MEASURING MFP WHEN RATES OF RETURN ARE EXOGENOUS

Paul Schreyer
OECD
21 October 2003/Rev1
Draft – do not quote without permission of the author

Paper presented at the Workshop on Rethinking Total Factor Productivity Measurement

*Organised by Alice Nakamura and Erwin Diewert
October 27-28, 2003
Ottawa

TABLE OF CONTENTS

1. Introduction: gross operating surplus and the remuneration of capital

Official statistics do not normally provide direct observations on the price and volume of capital services. What is available from the national accounts is a measure of gross operating surplus, i.e., a measure of business profits from normal business activity, including mixed income, the income of self-employed persons. Thus, national accounts provide the researcher with data according to the following accounting identity:

(1) $$PQ = wL + GOS,$$

where PQ is current price output, wL is the remuneration of labour with a wage component w and a volume component L, and GOS is the gross operating surplus. For simplicity, it will be assumed here that mixed income is either zero or has been split up between the labour and the GOS component.

The national accounts provide no indication as to exactly which factor of production is remunerated through GOS. Fixed assets are certainly among them but they are not necessarily the only ones. The business literature offers a wealth of discussions about the importance of intangible assets, and there are good reasons to argue that such assets account at least for part of GOS. While this may appear a minor point, it puts in question an assumption routinely made by analysts of productivity and growth, namely that GOS exactly represents the remuneration of the fixed assets recognised in the System of National Accounts (SNA), or the value of the services of these assets. Assume that $\sum_{i=1}^{N} u_i^* K_i$ is a value for services from N capital goods, with a price component u_i^* (the user costs of capital) and a quantity component K_i (the flow of capital services of those assets identified in the national accounts), then the assumption is typically made that $GOS = \sum_{i=1}^{N} u_i^* K_i$.

In other words, it is assumed that remuneration of capital services exactly exhausts gross operating surplus. Empirically, the equality $GOS = \sum_{i=1}^{N} u_i^* K_i$ is obtained by choosing an appropriate value for the net rate of return to assets, which is part of the user costs[1]. Thus, the rate of return has been fixed *endogenously*.

This assumption is convenient because it is consistent with competitive behaviour on product and factor markets and a production process that exhibits constant returns to scale. Conditions such as these are required if (1) should be re-written as

(2) $$PQ = wL + \sum_{i=1}^{N} u_i^* K_i$$

[1] In a simple continuous-time formulation without taxes, the user costs or rental price of an asset (Jorgenson 1967) is given by $u^i = p_K^i \left(r + \delta^i - d\ln p_K^i / dt\right)$ where $p_K^i(t)$ is the purchase price of a new asset of type i, $r(t)$ is the net rate of return, δ^i is a rate of depreciation, $d\ln p_K^i / dt$ is rate of change of p_K^i .

i.e., if gross operating surplus should correspond exactly to the remuneration of fixed assets. Counter to claims sometimes made in the literature (Hsieh 2002 is a recent example) (2) is *not* an accounting identity. For example, for (2) to hold,

- the set of assets K has to be complete in the sense that all assets are observed by the statistician who compiles the national accounts;
- the ex-post rate of return on each asset (implicitly observed by the national accountant as part of GOS) equals its ex-ante rate return, the economically relevant part in the user costs of capital services;
- there has to be absence of purely residual profits such as mark-ups that arise in the presence of market power.

It is not immediately obvious that such conditions hold empirically. What happens in the more general case? Under exogenous rates of return to capital, total costs, i.e., the sum of labour remuneration and capital remuneration, will not necessarily equal total revenues or total value added at current prices. This raises the question how under such circumstances, indices of multi-factor productivity (MFP) should be computed and how growth accounting exercises should be carried out. The present note discusses some of the options, provides interpretations of various productivity measures and expresses preference for one particular MFP measure. Few things are new among the ideas expressed below except perhaps that they have been brought together in one place. For a broader and more rigorous treatment of production theory, we refer to Fuss and McFadden (eds.) (1978), and for productivity measures and index number theory to Balk (1998) and Diewert and Nakamura (2003). Further, well-known work on the effects of different assumptions about rates of return on productivity measure has been carried out by Harper et al. (1990).

2. Why GOS may differ from remuneration of capital

This paper considers a more general formulation and assumes that $PQ - wL = GOS = \sum_{i=1}^{N} u_i K_i + M$. Thus there is a term M that corresponds to the difference between the observed or imputed remuneration of assets and the value of gross operating surplus or non-labour input. If one uses $C = \sum_{i=1}^{N} u_i K_i + wL$ as shorthand for observed factor payments, one could also say that gross operating surplus is split into one component that reflects observable factor remuneration and another, residual one with several possible interpretations ($GOS = C + M$). In principle, there is no restriction on the sign of M. However, a negative term over an extended period of time is economically implausible and shall not be considered further. Consequently, in what follows, non-negativity of M is assumed.

We propose two interpretations of M. The first one is the existence of a mark-up of prices over total costs. Such a mark-up may exist if output markets are not fully competitive or if there is risk involved in producing and selling output so that M reflects a risk premium for financial capital invested in the firm or remuneration of entrepreneurship.

The second interpretation puts forward the existence of unobserved inputs and essentially describes a statistical phenomenon. It arises when: (a) not all capital inputs that give rise to operating surplus (itself estimated residually in the national accounts as non-labour income) are recognised in the national accounts; (b) when production processes exhibit decreasing returns to scale. Decreasing returns to scale constitute a special case of a non-observed input (Diewert and Nakamura 2003) and can therefore be treated in a similar way. However, in contrast to (a), this particular unobserved input would remain unobserved even if national accounts were exhaustive in capturing all types of market assets.

3. Production technology and producer behaviour

We start by introducing a production function that relates capital services from N different observable assets $K_1,...K_N$ and from one non-observable asset D as well as labour services L to a single[2] output Q:

(3) $Q = F[K_1,..., K_N, D, L, t]$

The production function contains a time variable t and as usual, Hicks-neutral technical change is measured by the shift of the production function over time. To derive a measure for this shift, F is differentiated totally and technical change is then given by the partial derivative of the production function with respect to the time variable:

(4) $$\frac{\partial F}{\partial t}\frac{1}{F} = \frac{\partial \ln F}{\partial t} = \frac{d\ln Q}{dt} - \sum_{i=1}^{N}\frac{\partial F}{\partial K_i}\frac{K_i}{F}\frac{d\ln K_i}{dt} + \frac{\partial F}{\partial D}\frac{D}{F}\frac{d\ln D}{dt} + \frac{\partial F}{\partial L}\frac{L}{F}\frac{d\ln L}{dt}.$$

A profit-maximising firm will hire labour and capital up to the point where marginal cost equals marginal revenue. When allowing for imperfect product markets but assuming competitive factor markets, profit maximization by a producer implies that there may be a mark-up (captured by the term[3] $0 < \mu \leq 1$) of prices over costs. Marginal revenues from capital service K_i are then given by $\mu P \frac{\partial F}{\partial K_i}$ where P denotes the output price. Marginal costs are given by the user cost term u_i so that the optimality condition for observable capital inputs reads as $\mu P \frac{\partial F}{\partial K_i} = u_i$ i=1,...N and the elasticity of production of capital good

[2] This is for convenience only: explicit extension to a vector of outputs is possible.

[3] When producers face an elastic demand function for their product, μ would be given by $\left(1 + \frac{\partial P}{\partial Q}\frac{Q}{P}\right)$ where $0 \geq \frac{\partial P}{\partial Q}\frac{Q}{P} > -1$ is the inverted price elasticity of demand.

i is $\frac{\partial F}{\partial K_i}\frac{K_i}{F}=\frac{u_i K_i}{\mu PQ}$. By a similar reasoning, the optimality conditions for the non-observed input D and for labour L can be established, where ϕ is the user cost of the unobserved input:

(5) $$\frac{\partial F}{\partial D}\frac{D}{F}=\frac{\phi D}{\mu PQ}\ \frac{\partial F}{\partial L}\frac{L}{F}=\frac{wL}{\mu PQ};\ \frac{\partial F}{\partial K_i}\frac{K_i}{F}=\frac{u_i K_i}{\mu PQ}\ \text{i=1,...N.}$$

Finally, it is assumed that F is linear homogenous in its capital and labour components. Thus, there are constant returns to scale in observed and non-observed inputs and, by Euler's theorem $F=\sum_{i=1}^{N}\frac{\partial F}{\partial K_i}K_i+\frac{\partial F}{\partial D}D+\frac{\partial F}{\partial L}L$. In conjunction with the optimality conditions above, this yields

(6a) $$\mu PQ=\sum_{i=1}^{N}u_i K_i+\phi D+wL.$$

At the same time, it has been assumed that $PQ=\sum_{i=1}^{N}u_i K_i+wL+M$ so that

(6b) $$M=PQ\left(1-\mu+\frac{\phi D}{PQ}\right).$$

(6b) shows how the difference M between GOS from the national accounts and the sum of payments to observed factors reflects market power or non-technological effects (captured by μ) and the influence of technologically relevant unobserved capital inputs (captured by ϕD).

4. Technical change

To find a computable expression for the technical change parameter $\frac{\partial \ln F}{\partial t}$ we use (4) and (5) to obtain the following theoretical index:

(7) $$\frac{\partial \ln F}{\partial t}=\frac{d\ln Q}{dt}-\frac{1}{\mu}\left\{\sum_{i=1}^{N}\frac{u_i K_i}{PQ}\frac{d\ln K_i}{dt}+\frac{\phi D}{PQ}\frac{d\ln D}{dt}+\frac{wL}{PQ}\frac{d\ln L}{dt}\right\}$$

In (7), each input is weighted by its share in total cost because $\mu PQ=\sum_{i=1}^{N}u_i K_i+\phi D+wL$ as shown in (6). Turned around, the growth accounting form of (7) is:

(8) $$\frac{d\ln Q}{dt}=\frac{1}{\mu}\left\{\sum_{i=1}^{N}\frac{u_i K_i}{PQ}\frac{d\ln K_i}{dt}+\frac{\phi D}{PQ}\frac{d\ln D}{dt}+\frac{wL}{PQ}\frac{d\ln L}{dt}\right\}+\frac{\partial \ln F}{\partial t}.$$

(8) delivers an explicit expression for the change in aggregate inputs, and if the mark-up parameter μ were known, and if there were no unobserved factor D, (7) could readily be implemented. In the presence of these unknown parameters, things are more

complicated. We start with a proposal for a computable MFP measure and follow with a discussion of its interpretation.

5. Deriving computable measures

There are essentially two strategies for the implementation of (8): (i) to introduce additional, and typically restrictive hypotheses about the size or nature of the unknown variables until an expression has been derived that is both computable and that offers a (seemingly) clear interpretation of productivity growth; (ii) to stay away from invoking additional hypotheses, and define a computable measure of productivity growth but allow for the fact that it may reflect more than pure technology shifts.

5.1. Apparent multi-factor productivity

We first follow avenue (ii) and propose a measure of 'apparent multi-factor productivity'. Then, additional hypotheses will be used to discuss strategy (i).

For the purpose at hand, let there be an aggregator X that combines quantities of observable inputs K and L. Specifically, define

$$(9)\qquad \frac{d\ln X}{dt} \equiv \sum_{i=1}^{N} \frac{u_i K_i}{C} \frac{d\ln K_i}{dt} + \frac{wL}{C} \frac{d\ln L}{dt}$$

as a Divisia quantity index of observable inputs, noting that weights correspond to the shares of each input in total *observable* inputs, as $C \equiv \sum_{i=1}^{N} u_i K_i + wL$. Next, define the rate of apparent multi-factor productivity growth (AMFP) as the difference between a Divisia quantity index of output and the quantity index of observable inputs as laid out in (9):

$$(10)\qquad AMFP \equiv \frac{d\ln Q}{dt} - \frac{d\ln X}{dt}.$$

The growth accounting equation that corresponds to (10) is then

$$(11a)\qquad \frac{d\ln Q}{dt} = \sum_{i=1}^{N} \frac{u_i K_i}{C} \frac{d\ln K_i}{dt} + \frac{wL}{C} \frac{d\ln L}{dt} + AMFP$$

$$(11b)\qquad \text{where } AMFP = \frac{\partial \ln F}{\partial t} + \frac{\phi D}{C + \phi D}\left[\frac{d\ln D}{dt} - \frac{d\ln X}{dt}\right].$$

According to (11a), the direct growth contribution of observed capital inputs and labour is given by the rate of the change of these variables, weighted by their average share in observed costs C. The productivity term AMFP reflects several factors: pure technical change or the shift of the production function and a term that captures effects of the non-observed variable D. Consider the following special cases:

- If there is no unobserved input (D=0), the second term in AMFP disappears and AMFP correctly captures technical change.
- If the volume change of the unobserved input equals the volume change of observed inputs, the second term disappears also and AMFP reflects only technical change.
- If there is an unobserved input but its quantity is invariant over time (D>0, $dD/dt = 0$), the second term in (11b) turns out to be equivalent to an expression that would be obtained for a production technology with observed inputs only but with decreasing returns to scale in these inputs. This is explained in greater detail in section 5.2 (d).

We conclude that – whatever the exact nature of the unobserved factor D – the AMFP computation will capture 'pure' technical change, the growth contributions of unobserved assets and scale effects. Thus, it preserves a technology-related interpretation as one would expect in conjunction with a measure of productivity. One notes, however, that these technology-related effects are not path independent – they vary with the levels and growth rates of observed and non-observed inputs, and the latter depend in turn on prices of inputs and outputs as well as on the size of mark-ups.

The contribution of productivity change to output growth is given by AMFP. Clearly, the interpretation of AMFP has to be kept in mind: it reflects the combined effects of technical change, of non-observed inputs, of non-constant returns to scale and, indirectly, of deviations from perfect competition in product markets. In other words, AMFP is a true 'residual'. But for many practical purposes, it will fulfil its role as a multi-faceted measure of productivity growth. We note in passing that AMFP could also serve as a useful measure of productivity growth when technical change is of a more general nature, and not necessarily Hicks-neutral.

If one wants to extend the analysis, an additional analytical step could be taken to deconstruct AMFP into its technical change component and into other effects. However, this requires invoking parametric methods of estimation if one does not want to impose competitive behaviour on product markets.

5.2. Invoking additional assumptions

This section follows the approach (i) outlined earlier: additional hypotheses are invoked to deal with the possible presence of unobserved inputs and mark-ups. Each set of hypotheses is designed so as to lead to a 'correct' measure of MFP in the sense that it reflects Hicks neutral technical change *if the hypotheses hold.* In addition, it is discussed how under these circumstances, the pragmatic AMFP measure would fare: would its use imply an upwards or downward bias, or would it coincide with the MFP result?

a) Defining away the unobserved input

If one assumes away unobserved inputs (D=0), the only possibility to explain a difference between total costs of observed measures and GOS is to allow for a mark-up that corresponds to this difference. In this special case, the mark-up factor can be computed directly as $\mu = 1 - M/_{PQ}$, and is determined by the ratio M/PQ, where M corresponds to the difference between non-labour income (GOS) and the sum of observed capital costs.

In this case, the growth accounting equation above becomes

(12a) $$\frac{d\ln Q}{dt} = \frac{C}{\mu PQ}\left\{\sum_{i=1}^{N} \frac{u_i K_i}{C}\frac{d\ln K_i}{dt} + \frac{wL}{C}\frac{d\ln L}{dt}\right\} + MFP1$$

(12b) with $MFP1 = \frac{\partial \ln F}{\partial t}$.

MFP1 thus correctly traces technical change, provided the assumption D=0 is true. As $C = \mu PQ$, the calculation of MFP1 is simply $MFP1 = \frac{d\ln Q}{dt} - \frac{d\ln X}{dt}$. Thus, if the assumption above holds, the true productivity measure MFP1 would coincide with the result obtained from applying an AMFP measure. This was already pointed out under the first bullet in the discussion of AMFP.

b) Defining away mark-ups and assuming proportionality of D and K

A second possibility is to allow for an unobserved factor but to impose fully competitive output markets ($\mu = 1$). It follows that the entire difference between GOS and the sum of observed payments to capital is identified with payments to the unobserved input: $M = \phi D$. To measure (8), an additional assumption is needed, for example that the rate of change of the unobserved input D equals that of observed capital inputs: $\frac{d\ln D}{dt} = \frac{d\ln K}{dt}$ where $\frac{d\ln K}{dt} = \sum_{i=1}^{N} \frac{u_i K_i}{uK}\frac{d\ln K_i}{dt}$ is a Divisia quantity index of observed assets and $uK = \sum_{i=1}^{N} u_i K_i$ is the value of their capital services.

Under these conditions, the growth accounting equation (8) becomes

(13a) $$\frac{d\ln Q}{dt} = \left\{\frac{uK}{PQ}\frac{d\ln K}{dt} + \frac{M}{PQ}\frac{d\ln K}{dt} + \frac{wL}{PQ}\frac{d\ln L}{dt}\right\} + MFP2 \text{ and}$$

(13b) with $MFP2 = \frac{\partial \ln F}{\partial t}$.

MFP2 again traces technical change correctly, as long as the additional hypotheses hold. The measured growth contribution of the observed capital inputs merits further discussion. It is easily verified that $\frac{uK}{PQ} + \frac{M}{PQ} + \frac{wL}{PQ} = 1$ so that the weight that now attaches to the observed capital inputs $\frac{uK}{PQ} + \frac{M}{PQ} = \frac{GOS}{PQ}$ equals the share of GOS in total production or income which in turn is the complement to the labour share in total income. Thus, the income of the unobserved factor D has been distributed to the observed capital inputs. (13a) can thus be rewritten as:

(13c) $$\frac{d\ln Q}{dt} = \left\{ \frac{GOS}{PQ}\frac{d\ln K}{dt} + \frac{wL}{PQ}\frac{d\ln L}{dt} \right\} + MFP2$$

It is interesting to note that (13c) bears a strong resemblance to a model with endogenous net rates of return. The overall rate of growth of capital services, $d\ln K/dt$ as defined earlier, enters with the same weight – one minus the labour share in total income (see below, case (c)). Of course, in the endogenous model, the growth rate of observed capital services $d\ln K/dt$ may itself may be different, as each asset's user cost term reflects an endogenous rather than an exogenous rate of return.

Suppose the above assumptions are true but an AMFP measure is applied, what would be the resulting bias with regard to the 'true' MFP2 measure? After some manipulations, it can be shown that

(13c) $$AMFP = MFP2 + \frac{M}{PQ}\frac{d\ln K}{dt}.$$

Thus, AMFP will overstate MFP2 if the growth rate of observed capital assets – and by assumption the growth rate of the unobserved asset – is positive. In the empirical example presented in section 6, this is the case and AMFP turns out to be consistently higher than MFP2.

c) Defining away mark-ups and unobserved input

These are the assumptions invoked when MFP computations rely on endogenous rates of return. The endogenous approach was developed in detail by Christensen and Jorgenson (1969), and applied in many studies of productivity growth, including by statistical offices (BLS 2003). It is the most widely-used methodology but also the one that requires the most restrictive set of assumptions necessary to justify the use of endogenous rates of return[4]. There must be no unobserved inputs (D=0) and no mark-ups ($\mu = 1$) and perfect anticipation of asset price changes and depreciation. This implies that $PQ = C = \sum_{i=1}^{N} u_i^* K_i + wL$. The growth accounting model becomes

(14a) $$\frac{d\ln Q}{dt} = \left\{ \sum_{i=1}^{N} \frac{u_i^* K_i}{PQ}\frac{d\ln K_i}{dt} + \frac{wL}{PQ}\frac{d\ln L}{dt} \right\} + MFP3$$

(14b) where $MFP3 = \frac{\partial \ln F}{\partial t}$.

[4] The endogenous rate of return is computed by choosing that net rate of return that just equalises the sum of user costs of observed assets with non-labour income (GOS for simplicity) . Using the same notation for user costs as in footnote 1, this means $GOS = \sum_{i=1}^{N} p_K^i (r^* + \delta^i - d\ln p_K^i / dt) K^i$.

As before, the additional restrictions imply that measured productivity change corresponds to the shift of the production function. As pointed out above, there is similarity with (b) because (14a) can be re-written as

$$\text{(15)}\quad \frac{d\ln Q}{dt} = \left\{ \frac{GOS}{PQ}\frac{d\ln K^{*}}{dt} + \frac{wL}{PQ}\frac{d\ln L}{dt} \right\} + MFP3\,,$$

so that the contribution of capital assets is the product of the rate of growth of observed capital services and the share of GOS in total output or cost.

If the assumptions above are correct, and if an endogenous rate of return is used, a measurement of productivity with AMFP* would yield the correct result, as $AMFP^{*} = \frac{\partial \ln F}{\partial t} = MFP3$ in this case. We have marked AMFP* with an asterisk here to point out that AMFP is based on a capital measure that reflects endogenous rates of return. If AMFP is computed on the basis of exogenous rates, it would clearly differ from MFP3. This is also borne out in the empirical example below. However, no a-priori statement can be made as to the sign of this difference.

d) Defining away mark-ups and assuming time-invariance of the unobserved input or: a primal productivity measure when returns to scale are decreasing

This constitutes yet another possibility to deal with the unknown parameters D and μ. As before, μ is set to equal unity but this time the unobserved factor D is considered a fixed factor – positive but time invariant. Thus, the entire difference between GOS and observed asset rental payments is ascribed to the unobserved factor: $M = \phi D$. Because D is taken as time-invariant, any change in M implies a proportional change in ϕ. One obtains from (7) and (8) that

$$\text{(16a)}\quad \frac{d\ln Q}{dt} = \frac{C}{PQ}\left\{ \sum_{i=1}^{N} \frac{u_i K_i}{C}\frac{d\ln K_i}{dt} + \frac{wL}{C}\frac{d\ln L}{dt} \right\} + MFP4$$

(16b) where $MFP4 = \frac{\partial \ln F}{\partial t}$.

Some more discussion is useful here. First, (16a) can be written as

$$\text{(17)}\quad \frac{d\ln Q}{dt} = \frac{C}{PQ}\frac{d\ln X}{dt} + MFP4\,.$$

It can be shown that this growth accounting equation and/or productivity measure MFP4 is equivalent to the result that would be obtained if a model had been set up with a production function in observed inputs only, but allowing for non-constant returns to scale[5]. In the absence of market power, the ratio between costs and the value of output is readily interpreted as the elasticity of scale of the production function and (17) shows that the contribution of the observed assets correspond to their contribution to the input aggregator X, weighted by the scale elasticity C/PQ.

Note that to derive (16) from the general model (7) and (8), we have implicitly maintained the assumption of an optimal choice of D, $\frac{\partial F}{\partial D_i}\frac{D}{F}=\frac{\phi D}{\mu PQ}$. In other words, it has been assumed that the price of the unobserved factor D is always such that the optimality condition holds even though the quantity of D remains unchanged. Only by maintaining this condition can it also be ensured that a dual approach towards productivity measurement (assessing a shift of the cost function) yields the same result as the primal approach presented here. If we had an explicit model of production with non-constant returns to scale, the dual productivity measure would inevitable differ from the primal one (see point (e) below).

If the assumptions above are correct, how does MFP4 relate to AMFP? It is easily established that under these circumstances, AMFP would capture $AMFP=\frac{\partial \ln F}{\partial t}-(1-\frac{C}{PQ})\frac{d\ln X}{dt}$. If returns to scale are non-increasing ($C/PQ \le 1$), and if the quantity of input increases, AMFP would turn out to be smaller than MFP4 as AMFP captures both the effects of pure technical change and non-constant returns.

e) Defining away mark-ups and unobserved factors but assuming decreasing returns to scale: a dual productivity measure

The final option presented here is one without unobserved factors and a production technology that exhibits decreasing returns to scale, say $Q=G[K_1,\ldots,K_N,L,t]$. Under these circumstances, condition (6a) becomes $\lambda PQ=\sum_{i=1}^{N}u_iK_i+wL$ or $C/PQ=\lambda$ where λ is the scale elasticity of the production function G. If one follows the same procedure as in (7) and (8), a primal productivity measure $\frac{\partial \ln G}{\partial t}$ can be derived that has the same magnitude as MFP4. Put differently, $\frac{\partial \ln G}{\partial t}=\frac{\partial \ln F}{\partial t}$, and there is equivalence with the case described in section (d) above.

It is well known, however, that a production technology with non-constant returns to scale gives rise to several productivity measures (see, for example, Balk (1998)), in particular a primal one (the shift of the production function over time) and a dual one (the shift of a cost function over time). To derive the dual measure, a cost function $C=C(Q,u_1,u_2,\ldots u_N,w,t)$ is used that is dual to the production function G. Dual technical

[5] See also Diewert and Nakamura (2003) who introduce an unknown variable into a cost-function to deal with decreasing returns to scale.

change can then be presented as the (negative) rate of change of cost that is only associated with the passage of time, $\partial \ln C/\partial t$. It can be shown (Appendix 1) that this shift in the cost function corresponds to the following expression:

$$(18) \qquad -\frac{\partial \ln C}{\partial t} = \frac{PQ}{C}\frac{d\ln Q}{dt} - \frac{d\ln X}{dt} \equiv MFP5$$

Obviously, the dual approach cannot be applied to assess the contributions of primal factors to output growth – although it could be used to assess the contributions of primal factor prices to overall price change – but the dual approach yields yet another valid productivity measure, MFP5, as defined in (18). Under decreasing returns to scale and competitive factor markets (PQ>C), the dual productivity measure MFP5 will always exceed the primal one, MFP4. Furthermore, MFP5 will always exceed AMFP when returns to scale are decreasing.

6. Does it matter empirically?

At this point we shall take a look at the empirical implications of the different choices for productivity measures and growth accounting. Results are based on the OECD Capital Services Database (forthcoming) that is described in greater detail in Schreyer et al. (2003). For the present paper, only results for Canada are shown to illustrate the quantitative consequences of moving between different models. At a later stage, more countries will be examined to see whether the empirical conclusions hold more generally.

All results relate to the total economy. Output is measured as total GDP, labour input are total hours, and observed capital input measures are based on the perpetual inventory method and aggregated from six different asset types, IT hardware, communications equipment, other machinery and equipment, transport equipment, structures and intangible assets (software). For the results on the basis of exogenous rates of return, a long-term real interest rate of 4.8% was used to derive user costs of capital. The real interest rate reflects nominal interest rates of different maturity[6] for the period 1980-2000, deflated by the consumer price index.

[6] These are the average bank rate, the bank rate on prime loans, long-term government bond yields, short-term government bond yields, the interest rate on a 90 day bank fixed deposit and the treasury bill rate.

Table 1: Alternative MFP estimates, Canada
Average annual percentage changes, total economy

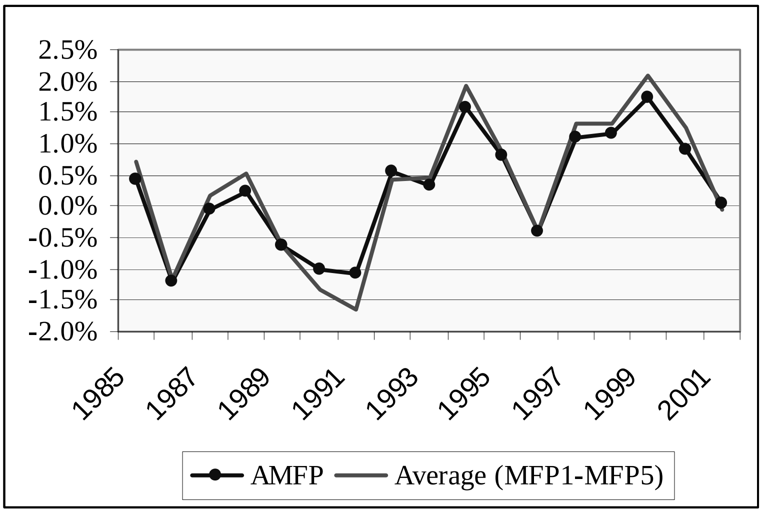
	1985-90	1990-95	1995-01	1985-01
Output	2.8%	1.7%	3.5%	2.7%
Labour input	2.5%	0.3%	2.0%	1.6%
Capital services (exogenous RoR)	6.2%	4.1%	4.9%	5.1%
Capital services (endogenous RoR)	5.7%	3.8%	4.2%	4.5%
MFP1	-0.5%	0.4%	0.8%	0.3%
MFP2	-1.2%	0.0%	0.3%	-0.3%
MFP3	0.1%	0.1%	0.6%	0.0%
MFP4	0.1%	0.7%	1.6%	0.7%
MFP5	0.1%	0.9%	1.6%	0.9%
Average (MFP1-5)	-0.3%	0.4%	1.0%	0.3%
AMFP	-0.5%	0.4%	0.8%	0.3%

Table 1 summarises empirical results for Canada. It shows the rates of change of output (GDP) and labour input as well as the volume changes in capital services, alternatively based on exogenous and endogenous rates of return. The first observation is that moving from an endogenous to an exogenous rate of return leads to a rise in the observed measure of capital input – at least in the case of Canada and for the period at hand.

The second panel in the table reviews results for the five alternative MFP measures presented in the text above. It is immediately apparent that the different options – each associated with a particular set of assumptions about market structures or production technology – leads to considerable variation in the resulting MFP measures. Unless one employs econometric techniques or unless a-priori knowledge about technology and market structure are available, it will be difficult to choose between the different options. Needless to say that every different MFP measure implies a different message about the relative contribution of capital services to output growth.

The third panel reproduces the pragmatic AMFP measure proposed in the text. It does not hinge on any particular hypothesis about perfect competition in output markets or about technology. It is also remarkable that a simple average across the five specific MFP measures yields a time series that is very close to the AMFP result. This can be seen from Figure 1 that plots the entire time series of AMFP against the time series of averages across the five other measures.

Figure 1: Apparent MFP and average of alternative rates, Canada
average annual percentage changes, total economy



7. Conclusions

This paper examined productivity and growth accounting measures when rates of return to capital inputs are exogenously determined. Several hypotheses about competition on output markets and about technology are invoked, each of which is compatible with exogenous rates of return. The following conclusions emerge:

The endogenous case – widely used in empirical research – imposes quite stringent assumptions: constant returns to scale and fully competitive output markets, all capital inputs are fully observed and there is perfect foresight by producers concerning expected changes in prices, and concerning rates of return and depreciation.

Different hypotheses entail different MFP measures. In the absence of further a-priori information or recourse to parametric techniques, there is no obvious way of discriminating between different hypotheses and to make an informed choice between productivity and growth accounting measures.

Empirically, the differences matter. This is evidenced for Canada in the present document but equally holds for other countries. In total, five different productivity measures were computed, each consistent with a particular set of assumptions.

The paper then goes on and suggests a pragmatic way forward – that of using an 'Apparent' MFP measure that is simply a ratio between a volume index of output and a volume index of observed aggregate inputs. It makes no claim to be a pure measure of technical change in the sense that it would represent a path-independent shift in the production or cost function. However, it is shown that AMFP

- only picks up technology-related effects in a broad sense (scale effects and shifts in the production function but not mark-ups);

- is a measure that is close to the average of other measures;
- can easily be communicated due to its simplicity;
- can also be applied when assumptions about the nature of 'pure' technical change are relaxed, allowing, for example, for a formulation that encompasses neutral and biased technical change;
- empirically traces closely the average of the five alternative measures derived in this paper. This has to be confirmed for a broader number of countries, however.

Clearly, the interpretation of AMFP has to be kept in mind: it reflects the combined effects of technical change, of non-observed inputs, of non-constant returns to scale and, indirectly, of deviations from perfect competition in product markets. In other words, AMFP is a true 'residual'. But for many practical purposes, it will fulfil its role as a multi-faceted measure of productivity growth.

References

Balk, Bert (1998) ; *Industrial Price, Quantity, and Productivity Indices*; Kluwer Academic Publishers.

Bureau of Labor Statistics (2003); Multi-factor productivity measures, available from http://www.bls.gov/mfp/home.htm

Christensen, Laurits R. and Dale W. Jorgenson (1969); "The Measurement of U.S. Real Capital Input, 1919-67"; Review of Income and Wealth, Series 15, No 4, pp 293-320.

Diewert, Erwin D. and Alice Nakamura (2003); "The Measurement of Aggregate Total Factor Productivity Growth"; forthcoming in Heckman and Leamer (eds.) Handbook *of Econometrics*, Volume 6.

Diewert, Erwin D. (1976); "Exact and Superlative Index Numbers"; *Journal of Econometrics* (4).

Fuss, Melvin and Daniel McFadden (eds.) (1978); *Production Economics: A Dual Approach to Theory and Applications*, Amsterdam, North Holland Publications.

Harper, Michael J., Ernst R. Berndt and David O. Wood (1990); "Rates of Return and Capital Aggregation Using Alternative Rental Prices"; in Dale W. Jorgenson and Ralph Landau (eds.); *Technology and Capital Formation*; MIT Press.

Hsieh, Chang-Tai (2002); "What Explains the Industrial Revolution in East Asia? Evidence from the Factor Markets"; *American Economic Review*, Volume 92, Issue 3, June, Pages 502-526.

Hulten, Charles (1990); "The Measurement of Capital"; in Ernst R. Berndt and Jack Triplett (eds.), *Fifty Years of Economic Measurement*, NBER.

Schreyer, Paul; Emmanuel Bignon et Julien Dupont (2003); "OECD Capital Services Measures: Methodology and a First Set of Results"; *OECD Statistics Working Papers (forthcoming)*.

Shephard, Ronald W. (1953); *Theory of Cost and Production Functions*. 1970 edition, Princeton, NJ: Princeton University Press.

Appendix 1

This appendix derives the dual productivity measure $\partial \ln C/\partial t$ discussed in section (5e). We start by defining the cost function[7] $C = C(Q, u_1, u_2, \ldots u_N, w, t)$ as $C = \min_{K_1,..K_N,L} \left\{\sum_{i=1}^{N} u_i K_i + wL\right\}$ s.t. $Q = G(K_1,\ldots,K_N,L,t)$. Differentiating C totally yields:

(A.1) $$dC = \sum_{i=1}^{N} \frac{\partial C}{\partial u^i} du^i + \frac{\partial C}{\partial w} dw + \frac{\partial C}{\partial Q} dQ + \frac{\partial C}{\partial t} dt$$ and further

(A.2) $$\frac{\partial C}{\partial t}\frac{1}{C} = \frac{\partial \ln C}{\partial t} = \frac{d\ln C}{dt} - \sum_{i=1}^{N} \frac{\partial C}{\partial u^i}\frac{u^i}{C}\frac{d\ln u^i}{dt} - \frac{\partial C}{\partial w}\frac{w}{C}\frac{d\ln w}{dt} - \frac{\partial C}{\partial Q}\frac{1}{C}\frac{d\ln Q}{dt}.$$

For a cost-minimising producer who faces competitive factor market, Shepard's (1953) lemma holds: $\frac{\partial C}{\partial u^i} = K^i$ and $\frac{\partial C}{\partial w} = L$. If further the assumption of profit maximisation is maintained, output prices equal marginal cost, and $\frac{\partial C}{\partial Q} = P\mu$ so that the shift of the cost function over time is given by:

(A.3) $$-\frac{\partial \ln C}{\partial t} = \frac{\mu PQ}{C}\frac{d\ln Q}{dt} - \left(\frac{d\ln C}{dt} - \sum_{i=1}^{N} \frac{u^i K^i}{C}\frac{d\ln u^i}{dt} - \frac{wL}{C}\frac{d\ln w}{dt}\right).$$

Next, use the shorthand symbol $\frac{d\ln X}{dt} \equiv \sum_{i=1}^{N} \frac{u_i K_i}{C}\frac{d\ln K_i}{dt} + \frac{wL}{C}\frac{d\ln L}{dt}$ for the Divisia quantity index of inputs, and $\frac{d\ln P_X}{dt} \equiv \sum_{i=1}^{N} \frac{u_i K_i}{C}\frac{d\ln u_i}{dt} + \frac{wL}{C}\frac{d\ln w}{dt}$ for the corresponding Divisia price index of inputs to obtain the following relationship:

(A.4) $$\frac{d\ln X}{dt} = \frac{d\ln C}{dt} - \sum_{i=1}^{N} \frac{u^i K^i}{C}\frac{d\ln u^i}{dt} - \frac{wL}{C}\frac{d\ln w}{dt} = \frac{d\ln C}{dt} - \frac{d\ln P_X}{dt}.$$

We can now establish the link between the primal and the dual measure of productivity change:

(A.5) $$-\frac{\partial \ln C}{\partial t} = \frac{\mu PQ}{C}\frac{d\ln Q}{dt} - \frac{d\ln X}{dt},$$ and

(A.6) $$-\frac{\partial \ln C}{\partial t} = \frac{\mu PQ}{C}\left(\frac{d\ln Q}{dt} - \frac{C}{\mu PQ}\frac{d\ln X}{dt}\right) = \frac{\mu PQ}{C}\frac{\partial \ln F}{\partial t}.$$

Consequently, if there are decreasing returns to scale ($\mu PQ > C$, $\lambda < 1$), the primal measure of technical change will be smaller than the dual measure. The opposite holds for increasing returns and measures are equal for constant returns to scale.

[7] Key properties of the cost function were first derived by Shephard (1953) – many other developments followed. For an overview, see Diewert (1987).

Whichever way one approaches the measurement of technical change, though, the general case of the presence of market power and non-constant returns to scale imposes a measurement problem for the statistician who follows a non-parametric approach and only disposes of a set of price and quantity observations. Thus, to compute the scale elasticity empirically, the assumption was made in section (5e) above that there is no additional mark-up, i.e., $\mu = 1$.