Total factor productivity and the measurement of technological change

Richard G. Lipsey *Department of Economics, Simon Fraser University*
Kenneth I. Carlaw *Department of Economics, University of Canterbury*

Abstract. TFP is interpreted in the literature in different, mutually contradictory ways. Changes in TFP are shown to measure not technological change, only the super-normal returns to investing in such change – returns that exceed the full opportunity cost of the activity. Thus, in the limit, technological change can proceed with unchanged TFP. Measuring the effects of technological change instead requires counterfactual estimates. Reasons why changes in TFP are imperfect measures of super normal returns are also studied – reasons connected with the timing of output responses, the treatment of R&D in the national accounts, the omission of resource inputs, and two types of aggregation.

La productivité totale des facteurs et la mesure du changement technologique. La productivité totale des facteurs (PTF) est interprétée dans la littérature de manières différentes et mutuellement contradictoires. On montre que les changements dans la PTF ne mesurent pas le changement technologique, mais seulement les rendements supra-normaux sur les investissements dans de tels changements – des rendements qui dépassent le plein coût d'opportunité de cette activité. Donc, à la limite, le changement technologique peut procéder sans que la PTF change. Mesurer les effets du changement technique réclame des évaluations d'hypothèses de rechange. Les raisons pour lesquelles les changements dans la PTF sont des mesures imparfaites des rendements supra-normaux sont aussi étudiées – raisons connectées avec la réponse des niveaux de

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 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).
Email: rlipsey@sfu.ca; kenneth.carlaw@canterbury.ac.nz

Canadian Journal of Economics / Revue canadienne d'Economique, Vol. 37, No. 4
November / novembre 2004. Printed in Canada / Imprimé au Canada

0008-4085 / 04 / 1118–1150 / © Canadian Economics Association

production, le traitement de la R&D dans les comptes nationaux, l'omission des intrants de ressources, et deux types d'agrégation.

Economic historians and students of technology agree that technological change is *the* major determinant of very long-term economic growth.[1] If we knew no more than the Mesopotamians, or Medieval Europeans, our living standards would not be far above theirs, a bit more, owing to things such as capital accumulation but not much. Yet over shorter periods of time, there is debate over what proportion of economic growth is due to technological change and what to other forces, such as the accumulation of physical and human capital. Such debates imply that we are able to separate the effects of technological change from those of the other determinants.

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 total factor productivity (TFP) approach to measuring changes in technology. We argue that, contrary to widely held views, changes in total (or multi-) factor productivity do not measure technological change. Ideally, they measure only the super normal gains associated with such changes. Next, we discuss the counter-factual measures that are needed. Finally, since TFP is commonly used for many different purposes in theory, history, and policy, we look at some problems concerning its measurement and interpretation.[2]

1. Measuring technological change through TFP

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.

'A growth-accounting exercise [conducted by Alwyn Young] produces the startling result that Singapore showed no technical progress at all.' Krugman

1 Lipsey has argued this in several publications, for example, Lipsey (1992, 1993 and 1994). Of course, technological change and investment are interrelated, the latter being the main vehicle by which the former enters the production process.

2 Another important reason for concentrating on TFP is that low TFP growth rates have been used by several authors to express scepticism about the existence and importance of the New Economy brought about by the ICT revolution. (See, e.g., Gordon 2000.) See Lipsey and Bekar (1995) for an early statement of why we do not accept this argument and Lipsey (2002) for a recent one.

(1996, 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, 50)

'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, 50–1, italics added)

'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, 119)

'It is clear that British capabilities for the transfer and improvement of technology were strong and improving during the first industrial revolution, and this no doubt was central to the (otherwise surprising) steady acceleration in TFP growth.' Crafts (1966, 200)

'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, 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, 54)

'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 off by research projects, or it may simply reflect inspiration and ingenuity.' Hulten (2000, 61)[3]

'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)

'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'

3 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.

measures of technical progress. Abramovitz labelled the resulting index "a measure of our ignorance."' Griliches (1995, 5–6) quoting Abramovitz (1956, 8)

These quotations illustrate three main positions. One group holds the view that changes in TFP measure the rate of technological change – Krugman, Young, Crafts,[4] Statistics Canada, Barro. The second group holds that TFP change measures only the 'free lunches' of technological change, which they argue are mainly associated with externalities and scale effects – Jorgenson and Griliches in (5) and Hulten. The third group is sceptical that TFP measures anything useful – Abramovitz and Griliches in (7&8)[5]

Surely it is close to a scandal that a measurement relied on so widely is so variously interpreted. Although our position is close to the 'free lunch view,' we argue that there are important ambiguities surrounding that concept. Also, we do not accept that TFP growth should be close to zero, as Jorgensen and Griliches argued in their classic 1967 article. We hope in this article to go some way towards responding to Prescott's (1998) call for a theory of TFP. He implicitly agrees with our position by arguing that the sources of international TFP differences are more than just differences in employed technologies. We would, however, add to his list of other sources.[6]

1.1. Two measures of TFP

We start with a brief survey of the two most common methods of measuring TFP.

1.1.1. The growth accounting method

The growth accounting or residual method of TFP measurement originates in Solow's 1956 and 1957 articles. A production function is used to relate measured inputs to measured output. Any output growth not associated statistically with the growth in measured inputs is assumed to result from technological change (and other causes such as scale economies). Critical to this approach is the concept of a production function that is valid at whatever level of aggregation the calculations are to be made. This poses two sets of conceptual problems.

4 In the major debate among economic historians regarding of the timing of the Industrial Revolution, the behaviour of measures of TFP has often been used as a measure of the timing. (See, e.g., Crafts and Harley 1992; Berg and Hudson 1994.)

5 Although we give only two 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 1950s, attacked the validity of the concept of an aggregate production function. See also Metcalf (1987, 619–20).

6 Prescott suggests that international differences are explained by resistance of special interests to the adoption and efficient use of technologies currently used elsewhere. Studies such as Pack and Westfall (1986), Westphal (1990), and Wade (1990) provide evidence related to the many other reasons why TFPs differ among countries.

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

$$Y = AF(K,H,L), \tag{1}$$

where Y is output, K is physical capital, H is human capital, and L labour. To calculate TFP over long periods of time by production function based methods, we must make the 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 redesign of the factory layout. We must also assume that we can measure factor inputs over these long periods. In what units, for example, should we compare the amount of capital invested in a Victorian, steam-driven, manually controlled factory making stage coaches with that in an electrically powered, largely robot-controlled, modern factory making diesel electric passenger trains?[7]

A second set of problems concerns 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 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, in the Austrian tradition competition is seen 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 they be aggregated from a set of markets that contain the mixture of 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 oligopoly matters. To make contact with the existing literature, we will proceed as if the aggregate production function is a meaningful concept over the time period in which we are interested.

7 Jorgenson, Gollop, and Fraumeni (1987) attempt to control for some of these problems by accounting for changes in things such as input quality, legal forms of organizations, capital asset classes, and sectoral substitution.

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, 255–6).

To define TFP, we use the Cobb-Douglas version of the aggregate function[9]

$$Y = AL^{\alpha}K^{\beta},\ \alpha + \beta = 1,\ \text{and}\ (\alpha,\beta) \in (0,1), \tag{2}$$

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 utilization 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 utilization 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):

$$\frac{\dot{A}}{A} = \frac{\dot{Y}}{Y} - \alpha\frac{\dot{L}}{L} - \beta\frac{\dot{K}}{K} = \frac{T\dot{F}P}{TFP}. \tag{3}$$

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.[10]

9 Many TFP studies use the more flexible translog function. Also, Jorgenson (2001) uses a production possibility frontier approach rather than an aggregate production function. But we lose little at our conceptual level by using the Cobb-Douglas formulation.

10 This aggregate production function approach is a simplification of a much broader concept of the aggregate production function that allows for the resource consuming activities of R&D. Examples include Jorgenson and Griliches (1967), Jorgenson (2001), and Barro (1999), whose approaches involve intermediate production functions, or a meta production possibilities frontier. Two critical features of these approaches are their treatment of R&D as an input and the returns to scale in production functions (or overall production possibilities set). Jorgenson and Griliches (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 changes in TFP. In contrast, Barro (1999) uses production functions that allow R&D to generate expanding product variety or quality with increasing returns to the intermediate R&D inputs. In Barro's case, because of the increasing returns to the intermediate R&D input, there is a Hicks-neutral, 'manna from heaven' component of technological change that is measured by changes in TFP and a component of the endogenous technological change generated from costly R&D that is not. All of this leaves open the questions about the meta or all-encompassing notion of the aggregate production function and about the appropriate formulation of R&D and knowledge production.

Many economists have identified problems, of both concept and measurement, associated with growth accounting. Key references are Griliches (1987, 1994, and 1995), where he considers many sources of error in TFP measurements.[11] We do not review most of these issues because the problems are well understood.

1.1.2. The index number method

The index number method is an extension of, and complement to, growth accounting. Both use indexes and involve similar problems. However, the index number approach does not require an aggregate production function, though an appropriate index can be selected via the economic approach for some specified production function. Index number theory provides explanations for what is and is not possible in aggregation and acts as a check on some measurement problems in accounting. The method is to divide an output index by an input index: $A_t = Y_t/I_t$, where A_t is a level measure of total factor productivity, Y_t is an index of real output, and I_t is an index of the factor quantities used in production. This is a straightforward calculation, given the indexes of output and input.

Two approaches, the economic and the axiomatic, are commonly used for selecting from the many different index numbers that can be used in his type of measurement and that are surveyed by Diewert (1987) and Diewert and Lawrence (1999). In the economic approach, particular production functional forms can be linked to particular indexes. 'For example, the Törnqvist index used extensively in past TFP studies can be derived assuming the underlying production function has the translog form and assuming producers are price taking revenue maximizers and price taking cost minimizers' (Diewert and Lawrence 1999, 9). In contrast, the axiomatic approach compares the properties of the index number formulations with 'desirable properties,' and the index number that has the largest number of desirable properties is then used to calculate TFP.

One application of the index number approach is distance function analysis (DEA), which makes the strong claim of being able to separate TFP into two parts, one due to increased efficiency in resource use and one due to technological change. (See, for example, Fare et al. 1994 and Fare, Grosskopf, and Margaritis 1996.) The method uses a Malmquist index and compares ratios of outputs with inputs (TFP) across units. It requires the assumption that all the units being compared, which may be firms, industries, or whole countries, have identical production functions. Although this heroic assumption might be

11 Griliches (1987, 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 inappropriate data, (6) weights for indices, (7) theoretical specifications of relations between inputs, technology, and aggregate production functions, and (8) aggregation over heterogeneity. Concerning point (6), Diewert (1987, 767–80) shows that very restrictive assumptions have to be satisfied to generate these indices of output and input.

applicable when one industry is being compared across similar countries, it is not credible when the comparison is made among different industries or even firms within one industry. If production functions are not the same (for evidence see Jorgenson, Gollop, and Fraumeni 1987), there is no reason why relative average products should relate to relative efficiencies, since efficiency requires equality of the marginal not average products of each resource in its various uses. Commenting on the study by Fare, Grosskopf, and Margaritis (1996), Diewert and Lawrence (1999) state: 'Evidently, they just used a variant of DEA analysis, *assuming that the value added outputs of each industry can be produced by every other industry*. This seems to be a rather untenable assumption to say the least and hence we suspect that their measures of efficiency change and technical progress are essentially worthless.'

1.2. TFP and costly technological change

We argue that changes in TFP do not measure technological changes but do, ideally at least, measure the associated super-normal profits, externalities and 'free lunches'.[12] Virtually all technological change is embodied in one form or another: new or improved products, 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 on process and organizational technologies. For concreteness, we focus on capital goods although any embodied technology would do.

Although much theory proceeds as if these changes appear spontaneously, most of them result from resource-using activities. The costs involved in creating these technological changes are more than just conventional R&D costs. 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.'

Jorgenson and Griliches (1967) made a path-breaking contribution when they argued that changes in TFP would measure only the gains in output that were over and above the development costs of the technological advances that brought them about. Unfortunately, because they argued that these gains would, when properly measured, be close to zero and because Jorgenson subsequently spent a good deal of time trying to produce this zero result empirically, the debate that followed the publication of their paper centred on whether the measure should be zero, obscuring their important point that changes in TFP did not measure technological change.[13]

12 Others who have argued positions include Nelson (1964), Jorgenson and Griliches (1967), Rymes (1971), who discusses the need to measure technological change in a dynamic framework, and Hulten (1979 and 2000).

13 Jorgenson subsequently changed his view that TFP should be equal to zero and has developed a number of refinements for measuring TFP. (See, e.g., Jorgenson, Gollop, and Fraumeni 1987; Jorgenson and Stiroh 2000.)

1.2.1. 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 implies that the price of the capital 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, $T\dot{F}P < 0$. (2) If $w = v$, costs are just covered, the rise in the cost of the machine just equals the value of the new output, and $T\dot{F}P = 0$. (3) If $w < v$, profits are made and $T\dot{F}P > 0$. 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 technological change per se but only the profits that it produces (as well as some free lunch externalities). Thus, zero change in TFP does not mean zero technological change. It means only that investing in R&D has had the same marginal effect on income as investing in existing technologies (investment with no technological change) and that there are no external effects that show up in current increases in output elsewhere in the economy without corresponding current increases in inputs.[14]

If the marginal productivities of investing in new and existing technologies are the same, the new technology might seem to confer no benefit. In section 2, below, and in more detail in Carlaw and Lipsey (2002a), 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.

1.2.2. Human capital

Now consider the accumulation of human capital. 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

14 The study paper version of this paper (available from http://sfu.ca/~rlipsey) contains an appendix in which this case is investigated in much more detail.

knowledge that was available in, say, 1900. 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 such as 1900. Holding population constant, the difference between the total contribution of human capital and the technology-constant contribution 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 the effects on output of accumulating more human capital from those of acquiring new technological knowledge. 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 increases in TFP. Whether or not we do this is largely a matter of taste and convenience. But if we do, we are not justified in concluding that technological change accounts for little of the observed increases in outputs just because increases in inputs, including increases in measured human capital, explain most of the growth in output statistically.

Similar problems arise in disentangling the influence of pure human capital from the influence of the technological knowledge embodied in it when dealing with cross-sectional data. To make useful cross-section comparisons, we must understand not only *how much* is known, but also *what* is known. For example, six years of schooling in Marxist philosophy and the sayings of Chairman Mao would produce less valuable human capital than six years' studying the 'three Rs.'[15]

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. Thus, they cause understatements of the contribution of technological change to economic growth.[16]

15 Pomeranz (2000) argues that the eighteenth-century Chinese had a literacy level similar to that of Europe. But a quantity of human capital equal to that of the Europeans does not imply that the Chinese were on the verge of the Industrial Revolution. Bekar and Lipsey (forthcoming) argue that English human capital provided the knowledge needed for 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 differences in the input of human capital *do not* demonstrate that *A*'s technology enhancing policies are ineffective. (Arguments that they do are common in the development literature.)

1.2.3. Disembodied technological change

None of the above conclusions would be altered if the technological changes were disembodied, because all that matters for changes in TFP is whether there is an increase in inputs to offset any observed increase in outputs. Thus, contrary to what is often stated in the literature, disembodied technological change does not necessarily raise TFP. We suspect that the presumption in the literature that it does is due to an often-made implicit, assumption that all disembodied changes are costless, which, of course, they are not. For example, the reorganization of the factory that followed the introduction of the unit drive electric motor was not embodied in any physical capital, yet it entailed some heavy development and learning costs, which, in the limit, could have been equal to the resulting gain in productivity thus reducing measured TFP growth to zero. This is a limiting case, but it illustrates that, just as with embodied changes, it is the margin of increases in outputs over increases in inputs that matters, not the nature of the technological change itself.

Another possible source of confusion is concentration in the theoretical literature on embodied technological changes that are Harrod neutral rather than on the common case in which a new machine is absolutely saving on both labour and capital. Lean production is one of many examples. (See Womack, Jones, and Roos 1990.) Hicks-neutral *embodied* change that raises the efficiency of labour and capital in equal proportions is analytically indistinguishable from Hicks-neutral *disembodied* technological change that raises the efficiency parameter A in our production function.[17]

1.2.4. 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 third parties. This does not capture the full range of benefits that TFP measures. 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 arise only 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 characterizes real-world technological change in which new technologies are developed under conditions of Knightian uncertainty, with knowledge at least partially appropriable by those who create it, investments in new technology can, and often do, yield

17 More generally, an embodied technological change that raises the efficiency of K by $x\%$ and L by $y\%$ ($y < x$) in equation (2) is empirically indistinguishable from a change that raises A by $y\%$ and K by $(x - y)\%$.

returns well above the going rate of return in the rest of the economy. An entrepreneur who received a return of say 30% on innovating an unproven new technology when the normal return on investing in exiting technology was 15% would be surprised 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 major 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 embodying existing technology as the firm's 'super normal profits.'

To allow for these returns to uncertainty, as well as for genuine free lunches, we define the 'super normal benefits' of technological change 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, the full super normal benefits could be appropriated by the developers, yielding super normal benefits without externalities. If the developers cannot appropriate all the gains for themselves, some of the super normal benefits may become externalities. Furthermore, because of technological complementarities, major innovations in one sector create opportunities for profitable, resource-using innovations elsewhere that do not show up as conventional externalities.[18] If exploiting these opportunities yields a return above full costs, they are also external benefits.

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. If one understands that TFP includes part of the return on innovation, there is no problem in calling it 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 easy to misinterpret what changes in TFP do measure, once it is accepted that they are not a measure of technological change. For example, Hulten writes that the Hicksian shift parameter, A_t, '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_t' (2000, 9n.5). In contrast, we argue that the (often large) part of the gains from R&D that is in excess of the 'normal' rate and that

18 Carlaw and Lipsey (2002a) define a class of spillovers arising from new technologies that is much broader that conventionally defined externalities. They call these 'technological complementarities' and define them as arising 'in any situation in which the past or present decisions of the initiating agents with respect to their own technologies affect the value of the receiving agents' existing technologies and/or their opportunities for making further technological changes.'

accrues to the innovators for undertaking uncertain investment, and without which invention and innovation would not occur, is so captured.

2. Counter factual measures

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

2.1. The conceptualization of technological change

One big problem in measuring the effects of technological change is to separate them from the capital accumulation that embodies much of them. In practice, this is difficult, possibly impossible. 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. (1) Feeding 6 billion people with the agricultural technologies of 1900 would have been impossible.[19] (2) Pollution would have become a massive problem. (3) Since most new technologies save on all inputs – a process that Gubler (1998,240) calls 'dematerialization' – to produce the value of today's output with 1900 technologies would have required vastly more resources than are currently being used, thus exhausting many of them. 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. (4) 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.

While this is, of course, speculation, it shows that growth of labour and capital at twentieth-century rates with truly constant technology would have produced massive problems.

19 Since population is endogenous, it is not clear how much population would have increased if food-producing technologies had remained frozen at 1900 levels.

2.2. GPT-driven growth[20]

We view economic growth as being driven by a succession of general purpose technologies (GPTs) that create opportunities for profitable investments in new product, process, and organizational technologies.[21] The opportunities come from the technological complementarities created by radically new technologies and would not have existed without the GPT.

The evolution of major new technologies, particularly GPTs, prevents 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). 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.[22] 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.

These considerations suggest two important conclusions. First, because the concept of the super normal gains measured by changes in TFP is based on what is happening to current output and current costs, it does not 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 changes. Second, the economic benefit of new technologies is in

20 This section draws on Carlaw and Lipsey (2002, and forthcoming).

21 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.)

22 Freeman and Louca (2001) provide an analysis of this phenomenon built around the concept of technoeconomic paradigms rather than GPTs.

the *future path of returns* rather than in gains on the current margin between the new and the old technologies. New technologies in general, and new GPTs in particular, sustain the growth process, although they do not necessarily accelerate it.

2.3. The counterfactual measurement of technological change

The important message is that there need be no observable impact of a new technology on current 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.[23]

This phenomenon is illustrated in figure 1, which gives two time paths for the return on capital. The first is constant along the arrowed curve MP_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. The second curve, MP_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.[24] The gain from technological change is measured by the gap between MP_1 and MP_2.[25]

So even if there are no super normal gains from new technology, even if all the innovations enabled by each new GPT just covered their full development costs, and even if measured TFP were constant, economic growth would still be sustained by a succession of GPTs that held the return on investment in embodied technologies above what it would have been if technology had remained static.[26]

23 Several authors are currently investigating the impact 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, 2000) and IMF (2000). This research attempts to measure directly the new technology that is embodied in capital goods. Some findings for Canada by Carlaw and Kosempel (2004) show that in the period between 1974 and 1996, while the rate of TFP growth was declining, investment-specific technological change was growing. This research provides a proximate measure of technological change taken from independently measured data.

24 The argument does not depend on the GPT concept, only that holding all technological knowledge constant would lead to a negatively sloped MP_2 curve.

25 This type of historical counterfactual is not to be confused with the limited counterfactual commonly used in cleometrics and effectively criticized in chapter 1 of Freeman and Louca (2002). Here, we are considering a world with and without all new GTPs and all the innovations that they enable.

26 A formal model of GPT-driven growth in which sustained growth proceeds indefinitely with zero TFP growth is discussed in Carlaw and Lipsey (forthcoming).

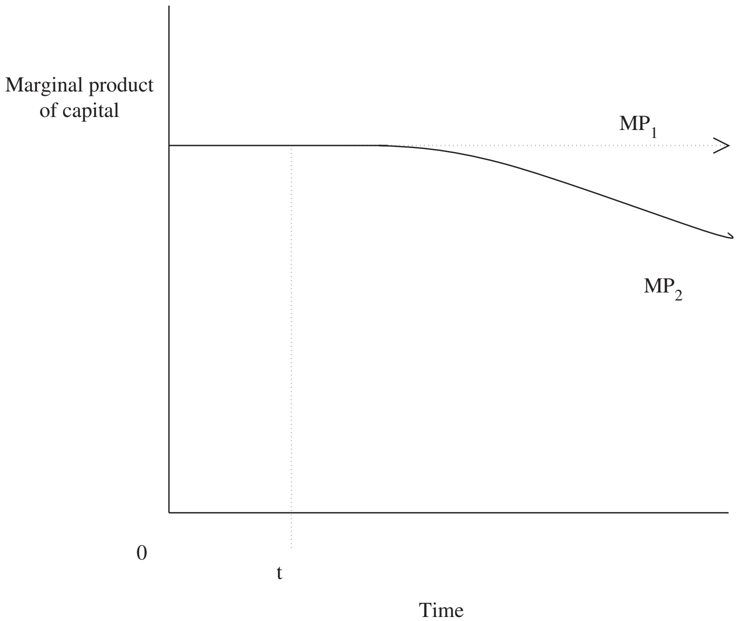

FIGURE 1

3. How well do TFP changes measure the super normal benefits of technological change?

If TFP growth is a measure of the super normal benefits associated with technological change, we may ask how well it measures them. To isolate our analysis from the issues discussed in section 1, we assume that all technological changes considered in this section occur costlessly. Thus in an ideal measure, all of the output changes induced by changes in technology should show up as changes in TFP.

3.1. Timing of quantity responses

First, consider an example of the importance of the timing of the quantity response to a price reduction that is brought about by a technologically induced cost reduction. A product of unchanged quality costs \$4,000 in year 1 and \$540 in year 20 (an average cost fall of 20% per year), and 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 1, the unit costs fall by 10% per year, while sales rise at 20% each year, so that total costs of production and consumers' expenditure on the product rise at 10%. Thus, the industry's TFP will be rising at 10% 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 0.02, the final weight will be 0.131.[27] Thus, the contribution to the nation's TFP change will rise steadily from 0.2% in the first year to 1.25% in the final year, using a Törnqvist growth rate index.

In case 2, all of the cost reductions come in the first year and sales, *s*, expand at 20% each year. In the first year, the industry's total costs fall from (4,000) (*s*) to (540) (1.2*s*), making its index of total costs fall from 100 in the base year to 16.2 in the following year. With an index of output of 120, the TFP, calculated using the index number method, is 120/16.3 or 7.36 (while the TFP for the base period is unity by definition since both index numbers are set at 100 in the base year). If we take into account the change in the share weight implied by the change in sales and costs (i.e., it drops from 0.02 to 0.0033) and use a Törnqvist growth rate index, we get a TFP change of ln (7.36) (0.012) (100%) = 2.4% in the year of the free gift and a zero thereafter.

We have 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 1, the contribution to the increase in national TFP in the last two years, when only $1/10^{th}$ of the total change in productivity occurs, is approximately the same as the contribution in case 2 in the first year when *all* of the 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.[28]

This is a common 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 infrastructure of petroleum refining, distribution, roads, motels, and so on to be developed.

27 If value of sales increases at 10% per year, it is 7.4 times as large in 20 years (i.e., after 21 periods of compounding). Normalizing the first period's price times quantity at unity then implies that for a 0.02 share weight the rest of the economy must be 49 (i.e., 0.02 = 1/(1 + 49) in the denominator for the share weight, and thus the share weight in the last period is 7.4/(7.4 + 49) = 0.131.

28 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. (All data are from 'The Model T Ford Club of America' http://www.mtfca.com/encylo/fdprod.htm.)

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, U.S. rural electrification came in 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 consumer and producer 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 expands only slowly over years, and even decades, cannot be dismissed as a theoretical curiosity. In such cases, national 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.

3.2. 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 is 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.[29] 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.[30] Whatever else we may think about 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. By definition, not only will it show a negative contribution to TFP, but it will also 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 itself is sold to a foreign multinational.[31] If the patent or the firm is sold to another domestic

29 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.'

30 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.

31 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.

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 domestically 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. Many recognize these limitations, but, as our initial quotations show, not everyone does.

3.3. Omitted inputs: natural resources made explicit

Failure to measure any input can bias TFP measurements. 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, and so on.[32] Almost invariably, however, everything that is subsequently done 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.[33] The first is the prediction derived from the standard formulation in equation (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 resource inputs that would have to increase at the common rate include 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 in standard theoretical treatments and in 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, owing to new techniques.)

32 Solow (1957) warned about the bias caused by omitted variables in the measurement of the residual. Hulten (2000, 51) discussed the problem of omitted environmental variables, concluding that solving it 'is an impossibly large order to fill.'

33 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.

To illustrate some of the problems associated with the omission of natural resources, let the underlying production function be

$$Y = BK^{\alpha}(nR)^{1-\alpha} \quad \alpha \in (0,1), \tag{4}$$

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 $\dot{Y}/Y = \dot{B}/B + \alpha\dot{K}/K + (1-\alpha)(\dot{R}/R + \dot{n}/n)$.
Let those measuring TFP assume the production function to be

$$Y = AK, \tag{5}$$

so that

$$\frac{\dot{Y}}{Y} = \frac{\dot{A}}{A} + \frac{\dot{K}}{K},$$

and measured TFP becomes

$$\frac{T\dot{F}P_m}{TFP_m} = \frac{\dot{A}}{A} = \frac{\dot{Y}}{Y} - \frac{\dot{K}}{K} = \frac{\dot{B}}{B} + \alpha\frac{\dot{K}}{K} + (1-\alpha)\left(\frac{\dot{R}}{R} + \frac{\dot{n}}{n}\right) - \frac{\dot{K}}{K}. \tag{6}$$

We assume that we have the same measures of Y and K in (4) and (5) so that all that differs is the actual productivity coefficient B and its estimated value A.

First, let the growth of the capital stock and resource inputs measured in physical units be zero. It is clear from (6) that measured TFP then correctly picks up changes in the underlying productivity parameters, B and n. Next, let the only variable that is growing be the unmeasured resource inputs, R. In this case, (6) shows that TFP_m rises by the amount of the extra resource consumption and is, therefore, biased upwards. Finally, let the measured accumulating factors, K, be the only independent variable that is growing. Output, Y, is then growing at the fractional proportion α of the growth of K. Since the assumed equation (5) gives K a coefficient higher than α, the increase in output is expected to be more than the actual increase, causing TFP_m to be negative although technology is constant. In summary, increases in the use of unmeasured inputs (natural resources in our example) will bias measured TFP upwards, while growth in the measured accumulating factors will bias it downwards.

3.4. Aggregation of inputs[34]

Next, consider the problems of obtaining a measure of inputs for the production function at whatever the level of aggregation over which that

34 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.

function is defined. Let the firm's real microeconomic production function be

$$Y = BM^{\gamma}N^{\delta}P^{\varepsilon}R^{\eta} \quad (\gamma,\delta,\varepsilon,\eta) \in (0,1) \text{ and } \gamma + \delta + \varepsilon + \eta = 1, \tag{7}$$

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. Assume that the aggregate production function is used to calculate the firm's TFP is

$$Y' = AK^{\alpha}L^{\beta} \quad (\alpha,\beta) \in (0,1)\alpha + \beta = 1. \tag{8}$$

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 $K = p_m M + p_n N$, where prices are equal to marginal products:

$$K = (\gamma BM^{\gamma-1}N^{\delta}P^{\varepsilon}R^{\eta})M + (\delta BM^{\gamma}N^{\delta-1}P^{\varepsilon}R^{\eta})N = (\gamma + \delta)Y.$$

Similarly, let the firm's aggregate labour be

$$L = p_p P + p_r R, = (\varepsilon + \eta)Y.$$

Now let B in (7) change continuously through time with unchanged inputs of M, N, P, and R. Then $\dot{Y} = (dY/dB)\ (\dot{B}) = (M^{\gamma}N^{\delta}P^{\varepsilon}R^{\eta})\ (\dot{B})$. Any change in B now shows up as changes in the two aggregate inputs, K and L:

$$\dot{K} = (dK/dY)(dY/dB)(\dot{B}) = (\gamma + \delta)(dY/dB)(\dot{B}), \text{ and}$$
$$dL/dt = (dL/dY)(dY/dB)(dB/dt) = (\varepsilon + \eta)(dY/dB)(dB/dt).$$

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

$$\dot{Y}' = (dY'/dB)(\dot{B}) = (dY'/dK)(\dot{K}) + (dY'/dL)(\dot{L}).$$

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

$$\frac{T\dot{F}P}{TFP} = \frac{\dot{Y}'}{Y'} - \alpha\frac{\dot{K}}{K} - \beta\frac{\dot{L}}{L} = 0. \tag{9}$$

If instead, we had calculated a Divisia index from the firm's disaggregated production function, (7),

$$\frac{T\dot{F}P}{TFP} = \frac{\dot{Y}}{Y} - \alpha\frac{\dot{M}}{M} - \beta\frac{\dot{N}}{N} - \gamma\frac{\dot{P}}{P} - \delta\frac{\dot{R}}{P} = \frac{\dot{B}}{B}, \tag{10}$$

we would have obtained the correct answer that, in this case, the rate of productivity growth is equal to the rate of output growth 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 are typically used to aggregate the capital (whether stocks or service flows) used in any real production process. Consider the calculation based on (8) and (9) if L had been aggregated by physical units and K by its marginal product. Then $\dot{L}/L = 0$ and TFP would increase by $[1-(\gamma+\delta)]\,\dot{B}/B = (\varepsilon+\eta)\,\dot{B}/B$, while the rest of the increase would be ascribed to an increase in capital. Since different kinds of labour are often expressed in common units such as labour hours, while 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).[35]

So the effects of technological changes that are felt below the levels of aggregation at which the production function is 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 technological 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, [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' (161, 166). Our analysis shows, however, that even when calculations are made at the industry level, substantial amounts of technological change will show up as increases in the industry's measured inputs.

3.5. Aggregation when calculating a TFP index

Now assume that a correct measure of the quantity of each type of input is available at whatever the level of aggregation TFP is being calculated. Percentage changes in each input can then be weighted and summed to get the overall percentage change in inputs to be compared with the percentage change in output so as 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.

35 An identical argument applies if we write equation (6) as $Y = B(mM)^{\gamma}(nN)^{\delta}(pP)^{\varepsilon}(rR)^{\eta}$, where lower case letters are efficiency parameters, and we allow one of more of them to change with B held constant.

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. 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 overestimate 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., hypothesis that the rate of productivity growth is independent of quantities of intermediate, capital, and labour inputs, and that the value shares are independent of time) is rejected in 39 of the 45 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 asks 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. The appendix provides a stylization of an economy comprising a primary sector and a manufacturing sector. Each sector is described by a Cobb-Douglas production function that uses both labour and capital as inputs, and there is a fixed total constraint on the amount of labour and capital available (i.e., there is no growth in total labour and capital). The economy is assumed to be in perfect competitive equilibrium initially. Hence the prices can be determined by the marginal products, and we have the normal accounting identity.

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, then TFP will exactly measure the free lunch associated with the change. (See equation (A1) in the appendix.)

Now, consider the second type of equilibrium. When the productivity change occurs, marginal products are driven out of equilibrium, but prices do not instantaneously adjust, because labour and capital are not instantaneously mobile. We can determine the direction of bias in the prices, given that we know productivity in manufacturing has risen relative to that of primary production.

Equation (A2) of the appendix shows the algebraic result where the difference between TFP calculated from the transition equilibrium denoted by the primes and TFP calculated from the instantaneous adjusted equilibrium is negative. This implies a negative bias in TFP calculations where marginal products are in transitional equilibrium. This bias exists for decreasing returns to scale, constant returns to scale, and even some parameterizations leading to increasing returns to scale. We can also see that there is no bias if all of the exponents on the inputs of all production functions are just equal to unity, which is a sufficient condition for the production functions to have increasing returns to scale. However, for sufficiently large increasing returns to scale, there is an upward bias in measured TFP. The algebra of the appendix thus allows us to specify the conditions under which TFP will be biased whenever the long run conditions of perfectly competitive equilibrium do not hold. These results encompass the results of Basu and Fernald (1995), who show an upward bias in TFP for imperfect competition under conditions of increasing returns to scale. In addition, the varying results of Jorgenson, Gollop, and Fraumeni (1987) can potentially be explained as differences in scale within the industries they measured.

Summary of the last two sections: On aggregation, we have found two important sets of circumstances in which free 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 does not hold and the share weights used in calculating TFP growth are biased by the existence of differing marginal products for the same input in different uses. This will bias TFP downward when all lines of production activity have decreasing or constant returns to scale, and even when they have mildly increasing returns.[36]

36 Hulten (2000, 34–5) advances a further reason why, in his opinion, TFP may assign some of the income increasing effects of technological change to increases in the capital stock. He 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, in Hulten's view, be incorrectly 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 of some of the growth to increases in the capital stock is incorrect. If the capital stock were to be held constant by fiat, while technological change continued, output would be growing only 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.

4. Conclusions

Changes in total or multi-factor productivity are correctly described as measuring changes in the difference between *measured* outputs and increases in *measured* inputs. Changes in TFP do not measure technological changes, since part of the return to innovation reimburses the (widely defined) development costs and thus show up as offsetting input costs. Changes in TFP are correctly interpreted as being an imperfect measure of the returns to investing in new technologies that are in excess of the return to investing in existing technologies, that is, the super normal gains of technological change. It is conceptually possible, therefore, to have sustained, technologically driven, economic growth with zero changes in TFP.

Not all of the super normal gains that are ideally measured by TFP growth are mere Manna from heaven, since the possibility of earning super normal profit is a needed incentive to undertake the uncertainties involved in invention and innovation rather than investing in existing technologies. Not all of the gains are captured by the initiating firms, since there are many spillovers to other agents in the form of unpaid-for increases in output or decreases in costs. Not all of the vast set of spillovers that extend over time and space are measured by TFP changes even ideally, since it measures only contemporaneous effects.

TFP growth is only an imperfect measure of the contemporaneous super normal gains associated with technological change because, among other things: (i) the same technological advance will have different effects on aggregate TFP, depending on the timing of the increases in output associated with the decreases in prices; (ii) the treatment of R&D in the national accounts tends to cause some innovative activity to reduce measured TFP, while the results may or may not show up as future increases in TFP; (iii) whenever there are unmeasured inputs, increases in their the use will bias measured TFP growth upwards, while increases in the use of measured inputs will bias it downwards; (iv) technological changes that occur below the level of aggregation at which TFP is calculated, whether for the firm, the industry or the economy, tend to show up as increases in the amount of inputs used rather than as increases in their efficiency; (v) when current marginal products are not reflected in the weights used in aggregating sectoral measures of TFP growth to obtain national measures, biases of unknown amounts (and if there are increasing returns to scale, of unknown direction) are introduced.

Many research projects are suggested by this paper, but space permits us briefly to mention only four. First, research into the quantitative nature of some of these biases is called for. Second, research is needed into the possibility of developing counterfactual measures of the impact of technological change. (The present authors are working on this problem.) Third, research is needed into the possibility that the biases differ sectorally. For example, TFP measures currently assign much U.S. growth to increases in the capital stock in the manufacturing sector, while much Canadian growth is assigned to growth of TFP in the service sector, (see, e.g., Tang, Rao, and Sharpe, forthcoming), which is often misinterpreted as being the result of technological change in that sector. It is an interesting possibility that this result, which is much discussed by policy-makers, is an artefact of differences between services and manufacturing in the bias for technological change to shown up as increases in the measured amount of capital. Fourth, independent measures of technological change and diffusion are needed to help to interpret results from the previous three lines of research.

In the meantime, the major differences in interpretations of TFP among those who use it should give us pause when TFP measures are used to support assertions about things such as the arrival or non-arrival of the 'New (ICT-driven) Economy,' or the lack or presence of technological dynamism at the time of the First Industrial Revolution, or the success or failure of the growth policies of the Asian Tigers. Confusion over the interpretation of TFP would be much reduced if each report of a TFP measurement carried the caveat: *there is no reason to believe that changes in TFP in any way measure technological change*.

Appendix

Consider the following stylization of a two-sector economy. Let the primary production sector be

$$X = A(L_x)^{\alpha}(K_x)^{\beta} \quad (\alpha,\beta) \in (0,\infty).$$

Let the manufacturing production sector be

$$Y = B(L_y)^{\gamma}(K_y)^{\sigma} \quad (\gamma,\sigma) \in (0,\infty).$$

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 numeraire 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:

$$\frac{P_x}{P_y} = \frac{MP_{L_y}}{MP_{L_x}} = \frac{MP_{K_y}}{MP_{K_L}}$$

$$\frac{w_y}{p_y} = \frac{w_x}{p_x} = MP_{L_x}$$

and

$$\frac{r_y}{p_y} = \frac{r_x}{p_y} = MP_{k_y}.$$

Simplify further by normalizing P_y 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:

$$\frac{T\dot{F}P}{TFP} = \left(\frac{P_x X}{P_x X + Y}\right)\frac{\dot{X}}{X} + \left(\frac{Y}{P_x X + Y}\right)\frac{\dot{Y}}{Y} - \left(\frac{w_x L_x}{w_x L_x + w_y L_y}\right)\frac{\dot{L}_x}{L_x} - \left(\frac{w_y L_y}{w_x L_x + w_y L_y}\right)\frac{\dot{L}_y}{L_y} - \left(\frac{r_x K_x}{r_x K_x + r_y K_y}\right)\frac{\dot{K}_x}{K_x} - \left(\frac{r_y K_y}{r_x K_x + r_y K_y}\right)\frac{\dot{K}_y}{K_y},$$

which is simply the Divisia index for TFP.

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

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

From the resource constraints we know that

$$\dot{L}_y = -\dot{L}_x, \text{ and } \dot{K}_y = -\dot{K}_x,$$

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

$$\frac{T\dot{F}P}{TFP} = \frac{1}{P_x X + Y}\left[\frac{\dot{B}}{B}Y\right]. \tag{A1}$$

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:

$$\frac{P_x}{P_y} < \frac{MP_{L_y}}{MP_{L_x}}, \text{ and } \frac{P_x}{P_y} < \frac{MP_{K_y}}{MP_{K_x}}.$$

Again

$$\frac{w_y}{P_y} = MP_{L_y}, \text{ and } \frac{r_y}{P_y} = MP_{K_y}.$$

However, now,

$$\frac{w_x}{P_y} = \frac{P_x}{P_y} MP_{L_x} < \frac{MP_{L_y}}{MP_{L_x}} MP_{L_x}, \text{ and } \frac{r_x}{P_y} = \frac{P_x}{P_y} MP_{K_x} < \frac{MP_{K_y}}{MP_{K_x}} MP_{K_x}.$$

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

$$\frac{P_x}{P_y} = \frac{G(MP_{L_y})}{MP_{L_x}}, \ \frac{P_x}{P_y} = \frac{\hat{G}(MP_{K_y})}{MP_{K_x}}$$

and

$$\frac{w_x}{P_y} = \frac{P_x}{P_y} MP_{L_x} = \frac{G(MP_{L_y})}{MP_{L_x}} MP_{L_x} = G(MP_{L_y}),$$
$$\frac{r_x}{P_y} = \frac{P_x}{P_y} MP_{K_x} = \frac{\hat{G}(MP_{K_y})}{MP_{K_x}} MP_{K_x} = \hat{G}(MP_{K_y}).$$

Once again, normalize P_y to be one. The change in TFP is expressed as

$$\frac{T\dot{F}P'}{TFP'} = \frac{P_x X}{P_x X + Y}\frac{\dot{X}}{X} + \frac{Y}{P_x X + Y}\frac{\dot{Y}}{Y} - \frac{w_x L_x}{w_x L_x + w_y L_y}\frac{\dot{L}_x}{L_x} - \frac{w_y L_y}{w_x L_x + w_y L_y}\frac{\dot{L}_y}{L_y} - \frac{r_x K_x}{r_x K_x + r_y K_y}\frac{\dot{K}_x}{K_x} - \frac{r_y K_y}{r_x K_x + r_y K_y}\frac{\dot{K}_y}{K_y}.$$

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

$$\frac{T\dot{F}P'}{TFP'}=\frac{1}{P_xX+Y}\left[G(MP_{L_y})\dot{L}x+\hat{G}(MP_{K_y})\dot{K}_x+MP_{L_y}\dot{L}_y+(MP_{K_y})\dot{K}_y+\frac{\dot{B}}{B}Y\right]$$
$$-\frac{1}{w_xL_x+w_yL_y}[G(MP_{L_y})\dot{L}_x+MP_{L_y}\dot{L}_y]-\frac{1}{r_xK_x+r_yK_y}[\hat{G}(MP_{K_y})\dot{K}_x+MP_{K_y}\dot{K}_y].$$

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

$$\frac{T\dot{F}P'}{TFP'}-\frac{T\dot{F}P}{TFP}<0,$$

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

First, we note that $T\dot{F}P/TFP$ is a positive term in $T\dot{F}P'/TFP'$ so that we can eliminate $1/(P_xX+Y)$ $[(\dot{B}/B)\mathrm{Y}]$ from both TFP calculations leaving just the remaining terms in $T\dot{F}P'/TFP'$. Making use of the fact that $\dot{L}_x=-\dot{L}_y$ and $\dot{K}_x=-\dot{K}_y$, we get the following:

$$\frac{T\dot{F}P'}{TFP'}-\frac{T\dot{F}P}{TFP}=\left[\frac{G(MP_{L_y})-MP_{L_y}}{P_xX+Y}-\frac{G(MP_{L_y})-MP_{L_y}}{w_xL_x+w_yL_y}\right]\dot{L}_x$$
$$+\left[\frac{\hat{G}(MP_{K_y})-MP_{K_y}}{P_xX+Y}-\frac{\hat{G}(MP_{K_y})-MP_{K_y}}{r_xK_x+r_yK_y}\right]\dot{K}_x \qquad \text{(A2)}$$

or

$$=\left[\frac{1}{P_xX+Y}-\frac{1}{w_xL_x+w_yL_y}\right][G(MP_{L_y})-MP_{L_y}]\dot{L}_x$$
$$+\left[\frac{1}{P_xX+Y}-\frac{1}{r_xK_x+r_yK_y}\right][\hat{G}(MP_{K_y})-MP_{K_y}]\dot{K}_x.$$

We want to sign this to determine if there is bias in the Divisia index when the marginal products are not fully adjusted to long-run competitive equilibrium. To do this, note the following:

$$[G(MP_{L_y})-MP_{L_y}]\dot{L}_x>0,\text{and}[\hat{G}(MP_{K_y})-MP_{K_y}]\dot{K}_x>0,\text{ since}$$
$$[G(MP_{L_y})-MP_{L_y}]<0,\dot{L}_x<0,[\hat{G}(MP_{K_y})-MP_{K_y}]<0,\text{and }\dot{K}_x<0.$$

Thus, the two expressions left to evaluate are

$$\left[\frac{1}{P_xX+Y}-\frac{1}{w_xL_x+w_yL_y}\right]$$

and

$$\left[\frac{1}{P_xX + Y} - \frac{1}{r_xK_x + r_yK_y}\right].$$

By evaluating the expressions around zero we get

$$\left[\frac{1}{P_xX + Y} - \frac{1}{w_xL_x + w_yL_y}\right] = G(MP_{L_x})\left(1 - \frac{1}{\alpha}\right) + \left(MP_{L_y}\right)\left(1 - \frac{1}{\gamma}\right) < 0 \quad \text{if } \alpha < 1 \text{ and } \gamma < 1$$

and

$$\left[\frac{1}{P_xX + Y} - \frac{1}{r_xK_x + r_yK_y}\right] = \hat{G}(MP_{K_x})\left(1 - \frac{1}{\beta}\right) + \left(MP_{K_y}\right)\left(1 - \frac{1}{\sigma}\right) < 0 \quad \text{if } \beta < 1 \text{ and } \sigma < 1.$$

This says that

$$\frac{T\dot{F}P'}{TFP'} - \frac{T\dot{F}P}{TFP} < 0 \text{ for } \alpha < 1, \gamma < 1, \beta < 1 \text{ and } \sigma < 1.$$

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. This implies that there is an identifiable bias in TFP calculations where marginal products are in transitional equilibrium. This negative bias exists for decreasing returns to scale, constant returns to scale, and even some parameterizations leading to increasing returns to scale. Furthermore, for sufficiently large increasing returns to scale there will be an upward bias of measured TFP.

References

Abramovitz, M. (1956) 'Resource and output trends in the U.S. since 1870,' *American Economic Review* 46, 5–23

Barro, R.J. (1999) 'Notes on growth accounting,' *Journal of Economic Growth* 4, 119–37

Basu and Fernald (1995) 'Are apparent productive spillovers a figment of specification error?' *Journal of Monetary Economics* 36, 165–88

— (1997) 'Returns to scale in U.S. productions: estimates and implications,' *Journal of Political Economy* 105, 249–83

Bekar, Clifford, and Richard G. Lipsey (forthcoming) 'Science, institutions, and the Industrial Revolution,' *Journal of European Economic History*

Berg, Maxime, and Pat Hudson (1994) 'A Comment on the Crafts-Harley View of the Industrial Revolution,' *Economic History Review* NS 47, 147–9

Blaug (1997) 'Competition as an end-state and competition as a process,' in *Trade Technology and Economics: Essays in Honor of Richard G. Lipsey*, ed. B. Curtis Eaton and Richard G. Harris (Cheltenham, U.K.: Edward Elgar)

Caballero, R., and R. Lyons (1990) 'Internal and external economies in European industries,' *European Economic Review* 34, 805–30

— (1992) 'External effects in the U. S. procyclical productivity,' *Journal of Monetary Economics* 29, 209–26

Carlaw, Kenneth, and Steven Kosempel (1994) 'The sources of economic growth in Canada,' *Economics of Innovation and New Technology* 13, 299–309

Carlaw, Kenneth, and Richard G. Lipsey (2002) 'Externalities, technological complementarities and sustained economic growth,' *Research Policy* 31, 1305–15

— (forthcoming) 'A model of GPT-driven, sustained growth,' *Economic Journal*

Crafts, N.F. (1996) 'The first Industrial Revolution: a guided tour for growth economists,' *American Economic Review* 86, 197–201

Crafts, N.F., and C. K. Harley (1992) 'Output growth and the industrial revolution,' *Economic History Review* 45, 703–30

Diewert, E., (1987) 'Index numbers,' in Eatwell, Milgate, and Newman (1987)

Diewert, E., and D. Lawrence (1999) 'Measuring New Zealand's productivity,' New Zealand Treasury Working Paper http://www.treasury.govt.nz/workingpapers/1999/99-5.asp.

Diewert, W.E., and A.O. Nakamura (2003) 'Index number concepts, measures and decompositions of productivity growth,' *Journal of Productivity Analysis* 19, 127–59

Eatwell, J., M. Milgate, and P. Newman, eds (1987) *The New Palgrave, a Dictionary of Economics* (London: Macmillan)

Fare, Grosskopf, and Margaritis (1996) 'Productivity growth,' in *A Study of Economic Reform: The Case of New Zealand*, ed. B. Silverstone, A. Bollard, and R. Lattimore Contributions to Economic Analysis, vol. 236 (Amsterdam: Elsevier)

Fare, Grosskopf, Norris, and Zhang (1994) 'Productivity growth, technical progress, and efficiency change in industrialized countries,' *American Economic Review* 84, 66–83

Freeman, Chris, and F. Louca (2002) *As Time Goes By: From the Industrial Revolution to the Information Revolution* (Oxford: Oxford University Press).

Gordon, R.J. (2000) 'Does the "New Economy" measure up to the great inventions of the past?,' NBER Working Paper No. 7833, Washington, DC

Greenwood, J., Z. Hercowitz, and P. Krusell (1997) 'Long run implications of investment-specific technological change,' *American Economic Review* 87, 342–62

— (2000) 'The role of investment-specific technological change in the business cycle,' *European Economic Review* 44, 91–115

Griliches, Z. (1987) 'Productivity: measurement problems,' in Eatwell, Milgate, and Newman (1987)

— (1994) 'Productivity, R&D, and the data constraint,' *American Economic Review* 84, 1–24

— (1995) 'The discovery of the residual,' NBER Working Paper No. 5348, Washington, DC

Grubler, Arnulf (1998) *Technology and Global Change* (Cambridge: Cambridge University Press)

Hall, Robert E. (1988) 'The relation between price and marginal cost in U.S. industry,' *Journal of Political Economy* 96, 921–47

Harberger, A. (1998) 'A vision of the growth process,' *American Economic Review* 88, 1–32

Helpman, Elhanan, ed. (1998) *General Purpose Technologies and Economic Growth* (Cambridge, MA: MIT Press)

Hulten, Charles R. (1979) 'On the importance of productivity change,' *American Economic Review* 65, 126–36
— (2000) 'Total factor productivity: a short biography,' NBER Working Paper No. 7471, Washington, DC
IMF Staff (2000) Country Report No. 00/34
Jorgenson, Dale (1995) *Productivity*, Vol. 1, *Postwar U.S. Economic Growth* (Cambridge, MA: MIT Press).
— (2001) 'Information technology and the U.S. economy,' *American Economic Review* 91, 1–32
Jorgensen, Dale, and B.M. Fraumeni (1995) 'The accumulation of human and nonhuman capital 1948–1984,' in Jorgensen (1995)
Jorgensen, Dale, and Z. Griliches (1967) 'The explanation of productivity change,' *Review of Economic Studies* 34, 249–83
Jorgensen, Dale, and K.J. Stiroh (2000) 'U.S. economic growth at the industry level,' *American Economic Review, Papers and Proceedings* 90, 161–7
Jorgensen, Dale, Frank Gollop, and B.M. Fraumeni, (1987) *Productivity and U.S. Economic Growth* (Cambridge, MA: Harvard University Press)
Krugman, Paul (1996) 'The myth of Asia's miracle,' in *Pop Internationalism* (Cambridge, MA: MIT Press).
Lipsey, Richard G. (1992) 'Global change and economic policy,' in *The Culture and Power of Knowledge: Inquiries into Contemporary Societies*, ed. Nico Stehr and Richard V. Ericson (Berlin: Walter de Gruyter)
— (1993) 'Globalisation, technological change and economic growth,' Annual Sir Charles Carter Lecture, Northern Ireland Economic Council, Report No. 103; reprinted in Lipsey (1997)
— (1994) 'Markets, technological change and economic growth,' *Pakistan Development Review* 33, 327–52; reprinted in Lipsey (1997)
— (1997) *The Selected Essays of Richard Lipsey*, Vol. I, *Microeconomics, Growth and Political Economy*; Vol. II, *Macroeconomic Theory and Policy* (Cheltenham, U.K.: Edward Elgar)
— (2002) 'The productivity paradox: a case of the emperor's new clothes,' *ISUMA: Canadian Journal of Policy Research*, Special Issue on Policies for the New Economy, 3, 120–6
Lipsey, Richard G., and Clifford Bekar (1995) 'A structuralist view of technical change and economic growth,' in *Bell Canada Papers on Economic and Public Policy* (Kingston, ON: John Deutsch Institute)
Lipsey, Richard G., Clifford Bekar, and Kenneth Carlaw (1998a) 'What requires explanation,' in Helpman (1998)
— (1988b) 'The consequences of changes in GPTs,' in Helpman (1998)
Metcalf, S. (1987) 'Technical change,' in Eatwell, Milgate, and Newman (1987)
Nelson, Richard R. (1964) 'Aggregate production function and medium-range growth projections,' *American Economic Review* 54, 575–606
Pack, H., and L.E. Westphal (1986) 'Industrial strategy and technological change: theory or reality,' *Journal of Development Economics* 22, 87–128
Pomeranz, Kenneth (2000) *The Great Divergence: China, Europe and the Making of the Modern World Economy* (Princeton, NJ: Princeton University Press)
Prescott, E.C., (1998) 'Needed a theory of total factor productivity,' *International Economic Review* 39, 525–51
Rosenberg, Nathan (1982) *Inside the Black Box: Technology and Economics* (Cambridge: Cambridge University Press)
Rymes, Thomas K., (1971) *On Concepts of Capital and Technical Change* (Cambridge: Cambridge University Press)

Solow, R. (1956) ‘A contribution to the theory of economic growth,’ *Quarterly Journal of Economics* 70, 65–94

— (1957) ‘Technical change and the aggregate production function,’ *Review of Economics and Statistics* 39, 312–20

Statistics Canada (1998) Catalogue No. 13–568

Tang, J., S. Rao, and A. Sharpe (forthcoming 2005) ‘Productivity growth in service industries: a Canadian success story,’ in *Service Industries and the Knowledge-Based Economy*, ed. R.G. Lipsey and A. Nakamura (Calgary: University of Calgary Press)

Wade, Robert (1990) Governing the market: economic theory and the role of government in East Asian industrialization, (Princeton, NJ: Princeton University Press)

Westphal, L. (1990) ‘Industrial policy in an export-propelled economy: lessons from South Korea’s experience,’ *Journal of Economic Perspectives* 4, 41–60

Womack, J.P., D.J. Jones, and D. Roos (1990) *The Machine that Changed the World* (New York: Rawson)

Young, Alwyn (1992) ‘A tale of two cities: factor accumulation and technical change in Hong Kong and Singapore,’ *NBER Macroeconomic Annual* (Cambridge, MA: MIT Press)

— (1995) ‘The tyranny of the numbers,’ *Quarterly Journal of Economics* 110, 641–80