Growth, Capital Shares, and a New Perspective on Production Functions

Preliminary — Comments appreciated

Charles I. Jones*

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

June 12, 2003 – Version 1.0

Standard growth theory implies that steady-state growth in the presence of exponential declines in the prices of computers and other capital equipment requires a Cobb-Douglas production function. Conventional wisdom holds that capital shares are relatively constant, so that the Cobb-Douglas approach might be a good way to model growth. Unfortunately, this conventional wisdom is misguided. Capital shares exhibit substantial trends and fluctuations in many countries and in many industries. Taken together, these facts represent a puzzle for growth theory. This paper resolves the puzzle by (a) presenting a production function that exhibits a short-run elasticity of substitution between capital and labor that is less than one and a long-run elasticity that is equal to one, and (b) providing microfoundations for why the production function might take the Cobb-Douglas form in the long run.

Key Words: Growth, Leontief, Pareto, Capital Share, Equipment Prices, Steady State
JEL Classification: O40, E10

*I am grateful to Pete Klenow, Sam Kortum, and Alwyn Young for helpful conversations during the genesis of this paper and to seminar participants at Arizona State University, the Chicago GSB, Stanford, and U.C. Berkeley for comments on an earlier draft. This research is supported by NSF grant SES-0242000.

This paper is motivated by a collection of stylized facts that, when taken together, are puzzling from the standpoint of modern growth theory. The stylized facts are these:

1. Growth rates in U.S. per capita GDP do not show a trend for the last 125 years; see, for example, Jones (1995b).

2. Contrary to conventional wisdom, aggregate payments to capital and labor as a share of GDP are not constant over time, and in fact the capital share shows a substantial trend in many countries and industries. This fact has been documented in several recent papers and will be discussed in more detail in Section 2 below.

3. Estimates of the elasticity of substitution between capital and labor in production are often less than unity. Studies documenting this fact have been surveyed by Hammermesh (1993) and Antràs (2001). Recent estimates supporting this fact that also distinguish between skilled and unskilled labor can be found in Krusell, Ohanian, Rios-Rull and Violante (2000) and Caselli and Coleman (2000). No individual study is especially compelling, and even the results taken as a whole are not conclusive. But the typical result seems to be that the elasticity of substitution between capital and labor is less than one.[1]

4. The price of capital goods in the "equipment" category — computers, machine tools, turbines, mini-mills — have been falling relative to the price of nondurable consumption. This fact was documented carefully by Gordon (1990) and has been emphasized more recently in a series of papers including Greenwood, Hercowitz and Krusell (1997) and Whelan (2001). The falling relative price is taken as evidence of a faster rate of technological change being embodied in these capital goods than in consumption, and

[1] Antràs (2001) provides some insight into the wide range of the estimates. In particular, he notes that estimates close to unity typically come from time series studies that ignore the possibility of technological change, which biases the estimate toward one.

this phenomenon is called investment-specific technological change. One manifestation of this declining price is a rising ratio of real capital equipment divided by real GDP.

These stylized facts are puzzling from the standpoint of modern growth theory. To see this, recall a well-known result, which will be referred to as the steady-state growth theorem: if a neoclassical growth model is to possess a steady state with positive growth and a positive capital share, either technological change must be labor augmenting or the production function must be Cobb-Douglas in capital and labor. (This result is stated and proved formally in Appendix A).

Motivated by facts such as Fact 1 above, growth theorists typically want steady states to be possible in our growth models. The presence of investment-specific technological change and a rising equipment-output ratio suggest that the production function should be Cobb-Douglas if this is to be the case. However, the Cobb-Douglas production function is inconsistent with Fact 2, the variation and trends in capital shares, and Fact 3, the estimates of the elasticity of substitution between capital and labor. Alternatively, if one chooses a non-Cobb-Douglas production function to match Facts 2 and 3, the presence of investment-specific technological change will render steady-state growth with a positive capital share impossible.

A literature from the 1960s approaches this tension by seeking to uncover economic forces that would lead technological change to be entirely labor-augmenting in the long run. This is the approach taken originally by Kennedy (1964), Samuelson (1965), and Drandakis and Phelps (1966) and examined recently using the tools of new growth theory by Acemoglu (2001). There are two important problems with this approach. First, the spirit of the approach appears to be contradicted empirically by Fact 4: the declining relative price of equipment suggests that there is a substantial amount of technological change that has been embodied in capital in recent decades. Second, the careful microfoundations of the approach provided by Acemoglu (2001) indicate that the outcome in which technological change is labor augmenting only is a knife-edge case.[2]

This paper provides a possible resolution of this puzzle. Section 1 outlines the evidence supporting Fact 2 above, the trends and variation in capital shares. Section 2 begins the resolution of the puzzle. We introduce a new production function that is potentially consistent with the four facts documented above. Intuitively, factor shares and estimates of the elasticity of substitution are driven by short-run properties of the production function, while requirements for steady-state growth are driven by the shape of the production function in the long-run. We propose a production function that exploits this difference. The elasticity of substitution is very low in the short run, and it is hard to substitute capital and labor. In the long run, however, substitution possibilities open up and production takes the Cobb-Douglas form.

Of course, this last step is somewhat arbitrary. There are lots of reasons why the short-run elasticity of substitution in production might be smaller than the long-run elasticity. However, why should the long-run elasticity exactly equal one? This is the analog to the question asked earlier in the 1960s about the direction of technical change: why should it be labor-augmenting? The work by Acemoglu (2001) and a consideration of the evidence on computer prices and the price of equipment capital more

[2]The production functions for capital-ideas and labor-ideas must be parameterized "just so." In particular, let N denote the stock of labor-augmenting ideas. Then the cost of producing new labor-augmenting ideas relative to the cost of producing new capital-augmenting ideas must decline at exactly the rate $\dot{N}/N$. Plausible specifications — such as one in which the output good itself is the main input into the production of new ideas (in which case the relative cost of producing labor and capital ideas is constant) or the idea production function employed by Jones (1995a) to remove scale effects from the growth rate (in which case the relative cost of producing labor ideas declines with N^{ϕ}) — lead to a model that does not exhibit a steady state with a positive capital share.

generally suggests that there is no good answer to the labor-augmenting question. To be consistent with steady growth, then, the production function must be Cobb-Douglas in the long-run. But why?

Section 3 provides an answer to this question. Production techniques are ideas. A particular production technique is "appropriate" at a given mix of inputs, but if the input mix is changed, the production technique is substantially less effective: this is consistent with a short-run elasticity of substitution between inputs that is less than one. In the long run, however, an economy that wishes to produce with a different input mix (for example because of growth in capital per worker), may discover new production techniques that are appropriate at the new input mix. While the short-run elasticity of substitution between inputs is driven by the properties of a single technique, the long-run elasticity of substitution is governed by the distribution of ideas — the ease with which new ideas appropriate for a different input mix can be discovered.

Building on an insight from Houthakker (1955–1956), this section shows that if ideas are distributed according to a Pareto distribution, then the long-run elasticity of substitution between inputs is unity. That is, the long-run production function is Cobb-Douglas.

1. THE FACTS ABOUT CAPITAL SHARES

The conventional wisdom about capital's share of income being relatively constant dates back at least to Solow (1957) and was one of the facts emphasized by Kaldor (1961).[3] At least as an approximation, it is perhaps not too far from the mark for the United States and Great Britain. Gollin (2002) reports employee compensation as a share of GDP for these two

[3] Interestingly, Solow (1958) clarifies his stand on this issue by spending an entire paper expressing skepticism about the constancy of factor shares.

economies going back to 1935 and argues that the conventional wisdom is supported.

On the other hand, capital shares are clearly not constant over time, even in advanced economies, as several recent papers have documented. Blanchard (1997) emphasizes that capital shares in France, Germany, Italy, and Spain exhibit large increases starting in the early 1980s and continuing through the 1990s; the magnitude of the increase is approximately from 0.32 to 0.40. Acemoglu (2001) displays data on the capital share for the United States and France dating back to the early 20th century. While there is no strong trend in these data, the medium-term fluctuations can be substantial. Harrison (2003) looks at labor's share for a large number of countries using the national accounts data from the United Nations and shows that there are large movements over time for many countries.[4]

Apart from Harrison (2003), these papers typically do not focus on a large number of countries. One reason is the difficulty adjusting for self-employment income, an issue discussed at length by Gollin (2002). Gollin shows that differences in employee compensation as a share of GDP across countries are largely explained by differences in self-employment rates. When he corrects the employee compensation shares for differences in self-employment, he finds significantly smaller differences in labor shares across countries.[5]

Still, in some ways this is a question of the glass half-full versus half-empty. While the shares move closer together when one makes Gollin's correction, they are still substantially different. The purpose of this section of the paper is to present some evidence on capital's share for OECD

[4]See also Caballero and Hammour (1998), Batini, Jackson and Nickell (2000), and Bentolila and Saint-Paul (2003).

[5]The working paper version of Gollin's paper suggested that the same finding applied over time: trends in employee compensation as a share of GDP often disappear when one corrects for self-employment. However, he had limited data in the time series dimension.

countries. We begin with the United Nations National Accounts data on employee compensation as a share of GDP. We then follow Gollin's preferred correction by dividing this share by the fraction of total employment accounted for by employees. The implicit assumption is that the average wage of the self-employed is equal to the average wage of employees. Finally, we compute the capital share as one minus this labor share.[6]

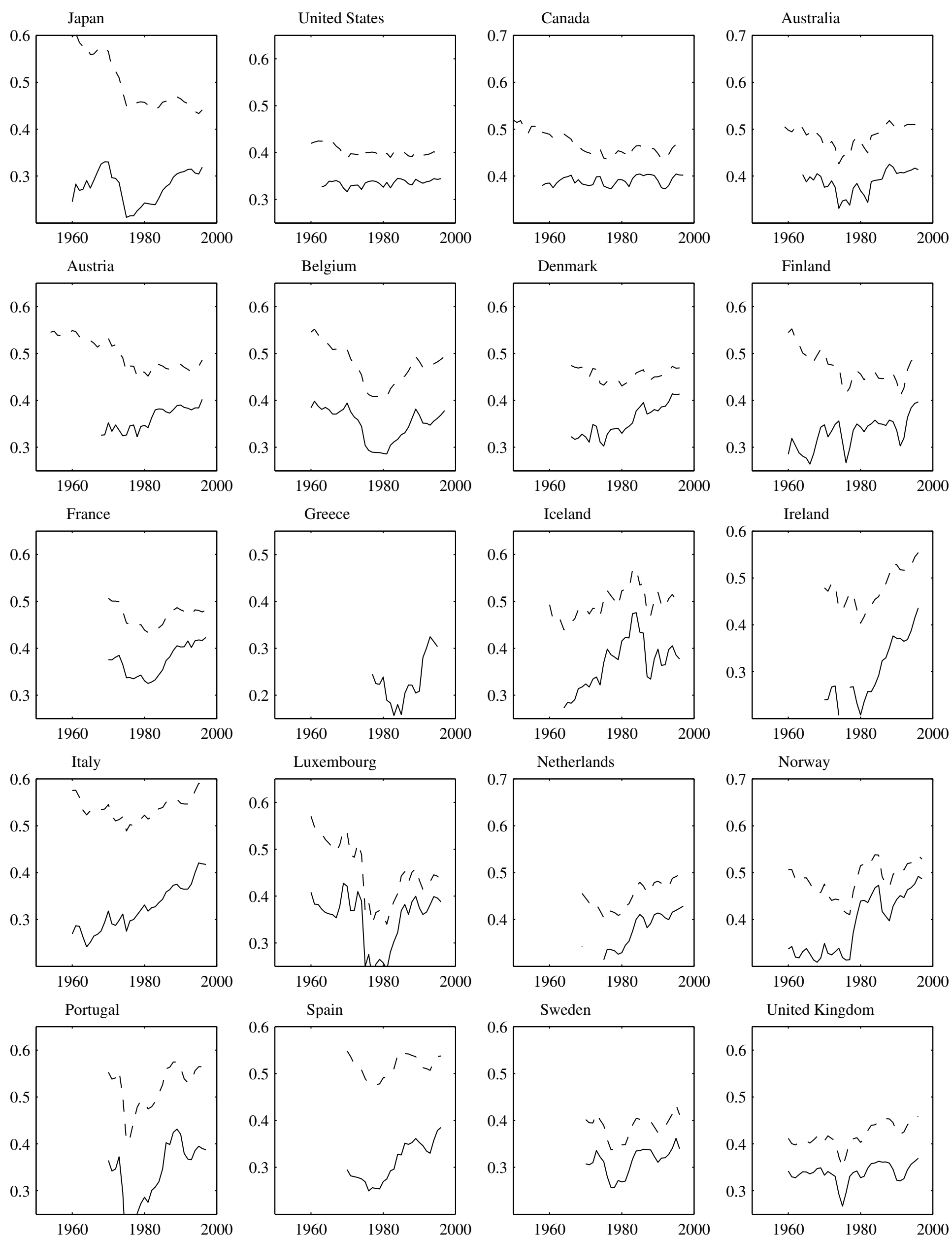
The evidence on capital's share for OECD countries is reported in Figure 1. The solid line in the figure is the preferred measure, which incorporates the correction for self-employment. The dashed line is the naive measure, constructed as one minus employee compensation divided by GDP.

Several features of the figure stand out. First, while the capital share for the United States and Canada are relatively smooth, this is not the case for capital shares in most other countries. In many countries there is an upward trend in the capital share, and the medium-frequency movements are often large. Denmark's capital share, for example, rises from about 0.3 to about 0.4 over the last quarter of the century. And the rises identified by Blanchard (1997) for France, Italy, and Spain are noteworthy. The large negative trend in the naive version of the capital share for Japan gets undone by a large decline in the fraction of the population that is self-employed, but there are still significant movements over time.

Another useful piece of evidence on variation in capital shares comes from looking at industry-level data in the United States, as reported in Table 1. Of the 35 two-digit industries, 22 (63%) exhibit significant trends in the capital share of at least a tenth of a percent per year, while 16 (46%) exhibit significant trends of at least two-tenths of a percent per year. Overall, there is a slight upward trend in the capital share in manufacturing and a slight downward trend in services. Agriculture, Petroleum, and Leather

[6] See Appendix B for more detail on the data sources.

FIGURE 1. Capital Shares in OECD Countries



Note: Solid line = With self-employment correction. Dashed line = Without correction. Capital shares are calculated as one minus the labor share. For the dashed line, the labor share is employee compensation as a share of GDP. For the solid line, this labor share is corrected by dividing by the employee share of total employment, as suggested by Gollin (2002). See the Appendix B for more detail.

TABLE 1.
Capital Shares for 2-Digit U.S. Industries

Industry	1960	1970	1980	1990	1996	Trend	t-stat
Motor vehicles	40.9	29.3	23.7	23.4	21.7	-0.5425	-5.52
Stone, clay, glass	32.1	24.1	19.1	18.4	24.2	-0.3730	-8.58
Other Services	32.7	32.4	29.7	28.1	25.2	-0.2544	-8.78
Primary metal	31.1	24.7	21.6	22.2	22.3	-0.2113	-4.72
Transportation	31.1	27.2	24.8	22.9	25.1	-0.1876	-8.01
Instruments	17.9	19.4	17.0	18.0	13.4	-0.1573	-3.79
Machinery, non-electi	23.0	25.1	22.4	24.4	26.8	-0.1452	-3.62
Finance Insurance and	59.3	59.0	54.7	55.2	58.1	-0.1273	-2.12
Electric utilities	68.1	65.2	60.9	66.0	64.3	-0.0884	-2.70
Rubber and misc plast	23.7	18.3	16.0	19.0	25.7	-0.0851	-2.64
Chemicals	45.6	40.5	35.2	44.0	47.8	-0.0728	-1.48
Transportation equipm	8.9	11.1	9.3	9.2	12.7	-0.0383	-1.34
Trade	20.8	21.9	22.2	19.9	22.7	-0.0276	-1.42
Textile mill products	21.7	25.3	18.6	22.9	25.3	-0.0015	-0.04
Oil and gas extractio	72.3	72.2	73.2	78.9	72.1	0.0037	0.10
Construction	8.5	12.3	14.2	12.4	9.4	0.0080	0.35
Furniture and fixture	15.8	16.2	16.8	14.1	21.5	0.0117	0.34
Gas utilities	67.7	66.2	65.4	69.0	71.1	0.0747	3.85
Paper and allied	35.5	30.1	28.7	38.1	40.7	0.0817	1.59
Printing, publishing	20.4	21.4	20.6	25.4	23.4	0.0993	3.36
Communications	49.0	48.0	45.3	53.2	51.7	0.1101	2.59
Coal mining	22.0	38.6	22.9	32.1	34.3	0.1121	1.22
Non-metallic mining	40.8	39.3	47.0	43.5	49.8	0.1505	3.38
Metal mining	41.7	37.8	43.6	46.8	47.6	0.2383	3.19
Government enterprise	31.9	30.5	42.2	37.9	48.4	0.2415	4.16
Lumber and wood	22.0	21.9	27.5	29.9	31.3	0.2648	4.76
Apparel	9.6	14.0	14.3	19.4	18.4	0.2868	13.31
Fabricated metal	15.8	16.8	22.2	25.7	36.0	0.3070	9.20
Electrical machinery	21.5	19.5	21.3	33.7	42.6	0.3916	5.06
Agriculture	24.5	30.2	37.6	41.3	41.5	0.4151	6.36
Tobacco	61.7	64.8	62.8	77.8	75.4	0.4183	6.22
Food and kindred prod	26.8	29.7	28.1	42.8	46.1	0.4558	10.03
Petroleum and coal pr	43.2	45.6	69.8	67.7	63.1	0.5533	6.46
Misc. manufacturing	17.5	18.7	20.5	40.8	41.1	0.5875	10.71
Leather	12.8	12.8	24.0	34.9	45.9	0.8074	12.50
Manufacturing	26.1	24.5	24.0	30.0	33.2	0.0808	2.97
Services	33.8	34.5	33.1	32.5	32.7	-0.0822	-3.33
Total	30.9	31.0	31.6	33.1	33.6	0.0347	2.32

Note: These capital shares are calculated as payments to capital as a share of value-added, using Dale Jorgenson's data on 35 2-digit industries. According to Jorgenson, Ho and Stiroh (2002, p. 11), these data include a correction for self-employment. Data downloaded from Jorgenson's web page on 11/28/01.

are examples of industries with large positive trends in the capital share (for example, the capital share in agriculture rises from about 25 percent to more than 40 percent). Stone, clay and glass, Primary metals, Motor vehicles, and Transportation are examples of industries with large negative trends. More generally, this evidence on sectoral variation is supported by Bentolila and Saint-Paul (2003) who document large sectoral movements in labor shares for the OECD countries.

Overall, both the country-level and the industry-level evidence — in this paper and in other papers — sharply call into question the stylized fact that capital shares are smooth, stable, and do not exhibit medium-run trends. This fact is roughly true for some countries, but it is strongly contradicted in others. Even in the United States, a country typically used to justify the stylized fact, the industry-level evidence suggests there are substantial changes in capital shares over time.

2. A RESOLUTION OF THE PUZZLE

At first glance, matching the four facts that began this paper appears difficult. Steady growth in the presence of investment-specific technological change, like that which drives down equipment prices, requires a Cobb-Douglas production function. However, the evidence of large movements in capital shares documented in the previous section is inconsistent with the constant-share prediction of a Cobb-Douglas function, as are the empirical estimates of elasticities of substitution in the literature.

One possible resolution of this puzzle, at least if one is willing to ignore the not-entirely-persuasive evidence on elasticities of substitution, is that production functions are Cobb-Douglas but wages are not equal to marginal products. In this case, changes in factor shares could reflect changes in the markup or changes in bargaining power. Blanchard and Giavazzi (2002) explore an explanation along these lines.

If one wishes to maintain the connection between factor prices and marginal products, a second possible resolution is suggested by the insight that capital shares and substitution elasticity estimates are driven by the short-run properties of the production function. In contrast, the requirement that production take a Cobb-Douglas form in order for steady growth to be possible in the presence of investment-specific technological change is really a statement about the shape of the production function in the long run.

What is needed, therefore, is a production function that (a) makes a distinction between the short-run elasticity of substitution (σ^{SR}) and the the long-run elasticity of substitution (σ^{LR}), and (b) has a long-run elasticity of substition equal to one ($\sigma^{LR} = 1$).

A class of models that meets the first requirement are the putty-clay vintage capital models of Caballero and Hammour (1998) and Gilchrist and Williams (2000). In these models, the technology level and the capital-labor ratio get embodied in a machine when the machine is built, and the technology for a given machine is ex post Leontief. Ex ante, however, the capital-labor ratio is a choice variable and a more flexible production function applies. If an economy with an existing stock of machines wants to adjust its aggregate capital-labor ratio, then, the presence of a large number of Leontief machines gives it a relatively low elasticity of substitution in the short-run. Over the long-run, this elasticity is much higher. These putty-clay vintage capital models, then can potentially reconcile the body of facts that began the paper, at least if we are willing to assume that the long-run elasticity in the model is equal to one (see Section 3 below).[7]

[7]Blanchard (1997) has a related model. Instead of the putty-clay structure, he assumes a firm consists of one unit of capital and n units of labor. There are ad hoc costs of adjustment for labor and capital (the number of firms) that create a wedge between the marginal revenue product of labor and the wage, allowing the capital and labor shares to vary over time even if the technology is Cobb-Douglas.

The remainder of this section will consider briefly an alternative story for this distinction between σ^{SR} and σ^{LR}, and there are surely others. The story told here is useful for three reasons. First, it is perhaps an interesting economic alternative. Second, it is much more tractable than the putty-clay vintage capital approach. And third, it is related to some microfoundations that follow later in the paper. Note, however, that nothing in the paper requires this particular story to be true, and I am equally happy with the bargaining explanation and the putty-clay explanation.

The story is as follows. Consider the following production function:

$$\begin{aligned} Y_t &= F(K_t, L_t; K_t^*, L_t^*, A_t^*) \\ &= \left(\alpha \left(\frac{K_t}{K_t^*}\right)^\rho + (1-\alpha)\left(\frac{L_t}{L_t^*}\right)^\rho\right)^{1/\rho} K_t^{*\alpha}(A_t^* L_t^*)^{1-\alpha}, \qquad (1) \end{aligned}$$

where $\rho < 0$. In this setup, quantities without an asterisk represent output, capital, etc., used in the economy. Quantities with an asterisk are parameters of the production technology. For example, K^* and L^* indicate the levels of capital and labor that are most "appropriate" for this technology at time t, and A^* denotes the productivity level of this technology.[8]

The CES term in equation (1) represents the short-run production function. If the economy changes its inputs, this CES term indicates how output varies. With $\rho < 0$, this production function exhibits a (short-run) elasticity of substitution that is less than one. On the other hand, at the "appropriate" levels of the inputs, this CES term is equal to one and the level of output is determined solely by the Cobb-Douglas function of K^* and L^*.

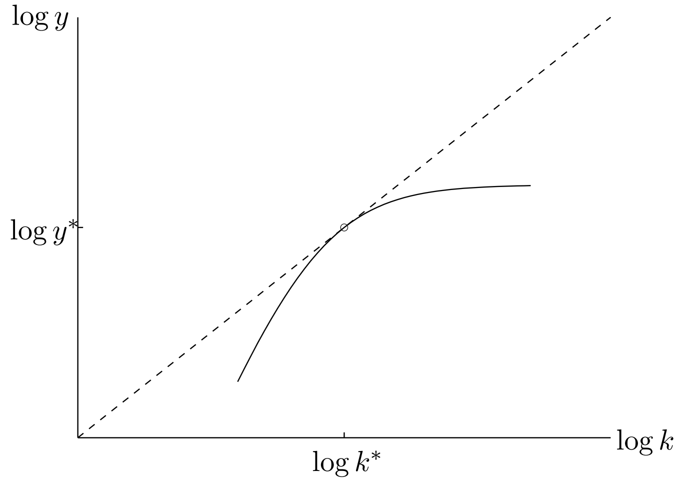
It is useful to rewrite this production function in per worker terms, as

$$y_t = \left(\alpha \left(\frac{k_t}{k_t^*}\right)^\rho + 1 - \alpha\right)^{1/\rho} y_t^*, \qquad (2)$$

where $y_t \equiv Y_t/L_t$, $k_t \equiv K_t/L_t$, $k_t^* \equiv K_t^*/L_t^*$, and $y_t^* \equiv k_t^{*\alpha} A_t^{*1-\alpha}$. When the capital-labor ratio is equal to its appropriate value k^*, output

[8] This use of "appropriate" technologies builds on Basu and Weil (1998).

FIGURE 2. The Production Function



The solid line plots the production function for a fixed value of k^*. The dashed line represents the Cobb-Douglas relation with slope α.

per worker is equal to y^*. When the capital-labor ratio varies from this appropriate value, output per worker varies from y^* according to a CES relation with a substitution elasticity less than one.

Figure 2 shows this production function graphically. Notice that the axes in this figure are $\log y$ and $\log k$. The dashed line, then, shows a Cobb-Douglas production relation with slope α. The solid line plots $\log y$ as a function of $\log k$ for a given level of k^*. Because the elasticity of substitution is less than one, this production function lies below the Cobb-Douglas relation everywhere, except obviously at the point at which the technology is appropriate.

Notice also that the slope of this production function in log-space corresponds to the capital share $\partial \log y / \partial \log k = \frac{\partial y}{\partial k}\frac{k}{y} = \frac{1}{1+\frac{1-\alpha}{\alpha}(\frac{k_t}{k_t^*})^{-\rho}}$. Below the appropriate capital-labor ratio, the capital share is higher than α, and above k^* the share is lower than α.[9] If the economy increases its capital-labor ratio for a given technology, the capital share declines because of the low substitution elasticity. In the long run, however, suppose that k^* itself rises. In this case, the decline in the capital share that would occur if the same technology were used is offset by an improvement in technology. By using a new technology that is appropriate at the higher capital-labor ratio, the capital share can be supported.

3. WHY MIGHT PRODUCTION FUNCTIONS BE COBB-DOUGLAS IN THE LONG RUN?

The previous section suggests a resolution to the puzzle that began the paper. For various reasons, the elasticity of substitution between capital and labor may be substantially lower in the short run than it is in the long run. If it happens to be equal to one in the long run, so that the long-run production function is Cobb-Douglas, then the facts can be reconciled. Factor shares and estimates of the elasticity of substitution are driven by the short-run production function, while the possibility of a steady state is ensured by the long-run Cobb-Douglas form. Of course a key unresolved question in this story is why the long-run elasticity of substitution should happen to equal unity. Why might production functions take the Cobb-Douglas form in the long run?

Consider for a moment what a production function is. At its most primitive level, one could imagine a Leontief production technology that says "for each unit of labor, take k^* units of capital, and you will get out y^* units of output." The values k^* and y^* are parameters of this production technology, and this production technology might be thought of as an idea. If one wants to produce with a very different capital-labor ratio, it may well be the case that this idea is not particularly helpful, and one needs to discover a

[9] Recall that since the elasticity of substitution is less than one, we have $\rho < 0$.

new idea "appropriate" at the higher capital-labor ratio. Notice that one can replace the Leontief structure with a production technology that exhibits a low elasticity of substitution, and this statement remains true: to take advantage of a substantially higher capital-labor ratio, one really needs a new technique targetted to that capital-labor ratio. According to this view, the standard production function that we write down, mapping the entire range of capital-labor ratios into outputs, is a reduced form. It is not a single technology, but rather represents the substitution possibilities across different production techniques. The elasticity of substitution for this long-run production function depends on the ease with which new techniques that are appropriate at higher capital-labor ratios can be discovered and implemented. That is, it depends on the distribution of ideas.

This insight is formalized below. We develop conditions under which the long-run production function is asymptotically Cobb-Douglas and interpret these conditions.

3.1. Setup

Let a particular production technique (call it technique i) be indexed by two parameters, a_i and b_i. With this technique, output can be produced according to a production function

$$Y = F(b_i K, a_i L). \tag{3}$$

Following the motivation given above, we assume that $F(\cdot, \cdot)$ exhibits a low elasticity of substitution between K and L.

This production function can be rearranged to give

$$Y = a_i L F(\frac{b_i K}{a_i L}, 1), \tag{4}$$

so that in per worker terms we have

$$y = a_i F(\frac{b_i}{a_i} k, 1). \tag{5}$$

Now, define $y_i \equiv a_i$ and $k_i \equiv a_i/b_i$, and the production function can be written as

$$y = y_i F(\frac{k}{k_i}, 1). \tag{6}$$

If we choose our units so that $F(1,1) = 1$, then we have the nice property that $k = k_i$ implies that $y = y_i$; an example of such a relation is that given earlier in equation (2). Therefore, we can think of technique i as being indexed by a_i and b_i, or, equivalently, by k_i and y_i.

As described above, in the long-run, the shape of the production function is driven by the discovery of new techniques rather than by the shape of the short-run production function that applies for a single technique. For this reason — and because it results in a more tractable problem that yields analytic results — we will make the extreme assumption that the short-run production function is Leontief. In the simulation results that follow the theory, we show how this requirement can be relaxed.

ASSUMPTION 3.1. *The production function for a given technique is Leontief. That is,*

$$Y = F(b_i K, a_i L) = \min\{b_i K, a_i L\}.$$

With this assumption, the per worker production function in (6) simplifies nicely: in per worker terms, we have

$$y = y_i \min\{\frac{k}{k_i}, 1\}.$$

That is, k_i units of capital for each worker produces y_i units of output for each worker, and with a given technique, additional capital is unproductive.

Research results in the discovery of new techniques. Let N denote the total number of techniques that have been discovered, and let i denote the i^{th} technique. This discovery process can be thought of as the drawing of balls out of a hat, where each ball has two numbers written on it, an a_i and a b_i. We make the following important assumption about the distribution of these numbers:

ASSUMPTION 3.2. *A research draw i results in the discovery of a new technique i, characterized by two numbers, a_i and b_i. These numbers are random variables drawn from independent Pareto distributions:*

$$a_i \sim G_1(a) = 1 - \left(\frac{a}{\gamma_a}\right)^{-\alpha}, \;\; a \geq \gamma_a > 0, \;\; \alpha > 0 \tag{7}$$

and

$$b_i \sim G_2(b) = 1 - \left(\frac{b}{\gamma_b}\right)^{-\beta}, \;\; b \geq \gamma_b > 0, \;\; \beta > 0. \tag{8}$$

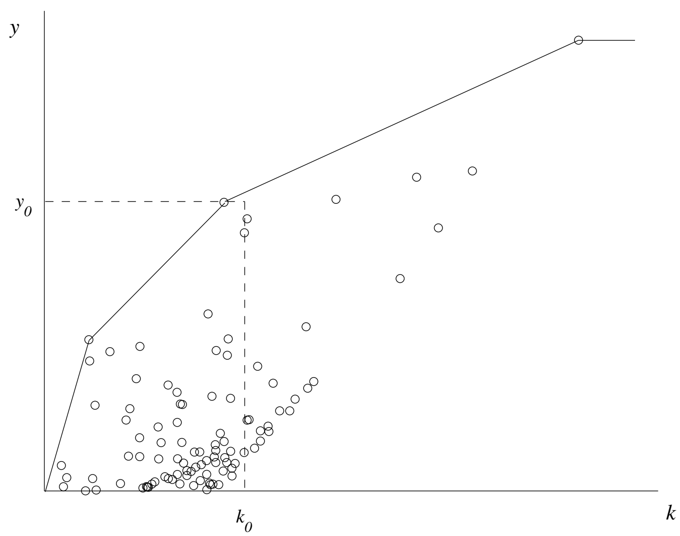
This research technology features two desirable properties: First, a technique that is parameterized by a high level of capital per worker $k_i = a_i/b_i$ tends to produce a high level of output per worker $y_i = a_i$. Second, such productive techniques are relatively hard to come by; it is easier to discover less productive techniques.

These properties yield a familiar-looking production set, as shown by the example in Figure 3. The circles denote techniques that have been discovered — the set of (k_i, y_i) pairs — and the lines joining the circles at the edge of the production set fill in the convex hull (these are production outcomes that could be obtained by using two techniques at once in varying proportions).

The key question we'd like to answer is this. With this setup, what is the functional relation $y^* = f(k^*)$? That is, given a particular level of capital per worker k^*, what is the maximum amount of output per worker that can be produced using the available techniques?

Again so that we can derive analytic results, we will actually be slightly more restrictive. First, we will answer this question assuming only one

FIGURE 3. An Example of the Long-Run Production Function



technique can be used at a time. In Figure 3, for example, the answer to this question if $k^* = k_0$ is given by y_0 (notice that free disposal is allowed, so that we can throw away some capital to take advantage of a superior technique). Second, while we allow free disposal of capital, we assume that all labor must be employed.[10] We will show in the simulations that follow the theory that these simplifying assumptions are not crucial.

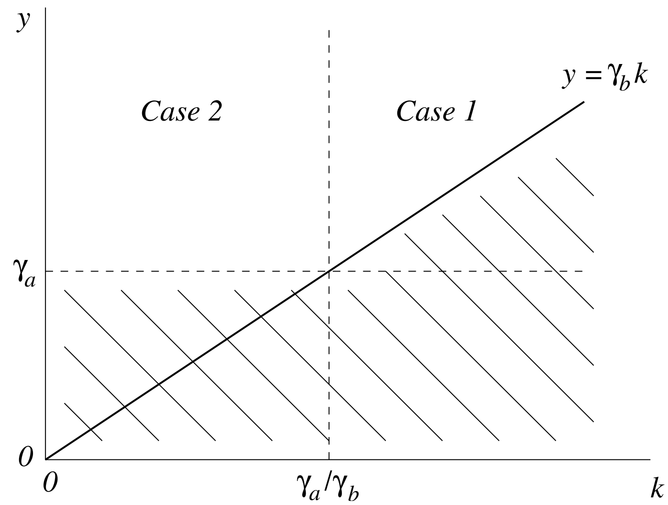
3.2. Derivation

Formalizing these insights, let this production function $y^* = f(k^*)$ be defined as follows

$$y^* = f(k^*) = \max_i \{a_i : \frac{a_i}{b_i} \leq k^*\}, \quad 0 \leq i \leq N. \tag{9}$$

[10]By forcing firms to use a single technique, they may want to throw away labor or capital. If they are allowed to choose, then we must value the labor and capital, so that prices and general equilibrium come into the problem, making it much more complicated. Importantly, notice that if we were not ignoring the convex hull, this would not be an issue. Assuming that only one factor can be thrown away mimics the full problem in that prices do not enter the calculation.

FIGURE 4. The Model's Two Cases



That is, y^* is given by the highest level of a_i from among the set of techniques with $k_i \leq k^*$.

The argument proceeds in the following way. We first characterize the distribution of the a_i's in equation (9). Then we apply a result from Extreme Value Theory to characterize the distribution of the maximum. This yields the Cobb-Douglas result.

To begin, define the conditional distribution of the a_i's as

$$F(y;k^*) \equiv \text{Prob}\,\{a_i \leq y \mid \frac{a_i}{b_i} \leq k^*\}. \tag{10}$$

Given the setup of the model, there are two separate cases that need to be considered. These are described in Figure 4. All of the draws of techniques satisfy two inequalities. First $y_i \geq \gamma_a$. This is straightforward since $y_i \equiv a_i$ and $a_i \geq \gamma_a$ as part of Assumption 4.1. Second, $y_i/k_i \geq \gamma_b$. This is true since $y_i/k_i = b_i$ and $b_i \geq \gamma_b$. This means that all techniques that are discovered will lie either in the region labeled "Case 1" or the region labeled "Case 2" in Figure 4.

For expositional purposes, we will develop the results first for Case 1 and then show that one gets an identical long-run production function in Case 2. To begin, we have the following result (all proofs are given in Appendix C):

PROPOSITION 3.1. *For $k^* > \gamma_a/\gamma_b$ and $y \geq \gamma_a$ (Case 1), the distribution of a_i conditional on the set of techniques that can be used with capital per worker k^* is given by*

$$F(y; k^*) = 1 - \theta(k^*) y^{-(\alpha+\beta)},$$

where

$$\theta(k) \equiv \frac{\frac{\alpha}{\alpha+\beta}\gamma_a^{\alpha}\gamma_b^{\beta}k^{\beta}}{1 - \frac{\beta}{\alpha+\beta}\left(\frac{\gamma_a}{\gamma_b}\right)^{\alpha} k^{-\alpha}}.$$

This proposition says that the distribution of output per worker that can be produced using appropriate technologies with capital per worker below k^* is a Pareto distribution. It is straightforward to show that $\theta'(k) > 0$ in the relevant range: if a particular technique is associated with a high level of capital per worker, it is more likely to produce a high value of output per worker.

The next proposition characterizes the distribution of $k_i^* \equiv a_i/b_i$ itself.

PROPOSITION 3.2. *Define $G(k^*) \equiv \text{Prob}\{a_i/b_i \leq k^*\}$. Then for $k^* > \gamma_a/\gamma_b$ (Case 1),*

$$G(k^*) = 1 - \frac{\beta}{\alpha+\beta}\left(\frac{\gamma_a}{\gamma_b}\right)^{\alpha} k^{*-\alpha}.$$

That is, the distribution of the "appropriate" capital requirements for ideas is also Pareto. The associated density function is decreasing in k, so that it is more difficult to find ideas that work with a larger amount of capital per worker.[11]

[11] One might wonder about the intuition for why the exponent in the distribution depends only on α rather than on some combination of α and β. Since $k_i \equiv a_i/b_i$, in order to get

Together, these two propositions characterize the research process. It is hard to find ideas that use a large amount of capital per worker. But conditional on finding one of these ideas, it is likely to produce a large amount of output per worker when used.

The production relation in equation (9) says that y^* is an extreme value of the $F(y;k^*)$ distribution, as suggested back in Figure 3. Because this distribution has an upper tail that is a power function (i.e. because the upper tail behaves like a Pareto distribution), one can apply Theorem 2.1.1 from Galambos (1978), which says that the maximum extreme value from a distribution with a power function tail, appropriately normalized, obeys a Frechet distribution:

PROPOSITION 3.3. *Let $N(k^*)$ be the number of techniques that require capital per worker less than k^*. Then for $k^* > \gamma_a/\gamma_b$ (Case 1),*

$$\lim_{N(k^*)\to\infty} \text{Prob}\left\{(N(k^*)\theta(k^*))^{-\frac{1}{\alpha+\beta}}\, y^* \le x\right\} = \exp(-x^{-(\alpha+\beta)}). \tag{11}$$

Or, stated differently,

$$(N(k^*)\theta(k^*))^{-\frac{1}{\alpha+\beta}} y^* \overset{a}{\sim} \textit{Frechet}(\alpha+\beta). \tag{12}$$

This proposition has the following interpretation. The random variable $y^*(k^*)$ is the maximum amount of output that can be produced with capital k^*. As the number of techniques that are relevant at k^* grows, this output goes to infinity. Hence, we need to normalized by some function of $N(k^*)$, and it turns out that the right normalization is $N(k^*)^{1/(\alpha+\beta)}$. Of course, even the normalized output will still depend on the amount of capital. But one can imagine that there is some function of k^* that we could divide by

a technique appropriate for a high level of k, we need a high value of a_i and a low value of b_i. However, $b_i \ge \gamma_b$ by assumption, so large values of k_i ultimately depend on large values of a_i.

to eliminate this dependence. This function turns out to be $\theta(k^*)^{1/(\alpha+\beta)}$. By appropriately normalizing y^*, we get a stable asymptotic distribution for the normalized random variable.

To push the interpretation further, let ϵ be an i.i.d. random variable drawn from this Frechet distribution. Then the production possibilities frontier can asymptotically be thought of as

$$y^* = (N(k^*)\theta(k^*))^{1/(\alpha+\beta)}\epsilon.$$

The last step in getting the production function is to consolidate the terms that depend on k^*. Since N denotes the total number of production techniques that have been discovered,

$$plim_{N\to\infty}\frac{N(k^*)}{N} = \text{Prob}\,\{\frac{a_i}{b_i} \le k^*\} = G(k^*) = 1-\frac{\beta}{\alpha+\beta}\left(\frac{\gamma_a}{\gamma_b}\right)^{\alpha} k^{*-\alpha}.$$

That is, the fraction of ideas that are relevant at k^* in a sample of N draws is, roughly speaking, $G(k^*)N$. But $G(k^*)$ is just the denominator in $\theta(k^*)$, so the long-run production function behaves asymptotically like

$$y^* = AN^{\frac{1}{\alpha+\beta}}k^{*\frac{\beta}{\alpha+\beta}}\epsilon \tag{13}$$

or, multiplying both sides by the number of workers L,

$$Y^* = AN^{\eta/\beta}K^{*\eta}L^{*1-\eta}\epsilon$$

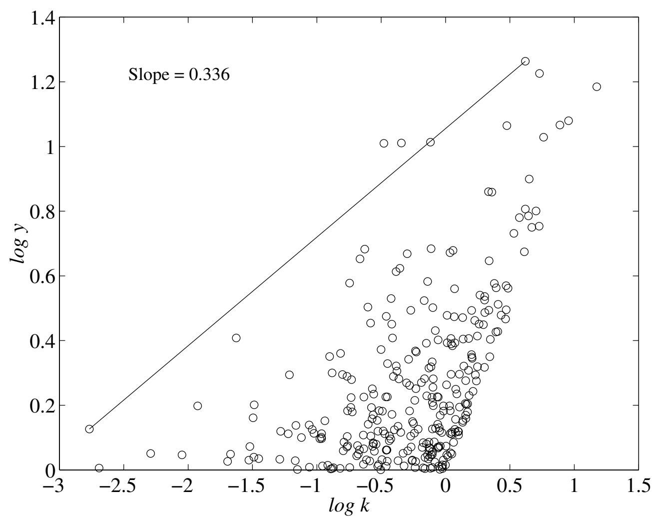
where $\eta \equiv \beta/(\alpha+\beta)$ and $A \equiv \left(\frac{\alpha}{\alpha+\beta}\gamma_a^{\alpha}\gamma_b^{\beta}\right)^{\frac{1}{\alpha+\beta}}$.

We have derived this result for the case where $k^* \ge \gamma_a/\gamma_b$, i.e. for Case 1 in Figure 4. As Proposition C.1 in Appendix C shows, however, one gets exactly the same production function in Case 2.

3.3. Simulations

Figure 5 begins by presenting a simple numerical simulation of the production function result. In the simulation, 300 techniques are drawn from

FIGURE 5. Simulation of the Long-Run Production Function



Note: In the simulation, $N = 300$, $\alpha = 4$, $\beta = 2$, and $\gamma_a = \gamma_b = 1$. All 300 techniques are plotted in log-log space, and a slope is computed between the technique that generates the highest output and the technique that employs the least amount of capital.

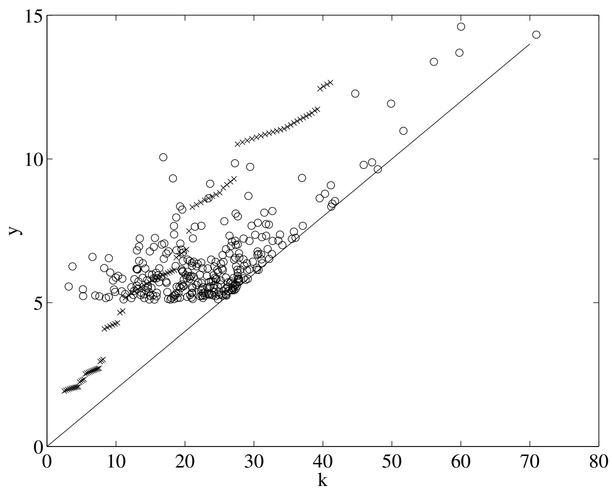
the Pareto distributions, with $\alpha = 4$ and $\beta = 2$. According to the model, the convex hull of the production function should approximately be Cobb-Douglas, with a capital elasticity of $\beta/(\alpha + \beta) = 1/3$. This is indeed what is found.

Next, we turn to a full simulation of the model. To do this in the simplest way, we embed this approach to modeling production and ideas in a standard Solow growth model. In addition, we relax the assumption that individual techniques are Leontief and specify production as a CES function. The full setup looks like:

$$Y_t = F(b_i K_t, a_i L_t) = (\alpha(b_i K_t)^\rho + (1 - \alpha)(a_i L_t)^\rho)^{1/\rho} \tag{14}$$

$$K_{t+1} = (1 - \delta)K_t + sY_t \tag{15}$$

FIGURE 6. Simulation of the Production Function



Note: Circles indicate ideas, plus signs indicate capital-output combinations that are actually used. The model is simulated for 100 periods with $N_0 = 50$, $\alpha = 5$, $\beta = 2.5$, $n = .10$, $\gamma_a = 1$, $\gamma_b = 0.2$, $k_0 = 2.5$, $s = 0.2$, $\delta = .05$, and $\rho = -1$.

$$N_t = N_0 e^{nt} \tag{16}$$

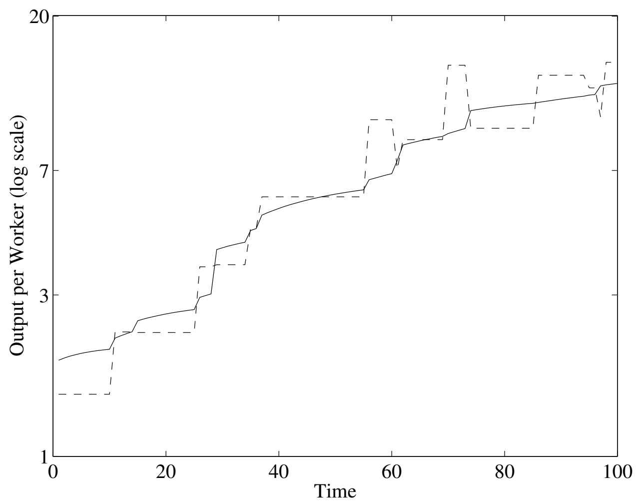
where N_t denotes the total number of techniques that have been discovered as of date t, assumed to grow exogenously at rate n. The research process is that specified earlier, and from among the N_t techniques available at time t, we choose the technique that produces the highest amount of output per worker.[12]

We simulate this model for 100 years and plot the results in several figures.[13] Figure 6 shows a subset of the nearly 1 million techniques that

[12]Computing the convex hull of the overlapping CES production functions seems to be an extremely hard problem. To simplify, we continue to assume that firms can use only a single technique at each point in time. To find the technique they use (from among nearly 1 million techniques by the end of the simulation), we first compute the convex hull of the (k_i, y_i) points. Then, considering only these points on the convex hull, we find the highest amount of output that can be produced using the full CES production function with each technique.

[13]The parameter values used in the simulation are listed in the notes to Figure 6.

FIGURE 7. Output per Worker over Time

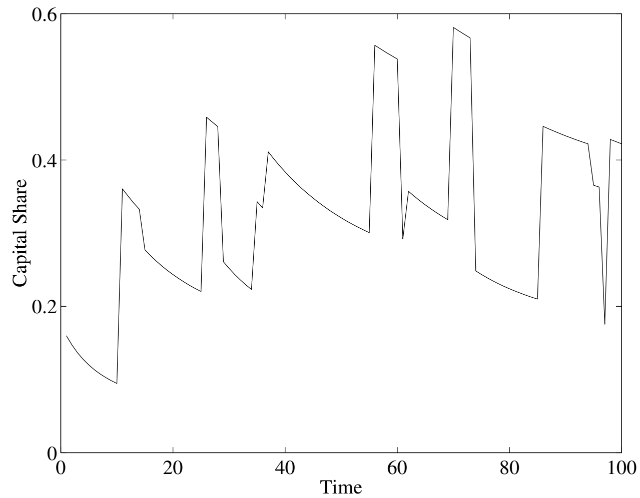


Note: See notes to Figure 6. The solid line plots output per worker y_t at each date. The dashed line in the figure plots y_{it}, i.e. the level of y_i corresponding to the technique used at each point in time.

are discovered over these 100 periods. In particular, we plot only the 300 points with the highest values of y (these are shown with circles "o"). Without this truncation, the lower triangle in the figure that is currently blank but for the plus signs is filled in as solid black. In addition, the capital-output combinations that are actually used in each period are plotted with a plus sign ("+"). When a particular technique is used for a large number of periods, the points trace out the CES production function. Then, as the economy switches to a new technique, the capital-output combinations jump upward.

Figure 7 shows output per worker over time, plotted on a log scale. The average growth rate of output per worker in the simulation is 1.92 percent, close to the theoretical value of 2 percent implied by the parameter

FIGURE 8. The Capital Share over Time



Note: See notes to Figure 6.

values, given by n/α.[14] The dashed line in the figure plots the level of y_i corresponding to the technique used at each point in time. Sometimes the economy is above this level, and sometimes below, with the obvious implication for the capital share, explored next.

Figure 8 plots the capital share $F_K K/Y$ over time. Even though the economy grows at a stable average rate over time, the capital share exhibits large movements over time. The movements upward are jumps, while the movements downward are gradual. This means the model as is cannot provide a complete explanation for the behavior of capital shares (but that is not a primary goal of the paper).

[14] We compute the average growth rate by dropping the first 20 observations (to minimize the effect of initial conditions) and then regressing the log of output per worker on a constant and a time trend.

3.4. Equipment Prices and Capital-Augmenting Technology

To explain the facts that began this paper — including the steady decline in equipment prices — one needs a production function that is Cobb-Douglas in the long-run. We have provided a theory suggesting why this might be the case. However, the extent to which this theory really is consistent with capital-augmenting technological change is worth discussing in more detail.

First, notice that the production function for a given technique, as assumed in equation (3), is

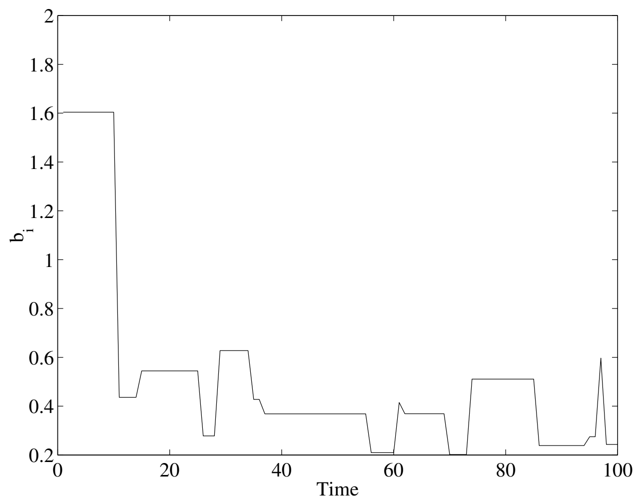
$$Y = F(b_i K, a_i L).$$

If we apply the steady-state growth theorem to this production function, then we know that in the long-run either b_i must be constant or the production function $F(\cdot,\cdot)$ must be Cobb-Douglas. Since we've assumed that the production function for an individual technique is not Cobb-Douglas, we are left with the result that b_i cannot grow in the long run if we are to have steady-state growth. But then how are we able to get a steady state in the model in Section 3?

The answer can be seen in Figure 9. Recall that $y_i \equiv a_i$ and $k_i \equiv a_i/b_i$. Therefore $y_i/k_i = b_i$. If y_i/k_i is to be stationary, as one would expect would be required in steady state, then the value of b_i for the technique that is chosen at each date must also be stationary, and in fact Figure 9 suggests this is the case. Notice that this is true despite the fact that the largest value of b_i discovered goes to infinity as the number of draws goes to infinity.

What is going on? The intuition is most easily seen if the production function for a given technique has a very low elasticity of substitution, e.g. Leontief. On the one hand, a lower $k_i \equiv a_i/b_i$ is better since a firm can always throw k away to take advantage of a more productive technique; this says a high value of b is desirable. On the other hand, it is inefficient to

FIGURE 9. b_i^* Over Time



Note: The figure plots the value of b_i for the technique that is chosen at each date. See also notes to Figure 6.

throw capital away, so an economy with a given capital-labor ratio k would like to use a technique that is appropriate at that level, i.e. $k_i = k$. For a large value of k, this is more easily accomplished with a low value of b_i. At some level, it is the offsetting of these two forces that leads to a stationary distribution of b_i being chosen.

The question this issue obviously raises is whether or not the model can be consistent with falling relative prices of equipment, such as computers. This is especially important since falling equipment prices is one of the motivations for the paper in the first place.

The answer to the question is "Yes," and the answer itself raises some interesting issues. Suppose that a particular computer technology, say the Pentium III processor, is a particular technique — a (k_i, y_i) pair — and a new computer technology, say the Itanium processor, is a new technique. It should be clear intuitively (and playing around with some algebra can formalize this intuition) that the introduction of a better technique leads the price of an existing technique to fall.

This approach suggests treating new computers as new ideas and implies that the price of a computer incorporates the productivity improvement associated with the idea. In equilibrium it may well be the case that the only ideas that get used are those consistent with a stationary y_i/k_i, and yet the quality-adjusted price of a particular idea can decline over time. Moreover, ideas with higher and higher values of b_i are always being discovered, so there is no sense in which research gets restricted to only the labor-augmenting direction.

This reasoning has an important implication for the current hedonic approach to prices used by the BEA in the construction of the National Income and Product Accounts. The current approach treats better computers as equivalent to more computers, so technological change leads mechanically to a rising real ratio of equipment to GDP. The approach here suggests that any given idea has a pair of numbers — an a_i and a b_i associated with it, and that these numbers are characteristics that determine prices. There is no reason that these characteristics should be lumped into a single dimension.

4. DISCUSSION

The basic theory result in this paper is that the long-run production function is asymptotically Cobb-Douglas under the crucial assumption that the underlying parameters characterizing the production techniques obey Pareto distributions. A number of remarks on this result, as well as some important directions for further research, are appropriate.

The first main comment is that this result is related to a classic result by Houthakker (1955–1956). Houthakker considers a world of production units (e.g. firms) that produce with Leontief technologies where the Leontief coefficients are distributed across firms according to a Pareto distribution. Importantly, each firm has limited capacity, so that the only way to expand output is to use additional firms. Houthakker then shows that the aggregate production function across these units is Cobb-Douglas.

The result here obviously builds directly on Houthakker's insight that Pareto distributions can generate Cobb-Douglas production functions. The result differs from Houthakker's in several ways, however. First, Houthakker's result is an aggregation result; it applies exactly at all points in time. It would, for example, suggest that capital shares at the 2-digit industry level should be constant. Here, the result is an asymptotic result for a single production unit (be it a firm, industry, or country), and the result applies to the long-run production function, i.e. to the shape of the production function that is relevant if one looks across techniques. Second, the Leontief restriction in Houthakker's paper is important for the result; it allows the aggregation to be a function only of the Pareto distributions. Here, in contrast, the result is really about the shape of the long-run production function, looking across techniques. The local shape of the production function does not really matter, so that no restriction to the Leontief form for the shape of a particular technique is needed. Finally, Houthakker's result relies on the presence of capacity constraints. If one wants to expand output, one has to add additional production units, essentially of lower "quality." Because of these capacity constraints, his aggregate production functions are characterized by decreasing returns to scale. In contrast, the result here shows how the nonrivalry of ideas and the constant returns to rivalrous factors that is implied can be respected: a given technique can be used at any scale of production.

The second main comment is that Pareto distributions are crucial to the result. Is it plausible that the distributions for ideas are Pareto?

To begin, let's develop some intuition for why the Pareto distribution is important by considering a simple example. Imagine drawing social security numbers for the U.S. population at random, and for each person drawn, record their income and their height. Let y^{max} denote the maximum income drawn in a given sample and let h^{max} denote the maximum height found in the sample. Then consider the following conditional probability: Prob $\{X \geq \gamma x^{max} \mid X \geq x^{max}\}$ for $\gamma > 1$, where x stands for either income or height. This probability answers the question: "Given that the tallest person observed so far is 6 feet 6 inches tall and given that we just found someone even taller, what is the probability that this new person is more than 5 percent taller than our 6 foot 6 inch person?" Clearly as h^{max} gets larger and larger, this conditional probability gets smaller and smaller — there is no one in the world taller than ten feet.

In contrast, consider the income draws. Now, the probability answers the question: "Given that the highest-earning person observed so far has an annual income of $240,000 and given that we just found someone who earns even more, what is the probability that this new person's earnings exceed the previous maximum by more than 5 percent?" It turns out empirically that this probability does not depend on the level of y^{max} being considered. Indeed, it was exactly this observation on incomes that led Pareto to formulate the distribution that bears his name: the defining characteristic of the Pareto distribution is that the conditional probability given above is invariant to x^{max}.

In applying this example to growth models, one is led to ask whether the distribution of ideas is more like the distribution of heights or the distribution of incomes. An important insight into this question was developed by Kortum (1997). Kortum formulates a growth model where productivity levels (ideas) are draws from a distribution. He shows that this model generates steady-state growth only if the distribution has Pareto tails. That is, what the model requires is that the probability of finding an idea that is 5 percent better than the current best idea is invariant to the level of productivity embodied in the current best idea. Of course, this is almost the very definition of a steady state: the probability of improving economy-wide productivity by 5 percent can't depend on the level of productivity. This requirement is satisfied only if the tails of the distribution are power functions, i.e. only if the tails look like the tails of the Pareto distribution.[15]

A literature in physics on "scale invariance" suggests that if a stochastic process is to be invariant to scale, it must involve Pareto distributions. Steady-state growth is simply a growth rate that is invariant to scale (defined in this context as the initial level of productivity). Whether incomes are at 100 or 1000, steady-state growth requires the growth rate to be the same in both cases.

Additional insight into this issue emerges from Gabaix (1999). Whereas Kortum shows that Pareto distributions lead to steady-state growth, Gabaix essentially shows the reverse in his explanation for Zipf's Law for the size of cities. He assumes that city sizes grow at a common exponential rate plus an idiosyncratic shock. He then shows that this exponential growth generates a Pareto distribution for city sizes.[16]

These papers by Kortum and Gabaix suggest that Pareto distributions and exponential growth are really just two sides of the same coin. The result in the present paper draws out this connection further and highlights the additional implication for the shape of production functions. Not only are Pareto distributions necessary for exponential growth, but they also imply that the long-run production function takes a Cobb-Douglas form.

[15]Kortum also shows that if the tails of the distribution are thinner than Pareto, as is the case for the log normal or exponential distributions, then exponential growth rates decline to zero. If the tails are thicker, then presumably growth rates rise over time, but it is not clear to me that this case is fully considered by Kortum.

[16]An important additional requirement in the Gabaix paper is that there be some positive lower bound to city sizes that functions as a reflecting barrier. Otherwise, for example, normally distributed random shocks results in a log-normal distribution of city cizes.

5. CONCLUSIONS

This paper outlines four stylized facts with which growth models should be consistent: steady states, capital shares that are not necessarily constant, an elasticity of substitution between capital and labor that is less than one, and the presence of technological progress that is embodied in some kinds of capital like computers. Superficially, these facts appear to pose a puzzle for growth theory given the well-known theorem that steady-state growth requires labor-augmenting technical change or Cobb-Douglas production.

This puzzle can be resolved by recognizing that the short-run and long-run elasticities of substitution between capital and labor are likely to be different, with the elasticity being much lower in the short run. A simple example of a production function with these characteristics is the SR/LR production function introduced in equation (1), which is related to the literature on appropriate technologies.

Motivated by this production function, the paper emphasizes a particular view of production functions. A production technique is an idea that indicates how to produce with a particular amount of capital per person. While it is possible to use the technique with a different capital-labor ratio, diminishing returns sets in quickly as the elasticity of substitution is less than one. If one wants to produce with a much larger capital-labor ratio, one really needs a new technique. Therefore, the long-run shape of the production function is intimately tied to the distribution of ideas.

Finally, the paper builds on insights by Houthakker (1955–1956) and Kortum (1997) to show that if the distributions related to ideas are Pareto distributions, then the long-run production function has a Cobb-Douglas form, where the exponents in the production function depend on the parameters of the idea distributions.

APPENDIX A

The Steady-State Growth Theorem

That steady-state growth implies either a Cobb-Douglas production function or labor-augmenting technical change is known at some level by all growth economists. The exact nature of the result, however, is sometimes unclear, or at least it was to me. For example, Barro and Sala-i-Martin (1995) state and prove the result in the Appendix to Chapter 2, but only in the following way: if there is a steady state and if technical change is factor augmenting, then it must be labor augmenting. This seems to open the door to the possibility that one can get steady-state growth with non-factor-augmenting technological change. As it turns out, this is not true.

To show what is and is not possible, we begin with the factor augmenting version of the theorem and then explain how the more general theorem is proved.

THEOREM A.1. *Suppose $Y_t = F(B_t K_t, A_t L_t)$ is the production function for a neoclassical growth model, where t indexes time and B_t and A_t represent exogenous technological change at constant exponential rates.*[1] *If the model exhibits a steady state with a constant rate of growth of Y/L, and constant, nonzero factor shares $F_L L/Y$ and $F_K K/Y$, then along the balanced growth path either B_t is constant or the production function is Cobb-Douglas.*

[1]By neoclassical growth model, we mean that there is a standard capital accumulation equation, requiring the growth rate of Y to equal the growth rate of K in the steady state. In addition, the production function exhibits constant returns to scale in K and L and satisfies $F_K \geq 0$, $F_L \geq 0$, $F_{KK} \leq 0$, $F_{LL} \leq 0$. In addition, we will require an Inada condition $\lim_{K\to\infty} F_K = 0$. This condition is needed to rule out the possibility that the production function is asymptotically linear in capital.

Proof of Theorem A.1[2]

Differentiating the production function leads to the standard growth accounting formula

$$\hat{Y}_t = \alpha_{Kt}(\hat{B}_t + \hat{K}_t) + \alpha_{Lt}(\hat{A}_t + \hat{L}_t),$$

where hats denote growth rates and $\alpha_{xt} \equiv \frac{F_x x}{Y}$ denotes the factor share for x. Since $\hat{Y} = \hat{K}$ in steady state and since constant returns leads to factor shares that add to one, this equation reduces in steady state to

$$\hat{Y} = \frac{\alpha}{1-\alpha}\hat{B} + \hat{A} + \hat{L} \tag{A.1}$$

where $\alpha = \lim_{t\to\infty} \alpha_{Kt} \in (0,1)$ denotes the limiting value of the capital share. Therefore

$$\tilde{y}_t \equiv \frac{Y_t}{B_t^{\alpha/(1-\alpha)} A_t L_t}$$

must be constant in steady state.

Since K/Y and α are constant in steady state, the marginal product of capital must be constant as well. Indeed, this is the key to the proof: unless either B_t is constant or the production function is Cobb-Douglas, the marginal product will not be constant. To see this, let $f(x) \equiv F(x,1)$ denote the production function describing the amount of output that can be produced with 1 unit of labor in efficiency units and x units of capital in efficiency units. Then, we have

$$\begin{aligned}
\frac{\partial Y_t}{\partial K_t} &= B_t F_1(B_t K_t, A_t L_t) \\
&= B_t F_1(\frac{B_t K_t}{A_t L_t}, 1) \\
&= B_t f'(\frac{B_t K_t}{A_t L_t}) \\
&= B_t f'(B_t^{\frac{1}{1-\alpha}} \tilde{k}_t),
\end{aligned}$$

[2]This proof is motivated by the one in Barro and Sala-i-Martin (1995), but contains more intuition and emphasizes the importance of nonzero factor shares.

where $\tilde{k}_t \equiv \frac{Y_t}{B_t^{\alpha/(1-\alpha)} A_t L_t}$.

Then the marginal product of capital can be constant in two ways. First, B_t could be constant, in which case technical change is labor augmenting. Second, the production function $f(\cdot)$ could be such that the exponential rise in B exactly cancels. It is easy to see that this can occur only if $f'(x) = Cx^{-(1-\alpha)}$, where C is some constant. Therefore, $f(x) = \bar{C}x^{\alpha}$, where $\bar{C} = C/\alpha$.[3] Since $Y/AL = f(BK/AL)$, the production function must, at least in steady state, be given by $Y = \bar{C}(BK)^{\alpha}(AL)^{1-\alpha}$. QED.

Remarks

A couple of clarifying remarks about this result are helpful. First, the intuition. For the capital share to be positive and constant, the marginal product of capital must be positive and constant. The only way this can happen is either if B_t is constant or if the production function is Cobb-Douglas. Otherwise, the rise in B_t causes the marginal product of capital to trend (e.g. to fall to zero if the elasticity of substitution in production is less than one).

Second, the theorem as stated and proved shows only that *if* technological change is factor-augmenting, then it must be labor-augmenting. This seems to suggest the possibility that technological change that is not factor augmenting might give rise to a steady state, at least under some circumstances. This is not correct. The correct statement of the full theorem applies to a more general production function $F(K, L, t)$. The full theorem is proved by Uzawa (1961) in the following way. First, consider the following definition, due to Harrod (1948):

DEFINITION A.1. Technical change is said to be *Harrod-neutral* if, holding constant K/Y, the technical change leaves factor shares (e.g. $F_K K/Y$) unchanged.

[3]This can be proved formally by solving the implied differential equation for $f(x)$.

Then, if we want a steady state with constant factor shares, by definition the technical change must be Harrod-neutral. Uzawa (1961) establishes the now well-known equivalence between Harrod-neutral technical change and labor-augmenting technical change. Therefore, to get a steady state with constant factor shares, it must be possible to write the production function so that technical change is labor augmenting.

Third, the requirement in the theorem that factor shares be nonzero is important, as the following simple example shows. Suppose the production function is $Y = F(BK, AL)$, where $F(\cdot, \cdot)$ is a standard CES production function with an elasticity of substitution less than one. If B and A grow at constant exponential rates, then a neoclassical growth model with this production function *does* converge to a balanced growth path. The capital share falls to zero and the economy behaves asymptotically as if the production function were simply $Y = AL$.[4]

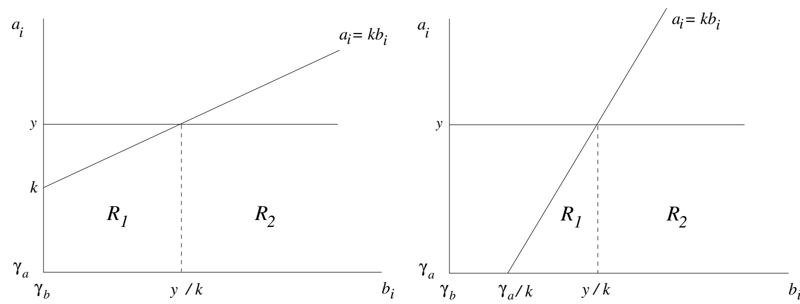
APPENDIX B

Data Sources for Capital Shares

Two data sources are used for computing the capital shares. Employee compensation as a share of GDP is calculated using the United Nations National Accounts database, purchased from the United Nations Statistics Division. This database contains, in electronic format, the data corresponding to the National Accounts Statistics, Part 1, Main Aggregates and Detailed

[4]Krusell et al. (2000) provide an interesting example of this. They assume a four-factor production function that depends on structures, equipment, skilled labor, and unskilled labor, with a general nested CES structure. The equipment-specific technological change that they assume essentially means that the equipment share of factor payments goes to zero: equipment and skill are more complementary than Cobb-Douglas, so the equipment-specific technical change makes that particular CES function ultimately proportional to just the quantity of skilled labor. Hence this paper gets steady state growth with a non-Cobb-Douglas production function, but only in the limit as the equipment share of factor payments to goes to zero.

FIGURE 9. The Two Cases in (a_i, b_i) Space



(a) Case 1: $\frac{k}{y} \leq \frac{1}{\gamma_b}$ and $k \geq \frac{\gamma_a}{\gamma_b}$ (b) Case 2: $y \geq \gamma_a$ and $k < \frac{\gamma_a}{\gamma_b}$

Tables for the years 1950–1997. The 1997 end date occurs because 1997 is the last year that the 1968 version of the System of National Accounts was used. The data come from Table 103. Employee compensation is line 5, while GDP is line 13 of the table.

The correction for self-employment is done using data from the OECD *Labour Force Statistics* publication. The underlying data series is wage earners and salaried employees as a share of total civilian employment and comes from Section IIIA of the publications. Data from three different editions of this publication is compiled, the 1971 edition (years 1958–1969), the 1985 edition (years 1963–1983), and the 2002 edition (years 1981–2001). Data are spliced in 1963 and in 1981, using a constant factor of proportionality based on the values in those years.

APPENDIX C

As discussed in the text, there are two cases in general that need to be considered in the model: $k \geq \gamma_a/\gamma_b$ and $k < \gamma_a/\gamma_b$. Figure 9 shows these cases graphically. The first three propositions prove the results for Case 1.

Proposition C.1 then gives the results for Case 2.

Proof of Proposition 3.1

This is the proof of the proposition that characterizes the distribution of output per worker y that can be produced given a technique that operates with capital per worker k, for Case 1.

By Bayes' Law,

$$\begin{aligned} F(y;k) &\equiv \text{Prob}\,\{a_i \le y \mid \frac{a_i}{b_i} \le k\} \\ &= \frac{\text{Prob}\,\{a_i \le y \text{ and } \frac{a_i}{b_i} \le k\}}{\text{Prob}\,\{\frac{a_i}{b_i} \le k\}}. \end{aligned} \tag{C.1}$$

Call the numerator of this expression $H(y,k)$ and call the denominator $G(k)$. That is, $H(y,k)$ is the relevant joint probability and $G(k)$ is the unconditional probability that a technique can be used with capital per worker equal to k. We now calculate these two probabilities.

First, consider $H(y,k)$. Let $g_1(a) \equiv G_1'(a)$ and $g_2(b) \equiv G_2'(b)$ be the densities of the Pareto distributions, and let $g(a,b) \equiv g_1(a)g_2(b)$ be the joint density (since the distributions are independent).

Figure 9 shows graphically how to calculate the probability. For Case 1, we have

$$\begin{aligned} H(y,k) &\equiv \text{Prob}\,\{a_i \le y \mid \frac{a_i}{b_i} \le k\} \\ &= \int\int_{R_1} g(a,b)dadb + \int\int_{R_2} g(a,b)dadb \\ &= \int_{\gamma_b}^{y/k}\int_{\gamma_a}^{kb} g_1(a)dag_2(b)db + \int_{y/k}^{\infty}\int_{\gamma_a}^{y} g_1(a)dag_2(b)db. \end{aligned}$$

Straightforward evaluation of the integrals reveals that

$$H(y,k) = 1 - \frac{\beta}{\alpha+\beta}\left(\frac{\gamma_a}{\gamma_b}\right)^{\alpha} k^{-\alpha} - \frac{\alpha}{\alpha+\beta}\gamma_a^{\alpha}\gamma_b^{\beta}k^{\beta}y^{-(\alpha+\beta)}. \tag{C.2}$$

Next, consider $G(k)$. By a similar argument,

$$\begin{aligned} G(k) &\equiv \text{Prob}\left\{a_i \leq kb_i\right\} = \int_{\gamma_b}^{\infty}\int_{\gamma_a}^{kb} g_1(a)da g_2(b)db \\ &= \int_{\gamma_b}^{\infty} G_1(kb)g_2(b)db \\ &= 1 - \frac{\beta}{\alpha+\beta}\left(\frac{\gamma_a}{\gamma_b}\right)^{\alpha} k^{-\alpha}. \end{aligned} \tag{C.3}$$

Finally, combining the results from equations (C.1), (C.2), and (C.3), we have the desired result for Case 1:

$$F(y;k) = 1 - \theta(k) y^{-(\alpha+\beta)}, \tag{C.4}$$

where

$$\theta(k) \equiv \frac{\frac{\alpha}{\alpha+\beta}\gamma_a^{\alpha}\gamma_b^{\beta}k^{\beta}}{1 - \frac{\beta}{\alpha+\beta}\left(\frac{\gamma_a}{\gamma_b}\right)^{\alpha} k^{-\alpha}}. \tag{C.5}$$

Q.E.D.

Proof of Proposition 3.2

This proposition provides the distribution of techniques that can work with a capital per worker level less than k. It was proven above, as part of the proof of Proposition 3.1; see equation (C.3) above.

Proof of Proposition 3.3

This proposition applies an extreme value theorem of Galambos (1978) to show that the distribution of the largest amount of output per worker y^* that can be produced with a given quantity of capital per worker k^* is distributed according to the Frechet distribution. To begin, it is useful to state the extreme value theorem itself, Theorem 2.1.1 of Galambos (1978).[1]

[1] See also Kortum (1997) and Billingsley (1986) for discussions of this theorem.

THEOREM C.1 (2.1.1 of Galambos 1978). *Let x be a random variable drawn from a distribution $F(x)$, with* $\sup\{x : F(x) < 1\} = +\infty$. *Assume there is a constant $\gamma > 0$ such that, for all $x > 0$,*

$$\lim_{\tau\to\infty} \frac{1 - F(\tau x)}{1 - F(\tau)} = x^{-\gamma}. \tag{C.6}$$

Consider n draws of x from the distribution F, and let Z_n denote the maximum value from these n draws. Then there is a sequence $b_n > 0$ such that,

$$\lim_{n\to\infty} \text{Prob}\,\{Z_n < b_n x\} = H_{1,\gamma}(x), \tag{C.7}$$

where

$$H_{1,\gamma} = \begin{cases} \exp(-x^{-\gamma}) & \textit{if } x > 0 \\ 0 & \textit{otherwise}. \end{cases} \tag{C.8}$$

The normalizing constant b_n can be chosen as

$$b_n = \inf\{x : 1 - F(x) \leq \frac{1}{n}\}. \tag{C.9}$$

This theorem has the following interpretation. The (normalized) distribution of the maximum value drawn from a distribution with a Pareto (i.e. power function) upper tail converges to the Frechet distribution as the number of draws goes to infinity. The normalizing factor is given by b_n in the theorem.

This theorem applies directly to our problem. The distribution $F(y; k)$ in the upper tail (i.e. for any $y \geq k\gamma_b \geq \gamma_a$, which is just Case 1) is exactly Pareto. Let $y^* \equiv max_i\{a_i : a_i/b_i \leq k^*\}$. Let $N(k^*)$ denote the number of techniques that can be used with capital per worker k^*. Then, as a direct application of the Galambos theorem,

$$\lim_{N(k^*)\to\infty} \text{Prob}\,\{\frac{1}{b_N} y^* \leq x\} = \exp(-x^{-(\alpha+\beta)}) \tag{C.10}$$

for $x > 0$.

The normalizing sequence b_N is given by

$$\begin{aligned} b_N &= \inf\{y : 1 - F(y; k^*) \leq \frac{1}{N(k^*)}\} \\ &= \inf\{y : \theta(k^*) y^{-(\alpha+\beta)} \leq \frac{1}{N(k^*)}\} \\ &= (\theta(k^*) N(k^*))^{1/(\alpha+\beta)}, \end{aligned}$$

which completes the proof the proposition. Q.E.D.

The Results for Case 2

We now state and prove the proposition that derives the analogous results in Propositions 4.1 to 4.3 for Case 2.

PROPOSITION C.1. *In Case 2, i.e. when $y \geq \gamma_a$ and $k < \gamma_a/\gamma_b$, we have the following results:*

1. The important probabilities are given by

$$\begin{aligned} H(y,k) &\equiv \text{Prob}\,\{a_i \leq y \mid \frac{a_i}{b_i} \leq k\} \\ &= \frac{\alpha}{\alpha+\beta}\left(\frac{\gamma_b}{\gamma_a}\right)^{\beta} k^{\beta}\left(1 - \left(\frac{y}{\gamma_a}\right)^{-(\alpha+\beta)}\right). \end{aligned} \tag{C.11}$$

$$\begin{aligned} G(k) &\equiv \text{Prob}\,\{a_i \leq k b_i\} \\ &= \frac{\alpha}{\alpha+\beta}\left(\frac{\gamma_b}{\gamma_a}\right)^{\beta} k^{\beta}. \end{aligned} \tag{C.12}$$

$$F(y;k) = 1 - \left(\frac{y}{\gamma_a}\right)^{-(\alpha+\beta)} \tag{C.13}$$

2. The Galambos theorem applies so that

$$\lim_{N(k^*)\to\infty} \text{Prob}\,\{\frac{1}{b_N} y^* \leq x\} = \exp(-x^{-(\alpha+\beta)}) \tag{C.14}$$

for $x > 0$.

The normalizing sequence b_N *is given by*

$$\begin{aligned} b_N &= \inf\{y : 1 - F(y; k^*) \leq \frac{1}{N(k^*)}\} \\ &= \gamma_a (N(k^*))^{1/(\alpha+\beta)}, \end{aligned}$$

3. Finally, letting N *denote the total number of techniques, the production function behaves asymptotically as*

$$y^* = AN^{\frac{1}{\alpha+\beta}} k^{*\frac{\beta}{\alpha+\beta}} \epsilon$$

where $A \equiv \left(\frac{\alpha}{\alpha+\beta}\gamma_a^\alpha \gamma_b^\beta\right)^{\frac{1}{\alpha+\beta}}$, *and where* ϵ *is an i.i.d. random variable drawn from the Frechet distribution given above.*

The proof of this proposition follows the same logic as the proof of Propositions 4.1 through 4.3, and the reasoning given in the text to derive the behavior of the production function in the long run.

REFERENCES

Acemoglu, Daron, “Labor- and Capital-Augmenting Technical Change,” July 2001. MIT mimeo.

Antràs, Pol, “Is the U.S. Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution,” February 2001. MIT mimeo.

Barro, Robert J. and Xavier Sala-i-Martin, *Economic Growth*, McGraw-Hill, 1995.

Basu, Susanto and David N. Weil, “Appropriate Technology and Growth,” *Quarterly Journal of Economics*, November 1998, *113* (4), 1025–1054.

Batini, Nicoletta, Brian Jackson, and Stephen Nickell, “Inflation Dynamics and the Labour Share in the UK,” November 2000. Bank of England MPC Discussion Paper No. 2.

Bentolila, Samuel and Gilles Saint-Paul, “Explaining Movements in the Labor Share,” 2003. CEMFI mimeo.

Billingsley, Patrick, *Probability and Measure*, New York: John Wiley & Sons, 1986.

Blanchard, Olivier J., "The Medium Run," *Brookings Papers on Economic Activity*, 1997, *1997* (2), 89–141.

___ and Francesco Giavazzi, "Macroeconomic Effects of Regulation and Deregulation in Goods and Labor Markets," October 2002. MIT mimeo.

Caballero, Ricardo J. and Mohamad L. Hammour, "Jobless Growth: Appropriability, Factor Substitution, and Unemployment," *Carnegie Rochester Conference Series on Public Policy*, 1998, *48*, 51–94.

Caselli, Francesco and Wilbur John Coleman, "The World Technology Frontier," September 2000. NBER Working Paper No. 7904.

Drandakis, E.M. and Edmund S. Phelps, "A Model of Induced Invention, Growth, and Distribution," *Economic Journal*, December 1966, *76* (304), 823–840.

Gabaix, Xavier, "Zipf's Law for Cities: An Explanation," *Quarterly Journal of Economics*, August 1999, *114* (3), 739–767.

Galambos, Janos, *The Asymptotic Theory of Extreme Order Statistics*, New York: John Wiley & Sons, 1978.

Gilchrist, Simon and John C. Williams, "Putty Clay and Investment: A Business Cycle Analysis," *Journal of Political Economy*, October 2000, *108* (5), 928–960.

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

Gordon, Robert J., *The Measurement of Durable Goods Prices*, Chicago: University of Chicago Press, 1990.

Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell, "Long-Run Implications of Investment-Specific Technological Change," *American Economic Review*, June 1997, *87* (3), 342–362.

Hammermesh, Daniel S., *Labor Demand*, Princeton, NJ: Princeton University Press, 1993.

Harrison, Ann E., "Has Globalization Eroded Labor's Share? Some Cross-Country Evidence," May 2003. U.C. Berkeley mimeo.

Harrod, Roy F., *Towards a Dynamic Economics*, London: Macmillan, 1948.

Houthakker, Henrik S., "The Pareto Distribution and the Cobb-Douglas Production Function in Activity Analysis," *Review of Economic Studies*, 1955–1956, *23* (1), 27–31.

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

___ , "Time Series Tests of Endogenous Growth Models," *Quarterly Journal of Economics*, May 1995, *110* (441), 495–525.

Jorgenson, Dale W., Mun S. Ho, and Kevin J. Stiroh, "Growth of U.S. Industries and Investments in Information Technology and Higher Education," October 2002. Harvard University mimeo.

Kaldor, Nicholas, "Capital Accumulation and Economic Growth," in F.A. Lutz and D.C. Hague, eds., *The Theory of Capital*, St. Martins Press, 1961, pp. 177–222.

Kennedy, Charles M., "Induced Bias in Innovation and the Theory of Distribution," *Economic Journal*, September 1964, *74* (295), 541–547.

Kortum, Samuel S., "Research, Patenting, and Technological Change," *Econometrica*, 1997, *65* (6), 1389–1419.

Krusell, Per, Lee Ohanian, Jose-Victor Rios-Rull, and Giovanni Violante, "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, September 2000, *68* (5), 1029–1053.

Samuelson, Paul A., "A Theory of Induced Innovations Along Kennedy-Weisacker Lines," *Review of Economics and Statistics*, 1965, *47*, 343–356.

Solow, Robert M., "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, August 1957, *39* (3), 312–320.

___ , "A Skeptical Note on the Constancy of Relative Shares," *American Economic Review*, September 1958, *48* (4), 618–631.

Uzawa, Hirofumi, "Neutral Inventions and the Stability of Growth Equilibrium," *Review of Economic Studies*, February 1961, *28* (2), 117–124.

Whelan, Karl, "Balance Growth Revisited: A Two-Sector Model of Economic Growth," 2001. Federal Reserve Board of Governors mimeo.