OECD EXPERT WORKSHOP ON PRODUCTIVITY: INTERNATIONAL COMPARISON AND MEASUREMENT ISSUES

SUMMARY AND CONCLUSIONS

by

Paul Schreyer, Directorate for Science, Technology and Industry, OECD

Background

The Expert Workshop on "Productivity: International Comparison and Measurement Issues" was held under the auspices of the OECD Working Party No. 9 on Industrial Statistics and organised by the Economic Analysis and Statistics Division of the OECD Directorate for Science, Technology and Industry. The Workshop addressed delegate experts from Member countries as well as other experts invited by the OECD Secretariat.

The assessment of the performance of economies, industries or firms frequently involves the measurement of productivity trends and levels. Despite, or because of, the widespread use of productivity measures, there has been considerable debate about concepts of productivity measurement, about data requirements and about biases due to data constraints. Key theoretical and empirical issues associated with productivity measurement include the quality adjustment of price indices, the reduction of base-year biases of index numbers in the measurement of real output, the derivation of satisfactory conversion factors to compare industry output internationally, and productivity measurement in service industries. The objectives of the workshop were to examine some of these issues and:

- assess how they affect the international comparison of productivity growth rates and productivity levels in individual industries;
- provide an overview of recent developments, both in terms of the advancement of academic work and in terms of current practices by statistical agencies in Member countries;
- draw conclusions for possible statistical and analytical work at the international level.

Given the breadth of the topic "productivity measurement", the Workshop had to focus on a limited number of issues in order to provide a platform for fruitful discussions. These three issues were: *i)* the quality-adjustment of product- or industry-level deflators and the potential impact of this adjustment on measures of productivity; *ii)* the choice of appropriate index number formulae for the measurement of productivity over time and across space; *iii)* the choice of the appropriate conversion factor for international comparisons of levels of output and productivity.

These three specific issues were placed in a more general context during the first session of the Workshop which provided an overview of productivity measurement issues. The papers presented by van Ark and by Kuroda *et al.* show the impressive number of issues involved in productivity measurement. These include:

- The basic *choice between partial and total factor productivity measures*, which is typically a trade-off between measurability (which favours partial measures such as labour productivity) and theoretical appropriateness (which clearly suggests measures of total factor productivity).

- The appropriate *choice of the measure of output* (gross output, gross output adjusted for intra-industry flows, value-added, physical measures) and its aggregation over products or industries – an issue also taken up by Kuroda *et al.* who provide a coherent theoretical framework that embeds sectoral and aggregate productivity measurement in an input-output framework. The same topic is also picked up in the Workshop's session on index numbers.

- The *measurement of labour and capital inputs*. The latter has given rise to particularly broad coverage in theoretical and empirical research as the measurement of services from the capital stock in the production process is fraught with conceptual and measurement difficulties. A point in case is the treatment of quality changes in capital goods which carries over to measurement problems of real output.

- The *international comparison of productivity levels*, which entails a whole series of theoretical and practical questions on how to derive conversion factors for the comparison of figures of real output and productivity in a common currency, or which index numbers to use to ensure coherent aggregation over time, across space and between more than two countries.

Quality adjustment of price indices

The debate on quality-adjusted price indices was mainly triggered by the case of computers which experienced unparalleled jumps in performance, only inadequately reflected in traditional deflators. A substantial number of "hedonic" estimates has been developed to construct performance-related computer prices (for a survey, see the background paper by Triplett). From an international perspective, there are at least two interesting questions here: *i)* what can we learn from the experience of those countries who have actually developed hedonic computer price indices? and *ii)* do these adjustments matter for sectoral and aggregate productivity comparisons?

Country experience

On this first aspect, the Workshop provided an overview of the techniques for quality-adjustment of computer prices applied in three countries [United States (Triplett), France (Moreau) and Canada (Lowe)]. Quality adjustment is based on hedonic methods: in a simplified way, these methods have in common the use of regression analysis to estimate a hedonic function $p = h(X_1, X_2, .. X_N)$, relating observed prices *p* of computer models to quantities of the associated characteristics *X* such as speed or memory. This hedonic function permits to control for those price changes induced by quality changes (*i.e.* a change in the quantity of characteristics) and to separate them from residual price changes unexplained by shifts in characteristics. Although the precise techniques to specify and estimate the hedonic function and to derive quality-adjusted deflators vary between countries, several common conclusions arise:

- First, series of computer price indices based on hedonic functions vary dramatically from their counterparts based on traditional methods of deflation. For example, when Canada introduced hedonic measures for computers and reworked its historic series, an index which had doubled between 1971 and 1985 was replaced by one which had fallen below 20 (1971=100) (Lowe). Similar changes were observed in the United States and in France.

- Second, setting up and maintaining hedonic price measures for computers is resource intensive, in particular the field work of gathering primary observations on prices and characteristics. Given the costliness of constructing hedonic price measures, a natural question is whether indices developed in one country could be used elsewhere? It is difficult to assess whether the distortion created by the introduction of a hedonic measure conceived for a different country is smaller or larger than the distortion due to the use of a traditional rather than a hedonic index but it appears that for countries whose markets are closely integrated, the first bias might be much less severe than the second.

- Third, not all countries that have developed quality-adjusted price indices for computers have incorporated these new series in their national accounts framework. Currently, these indicators are only integrated into the national accounts in the United States and Canada.

Effects on sectoral and macroeconomic productivity measures

This second aspect was treated in the papers by Triplett and by Schreyer. Both start from the observation that, for productivity measures, it is insufficient to limit quality adjustment to measures of the output side – measures of inputs need equal consideration. For example, production processes for computer equipment use semiconductors as inputs – quality adjustment of the price index for computers alone would overstate productivity growth in the computer industry unless measures of semiconductor input are also quality adjusted.

The model presented by Triplett considers interrelated production processes involving three industries – computer equipment, semiconductors and semiconductor manufacturing equipment. Triplett then formulates the model in terms of product characteristics such as semiconductor or computer speed and assesses empirically the potential effects on price and productivity measures when both output and input measures are quality adjusted. To quote only one of the impressive examples: over the past 20 years, quality-adjusted prices of computer equipment declined at a rate of about 14 per cent a year. At the same time, semiconductor prices fell by a rate of more than 30 per cent a year. These 30 per cent have to be weighted by the cost share of semiconductors in the computer industry (a figure between 10 and 30 per cent) to obtain a measure of the percentage contribution of semiconductors to computer costs. Then, using the 10 per cent share, the ratio between output and input prices in the computer industry changed at least by 14 – 30*0.1 = 11 per cent (under the 30 per cent cost share, this figure amounts to 14 – 30*0.3 = 5 per cent). As the ratio between input and output prices is only a different way to express changes in relative productivity, it follows that productivity measures for the computer equipment industry are quite sensitive to introducing quality-adjusted measures of semiconductor input. A second point arising from Triplett's paper is that it shows how much faster semiconductor prices have fallen than computer prices.

While Triplett develops an inter-industry model that essentially deals with the effects on productivity measures of price changes of capital goods in three well-defined high-tech industries, Schreyer takes the stylised fact of mismeasurement of gross output in information and communication technology (ICT) industries as a starting point for his analysis to trace the impact of such mismeasurement on measures of real value added in other industries and at the aggregate level.

His assumption is that transactions and flows at current prices are correctly measured, but that their breakdown into price and quantity components is not, unless quality-changes are adequately accounted for, *e.g.* through hedonic functions. The analysis of the potential size of mismeasurement is conducted for five OECD countries and the period 1985-90. Using input-output tables it is demonstrated that, in the absence of quality adjustment of the price indices in ICT industries, the average annual growth rate of real business sector GDP is underestimated by between 0.2 and 0.6 percentage points, depending on country and aggregation method. Thus, neglecting quality adjustment of deflators in sectors that experience rapid quality change, may lead to non-negligible effects on measures of aggregate and inter-industry productivity.

Aggregation and comparison: Fixed-weight versus flexible-weight indices

Closely connected to the question of the effects of quality adjustment on price indices is the issue of selecting the right index number formula to aggregate across products or industries. The paper by Dean examines the role of index numbers of inputs and outputs in productivity measurement and, more generally, describes the US Bureau of Labor Statistics' (BLS) approach to sectoral and aggregate productivity measurement.

Traditionally, indices of real output have been based on the Laspeyres or Paasche formula with fixed base-year prices as weights. If prices change rapidly, as is for example the case for information and communication technology products, these weights become rapidly obsolete and can introduce significant biases in the measurement of prices and quantities. Superlative index numbers such as the Fisher Ideal Index or the Törnqvist Index avoid such problems as they allow for changing weights and are able to accommodate substitution in production. In 1995 BLS changed its index number method for industry output data to a changing-weight method, along with the Bureau of Economic Analysis (BEA) which published a Fisher-based index as the real GDP series. In BLS' productivity calculations, changing-weight indices are used for both outputs and inputs, as they have superior statistical qualities and are founded in microeconomic theory.

Dean demonstrates that the adaptation of changing-weight indices can produce substantial effects on trends in inputs, outputs and productivity. While over a very long period (1948-92), differences between methodologies are small, multifactor productivity measures for the early 1990s underwent a substantive downward revision (0.4 instead of 1.7 per cent annual growth rate in 1990-92 in the private business sector) under the improved methodology. These effects are consistent with the impact of using contemporaneous computer prices, in place of 1987 ones, to value output in 1990-92. Concerning international comparability, Dean points out, however, that the adoption of the new measures by BLS may have rendered US productivity data less comparable now than they formerly were to data for countries that use fixed-weight indices.

While the merits of flexible index numbers have been discussed in a context of inter-temporal productivity measurement, they are equally applicable to questions of inter-spatial comparisons – a point addressed in the paper by Fujikawa and Milana. The authors provide some theoretical foundations for the choice of index numbers and then deal with two specific issues: *i)* how to ensure transitivity in multilateral comparisons of productivity levels; and *ii)* how to assess direct and indirect productivity measures by means of input-output models. Problem *i)* can be solved by defining a hypothetical reference country, in fact a geometric average of all bilateral ratios. Fujikawa and Milana show this with a trilateral example United States-Germany-Japan for which they present empirical results. Question *ii)* rests on the observation that typical accounting procedures only consider direct input costs and leave "unexplained" those indirect input costs that are incorporated

into intermediate inputs. Fujikawa and Milana felt that this was a rather strong limitation of the analysis because technical progress and quality change are notoriously embedded in the goods and services used as inputs of production. They integrate flexible index numbers into an input-output framework which allows direct and indirect productivity effects for different industries to be disentangled in a comparison between the United States and Japan.

International comparisons of productivity levels

A recurrent issue in the comparison of output and productivity *levels* between countries is the choice of the accurate conversion factor to express figures of output or productivity in a common currency. Thus, if Q_i and Q_j denote the value of country *i*'s and country *j*'s output in local currencies, which price levels p_i and p_j should be applied to obtain a measure *R* that compares real output in the two countries?

$$R = \frac{Q_i}{Q_j} / \frac{p_i}{p_j}$$

At the industry level, the appropriate measures are relative output prices of the produced commodities or services. However, these are not normally readily available from official statistics. Therefore, empirical studies often use exchange rates or expenditure-based purchasing power parities for the entire economy as proxies with a considerable risk of arriving at biased productivity comparisons. Several papers (by O'Mahony, Hooper, Gersbach and van Ark) presented in the Workshop deal with the most frequently used approaches to improve on such crude proxies and to derive better measures of relative output prices: *i)* the expenditure purchasing power parity approach (EPPP); *ii)* the unit value ratio (UVR) approach; and *iii)* the use of physical measures.

The EPPP approach builds on the information available from well-established statistical work by the United Nations, Eurostat and OECD that compare prices of detailed final expenditure categories across countries. While detailed and carefully defined, the prices compared relate to consumer prices and to final goods only. The required industry prices should, however, relate to producer prices and include both intermediate and final goods. Thus, the strategy of the EPPP approach consists of correcting the measures of expenditure-related prices where possible to improve the approximation to industry producer price ratios.

The UVR approach is based on information in each country on the sales values and quantities of goods produced. Unit values are computed by dividing the value of production at the sub-industry level by physical output measures. Aggregation of the so-derived unit values to the industry level yields and comparison between countries yields unit value ratios (UVRs). One immediate advantage of UVRs is that they compare proxies of producer prices and are, in principle, able to include price ratios of intermediate goods.

Hooper describes the different steps in the process of deriving industry EPPPs from the initial consumer expenditure price comparisons: he adjusts initial expenditure PPPs for distribution margins, for net indirect taxes and for the effects of international trade. A very useful summary comparison of his measures with other approaches reveals large differences between results (Table 1).

O'Mahony develops a practical grid of criteria (Table 2) against which she assesses the two approaches. Her overall conclusion is that, since both approaches have advantages and drawbacks,

the most sensible procedure would be to use a mixture of the two, ideally combined with a third, independent price comparison.

Table 1. Output price levels
United States = 100

Method	Japan	Germany	France	Italy	United Kingdom	Canada
EPPPs	150	146	149	168	142	122
UVR	107	133	129	..	132	..
PPP for total GDP	154	155	135	110	101	105

Source: Hooper (1996).

Gersbach describes a project on the development of conversion factors that combines information gained from various methods, including the derivation of new international price ratios from specific surveys, the adjustment of existing product UVRs for differences in product mix and product quality on the basis of industry experts' opinions; the adjustment of expenditure PPPs for distribution margins and taxes; the adjustment of relative prices for price differentials of intermediate inputs and the derivation of implicit PPPs from comparisons of physical productivity measures.

Table 2. Properties of EPPPs and UVRs

Criterion	EPPPs	UVRs
Conceptually correct measurement = compatibility of price ratios with concept of output employed	Based on final consumption prices – correct only if output consists of final retail sales; for other sectors, adjustments necessary for distribution margins, net indirect taxes and international trade	UVRs measure producer prices
Coverage of sectors and coverage of products within a sector	Intermediate products excluded but those prices that are covered are based on large numbers of individual items with very detailed specifications	Intermediate products can be included but relatively small number of items covering only about 20% of manufactured output; matching problems in specific areas such as machinery
Intertemporal consistency	Not generally guaranteed – assessment shows a number of inconsistencies between development of price ratios in benchmark years and the trends implied by national price indices	Not generally guaranteed and difficult to assess but there is some evidence of greater stability of UVRs
Quality adjustment	Detailed specification allows better accounting for quality differences	Quality differences may be inadequately taken into account as only broad ranges of goods are matched
Index number properties	Multilateral schemes are used to ensure transitivity of bilateral results	Generally not transitive

Source: OECD, extracted from O'Mahony.

Conclusions

Daniel Malkin, Head of the EAS Division pointed out the importance of interaction between users and producers of data and indicators in permitting OECD to fulfil its role of monitoring and analysing structural change and industrial performance in Member countries. The Workshop, which brought together producers of data from statistical offices, country delegates and independent experts,

certainly advanced the dialogue in this respect. It also matched well with policy makers' requests for a better understanding of the relationship between microeconomic and macroeconomic issues. This fits in with an increasing convergence in the understanding of macroeconomic processes and with greater prominence of microeconomic performance measures such as industry and firm productivity and its determinants.

The choice of the quality adjustment of price indices, especially in relation to high-tech sectors and products, as a central issue in the Workshop, is at the heart of the concern of the Directorate for Science, Technology and Industry. Improved measurement techniques as implemented in the United States clearly constitute significant progress but also present a new challenge for international comparisons as the new methods may render US data less comparable – an issue that was discussed, among others, by Dean and van Ark. The conclusion is, of course, to incite other countries to follow suit in their effort to introduce new measures and index numbers. It was, however, recognised that any such introduction is a resource-intensive process with a substantial fixed-cost element that can more easily be borne by large countries than by small ones.

Another institutional question is how newly developed indicators and measures find their way into national accounts. There is a trade-off between preserving inter-temporal and international consistency of national accounts and exploiting the advantages of new measures, for example hedonic price indices.

The Workshop was also very useful in presenting the state of the art concerning conversion factors for international comparisons of levels of output and productivity. There seems to be agreement that there is no universal best practice in the field. Rather, the choice of one method or another depends on the nature of the question investigated and, more pragmatically, on data availability. However, certain avenues for fruitful further developments were identified, in particular the use of input-output tables to improve adjustments of expenditure-based purchasing power parities.

Malkin pointed out that increased sophistication in index numbers and measurement techniques cannot substitute for weaknesses in statistical groundwork for the collection of primary data on output, labour and capital input. For example, OECD databases still lack satisfactory data on hours worked by detailed industry – a major impediment to productivity measurement. Along the same vein, comparability between input and output data is not always ensured, with different variables stemming from different surveys that may prove difficult to match.

Future OECD work in the field could possibly evolve around three areas:

- The role of OECD as a clearinghouse for methodological issues provides a singular framework for the *exchange of views between government representatives and outside experts.* Further workshops may be useful, focusing, for example, on the question of productivity measurement in services.

- It should be possible to capitalise on progress made in some countries and provide a wider international dissemination, in co-operation with other parts of OECD and Eurostat, of information on *best practices in productivity measurement*. The improvement of international comparability of certain data series could be a possible useful side-effect.

- Possibilities should be explored to launch a *pilot activity for the development of industry-related PPPs* for international comparisons of levels of output and productivity. Any such activity would build on and complement ongoing work in OECD and in other institutions.

Annex

AGENDA OF THE WORKSHOP

Thursday, 2 May-1996

Opening Remarks by Thomas Andersson, Deputy Director, Directorate for Science, Technology and Industry, OECD

Session 1: Overview
Chair: Angus Maddison

Bart van Ark (University of Groningen, Netherlands)
Issues in Productivity Measurement: Statistical Problems and Policy Links

Masahiro Kuroda (Keio University, Japan), Kazuyuki Motohashi (Directorate for Science, Technology and Industry, OECD) and Shimpo Kazushige (Keio University, Japan)
Issues on the International Comparison of Productivity: Theory and Measurement
Discussant: René Durand (Statistics Canada)

Session 2: Quality Adjustment of Price and Quantity Indices
Chair: Masahiro Kuroda (Keio University, Japan)

Antoine Moreau (INSEE, France)
Méthodologie de l'indice des prix des micro-ordinateurs et des imprimantes en France
(Methodology of Computer and Printer Price Indices in France)

Jack E. Triplett (Bureau of Economic Analysis, United States)
Industry Productivity Measures and Hedonic Price Indexes: Do They Fit?

Robin Lowe (Statistics Canada)
Handling Quality Change in the Canadian National Accounts Price Deflators
Discussant: Anna Hallgren (Swedish National Board for Industrial and Technical Development)

Paul Schreyer (Directorate for Science, Technology and Industry, OECD)
Quality-adjustment of Price Indices in Information and Communication Technology Industries: Simulation of Effects on Measured Real Output in 5 OECD Countries
Discussant: Kiyoshi Fujikawa (Osaka University of Economics, Japan)

Friday, 3 May 1996

Session 3: Aggregation and Comparison: Fixed-weight *vs*. Flexible-weight Indices
Chair: René Durand (Statistics Canada)

Edwin R. Dean (Bureau of Labor Statistics, United States)
Productivity Measurement with Changing-weight Indices of Outputs and Inputs
Discussant: Bert Balk (Central Bureau of Statistics, Netherlands)

Kiyoshi Fujikawa (Osaka University of Economics, Japan) and Carlo Milana (Institute of Studies for Economic Planning, Italy)
Bilateral and Multilateral Comparisons of Productivity in Input-output Analysis Using Alternative Index Numbers
Discussant: Prasada Rao (University of New England, Australia, and Notre Dame University, United States)

Session 4: International Comparison of Productivity Levels
Chair: Prasada Rao (University of New England, Australia)

Mary O'Mahony (National Institute for Economic and Social Research, United Kingdom)
Conversion Factors in Relative Productivity Measurement: Theory and Practice
Discussant: Dirk Pilat (Department of Economics, OECD)

Peter Hooper (Board of Governors of the Federal Reserve Bank, United States)
Comparing Manufacturing Output Levels Among the Major Industrial Countries
Discussant: Bernd Görzig (German Institute for Economic Research, Germany)

Hans Gersbach (Universities of Heidelberg and Basel, Switzerland)
International Comparison of Productivity at the Industry Level
Discussant: Douglas Koszerek (Eurostat)

Conclusions by Daniel Malkin, Head of the Economic Analysis and Statistics Division, Directorate for Science, Technology and Industry, OECD