STUDYING THE GROWTH PROCESS:
A PRIMER

Arnold C. Harberger
University of California, Los Angeles

Paper Prepared For A Conference on
Capital Formation and Economic Growth
The Hoover Institution
Stanford University

October 1997

STUDYING THE GROWTH PROCESS:

A PRIMER

Arnold C. Harberger

University of California, Los Angeles

NOTE TO READERS

This paper has a rather special story. It had been my plan for some time to take "A Vision of the Growth Process" as the title of my presidential address to the American Economic Association. No real draft exists even now, but I had been making presentations under that or similar titles for some time -- a very early one at Stanford, a few years ago, others at conferences at Cornell and in Kuala Lumpur, others at several Latin American institutions (the Catholic University of Chile, the University of Chile, the Center for Public Studies in Santiago, the Higher Institute of Government Economists in Buenos Aires. The most recent presentations on similar topics were at Texas A&M University, at the Western Economic Association meetings (Seattle, July, 1997) and at the Latin American meetings of the Econometric Society (Santiago, August, 1997). I want to express my thanks for the feedback I received on all these occasions.

Then, in, I believe September 8, I received from Michael Boskin an invitation to participate in a Conference on Capital Formation and Economic Growth, to be held a month later at the Hoover Institution in Stanford. The guest list read almost like a Who's Who in the field of growth. It would be hard to imagine a group of equal size that could give me more relevant feedback, prior to the drafting of my presidential address.

So I jumped at the chance to present a paper at the conference. Preparing it was another matter. I already had travel plans for most of September, and a complicated schedule after my return. Nonetheless I worked as hard as I could, in all the time available, and tried to lay out the underpinnings of my vision of the growth process. My only regret is that, as the paper was written virtually out of a suitcase, I did not have the kind of access to libraries, or the time to spend a few days in the stacks, that I would have liked.

I am very eager to get feedback -- as much as possible and as soon as possible. I would like both substantial and bibliographic feedback -- particularly references to others who have expressed similar ideas. Feedback will be of use to me personally at any time, but the deadline for my paper for the March AER is almost immediate. Many thanks for all the help you can give.

INTRODUCTION AND SUMMARY

1. This paper presents a vision of the growth process that differs, at least in its points of focus and emphasis, from much of the recent literature in the field. The key points of focus are: a) the economic rate of return to capital as an important element in the breakdown of growth and b) the interpretation of the TFP "Residual" in the growth breakdown as representing real cost reductions of all kinds.

Both of these foci lead us to look to the firm as the stage where the action takes place. The quest for profits is the motivating force for the economic activities of firms; seeking high rates of return on investment (even though the margins may be determined by some perceived cost of capital) and seeking real cost reductions of all kinds are natural avenues for implementing this quest. They are what all firms do (or at least should do if they expect to succeed).

I believe economists have focused too much on explaining the residual as stemming from broad, economy-wide externalities, and not enough on the role of the volition and effort of economic agents at the level of the firm. Without denying the existence of broad economy-wide externalities that influence the residual, I believe that closer observation of what happens at the levels of the firm and the industry will show that such macro-externalities are only a small part of the story, compared with the decisions and actions of the firms themselves.

2. This vision of the growth process leads one almost inevitably to some methodological twists -- twists which, if they are not new, at least differ in significant ways from what I take to be the most common practice in breaking economic growth down into components.

a) To measure the real rate of return to capital, one must express the numerator (real dollars of return) and the denominator (the capital stock) in the same units. The most efficient way to do this is to measure both output (value added) and the capital stock in units of the GDP deflator. That way one is sure that the outputs of all the sub-aggregates in the economy add up to the GDP, and one also satisfies the requirement that the capital and return be measured in the same units. When this is done at the aggregate level the contribution of capital to growth is (+) K where is an aggregate rate of return to capital the depreciation rate, and K the net increment to the capital stock in the period in question. At a subaggregate level the contribution of capital to growth in activity j is $(_j + _j) K_j$. At both levels, we find that high rates of return are an important component of most successful growth episodes.

b) To capture the great diversity of the labor factor, we would like to have a very fine breakdown of labor into categories (indexed by i). The labor contribution is then $_i w_i L_i$, where w_i represents the real wage of category i and L_i the change in hours worked by category i. Since the number of relevant categories of labor is huge, any such breakdown is difficult, and gets more difficult as one disaggregates from economy to sector to branch to industry and to firm. This Gordian knot can be cut by a simple assumption, similar to what is done in most countries to convert residential construction to real terms. There, one builds a price index of a "standard house" p, and then obtains a quantum of construction C* by dividing total construction outlays by the price of the standard house. In the resulting aggregate, each individual residence gets attributed a quantum of housing equal to p_i/p. In our work we define a standard wage w*, which we assign to "standard labor" or "raw labor". The excess of anybody's actual wage over w* we attribute to human capital. The returns to natural ability, as well as to formal education, training, and experience belong there, under this interpretation. High returns due to a distorted wage structure are not appropriately attributable to human capital, but the methodology would nonetheless be correct in attributing to the affected labor a marginal productivity that is measured by the distorted high wage.

The "labor contribution" as measured by w* L* is equal to

(1)

```
:{ w* left (  _i ` left ({w_i} over {w*}  right )
_i ~ + ~  _i L_i `   left ( {w_i} over {w*}
t ) right ) .  }
```

The second term will be zero if the structure of relative wages remains constant, or even if the weighted average premium does not change. Any changes in the weighted average premium will cause the calculated residual to be different from those calculated by other methods.

3. The "Two-Deflator Method" is characterized by the use of a single numeraire-deflator (say, the GDP deflator), by the treatment of the quantum of output as value added divided by the numeraire-deflator, and by the use of a standard wage w* and a quantum of labor L* equal to the wages bill divided by w*.

It goes without saying that the 2-deflator method is rough. But it is also tremendously robust and easily applied. I think of it as being really designed for use at the firm level, where very commonly we can get data in value added, on gross investment, and on the wages bill, but know nothing (from standard sources) about the quantum of output nor about the number of total hours worked (or even the total number of employees used). This opening of wide new vistas, of huge new data sets, is what I consider the big argument for the 2-deflator approach.

But there are other pluses as well. First, at the aggregate economy level, the 2-deflator approach comes very close to the traditional approach. In rate-of-growth terms, we have $(R^*/y) = g_y - s_k g_{k^*} - s_L g_{L^*}$, compared with $(R/y) = g_y - s_k g_k - s_L g_N$, where g refers to growth rate, and y to GDP. K* differs from K in being built up from gross investment deflated by the GDP deflator, while K is built up from gross investment deflated by the investment deflator of GDP. L* is in principle much more refined than N (number of workers), but its measurement can be influenced by a widening or narrowing skill premium. (R^*/y) likewise differs from Jorgenson's residual

(2) $$R''/y = g_y - {}_j\, s_{kj} g_{kj} - {}_i\, s_i g_i$$

mainly in his use of different capital deflators for different categories of capital. (The implicit labor breakdown in L* is even finer than Jorgenson's but Jorgenson does not have to contend with the possibilities of widening or narrowing skill premiums -- at least not among the labor categories he works with.)

The bottom line is that when Harald Beyer compares his 2-deflator results, at the national aggregate level, with those of others using different methods, he finds on the whole only modest differences.

When one uses the 2-deflator method at the industry level, one often has the possibility to adjust the quantity variable so as to make it correspond with that of the more traditional approach. Thus, we may start by using dy $(= p_j dy_j + y_j dp_j)$ as the quantity variable, and calculate a residual R using that concept. Then we may get an adjusted residual by taking $R - y_j\, dp_j$. This is easy to do so long as we have decent data on dp_j the relative price of j, which we often do at the branch or industry level. Figure 3 shows the distributions of the differences between Jorgenson's residuals and the 2-deflator residuals, with and without price adjustment. I find the differences without adjustment to be small enough to be quite acceptable. With adjustment the degree of agreement between the two approach is quite notable (85% of differences less than one percentage point of per annum growth).

When one gets down to the firm level, the 2-deflator method is in its element. Rarely do we have decent time series on the price index of a firm's output. Treating many firms in one industry, one might then give all of them the same price index -- that of the industry. At that point the distribution of adjusted TFP residuals among the firms would end up differing from the original 2-deflator distribution by a constant. When expressed in percentage terms we would have

$$R/y = g_y - s_{kj}g_j - s_j g_{*j} \qquad (3)$$

for each firm j without adjustment, while with adjustment we would have R/y_j = the same expression minus g, the rate of growth of the industry's relative price index (the same for all firms).

So in the great bulk of cases one ends up with something quite like the 2-deflator method when working at the individual firm level. The consolation is that the residual terms of individual firms, calculated for the whole economy, add up to a residual term for the aggregate -- in the sense that outputs sum to GDP, the L sum to L^* for the economy, the K to K^* for the economy, etc.

4. The outstanding characteristic of TFP residuals calculated in a disaggregated way is the number (and percentage) of them that are negative. This phenomenon is widespread -- it applies to all the distributions of TFP by branch and industry that I know of. For many years I thought of this as an industry phenomenon, with some industries enjoying positive TFP, presumably because they share in new technical advances, while others experienced negative TFP, say under the pressure of reduced consumer demand, or increased foreign competition, or whatever. Thinking of TFP experiences being an industry characteristic in these terms by itself casts doubt on certain global, macro-externality explanations of TFP improvement. In my analogy, one would expect such global causes as externalities due to the overall scale of the economy, or economies of scale linked to the total stock of human capital, to work like yeast -- producing TFP expansions that were quite general throughout the economy. Instead, TFP seems to work more like mushrooms, with big gains popping up in unexpected places, not even concentrating on the same industries period after period, as one might expect on certain hypotheses.

This story is reinforced as one examines the data at the firm level. Here, it appears that only rarely do the firms constituting an industry share the same TFP experience. More often than not, less than half the firms in an industry are capable of accounting for that industry's full quota of real cost reduction. Maybe

60-70 percent of the firms in an industry will experience TFP improvement. Together they account for maybe 150 or 175 percent of the industry's total real cost reduction. Then come the firms with negative TFP changes. They pull the industry total back down to, at the end, 100%.

5. This picture of TFP at the firm level is largely taken from work by Leonardo Torre on a sample of some 2000 Mexican firms -- it is confirmed also by some work of Edgar Robles on the U.S. petroleum industry. But I would not be putting it forward if I did not believe that it will be robustly confirmed as we do further work on the individual firm level.

The picture is one of severe competition, of gains outweighing losses overall, but with plenty of negative components. Indeed, at least at the manufacturing industry level, the distribution of total TFP increases is quite tight, if one counts only the branches with positive TFP improvement over different time periods. What make for big differences across time periods are the branches with negative TFP experience.

I think we have a big research agenda ahead of us, trying to understand why this is so, and what are the forces behind the various negative TFP experiences of different industries within a sector, or of different firms within an industry.

But even before digging into that major task, I feel I see in the data the great insight of Schumpeter "in the flesh", as it were. Schumpeter spoke of "creative destruction" as the vital force of a competitive market economy in action. This is the message that I receive from the whole array of data I have looked at. Winning and losing are both integral parts of the economic process. Some losers may be lazy or incompetent, but on the whole they are trying, maybe just as hard as the winners, to stay afloat. Economies of scale may take on a new aspect, when one considers that negative TFP experiences can arise from firms being driven back up their short run average cost curves as competitive forces cause output to fall well below designed capacity.

Anyway, this vision of the growth process opens up many new vistas and gives us many new challenges. To me, it gives life to the residual, viewed as real cost reduction, in a way that remote macroeconomic externalities never did. It gives the residual body, in the sense that the number of dollars saved by real cost reduction is a tangible and measurable quantity. It gives the residual a name (real cost reduction), an address (the firm), and a face (the face of the entrepreneur, the CEO, the production manager, etc.) And finally we see that there can be vastly different expressions on that face, even as we move from firm to firm in a given industry, as the TFP experience of a period moves from sharply positive to devastatingly negative.

6. Finally, there are policy lessons that flow from this vision of the growth process -- policy lessons that are different from and additional to the lessons that come from thinking in terms mainly of scientific advances plus macro-externalities. These are lessons that focus on the economic agents who are trying to reduce real costs, in any way possible, at the level of the firm. These lessons include the control of inflation, whose worst costs stem from its blurring economic agents' perceptions of relative prices. It is much harder to reduce real costs when one perceives relative prices with great error than when one perceives them much more precisely. The lessons also include eliminating important price distortions, which cause agents to minimize the wrong costs instead of the right ones. The lessons also include reducing the real costs of compliance with rules, regulations and procedures that have little economic justification or purpose. Privatization can play a major role in giving private agents a much freer hand then public sector managers typically have in seeking least-cost solutions and growth-producing innovations. Finally, one can easily motivate, in terms of their likely impact on growth the types of policies that have characterized recent major takeoffs -- policies that John Williamson has aptly summarized as "macroeconomic prudence, outward orientation, and domestic liberalization." Such policies are not only likely to lead to greater economic efficiency, looked at from a comparative static point of view. They also facilitate the process of real cost reduction, thus greasing the wheels of this engine of growth.

FOCUSING ON THE RATE OF RETURN

7. Our literature has not put enough stress on the importance of the rate of return to capital in the growth process. Maybe this is because, in a sense, it got buried in our notation. Consider the simple case:

(4) y = L + (+) K + R

I have always liked to translate this to a growth rate in the following way:

(4a)

:{ { y} over {y} ~ = ~ {Lw} over {y} ` left
L} over {L} right)~ + ~ (+) ` left ({ K}
: {y} right) ~ + ~ R over {y} }

This way of doing it calls attention to the role of in accounting for growth. It has a role completely symmetrical to the rate of net investment (K/y). In point of fact, we often see as much or more variability in as we do in K/y. And, among all the determinants of growth may be one of the most amenable to policy influences, either for good or for ill.

In contrast, the standard translation of (4) into a growth rate looks like this:

(4b)

:{ { y} over {y} ~ = ~ left ({Lw} over {y} right) `left
_} over {L} right) ~ + ~ left ({(+)K} over {y} = s
t)` left ({ K} over {K} right) ~ + ~ R over {y} }
:{ { L} over {L} ~ +
_k ~ { K} over {K}
~ R over {y} }

Here the rate of return is "buried" in s_k, the share of capital in GDP. It's not <u>absent</u>, just hidden from view. But that may have been sufficient to cause us, as a profession, to give it less than its due weight in our discussions of the growth process.

<u>8.</u> There is a very close connection between the rate of return and the residual R". [I will use R" here to emphasize that I am talking about real cost reduction, not improvement in labor quality. This connection comes from the fact that R has a "dual" representation

(5a) $R = p_j\ y_j - {}_i\ w_{ij}\ L_{ij} - ({}_j + {}_j)\ K_j$, and

(5b) $R = {}_i\ L_{ij}\ w_{ij} + K_j\ ({}_j + {}_j) - y_j\ p_j$

Thinking of R as real cost reduction, (5a) says that it can be measured as a growth in output in excess of the growth in inputs (i.e., an increase in output per unit of input). The "dual" representation (5b) says that the effects of real cost reduction will be transmitted either to labor in the form of increased real wages (w_{ij}), or to capital in the form of increased real profits (${}_j$) or to consumers in the form of lower real product prices (- p_j), or some combination of the three.

<u>9.</u> Of the three terms on the right hand side of b), $K_j\ ({}_j + {}_j)$ is the one that burns brightest at the level of the firm. When entrepreneurs and business managers find ways to generate real cost reductions, they are mainly thinking about increased profits. Changes in wage rates (w_{ij}) and in product prices (p_j) are mostly the result of market forces beyond the control of the firm -- they may be welcome boons or extremely distasteful shocks, but they end up as something the firm either enjoys or has to cope with. In contrast, a rise in profitability is something that every firm is (or should be) actively seeking all the time. Hence it should come as no surprise that when we calculate TFP indexes by methods that involve the explicit calculation of rates of return, there is typically a quite high (and in my experience always positive) correlation between the TFP index (really, an accumulated residual term, in some form) and the rate of profit (an accumulation of the $_j$). In short, the residual R (or even R_j for that matter) tends to be high when the profit rate $_j$ is rising, and to be low (or even negative) when the profit rate is falling. But this is not the right way to think about it. The right way is to recognize that real cost reductions are more often than not reflected in higher profit rates, and vice versa for real cost increases. The causality in most sensible economic scenarios, goes from R_j to $_j$, not the other way around.

<u>10.</u> This is the point where one must pound home the <u>truth</u> that no simple, single label like economies of scale or externalities from human capital, or inventions resulting from R&D, can conceivably cover more than a small fraction of the real cost reductions that we see in the real world. Ask me why, and I'll say we know better. Ask me how, and I'll say by talking to businessmen, by viewing business operations at first hand, by reading about how struggling companies were turned around, etc., etc. I know of one case where installing music in an apparel plant caused a 20% increase in productivity, and of another where merging two refrigerator factories and reducing the number of separate models from two dozen to just two, led to a more than doubling of output using less than half the initial labor force, with a consequent huge increase in profitability (and total factor productivity). We all know of cases where changing managers enhanced productivity, even without any accompanying change in the technical mode of production (i.e., machines, processes, etc.) We all know, too, that a large fraction of investments are aimed at making things "cheaper and better", but the investment may be anything from cellular phones for salesmen to longer lasting tires for one's trucks to modern, computerized inventory systems. Why should we even <u>try</u> to find a better umbrella than "real cost reduction" to cover all these cases? Of course, this should in no way preclude us from later taking a serious look under the umbrella, to try to better grasp the nature and variety of what we see there.

<u>11.</u> Real cost reduction (RCR) is, to me, a beautiful label, because is so clearly evokes the idea of <u>volition</u> on the part of entrepreneurs, managers, engineers, businessmen in general. To me, it is almost coterminous with the idea of increasing profits. So I am quite happy with a residual R" that measures (or tries to measure) RCR. I do not at all feel it is a "measure of our ignorance", or anything else about which we should feel embarrassed. On the contrary, I see RCR as almost on a par with the search by consumers of ways to get more utility out of the resources at their command.

This is not to say I am averse to efforts to learn more about the many incarnations of RCR, its many facets and reflections. Indeed, that is a direction in which I would very much like to go. But we do not learn more by making assumptions that in one way or another impute a single underlying cause (or even two or three). Such assumptions flout the essential nature of RCR. (Much more on this will come later.)

MEASURING THE REAL RATE OF RETURN

12. By now the essentials of measuring the real rate of return to capital are well established. The capital stock is formed by accumulating real net investment year by year. Where one has a clear starting date (as in an investment project) one starts from that point. Where one is dealing with a continuing operation, one has to estimate the initial capital stock as of the starting date. Various methods exist for making plausible estimates of initial stock; they become less important, in a sense, the father back in the past the starting date is located, for that means that a larger fraction of the initial value will have been already depreciated by any more recent time on which we are really focusing. Most economists have opted to ignore accounting depreciation because it so often fails to reflect true real depreciation, at least where significant inflation appears over the period of measurement and/or where tax laws exert a strong influence on the depreciation allowances the firms choose to claim. The standard approach that has dominated actual practice is the perpetual inventory method. Each year's nominal investment is deflated by a chosen deflator; the result is added to the beginning-of-period capital stock, after the latter has been adjusted for depreciation during the period. Thus we get $K_t = [K_{t-1}(1- \ ) + I_t]$ where K and I are both expressed in units of the chosen deflator and where is the assumed real depreciation rate per period. Sometimes one has no choice but to consider the capital stock as a whole, but when one is lucky one finds separate investment data on, say, structures, machinery and equipment, and inventories. Here one can use appropriately different depreciation rates for the different categories, including zero depreciation, of course, for inventories. When one has to treat the capital stock as a whole, one tries to pick (or guess at) a number that represents a sensible weighted average of the rates of depreciation of the various components of the capital stock in question. In many exercises carried out over a period of years, I have used 2 1/2% per annum for structures, 8% for machinery and equipment and zero percent for inventories, with 5% being the global figure used whenever a breakdown by type was not available. It is always prudent to perform sensitivity tests on rough assumptions such as these; my standard ones use 2, 2 1/2, and 3 percent for structures and 6, 8, and 10 percent for machinery and equipment.

The end result of applying the perpetual inventory method is a time series of capital stock expressed in real units. These units will be consumption baskets if the chosen deflator used is the CPI, production baskets if the chosen deflator is the GDP deflator, and investment baskets if the chosen deflator is the investment deflator of the national accounts.

13. In measuring the rate of return $_j$ we want to put in the numerator a real flow of returns to capital expressed in the same units as the denominator K_j. This pretty much automatically rules out using the investment deflator of the national accounts as the standard deflator (who wants to think of measuring the GDP and its components in standard baskets of factories, machines, and inventories?). A better case can be made for using the CPI as the standard deflator. Capital is then measured as a homogeneous blob of consumption baskets foregone through years of past investment and not yet recovered via subsequent real depreciation. But this has the disadvantage of restating something very familiar (the GDP) in a new metric. All in all, I prefer (at least for the purposes of this paper) to stick with the GDP deflator as the standard numeraire, to be carried through the entire exercise.

Thus, we have $_j$, the real net-of-depreciation income accruing to capital in activity j, expressed as

$$_j = p_j y_j - {}_i L_{ij} w_{ij} - {}_j K_{j,-1}$$

where

$p_j y_j$ = GDP originating in activity j, (i.e., gross value added), expressed in units of the global GDP deflator.

p_j = the relative price of value added in activity j. This is simply the price of the product, where there are no purchased inputs. Where there are such inputs, p_j is the relative price of the product minus ${}_h a_{hg} q_h$, where a_{hj} are the fractions of total cost spent on input h, and the q_h are the input prices, each deflated by the standard (GDP) deflator.

L_{ij} = the amount of labor of type i used in activity j (expressed hopefully, as a total of hours worked by such labor) and

w_{ij} = the wage paid to labor of type i in activity j.

For ease of understanding, it is best to simply think of y_j as the product of activity j and p_j as its relative price -- i.e., not to become embroiled in the complications of expressing real value added of each activity as a product of a quantum and a price. I do not believe that any essential problems are being glossed over when one thinks in these terms.

Since $_j$ and K_j, as defined are expressed in the same units it is straightforward to calculate $_j$ simply as their ratio. This, at least, is the basic concept.

14. Some subtleties do enter, however, in the calculation of $_j$. First, if the time series on K_j is built up as described in para. 12 above, it counts only reproducible capital, or, in a more mechanical sense, whatever is counted under investment in the national accounts or any other source that is being used. Sticking for the moment with the national accounts, we have somehow to recognize that the income accruing to capital, as measured there, also includes income from land. There are two ways to deal with this problem, as one tries to estimate $_j$. One is to pare down the numerator in an effort to exclude income attributable to land, and the other is to build up the denominator in order to include land in the capital stock. Of the two, the first is by far the one to be preferred. I do not believe there is any country in which cadastral information exists that comes even close to giving the current nominal or real market value of land. On the other side it is significantly easier to estimate a rough fraction of the value added in agriculture, forestry and cattle raising that can plausibly be attributed to land, and another fraction of the value added in the sector producing the services of dwelling units, which might there be reasonably attributed to land. Hard work will in most cases produce inputs such as the typical fraction of agricultural output going for rent, when land is rented; the typical fraction that land cost occupies in the cost of dwelling units; the typical rate of real appreciation of such land (because we want to attribute to land only that part of its return that is reflected in real or imputed rents), etc. On the basis of such information, reasonably refined estimates can be made of the two fractions mentioned above. If rough-and-ready assumptions are all that one can manage, I would suggest assigning one third of GDP originating in agriculture and one-tenth of GDP originating in the rental services of dwelling units as plausible guesses. At least they have served their purpose quite well for me, in a multitude of applications over a period of some 20 years.

15. A second subtlety concerns dividing the income from unincorporated enterprises into a fraction due to labor of all types and another due to capital services. Usually this income is treated in the accounts as if it were a return to capital (in most Latin American countries it is added to corporate profits under the label of "excedente de explotacion" or something similar). So the task is to sequester a piece of this income and impute it to the labor factor. Different techniques have been used to impute wage rates to owners of unincorporated enterprises and their family members, the simplest being to assign to them the average wages of the hired workers in the sector or activity in question. Careful identification of just who the family members are, and just how often and how much they work (e.g., just at planting and harvest times?) can improve significantly on the first type of imputation. Once again, there exists the alternative of simply making an informed guess as to the fraction of "income of unincorporated enterprises" that can plausibly be assigned to the "unpaid" labor of owners and their families. If one knows the setting, it is not too hard to build up a reasonable estimate of this fraction. (How capital-intensive are the activities in question? What fraction of the total labor used is represented by "unpaid" owners' and family labor?; Are they well-educated plantation owner-managers, or the wives and children of poor peasants? In different situations that I have confronted, I have ended up using fractions as low as 15% and as high as 50%, sometimes as limits to a plausible range (like 15-30 percent) which is later explored in sensitivity tests.

16. A third subtlety has to do with taxes. I firmly believe that for growth accounting purposes we should try to come as close as we can to the marginal products that a production manager would have in mind (i.e., we are concerned with the <u>cost</u>, per hour of labor, not the net wage the worker receives). Hence we want labor's wage to be gross of payroll taxes, and capital's return to be gross of corporation income and property taxes. In principle, value added taxes should be assigned to labor and capital in accordance with their contributions to value added. Sumptuary and excise taxes could be divided in the same way, and assigned to the value added of the sector whose processes create the relevant taxable event (e.g., distilling liquor; manufacturing cigarettes, etc.) But this is one place where experience has led me to sacrifice a little purity of concept in favor of simpler and more communicable results. (The problem is that when capital is assigned its share of the excise taxes on liquor and tobacco, those industries invariably end up with huge calculated real rates of return $_j$, (which we then have to elucidate). A certain conceptual clarity can be restored by defining a GDP net of all indirect taxes falling on final products (ref. Christensen-Jorgenson, in Jorgenson (1995), but this in turn creates the problem of dealing with an unfamiliar aggregate as we explain the growth of total outputs. I hope readers can sense that I have no strong opinion here.

USING THE RATE OF RETURN IN GROWTH ACCOUNTING

17. In spite of the subtleties mentioned in the last three paragraphs, I believe I have described a simple and straightforward way of estimating the rate of return to capital, one that has in some variant or other been widely used by many economists in studies aiming at measuring the "social" or "economic" rate of return to capital in different activities and for an economy as a whole.

At this point I want to explore how to apply the rate of return in growth accounting. No issue with respect to the labor term is involved here, so for the work of this section I will simply treat that term as $_{ij}$ L_{ij}, implying that differences in labor quality are adequately taken into account and that the resulting residual term R in principle measures the number of GDP baskets saved during the period in question by real cost reductions.

18. The special characteristic of a rate of return approach was already hinted at in para. 13 -- the "product" in a genuine sense is the real value of output if there are no purchased inputs, and is simply real value added (nominal value added divided by, in our case, the GDP deflator) in the general case. And this has a powerful implication. The product of activity j can rise by 10% with a rise of y_j by that percentage, p_j remaining constant, or by a rise of 10% in p_j, with y_j constant.

Before reacting in horror to this implication, walk with me through a simple ex-post evaluation of an investment project. We buy copper bars at $1 a pound, hold them for a year and sell them at $1.50 a pound. Good economics tells us that this investment was both privately and socially profitable, provided that the relevant discount rate is less than 50% and that storage costs are negligible. Alternatively, we plant forests to two varieties of trees; at the end of a given period one produces 1 million board feet of lumber at a price of $2 per board food, while the other produces 2 million board feet at a price of $1 per board foot. If the relevant profiles of costs for the two projects are the same, we should be indifferent between them, both from a private and a social point of view.

Some important lessons that students learn in courses in project evaluation are that

a) project profiles must be defined in real terms

b) It is not ever correct to use a Laspeyres index approach to measuring the profile (i.e., $F_{jt} = p_{jo}y_{jt} - {}_{i}\, w_{ijo}\, L_{ijt}$, where p_o and w_o are nominal prices at time 0, is simply wrong).

c) The correct way to measure the flows that constitute the project profile is by deflating nominal flows by a general deflator, yielding

(6)
$$F_{jt} = (P_{jt}y_{jt} - {}_{i}\, W_{ijt}L_{ijt})/{}_{t}$$
$$= p_{jt}y_{jt} - {}_{i}\, w_{ijt}L_{ijt},$$

where p_{jt}, as before, is $P_{jt}/{}_{t}$ and w_{ijt} is $W_{ijt}/{}_{t}$.

These principles must rule in the calculation of rates of return, just as they must in ex-post project evaluations.

But when they do rule for calculating , and we move to a growth accounting framework, are we not bound to stick with the definition of product as (py) and not as simply (y)? This question is explored in the following paragraphs.

19. The above question can be answered in three different ways, depending on the level of aggregation at which we are working. First, suppose we are explaining growth in GDP itself (a national or sometimes regional aggregate). The key to our story in this case is that, so long as the corresponding GDP deflator is used as our standard deflator throughout the analysis, there is no distinction between (py) and (y), for p -- the nominal GDP deflator deflated by itself -- is identically equal to one. This means that, at the aggregate level we have

(7a) $R' = (py) - {}_{i}\,{}_{i}\, L_i - (+)\, K$
$$= y - {}_{i}\,{}_{i}\, L_i - (+)\, K, \text{ and}$$

(7b) $R' = {}_{i}\, L_i\, {}_{i} + K\, (+)$

This reduces to a quite standard formulation. The only difference that I can see is that the conceptual framework for using had led us to express our capital stock in production units rather than investment units, (i.e., we have used the general GDP deflator to deflate all the annual investment flows that were used to derive K via the perpetual inventory method).

Table 1 is drawn from the Ph.D. dissertation work of Harald Beyer (1996). It shows the rates of growth of aggregate capital stock for a number of countries using on the one hand the methodology set out above, and on the other the methodology developed and applied by Jorgenson and his collaborators in recent years. This methodology is much more sophisticated than what I would call the "rate of return" approach. In my opinion, it can be said to represent current best practice. That approach uses several investment deflators for different parts of the capital stock, together with the standard GDP deflator for output.

The results of Table 1 are plotted in Figure 1. I believe that the differences in concept and methodology that are being explored here are simply "no cause for alarm". The attributions of components of growth and the calculations of TFP indexes as between one methodology and the other are not going to be different enough to worry about, regardless of which methodology one might, at bottom, prefer.

20. Second, suppose we are explaining growth at the level of sectors, industries or branches for which GDP originating (gross value added) is already broken down into a quantum (y_j) and a price (p_j). Here we have an option, but no real problem. If we want to maintain the "project evaluation" flavor of the rate-of-return approach as I have described it, we use:

(8a) $R = (p_j y_j) - {}_i {}_{ij} L_{ij} - ({}_j + {}_j) K_j$

(8b) $R = {}_i L_{ij} {}_{ij} + K_j (+)$.

Note that when we do this, there is no term for change in the relative price of the product (value added) of the sector. Some may consider this a serious flaw, but first one should note

TABLE 1

Comparison of Rates of Growth of Capital Stock*

(percent per year, 1971-91)

Country Stock	Using Investment Deflator	To Build Capital Stock	Rate of Growth of Capital Stock Using General GDP Deflator To Build Capital
Canada		5.3	4.7
France	3.7	3.9	
Germany		2.8	2.7
Greece	4.5	5.6	
Ireland	5.2	5.5	
Italy		4.6	4.6
Japan		5.4	5.1
Netherlands		3.2	3.4
Norway		3.2	4.3
Portugal		5.4	6.5
Spain		5.7	5.5
Sweden		3.0	2.5
United Kingdom		2.9	2.7
United States		3.5	3.5
Taiwan	9.1	9.4	

*In developing these estimates, Beyer used the following depreciation rates, obtained as "unweighted averages of the depreciation rates of more detailed asset types presented in Hulten and Wykoff (1981, Table 2) and Jorgenson and Sulivan (1981, Table 1)".

Residential Buildings	1.3%
Non-Residential Buildings	2.9%
Other Durable Structures	2.9%
Transport Equipment	18.2%
Machinery & Other Equipment	13.;8%
Inventories	0%

He notes that "The results are relatively insensitive to the depreciation rate". For excluding the income from land from the calculation, the assumptions mentioned earlier in this paper were used, imputing to land one third of the GDP originating in agriculture and one tenth of the actual and imputed rents to residential dwellings.

Source: Beyer (1996).

FIGURE 1

Comparison of Rates of Growth of Capital Stock that, over the whole GDP, $_j$ $(p_j y_j)$ should exactly equal the change in real GDP. And R, again summed over the whole GDP, should give us a perfectly sensible and respectable estimate of the impact of real cost reduction on GDP growth in the period in question.

But we have another option, which can bring us into a virtual alignment with the traditional methodology. This option is to break $(p_j y_j)$ into $(p_j\ y_j + y_j\ p_j)$, which can easily be done if we have data on p_j and y_j separately. Then we can calculate

(8c) $R = p_j\ y_j - {}_i\ w_i;\ L_{ij} - ({}_j + {}_j)\ K_j$

(8d) $R = {}_i\ L_{ij}\ w_{ij} + K_j\ ({}_j + {}_j) - y_j\ p_j$

Comparing (8a) with (8c) and (8b) with (8d), we quickly see that $R = R - y_j\ p_j$. Thus, if we don't like the presence of a relative price change term in $(p_j y_j)$ in (8a), we simply take it out. And if we don't like the absence of a relative price term (reflecting the benefit of real cost reduction in industry j that is passed on to consumers of its product), we simply append the term $- y_j\ p_j$ to the R we may have already calculated in, say, (8b).

But now consider adding up expressions for R of the form (8d), over all the activities composing the GDP. What we should get is a $_i R$ that is just equal to the R that we would get, working from (8b). That is to say, $_j y_j\ p_j$ should, in principle, be identically zero, since the weighted sum of relative prices should always equal one. The benefits that consumers (or users) get from relative price reductions of some components of GDP have to be balanced by the costs that consumers (or users) taken as a group have to bear as a consequence of increases in the relative prices of other components of GDP.

21. As between (8a) and (8d) I much prefer the latter, so long as I feel confident that the relative price term $-y_j\ p_j$ is reasonably accurate. My preference here is due to the fact that consumers in the end <u>are</u> the main beneficiaries of technical progress. Indeed, equations like (8d) provide the way to explain most trends in relative prices over time. Assume there is not much long-term trend in , so real cost reduction will either be reflected in real wages of in p_j. But, as mentioned before, real wages are determined in a broad "aggregate" labor market -- wage rates rise for lucky and unlucky industries alike. So on the whole rising real wages operate to press all relative prices up, while specific real cost reductions in each industry operate to press its relative price down. Since overall the average relative price change has to be zero, this means that "typical" products will have no relative price trend, those with RCR less than average (haircuts, taxi rides, building construction) will have relative prices that rise through time, while those with RCR greater than average (television sets, computers) will have relative prices that fall through time. This vision of relative price formation is intimately linked to the formulation (8d), and helps explain my preference for it.

22. But the case in favor of formulation (8d) fades as one comes to have more doubts as to the validity of p_j as an appropriate measure of the genuine relative price of y_j. In most countries, deflators by sector and industry do not even exist in the national accounts. And where they do, they are often the result of investigators simply grasping at whatever straws are to be found -- prices of two or three products representing "electrical machinery", or prices of plate glass, bricks and marble tiles being used to represent "stone, clay and glass products". Such extremely rough approximations can spell trouble even when we are analyzing the growth of the industry in the label; it is far worse when we are analyzing very diverse component branches of that industry and still worse when we get down to the level of the firm. What sense does it make to use the same "machinery price index" to deflate the output of three separate firms, one making turbines, the other elevators, and the third precision lathes? Nor would it make sense to use a non-ferrous metals price index to deflate the outputs of four different mines, one producing aluminum, another copper, a third lead and a fourth zinc. (This would not be a real world problem because separate prices of these metals are readily available, but it well illustrates the problem because there is an enormous amount of divergence in the time paths of the relative prices of individual metals.)

23. Once we are at the firm level, we also see quite clearly that the capital assets of firms differ vastly, even among firms flying under the same industry label. Certainly it makes little sense to use the same capital goods price indexes to deflate the investments of each and every firm in an economy, seeking to generate a series on the quantum of capital goods used by each of them.

What a relief it is, then, to find a conceptual framework that simply does not require separate price indexes covering: a) the mix of capital goods employed by a firm and b) the mix of product items that constitute its output. For if we needed such indexes, we'd have a devil of a time creating them. Vastly different output and input indexes would be needed for Chrysler than for General Motors, for Monsanto than for DuPont, for Microsoft or even Compaq than for IBM. And these differences would pervade both the capital goods used by the firms as well as the output mixes produced by them. How do we measure the quantum and price of the services of the travel agents that sell us tickets or of the security firms that patrol our neighborhoods?

As I said, what a relief it is to simply be able to take each company's purchases of capital assets plus its net cash flow spent on increased inventories, and deflate it by the general GDP deflator as one goes about building the real capital stock of the firm. And yet another relief comes from being able to define the firm's flow of output by simply deflating its annual nominal flow of gross value added by the same general GDP deflator.

And, finally, there is the sense of comfort one gets from knowing that the framework one is using has deep roots in the theory of capital and in the longstanding practice of economic project evaluation. Applied over all activities, the outputs add up to the GDP and the individual measures of R for all component activities add up to a self-standing measure of RCR for the economy as a whole; one that is equal in principle to $_jR$ obtained using the more traditional methodology of (8d), a methodology we might like to use but are precluded from doing so except in rare instances at the firm level, simply for lack of adequate data.

24. But we should bear in mind that everything has its price -- every silver lining has its cloud, as it were. When we use the approach of (8b), we obviously have a residual that, through it has meaning, is different from the more familiar one at all levels lower than the aggregate. R and R represent two different ways of distributing the same overall real cost reduction among sectors, industries and firms. Working with R thus calls for a process of learning by doing, as we attempt to capture the subtleties and nuances of a different way of doing things.

Beyond that, we should recognize that in most straightforward applications of the rate-of-return methodology to firms, we would probably quite naturally count acquisitions of land along with buildings and machinery as part of the investment process, and we would correspondingly include the income imputable to land as part of the income accruing to capital for purposes of measuring the rate of return. This represents the opposite choice from the one I prefer to work with when using national income accounts (see para. 14 above), but it is a perfectly sensible choice, nonetheless.

Then there are all the complications posed by financial assets and liabilities when we are working at the firm level. Somehow the national income accountants solve that problem for us before presenting us with data on GDP originating by sector and industry, but it is a problem we have to face for ourselves when we deal with individual firms. My own prudential sense has led me to avoid firms with heavy loads of financial assets and/or liabilities in my own experiments at the firm level. But this is only so that I can build up a backlog of experience before tackling the knottier cases. In the end one must find a way to handle such cases, either by developing clear conventions or by applying certain broad principles on a case-by-case basis.

25. The rate-of-return approach may also bring us new vistas. I am particularly fascinated by the role of working capital in our overall story. And by working capital I do not mean financial assets, even through working capital can sometimes take this form. I really mean the difference between an investment profile developed using standard cash-flow analysis on the one hand, and an investment profile that counts only purchases of durable assets (including net purchases of inventories). I can think of business starting up that own nothing more than office equipment. Yet they may take years to earn a profit. What abut all the negative cash flows sustained during those years? To me, they are just as much "real capital investment" as are the buildings and machines that a different firm might have bought to generate a similar profile of negative real cash flows.

The above example is a particularly easy case, because it is easy to identify the negative cash flows that represent investment. But if similar "investments" take place as part of the ongoing activities of a profitable going concern, they will normally be reflected in lower current profits and will be hard to identify as investments. Is there any chance of doing with working capital outlays generally the same sort of thing that some authors have done with R&D expenditures -- i.e., take them out of current expense and treat them as additions to the capital stock (with or without subsequent depreciation)?

I leave these thoughts as something worthy of future reflection and experimentation. My only point here is that the rate-of-return framework provides a natural base from which to start such pursuits.

26. A further observation concerns the backward-looking nature of the rate-of-return approach, and I believe all (or nearly all) other approaches to growth accounting. When we measure $_{jt}$ it is the return to capital accruing in period t, divided by the cumulated net investment built up over previous years. This rate of return can be 25% for decades on end, and it can also turn out to be 2.5% for year after year. Does this mean that firm A is likely to be truly generating a 25% return, and firm B a 2.5% return on new investments? Hardly likely. More plausible is the interpretation that firm A made some wise (or lucky) choices early on, and generated cash flows that were high in relation to its cumulated investment. If firm A were sold, it might sell for 2 or 2 1/2 times the value of K that we calculate for it. On the other hand, firm B's situation likely reflects past mistakes (or bad luck), with the result that firm B's owners might be able to sell it on the marketplace for only maybe a fifth or a quarter of the value of K that we develop for it.

I do not propose modifying the rate-of-return method on the basis of the above observation. To try to build in the market's revaluations of "the value of the firm" would open a Pandora's Box of problems to which no simple solution could be found. The rate of return method as it stands is as firm as a three-legged stool, being solidly rooted in the techniques of ex-post project evaluation. What we have to do is recognize that low rates of return can reflect past investments that went bad, and high rates can reflect past investments that thrived.

27. The attribute of ex-post calculation applies not just to the rate-of-return approach but to any method based on a formulation like $d\log y - s\ d\log L - s_k\ d\log K = d\log (TFP)$, unless the factor shares s and s_k are imposed rather than derived from the data. Most of such formulations of which I am aware follow a Divisia-type procedure in which a time path is divided into a series of small steps, each of which starts with a fresh calculation of s and s_k at the beginning of the step (or averages s and s_k from the beginning and the end of the step). Suppose that a normal rate of return is 10%, which in turn engenders a capital share equal to 40% of output (of the firm or industry being studied). In a series of years of negative real rate of return, that 0.4 share might be cut to only 0.15, and attributions of a capital contribution equal to 0.15 dlog K will in a sense be attributing to the new investment that takes place the very same negative real rate of return that corresponds to a capital share of 0.15. The rate-of-return method will do exactly the same thing when it attributes to new investment a $({}_j+{}_j)$ equal to, say, a "measured" 3% rate (implying ${}_j$ = -.02) compared with, say, a "normal" 15% (implying ${}_j$ = .10).

It seems to me that we are prisoners of this problem as long as we use a Divisia-type approach based on actual contemporaneous measurements of either (+) or s_k. The natural alternative is to find some way to attribute to new investment a rate of return different from that on the contemporaneous capital stock, as built up from historical costs by the perpetual inventory method.

Let us stop to think about this problem. Suppose we calculate a real rate of return of, say, 50% (or its counterpart in a share of capital equal to, say, 70%) for a firm. What do we really gain by attributing such a high rate of return to the new investment that is currently taking place? And if we calculate a negative real rate of return (or its corresponding very small share of capital) what do we gain by attributing that negative rate of return to new investments that are being made? In point of fact, most investments made by companies with very high observed rates of return on accumulated historical investment are made, not with a management turning down all investments that promise less than a high rate of return, but rather with a calculation of whether such investments will bring in more than they cost (i.e., investments are made under a criterion of net present value at something which the firm views as its "real cost of capital".) The real cost of capital will not be 50% for the firms that are super-profitable, nor will it be negative for firms currently making losses. But what will it in fact be? I certainly would not want to assume that all firms in the whole economy are aiming at an after tax real return similar to that on government bonds (from which we could build separate estimates of ${}_j$ based on the different tax structures of the different firms, industries, or sectors). Maybe we could come closer by measuring a broad historical real rate of return for each major sector (corporate, noncorporate, housing, etc.) and attribute this rate of return to all new investments everywhere in that sector.

<u>**28.**</u> My inclination at the moment, however, is to think of simply imposing sensible "expected rates of return", telling readers what one is doing, and providing ample explorations of the sensitivity of results to the assumptions. One might, for example, impose a real rate of 4% in the housing sector, of 8% in the noncorporate sector and of 12% in the corporate sector, and then try again with 6, 12, and 18 percent to test for sensitivity.

On the whole I feel this is a very promising additional line of approach to the decomposition of economic growth and the measurement of TFP. But in order to appreciate it one should first go through an exercise deriving the "dual" expression for the residual. Consider the "traditional" equations:

(5a) $R = p_j\ \ y_j - {}_i w_{ij}\ \ L_{ij} - ({}_j + {}_j)\ K_j$

(5b) $R = {}_i\ L_{ij}\ \ w_{ij} + K_j\ \ ({}_j + {}_j) - y_j\ p_j$

Now, for the reasons set out above, impose a sensible target rate of return $_j$ in place of the measured $_j$:

(9a) $R = p_j\ y_j - \ w_{ij}\ L_{ij} - ({}_j + {}_j)\ K_j$

This is saying it makes more sense to attribute, say a 10 or 15 percent rate of return to new investments than either a 50% or a negative rate, when these are what emerge from our calculations.

Now consider the dual representation of (9a).

(9b) $R = \ L_{ij}\ \ w_{ij} + (K_j\ ({}_j + {}_j) + ({}_{j-j})\ K_j - y_j\ P_j$

The underlined segment <u>must</u> be included in (9b), because in moving from (5ab) to (9ab), we obviously set things up so that $R = R + ({}_{j-j})\ K_j$.

To me, the "dual" representation (9b) was a lot less easy to swallow than the original (9a). In particular, I worried about the added term $({}_{j-j})\ K_j$. If $_j$ is 20% and $_j$ is 12%, does this mean that we are "imposing" a constant extra increment of .08 K_j/p_jy_j to the "regular" TFP growth rate? The answer to this is mainly no. If we take $_j$ as a correct measure of the marginal productivity of K_j, the answer is emphatically no. What this extra term does is correct for a false attribution of the contribution of K_j to growth, obtained when we use the calculated $_j$ (or the calculated s_{kj}). The answer is less emphatic, of course, to the extent that the imposed $_j$ measures the true marginal productivity of K_j with error. The answer to the question becomes a pure Yes if the $_j$ was really a correct measure to begin with, and substituting $_j$ for it was a complete mistake. Certainly it is dumb to impose a $_j$ unless one has reasonable confidence that it is better than $_j$ as a measure of mpk. To the extent that it is, it helps correct errors in the estimates of R based on $_j$'s or s_{kj}'s that do not adequately reflect the likely mpk of new investment.

<u>29.</u> Let me now try to give an <u>overview of some key points</u> we have covered:

a) Most <u>traditional calculations</u> have built up time series on capital stock using a deflator representing capital goods prices. Rates of return can be calculated from those capital stocks, but only after making price adjustments to express capital and the income from it in the same units. Typically rates of return have not been calculated, let alone emphasized in this literature.

b) The Rate of Return Approach: Shifting to a scheme in which a single standard deflator is used to express investment, capital and output in the same units puts the calculation of the real rate of return on center stage and permits us to express capital's contribution to the GDP growth rate as (+)(K/y), where is the rate of return to capital and (K/y) is the rate of net investment out of GDP. At the aggregate level we can do the necessary calculations working from the nominal national accounts, with no relative price (as distinct from wage) information beyond the GDP deflator.

We can also work without relative output prices at subaggregate levels like sectors, industries and firms. Where the entity we work with has good information on relative prices we can use that information to generate a residual that is quite "traditional":

(5a) $$R = y_j - {}_i w_i L_i - ({}_j + {}_j) K_j$$

If we do not have adequate information in relative product prices, the rate of return approach will give us

(8a) $$R = (y_j p_j) - {}_i w_i L_i - ({}_j + {}_j) K_j$$

The two residuals R and R are related by the equation

(10) $$R = R - y_j p_j.$$

The interesting thing is that ${}_j y_j p_j = 0$, when summed over all components of gross value added in the national product accounts. Hence both R and R represent sensible alternative ways to split up aggregate real cost reduction among sectors, industries, and firms.

The rate of return approach has its best justification at highly disaggregated levels where there is very little chance of getting decent data on p_j at a reasonable cost. I think of it as the "method of choice" for doing breakdowns of growth for large numbers of firms in an expeditious and easily communicable way.

c) Imposing rates of return in assigning capital's contribution to growth. There has been, to my knowledge, little discussion of this alternative. I feel there is little need for it at the aggregate level where we mainly capture strong central tendencies. At the disaggregated level, both a) and b) attribute a contribution of newly added capital on the basis of the observed $_j$ or the observed share of capital s_{kj}. These vary very widely in disaggregated data, and we can properly doubt that investment decisions are taken using the observed $_j$ (or the one implied by an observed s_{kj}) as a target or criterion rate of return. When we have serious doubts on this score, we should consider imposing what we think is a sensible rate $_j$ reflecting the expected marginal productivity of new investment. Work actually doing this lies pretty much in the future, but thinking along these lines is enough to give us some humbling second thoughts concerning the limitations of approaches a) and b).

THE LABOR CONTRIBUTION TO GROWTH AND THE 2-DEFLATOR METHOD

30. I have characterized the rate of return approach by the use of a single numeraire or standard deflator to deflate investment, value added, income, capital, etc.; that numeraire is also the metric in terms of which relative prices p_j and real wages w_i are measured.

Going one step farther converts the rate-of-return approach into what I have called the 2-deflator approach. This approach settles upon a basic or standard real wage w^*, ideally applicable to a low-skilled type of labor that we call "the standard worker". The fundamental idea is to measure the quantum of labor in these standard units. The variable w^* (the real wage of the standard worker) is the second deflator of the 2-deflator method.

31. Let us start with the oldest treatment of labor's contribution to growth. This is simply L where is the observed average wage (over the whole labor force) and L is the change in number of manhours. This, as mentioned earlier, "assumes" that the average quality of the new workers is the same as the average quality of the pre-existing labor force. Hence any change of quality ends up in the residual term.

Disaggregation is the preferred way to get around this. When labor is broken down into more homogeneous cells we can express its contribution to growth as $_i w_i L_i$. All sorts of changes in the compositional mix of the labor force are captured in this measure of the labor contribution. This line of approach has been pioneered and brought to a high art by Dale Jorgenson and his associates [see Jorgenson (1995)]. Their efforts have led to a breakdown of labor by 2 genders, 5 educational categories, 20 industrial categories plus 10 groups for occupation, 8 for age, and 2 for employment status, in their work for the United States. I do not have the corresponding count for their work on other countries, but it is easy to see from the U.S. example how monumental a task is involved in going down this road. Not only is the task a huge one in terms of sheer work; its data requirements are such that many interesting cases of productivity analysis simply cannot be pursued for want of data.

I have always been an advocate of the approach followed by Jorgenson and his collaborators. However, an appreciation of the huge difficulties of implementing this approach on a wide scale is what led me to turn to the 2-deflator approach.

32. Let us suppose, initially, that it is our objective to get an absolutely ideal breakdown of labor into categories, and using that breakdown to define labor's contribution to growth as $_{ii} L_i$. Now this idealized breakdown would be much finer than Jorgenson's. Wages differ not only by industry, broad occupational groups, gender and age, but by many other dimensions. There are cheap lawyers and expensive lawyers, high-quality CPAs side-by-side with accountants of only moderate skill. If we now consider the wages bill of any entity (firm, industry, sector, economy) as consisting of $w_{it}L_{it}$, (where t represents the time period), and we divide this by the wage of a standard worker w, we get $L = L_{it}(w_{it}/w)$ L_{it} would here measure the man hours worked by a given narrow category of workers, and (w_{it}/w) would measure their contribution per man hour, as a multiple or fraction of that of the standard worker. A highly skilled medical specialist might represent 20 standard labor units, an ordinary medical practitioner might represent 10, an average nurse might represent 3, a very experienced nurse might count as 5, the nurse's aide might count as 2, the hospital orderly as 1, and the sweeper perhaps as only 2/3 of a standard labor unit.

It does not matter how many different relevant labor categories there might be in the entity under study; there may be ten, or there may be a thousand. When we divide the wages bill by w^* we are counting each individual worker on the basis of his or her wage, as contributing the number of labor units represented by that particular (w_j/w^*).

The growth equation for real gross domestic product corresponding to (4) would now look like

(11) $$y = w^* L^* + (+) K^* + R^*$$

Here w^* and L^* are as defined above, while K^* is the capital stock obtained using the single numeraire-deflator $_d$ and R^* is the unexplained residual that emerges when the 2-deflator approach is used.

We are here concentrating on the first term in (11). It can quickly be seen by differentiating (11) that

(12) $$w^* L^* = w^* \,_i(w_i/w^*) L_i + w^* \,_i L_i \, (w_i/w^*)$$

In the happy event in which the relative wages (w_i/w^*) of the different categories do not change, this reduces to:

(12') $$w^* L^* = w_i L_i, \text{ when } L_i (w_i/w^*) = 0.$$

As can be seen, (12') does not require that each and every (w_i/w^*) remain unchanged, only that their weighted sum is zero.

33. I am most of the time quite content to work with the assumption that the weighted average of the wage premia (w_i/w^*) does not change very much. But one should nonetheless address the question of how to interpret the change in the weighted average premium if it is significant. Quite obviously if $_i L_i$

(w_i/w^*) is positive, that fact makes labor's contribution to growth larger, and the residual correspondingly smaller. If the residual is interpreted as a grab-bag for all sorts of real cost reductions, then the increase in real labor costs (relative to the standard real wage w^*) operates as an offsetting force to whatever other real cost reductions are taking place. In short, the residual R^* can still be interpreted as reflecting net "real cost reductions", if one counts the rise in the average wage premium as an increase in real cost.

It should be clear from the above that the use of the standard labor unit to obtain L^* produces genuinely <u>desired</u> results in giving us $_i(w_i/w^*)\, L_i$ as a part of $w^* L^*$, and where it gives us more than this, the results are not hard to interpret and are fully compatible with the view that R^* is mainly a reflection of net real cost reductions. At the same time one quickly sees how the results one would get from this analysis would be different, if one chose the basic labor unit differently.

Consider a simple case in which we have just three categories of labor, unskilled, skilled, and professional. Suppose, too, that in the period in question the skill premiums increase. Let me assume that if we choose the middle category as our standard unit of labor, the summation $L_i\, (w_i/w^*)$ comes out to be exactly zero. The contribution of labor is then represented by (12'), and there is no term contributing to the change in real costs.

If we instead choose unskilled labor as our standard unit the second term in (12) will be positive. More of the increase in output will be counted as due to increased labor, less will be attributed to a net reduction in real costs.

By the same token, if we choose professional labor as our standard unit, the second term in (12) will be negative. Now the reduction in "value" of skilled and unskilled labor, relative to the professional wage (now the standard) will increase the size of the residual R^* and should be interpreted as one more "real cost reduction".

Readers should recognize that the underlying reality does not change. What we have above are three different ways of reflecting the same reality depending on our choice of the standard labor units. This type of problem is typically present whenever one can choose among alternative numeraires, alternative types of index number, etc. If we calculated GDP or real personal income in terms of a sugar-price numeraire, we would have highly volatile time series for real output and real income. It would be very wasteful to do it this way because we would end up explaining most of the changes in output and income, thus calculated, in terms of what happened to the relative price of sugar!! To see ordinary cyclical fluctuations properly we would first have to correct for the vagaries of the sugar market. But note -- if we did so, and did so properly -- we should come out in the end with the same analysis of, say, cyclical movements. I repeat, the choice of numeraire does not change the underlying reality or its true explanation, but it can make our task of analysis much easier (if we choose the numeraire well) or much more cumbersome and difficult (if we choose the numeraire badly).

34. Here is a list of a few candidates for the basic labor unit.

a) The average wage of a well-specified, important group of "unskilled" workers. Consider, for the U.S., the category of male workers, 30-40 years old, with 10-12 years of education. For a middle-income country, change the education level to "complete primary". For a poor country, drop it to 0-4 years. Such a wage will rarely be available in annual time series, so it would have to be interpolated using one of the alternatives which follow.

b) The average wage of workers in a particular industry or industries, ideally in certain job classifications that pretty much guarantee a low level of skill. Textile operatives, workers in the apparel and/or shoe industries are good examples. The ILO provides a fairly good set of time series on wages by industrial category; it is available for many, but by no means all developing countries.

c) A "formula wage" aimed at reflecting what our unskilled worker might approximately earn. We have used 2/3 of one annual per capita GDP of the country in question in quite a number of exercises, without any snake rising up to bite us, as it were. This number was picked as the sort of full-time annual earnings that would clearly be identified with a very low skill level. The advantage is that one can use it as a standard assumption, even in countries for which good labor market data are not available. The disadvantage is that it is not drawn directly from the labor market and does not in fact track the earnings history of any well-specified group of workers.

d) For the United States, a fourth alternative has come to our attention. It is a wage index developed by Finis Welch and Kevin Murphy, geared as a constant-quality index covering the entire labor force -- i.e., built up from a weighted average of the percentage periodic increments of wages spanning the whole spectrum of skills. Such an index could be directly used in the definition of L*, or it could be calibrated to the base period wage of a specific category of unskilled workers.

35. Whatever choice is made for the standard labor unit, one must bear that specific choice in mind when interpreting the results. But note too that whatever choice is made for w*, the expression w* L* always includes the first term of (12), which is the exact disaggregated measure that incorporates attributing to each new worker a marginal product equal to his or her corresponding wage, and dealing with compositional shifts (including those due to education and training) in exactly the same way. This term taken by itself provides a solution to the disaggregation problem which is far more detailed and refined than even Jorgenson's heroic efforts have given us.

One of the significant uses of w* is to permit one to quickly divide the labor contribution into a part (w* N) due to the increment of "raw labor power", or "brute force", (here N represents the labor force measured ideally in simple man hours or man years, unadjusted for skill), and a second part w*(L*- N) due to the increment of human capital, very broadly interpreted. This latter component includes the greater market productivity that stems from education and training, and it also incorporates increases due to improvements in aptitude, and to the greater scarcity value of some skills or employments. At the same time it also builds in the loss of market value of specific skills (like super-accurate typing, which lost much of its marketability when easy correction became possible on the word processor).

Out of the term w*(L*- N) we can also separate what we call a "maintenance component" w*[(L*-N)/N] N, that endows new labor units (N) with the pre-existing average "human capital contribution" [(L*-N)/N]. The remaining piece of w*(L*- N) can then be interpreted as the "quality improvement component", once again very broadly construed.

36. Although some people find our use of L* and w* odd, I think it fits quite nicely into a more traditional way of dealing with labor's contribution. Let us consider that if people have a problem with the w* approach, it probably concerns cases of significant widening or narrowing of skill premia, or, more generally cases where $_i L_i(w_i/w^*)$ differs from zero to an important degree. An obvious remedy is to disaggregate the labor force into components. In some Latin American countries, firms actually keep separate payrolls for blue collar (obreros) and white-collar (empleados) people, so there it would be easy to work with w, and w which would end up giving us an R = K* * + L w + L w, and an R equal to the same thing minus y_j p_j (if we had a reliable p_j to work with).

Then comes the question, if two w*'s are okay, then why not 10, or 20, or 200? This would certainly be difficult at the firm level, but not necessarily at the national level in some advanced countries. Consider the United States, and suppose that we could divide the total wages bill (augmented by the labor of self-employed and family members) into, say, 20 categories. Suppose, too, that we were still worried about the significant heterogeneity of labor within each category, and that in addition we didn't have a good sense of the actual number of hours worked within each category (problems of part-time and full-time, and also categories with high turnover where the data show the number of individuals that received payment in a period, not the number of "jobs".) What better solution for such a case then to find twenty hourly wage rates w, and divide each of them into the corresponding wages bill to form a quality-corrected quantum of labor L? The expression $_i$ w L would in this case give us an exact representation of the labor contribution if within each of the 20 cells, the structure of relative wages had remained constant, or even if just the within-cell "average premia" ($_i$/w) had stayed the same.

My question now is, what would come out if we were to perform such an exercise for the U.S., following exactly the categories of labor used by Jorgenson and his associates, and employing the same numbers as they use as representative wages within each cell (defining a cell as that unit, within which they assume the structure of relative wages to be constant)? My guess is that the results of our calculation of $_i$ w L would be very close (if not identical) to their estimate of the labor contribution. (One of my hopes is to actually get to perform such an experiment in collaboration with Jorgenson.)

If I am right in this last conjecture, then the main difference between my use of a single w* in the 2-deflator approach and the well-established methodology of Jorgenson consists in the number of cells that one works with. It is not a profound difference of principle.

37. Table 2 and Figure 2 give a comparison of results as between the 2-deflator approach and more traditional methods. On the whole the comparison is quite reassuring. Recall that I am not putting the 2-deflator method forward as being more refined and accurate than the traditional methodology (particularly in its best applications). Rather, I present it as opening new vistas because of its ease of application, particularly its lower data requirements.

Note that the differences in estimated rates of TFP growth exceed one half percent only for Korea, Taiwan and the United Kingdom. And even these three countries do not appear as glaring outliers in the scatter diagram of Figure 2. All in all, I feel that these comparisons

TABLE 2

Comparison of Rates of TFP Increase

(11 Countries; 2-Deflator Method Compared With Other Studies)

	2-Deflator Method (Beyer)	Other Studies
Canada	.0064	.0046[a]
Colombia	.0074	.0120[b]
France	.0099	.0145[a]
Germany	.0129	.0158[a]
Hong Kong	.0229	.0230[c]
Italy	.0127	.0177[a]
Japan	.0150	.0196[a]
Korea	.0238	.0170[c]
Taiwan	.0368	.0260[c]
United Kingdom	.0022	.0130[a]
United States	.0023	.0041[a]

Sources: a) Dougherty (1991) as quoted by Barro and Sala-I-Martin (1995). This study covers the period 1960-90.
b) Elias (1990). This study covers the period 1940-80.
c) Young (1995). This study covers the period 1966-90 for Taiwan and Korea, and 1966-91 for Hong Kong.

Most of the differences can be explained by the different periods that the several estimation cover. The 1960s were in general a period of larger growth than the decades that came afterwards. In general our estimations are consistent with the so-called productivity slowdown in the 1970s. The contribution to TFP is lower in our study than in studies with time spans that reach years in the decades before the 1970s. The differences in the results exceed one half percent per year only for Taiwan, Korea and the U.K. Differences in time period and in sector coverage may account for some of the discrepancy for the cases of Korea and Taiwan. Data in the Tables are taken from Beyer (1996) who used average wages of males working in the manufacture of wearing apparel (ISC322).

FIGURE 2

Rates of TFP Increase

2-Deflator Method Compared With Other Studies give ample support to the further use of the 2-deflator methods. One should bear in mind, of course, that Table 2 and Figure 2 do not present ideal comparisons. These would entail using exactly the same data and concepts wherever possible, trying to develop a comparison in which the only sources of difference would be methodological. Such comparisons are on our agenda for the future.

38. Table 3 helps to illustrate some of the advantages of focusing on the rate of return to capital. Let me say from the outset that one can also get measures of the rate of return starting from the traditional methodology -- but one does not need them if one works with

$$\text{dlog TFP} = \text{dlog } y - s \;\text{ dlog } L - s_K \text{ dlog } K.$$

And, in fact, few authors using the traditional approach report having calculated rates of return. The rate of return approach and the 2-deflator approach (which is sort of an augmented rate of return approach), on the other hand, directly involve the calculation of the rate of return, and thus invite us to focus more attention on it.

Table 3, like the preceding two, is adapted from Beyer's work. He carried out an analysis of the growth experiences of 32 countries ranging from Sri Lanka to the U.S. on the income scale, and from Iceland to Australia in the geographic scale. In Table 3 we present results for his 10 countries with the highest and for his 10 countries with the lowest GDP growth rates from 1971 to 1991. In the second column we show the calculated average annual rate of return. In the third column we have capital's contribution to the growth rate, and in the final column the estimated average annual rate of TFP growth, all over the same time period.

This table highlights the usefulness of looking at the rate of turn in analyzing growth. (There is, by the way, much less reason to worry at the aggregate level about the potential great discrepancies between calculated rates of turn and the expected marginal productivity of new

TABLE 3

Growth Rates, Rates of Return, and Rates of TFP Improvement

(Selected from a sample of 32 Countries)

Ten Fastest-Growing Countries	GDP Growth Rate	Rate of Return	Capital Cont. To Growth Rate	TFP Growth Rate
Taiwan	8.83	15.0	3.81	3.68	
Korea	8.47	13.2	4.30	2.38
Thailand	7.65	12.5	3.68	2.96
Hong Kong	7.91	20.0	3.56	2.28
Ecuador	5.58	14.0	2.70	0.36
Cyprus	5.12	10.6	2.99	1.92	
Zimbabwe	4.62	13.6	2.42	0.97
Colombia	4.43	11.3	1.99	0.74
Iceland	4.35	9.4	1.95	1.77	
Ireland	4.12	6.7	1.70	0.36	
Median	5.35	12.85	2.84	1.83	
Mean	6.10	12.63	2.91	1.74

Ten Slowest-Growing Countries	GDP Growth Rate	Rate of Return	Capital Cont. To Growth Rate	TFP Growth Rate
Austria	2.87	5.1	1.13	1.29	
France	2.80	6.1	1.21	0.99	
Germany	2.60	6.3	0.97	1.29
Belgium	2.56	6.8	1.06	1.60
Netherlands	2.52	7.0	1.12	0.83
United States	2.52	9.1	1.20	0.23
South Africa	2.16	7.5	1.58	-0.97
Denmark	2.15	7.5	1.01	0.82
United Kingdom	2.12	9.6	0.95	0.22
Sweden	1.84	4.3	0.66	0.24
Median	2.52	6.90	1.09	0.825	
Mean	2.41	6.93	1.09	0.661

Source: Beyer (1996), Tables III.1.1 through III.1.32; also Appendix I for rates of return.

investment than there is at the level of the firm. See paras. 27 and 28.) We found here an unequivocal tendency for fast growing countries to be experiencing high rates of return as well as high capital contributions and high rates of TFP improvement. This is all the more interesting because in the calculation of TFP a higher level of the rate of return operates to reduce the calculated TFP (i.e., is a positive component of R' and should presumably be positively correlated with it, but R' is found by subtracting K from y; hence in a sense the level of should presumably be negatively correlated with R'). What we are seeing here, in my opinion, is a genuine syndrome in which all sorts of good things go together. Strong real cost reductions and high rates of return create attractive investment opportunities, which, when acted upon bring about a high capital contribution to growth. It should be no surprise that under such circumstances the GDP growth rate itself tends to be high. It should likewise be no surprise that the opposite syndrome -- with weak real cost reductions and low rates of return producing fewer interesting investment opportunities -- should end up being associated with a low capital contribution and a low GDP growth rate.

COMPARING TWO METHODS

39. As mentioned earlier (para 22), a breakdown of the analysis at a fairly broad industry level gives us the possibility of using the 2-deflator method, yet ending up with an estimate of R, not R. This possibility arises when we have access to a time series in p_j (the relative price of the industry j's output), or alternatively to an output series for y_j which then can be used to replace the 2D output variable $(y_j p_j)$.

Such a comparison can be made on the basis of work done by Edgar Robles (1997). Robles took as his starting point the productivity analysis of U.S. manufacturing given in Jorgenson, Gollop and Fraumeni (1987). Specifically he used their data to obtain nominal GDP originating in each manufacturing industry, but he deflated this number by the general GDP deflator rather than the specific industry deflator used by Jorgenson, et al. Robles also notes a conceptual difference in the treatment of taxes and subsidies, where

Jorgenson operates from the producer's point of view (taxes being treated as costs), while Robles operates from a global economy point of view. Table 4 shows Jorgenson's TFP estimates in column (1), and those emerging from Robles's application of the 2-deflator method in column (2). The differences between these are shown in column (4), labelled "raw differences". Part of these raw differences obviously stems from the difference in output definitions, but this source of discrepancy can be easily eliminated. The differences in rates of output growth are shown in column (4) of Table (3) and the difference in TFP growth rates, adjusted to a common output definition, is shown in column (5).

The results of Table 4 are shown as histograms in Figure 3. To my mind, the 2-deflator method comes out quite well, even in the raw, or unadjusted form. The two outliers (petroleum and tobacco) are probably there simply because of a different conceptual treatment of taxes; certainly they merge with all the rest once differences in output concept are taken into account (adjusted differences). To put the story a different way, we see that 3/4 of the differences in TFP growth rates are within ± 1% per annum in the raw form, and fully 85% once one adjusts for differences in output concept. This is close enough to spur me to continue studying and using the 2-deflator approach, particularly so when one realizes that it is much easier to apply, and can be applied readily to a much broader range of data, than the Jorgenson approach. Having been in favor of the Jorgenson approach for decades, I do not want to appear antagonistic to it. The analogy that comes to mind is of Wedgewood china with Corelle dinnerware. Each piece of Wedgewood china is almost a work of art in its own right. Meticulous care is

TABLE 4

2-Deflator and Jorgenson Methods Compared

(U.S. Manufacturing Industries, 1948-1979)

Industry	Rate of TFP Growth Jorg. (1)	2D (2)	Difference in Of Output Growth (Jorg.-2D) (3)	Difference In TFP Growth Raw (1)-(2) (4)	Adjusted (4)-(5) (5)
All Mfs.	1.22	1.06	0.19	0.16	-0.03
Food, etc.	1.90	0.31	1.40	1.53	0.13
Tobacco, etc.	1.36	-1.46	1.30	2.84	1.54
Textile Mfgr.	2.44	1.11	2.50	1.33	-1.17
Apparel	2.31	1.65	1.82	0.66	-1.16
Lumber/Wood	1.31	1.70	-1.10	-0.39	0.71
Furniture, etc.	1.18	1.13	0.36	0.05	-0.31
Paper, etc.	1.13	0.71	0.24	0.42	0.16
Printing, etc.	1.02	1.67	-0.59	-0.65	-0.08
Chemicals	1.27	0.87	0.98	0.50	-0.48
Petroleum, etc.	0.03	3.06	-3.28	-3.03	0.25
Rubber, etc.	0.76	1.42	0.01	-0.66	-0.67
Leather, etc.	0.64	1.54	-0.27	-0.90	-0.63
Stone, Clay, Glass	0.62	0.99	-0.48	-0.37	0.11
Primary Metals	0.18	1.55	-2.07	-1.37	0.72
Fabricated Metals	0.76	1.49	-0.71	-0.73	-0.02
Machinery; nonelec.	0.63	1.52	-0.76	-0.89	-0.13
Elec. Machinery	2.16	1.35	1.25	0.81	-0.44
Transport Eq.	1.74	1.07	-0.25	0.67	0.92
Photographic Eq.	0.97	1.84	-0.08	-0.87	-0.79
Misc. Mfg.	0.54	0.96	0.07	-0.42	-0.49

FIGURE 3

Histograms of Differences in Rates of TFP Growth

(Jorgenson Minus 2-Deflator)

Range of TFO Growth Rate	Raw Differences	Adjusted Differences
-3.5% to -3.0%	X (petroleum)	
-3.0% to -2.5%		
-2.5% to -2.0%		
-2.0% to -1.5%		
-1.5% to -1.0%	X	XX
-1.0% to -0.5%	XXXXXX	XXX
-0.5% to zero	XXXX	XXXXXXX
zero to 0.5%	XX	XXXX
0.5% to 1.0%	XXX	XXX
1.0% to 1.5%	X	X
1.5% to 2.0%	X	
2.0% to 2.5%		
2.5% to 3.0%	X (tobacco)	

given to every stage of the process. It is also costly in terms of the resources used. Contrast that with Corelle dinnerware, mass-produced cheaply and efficiently by Corning and highly serviceable for the day-to-day needs of millions of families. The two products are not really competitive with each other; each has its own justification and place.

YEAST VS. MUSHROOMS -- PART I

40. Although I have been aware since the 1950s that real cost reduction tends to be concentrated rather than widely dispersed among industries, I did not really appreciate the extent and implications of that concentration until 1990, when I wrote a background paper (Harberger, 1990) for the World Bank's World Development Report, 1991. The key that opened the door to new insights was thinking about TFP growth (or, as I prefer, real cost reduction) not only in terms of percentage rates per year, but also in terms of the number of real dollars saved via RCR in different activities.

Table 5 illustrates how this modest shift in focus helps us see certain things more clearly. Column (1) presents the familiar measure of the percentage by which TFP grew, or real costs were reduced, during the period in question (note that the percentages apply to the period 1958-67 as a whole; they are not annual rates). To turn these percentages into dollar amounts of real cost saving over the period, we multiply them by base-period real GDP [col. (4)]. The results are shown in column (2). Columns (3) and (5) are the cumulative sums of columns (2) and (4) respectively. Working with these figures one can make statements like those at the bottom of the table -- i.e., the top 10% of industries accounted for 30% of total real cost reduction; the top 22% of industries (measured by initial value added) accounted for more than half of total real cost reduction.

TABLE 5

Concentration of TFP Growth Among U.S. Industries 1958-1967

(Columns (2) to (5) in Billions of 1958 Dollars)

	TFP Growth Over Period (1.0=100%) (1)	Absolute Amount of Real Cost Reduction [(1)×(4)] (2)	Cum. Sum of (2) (3)	GDP by Industry 1958 (4)	Cum. Sum of (4) (5)
Lumber and Wood Products	0.72	2.51	2.51	3.50	3.50
Railroad Transport	0.63	5.52	8.03	8.70	12.20
Textile Mill Products	0.61	2.49	10.52	4.10	16.30
Electrical Machinery	0.55	5.10	15.66	9.30	25.60
Transport Equipment	0.46	7.05	22.71	15.20	40.80
Chemicals	0.44	3.97	26.68	9.10	49.90
Public Utilities	0.42	4.65	31.33	11.00	60.90
Petroleum and Coal	0.41	1.27	32.60	3.10	64.00
Rubber and Products	0.41	1.23	33.83	3.00	67.00
Mining	0.41	5.20	39.03	12.60	79.60
Communication	0.40	3.61	42.64	9.00	88.60
Trade	0.33	24.93	67.57	76.40	165.00
There follow 18 more industries whose combined results are	0.03	7.53	75.10	239.80	404.80

Top 10%[b] of industries account for 30% of total TFP contribution
Top 22%[b] of industries account for 52% of total TFP contribution
Top 40%[b] of industries account for 70% of total TFP contribution

Source: John W. Kendrick and Elliot S. Grossman, Productivity in the United states: Trends and Cycles, Baltimore: Johns Hopkins University Press, 1980. GDP data from U.S. national accounts.

[b]These percentages are contributions to GDP of industries ranked according to their percent rate of TFP growth over period.

Readers will notice that at the foot of each column in the table is an entry referring to 18 additional industries, whose aggregate contribution to TFP was only 10% of the total, while their aggregate contribution to initial output was almost 60% of the total. (I have many times regretted lumping all these industries together in a single residual category -- more will be said on this point later.)

The analogy with yeast and mushrooms comes from the fact that yeast causes bread to expand very evenly, sort of like a balloon being filled with air, while mushrooms have the habit of popping up, almost overnight, in a fashion that is not easy to predict. The results of my calculations using the Kendrick-Grossman data pointed very clearly to a "mushrooms" interpretation. Not only were the contributions to RCR highly concentrated in a relatively few industries, these industries also were very different as one shifted from decade span to decade span. The top 4 branches in percentage of real cost reduction during 1948-58 were Communications, Public Utilities, Farming, and Miscellaneous Manufacturing. In 1958-67 they were Lumber, Railroad Transport, Textile Mills, and Electrical Machinery. In 1967-76 they were Finance, Insurance & Real Estate, Apparel, Communications, and Chemicals. Only Communications appears twice among these twelve listings.

Now to my mind, this already brings evidence to bear on a number of possible hypotheses concerning the nature of TFP improvement. Certainly some ways of interpreting a generalized externality due to improved education would be hard to justify using evidence like this. Strong links of the residual term to R&D expenditures would suggest a high degree of persistence among the leaders in TFP improvement. So also (probably) would economies of scale associated with the scale either of the firm or of the industry. Such economies are not likely to jump wildly around from one industry to the next, from period to period. One would expect them to embody characteristics of the productive process which would be relatively stable over time; hence they should show a reasonably high degree of persistence, over time, in terms of the TFP experience of particular industries.

41. No economist can look at Table 5 without thinking of its close analogy with a Lorenz curve. That indeed was the next step we took in trying to represent the degree of concentration of real cost reduction. Figure 4 (drawn from Robles (1997), shows the quasi-Lorenz curves for a 20-industry breakdown of the U.S. manufacturing sector over 4 successive 5-year periods.

What strikes one immediately about Figure 4 is the characteristic "overshooting". I have marked with the first heavy vertical line the point where the rising curve crosses 100% on the vertical axis. The interpretation is that in 1970-75 the cumulative real cost reduction of just 25% of manufacturing industries (measured by initial value added) was equal to the total RCR for manufacturing as a whole. After that there are other industries producing another 40% of the total, but their contribution is offset by still other industries with negative RCR during the period.

Corresponding to the 25% figure for 1975, we have 12% for 1975-80, 48% for 1980-85, and 40% for 1985-91. These are the fractions of manufacturing industry which by themselves were able to account for the full amount of real cost reduction during the respective period, in the manufacturing industry as a whole.

The second heavy vertical line in each panel of Figure 4 marks the maximum point of the curve. The interpretation is that about 64% of industries enjoyed real cost reduction during 1970-75, with the remaining 36% suffering real cost increases (declining TFP). For the subsequent periods the corresponding figures are 65(35) percent 78(22) percent, and 82(18) percent. Here the first figure is the percent of industries enjoying real cost reductions; the

FIGURE 4

Figure 4 (cont.)

figures in parentheses represent those experiencing declining TFP.

Some interest attaches to the ordinate of the maximum point on each curve. In the first period, TFP growth ended up accounting for close to 170% of the RCR for total manufacturing. In 1975-80 this figure was about 240%, in 1980-85 only about half that and in 1985-91 a little more than 125%. The trouble is that when the aggregate TFP contribution is relatively small, the cumulative total of the positive contributions is a large multiple of that aggregate, while when the aggregate is large, this multiple tends to be smaller. Thus, for 1970-75, and for 1975-80, the total RCR in manufacturing as a whole was only about 2.3% of initial manufacturing value added. In contrast, the total RCR for all manufacturing was almost 10% of initial manufacturing value added in 1980-85, and about 7.5% in 1985-91.

The problem obviously becomes greatly compounded of the real cost reduction for the aggregate (in this case total manufacturing) turns out to be negative. Special conventions would have to be established to make clear the interpretation of Lorenz-like diagrams in such cases.

I believe I have hit on a felicitous way of solving all these problems, and at the same time creating an even better, clearer visual representation of the degree of concentration or dispersion of real cost eduction among the components of an aggregate. The idea is simply to relabel the vertical axis of the Lorenz-like diagram, making it represent an annual growth rate. For simplicity, think of 30 line as representing 1% per annum of TFP growth. The rest of the vertical axis would be calibrated accordingly. Thus, by looking at the slope of a simple chord, we could visually assess how rapid was the TFP growth, of the aggregate in question.

Figure 5 is presented simply for didactic purposes. Here we have a hypothetical industrial branch made up of four industries, A, B, C, and D. First, we order the industries in descending order, according to their rates of TFP increase in the period. Then we calculate cumulative

FIGURE 5

real cost reduction (a real dollar amount) and plot it against cumulative initial real value added. Then we scale the vertical axis so as to comply with whatever metric we have decided upon for the TFP growth ate (in our example, a 30 line representing a 1% annual TFP growth rate), and the horizontal axis so as to add up to 100%.

In the lower panel of Figure 5 I give a few examples to show how these diagrams cope with the problems of a low TFP growth rate (the overshoot for the case of 0.25% growth would show up peaking at something like 800%, in a Lorenz-type diagram) and of negative TFP growth (where it is hard to even conceptualize a Lorenz-type picture).

I first presented these diagrams before a large audience at the Western Economic Association meetings in Seattle (July, 1997), and for that presentation coined the label of "sunrise diagrams", on their analogy with the sun rising over a hill. That same evening Yoram Barzel suggested that where the aggregate slope is negative, we apply the term "sunset diagrams", which I immediately accepted.

42. Figure 6 presents a set of sunrise-sunset diagrams based on Jorgenson, Gollop and Fraumeni (1987, pp. 188-90). These cover 32 industrial sectors (their 35 minus agriculture, trade, and government enterprises). I think the utility of sunrise-sunset diagrams needs no further championing, once these pictures are examined and digested. Practically all variants are represented in these real-world cases: low TFP growth with a huge overshoot (1953-57 and 1969-73); negative growth with large and moderate overshoots (1966-69 and 1973-79); moderate growth with small (1979-85) medium (1960-66) and large (1948-53) overshoots.

One striking fact that emerges from this set of pictures is how variable across periods is the negative contribution of the losers. If the losers had only contributed zero change in TFP, we would have had cumulative TFP contributions of about 0.8% per annum in 1948-53, in 1953-

FIGURE 6

Figure 6 (cont.)

Figure 6 (cont.)

57, in 1957-60, and in 1960-66. And the other periods would not have been much different: about 0.,7% in 1969-73, 0.6% in 1966-69, and 0.5% in 1973-79. Instead of this narrow range of cumulative contributions, we have an actual distribution that goes from -0.9% in 1973-79 through around 0.1% in 1953-57 and 1969-73 to over 0.5% in 1960-66 and 1979-85.

Does this not suggest that we make a major research push trying to improve our understanding of the phenomenon of negative TFP growth? What syndromes characterize the firms and industries experiencing it? How much of it stems from external shocks like international prices? How much from competition within the industry? How much of it represents firms struggling to survive, yet experiencing output levels well below their previous peaks (and presumably below installed capacity)? How much of it represents things like "labor hoarding" as firms go through periods of adversity?

YEAST AND MUSHROOMS -- PART II

43. I hope that in the previous section I have made a convincing case concerning: a) the usefulness of sunrise-sunset diagrams, b) the aptness of the "yeast vs. mushrooms" dichotomy and c) the pervasiveness with which the mushroom side of that dichotomy seems to come out ahead when the GDP is broken down into industries or industrial branches for TFP analysis. The grand design that emerges from the studies reported here, and from just about all the other industry breakdowns that I recall having seen, is that: i) a small to modest fraction of industries can account for 100% of aggregate real cost reduction in a period, ii) the complementary fraction of industries contains winners and losers whose TFP contributions cancel each other, iii) the losers are a very important part of the picture most of the time, and contribute greatly to the variations we observe in aggregate TFP performance, and iv) there is little evidence of persistence from period to period of the leaders in TFP performance.

44. The above results are, I think, very interesting (in the sense of piquing our curiosity), very strong (in terms of their implications about the nature of the growth process) and very robust (in the sense that they have wide applicability over different data sets analyzed by different authors using at least somewhat different methods. But these results, so far, are quite compatible with what I might call an "industry view" of the TFP story. This is the way I myself looked at the growth process until quite recently. I used to tell my classes that the heros of the TFP story were rubber tires and autos in the 1920s, refrigerators and other household appliances in the 1930s, pharmaceuticals (sulfa drugs, penicillin, etc.) in the 1940s, television in the 1950s, etc., etc., up to computers and telecommunications in recent decades. The image that I had in mind was one of yeast within each industry, and mushrooms between industries -- a commonality of TFP experience by firms within an industry, depending on that industry's luck in the technological draw, side by side with highly diverse experience between industries because the distribution of technical advances had wide dispersion, even for periods as long as a decade.

Getting access to data at the firm level permits one to explore whether this view is compatible with the actual experiences of firms and industries. We are just in the early stages of this exploration, but I think the result is quite clear already, namely, the "mushrooms" story prevails just as much among firms within an industry as it does among industries within a sector or broader aggregate. I will present here only a taste of the evidence from the United States (on which our systematic work just recently got started), our massive evidence comes from the Mexican manufacturing sector, for which Leonardo Torre has analyzed data from a sample of over 2000 firms. A small fraction of these firms were lost owing to missing data, but some 1900-firms remained in the sample that Torre finally worked with. These firms were divided into 44 branches of industry, so that on average we have about 43 firms per branch.

45. There are really too many ways to present such a mass of information as is contained in Torre's study. What we will do here is give the aggregate picture in Figure 7, and then show in Figure 8 three fast-growing branches, three of around median growth, and three from among the slowest-growing branches.

We also provide in Figure 9 a breakdown of firms by size group, to explore whether the TFP experience appears to be much influenced by that variable.

To complement these figures, we finally present, in Figure 10, certain summary statistics from the sunrise-sunset diagrams of the 42 branches that Torre studied. Here the first panel gives the distribution of average rates of TFP growth among the 42 industries. The second panel shows the distribution of peak cumulative contributions, i.e., what the TFP contribution would have been had all the negatives been zeros. The third panel shows, from branches with positive TFP growth, the percentile of firms which by themselves account for 100% of the industry's TFP growth. And finally, the fourth histogram displays the percentile of firms (by initial value added) marking the borderline between positive and negative TFP growth.

This evidence almost seems to replicate, for firms within an industry, what was found in the previous section for industries within the economy -- rampant overshooting of sunrise-sunset diagrams, great influence of firms with negative TFP growth in determining the TFP outcome for an industry, a small or moderate fraction of firms accounting for 100% of the TFP growth of an industry (when that growth is positive) with the complementary fraction being winners and losers whose efforts end up just offsetting each other. It remains to try to give some interpretation to those results.

FIGURE 7

FIGURE 8

Figure 8 (cont.)

Figure 8 (cont.)

FIGURE 9

FIGURE 10

Figure 10 (cont.)

"JUST ERRORS" OR "IT'S A JUNGLE OUT THERE"?

46. The first question that will enter the mind of many economists, on looking at the evidence presented so far, is: how much of what we have seen and emphasized might simply be the result of errors of observations? This is by no means a frivolous question. For one can actually create frequency distributions of rates of TFP increase which contain exactly the same information as the sunrise-sunset diagrams previously presented. The only trick is to count as the unit of frequency not one firm (out of an industry aggregate) or one industry (out of some larger aggregate), but instead, say, 1% of the total value added of the aggregate. Thus a firm with 20% of the value added of an industry would appear with 10 times the weight of a firm accounting for 2% of the value added of that industry. In such a chart, the cumulative frequency (say 68%) above TFP = 0 would represent the projection on the horizontal axis of the maximum point on a sunrise diagram. Its complement (32%) would represent the initial value added associated with negative TFP performance during the period.

If, then, all the information could be generated by a properly designed frequency distribution of rates of TFP growth, could it not all be the result of chance alone -- more specifically, of errors of measurements? I really think not -- my favorite quip on this is that "white noise does not sing a tune". That is, if we can rationalize what we see in terms of an analytical framework which embodies well-established economic principles and sensible presumptions about underlying relationships and facts, this is itself strong evidence against the white noise hypothesis.

47. Nonetheless, we have to face the fact that errors of observation of some magnitude certainly do exist, and must recognize that they can cloud our perceptions and bias our results. What I am going to do here is consider frequency distributions of firms. TFP is measured in two ways -- one using value added by a set of firms on the one hand and the other using "output" by those same firms, measured through dividing value added by separate estimated firm-by-firm price indexes p_j. For these purposes we can conveniently think in terms of logarithms, so let:

v_j = observed value added of firm

p_j = estimated firm-level price index

$y_y = v_j - p_j$ = estimated output

$v_j = T_j + e_j$ [T_j = true value added]

$p_j = \ _j + u_j$ [$_j$ = true price index]

$q_j = T_j - \ _j$ [true output of firm]

We would like to have data on $_j$ and its variance

:{ stack {2 # q dot} ~

stack {2 # T dot } ~ - If we simply work with observed value added as our quantity variable, we get

` _{ dot T dot} ~ + ~

k {2 # dot}}

:{ stack {2 # v dot} ~ = ~ stack {2 # T If we worked with the measured y_j, we get

} ~ + ~ stack {2 # e dot} ~~~~

uming ~~ _{T dot e dot}~ = ~ 0)} (assuming

:{ _y^2 ~ = ~ stack {2 # T dot} ~ to be strictly random)

stack {2 # e dot } ~ + stack {2

lot} ~ + ~ stack {2 # u dot } ~ - ~ My presumptions are as follows:

[T dot dot}}

i) We can estimate value added quite accurately at the firm level. Hence the presumption that

:{ u dot

and ~~ e is small.

}

func{ stack {2 # e dot} }

ii) In most industries, there is considerable variety among the firms and their products. Hence, except in cases of industries with very homogeneous products, we should not expect to be small. Hence, I expect

func{ stack {2 # dot } }

iii) Finally, we have the presumption that, at least at the level of firms within an industry, $_T < 0$. We know that firms choose to operate in regions of the demand curve where they consider the elasticity facing them to be greater than one. But also, in an analysis of the growth process, one would expect the big gains in value added to accrue to those firms in an industry which were passing along to consumers some of the fruits of current or past real cost reductions.

These three presumptions lead me to the conclusion that is likely to understate the true variance of output

:{ stack { dot} ~ > stack {2 # ot }. }

func{ stack {2 # v dot } }

func{ and that
stack is likely to overstate
{2 # q And since Torre worked with real value added as his quantity variable, this suggests that if anything the
dot },
} substitution of the "true quantity variable" q for observed value added v would have given results with
func{
stack greater dispersion of TFP, and consequently greater overshooting in his sunrise-sunset diagrams.
{2 #
y dot **48.** The above demonstration should be taken as merely suggestive. It is not important to me that
} }
func{ Torre's results underestimate the variability of the different firms' TFP experience. It is only important that
stack
{2 # q measurement error should not be the principal determinant of those results. On this I feel very confident. In
dot }.
} my view, it really is "a jungle out there", with winners and losers in every period -- good as well as bad.

As I have noted earlier, we are only just beginning a systematic study of TFP among U.S. firms, so I can offer no display comparable to Torre's.

However, Robles did examine the experience of 12 firms in the U.S. oil industry. His results are summarized in Figure 11. (Note that Robles's diagrams are Lorenz-type rather than sunrise-sunset.) But Robles tells basically the same story as Torre. Three firms out of the twelve were more than sufficient to generate the real cost reduction experienced by the total group. Half (or almost half) of the firms had negative TFP growth in each period. And the

FIGURE 11

cumulated amount of this negative TFP growth was sizeable when measured against the total TFP performance of the industry.

What I see in TFP performance is quite analogous to what I see in the stock market pages of the newspaper. There are winners and losers every day, every month, every year. The gains and losses come from all sorts of causes. World price shocks can drive firms into negative TFP performance if the consequent output reductions are greater than the reductions of inputs. So too can cyclical or secular declines in demand, including those caused by the successful actions of competitors.

When firms are under stress, they typically fight to stay alive. Maybe they fight for too long in some cases, in the sense that less of society's resources would be wasted if they were to quit earlier in response to a challenge that turns out to be deadly. But they don't recognize the challenge as deadly, so they keep struggling to survive. I believe this is part of the nature of entrepreneurs, CEOs, and business leaders in general. They wouldn't be where they are, doing what they are doing, if they were ready to quit at the first sign of a challenge. They're fighters by nature, and they probably wouldn't have achieved success if they weren't.

Firms with negative TFP growth may even be innovators. New challenges come and different firms think of different ways to respond to them. Some (like Intel and Microsoft) end up winners; others (Montgomery Ward and Apple?) end up losing. But if may not be that they just waited passively and tried to fight to survive in the face of negative shocks. They may have had quite innovative ideas, with decent prior probabilities of success, but in the end success didn't come. Thus, negative TFP performance can, and I believe often does come, simply from "backing the wrong horse".

49. To me, Schumpeter's vision of "creative destruction" captures much of the story. What he's saying is, yes, it's a jungle out there, but the processes of that jungle are at the core of the dynamics of a market-oriented economy. They are what got us to where we are, and they hold the best promise for further progress in the future.

In my opinion, Schumpeter saw through to the essence of the problem, but it is not wise for us to be fatalistic in accepting his vision. We cannot lose by making a major effort to understand the process of TFP improvement where it happens -- at the level of the firm. All the more so, because of the pervasiveness of negative as well as positive TFP performance among the components of almost any aggregate. By learning more about this aspect of the aggregate picture, we may stumble upon ways to "accentuate the positive, eliminate the negative" parts of the TFP story. But that is too quixotic a goal to take as our point of focus right now. To me, our present task is simply to get hold of the huge mass of information that is available at the firm level, and squeeze it hard enough to wring out as much understanding, as much insight as we can.

50. I see the 2-deflator approach as playing a major role in this process; it is the very basic simplicity of this method that opens the door to the vast array of information that is available at the firm level. And if we have qualms about applying this method, we should take heart in the fact that the parts add up to the whole. That is, firm output as defined by the 2-deflator method adds up to industry output, industry output to sector output, sector output to the GDP. So, too, for the separate inputs, and for the TFP contribution, which I like to call Real Cost Reduction. Of course, the growth residuals, R, obtained using the 2-deflator method will differ from those obtained using other methods. But we have seen, in Table 4 and in Figure 3, that the differences are quite modest in size, fully acceptable as the price we pay for a key that opens the door to literally tons of new information, even if we may accept that in some cases, as with Jorgenson's meticulous artwork, there is an alternative method that gives better results.

CONCLUDING OBSERVATIONS: METHODS AND RESEARCH

51. What I am about to say in this section is not meant to consist of direct implications of what has gone before. Instead, I think of the earlier parts of this paper as building a case for a certain vision of the economy, and of how the forces of growth work within it. This vision in turn leads one to think in different ways not only about the growth process itself but about how we as economists might best advance our study and understanding of it, and how policies might be molded so as better to promote it.

52. It is always wise to study the components of growth separately. The rate of investment, the rate of return on capital, the rate of growth of the labor force in numbers or in hours worked, the contribution of human capital or of the increment in average quality of labor, and the residual representing real cost reduction -- all these are sufficiently different, and potentially sufficiently disjoint from each other, to merit their being treated separately. If I have three things to point to here, they are:

i) the worthwhileness of measuring the rate of return and emphasizing its role in the growth process

ii) the importance of focusing on investment rather than saving in studying the process of growth. Saving is an interesting topic in its own right, but the more "open-economy" is the situation being studied, the less saving has to do with investment. Saving takes on importance in closed economy models focused on aggregate growth, in which case it is equal to investment.

iii) the importance of viewing the residual as an umbrella covering real cost reductions of all kinds, and of recognizing that we are closer to home thinking that RCR takes 1001 forms than that it can be well represented by one or two or three aggregate-style variables.

53. In principle, the accumulation of human capital by the labor force should be represented in the labor contribution of the growth equation, or in a bifurcation of this contribution into one due to raw labor, the other to human capital. It is in a term like $_{i}\, w_i L_i$ that one captures the shifting skill composition of the labor force. In particular, we capture here the higher wages that are the fruits of investment in education and training, which are the benefits that the workers themselves perceive. These should be kept separate from any externalities education might have.

It is important to try to keep this internalized part of the story out of the residual, so that we can straightforwardly interpret the residual as Real Cost Reduction.

54. To study externalities due to education, training, or human capital, we should not be content with broad generalizations such as "TFP growth is higher in entities with lots of human capital per worker." We should try to figure out how this externality works. Is it higher for firms with high incidence of human capital? For industries or sectors? Or are human capital externalities more spatial in nature, making more efficient the economic life of the cities, provinces, states or nations which have high concentrations of human capital? And if this is a fruitful trail to pursue, at what type and size of geographical units do these externalities typically work?

55. The same goes for economies of scale. We should not be satisfied with vague attributions of economies of scale, say, at the level of the national economy. Instead, we should pursue the matter. If the economies of scale are national, through what channels do they work, and what evidence do we have to look at to see them in operation? In particular, what is their connection to real cost reductions where they really happen -- i.e., at the level of the firm? Economies of scale at the levels of the firm and the industry are easier to visualize. Here, too, however, the task is to check them out -- to see if the real cost reductions of firms are linked to the initial sizes of those firms themselves, or of the industries in which they operate, and of the direction (up or down) in which output is moving.

56. Perhaps most important of all, we should really try to take full advantage of evidence at the firm level. I think particularly of identifying considerable numbers of outstanding cases of TFP improvements and TFP decline, and studying them one by one to try to ferret out the sources of their big real cost reductions and real cost increases. You can be pretty sure, if there have been big real cost reductions in a firm, some people in that firm have a pretty good idea of where those reductions came from, and how they were accomplished. By capturing this grass-roots evidence, we can put some added discipline into our ruminations about the nature of TFP at the aggregate level. In general our aggregate story should be compatible with, and comfortably contain, what we see at the grass-roots level. In particular our overall TFP improvement should comfortably accept the overwhelming evidence of the "mushrooms" rather than "yeast" nature of the process.

57. Special urgency applies to the study of declining total factor productivity at both the firm and the industry levels. The pervasiveness of declining TFP is perhaps the most profound conclusion to emerge from the empirical links that I have reported here. As a profession, we obviously have been aware of its existence at the industry level, for virtually all studies that give a breakdown by industry reveal it. Yet to my knowledge, we have barely scratched the surface in studying it. I find it hard to think of more fertile soil for future research on the process of economic growth.

CONCLUDING OBSERVATIONS: POLICY IMPLICATIONS

58. The springboard for this listing of policy implications is the interpretation of the growth residual as representing real cost reduction and the ready acceptance that in the real world RCR comes in 1001 different forms. The first key observation is that people must perceive real costs in order to reduce them. Hence policies that impede the accurate perception of real costs are ipso facto inimical to growth. Inflation is the most obvious, probably the most pervasive, and almost certainly the most noxious of such policies. If

I have any expertise based on experience in economics it has to be in the first-hand observation of processes of serious inflation. So I ask you to take my word for it: the most serious cost of inflation is not a triangle or a trapezoid under the demand curve for real cash balances; it is not the inflation tax. The most serious cost of inflation is the blurring of economic agents' perceptions of relative prices. This happens because individual prices adjust in different ways and at very different rates. A high product price and a low input cost normally is an invitation clamoring for new investments to be made. Not so during a serious inflation, when such a signal can easily turn out to be "here today, gone next month" as both product and input prices continue on their separate paths of adjustment to the ongoing inflation. Without exception, in my own observations, the higher the rate of inflation, the worse is its effect in blurring agents' perceptions of relative prices. In an inflation at, say, 20% to 50% per year, people see prices as in a morning haze; in one of 20% to 50% per month, they see them as in a London fog. Many empirical studies exist showing that serious inflations are seriously inimical to growth. The clouding of perceptions of relative prices is an important reason why -- for it gets in the way of successful real cost reductions at the level of the individual firm.

59. A second policy implication is, in the words of my friend and longtime collaborator Ernesto Fontaine, avoid "prices that lie" (precios mentirosos). Talking about inflation, we focus on the blurring of the signals that the price system gives; here we focus on its giving wrong signals due to distortions that have been introduced, usually as a direct consequence of government policies. No good can or did come, in terms of economic efficiency, from tariffs of 50% and 100% and more, giving effective protection often of 200% and 300% and more. Nor can growth be fostered by heavy-handed price controls and interventions in credit markets.

I am not being a religious purist here -- just as big distortions have big costs, small distortions typically have small costs, and all economies are distorted to some degree. The message here is that economies have to pay the price for the level of distortions they choose to have, and that one of the important components of that price is that distortions create situations where what is truly a saving of private costs is not a genuine saving of costs from the point of view of the economy as a whole.

60. Just as bad, and even sometimes worse than direct distortions can be and often are the excess costs imposed on an economy by ill-conceived regulations and bureaucratic hurdles. Hernando DeSoto has made the exposure of these trammels in Peru into what has become virtually his life's work. Clear rules of the game are an essential and integral part of a well-functioning market economy, but all too easily these get supplemented by others that make investment, production, marketing, sales, new product development, etc., etc., more costly. Labor laws have been particularly troublesome, often adding artificially to the cost of labor and often giving firms a strong incentive to avoid hiring new workers, simply because of the high costs associated with any later dismissal of them.

But there are loads of other items -- the need for approvals, sometimes a dozen or more, before undertaking some investment or some new venture; regulations that one way or another impede new entry, so as to protect strong vested interests (small retailers being protected against supermarkets in many countries), the complexity of tax codes and their enforcement, which imposes large compliance costs on business firms and individuals. Somehow, countries interested in promoting growth should find ways of paring their regulatory framework down to those rules and requirements that are really justifiable in terms of their costs and benefits to the economy and society at large.

61. The recent wave of privatizations among both developed and developing economies may have important effects in enabling real cost reductions that otherwise might have been delayed, or not have happened at all. It is, I believe, fair to say that in most countries state-owned enterprises operate under a series of constraints that seriously get in the way of real cost minimization in a comparative static sense and of real cost reduction in a dynamic sense. These constraints sometimes limit the salaries of executives, sometimes impose onerous conditions on the firm as it employs lower-skilled workers, often limit the capacity of the firm to shut down inefficient lines of production, almost always make it difficult to fire workers, etc., etc. To my mind, however, perhaps the worst attribute of state-owned enterprises is the ethos that often evolves inside of them -- an ethos where middle managers are well advised to "leave well enough alone", "not rock the boat" and "not invite trouble". This ethos flies in the face of a vision of the growth process that gives a huge role to the search for real cost reductions at the grass-roots level, and that recognizes the tumult that accompanies "creative destruction" in all its forms. I thus must applaud the contemporary trend toward privatization. If I harbor any qualms in this connection, they concern the degree to which many privatizations have been motivated by purely fiscal considerations rather than by a general search for economic efficiency. This may have led to gratuitous transfers of wealth in some instances, and to the planting of newly private enterprises in soil that was not properly prepared (e.g., still lacking a sound regulatory framework for electricity rates, or intelligent rules promoting competitiveness in at least some aspects of telecommunications, etc.)

62. One cannot complete a list like this without mentioning something that most of us simply take for granted -- a sound legal and institutional framework in which individuals are protected against arbitrary incursions on their property and other economic rights. This very basic point -- recently much emphasized by Douglass North (1991) and Robert Barro (197) -- is at least potentially a vital element for a sustained process of successful economic growth. If it is true that spurts of growth have sometimes occurred in the absence of such a framework, it is also true that just about every case of sustained growth over long periods of time has benefitted from a sound institutional and legal environment.

63. Somewhat related to the above is the element of political consensus concerning the broad outlines of economic policy. We have learned from experience that very admirable policy reforms can take place, yet end up having little effect. This can happen because a new government comes in and reverses the reform. But it can also happen because people fear that a new government will come in and reverse the reforms later on. At the moment, the Chilean economy is one of the jewels of economic growth (and general economic success) in Latin America. Many people point to the thoroughness and pervasiveness of Chile's economic reforms over the last two decades or so. But not so many point to the fact that the reform package has remained essentially intact through several changes of ministers, and even more important, through two presidential elections in which the winners came from the opposite side of the political fence from the government that initiated the reforms. The confidence in the economic order of things instilled by this sequential endorsement of the basic framework of economic policy has to be one of the important reasons for Chile's continued, very impressive economic performance. And it is important, also, in the context of this paper. Living in a world in which real cost reductions are a key dynamic force producing economic growth, we must look to the motivations and preoccupations of those who take the critical decisions at the level of the firm. For these decisions, it is not only important that the policy framework be good now; the expectation that it will stay good in the future is also important. Otherwise investments will tend to be limited to those with short horizons and payment periods, and much soil, fertile with longer-term economic opportunities, will go unplowed.

64. As my final point, let me make explicit what has been implicit in many of the thoughts expressed in this paper -- that the justification for perfecting the functioning of the market system does not lie only in reducing the efficiency costs associated with each period's operation of the economy. Perfecting a country's economic policy does not only cause it to move from a path at around, say, 90% of its potential output to another equal to 95% of potential, with the time path of potential output being somehow given in advance. That gain would certainly be a worthwhile gain, and it would amply justify a lot of hard work involved in achieving it. But that gain is still fundamentally comparative-static in nature.

What I hope to have evoked in this paper is a sense that the perfecting of economic processes can also in nearly all cases be justified as greasing the wheels of the constant search for new avenues of real cost reduction. To the extent that economic reforms do so, they become vehicles for bringing an economy to a point where, year after year, new, cheaper and better ways are found of doing things, not just in so-called "production" but also in such mundane areas as merchandising, sales, finance, insurance and many more.

Some years ago, in a book that I edited called World Economic Growth, I wrote an essay called "Economic Policy and Economic Growth," in which I listed "thirteen lessons" that I thought followed from the papers presented in the volume, recounting the growth experiences of countries as disparate as Ghana and Taiwan, or Japan and Sweden. These lessons -- basically focused on thinking about policies in terms of their economic costs and benefits -- could easily be read as a reprise of the old comparative static story. But they were not meant as such -- it is quite relevant that they appeared in an essay concerned with world economic growth. The point was that these sensible policies emerged as part of the consensus of serious economists, each an expert in his particular country's history, focusing attention on the process of economic growth.

A few years later as the theme of a different concept, John Williamson coined the term, the "Washington consensus". Williamson listed 10 points, covering territory very similar to my thirteen lessons. He also produced a pithy summary that captures the essential thrust of both his and my listings: "Macroeconomic prudence, outward orientation, and domestic liberalization". He, too, and the members of the Washington professional establishment whose apparent consensus led to Williamson's list, were not just thinking of comparative static gains as they reached their conclusions about policy. They were thinking about ways to move economies from slow growth, stagnation, even in some cases negative growth, to a healthy, prosperous flowering of economic progress.

To me, the dynamics of real cost reduction are at least an important piece of what people have in mind when they list efficiency-oriented policies as essential ingredients of a program promoting economic growth. It is policies of this type that give the right signals to the CEOs and the managers down the line, that take away trammels that impede their quest for real cost reductions, and that create an environment in which Schumpeter's process of "creative destruction" can work its wonders.

REFERENCES

Abramovitz, Moses (1956). "Resources and Output Trends in the United States Since 1870." American Economic Review, 46, 2 (May), 5-23.

Aghion, Philippe and Peter Howitt (1992). "A Model of Growth Through Creative Destruction." Econometrica, 60, 2 (March), 323-51.

Barro, Robert J., and Xavier Sala-i-Martin (1994). Economic Growth, New York: McGraw-Hill.

Chenery, H., M. Syrquin, and S. Robinson (1986). Industrialization and Growth, Oxford: Oxford University Press.

DeSoto, H. (19??). The Other Path

Elias, V.J. (reference to come)

Griliches, Z. (1994). "Productivity, R&D, and the Data Constraint." American Economic Review, 84, 1 (March), 1-23.

Grossman, Gene M., and Elhanan Helpman (1991). Innovation and Growth in the Global Economy, Cambridge, MA and London: MIT Press.

______ and ______ (1994). "Endogenous Innovation in the Theory of Growth." Journal of Economic Perspectives, 8, 1 (Winter), 23-44.

Harberger, A. (ed.) (1984). "Economic Policy and Economic Growth." World Economic Growth, San Francisco, CA: ICS Press, 427-66.

______ (1990). "Reflections On the Growth Process" World Bank Working Paper.

Hulten, C. and F. Wykoff (1981). "The Measurement of Economic Depreciation," in C. Hulten, ed., Depreciation, Inflation and the Taxation of Income From Capital, Washington, DC: The Urban Institute Press, 81-125.

Jorgenson, D. (1995). Productivity, Vols. 1 and 2, Cambridge, MA: MIT Press.

Jorgenson, D. and Z. Griliches (1967). "The Explanation of Productivity Change." Review of Economic Studies, 34, 3 (July), 249-80.

Jorgenson, D., F. Gollop and B. Fraumeni (1987). Productivity and U.S. Economic Growth, Cambridge, MA: Harvard University Press.

Kendrick, J. (1961). Productivity Trends in the United States, Princeton, NJ: Princeton University Press.

Lucas, R.E., Jr. (1988). "On the Mechanics of Economic Development." Journal of Monetary Economics, 22, 1 (July), 2-42.

______ (1991). "On Making an Economic Miracle." Econometrica (reference to come).

North, D. (reference to come)

Robles, E. (1997). An Exploration into the Sources and Causes of Economic Growth in the United States and Fourteen Latin American Countries, UCLA unpublished doctoral dissertation.

Romer, P. (1986). "Increasing Returns and Long-Run Growth." Journal of Political Economy, 94, 5 (October), 1002-37.

______ (1990). "Endogenous Technological Change." Journal of Political Economy, 98, 5 (October), S71-S102.

______ (1994). "The Origins of Endogenous Growth." Journal of Economic Perspectives, 8, 1 (Winter), 3-22.

______ (1994). "New Goods, Old Theory, and the Welfare Costs of Trade Restrictions." Journal of Development Economics, 43, 5-38.

Solow, R. (1957). "Technical Change and the Aggregate Production Function." Review of Economics and Statistics, 39, 3 (August), 312-20.

______ (1994). "Perspectives on Growth Theory." Journal of Economic Perspectives, 8, 1 (Winter), 45-54.

Torre, L. (1997). "Concentration Patterns of the Contribution of TFP to Output Growth: Evidence From the Mexican Manufacturing Sector," paper presented at the meeting of the Western Economic Association, and research in progress.

Welch, F. and K. Murphy (1997) (reference to come)

Williamson, J. (ed.) (1990). Latin American Adjustment, Washington, DC: Institute for International Economics, 2.

Young, A. (1995). "The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience." Quarterly Journal of Economics, 110, 4 (August), 641-80.