THE MEASUREMENT OF TECHNOLOGICAL CHANGE

By
Richard G. Lipsey, Simon Fraser University
and
Kenneth I. Carlaw, University of Canterbury

Contacts
Richard Lipsey,
Emeritus Professor of Economics at Simon Fraser University
e-mail: rlipsey@sfu.ca,
home page: http://www.sfu.ca/~rlipsey

Kenneth Carlaw
Lecturer, University of Canterbury, New Zealand
e-mail: k.carlaw@econ.canterbury.ac.nz
ABSTRACT

We consider how to measure technological change, looking first at the residual approach that defines TFP. Here we first use quotations to show that TFP is interpreted in several mutually contradictory ways by well known writers on the subject. We then argue that, contrary to widely held views, TFP does not measure technical change. In so far as it measures anything, it measures the super normal gains that are associated with growth-creating technical change — gains that are similar, although not identical, to Jorgensen and Griliches’ concept of free lunches. In the next part, we show several reasons why TFP is an imperfect measure of these gains. First, a given technological change causes very different changes in calculated TFP depending on the timing of innovation. Second, the treatment of R&D in the national accounts causes some technological dynamism to escape TFP measurements. Third, the omission of certain inputs, particularly natural resources, can cause significant amounts of technological change that increase the efficiency of the use of these resources to be recorded as increases in other inputs rather than in TFP. Fourth, aggregation to obtain a measure of inputs to insert into the production function used to measure TFP tends to cause the effects of much technological change to be recorded as increases in the quantity of factors. Fifth, a similar result is associated with the expenditure weights used to calculate the TFP index when markets are not in full equilibrium. In the section that follows, we discuss the type of counter factual measures that we argue are needed to properly measure technical change. Finally, we illustrate the contentions made earlier in the paper by constructing a simple model of sustained endogenous growth in which there are neither externalities nor increasing returns and TFP never changes.

Key words: Total factor productivity, technological change, aggregation, economic growth, free lunches.
# TABLE OF CONTENTS

THE MEASUREMENT OF TECHNOLOGICAL CHANGE ................................................. 1

I. THE RESIDUAL APPROACH ..................................................................................... 2

I.2. THE BASICS OF TOTAL FACTOR PRODUCTIVITY ............................................ 4

I.2.1 The Production Function ................................................................................. 4
I.2.2 TFP Defined ...................................................................................................... 5
I.2.3 Recognized Problems With TFP ...................................................................... 6

I.3. TFP AND COSTLY TECHNOLOGICAL CHANGE .............................................. 6

I.3.1 Physical Capital ............................................................................................... 7
I.3.2 Human Capital .................................................................................................. 8
I.3.3 Free lunches and super normal benefits ....................................................... 9

I.4. HOW WELL DOES TFP MEASURE SUPER NORMAL BENEFITS ..................... 10

I.4.1 Timing of Cost Reductions ............................................................................. 10
I.4.2 The treatment of R&D .................................................................................. 11
I.4.3 Induced Investment ......................................................................................... 12
I.4.4 Omitted inputs Natural Resources Made Explicit ......................................... 12
I.4.5 Aggregation of Inputs .................................................................................... 13
I.4.6 Aggregation when calculating a TFP Index .................................................. 15

II. COUNTER FACTUAL MEASURES ................................................................... 19

II.1. The conceptualization of Technological Change ............................................ 19
II. 2 GPT-Derived Growth ..................................................................................... 20
II.3 The Measurement of Technological Complementarities ................................. 21
II.4 A Model of GPT-driven Growth with constant TFP ....................................... 22

III. CONCLUSIONS ................................................................................................. 24

APPENDIX .............................................................................................................. 26

BIBLIOGRAPHY ....................................................................................................... 30

FIGURE 1 ................................................................................................................ 33
THE MEASUREMENT OF TECHNOLOGICAL CHANGE

Economic historians and students of technology are agreed that technological change is the major determinant of very long-term economic growth. If we knew no more than the Mesopotamians, the Ancient Greeks, or Medieval Europeans, our living standards would not be far above theirs, a bit more, due to capital accumulation and other matters. But we are orders of magnitude better off materially than they were because of the advances in technological knowledge concerning the products we make, how we make them and how we organize our value-creating activities. Yet, over shorter periods of time, there is debate over how much of economic growth is due to technological change and how much is due to other forces such as the accumulation of physical and human capital.

In this, the first of three related papers, we consider how to measure technological change. We define technological knowledge as the idea set that specifies all activities that create economic value. It comprises knowledge about product technologies, the specifications of everything that is produced, process technologies, the specifications of all processes by which goods and services are produced, and organizational technologies, the specification of how productive activity is organized. All these are often referred to as just technology and we will follow that practice whenever there is no ambiguity.

We first consider the residual approach for measuring changes in technology. This is by far the most common measure. In the first part of this paper, we argue that, contrary to widely held views, total (or multi-) factor productivity (TFP) does not measure technical change. In so far as it measures anything, it measures the super normal gains that may be associated with growth-creating technical change. In the second part, we argue that TFP is an imperfect measure of these gains. Next, we discuss the counterfactual measures of technical change that we argue are needed. Finally, we illustrate the contentions made earlier in the paper with a simple model of sustained endogenous growth in a world with neither externalities nor increasing returns and zero TFP change. In the second paper, Carlaw and Lipsey (2002), we argue that the gains measured by TFP are only a small subset of the growth-inducing spillovers associated with technological change. In the final paper, Carlaw and Lipsey (2002, manuscript), we illustrate all of our earlier arguments by developing a much more detailed version of the model of endogenous growth briefly outlined at the end of this paper.

Our main purpose is to direct attention to the development of measurements of technological change other than TFP, measures that might be important for understanding what drives long run economic growth. Nonetheless, we pay much attention to TFP, not only because it is often assumed to measure technological change, but also because there is a debate within policy circles about what TFP

---

1 Earlier versions of this paper were presented at three successive workshops held by Statistics Canada and one by the Canadian Institute for Advanced Research. We are grateful for the many comments and suggestions provided by the members of these workshops and in particular to Alvero Pereira, John Baldwin, Erwin Diewert, Mel Fuss, Alice Nakamura, Andrew Sharpe, Manuel Trajtenberg, and Thomas Wilson. Some of the conclusions from earlier versions of the paper were summarized in International Productivity Monitor, Inaugural Issue, Fall 2000. Financial support for part of this research was provided by the Royal Society of New Zealand’s Marsden Fund (grant number UOC101).

2 We have argued this point in several publications, e.g., Lipsey, (1992, 1993 and 1994). Of course, technological change and investment are interrelated, the latter being the main vehicle by which much of the former enters the production process. It follows that the long-term rate of growth will be slowed by anything that slows either the development of new technologies or the rate at which new technologies are embodied through investment. Thus, the high correlation between new investment and growth does not demonstrate that investment is the main cause of long-term growth. Both technological change and investment are necessary.

3 We have given our reasons for adopting this definition in Lipsey, Bekar and Carlaw (1998b)
measurements imply about the appropriateness of policies designed to induce economic growth. Many of these debates focus on TFP as a measure of something and as a lever for policy makers to use in the process of encouraging economic growth. One of the major problems with this view is identified by Hulten (2000). “The various factors of TFP are not measured directly, but lumped together as a residual “left over” factor (hence the name). They cannot be sorted out within the pure TFP framework, and this is the famous epithet “measure of our ignorance.”” (61) A second problem with this view is that TFP, whatever it measures, is an outcome not a cause of anything and, therefore, not a lever for policy makers to use to affect growth rates. It may be an indicator but it is certainly not a policy instrument.

Another important reason for concentrating on TFP at the outset is that many growth economists whose insights have been honed on the aggregate production function in which technological change must raise TFP are sceptical of the existence of a New Economy. (See, for example, Gordon 2000.) In spite of a number of articles pointing out that productivity change is not an index of technological change, most commentators have argued one of two positions. The first extreme position is that there has been no New Economy because there has been no major acceleration in TFP change. The less extreme, but in our opinion no less misguided, position is that there is a paradox in the apparent evidence of a technological transformation and the lack of major TFP gains. In contrast, we have argued in a number of publications that there is no necessary relation between the rate of technological transformation of the economy on the one hand and the rate of productivity growth on the other hand. (See Lipsey and Beker (1995) for an early statement of this argument and Lipsey (2002) for a recent one.)

I. THE RESIDUAL APPROACH

The most commonly used measure of technological change is the change in total (or multi-) factor productivity, TFP. This is a residual approach that originates in Solow's seminal 1956 and 1957 articles. A production function is used to relate measured inputs to measured output. Any growth in output that cannot be associated statistically with the growth in measured inputs is assumed to be the result of technological change (and some other causes such as scale economies). The many different index numbers that can be used in his type of measurement are surveyed by Deiwert and Nakamura (2002).

The following quotations, which we list in descending order of the scope that they give to TFP, illustrate some of the different interpretations of TFP that are current in the literature.

(1) “A growth-accounting exercise [conducted by Alwyn Young] produces the startling result that Singapore showed no technical progress at all.” Krugman (1996, p.55) “Singapore will only be able to sustain further growth by reorienting its policies from factor accumulation toward the considerably more subtle issue of technological change.” Young (1992, p.50)

(2) “Improvements in technology – the invention of the internal combustion engine, the introduction of electricity, of semiconductors – clearly increase total factor productivity.” Law (2000, pp.6-7)

(3) “Technological progress or the growth of total factor productivity is estimated as a residual from the production function…. Total factor productivity is thus the best expression of the efficiency of economic production and the prospects for longer term increases in output.” Statistics Canada, (1998, pp. 50-51, italics added)

(4) “Growth accounting provides a breakdown of observed economic growth into components associated with changes in factor inputs and a residual that reflects technological progress and other elements.” Barro (1999, p. 119)

(5) “The defining characteristic of [total factor] productivity as a source of economic growth is that the incomes generated by higher productivity are external to the economic
activities that generate growth. These benefits “spill over” to income recipients not involved in these activities, severing the connection between the creation of growth and the incomes that result.” Jorgenson (1995, p. xvii.) “That part of any alteration in the pattern of productive activity that is ‘costless’ from the point of view of market transactions is attributed to change in total factor productivity.” Jorgenson and Griliches (1967), reprinted in Jorgenson (1995, p.54)

(6) The residual should not be equated with technical change, although it often is. To the extent that productivity is affected by innovation, it is the costless part of technical change that it captures. This “Manna from Heaven” may reflect spillover externalities thrown of by research projects, or it may simply reflect inspiration and ingenuity. Hulten (2000, p. 61)

(7) “The central organising concept…[is] the division of observed growth in output per worker into two independent and additive elements: capital-labour substitution, reflected in movements around the production function; and increased efficiencies of resource use, as reflected by shifts in this function. To maintain additivity, …the analysis…could not be applied cumulatively without introducing an interaction term between capital substitution and increased efficiency…. [The residual debate] never did attempt to answer the question, of what is the residual composed. This remains the dominant question.” Metcalfe (1987, p. 619-20)

(8) “Is there something possibly wrong with the way we ask the productivity question, with the analytical framework into which we force the available data? I think so. I would focus on the treatment of disequilibria and the measurement of knowledge and other externalities.” Griliches (1994)

(9) “All of the pioneers of this subject were quite clear about the tenuousness of such calculations and that it may be misleading to identify the results as ‘pure’ measures of technical progress. Abramovitz labelled the resulting index ‘a measure of our ignorance’.” Griliches (1995, pp.5-6) quoting Abramovitz (1956, p.8)

These quotations illustrate three main positions.

• One group holds the view that changes in TFP measure the rate of technical change—Krugman, Young, Law, Statistics Canada, Barro.
• The second group holds that TFP measures only the “free lunches” of technical change, which they argue are mainly associated with externalities and scale effects—Jorgenson and Griliches (hereafter J&G) in (5) and Hulten.
• The third group is sceptical that TFP measures anything useful—Metcalf, Abramovitz and Griliches in (8&9)

Surely it is something close to a scandal that a measurement that is relied on for so many purposes in theory, applied work, and policy assessment seems to be so variously interpreted.

Although our position is close to the “free lunch view,” we argue that there is some important ambiguity surrounding that concept. Also, we do not accept that TFP should be close to zero, as Jorgensen of Griliches argued in their classic 1967 article. We argue later in this paper that technological change has associated with it many gains that should yield positive TFP measures.

---

4 This notion is similar to that of Harberger (1998). His notion of “real cost reduction” is a catch-all much like “free lunch” and not narrowly interpreted as externalities.

5 Although we have only three representatives of this view in our quotations, it has many other well-known members, including the Cambridge (England) economists who for several decades, starting in the 1950’s, attacked the validity of the concept of an aggregate production function.
We do not dispute Prescott’s (2000) call for a theory of TFP. He implicitly agrees with our position by arguing that the sources of international differences in TFPs are more than just differences in employed technologies. We would, however, add substantially to his list of other sources.6

I.2. THE BASICS OF TOTAL FACTOR PRODUCTIVITY

I.2.1 The Production Function

Calculations of TFP require the concept of a production function that is valid at whatever level of aggregation the calculations are to be made.7 This poses two quite different sets of conceptual problems.

First, if we are to measure technological change over long periods of time, there must be a stable production function such that changes in output can be related to changes in factor inputs and changes in productivity (plus a number of other factors such as and scale effects that are often ignored in practice). Let us write

(I.1) \( Y = AF(K,H,L) \)

where \( Y \) is output, \( K \) is physical capital, \( H \) is human capital, and \( L \) labour. To calculate TFP over long periods of time, we must assume that this function remains stable with productivity-increasing changes in technology being registered solely by changes in the productivity factor, \( A \). It is an heroic assumption that such a function remains stable over changes in general purpose technologies such as the replacement of steam by electricity as the power source for factories and the associated redesign of the factory layout.

We must also assume that we can meaningfully measure the inputs of factors over these long periods. For example, we must be able to measure capital inputs over periods in which the technologies in use are embodied in widely varying types of physical goods. In what units, for example, do we compare the amount of capital invested in a Victorian, steam-driven, manually controlled factory making stage couches with that in an electrically powered, largely robot-controlled, modern factory making diesel electric passenger trains? (Jorgenson, Gollop and Fraumeni (1987) attempt to control for some of these problems by specifically accounting for changes in such things as input quality, legal forms of organizations, capital asset classes and sectoral substitution.)

A second set of problems concern the aggregation from the production functions for individual products to the function that is actually used. This is possible in standard neoclassical theory that treats

6 Prescott suggests that international differences are explained by “...the strength of the resistance to the adoption of new technologies and to the efficient use of currently operating technologies.” While this is undoubtedly one important reason, we believe that the evidence from such countries as the Asian NICs suggests other equally compelling reasons. We have space here for just one example. Learning how to adopt and adapt technologies in operation in other countries with different levels of economic sophistication is no easy matter. For a relatively backward country to acquire the necessary tacit knowledge to adopt and adapt production functions from more advanced countries is difficult—especially when “production” includes product design, shop floor operation, firm management, financial arrangements, distribution networks, marketing, and after sales service. There are many risks involved in making such jumps in poorly developed countries and rational agents may judge the risks to be too great to justify the attempt. In the NICs, government incentives were important in overcoming this rational reluctance. For example, the policy of export promotion forced producers to learn about standards in the global marketplace for design, quality, sales and after-sales servicing. Studies such as Pack and Westfall (1986) Westphal (1990) and Wade (1990) give important evidence related to the many other reasons why TFPs differ among countries.

7 For example, Alwyn Young (1995, p.8) uses five sub-categories of capital, residential buildings, non-residential buildings, other durable structures, transport equipment, and machinery and many categories of labour distinguished by sex, age and education, all of which are inputs into a single aggregate production function.
competition among firms as the *end state* that is perfectly competitive equilibrium. Production functions at any higher level of aggregation can be formally aggregated from firm production functions given a perfectly competitive world of single product firms that are in equilibrium. (If there are multi product firms the basic function is each single product.) In contrast, the Austrian tradition sees competition as a *process* that takes place in real time.\(^8\) Industry or nation wide production functions cannot, however, be aggregated formally from a set of producers that are in process competition, even if all agents are price takers. Neither could it be aggregated from a set of markets that contain the mixture of monopoly, oligopoly, monopolistic and perfect competition that characterizes real-world industrial structures, even if all firms were in end state equilibrium. The judgment of economists varies greatly as to how much the absence of end state competition and the presence of non-perfectly competitive market structures matters. To make contact with the existing literature, we will proceed as if the aggregate production function is a meaningful concept, stable enough to measure productivity changes by some residual method over the time period in which we are interested.

### I.2.2 TFP Defined

Consider the simple Cobb-Douglas version of the aggregate function
\[ Y = AL^\alpha K^\beta, \quad \alpha + \beta = 1, \quad \alpha \geq 0, \quad \beta \geq 0 \]
where the variables are as defined above. Changes in \( A \) indicate shifts in the relation between measured aggregate inputs and outputs and we assume that these changes are caused by changes in technology (or changes in efficiency and/or in the scale of operations of firms).

In theory, these inputs should be measured as flows of current services. In practice, the stock of available inputs is often used on the assumption that, over the long term, variations in capacity utilisation can be ignored. Formally, what is required is that there be a constant proportional relation between the stock and the flow such as would happen if the level and intensity of utilisation of each stock were unchanged. We make this assumption in what follows so that we can move between using stocks of capital and flows of capital services.

The geometric index version of TFP is calculated by dividing both sides of the production function by \( L^\alpha K^\beta \) to obtain
\[ TFP = Y/(L^\alpha K^\beta) = A. \]

The growth rate measure of TFP can then be calculated as an arithmetic index generated by taking time derivatives of both sides of the above TFP expression (where the dot superscript denotes the time derivative):

---

\(^8\) "...firms jostle for advantage by price and non-price competition, undercutting and outbidding rivals in the market-place by advertising outlays and promotional expenses, launching new differentiated products, new technical processes, new methods of marketing and new organisational forms, and even new reward structures for their employees, all for the sake of head-start profits that they know will soon be eroded. ...[in short] competition is an active process." Blaug (1997, p.255-6).

\(^9\) Many recent TFP calculations use the more flexible translog function. But we lose little at the conceptual level by using the simpler Cobb-Douglas function. Also Jorgenson (2001) uses a production possibility frontier approach rather than an aggregate production function approach, which allows for the explicit use of constant quality prices of information technology products in his TFP calculation. This resolves an important measurement problem and avoids the need to use the concept of embodiment to convert investment goods into output by explicitly using two prices. But again we loose little at our conceptual level by using the Cobb-Douglas formulation. The production possibilities approach is implicitly used in section I.3.2 and its appendix when we have more than one sector. But there we do not need to worry about output versus inputs prices as everything is measured in real output and input units and converted back and forth by the production functions.
This equation defines total factor productivity as the difference between the proportional change in output minus the proportional change in a Divisia index of inputs.\(^\text{10}\)

The aggregate production function approach provided here is a simplification of a much broader concept of aggregate production. Jorgenson and Griliches (1967) first introduced the notion that all lines of activity can be properly aggregated and measured, including the resource-consuming activities of research and development. Jorgenson (2001) introduces a production possibilities approach to measure TFP that explicitly incorporates various lines of production activity. Barro (1999) introduces aggregate production functions that also include R&D as in intermediate activity in production. Two critical features of these approaches are their treatment of R&D as an input and the returns to scale in the production functions (or overall production possibilities set). J&G (1967) and Jorgenson (2001) treat all lines of production activities as having constant returns to scale, which implies that the part of technological change that involves costly R&D is not measured by TFP. In contrast, Barro (1999) uses production functions that allow R&D to generate expanding product variety or expanding product quality that have increasing returns to the intermediate R&D inputs to production. In Barro’s case, TFP measures the exogenous (Hicks neutral, “Mana from heaven”) component of technological change and the endogenous technological change generated from costly R&D. However, his endogenous component is also of the free lunch variety since it results from increasing returns to all lines of production activities. All of this leaves open the questions about the meta or all encompassing notion of aggregate production and about the appropriate formulation of R&D and knowledge production in that framework. We return to the issue of R&D and TFP later in the paper.

I.2.3 Recognized Problems With TFP

Many economists have identified problems associated with TFP both as a concept and with its measurement. Key references for measurement problems are Griliches (1987, 1994 and 1995), where he considers the many sources of error in TFP measurements.\(^\text{11}\) We do not review most of these issues because the problems associated with them are well understood. Instead, we focus our attention on a handful of specific issues that underlie the mutually inconsistent interpretations of TFP made by the groups identified above. Although some of these have been discussed in the literature much confusion remains.

I.3. TFP AND COSTLY TECHNOLOGICAL CHANGE.

In this section, we argue that TFP does not measure technological change but does, ideally at least, measure the super-normal profits, externalities and other “free lunches” associated with such

\(^\text{10}\) Most work on TFP uses a Tornquist index, which is basically a discrete version of the continuous Divisia index. It is a percentage change index that averages base and given years weighted indexes, as does the Fisher Ideal index. For our purposes, we use the continuous Divisia index, which weights percentage changes in specific inputs by their share of total cost. None of the conclusions we reach would be seriously affected by the substitution of one index for the other, except where we specifically wish to measure empirically TFP using real world discrete data.

\(^\text{11}\) Griliches (1987, p.1010-13) outlines some conceptual and empirical problems concerning the measurement of TFP. These relate to the following issues: (1) a relevant concept of capital, (2) measurement of output, (3) measurement of inputs, (4) the place of R&D and public infrastructure, (5) missing or inappropiate data, (6) weights for indices, (7) theoretical specifications of relations between inputs, technology and aggregate production functions, (8) aggregation over heterogeneity. Concerning point (6), Diewert (1987, p.767-780) shows that very restrictive assumptions have to be satisfied to generate these indices of output and input.
Virtually all technological change is embodied in one form or another: new or improved products, new or improved capital goods or other forms of production technologies, and new forms of organization in finance, management or on the shop floor. Although much innovation is in product technology, we concentrate here on process and organizational technologies. For concreteness, we speak in this section of capital goods although any embodied technology would do.

Although much theory proceeds as if these changes appear spontaneously, most of them are the result of resource-using activities. The costs involved in creating these technological changes are much more than just conventional R&D costs. In addition, they include costs of installation, acquisition of tacit knowledge about the manufacture and operation of the new equipment, learning by doing in making the product, and learning by using it, plus a normal return on the investment of funds in development costs. We refer to the sum of these as “development costs.”

It was the great contribution of J&G to point out that TFP would only measure the gains in output that were over and above their development costs. Because they argued that these gains would, when properly measured, be close to zero, the subsequent debate centred on whether the measure should be zero, obscuring their really important contribution that TFP did not measure the full contribution of the new technology.

1.3.1 Physical capital

Firms that do not make mistaken investments in developing new technologies must recover all of their development costs in the selling price of the new capital good. This means that the price of the good, and thus the investment that the users must make in buying it, will capitalize all development costs. Let us say, for example, that an existing machine is improved so that it does more work on the same job than did its predecessor. Let the value of the fully perfected new machine’s marginal product in the user industry be \( v \). This is the maximum price that users will be willing to pay for each new machine. Let the price required to allow the producers just to recoup their full development costs be \( w \) and consider three cases. (1) If \( w > v \), foresighted producers will not develop the machine and if they do \( TFP < 0 \). (2) If \( w = v \), costs are just covered, the rise in the cost of the machine just equals the value of the new output, \( TFP = 0 \). (3) If \( w < v \), profits are made and \( TFP > 0 \).

In case (3), there is a return over and above what is needed to recover the development costs that created the innovation. This will be shared between the capital goods producers and the users in a proportion that will depend on the type of market in which the good is sold. In all three cases, we have technological change. This is the sense in which changes in TFP do not measure technical change _per se_ but only the profits that it produces (as well as free lunch externalities). Thus, zero...
change in TFP does not mean zero technical change. It only means that investing in R&D has had the same marginal effect on income as investing in existing technologies (investment with no technical change) and that there are no external effects that show up in increased output elsewhere without corresponding increases in inputs. In the Appendix, we consider such changes in more detail allowing for capital enhancing, labour enhancing and neutral technological change.

If the marginal productivities of investing in new and existing technologies are the same, the new technology might seem to confer no benefit. Later in this paper, and in more detail in Carlaw and Lipsey (2002), we argue that the gain under these circumstances is not to be found at the current margin. Instead, it is to be found in the difference between the time path of GDP if technology had remained constant and the path of its actual behaviour as technology changes.

I.3.2 Human Capital

Now consider the accumulation of human capital. To estimate the effects of accumulating more human capital while holding technology constant, we need to think of educating more people to the full level of knowledge existing at some base period, e.g., 1900. Time series data show more time spent in both formal and informal education today than in 1900, partly because there is more to learn for all levels of entry into the labour force. Today’s contribution of human capital to output would be much smaller than it actually is if the longer time in education were spent in learning only the knowledge that was available in 1900. Holding population constant, the difference between these two values measures the results of embodying new technological knowledge in human capital rather than the accumulation of “pure” human capital.

It is conceptually difficult, therefore, to separate what should be regarded as the output effects of accumulating more human capital from those of new technological knowledge in time series calculations. If we include the effects of this new technological knowledge as human capital, we will measure the effects of technological change as increases in human capital rather than as increased in TFP. Whether we do this or not is largely a matter of taste. But, if we do so, we are not justified in concluding that technological change accounts for little of the observed increases in outputs just because increases in inputs, including increased human capital, can do so statistically.

Similar problems arise in disentangling the influence of pure human capital from the influence of the technological knowledge that is embodied in it when dealing with cross sectional data. To make useful cross section comparisons, we need to understand not only how much is known but also what is known. For example, eight years in school studying Marxist philosophy and the sayings of chairman Mao would produce less valuable human capital than eight years studying the “three Rs.”

Correctly measuring the quantity of human capital and allowing for variations in it are important, particularly for TFP studies based on a single macro production function, which usually includes a single index number of human capital as an input. None of the measures that are currently used in practice can separate the accumulation of “pure” human capital from the accumulation of the technological knowledge that it embodies. They thus cause understatements of the contribution of technological change to economic growth.

---

15 Kenneth Pomeranz (2000) has shown that, the 18th century Chinese had a level of literacy more or less equal to that of Europe. But a quantity of human capital similar to that of Europeans does not imply that the Chinese were on the verge of the Industrial Revolution. As Bekar and Lipsey (2001 manuscript) argue, the content of that capital was very different: English human capital providing the knowledge required for successful mechanization of industry while Chinese human capital contained little scientific and engineering content.

16 For another illustration, consider two countries, A which has an elaborate set of technology enhancing policies and B which has none. Years of schooling are higher in A than in B because there is more technological knowledge to be learned. If we ascribe the superiority of A’s productivity over B’s to a higher quantity of human capital, we are measuring differences in available technologies as differences in human capital. Measures that produce similar TFP residuals and account for output differences by
I.3.3 Free lunches and super normal benefits

The concept of a free lunch has become associated with externalities and other unpaid for benefits that accrue to others. But we do not think this captures the full range of benefits that TFP does measure. In a perfectly competitive end-state equilibrium in which foresighted individuals invest in new technologies under conditions of risk, the expected return from all lines of expenditure are equal. Thus the expected returns to investing in a new technology just cover the opportunity cost of its R&D and are equal to the return to investing in new capital that embodies existing technologies. Additional returns would then only arise because of externalities. For this reason, Jorgenson and Griliches, and others who followed them, associated TFP with the “free lunches” of externalities.

In contrast, under the process competition that characterises the real world of technological change in which new technologies are developed under conditions of Knightian uncertainty, and knowledge is at least partially appropriable by those who create it, investments in new technology can, and often do, yield returns well above the going rate of return in the rest of the economy. An entrepreneur who received a return of say 30% on an investment in an unproven new technology when the normal return on investing in exiting technology was 15% would be incensed to be told that half of what he had earned was a “free lunch.” Indeed, it is not a free lunch but a return for undertaking the risks and uncertainties associated with investing in new technologies. We define the difference between the firm’s return to innovation and the return that can be obtained by investing in capital that embodies existing technology as “super normal profits.”

To allow for these returns to risk and uncertainty, as well as genuine free lunches, we define technological change’s “super-normal benefits” as the sum of all associated output increases and cost reductions accruing to anyone in the economy minus the new technology’s development costs. These are the sum of super normal profits that accrue to innovators plus the benefits of external effects in raising outputs elsewhere without corresponding increases in costs.

If a new technology is developed in an oligopolistic industry, as it often is, the full super normal benefits could be appropriated by those who develop it, in which case there are super normal benefits without externalities. If the developers of the new technology cannot appropriate all of the gains for themselves, some of the super normal benefits may become externalities that benefit others. Furthermore, because of technological complementarities, major innovations in one sector provide new opportunities for profitable innovation elsewhere. If these yield a return over and above full costs, they are also external benefits arising from the original innovation. (There are some difficult problems of attribution here that are discussed in Carlaw and Lipsey 2002.)

These considerations do not alter the measured value of TFP changes, which remain increases in output in excess of measured increases in inputs, but they do suggest an alteration in how we view them. As long as one understands that TFP includes part of the return on innovation, there is no problem in calling at a measure of “free lunches.” We prefer the term “super normal benefits” since this term avoids the impression that they are strictly Manna from heaven.

The distinction is worth making because it is so easy to misinterpret what TFP does measure once it is accepted that it is not a measure of technological change. For example, Hulten (p 9, n5) writes that the Hicksian shift parameter, A, “…captures only costless improvements in the way an economy’s resources of labor and capital are transformed into real GDP (the proverbial “Manna from Heaven”). Technical change that results from R&D spending will not be captured by A,…”. In fact,
that part of the gains from R&D that accrues to the innovators and that is in excess of the “normal” rate of return on investing in existing technologies is so captured; it is returns — increases in the value of output — arising from the undertaking of highly risky and uncertain investments in new technology.

I.4. HOW WELL DOES TFP MEASURE SUPER NORMAL BENEFITS

If TFP is a measure of the super normal benefits associated with technological change, we may ask how well it measures them. To isolate this analysis from the issues discussed in the previous section, we assume that all technological changes considered in this section occur costlessly. Thus, all of the effects on output should show up as changes in TFP.

I.4.1 Timing of Cost Reductions

Assume that some free technological advance allows cost reductions on all of some industry’s future output. The effect on measured TFP will depend on the timing of the innovation. Here is an illustrative example.

A product costs $4,000 in year 1 and $540 in year 20 (an average cost fall of 10% per year) and that the resulting increase in output over the same period is an average of 10% per year. Now consider two time paths for these changes.

In case I, the costs fall and the sales rise both at 10% each year. In this case, the industry's TFP will be rising at 20% each year and its contribution to the national TFP figure will be rising as its weight in total TFP rises. If, for example, we let total GDP be constant and the original weight be .02, the final weight will be .148. Thus the contribution to the nation's TFP change will rise steadily from .4% in the first year to 2.96% in the final year.

In case 2, all of the cost reductions come in the first year and sales then expand at 10% each year. In the first year, the industry's TFP increases by nearly 100% (a .865 reduction in costs and an increase in sales of .1). With a weight of .02, this makes its contribution to the increase in national TFP 2% in that year and zero thereafter.

Now this is very interesting. There is the same technology induced reduction in costs and the same increase in output in both cases. All that differs is the timing. Yet in case one, the contribution to the increase in national TFP in the last year, when only 1/20th of the total change occurs, exceeds the contribution in case two in the year when all of its cost reducing change occurs.

It is well known that large productivity increases in industries with small weights in total output do not contribute much to national changes in TFP. This is one reason why the early Industrial Revolution, which was concentrated in the textile industry, caused so little change in national TFP. (See e.g., Crafts and Harley 1992). But our literature search does not reveal any statement that the same increases in technologically driven cost reductions and the same resulting increases in output can give radically different national changes in TFP depending on the timing of the changes in costs and the changes in output.

This is more than a theoretical possibility. Something like this occurred in the automobile industry with the introduction of Henry Ford's Model T. The price of cars fell quickly while it took a decade for demand to respond fully.

---

18 Note that this timing problem is different from the one discussed in Helpman (1998) because the technological change considered here is costless. In contrast, Helpman’s timing problem produces TFP slowdowns because of the timing of the allocation of scarce resources differs from the timing of the benefit of innovation.

19 If a series increases at 10% per year, it is 7.4 times as large in 20 years; hence the weight goes from .02 to .148.

20 The model T was introduced in 1909. In the first year when sales of the most popular model, the touring sedan, went over 100,000 (1913) its price was $600. Sales reached a peak of just under 900,000 cars at a price of $380 ten years later in 1923.
This is a general phenomenon associated with the introduction of a new consumer’s durable that requires the development of many ancillary products and services as well as much time to persuade consumers that the new product is here to stay. With cars, it took decades for the full supporting infra structure of petroleum refining, distribution, roads, motels, etc. to be developed. Slowly over this time, Americans became mobilized, until by the late 1920s, the family without a car was the exception rather than the rule. For another case, US rural electrification came to the early 1930s. At first, demand hardly increased. But slowly, over the next decade, farmers bought electrical milking machines, refrigerators, cooling systems and many other electrically driven consumers and producer’s durables. As a result, by the end of the decade rural demand for electricity had expanded greatly.

So the case in which costs fall suddenly and demand expends only slowly over years, and even decades, cannot be dismissed as a theoretical curiosum. In such cases, TFP measures will be much smaller than if the identical technologically driven change in costs had occurred slowly over time accompanied by the same overall increase in demand.

I.4.2 The treatment of R&D

In the national accounts of many countries, R&D is recorded on the input side as a current cost and not given any direct output. Offsets appear only when, and if, the results of R&D are used to reduce the costs or increase the output of final goods. If it follows, for example, that if an established local firm shifts resources from making machines into R&D to design better machines, it will record a fall in output with no change in input costs and hence, ceteris paribus, a reduction in its TFP. Whatever else we may think about the desirability of having such a characteristic in TFP measures, the resulting fall in TFP does not measure any technological regression.

Also, a start-up firm that does only R&D in one year will have its input valued at cost and record an equal negative profit, since it has no sales. Thus, by definition, not only will it show a negative contribution to TFP, it will show no contribution to current output. Of course, it may be contributing to technological dynamism by producing valuable new patentable technologies. If the patent produced by the R&D is sold abroad, this is recorded as a capital transfer. No income is ever recorded and hence there is no TFP gain at any point in the process. This is also the case if the start up firm is itself sold to a foreign multinational. If the patent or the firm is sold to another domestic firm, this is regarded as a capital transfer and there is no possible effect on TFP until after the new technology is put to use.

So in these respects, TFP measures nothing systematic concerning the value created by R&D until the new technologies are used to reduce costs or increase the production of final goods and services. Furthermore, there is a potential for getting temporarily misleading TFP measures as the economy switches resources from investment in producing hardware to investment in producing ideas. The figures may be permanently misleading if the intellectual property is sold to foreigners.

---

21 We are indebted to officials at Statistics Canada for the following observation. “If a reporting firm does capitalize R&D, we would record it as an investment and remove associated costs from output so TFP on the production side might not be affected. In any event, we would not…attach an output to the investment other than the ‘real’ value of the inputs, so no [change in] TFP would be possible.”

22 Shifting significant amounts of real resources within one firm between production and R&D does not often happen but this is what the economy does and it is heuristically simpler to think of this happening within one firm.

23 Particularly in small countries, many firms engage in start up behaviour and then sell out to foreign multinationals, realizing the return on their R&D expenditures from the sale price. Indeed, tax advantages given to small firms often encourage such activities. None of this value-creating activity, often in the “new economy,” will show up as income, or as increases in TFP.
Many users of TFP statistics understand these limitations but, as our initial quotations show, not everyone does. TFP figures pick up none of these activities, although they are surely evidence of technological dynamism.

I.4.3 Induced Investment
Hulten (2000 p 34-5) argues that where technological change induces some capital investment, part of the effects will be incorrectly assigned to an increase in capital. He considers a case of a balanced growth path with Harrod-neutral technological change. All of the growth in output is due to technological change in the sense that if there were no such change, output would be constant. But because capital is also growing in order to maintain a constant ratio of capital to efficiency units of labour, a proportion of the rise in output equal to capital’s relative share will be attributed to more investment and only the proportion equal to labour’s share will be attributed to technological change. We do not think this attribution to increased capital and to technological change is incorrect. If the capital stock were to be held constant by fiat while technological change continued, output would only be growing at the rate of increase in efficiency units of labour weighted by labour’s share. The rest of the increase is due to more capital investment.24

I.4.4 Omitted inputs: Natural Resources Made Explicit
Failure to measure any input can bias TFP downwards.25 We illustrate with the important case of natural resources. Following Solow (1957), growth theorists typically define physical capital to include natural resources, land, minerals, forests etc.26 Almost invariably, however, everything that is done after that is appropriate for physical and human capital but takes no account of the characteristics of natural resources. For example, although the stocks of plant and equipment can be increased more or less without limit, the stocks of arable land and mineral resources are constrained within fairly tight limits.

The shortcomings of this treatment of resources can be seen in the contrast between two positions. The first is the prediction derived from the standard formulation in equation (I.2) above that measured capital and labour could have been increased at a common steady rate from say 1900 to 2000 with constant technology and no change in living standards. The second is the belief that the supply of some key natural resources and much of the environment’s capacity to handle pollution could not have survived a six fold increase in industrial activity with 1900 technology. To reconcile these conflicting positions, we need to recognize that the capital that would need to grow would include such resource inputs as acres of agricultural land, quantities of mineral and timber resources, available “waste disposal” ecosystems, supplies of fresh water, and a host of other things that are ignored by the standard theoretical treatments and most applied measurements of capital. (Since technology is assumed to be constant in the above exercise, this growth cannot be the result of increased efficiency in the use of natural resources due to new techniques.)

24 Would this be a serious problem in practice if Hulten’s argument for the under measurement is accepted? We think not because it is hard to locate any real technological change that enhances labour’s productivity when combined with more and more units of capital that embody static technology. This problem is not often pointed out because there is no tradition in either the macro growth or the TFP literature of relating such theoretical concepts as Harrod neutral technological change to the known facts about actual changes in technology as detailed in books such as Rosenberg’s Inside the Black Box (1982).

25 Hulten (2000) makes a similar point in specific reference to the omitted negative effects of economic growth on the environment.

26 Solow himself was aware of the omitted variables effect and warned about the bias this could cause in measurement of the residual (Solow, 1957). Also Hulten (2000: 51) discusses the problem of omitted environmental variables. He notes, however, that a solution to the omitted variables problem “is an impossibly large order to fill.” This may be so but it is not calculated to increase confidence in the usefulness of the TFP measure.
To illustrate some of the problems associated with the omission of natural resources, let the underlying production function be:

\[ Y = BK^\alpha (nR)^{1-\alpha}, \]

where \( K \) is accumulating factors, \( R \) is natural resources and \( n \) is an efficiency coefficient standing for the technology of resource use. Taking time derivatives yields proportional rates of change of

\[ \frac{\dot{Y}}{Y} = \frac{\dot{B}}{B} + \alpha \frac{\dot{K}}{K} + (1-\alpha) \left( \frac{\dot{R}}{R} + \frac{\dot{n}}{n} \right). \]

Let those measuring TFP assume the production function to be

\[ Y' = AK, \]

so that

\[ \frac{\dot{Y}'}{Y'} = \frac{\dot{A}}{A} + \frac{\dot{K}}{K}. \]

Now calculate TFP from equation (II.3):

\[ \frac{\dot{TFP}_m}{TFP_m} = \frac{\dot{A}}{A} = \left( 1-\alpha \right) \left( \frac{\dot{R}}{R} + \frac{\dot{n}}{n} \right) - \frac{\dot{K}}{K}. \]

Letting \( \frac{\dot{B}}{B} = 0 \), yields

\[ \frac{\dot{A}}{A} = \left( 1-\alpha \right) \left( \frac{\dot{R}}{R} + \frac{\dot{n}}{n} \right) - \frac{\dot{K}}{K}. \]

So

\[ \frac{\dot{A}}{A} = 0 \text{ if } \frac{\dot{R}}{R} + \frac{\dot{n}}{n} = \frac{\dot{K}}{K} \]

\[ \frac{\dot{A}}{A} > 0 \text{ if } \frac{\dot{R}}{R} + \frac{\dot{n}}{n} > \frac{\dot{K}}{K} \]

\[ \frac{\dot{A}}{A} < 0 \text{ if } \frac{\dot{R}}{R} + \frac{\dot{n}}{n} < \frac{\dot{K}}{K} \]

The value of measured TFP changes depend only on the above relation not on any cost associated with changing \( n \). To illustrate, let \( \frac{\dot{R}}{R} = 0 \) and \( \frac{\dot{n}}{n} = \frac{\dot{K}}{K} \) so that measured TFP remains constant, indicating, on the standard interpretation of TFP measures, that there is no technological change. First, assume that the technological change represented by \( \frac{\dot{n}}{n} \) is free, coming with no resource cost. Then all of the benefit is a free lunch, yet TFP remains constant. Second, assume that there is a positive resource cost equal to \( xK\% \) (\( 0 < x < 1 \)). Now the level of \( Y \) will be lower by \( xK\% \) but the rates of growth of \( Y \) and \( A \) will be unchanged so that (II.3) will still fit the data giving no change in TFP. In the case studied in section 3, TFP measured the free lunches associated with technological change. In this case, changes in TFP do not measure free lunches. Instead, they measure the extent to which the rate of increase in the use of resources measured in efficiency units exceeds the rate of increase in the accumulating factors, irrespective of the cost of increasing resource efficiency.

1.4.5 Aggregation of Inputs

Next consider the problems of obtaining a measure of inputs for the production function at whatever the level of aggregation over which that function is defined. To make our point, let the firm’s real microeconomic production function be

\[ Y = BM^\gamma N^\delta P^\tau R^\eta \]

where \( M \) and \( N \) are two types of capital and \( P \) and \( R \) are two types of labour used within the firm and \( B \) is a productivity parameter.

---

27 The absence of explicit resource inputs from the neo-classical growth model, poses no problem for the measurement of income because all of the value of consumed resources must show up as income for the labour and capital services involved in extracting and processing them.

28 What follows is a simple algebraic demonstration of the empirical findings of Jorgenson, Gollop and Fraumeni (1987) chapter 8. They find that failure to include quality effects results an upward bias in the contribution of inputs to output growth, when aggregation occurs.
Now we assume that an aggregate production function is used to calculate the firm’s TFP:

\[(f.4.2) \quad Y' = AK^aL^b + \alpha + \beta = 1.\]

Note that \(Y\) and \(Y'\) measure the same output but we wish to keep track of the production function (aggregated or disaggregated) on which we are taking derivatives. Let the firm’s aggregate capital be calculated as \(K = p_mM + p_nN\) where prices are equal to marginal products: \(K = (\gamma B M^{\delta} N^\delta P^\epsilon R^\eta) M + (\delta B M^{\delta-1} N^\delta P^\epsilon R^\eta) N = (\gamma + \delta)Y\). Similarly, let the firm’s aggregate labour be \(L = p_pP + p_rR\), or \(L = (\epsilon + \eta)Y\).

Now let \(B\) in (I.4.1) change continuously through time with unchanged inputs of \(M, N, P,\) and \(R\). Then \(Y' = (dY/dB)(B) = (M^\delta N^\delta P^\epsilon R^\eta)(B)\). Any change in \(B\) now shows up as changes in the two aggregate inputs, \(K\) and \(L\):

\[
\begin{align*}
K = (dK/dY)(dY/dB)(B) &= \gamma(dY/dB)(B), \\
L = (dL/dY)(dY/dB)(dB/dt) &= \epsilon(dY/dB)(dB/dt).
\end{align*}
\]

So, using the fact that \(Y'\) is homogeneous of degree one in \(K\) and \(L\):

\[
Y' = (dY'/dB)(B) = (dY'/dK)(K) + (dY'/dL)(L).
\]

Thus a Divisia index based on the two aggregated inputs in the firm’s aggregate production function gives:

\[(f.4.3) \quad \text{TFP} = \frac{Y'}{Y'} - \alpha \frac{K}{K} - \beta \frac{L}{L} = 0.\]

If instead, we had calculated a Divisia index from the firm’s disaggregated production function, we would have obtained the correct answer:

\[(f.4.4) \quad \text{TFP} = \frac{Y}{Y} - \alpha \frac{M}{M} - \beta \frac{N}{N} - \gamma \frac{P}{P} - \delta \frac{R}{R} = \frac{A}{A} = \frac{B}{B}.
\]

since all four percentage changes in the disaggregated inputs are zero.

Labour in different uses can be measured in physical units, such as labour hours, with different qualities being converted into labour hour equivalents. In contrast, capital in different uses is composed of physically different items that cannot be aggregated physically. So monetary units must be used to aggregate the capital (whether stocks or service flows) used in any real production process. Consider the calculation based on (I.4.2) and (I.4.3) if \(L\) had been aggregated by physical units and \(K\) by its marginal product. Then \(L/L = 0\) and TFP would increase by \([1 - (\gamma + \delta)]B(B)\) while the rest of the increase would be ascribed to an increase in capital. Since different kinds of capital are usually expressed in monetary units to make them comparable, this kind of mixed unit aggregation is a common case. Then, the increase in output due to a productivity increase will be divided between a measured increase in the quantity of capital (in proportion to capital’s share) and a measured increase in TFP (in proportion to labour’s share).29

So the effects of technological changes that are felt below the levels of aggregation at which the production functions are defined will tend to show up at least partially as changes in the quantities of inputs. Since some amount of aggregation of inputs must always take place before any TFP index is calculated, some amount of technical change will always be recorded as changes in the quantity of inputs, especially capital.

Jorgenson and Stiroh (2000) argue for making disaggregated measures at the industry level. “Productivity growth…differs widely among industries.” (p.161) so that disaggregation “…is especially critical in evaluating the validity of explanations of economic growth that rely on developments at the level of industries, such as technology-led growth.” (p.166). Even when

---

29 An identical argument applies if we write (I.4.1) as \(Y = B(mM)^\eta(nN)^\delta(pP)^\epsilon(rR)^\eta\), where lower case letters are efficiency parameters, and we allow one of more of them to change with \(B\) held constant.
calculations are made at the industry level, our analysis shows that substantial amounts of technological change will show up as increases in the industry's measured inputs.

I.4.6 Aggregation When Calculating a TFP Index

Now assume that a correct measure of the quantity of each type of input available at whatever level of aggregation is being used to calculate TFP. Percentage changes in each input can then be weighted, and summed to get overall percentage change in inputs to be compared with the percentage change in output to calculate an index of TFP changes. The weighting coefficients on each type of input are the relative shares in expenditure.

One assumption that is critical for the validity of this procedure is that the marginal products of each factor of production are equated in all of its uses. There are three possible reasons why this assumption may not hold. First, as we discuss below, the economy may be in a transition between one competitive equilibrium and another. Second, as Hall (1988) and Basu and Fernald (1997) discuss, the use of revenue shares in the presence of imperfect competition implies that marginal products will not be equated even when the system is in full equilibrium. For our purposes, market structure does not matter as much as the transitional marginal products associated with moving from one equilibrium to another. Furthermore, Basu and Fernald (1995) show that under conditions of imperfect competition, aggregation of the sort done by Caballero and Lyons (1990 and 1992) will over estimate the free lunches associated with TFP. We demonstrate the conditions under which these spillovers are also underestimated. The third possibility is some combination of the first two. Jorgenson, Gollop and Fraumeni (1987) conduct a comprehensive analysis of sectoral substitution finding a number of things including that the hypothesis for Hicks neutrality (i.e., the rate of productivity growth is independent of quantities of intermediate, capital and labour inputs and the value shares are independent of time) is rejected in thirty-nine of the forty-five industries studied. Some of their other findings show that biases for productivity growth with respect to the three inputs are varied depending on the input and the industry measured. However, none of the empirical tests conducted by Jorgenson, Gollop and Fraumeni ask the question we deal with below; How much of the super normal benefits associated with a technological change is picked up by TFP when that technological change drives marginal products away from their equilibrium values?

Consider two concepts of equilibrium. In full equilibrium, all adjustments have been made and no agent wishes to alter his or her behaviour from period to period. In transitional equilibrium, each agent does not wish to alter behaviour in the period in question but behaviour does alter from period to period. Consider the following stylization of a two sector economy. Let the primary production sector be

\[ X = A(L_x)^\alpha (K_x)^\beta \]

Let the manufacturing production sector be

\[ Y = B(L_y)^\gamma (K_y)^\delta \]

Let the resource constraints in the economy be

\[ L = L_x + L_y \text{ and } K = K_x + K_y \]

The aggregate accounting identity for this economy is

\[ P_x X + P_y Y \equiv w_x L_x + w_y L_y + r_x K_x + r_y K_y \]

The \( P_i \) are output prices and the \( w_i \) are input prices for each sector where \( i = (x, y) \). If we take \( P_y \) as the numerare price for the system, then

\[ \frac{P_x}{P_y} X + Y \equiv \frac{w_x}{P_y} L_x + \frac{w_y}{P_y} L_y + \frac{r_x}{P_y} K_x + \frac{r_y}{P_y} K_y \]

is the accounting identity.

Assuming full perfectly competitive equilibrium, we can relate prices to marginal products in the following well known way:
\[ \frac{P_x}{P_y} = \frac{MP_{L_x}}{MP_{L_y}} = \frac{MP_{K_y}}{MP_{K_y}} \]
\[ \frac{w_x}{p_y} = \frac{w_x}{p_y} = MP_{L_x} \]

and
\[ \frac{r_y}{p_y} = \frac{r_y}{p_y} = MP_{K_y} \]

Simplify further by normalizing \( P_x \) to be 1.

Now consider what happens when the productivity parameter for the manufacturing sector costlessly increases. There will be an immediate increase in the marginal product of labour in manufacturing causing labour to migrate from primary to manufacturing production. If we maintain the assumptions of the full equilibrium, implicitly assuming that the adjustment takes place instantaneously, we can calculate the change in TFP in the following way:

\[ \frac{\Delta TFP}{TFP} = \left( \frac{P_x X}{P_x X + Y} \right) \dot{X} + \left( \frac{Y}{P_x X + Y} \right) \dot{Y} - \left( \frac{w_x L_x}{w_x L_x + w_y L_y} \right) \dot{L}_x - \left( \frac{w_y L_y}{w_x L_x + w_y L_y} \right) \dot{L}_y \]

\[ - \frac{r_y K_y}{r_y K_x + r_y K_y} \dot{K}_y - \frac{r_y K_y}{r_x K_x + r_y K_y} \dot{K}_y \]

which is simply the Divisia index for TFP.

The assumption of full perfectly competitive equilibrium along with some straightforward algebraic manipulation implies

\[ \frac{\Delta TFP}{TFP} = \frac{1}{P_x X + Y} \left[ P_x \dot{X} + P_y \dot{Y} \right] - \frac{w_x}{w_x L_x + w_y L_y} \left[ \dot{L}_x + \dot{L}_y \right] - \frac{r_y}{r_x K_x + r_y K_y} \left[ \dot{K}_x + \dot{K}_y \right] \]

From the resource constraints we know that
\[ \dot{L}_x = -\dot{L}_x, \text{ and } \dot{K}_y = -\dot{K}_y \]

so the second and third terms are zero. Substituting the time derivatives of the production function and the definition of \( P_x \) implies:

\[ \frac{\Delta TFP}{TFP} = \frac{1}{P_x X + Y} \left[ \frac{\dot{B}}{B} \right] \]

In this case, TFP exactly measures the gains associated with the free productivity increase in sector \( Y \).

Now consider the second type of equilibrium. What happens if the transition is not instantaneous? When the productivity change occurs, marginal products are driven out of equilibrium. But now prices do not instantaneously adjust because labour and capital are not instantaneously mobile. We can determine the direction of bias in the prices:

\[ \frac{P_x}{P_y} < \frac{MP_{L_x}}{MP_{L_y}}, \text{ and } \frac{P_x}{P_y} < \frac{MP_{K_y}}{MP_{K_y}} \]

Again \( \frac{w_x}{p_y} = MP_{L_y}, \text{ and } \frac{r_x}{p_y} = MP_{K_y} \).

However now \( \frac{w_x}{p_y} = \frac{P_x}{P_y} \frac{MP_{L_x}}{MP_{L_y}} < \frac{MP_{L_x}}{MP_{L_y}} \), and \( \frac{r_x}{p_y} = \frac{P_x}{P_y} \frac{MP_{K_y}}{MP_{K_y}} < \frac{MP_{K_y}}{MP_{K_y}} \).
Let \( G \in (0,1) \) be the gap between the marginal products of labour and full equilibrium prices and \( \hat{G} \in (0,1) \) be the gap between the marginal products of capital and full equilibrium prices so that:

\[
\frac{P_x}{P_y} = \frac{G(MP_{L_x})}{MP_{L_x}}, \quad \frac{P_x}{P_y} = \frac{\hat{G}(MP_{K_x})}{MP_{K_x}}
\]

and

\[
\frac{w_x}{P_y} = \frac{P_x}{MP_{L_x}} = \frac{G(MP_{L_x})}{MP_{L_x}} = G(MP_{L_x}), \quad \frac{r_y}{P_y} = \frac{P_x}{MP_{K_x}} = \frac{\hat{G}(MP_{K_x})}{MP_{K_x}} = \hat{G}(MP_{K_x})
\]

Once again normalize \( P_y \) to be one. The change in TFP is expressed as:

\[
\frac{TFP'}{TFP} = \frac{P_y}{P_x} \hat{X} + \frac{Y}{P_x} \hat{Y}
\]

Now substituting in the new definitions of the prices and the time derivatives of the production functions yields:

\[
\frac{TFP'}{TFP} = \frac{1}{P_x + Y} \left[ G(MP_{L_x}) \hat{L}_x + \hat{G}(MP_{K_x}) \hat{K}_x + (MP_{L_y}) \hat{L}_y + (MP_{K_y}) \hat{K}_y + \frac{\hat{B}}{B} \right]
\]

We can now subtract the original TFP calculation from the second to determine if there is any bias between the full equilibrium and the transitional equilibrium. If

\[
\frac{TFP'}{TFP} < 0
\]

then the transitional calculation under estimates the gains. If the inequality is reversed, then TFP overestimates the gains.

First we note that \( \frac{TFP}{TFP} \) is a positive term in \( \frac{TFP'}{TFP} \) so that we can eliminate \( \frac{1}{P_x + Y} \left[ \frac{\hat{B}}{B} \right] \) from both TFP calculations leaving just the remaining terms in \( \frac{TFP'}{TFP} \). Making use of the fact that \( \hat{L}_x = -\hat{L}_y \) and \( \hat{K}_x = -\hat{K}_y \) we get the following:
We are interested in signing this expression to determine if there is any bias in the Divisia index when the marginal products are not fully adjusted to long run perfectly competitive equilibrium. To do this note the following $\hat{G}(MP_{L_x}) - MP_{L_x} < 0,\hat{G}(MP_{K_x}) - MP_{K_x} < 0$.

Thus the two expressions left to evaluate are:

\[
\frac{1}{P_x X + Y} - \frac{1}{w_x L_x + w_y L_y} \left[ G(MP_{L_x}) - MP_{L_x} \right] L_x,
\]

\[
\frac{1}{P_x X + Y} - \frac{1}{r_x K_x + r_y K_y} \left[ \hat{G}(MP_{K_x}) - MP_{K_x} \right] \hat{K}_x.
\]

By evaluating the expressions around zero we get:

\[
\frac{1}{P_x X + Y} - \frac{1}{w_x L_x + w_y L_y} < 0 \text{ if } \alpha < 1 \text{ and } \gamma < 1 \text{ and}
\]

\[
\frac{1}{P_x X + Y} - \frac{1}{r_x K_x + r_y K_y} < 0 \text{ if } \beta < 1 \text{ and } \sigma < 1.
\]

This says that $\frac{TFP'}{TFP} - \frac{TFP'}{TFP} < 0$ for $\alpha < 1, \gamma < 1, \beta < 1$ and $\sigma < 1$. We can also see that there is no bias if all of these parameters are just equal to one, which implies that the production functions have increasing returns to scale.

Summary of the last two sections: On aggregation, we have found two important sets of circumstances in which technological change tends to be recorded as increases in the quantity of factors. The first
occurs even when equilibrium relations hold, whenever inputs are valued at market prices in order to be aggregated up to the level at which the TFP index is calculated (as is typically the case with heterogeneous capital inputs). The second occurs when long run competitive equilibrium relations do not hold so that the value weights used in calculating the index of TFP differ according to the different marginal products of one factor in various uses. We have shown that this may bias TFP downward even where all lines of production activity are constant returns to scale and it also encompasses the findings of Basu and Fremani (1995) by showing where an upward bias occurs for increasing returns to scale.

II. COUNTERFACTUAL MEASURES

If TFP does not measure technological change, (in the sense that there can be income-increasing technological change with constant TFP), we must look for an alternative measure. We start by reconsidering technological change and then look at how its effects might be ascertained.

II.1. The conceptualization of Technological Change

The big problem in measuring the effects of the technological change, is to separate it from the capital accumulation that embodies much of it. This is a difficult, possibly an impossible, task—one that is not often considered in detail by those who assume that an aggregate production function of the type given in equation (I.2.2) can describe growth over long periods of time.

Conceptually, these two concepts are separated in two steps. First, we hold all technologies constant at what was known at some base period. Then we accumulate more physical capital that embodies those base period technologies, and more human capital in the form of more education only in what was known in the base period. The resulting change in output is due to “pure” capital accumulation. The difference between this change and the actual change in output is “due to” or “enabled by” technological change in the sense that it could not have happened without it. Measured over a period of a century or more, the difference due to technological change would be very large indeed.

Here are just a few illustrative examples of what the constant-technology experiment would reveal if conducted between now and a base period of 1900.

- Feeding 6 billion people with the agricultural technologies of 1900 would have been impossible.30
- Pollution would have become a massive problem.
- Exhaustion of specific resources would have become a serious problem. Since most new technologies are absolutely saving in resources,31 to produce the value of today’s manufacturing and service output with 1900 technologies would have required vastly more resources than are currently being used. Furthermore, with no changes in technological knowledge, the scope to replace materials that were becoming scarcer with more plentiful alternatives would have been greatly restricted.
- The marginal utility of income would have diminished rapidly as people accumulated larger and larger stocks of the 1900-design durable goods, and consumed increasing amounts of 1900-style services and perishables.

30 Of course, population is endogenous and it is not clear how much population would have increased if food-producing technologies had remained frozen at their 1900 levels. However, Western sanitation, health and medical practices had already lowered worldwide death rates and had led to large increases in life expectancies with a resulting population boom. Thus, some large population expansion would certainly have occurred, creating serious Malthusian pressures..

31 This is a process that Grubler (1998, p.240) calls “dematerialization.” Among the many illustrations that he quotes are these: during the period 1975-94 “Total materials requirements per unit of (constant) GDP have declined between 1.3% per year in Germany, 2% per year in Japan, and 2% per year in the Netherlands.”
While all this is, of course, speculation, the main point is that growth of labour and capital at 20\textsuperscript{th} century rates with truly constant technology would have produced massive problems in many dimensions.\textsuperscript{32}

II. 2 GPT-Driven Growth\textsuperscript{33}

We view economic growth as being driven by a succession of general purpose technologies (GPTs), which create opportunities for profitable investments in a large set of new product, process, and organisational technologies.\textsuperscript{34} Even if there are no free lunches; no externalities; no super-normal profits on investment in these new technologies,\textsuperscript{35} they bring great economic gain because the opportunities for these investments \textit{would not have existed} without the GPT. These opportunities come from the technological complementarities that are created by radically new technologies. They have been a major (we would say the major) source of growth over the last three centuries.

The evolution of major new technologies, particularly general purpose ones, prevent the marginal product of capital from declining continuously over time because one innovation enables another in an evolution that stretches over decades, even centuries. Even if the development costs of each of the new technologies that are enabled by a new GPT are just covered by sales revenues, the path dependency in which new inventions and innovations build on existing ones, implies that the marginal product of capital will eventually be higher than it would have been under conditions of static technology. Eventually, however, the possibilities for exploiting one particular GPT begin to peter out. Think, for example, what the range of new innovation possibilities and the rate of return on investment would now be if the last GPTs to be invented had been steam for power, the iron steam ship for transport, steel for materials (no man-made materials) and the telegraph for communication (the voltaic cell but no dynamo).

These considerations suggest two important conclusions. First, the concept of the super normal gains measured by TFP is based on what is happening to current output and current costs. It does not, therefore, cover the important cases where one innovation enables others, often in an indefinitely long future stream. Think, for example, of how many of today’s innovations depend on the dynamo or the computer chip. So the social benefits from specific technological changes go well beyond what can even ideally be measured by TFP.

Second, the economic benefit of new technologies is in the \textit{future path of returns} rather than in any gain on the current margin between the new and the old technologies. With the opportunities created by the new technology for further technological innovations that stretch over future decades, the actual rate of return may hold constant instead of falling as it would if technology had remained static.

\textsuperscript{32} In another demonstration of these points, the calculations of the Club of Rome in the 1970s showed the folly of believing that production could long be increased at current rates with unchanged technology. The Club’s predictions of doom were falsified by continued technological advance, which invalidated their calculations on resource exhaustion and unsustainable pollution. But these mistaken predictions do show for how few decades current world growth rates could be sustained in a world of static technology.

\textsuperscript{33} This section draws on Carlaw and Lipsey (2002).

\textsuperscript{34} General purpose technologies share some important common characteristics: they begin as fairly crude technologies with a limited number of uses; they evolve into much more complex technologies with dramatic increases in the range of their use across the economy, and in the range of economic outputs that they help to produce. A mature GPT is defined formally as a technology that is widely used, has many uses, and has many complementarities with other existing technologies. (See Lipsey, Bekar and Carlaw 1998a.)

\textsuperscript{35} Of course, we have no doubt that there sometimes are enormous returns over and above development costs.
The important message is that there needs to be no observable impact of the new technology on rates of return; instead the impact is between what actually happens to returns over some future time period and what would have happened in the absence of the technology. The need to make this counterfactual observation makes it difficult to observe the effects of major innovations on rates of return. Nonetheless, the benefit grows over time as the gap grows between what the rate would have been as it fell continuously under the impact of capital accumulation and constant technology and what it actually is.

Alternative measures are being researched by several authors who are currently investigating the role of investment-specific technological change (a measure of the quality of investment in machinery and equipment) on economic growth. For examples see Greenwood, Hercowitz and Krusell, (1997 and 2000) and IMF (2000). This research attempts to measure directly the new technology that is embodied in capital goods. Some preliminary findings for Canada by Carlaw and Kosempel (2001) show that in the period between 1974 and 1996, while the rate of TFP growth was declining, investment-specific technological change was growing. The interesting feature of this research is that it provides a proximate measure of technological change taken from independently measured data as an alternative to the residual measure of TFP.

II.3 The Measurement of Technological Complementarities

The evolution of a GPT prevents the marginal product of capital from declining continuously over time. One innovation enables another, as we see from studying the course of any major GPT. Eventually, however, the possibilities for exploiting one particular GPT begin to peter out—relatively quickly with some GPTs and only over centuries with others. Thus the time path of cumulative investment opportunities related to a particular GPT from its inception often resembles a logistic curve, rising slowly at first when the GPT is still in its crude specific use stage; then rising ever more rapidly as each innovation expands the space for further innovations at an increasingly rapid rate; then slowing as the possibilities for new technologies that are enabled by the GPT begin to be exhausted.

For simplicity, we assume that the cumulative curve has a linear portion at the outset and then eventually flattens as possibilities begin to be exhausted.

Even if the development costs of each of the new technologies that are enabled by a new GPT are just covered by sales revenues, the path dependency of invention and innovation, in which new ones build on existing ones, implies that the marginal product of capital would eventually be higher than it would have become under conditions of static technology. This important phenomenon, which was emphasized by Nelson and Winter (1982 p. 256), is illustrated in Figure 1, which gives two time paths for the return on capital. The first is constant along the arrowed curve MP\textsuperscript{1}, assuming a succession of overlapping GPTs. Along this trajectory, investments in successive innovations are assumed each to earn only their opportunity cost as measured by the return on investment in existing technologies. Changes in TFP will thus be zero (which is why TFP is emphatically not a measure of technological change).\textsuperscript{37} The second curve, MP\textsubscript{2} falls on the assumption that no new GPTs are invented after time \( t \) so that returns eventually fall as innovation possibilities get used up.\textsuperscript{38} The gain from technological change is measured by the gap between MP\textsuperscript{1} and MP\textsubscript{2}.\textsuperscript{39}

\textsuperscript{36} Freeman and Louca (2001) is built around phenomena of this type of relation.

\textsuperscript{37} In this case, the value of the output from the new technologies is just equal to the cost of all inputs including R&D valued at opportunity cost. The implications for misinterpreting TFP measures in such cases are investigated in Lipsey and Carlaw (2002).

\textsuperscript{38} Those who do not like the GPT concept can think of holding all technological knowledge constant, which would merely steepen the R\textsubscript{2} curve leaving the rest of the argument intact.

\textsuperscript{39} This type of historical counterfactual is not to be confused with the limited counterfactual commonly used in cleometrics and aptly criticized in Chapter 1 of Freeman and Louca. Here we are considering a world with and without all new GTPs and all the innovations that they enable.
So whether or not there are super normal profits or externalities in the form of technology transfers for which the recipients would have paid more than they actually did pay, and whether or not there is a discrepancy between private and social rates of return, the technological complementarities that arise from radically new technologies have been a major source of growth over the last three centuries.

II.4 A Model of GPT-driven growth with constant TFP

We illustrate the point that growth can be sustained in the absence of scale effects, externalities, free lunches and TFP changes by presenting an abbreviated version of a model of GPT-driven endogenous growth presented Carlaw and Lipsey (2002). There are three sectors in this model, consumer goods, applied R&D and general purpose R&D. There are three specific forms of technological complementarity in the model, none of which are externalities. First, the stock of knowledge created in the general purpose sector increases the marginal productivity of resources devoted to the applied R&D sector. Second, some optimally chosen portion of the stock of applied R&D knowledge increases the marginal product of resources used to produce consumer goods. Third, the remaining portion of the stock of applied R&D knowledge is allocated to the general purpose sector to increase the marginal productivity of resources in that sector.

The resource constraint is:

\[ R = r_{ct} + r_{at} + r_{gt} \]

\( R \) represents all non produced factors of production. The factors devoted to the consumer goods, the applied R&D and the general purpose R&D sectors are \( r_{ct}, r_{at} \) and \( r_{gt} \).

Consumer goods production is defined as:

\[ C_t = (\mu A_t) \left( r_{ct} \right)^\alpha \]

where \( \alpha \in [0,1] \) , \( \mu A_t \) is the portion of the stock of applied R&D devoted to the consumer goods sector.

Applied R&D is defined as:

\[ a_t = G_t r_{at}^\beta \]

\[ A_t = a_t - \varepsilon A_t \]

where \( \beta \in [0,1] \) and \( \varepsilon \) is the rate of obsolescence that applies to the stock of R&D-produced knowledge. Increases in \( B_t \) raise the marginal productivity of resources in applied R&D and this in turn affects the marginal productivity of resources in consumer goods.

The general purpose sector is defined as:

\[ g_t = (1 - \mu) A_t r_{gt}^\sigma \]

\[ G_t = g_t - \delta G_t \]

The infinite horizon maximization problem for this economy is:

---

40 This model is simplified from that of Carlaw and Lipsey in a number of ways. Two important changes are that the system is modelled in continuous time and there is no uncertainty.

41 It is not an externality because in order to achieve this higher productivity, consumption and applied R&D had to forgo the production associated with the resources that went into the GP sector to produce the new knowledge.
Without going through the tedious maximization exercise to find the stationary values of the costate variables, \( r_{c,t}, r_{a,t}, r_{g,t} \) and \( \mu \), we can show the conditions for sustained “balanced growth” by looking that the time derivatives of the production functions in each sector. If we assume that the system has stationary values for the costate variables in the balanced growth equilibrium, the derivatives of the production functions in each sector with respect to time are:

\[
\begin{align*}
\frac{\dot{c}}{c} &= \frac{\dot{A}}{A}, \quad \frac{\dot{a}}{a} = \frac{\dot{B}}{B}, \quad \text{and} \quad \frac{\dot{b}}{b} = \frac{\dot{A}}{A}.
\end{align*}
\]

Substituting the laws of motion for the two knowledge stocks we get:

\[
\begin{align*}
\frac{\dot{c}}{c} &= \frac{\dot{b}}{b} = \frac{B_t}{A_t}(r_{c,t})^\beta - \varepsilon \\
\frac{\dot{a}}{a} &= \frac{1}{B_t}(1 - \mu)A_t(r_{b,t})^\eta - \delta
\end{align*}
\]

Balanced growth occurs when the growth rates of output in all sectors are constant. To check this, take the derivatives of the growth rates and set them to zero to find balanced growth.

\[
\begin{align*}
\frac{\ddot{c}}{c} &= \frac{\ddot{b}}{b} = \frac{\dot{B}_t}{B_t} - \frac{\dot{A}_t}{A_t} = 0 \\
\frac{\ddot{a}}{a} &= \frac{\dot{A}_t}{A_t} - \frac{\dot{B}_t}{B_t} = 0
\end{align*}
\]

The only way that the above can hold is if \( \frac{\dot{B}_t}{B_t} = \frac{\dot{A}_t}{A_t} \), which is the balanced growth solution.

The important thing to note is that the stock of knowledge is created from resources with have an opportunity cost in other uses. And each resource receives its marginal product in every period of the model. Thus, there are no externalities or scale effects driving endogenous growth in this model. In this model, TFP will be zero even though economic growth in balanced growth equilibrium is sustained at a constant, positive rate. To illustrate we calculate a Divisia index for TFP as follows:

\[
\frac{\text{TFP}}{\text{TFP}} = w_c \frac{\dot{c}}{c} + w_a \frac{\dot{a}}{a} + w_b \frac{\dot{b}}{b} - q \frac{\dot{r}_c}{r_c} - q \frac{\dot{r}_a}{r_a} - q \frac{\dot{r}_b}{r_b} - \frac{\dot{A}}{A} - \frac{\dot{B}}{B}
\]

The \( w \)'s are the income share weights for output the \( q \)'s are the income share weights for resources and the \( v \)'s are the income share weights for the stocks of knowledge in the economy. The first thing to note is that is equilibrium all of the \( r \)'s have zero growth. So the TFP calculation reduces to:

\[
\frac{\text{TFP}}{\text{TFP}} = w_c \frac{\dot{c}}{c} + w_a \frac{\dot{a}}{a} + w_b \frac{\dot{b}}{b} - q \frac{\dot{r}_c}{r_c} - q \frac{\dot{r}_a}{r_a} - q \frac{\dot{r}_b}{r_b} - \frac{\dot{A}}{A} - \frac{\dot{B}}{B}
\]

The next thing to note is that \( \frac{\dot{c}}{c} = \frac{\dot{a}}{a} = \frac{\dot{b}}{b} = \frac{\dot{A}}{A} = \frac{\dot{B}}{B} \) and \( \sum_i w_i = 1, \sum_j v_j = 1 \). Therefore, \( \frac{\text{TFP}}{\text{TFP}} = 0 \).
The implication is that sustained growth is possible where there are no super-normal gains, no externalities and no increasing returns to accumulating technological knowledge. All that is required is complementarity among the different types of technological knowledge in the different production activities. TFP being equal to zero is not evidence of either zero growth or zero technological progress.

III. CONCLUSIONS
Here is what we conclude from our present study—the first three points relating to the quotations given at the outset of the paper.

1. All improvements in technology, such as the internal combustion engine, do not “clearly raise TFP.”
2. TFP cannot simultaneously measure all technological change and just the super normal benefits and/or the free lunches associated with technological change.
3. TFP does not measure “prospects for longer term increases in output” since, among other reasons, new GPTs tend to be associated with up-front costs and downstream benefits, and there is no reason why the TFPs associated with successive GPTs should stand in any particular relation to each other.
4. Given no other problems, TFP does not measure all technological change that is embodied in capital goods. It only measures the returns in investing in such embodied technologies that are in excess of the return on investing in existing technologies. These are the super normal gains of technological change.
5. All TFP calculations must use inputs that are to some extent aggregated from the full set of micro inputs. When aggregation uses values, as it must do for capital and may do for other inputs, technological change that goes on at levels below that at which the TFP index is calculated will show up as changes in the measured quantity of inputs. They will not, therefore, be measured as changes in TFP.
6. When the system is in transition between two competitive equilibria, the expenditure weights used in the calculation of a TFP index will cause biases in the estimated TFP changes. In cases of constant or decreasing returns to scale in production functions, changes in technology will be measured as changes in the quantity of inputs because they alter the expenditure weights that are used in any index of TFP changes. This is also true for some cases of increasing returns to scale. Thus, the super normal gains will be under estimated. For sufficiently strong increasing returns to scale the direction of the bias reverses and Basu and Fernald’s result obtains.
7. TFP does not adequately capture the effects on either the growth process or on consumers surplus of those technological changes that operate by lowering the cost of small industries, causing large increases in future sales and outputs.
8. Increases in the efficiency of exploitation of unmeasured resource inputs can be mismeasured as increases in capital. This may be a particularly serious problem when some resources are not explicitly included as inputs, as is often the case when TFP is calculated at high levels of aggregation.
9. There is reason to suspect that TFP does not adequately reflect the increase in a firm’s capital value created by R&D activities that are realized through sale of intellectual property. Yet these are often technological advances created by the use of valuable resources.
10. Since all of the biases that we have been able to identify (points 5-9 above) have the effects of increasing the measured amount of inputs and hence reducing the value of TFP, we conjecture, but have not proved, that TFP measurements will seriously overstate the increase in inputs and seriously underestimate the super normal benefits created by technological change.
11. Super normal benefits and/or free lunches are only a small subset of the technological complementarities that are the main benefits of technological change and hence, even if it
measured super normal gains perfectly, TFP would not measure the full spillover effects of technological change.

12. It follows from points 4 to 9 that low TFP numbers for the Asian Tigers do not mean they are in the same boat as was communist Russia; their numbers are quite compatible with successful technology enhancing policies and the technological transformation of a country through domestically generated or imported capital.

13. The effects of technical change is best measured conceptually by a counter factual: the accumulated difference over any time period between the actual return on investment and what that return would have been if the technical changes in question had not occurred.

14. It is possible to demonstrate the growth sustaining power of technological change (of GPTs in particular) in a model where TFP will be zero yet endogenous growth will continue indefinitely. The important property of such a model is that knowledge is complementary with other knowledge and continually opens opportunities for further knowledge accumulation. Externalities, scale effects or other things that create super-normal gains are not necessary.

Increases in total factor productivity are correctly described as measuring the difference between increases in measured outputs and increases in measured inputs. They are correctly interpreted as being an imperfect measure of the super normal gains associated with technological change, not all of which are mere Manna from heaven. The degree of confusion surrounding the interpretation of TFP measures should, however, give us pause when they are confidently used to support assertions about the arrival or non-arrival of the New Economy or about technological change more generally. Confusion over the interpretation of TFP measurements would be much reduced if each time one was reported, it carried the caveat: there is no reason to believe that changes in TFP in any way measure technological change.
APPENDIX

To further study the issue of embodied technological change, we assume two industries. The final goods industry produces a constant amount of some final good, $Y$, whose specifications never change. Its inputs are labour and the services of a capital good that takes the form of a machine delivering an unchanged service. The final goods industry has the following Cobb Douglas production function:

$$(3.1) \quad Y = A(mK)^a(nL)^{1-a},$$

where $K$ is capital measured in value of fully employed machines and $L$ is labour measured in physical units, while $m$ and $n$ are efficiency parameters attached to capital and labour respectively.

The machines are produced by a capital goods industry. The producers spend money on broadly defined development costs to create new technologies that alter the efficiency of the machine they produce. New machines are sold by their producers at a price that is just sufficient to recover the direct costs of production and the full development costs — the price of the improved capital good capitalizes the development cost that created it. Thus, the capital good producers will register no change in their TFP since their costs change in the same proportion as the value of their output and any change in TFP will occur in the user industry. 42 We consider four different types of technological change. Whenever possible, we assume that the development cost of the machine, and hence the rise in its price, is just equal to the saving in inputs that it generates so that there is also no TFP gain in the consumer goods industry.

**Case 1, a disembodied technical change in the final goods industry lowers both labour and capital costs by x%**. 43 Consider two polar cases. First, at the least costly extreme, let the organisational change be costless; it is an isolated stroke of genius with virtually no development costs. Now the value of output is unchanged while the physical quantities of both inputs fall by $x\%$. The industry’s TFP will rise by $x\%$.

Second, let the industry’s organisational change have positive development costs. These are incurred in period zero and, to set up the most costly extreme case, assume them to be equal to the discounted present value of the cost savings, which begin to accrue at period one. This makes the total development cost equal to

$$\sum_{j=1}^{n} \frac{r}{(1+i)^j},$$

where $i$ is the return to investing in the technology when it is organized in the original way (or the opportunity cost of R&D), $r$ is the old cost minus the new cost per period (or the net cost savings per period) 44 and $n$ is the number of periods over which the new organization is expected to be useful. Because R&D is treated in the national accounts as a current cost with no parallel output, there will be a reduction in the level of TFP when the R&D is being conducted (equal to the present value of the future stream of benefits from R&D). There will be a subsequent increase in the level of TFP from its pre-R&D level, starting when the new organization is in place and extending over its lifetime. (The extra TFP will be $r$ for $n$ periods.) Over the lifetime of the new organization, the extra TFP will just

---

42 We make this assumption purely for heuristic reasons. If we altered the example so that some or all of the gain accrued to the machine producers, the only conclusion that would change would be the location of the TFP gains.

43 The famous productivity-increasing reorganization of the factory that occurred when (unit drive) electric motors succeeded (central drive) steam engines in providing factory power provides an example. Although the change was embodied in a new organization of the machine tools on the factory floor so that the flow of goods through the factory could be continuous, there was no need to alter the amount of physical capital or labour inputs. So with unchanged measured inputs, output rose..

44 $r$ is related to $x$ in the following way: $r = \text{old} - \text{new}, x = (r/\text{old}) \times 100\%$. 
make up for the loss of TFP while the R&D was being conducted (everything appropriately discounted).

Now assume that these disembodied changes are going on continually. If they are costless, there will be a continuous rise in the TFP of the consumers’ goods industry. If they are costly, so that at each point in time there is a new expenditure on R&D and an accompanying benefit from the past R&D expenditures still generating increases in output, TFP will be constant with no blips as in the once-for-all case.

What these polar cases show is that, contrary to what is often stated in the literature, disembodied technical change does not necessarily raise TFP. What matters is how the discounted present value of the gains due to the fall in direct costs per dollar’s worth of output compares with the development cost of creating the disembodied change. We suspect that the contrary presumption in the literature is due to an implicit, and invalid, assumption that all disembodied changes are costless.

Case 2, a new embodied technology saves equally on both labour and capital costs: This is the common case in which the capital goods industry develops a new machine that is absolutely saving on both labour and capital. Lean production is one of many examples. (See Womack, et al (1990).)

Since we are assuming that the new technology saves equally on both labour and capital, we cannot assume a development cost that would leave TFP unchanged. An unchanged TFP would require an increase in the cost of the machine, which would violate the requirement that both costs fall in equal proportion. So we assume that the cost of producing the new machine is $2x\%$ less than the cost of producing the old one. The development costs are just covered when the new machine is sold for $x\%$ less than the old machine. The new machine is also assumed to use $x\%$ less labour to produce the same amount of output as the old machine. Thus, with constant output, inputs of capital and labour both fall by $x\%$ so that the industry’s TFP rises by $x\%$. This innovation shows up in equation (3.1) above as an increase in both of the efficiency parameters, $m$ and $n$, by $x\%$.

Notice that in equation (3.1), as in any CRS production function, this change is indistinguishable from an increase in the parameter $A$ by $x\%$. So this type of embodied technological change looks empirically like disembodied change. Note also, that the mixed case in which one efficiency parameter rises by $x\%$ and the other by $y\%$ ($x > y$), is indistinguishable in the Cobb-Douglas function from an overall increase in $A$ by $y\%$ combined with an increase of $(x - y)\%$ in the efficiency parameter that actually increases by $x\%$.

Case 3, the new embodied technology saves only on labour inputs: The new machine is assumed to produce $x\%$ more output using the same amount of labour, a type of invention that is common in the history of technological change. We assume that the development cost of the new machine can just be recovered by the capital goods industry if the machine sells for $x\%$ more than the old machine. When the consumers’ goods industry replaces the old machines with new ones, this uses savings equal to the additional value of the machines. So the industry’s capital stock will grow by $y\%$. It will have $y\%$ more measured capital and $x\%$ more output. To get the limiting case in which TFP is zero, $y$ must equal $x/\alpha$.  

---

45 Let the machine reduce the number of labour and machine hours required to produce a unit of output by $x\%$. If the cost of labour and machines were unchanged, TFP would rise by $x\%$. For TFP to be unchanged the cost of a machine hour would have to rise by $x^1/\alpha$ which violates the assumption that capital and labour costs fall by the same amount.

46 The same statement is also true for some specific parameters in the translog function but not for all.

47 For example, if capital costs are 25% of total costs, then the quantity of capital must increase by four times the increase in output if TFP is to be zero. If development costs would cause a greater increase, the innovation will not be made by foresighted individuals. If it is less than that, there will be some surplus over development costs and some increase in TFP.
Case 4, the embodied technical change saves only on capital inputs: Now assume that the new machines act as if there were $x\%$ more of them than the old machines, making output increase by $\alpha x$. Machines increase in efficiency but labour does not. The efficiency parameter on machines grows by $x\%$ and, by assumption, the sale price of each physical machine rises by just enough to cover development costs of $y\%$. To get the limiting case of no TFP change we assume that the development costs are such that $y = x$. (The percentage change in output minus the weighted percentage change in capital, is $\alpha x - \alpha x = 0$.) If development costs exceed this amount, foresighted firms will not develop the machine; if they are less, there will be a TFP gain.

In the discrete once-for-all cases of 2, 3 and 4, there is a fall in TFP when the development costs are incurred and then a subsequent increase in TFP. In the limiting cases of 3 and 4 the present discounted value of the TFP gain is equal to the R&D cost. This blip can be eliminated in either of two ways. First, the R&D can be capitalized and treated as a capital investment in the first period. Then it can be depreciated over the life time of the new machines. Thus in the first period there will be an output of R&D capital to match the loss of other output and no change in TFP while in subsequent periods the slightly higher output will be matched by the depreciation of R&D capital, again leaving TFP unchanged throughout the whole time. Second, technological improvements may continue period by period. Now there is a constant amount of R&D each period and there is a fall in costs in the consumer's goods industry of the same value each period, leaving TFP unchanged.

Case 5, no technical change: The assumption that development costs could just be covered by the increased price of the new machine in cases 3 and 4 implied a certain amount of capital investment in the new machine. We now calculate the increase in output that would have occurred if that amount of investment had been made when there was no technological change. In case 3, the measured amount of capital increases by $x/\alpha\%$ and output by $x\%$, giving no change in TFP. Now let capital increase by $x/\alpha$ in equation (3.1) with no technical change. Output then increases by $x\%$. So the results are the same. In case 4, capital increased by $x\%$ and output by $\alpha x\%$. With constant technology, an $x\%$ increase in $K$ causes output to increase by $\alpha x\%$. So, once again, the results of case 4 with development cost just covered are the same as when an equivalent amount of investment occurs with static technology.

We have now reached a number of conclusions.

- The TFP measures in each of the first four cases correctly reflect the net increase in output due to technical change — i.e., the difference between what the industry’s output would be with and without technical change for a given amount of investment.
- This TFP measure is not, however, a measure of technical change per se but only of returns over and above those needed to recover the broadly defined development costs that created the innovation. Thus, zero TFP does not mean zero technical change, only that, at the margin, investing in R&D has the same effect on output as investing in existing technologies (investment with no technical change).
- In theoretical treatments, the distinction between embodied and disembodied technical change is often made in terms of the parameter of the production function that is affected, a shift in $A$ being disembodied technical change, and a shift in either $m$ or $n$ (in equation 3.1) embodied change. However, we cannot distinguish empirically, using an aggregate Cobb Douglas production function, between genuine disembodied technical change and technological change that is embodied in a new machine that lowers all input costs in equal proportion (a common case). Both appear in the observed function as Hicks neutral growth. The empirical observation of Hicks neutral growth does not, therefore, imply that technological change is not embodied in new capital equipment. Also, even when technical change is biased towards saving one factor more than the other, this is indistinguishable in the Cobb Douglas function (but not necessarily in a translog function) from a case of Hicks neutral technical change of an
amount equal to the contribution of the factor that is saved in the smaller proportion, combined with a change that saves a smaller amount on the other factor.

- Chen (1997) argues that the size of the TFP number is affected by whether technology is treated as disembodied or embodied, the former always yielding the higher number. In contrast, we argue that it is not whether the technology is embodied that matters but whether or not the gains from the technological change more than cover its development costs. We suspect that Chen is stating a commonly held view and hence that ours is an important point.

Jorgenson and Griliches argued that TFP should be small or zero because they associated free lunches with externalities and expected these to be small. As soon as we recognize that the returns to innovation when oligopolistic firms are competing in innovations can be, and often are, large, there is no contradiction in holding that TFP measures only the free lunches and that TFP numbers can be quite large.

End of Appendix

---

48 Specifically, let $0 < \frac{\Delta m_a}{m_a} < \frac{\Delta n_a}{n_a}$, where the subscripts $a$ and $c$ stand for actual and calculated (or measured). These values are indistinguishable empirically from $\frac{\Delta A_c}{A_c} = \frac{\Delta m_a}{m_a}$, and $\frac{\Delta n_a}{n_a} = \frac{\Delta n_a}{n_a} - \frac{\Delta m_a}{m_a}$. 
BIBLIOGRAPHY


Carlaw, Kenneth and Steven Kosempel (2001 Manuscript): “The Sources of Productivity Growth in Canada”.


IMF Staff (2000): *Country Report No. 00/34*.


Statistics Canada (1998), 13-568
FIGURE 1

Marginal product of capital

TIME

MP₁

MP₂