g_N which may be different from $g_{\bar{N}}$, and this learning/matching process can explain the divergence of incomes across countries. In particular, divergence would occur if the richest countries in the world were actually the best at learning

[11] This is presumably because of some research effort that takes place throughout the world as in Romer (1990) or Jones (1995), but we will take this growth as exogenous in what follows.

to use and produce new intermediate goods. Poor countries may find it difficult to learn to produce new intermediate goods, and if $g_N < g_{\bar{N}}$, it is possible to explain divergence.

How large would the differences need to be? Notice that the ratios of rich to poor incomes seem to grow by something like a factor of 2 over the 40 year period between 1960 and 1999. Doubling over 40 years is consistent with a growth rate difference in N of 1.8 percent. So if $g_{\bar{N}} = .018$ and $g_N = 0$ for the poorest countries in the world, this would account for the divergence that we have observed.

6. WHERE THIS PROJECT IS HEADED...

Since Romer (1990), ideas and knowledge have played a central role in the study of economic growth. In the theory of economic development, knowledge has played a less important role, with recent work emphasizing institutions. What I have learned so far in this project is that there may be a central role for knowledge and information in the theory of economic development.

Relative to the models I've developed so far, probably the most important missing piece that I have in mind pertains to *matching*. Management teams need to be matched up with several kinds of complementary information in order for productivity to be high. At this point, I've modeled this matching function as a simple production function with exogenous differences in productivity. However, the matching interpretation suggests that how people obtain this information may depend on how many people "close" to them possess the needed information. Geography and trade may play an important role in determining the productivity levels, and improvements in information technology may help the spread of at least some kinds of information.

REFERENCES

Acemoglu, Daron and Simon Johnson, "Unbundling Institutions," *Journal of Political Economy*, 2005, *forthcoming.*

—, —, and James A. Robinson, "The Colonial Origins of Camparative Development: An Empirical Investigation," *American Economic Review*, December 2001, *91* (5), 1369–1401.

Aghion, Philippe and Peter Howitt, "A Model of Growth through Creative Destruction," *Econometrica*, March 1992, *60* (2), 323–351.

—, —, and David Mayer-Foulkes, "The Effect of Financial Development on Convergence: Theory and Evidence," *Quarterly Journal of Economics*, February 2005, *120* (1), 173–222.

Basu, Susanto and John G. Fernald, "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy*, April 1997, *105*, 249–283.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, July 1995, *63*, 841–890.

Brynjolfsson, Erik, "The Contribution of Information Technology to Consumer Welfare," *Information Systems Research*, September 1996, *7* (3), 281–300.

—, Yu Hu, and Michael D. Smith, "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers," *Management Science*, November 2003, *49* (11), 1580–1596.

Chevalier, Judith and Austan Goolsbee, "Measuring Prices and Price Competition Online: Amazon.com and BarnesandNoble.com," *Quantitative Marketing and Economics*, 2003, *1*, 203–222.

Diamond, Peter A., "Wage Determination and Efficiency in Search Equilibrium," *Review of Economic Studies*, 1982, *49* (2), 217–27.

Garber, Alan M., Charles I. Jones, and Paul M. Romer, "Insurance and Incentives for Medical Innovations," July 2005. U.C. Berkeley mimeo.

Gollin, Douglas, "Getting Income Shares Right," *Journal of Political Economy*, April 2002, *110* (2), 458–474.

Goolsbee, Austan and Amil Petrin, "The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV," *Econometrica*, March 2004, *72* (2), 351–381.

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

Hausman, Jerry A., "Exact Consumer's Surplus and Deadweight Loss," *American Economic Review*, 1981, *71* (4), 662–676.

___, "Valuation of New Goods under Perfect and Imperfect Competition," in Timothy F. Bresnahan and Robert J. Gordon, eds., *The Economics of New Goods*, University of Chicago Press, 1997.

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

Jones, Charles I., "R&D-Based Models of Economic Growth," *Journal of Political Economy*, August 1995, *103* (4), 759–784.

___, "The Shape of Production Functions and the Direction of Technical Change," *Quarterly Journal of Economics*, May 2005, *120* (2), 517–549.

___, "Growth and Ideas," in Philippe Aghion and Steven A. Durlauf, eds., *Handbook of Economic Growth*, New York: North Holland, forthcoming.

Klenow, Peter and Andres Rodriguez-Clare, "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, 1997.

Kremer, Michael, "The O-Ring Theory of Economic Development," *Quarterly Journal of Economics*, August 1993, *108* (4), 551–576.

Lucas, Robert E., Jr., "On the Size Distribution of Business Firms," *Bell Journal of Economics*, 1978, *9*, 508–523.

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

Manuelli, Rodolfo and Ananth Seshadri, "Human Capital and the Wealth of Nations," March 2005. University of Wisconsin working paper.

McFadden, Daniel, "conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka, ed., *Frontiers in Econometrics*, Vol. 3, New York: Academic Press, 1973.

Mortensen, Dale T, "Property Rights and Efficiency in Mating, Racing, and Related Games," *American Economic Review*, 1982, *72* (5), 968–79.

Nevo, Aviv, "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, March 2001, *69* (2), 307–342.

Parente, Stephen L. and Edward C. Prescott, "Monopoly Rights: A Barrier to Riches," *American Economic Review*, December 1999, *89* (5), 1216–1233.

Petrin, Amil, "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 2002, *110* (4), 705–729.

Pissarides, Christopher A, "Short-run Equilibrium Dynamics of Unemployment Vacancies, and Real Wages," *American Economic Review*, 1985, *75* (4), 676–90.

Prescott, Edward C., "Needed: A Theory of Total Factor Productivity," 1997. Federal Reserve Bank of Minneapolis, Staff Report 242.

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

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