READER ON PRODUCTIVITY LEVELS

by Bart van Ark and Marcel Timmer
Groningen Groningen Growth and Development Centre
University of Groningen
(version October 2006)

Cover Note – October 2006

In the fall of 2002, the Statistical Working Party of the Committee on Industry and Business Environment at the OECD asked the Groningen Growth and Development Centre to prepare a reader on Productivity Level Measurement. The purpose of the reader was to serve as a complement to the OECD Productivity Manual (OECD, 2001a) which was primarily focused on measurement of productivity growth.

A first draft of the Reader was presented to an Expert Committee at the end of 2002. Some significant changes were suggested, and since then the completion of the project got postponed for several reasons. One reason were the very substantive changes in the methodology to measure industry-specific PPPs and extension of coverage during 2003 and 2004. Subsequently, the work became integrated in the EU KLEMS project on growth and productivity accounts.

Time has now arrived to revisit the original draft of the Reader, process the comments which were made at the time, and include an update on the measures of industry PPPs and productivity levels.

The present version is not yet complete:

- Chapters 1, 2 and 4 of the reader are included here. These chapters are fully updated and reflect the comments of experts and discussants at the earlier expert workshop as well as other comments received later. Chapter 2 will be also reconciled with the methodological section (section 2) from the paper by Inklaar and Timmer (item 26 on the agenda). Chapter 4 will also include the methodological sections (section 2) from the paper by Timmer, Ypma and van Ark (item 25 on the agenda)
- Chapter 3 (on basic statistical sources for output and input levels) is not yet included because it still need to be updated from 2002 in the light of the text in the OECD Compendium on Productivity Measurement (OECD, 2005) and other recent changes in national accounts and other relevant sources
- Chapter 5 (on sector-by-sector methodology on deriving PPPs) will be updated and merged with the new work on industry PPPs that is now presented separately in section 4 of the paper by Timmer, Ypma, and van Ark (item 25 on the agenda)
- Chapter 6 (on practical applications) will be updated and merged with the new work on value added-based productivity levels that is presented in sections 4 and 5 of the paper by Inklaar and Timmer (item 26 on the agenda) and the new EU KLEMS productivity levels.
- The concluding Chapter 7 will be added upon finalization of the project.

Participants in the *OECD Workshop on Productivity Analysis and Measurement* in Bern are invited to provide their comments on the present work. Following the workshop, the authors will discuss with OECd how to proceed with completing the final version and the publication of this work.

Groningen, 12 October 2006

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

1.1 Main Focus and Outline of the Reader

1. A considerable amount of work by the Statistical Working Party of the Committee for Industry and Business Environment (CIBE) over the past years has devoted attention to the appropriate methodology for the measurement of productivity. The work has been carried out in close co-operation with the Statistics Directorate (STD). It has covered methodological work which, for example, has led to the publication of the *OECD Productivity Manual* (OECD, 2001*a*). Efforts have also been undertaken to provide better guidelines for the measurement of certain data for productivity measurement, such as capital stock and capital services (OECD, 2001*b*). Moreover, OECD has also made efforts to provide up-to-date measures of productivity growth, notably in the context of a project on economic growth (OECD, 2001*c*). This has also led to the development of the *OECD Productivity Database* (Pilat and Schreyer, 2004), of which the major series were published and described in the *OECD Compendium of Productivity Indicators* (OECD, 2005). In recent years, much analyical work has focused on the sources of economic growth, including studies on the contribution of ICT-using and ICT-producing sectors to growth, the role of services, and the contribution of ICT investment. This analytical work hasinformed the policy discussion in CIBE and other OECD committees (OECD, 2004, 2006).

2. Since the release of OECD estimates of productivity *growth* in the OECD Productivity Database, more attention has turned to the measurement of productivity *levels*, since these serve as a yardstick of economic performance in many OECD countries. Estimates of productivity levels are potentially of great interest for policy analysis. They allow policy makers and business to benchmark the productivity performance of industries in their own country to that of industries in other countries. Relative productivity levels indicate the gaps countries face compared with the productivity leaders at aggregate or at industry level, and hence inform policy making in the light of catch-up and convergence. Comparisons of productivity levels may also help shed light on the relationship between productivity and competitiveness. It also strengthens the analysis of the locus of technical progress, in particular when supplemented by micro-oriented investigation of variance in performance between industries and between average and best practice firms. Finally, productivity level measurement may inform the debate on policy reforms that may be needed to enhance productivity performance.

3. So far only limited efforts have been undertaken to develop guidelines or best practices in the measurement of productivity *levels* across OECD countries. Over the years some work has been done at OECD (and formerly at OEEC) on this issue in the past (Paige and Bombach, 1959; OECD, 1996; Pilat, 1996; Scarpetta, *et al.*, 2000; Schreyer and Pilat, 2001, OECD, 2005), but a methodological and statistical framework for this work remains missing. Some statistical agencies publish estimates of comparative productivity, *e.g.* the US Bureau of Labor Statistics (BLS), Statistics Canada and the UK Office of National Statistics (ONS), Eurostat and the International Labour Organisation, as do some academic institutions, such as the Groningen Growth and Development Centre, and private institutions, such as the Conference Board (McGuckin and van Ark, 1999, 2005; van Ark, 2006).

3. The most useful analytical applications of productivity level measurement, however, are at the level of individual industries. Because of significant methodological and data-related problems, industry

level comparisons have so far been mostly carried out by researchers outside statistical agencies, primarily at academic and business research institutes. These include studies at the Groningen Growth and Development Centre (Maddison and van Ark, 1989, 2002; van Ark, 1993, 1996; Timmer, 2000, van Ark and Timmer, 2001), the National Institute of Economic and Social Research (Rostas, 1948; Smith, Hitchens and Davies, 1982; O'Mahony, 1992, 1999, 2002), Jorgenson and collaborators (Jorgenson 1995b), the Centre d'études prospectives et d'informations internationales (Freudenberg and Ünal-Kesenci, 1994; Nayman and Ünal-Kesenci, 2000) and the McKinsey Global Institute (1992, 1993).

3. The estimation of productivity levels raises several theoretical and practical problems which go beyond the construction of measures of productivity growth. A first issue concerns the theory of productivity comparisons. Chapter 2 discusses a range of methodological notions on concepts of productivity and bilateral and multilateral models of production for comparisons across countries. A second issue concerns the comparability of the data on output or value added, as well as labour and capital input at the aggregate and industry level, the internal consistency of these variables, and their comparability across countries. As these problems have also been extenseively addressed in, for example, OECD (2005), and are only briefly discussed in Chapter 3 of this Reader.

5. The focus of this reader is on the third major issue in productivity level measurement, which concerns the conversion of output and factor inputs, expressed at their own national prices and denominated in national currencies, into values at common prices denominated in a common currency. This is considerably more complicated at the industry level than at the aggregate level, since the purchasing power parities (PPPs) that are collected for the comparison of final expenditure across countries cannot be directly applied to output by economic activity. In the latter case, industry-specific production PPPs are required. The specifics of expenditure and production PPPs are discussed in Chapter 4, which also considers related issues on the multilateralisation of price comparisons across countries and the extrapolation of PPPs and real output and productivity comparisons over time. IT also introduces a framework, based on supply-use tables, that guides scholars who want to do industry-level comparisons in selecting expenditure or production PPPs for gross output. Chapter 5 concerns the practical data choices and problems that have to be confronted by researchers and statistical offices estimating PPPs and productivity levels by industry. These issues are are applied to the aggregate economy and to key areas of economic activity (agriculture, manufacturing and key service sectors). Chapter 6 discusses measures of PPPs for intermediate inputs, labour input and capital inputs. The final chapter, Chapter 7, concludes, draws some implications for further research and statistical analysis, and brings together the main recommendations of the reader.

4. This reader provides an overview of methods and sources as well as a range of illustrative examples of comparisons of productivity levels. It is both directed at generators of such estimates as well as at intensive users. Generators need to know the 'state of the art' and current best practice. They include mainly international statistical agencies, such as Eurostat, OECD and ILO, although some national statistical agencies may also with to be involved in such comparisons. Generators may also include dedicated economic research institutes, *e.g.* the Australian Productivity Commission, as well as academic research groups. The reader addresses this audience by introducing the key issues and options and providing references to follow up for greater detail. The amount of technical detail on production functions is limited, but extensive references are made to the underlying literature.

5. Intensive users of productivity levels include in the first place specialised government departments (including ministries of finance, economic affairs and social affairs) as well policy research institutes, private institutions such as investment banks, as well as academics. Intensive users require guidance on interpretation to ensure that measures are sufficiently sound and that strengths and weaknesses are clear. Other users include the media, consultants and the general public. Much of the detail in this

reader goes beyond their needs. These readers are referred to more analytical documentation by the OECD on the key indicators of productivity and their interpretation (OECD, 2004, 2005, 2006).

1.2 Limitations in Scope of the Reader

6. As in the OECD Productivity Manual, this reader primarily deals with the index-number approach to productivity measurement. In this approach, the main issue is how outputs and inputs for different countries are converted into a common currency in order to enable interspatial comparisons. Chapter 2 provides a brief overview of other approaches to productivity measurement, however, including econometric and programming approaches towards analysis of productivity frontiers.

9. Even more than in the case of the Productivity Manual, this reader on productivity levels is not prescriptive in the sense that it makes definitive choices in terms of which methods should, and which should not be used. By discussing the state of the art on the measurement of productivity levels, however, the more succesfull approaches are distinguished from less feasible methods. At many places, however, the choices also depend on the purpose of the comparisons, the level of industry aggregation, and the detail of information available.

10. The tables and figures in this reader are only illustrative, and do not provide an up-to-date overview of the most recent results from international comparisons of productivity levels. For this one is referred to several of the studies quoted in paragraph 2, as well as to several websites, such as the Penn World Tables at the Center for International Comparisons at the University of Pennsylvania (http://pwt.econ.upenn.edu/), the Groningen Growth and Development Centre of the University of Groningen (http://www.ggdc.net/index-dseries.html) and the EU KLEMS growth accounts database (http://www.euklems.net).

11. The reader also only pays limited attention to micro- or firm-level studies of productivity. Some studies have analysed data collected from a single establishment or a sample of comparable plants in the same industry. Another important line of research not addressed here is the use of longitudinal data on productivity performance using establishment data derived from business statistics, including industrial surveys and censuses. However, only a limited number of those studies deal with international comparisons of levels of productivity and efficiency. For example, so far the OECD work with firm-level data has not touched on comparisons of productivity levels (Scarpetta, et al. 2002; OECD, 2002*a*). There are some exceptions, however, one being several studies conducted by the National Institute of Economic and Social Research (Prais, 1995).

12. Another important contribution to international comparisons of productivity using company information are a range of studies carried out by the McKinsey Global Institute during the 1990s (Baily and Solow, 2001; Lewis, 2004; Baily and Kirkegaard, 2004). These studies have taken an approach that combines aggregate and sectoral estimates based on statistical work, as described in this reader, with measures based on company specific analysis in selected industries. The approach and results from these studies are discussed at several places in the reader.

13. This report focuses on *international* comparisons of productivity levels. However, the issues considered here are equally applicable to regional comparisons within countries. In particular, the frequent omission of using purchasing power parities to correct for differences in relative price levels between regions is an important source of distortion of regional measures of productivity. It should also be stressed that comparisons of productivity between countries within a monetary union, such as the Euro area, are still relevant. The adoption of a single currency obviously does not mean that relative price differences between countries have disappeared.

14. Readers should note that certain areas that are relevant to the measurement of productivity levels are closely linked to the measurement of productivity growth. Since the measurement of productivity growth was discussed in great detail in the *OECD Productivity Manual* (2001) and the *OECD Compendium of Productivity Indicators* (2005), these points will not be elaborated in great detail in this Reader. However, references to the Manual are included where appropriate.

CHAPTER 2: METHODOLOGY TO MEASURE COMPARATIVE PRODUCTIVITY LEVELS

2.1 Measures of productivity and their uses

15. This reader is about comparisons of productivity levels across space (countries or regions). Comparative productivity levels require ratios of output as well as ratios of inputs between two (or more) countries. The simplest way to obtain output ratios is to compare physical output measures, such as weight, volume or area. For example, when a labourer in the cement industry in country A produces 100 tons of cement per year on average compared to 200 tons per labourer in country B, 'physical labour productivity' in the cement industry in country A is half that of country B. Nowadays the use of physical units of output for productivity measurement is mostly restricted to benchmarking the efficiency of a particular production process for a specified product or for a closely related group of products over time. When comparing productivity at firm or industry level, the heterogeneity of output and the large variety of products makes the use of physical units irrelevant. Moreover it is often difficult to exclusively allocate inputs to one single output. In services, the use of physical units is often not at all possible.

16. In practice, one is more likely to have only access to figures on the total values rather than quantities of output and inputs. For comparisons of productivity levels across space, value measures need to be corrected for differences in relative prices between countries. This correction can be made by using purchasing power parities (PPPs), which specify the ratio of the price for a good or service, or for a bundle of goods and services, between two countries. The use of PPPs for comparisons of productivity levels is discussed in Section 2.2. and in considerable more detail in Chapter 4.

17. The literature distinguishes many different measures and concepts of productivity, each of which has its particular meaning and use. Broadly, productivity measures can be classified into single-factor productivity measures (relating a measure of output to a single measure of input) and multi-factor productivity measures (relating a measure of output to a bundle of inputs). Another distinction is between productivity measures that relate gross output to one or several inputs and those that use a value-added concept to capture movements of output (Table 2.1). The measures in Table 2.1 can be applied to calculate productivity growth rates, but also to the calculation of productivity levels.

Table 2.1. **Overview of the main productivity measures**

Type of output measure:	*Type of input measure*			
	Labour	***Capital***	***Capital & labour***	***Capital, labour & intermediate inputs (energy, materials, services)***
Gross output	Labour productivity (based on gross output)	Capital productivity (based on gross output)	Capital – labour MFP (based on gross output)	KLEMS multi-factor productivity
Value-added	Labour productivity (based on value-added)	Capital productivity (based on value-added)	Capital – labour MFP (based on value-added)	-
	Single factor productivity measures		***Multi-factor productivity (MFP) measures***	

Source: OECD Productivity Manual (OECD, 2001a).

2.1.1 Labour Productivity

18. The choice between the various measures depends in first instance on the focus and the purpose of the comparison. Most comparisons have focused on comparisons of value added per person employed. Combined with information for wages, this measure allows the comparison of unit labour cost, which shed light on the relationship between productivity and competitiveness (ILO, 1999, 2001). At the aggregate level, labour productivity measures can be combined with measures on the utilisation of the labour potential (*e.g.* employment-population rates) to provide measures of cross-country differentials in per capita income levels (Van Ark and McGuckin, 1999; Scarpetta *et al.* 2000, OECD, 2004, 2005). At industry level, gaps in labour productivity levels indicate the scope for further catch up relative to the productivity leaders. They capture differences in technology in its broadest sense including differences in capital intensity, capital quality, human capital, economies of scale and intangibles such as management techniques.

19. The choice of gross output-based over value added-based measures of labour productivity depends on the level of analysis and on the availability of data. For comparisons of single factor productivity at firm level, or at disaggregated industry level, gross output measures may be preferred. From the producer's perspective there is no inherent difference in the primary inputs (labour and capital) and intermediate inputs (materials, energy and service inputs). Production decisions are made for all inputs simultaneously and substitution between all inputs can take place, making them non-separable. At the firm level single factor productivity comparisons are often more easily made for gross output than for value added, because a precise measure of intermediate inputs is often difficult to obtain.

20. At more aggregate levels, the value added-measure of labour productivity is to be preferred. Gross output-based measures of labour productivity are quite sensitive to differences in the degree of vertical integration than value added-based measures. When countries differ considerably in their ratio of intermediate inputs to gross output, gross output-based measures of relative labour productivity levels are difficult to interpret. Consider the example in Table 2.2, which shows the input-output structure of an industry in two countries, A and B. Rows (1), (2a) and (2b) in the Table give the physical (or volume)

measures of three inputs (labour, materials and services).[1] Although the industry in country B uses only half of the volume of intermediate services input, it uses twice the labour input of country A. This can be due to more outsourcing of services such as cleaning, administration and bookkeeping in country A. The following rows show the volume measures of gross output and value added. Value added is calculated as gross output value minus the value of intermediate inputs. As can be seen from rows (6a) and (6b) the gross output-based labour productivity measures suggest a better result for the country which has the highest outsourcing intensity (country A), whereas in terms of value added, the productivity levels are the same. This is because outsourcing implies that less value is created by the firm itself and more is purchased as intermediate inputs.[2]

Table 2.2. **The effect of differences in vertical integration on volume measures of output, input and labour productivity**

		Country A volume (a) (euro)	Country B volume (a) (euro)	Relative level B/A
(1)	Labour input	10	20	200%
(2)	Intermediate inputs:	30	20	67%
(2a)	Materials input	10	10	100%
(2b)	Service input	20	10	50%
(3)	Total inputs [(1)+(2)]	40	40	100%
(4)	Gross output [(3)=(4)]	40	40	100%
(5)	Value added [(4)-(2)]	10	20	200%
	Labour productivity			
(6a)	gross output based [(4)/(1)]	4	2	50%
(6b)	value added based [(5)/(1)]	1	1	100%

(a) the volume measures represent PPP-adjusted output and input measures

21. Most international comparisons of labour productivity levels have focused on comparisons of value added per person employed. This is characteristic of the pioneering work by scholars like Rostas (1948), Maddison (1952), Paige and Bombach (1959) and Smith, Hitchens and Davies (1982). These studies used a mix of physical quantity and PPP-converted comparisons of value added. More recent work on comparative labour productivity measures, both at aggregate level (e.g. Maddison, 1995, 2001; Scarpetta, *et al.*, 2000; Capdevielle and Sherwood, 2002; Richardson, 2001) and at industry level, in

1. For the sake of argument we assume that all measures are PPP-adjusted, so that the measures are adjusted for differences in relative price levels between countries A and B. We also assume that there is no capital used in both countries.

2. Of course, this example is by construction a special case of outsourcing in which the volume of service inputs being outsourced equals the volume of labour inputs saved. Obviously, outsourcing can have beneficial effects on labour productivity even when measured on a value added-basis. The main point in the example, however, is that even in cases where there is no clear advantage of outsourcing in terms of total input use, the gross output based measure of labour productivity still indicates better performance in the country with a higher degree of outsourcing. See for an elaborated discussion, OECD (2001*a*, pp. 25-29).

particular for manufacturing (O'Mahony, 1992; Pilat, 1996; van Ark and Timmer, 2002) are almost exclusively based on PPP-converted comparisons of value added, but they differ in the type of PPPs that are used, which may be either expenditure or production based (Chapter 4). Section 2.2 introduces comparative levels of labour productivity, which is the mostly used partial productivity measure.

2.1.2 Multi Factor Productivity

22. Measures of multi-factor productivity (MFP) provide an indication of relative levels of the efficiency of input use between countries. MFP levels are measured as the difference in output between countries when differences in all inputs have been accounted for. Based on the economic theory of production, differences in MFP levels can be interpreted as differences in the level of disembodied technology. Technology is defined here narrowly as level differences in the production functions of countries. MFP does not, for example, include technology differences that are embodied in the use of capital goods. However, this interpretation of MFP levels is only true under a stringent set of assumptions including an identical production function (i.e., the same relative proportions of inputs for producing a unit of output), constant returns to scale, competitive markets and technical and allocative efficiency. Moreover, input measures must adequately reflect differences in adjustment costs, cyclical effects and input quality between countries (see OECD 2001*a*, Chapter 2, for further discussion). Finally, the interpretation from economic theory of MFP as a measure of technology is based on firms as decision-making units involved in revenue maximisation or cost minimisation. In analyses of higher level aggregates such as industries, sectors or aggregate economies, this interpretation only applies under very restrictive aggregation conditions. Hence MFP is often called a "measure of our ignorance" and differences in MFP levels should therefore be interpreted with care.

23. Although comparisons of labour productivity levels are not based on these stringent assumptions, MFP and labour productivity measures are not independent from each other. Multifactor productivity measures can be used to explain single factor productivity differences. For example, differences in labour productivity levels can be explained by differences in the ratio of capital to labour, and differences in multi-factor productivity. These, and other, links have been established with the help of the economic theory of production (see section 2.3 below and OECD (2001*a*) Chapter 2 for further discussion). In practice, single and multifactor productivity measures are close when one input accounts for a large part of the total value of inputs. For example, in some service sectors, where labour makes up the bulk of input costs, labour productivity measures will be close to measures of MFP. However, this is not to be expected in all cases.

24. For multi-factor productivity comparisons, in theory, gross output MFP measures are to be preferred over value added MFP measures at all levels of aggregation.[3] As technological change often affects all inputs, and not only the primary factor inputs labour and capital, gross output-based MFP levels better reflect differences in technology than value added-based ones. Value added-based measures provide an indication of the importance of the productivity differences in factor inputs for the economy as a whole. In principle, when the value added-based measure is obtained by independent comparisons of gross output and intermediate inputs with their own PPPs (so-called "double deflation), the data requirements for value added and gross output-based comparisons of productivity are the same. However, in practice, single deflated value added measures (using the PPPs for gross output) are often used in MFP comparisons due to lack of detailed price and quantity data on intermediate inputs (see section 2.2).

---

3. In a closed economy aggregate gross output MFP will equal value added MFP. This is because at the aggregate level gross sectoral output and value added are the same as intermediate input flows cancel out, in the absence of imports. See below for a discussion of the concept of sectoral output.

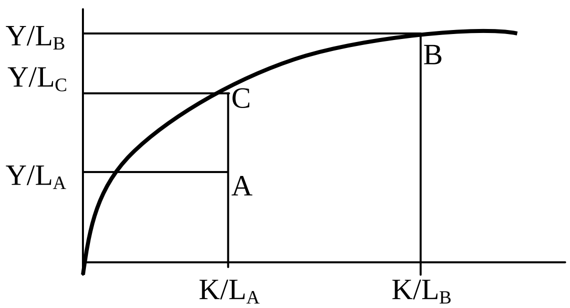
25. An additional problem in deriving meaningful gross-output based MFP comparisons at the industry level concerns the need for *sectoral output* measures, that is, output in an industry excluding the transactions that occur within the industry, *i.e.* intra-industry deliveries of intermediate products. The inclusion of intra-industry deliveries introduces double counting of industry output and makes productivity level measures dependent on the differences in the integration of different units or industries across countries. At the level of a single plant, gross output is the same as sectoral output. At the level of the entire economy, measures of sectoral output and of value added are close, although not identical in the presence of imported intermediate goods. At any other level of aggregation (industry or branch level), both measures will be different (see also discussion in OECD, 2001*a*, section 3.1.3). Although in a time-series perspective it might be expected that rates of growth of gross output and sectoral output do not differ greatly over short periods of time, there is less reason to expect such small differences for level comparisons, in particular between countries with different degrees of vertical integration. In practice, this problem can be exacerbated by differences in statistical practices between countries in measuring gross output. For example, in some countries output is measured at the level of a local unit or establishment (as in the U.S. Census of Manufactures), whereas in other countries (such as in Europe) a more integrated firm level or business unit concept is used. In the first case, the problem of double counting of output will be more severe. Consistent sectoral output measures across countries require intra-industry flows of intermediate products, which are difficult to derive empirically (see Chapter 3).

26. Multifactor-based productivity measures are well rooted in economic theory but due to their heavy data requirements (see Chapter 3) much less common in international comparisons of relative levels than single factor-based measures. Jorgenson and Yip (2001) provide relative MFP levels at the aggregate level for the G7 countries. At the industry level, interspatial comparisons of multifactor productivity at industry level are available for only a limited number of OECD countries. Dollar and Wolff (1992), van Ark and Pilat (1993), O'Mahony (1996) and O'Mahony and de Boer (2002) provide (single deflated) value added-based MFP comparisons. Applications of the gross output MFP methodology at industry level can be found in Jorgenson, Nishimizu and Kuroda (1987), Conrad and Jorgenson (1985) and, most recently Lee and Tang (2001) and Rao et al. (2004). In section 2.3.1 we provide a brief theoretical exposition of MFP-based productivity level comparison between two countries, which is followed by an extension towards multilateral comparisons in section 2.3.2. Even though a full-fledged level accounting framework puts a heavy demand on the data, a discussion of such models indicates the direction in which the research in this field might develop.

27. As mentioned above in Chapter 1, this reader focuses on the index number approach to productivity comparisons, because it is backed by theory (but under very stringent assumptions), because empirical implementation does not require as much information as other methods, and because it does not involve econometric estimations (see Box 2.1) which are usually not much applied by statistical agencies.

Box 2.1. **Frontier approaches to measuring multi factor productivity levels**

The index-number approach to productivity measurement described in this reader has a long history. More recently, other methods for productivity level comparisons have been suggested which can be used as an alternative and complement to the traditional approach. The biggest advantage of the index number approach is its easiness in calculation and the fact that only two observations of input-output combinations (including prices and quantities) are needed. However, this convenience comes at a cost which is not always fully appreciated. The interpretation of the multi factor productivity index as a measure of differences in disembodied technology presupposes the existence of competitive markets, such that prices reflect marginal costs and benefits, as well as full efficiency, including technical efficiency, allocative efficiency and scale efficiency. Put simply, allocative efficiency involves selecting the mix of inputs, which produce a given quantity of output at minimum costs. Scale efficiency indicates the maximum exploitation of economies of scale by producing at the (technically) optimal scale. Technical efficiency indicates that output is maximised given the level of inputs and the current state of technology. The assumption of technical efficiency in calculating MFP can be relaxed by using frontier methods. These methods do not require the assumption of cost minimisation or revenue maximisation and units (be it firms, industries or countries) can be technically inefficient. Production frontiers represent the maximum amount of output attainable from each input set, which need to be estimated. Frontiers allow for benchmarking a unit to best practice. If a unit operates below the production frontier it is said to be technically inefficient. To capture this idea consider the following figure:



Assume a simple world of one output (Y), two inputs capital (K) and labour (L) and constant returns to scale. The figure shows a production frontier, which indicates for each capital-labour ratio the maximum attainable labour productivity level. Two observations are provided, say countries A and B. Country B is on the frontier, indicating that it is technically efficient. Country A is operating below the frontier and is technically inefficient. Point C indicates the labour productivity level country A would attain when operating fully efficient. One way to decompose the difference in labour productivity levels of the two countries ($Y/L_B - Y/L_A$) is the following: one part of the difference is due to the difference in the amount of capital per worker ($Y/L_B - Y/L_C$), and the other part due to the level of technical efficiency ($Y/L_C - Y/L_A$).[4] As can be see, the actual decomposition of the labour productivity gap will depend on the particular shape of the production frontier.

Various approaches to the construction of production frontiers have been developed in the literature. They either belong to the 'econometric' approach or to the 'programming' approach. The econometric approach has a parametric nature and requires *a priori* specification of the functional form of the frontier (and the distribution of the inefficiency terms). In contrast, the programming approach is non-parametric and hence there is no need to specify a functional form. However, the biggest disadvantage of the programming approach is that it does not accommodate noise in the data. Initially frontier approaches were applied to firm-level data sets, but there is an increasing use of these techniques at the industry and aggregate level. Caves (1992) provides a good application of the stochastic approach for international comparisons of firm and industry performance. Färe *et al.* (1994) provide an application of the programming approach to international productivity measurement at the aggregate level. An additional advantage of frontier approaches is that, in contrast to the index-number approach, they do not require data on prices, but only information on input and output quantities. This can be particularly useful in the analysis of non-market activities were price information is missing.

4. Other decompositions are feasible as well. See Coelli, Rao and Battese (1998) for a non-technical introduction.

2.2 Labour productivity levels in international comparisons

28. The most commonly used single productivity measure for international comparisons of levels is labour productivity. This is generally defined as an output measure divided by a labour input measure. The labour input measure can be the number of persons employed, employees or hours worked (see section 3.2). The output measure can either be the physical quantity (or volume) of gross output or the volume of value added. For a binary comparison, that is a comparison between two countries., a gross output-based measure of labour productivity can be defined as follows. Let Y be a volume measure of gross output and Y^{AB} be a volume measure of gross output in country A relative to country B. Similarly, let L be a volume measure of labour input and L^{AB} be a volume measure of labour input in A relative to B. Then the volume measure of labour productivity in A relative to B based on gross output (LP_Y^{AB}) is given by:[5]

$$LP_Y^{AB} = Y^{AB} / L^{AB} \qquad (2.1)$$

29. The formula indicates that comparable volume measures of output and labour input for the two countries are needed. Labour is normally measured in physical units such as the number of workers or total hours worked. When a single output is being compared, physical measures, such as numbers of cars, are possible. However, when comparisons are made at the industry or aggregate level where output is not represented by a single product, output is given in value terms. In that case a correction for differences in relative price levels between countries A and B is needed. This is usually done with a purchasing power parity (PPP). A term "purchasing power parity" typically indicates relative levels of expenditure prices, but in practice is also used for relative levels of production prices between countries (see Chapter 4). Let P_Y be the price of a unit of gross output in country A, and PPP_Y^{BA} the price of gros output in country B relative to country A, then a volume measure of gross output in A relative to B (Y^{AB}) is given by:

$$Y^{AB} = \left(\frac{P_Y^A \times Y^A}{P_Y^B \times Y^B} \right) \times PPP_Y^{BA} \qquad (2.2)$$

30. As for gross output, value added-based volume measures of labour productivity also require a correction for differences in prices across countries. For value added two main approaches can be distinguished: single deflation and double deflation. According to the single deflation procedure, the PPP which is based on relative prices of gross output is used to obtain a volume measure of value added. If P_{va} is the defined (but unmeasured) price of value added, then a volume measure of single deflated value added in A relative to B ($VA(SD)^{AB}$) is given by:

$$VA(SD)^{AB} = \left(\frac{P_{VA}^A \times VA^A}{P_{VA}^B \times VA^B} \right) \times PPP_Y^{BA} \qquad (2.3)$$

and the single deflated relative labour productivity level ($LP_{VA\,(SD)}$) is given by:

$$LP_{VA(SD)}^{AB} = VA(SD)^{AB} \;/\; L^{AB} \qquad (2.4)$$

31. Single deflation has some important problems to it. It suffers from a terms-of-trade bias and a substitution bias. Firstly, the terms-of-trade bias arises when relative prices of output and intermediate inputs differ. For example, it might be the case that while the observed gross output prices in country A are

5. For simplicity, a time index t is omitted. All prices, quantities and values refer to the same point of time.

the same as in country B, the unobserved intermediate input prices are lower in A. In that case value added in A is overestimated relative to B, because the price gain from the lower input prices is now rel=flected in A's volume measures of value added. Secondly, differences in relative prices of primary factor inputs and intermediate inputs can lead to differences in the use of intermediate inputs, leading to substitution effects. Relatively lower intermediate input prices will lead to a higher use of intermediate inputs compared to the use of capital and labour inputs in producing output. When using single deflation,this substitution effect between intermediate inputs and value added is not reflected in the relative volume measure of value added. In practice, the terms-of-trade and the substitution biases are difficult to disentangle. The main point here is that as long as relative intermediate input prices do not move in tandem with relative output prices across countries, measures of single deflated value added will be biased. The double deflation procedure does not suffer from this bias.

32. The double deflation procedure provides a residual measure of the volume of value added in country A relative to country B, which is based on subtracting the volume index of intermediate inputs from the volume index of output for countries A and B. The volume index of intermediate inputs is based on a weighted average of the volume indices of intermediate products used by a given industry. An important issue here is to obtain the correct shares of intermediate inputs and value added in gross output. This can be done in various ways (see also OECD 2001*a*, p.102 *ff*). The two most frequently used procedures are the Laspeyres type index and the more general Törnqvist version. In the case of the Laspeyres index, the weights are expressed in values of the base country and enter the equation as arithmetic weights. In the Törnqvist formulation geometric weights are used which are based on average value shares in both countries. The Laspeyres type method can be used in cases of limited information on weights in countries other than the base country, but it leads to biases in the results as differences in substitution across countries are not taken into account. The Törnqvist double-deflated value added VA(DD) is given by:

$$\ln[\mathrm{VA(DD)}^{AB}] = \frac{1}{\bar{s}_{VA}}\left[\ln \mathrm{Y}^{AB} - \bar{s}_{M} \ln \mathrm{M}^{AB}\right] \quad (2.5)$$

where $\bar{s}_{VA}$ is the average value share of value added in gross output: $\bar{s}_{VA} = \frac{1}{2}(s^{A}_{VA} + s^{B}_{VA})$ where $s^{A}_{VA} = \frac{P^{A}_{Y} \times Y^{A} - P^{A}_{M} \times M^{A}}{P^{A}_{Y} \times Y^{A}}$ and similarly for B. $\bar{s}_{M}$ is the average value share of intermediate inputs in gross output in countries A and B; hence $\bar{s}_{M} = 1 - \bar{s}_{VA}$. M^{AB} indicates the volume index of intermediate input and is given by:

$$M^{AB} = \left(\frac{P^{A}_{M} \times M^{A}}{P^{B}_{M} \times M^{B}}\right) \times PPP^{BA}_{M} \quad (2.6)$$

with P_M and PPP_M as the price and PPP of intermediate inputs respectively. Double deflated relative labour productivity level ($LP_{VA\,(DD)}$) is given by:

$$\mathrm{LP}^{AB}_{VA(DD)} = \mathrm{VA(DD)}^{AB} \,/\, \mathrm{L}^{AB} \quad (2.7)$$

33. The use of double deflation procedures in international comparisons is in principle superior to the use of single deflation as it corrects for the terms of trade and substitution biases in the single deflation procedure (see above). However, in practice double deflation also has a number of problems to it. Firstly,

double deflation puts larger requirements on data, as besides PPPs for gross output, input PPPs are also needed. These are not usually readily available and must be constructed on the basis of input prices for the various countries in combination with input shares derived from input-output tables (see section 4.3). Secondly, double deflation introduces a new possible source of error not present in single deflation measures, *i.e.* errors associated with the measurement of input prices. In many sectors, material input prices are either unavailable or crudely measured. Hill (1971) suggests that the use of single deflation may be less misleading than using double deflation when material input prices are measured with error. This problem is aggravated by the fact that double deflated value added is defined as the output volume *minus* the intermediate input volume. A small percentage measurement error in the volume of gross output appears as a much larger percentage error in the volume of double deflated value added than is the case for the volume of single deflated value added .[6] One consequence of this is that the volume of double deflated value added, correcting for differences in price levels of output and inputs between two countries, may become negative even when nominal value added, which is not corrected for relative price differences, is positive.[7] Double deflation can also lead to measures of output that are unstable and erratic, especially at lower levels of aggregation when value added is a small share of gross output. This explains the ongoing popularity of single deflated value added in international comparisons of productivity levels, despite its theoretical inferiority.

2.3 Multi factor productivity levels in international comparisons

2.3.1 Bilateral MFP measures

34. The most commonly used measure of relative MFP levels is based on the volume index of value added. As for the case of labour productivity comparisons, the value added index can be based either on a single or a double deflation procedure. At the aggregate level, sectoral output equals value added so only single deflation is needed.[8] As before, $VA(SD)^{AB}$ is an index of the volume of value added in country A relative to country B using single deflation, and L^{AB} and K^{AB} are volume indices of labour and capital input respectively. Then under the assumption of constant returns to scale, the relative MFP level, $MFP^{AB}_{VA(SD)}$, can be written as

$$MFP^{AB}_{VA(SD)} = \frac{VA(SD)^{AB}}{(L^{AB})^{\bar{s}_L}(K^{AB})^{(1-\bar{s}_L)}} \qquad (2.8)$$

or by taking natural logarithms

6. Indeed if X is defined as the ratio of gross output to value added, the variance in the volume of value added is the sum of X^2 times the variance in the volume of gross output and (1- X^2) times the variance in the volume of intermediate inputs, suggesting larger variance for the double deflated volume of value added than single deflated volume of value added (Hill, 1971, p. 19).

7. When use is made of a Laspeyres-type deflation procedure, a negative volume of value added can even arise without measurement errors in either input or output volumes (see also OECD, 2001*a*, pp. 31 ff.) This is not necessarily problematic. It has a clear practical economic interpretation and the indicates inefficiency in the production process as an industry runs down its net worth, even at zero factor payments, when expressed in prices of another country.

8. Except for imported intermediate inputs, although few studies have addressed this issue in interspatial comparisons.

$$\ln\left[MFP^{AB}_{VA(SD)}\right] = \ln\left[VA(SD)^{AB} - \bar{s}_L \ln L^{AB} - (1-\bar{s}_L)\ln K^{AB}\right] \quad (2.9)$$

with $\bar{s}_L$ denoting the average value share of labour in value added at current national prices, $\bar{s}_L = \frac{1}{2}(s^A_L + s^B_L)$. Alternatively, the value labour share of one of either countries can be used, but this would not allow for differences in substitution elasticities across the two countries.

35. Gross output based MFP measures have been developed by Jorgenson and associates (Jorgenson, 1995) within the context of (bilateral) production models (see the appendix to this Chapter for a full exposition of this model). Based on the economic theory of production the difference in MFP levels between countries A and B can be expressed as the output volume index minus a weighted average of input volume indices known as the Törnqvist index:

$$\ln\left[MFP^{AB}_{Y}\right] = \ln\left[Y^{AB} - \bar{s}_L \ln L^{AB} - \bar{s}_K \ln K^{AB} - \bar{s}_M \ln M^{AB}\right] \quad (2.10)$$

with $\bar{s}_L$ denoting a simple average of the shares of labour in total input compensation: $\bar{s}_L = \frac{1}{2}(s^A_L + s^B_L)$ where $s^A_L = \frac{P^A_L \times L^A}{P^A_Y \times Y^A}$ and similarly for country B. The average compensation shares for capital input $\bar{s}_K$ and intermediate input $\bar{s}_M$ are defined correspondingly.

36. Formula (2.10) presents the derivation of gross output MFP based on the primal level accounting methodology. As an alternative, comparative estimates of multifactor productivity may be obtained on the basis of a dual or price function approach (see, *e.g.* Jorgenson and Kuroda, 1990; Lee and Tang, 2001). The point of departure in this approach is not the relative volume of output and inputs, but their relative prices. Analogue to the primary approach, differences in MFP levels between the two countries can be derived as the inverse of the relative output prices (PPP_Y) between countries A and B plus a weighted average of relative input prices, each weighted by their share in the gross output value:

$$\ln MFP^{AB}_{Y} = -\ln PPP^{AB}_{Y} + \bar{s}_L \ln PPP^{AB}_{L} + \bar{s}_K \ln PPP^{AB}_{K} + \bar{s}_M \ln PPP^{AB}_{M} \quad (2.11)$$

where PPP_Y^{BA} indicates relative price levels of gross output in country A relative to country B and similarly for labour, capital and intermediate input. These various PPP measures are discussed in more detail in Chapter 4 of this reader. Productivity level estimates based on the dual method can sometimes be obtained more easily, as input price data is often more abundant than information on input quantities. However, when applied to the same data set of prices and quantities, the primary and dual approach provide identical estimates of productivity.

37. Although value added-based and gross output-based MFP measures reflect different concepts, there is a direct theoretical relationship between the two MFP measures. For two countries with identical value added-gross output ratios, the difference in value-added based MFP equals the difference in gross-output MFP, multiplied by the inverse of the nominal share of value added in gross output (see OECD 2001*a*, pp 23 *ff.*). However, value added shares are generally not the same across countries and an adjustment reflecting the difference in the value added-gross output ratios is needed (see Baily, 1986 for the intertemporal case).

38. As discussed before, MFP measures can be used to explain differences in labour productivity levels. For example, the relationship between the gross-output based labour productivity and gross-output multifactor productivity measures is given by

$$\ln\left[LP_Y^{AB}\right] = \bar{s}_K \ln\left[k^{AB}\right] + \bar{s}_M \ln\left[m^{AB}\right] + \ln\left[MFP_Y^{AB}\right] \quad (2.12)$$

where lower case *k* and *m* indicate capital and intermediate input per unit of labour respectively. Similar relationships can be derived for single and double deflated value added-based productivity measures.

39. So far, this reader has dealt with labour, capital and intermediate deliveries as being homogeneous inputs. However, normally these input measures are aggregations over components, such as different asset types (for example, machinery and structures), different skill levels (high, intermediate and low educated workers), and different types of material and service inputs. If the aggregate capital, labour and material input volume indicators are translog functions of their components, the differences between logarithms of aggregate input volume indices for the two countries can be expressed as weighted differences between logarithms of the input volume indices of all components:

$$\ln\frac{K^B}{K^A} = \sum_k \bar{s}_k \ln\frac{K_k^B}{K_k^A}, \qquad (k = 1,2,...,r) \qquad (2.13)$$

$$\ln\frac{L^B}{L^A} = \sum_l \bar{s}_l \ln\frac{L_l^B}{L_l^A}, \qquad (l = 1,2,...,s) \qquad (2.14)$$

$$\ln\frac{M^B}{M^A} = \sum_m \bar{s}_m \ln\frac{M_m^B}{M_m^A}, \qquad (m = 1,2,...,u) \qquad (2.15)$$

with $\bar{s}_k$, the average of capital component *k* in total input compensation value and similarly for labour types *l* and intermediate input types *m*. *r* is the total number of capital asset types *k*, *s* the total number of labour types *l,* and *u* the total number of intermediate input types *m*.[9] These indices are referred to as the translog volume indices of relative capital, labour and intermediate inputs.[10]

2.3.2 Multilateral MFP measures

40. In contrast to the case of intertemporal comparisons for which the observations are a naturally ordered sequence, transitivity is an important requirement for spatial comparisons. An index I is transitive when a direct comparison between any pair of countries (A and B) gives the same result as indirect comparisons through any other country C, or $I^{AB} = I^{AC}/I^{BC}$. Caves, Christensen and Diewert (1982) extended the bilateral model of production to a multilateral setting, deriving translog multilateral output, input and productivity volume indices which satisfy the transitivity requirement. They define translog multilateral output, input and productivity indices by comparing each country strictly to a hypothetical

9. By definition, the sum of compensation shares of all capital types, s_k, is equal to the compensation share of aggregate capital s_K. The same holds true for labour and intermediate inputs.

10. Translog output aggregates can also be developed along the same lines, but we restrict ourselves to the case of one output for expository simplicity.

representative country, which is a simple average of all countries considered. In the case of one output and *n* inputs, the translog multilateral productivity index between country A and B is then defined as:

$$\ln\left[MFP^{AB}\right] = \ln\left[\frac{Y^B}{Y^A}\right] - \frac{1}{2}\sum_n (v_n^B + \bar{v}_n)\ln\left[\frac{X^B}{\bar{X}}\right] + \frac{1}{2}\sum_n (v_n^A + \bar{v}_n)\ln\left[\frac{X^A}{\bar{X}}\right] \qquad (2.16)$$

where Y is output and X is any type of input, and a bar indicates the arithmetic mean over the S countries (s = 1,.., S) of the variable under the bar and *n* the number of inputs, $\bar{v}_n = \frac{1}{S}\sum_{s=1}^{S} v_n^s$ with v^s_n the share of input *n* in total input compensation of country *s*.[11]

41. The multilateral production index is base-country invariant, which implies that output and inputs can be expressed as an index of any country *s*. This index sacrifices only a small amount of characteristicity and imposes no *a priori* restrictions on the structure of production. It has an intuitive interpretation as a comparison between two countries is obtained by first comparing each country with the "average" country and then comparing the differences in country productivity relative to this average country. Note that the imposition of the transitivity condition is also a constraint. Comparisons between two countries A and B are influenced by all other countries in the analysis, which implies that the results will change when new countries are added.

2.4 Aggregation of multi factor productivity levels

42. The link between industry-level and aggregate productivity level differences can be of considerable importance if one is interested in the contribution of individual industries to overall productivity level differences. The precise form of the aggregation procedure depends on the productivity measure used. When productivity is measured on a value added-basis, aggregation is straightforward. It involves weighting relative MPF levels by the industry's average value share in total value added, where the share is averaged over the two countries. As noted before, the components of this weighted average, *i.e.* value added-based productivity differences by industry, are in general not a valid representation of differences in the level of disembodied technology in individual industries.

43. Aggregation of differences in industry-level gross output based MFP measures is less straightforward. One possibility is to separately aggregate the gross output-based, value added-based and intermediate inputs-based comparisons by industry. However, this leads to inconsistent results between industry and aggregate level comparisons because the shares of intermediate inputs in output typically differ across countries **[not clear]**. Instead summation weights need to be based on each industry's gross output relative to economy-wide value added. Unlike the weights used for aggregation of value added-based MFP, these aggregation weights (known as Domar-weights) do not sum to unity, but exceed it. This is due to the links that exists between industries within a country through flows of intermediate inputs **[not clear]**. The flow of intermediate inputs allows productivity differences in successive industries to augment each other. Forming MFP measures at higher levels of aggregation entails integration of individual industries into increasingly larger units **[not clear]**. Output and inputs of the integrated industry should be measured as *sectoral output* and *sectoral input*, that is net of all intra-industry deliveries. Only in this case will aggregate MFP measures be similar to the weighted industry-level MFP measures. In the case of a

11. This index has been applied for example in Christensen, Cummings and Jorgenson (1981). For a generalisation of this approach, see Diewert (1986) and Fujikama and Milana (1996).

closed economy, sectoral output at the most aggregate level is identical to total value added. Moreover sectoral input equals total primary inputs as all intermediate input flows become intra-industry flows and should be excluded. For an elaborated discussion of these issues, see OECD (2001*a*, Chapter 8). However, for comparisons involving small, open economies, imported intermediate inputs can be significant. These are not intra-industry flows, even at the aggregate level, and should be taken into account in calculating aggregate MFP measures (Morrison and Diewert 1990).

Box 2.2.**Conclusions and recommendations**

Concepts: Most international comparative measures of productivity deal with labour productivity. Gross output per person employed or per hour worked may be the preferable measure at the firm or very detailed industry level. At higher aggregation levels, the use of gross output leads to double counting of industry output that is used for intermediate inputs. Value added is then the preferable output measure. For comparisons of levels of multi factor productivity (MFP) at the industry level, gross output is the preferable concept, as it simultaneously treats intermediate inputs and primary factor inputs. Nevertheless, due to lack of data on intermediate inputs, value added-based MFP measures are much more common. In a comparison between closed economies, overall MFP levels based on either gross output or value added will be the same.

Value added-based measures: The use of value added requires the choice to either deflate value added with a purchasing power parity (PPP) that reflects relative price levels for gross output (single deflation), or alternatively to separately deflate total gross output and intermediate inputs (double deflation). Single deflation suffers from the terms-of-trade bias as long as relative intermediate input prices do not move in tandem with relative output prices. However, lack of data on intermediate input prices and serious measurement issues often force the use of single deflation. For value added-based comparisons at the aggregate level in closed economies, single deflation will be sufficient as there are no intermediate input flows. These issues will be dealt with in more detail in Chapter 4.

Gross output-based measures: The optimal approach in a productivity level comparisons programme would be to use a full breakdown of output for all labour, capital, and intermediate inputs, and develop separate PPPs for output and for each input category. Comparisons of levels of labour productivity, capital productivity and multi factor productivity can then be shown separately, but in a consistent manner. Obviously a cross country comparison using full fledged production models puts a high demand on the data, as it requires PPPs for each of those categories (see Chapter 4), detailed and consistent information on outputs and input flows by industry, and the derivation of sectoral output and input measures (see section 3.1). In addition, MFP measures require specific assumptions concerning the remuneration of each of the inputs, which should be consistent with the production model used. Such assumptions need to be clearly spelt out in any cross-country comparison of MFP.

Aggregation: Value added-based measures can be aggregated by using industry shares in total value added as weights. The optimal approach for gross output-based comparisons is a Domar-like weighting system that uses each industry's sectoral output relative to total value added as in the intertemporal case.

Appendix 2A: Bilateral models of production

44. The methodology for productivity measurement has been developed most rigorously by Jorgenson and associates (Jorgenson, 1995) who have rooted it in the economic theory of production. The latter provides a firm basis for measuring productivity growth, but so far a theory of production for interspatial comparisons is lacking. Common practice therefore is to transform the methodology for measuring productivity growth directly into one for measuring productivity levels. Following this approach, instead of comparing two points in time, two countries are now being compared. Jorgenson and Nishimizu (1978) provide an exposition of this methodology on a value-added basis, which was extended by Jorgenson, Kuroda and Nishimizu (1987) to incorporate intermediate inputs as well. For the gross output added based MFP measure, consider the following production function at time t:

$$Y = f_t(K, L, M) \qquad (A.1)$$

where Y is gross output, K is capital input, L is labour input, and M is intermediate inputs. For a comparison between two countries A and B at a particular point in time the following translog production function can be formulated:[12]

$$\ln Y = \alpha_0 + \alpha_K \ln K + \alpha_L \ln L + \alpha_M \ln M + \alpha_D D + \frac{1}{2}\beta_{KK}(\ln K)^2 + \beta_{KL} \ln K \ln L$$
$$+ \beta_{KM} \ln K \ln M + \beta_{KD}(\ln K)D + \frac{1}{2}\beta_{LL}(\ln L)^2 + \beta_{LM} \ln L \ln M^{AB} + \beta_{LD}(\ln L)D$$
$$+ \frac{1}{2}\beta_{MM}(\ln M)^2 + \beta_{MD}(\ln M)D + \frac{1}{2}\beta_{DD}D^2 \qquad (A.2)$$

where D is a dummy variable equal to one for the base country (A) and zero for the other country (B). The production process is assumed to have common properties in both countries, but the dummy variable catches the country specific deviations. The difference in technology between the two countries can be defined as the logarithmic difference between levels of output of both countries, holding capital, labour and intermediate inputs constant:

$$v_D = \frac{\partial \ln Y}{\partial D}(K, L, M, D) = \alpha_D + \beta_{KD} \ln K \ln D + \beta_{LD} \ln L \ \ln D + \beta_{MD} \ln M \ln D + \beta_{DD} D \qquad (A.3)$$

45. In order to estimate differences in productivity levels between the two countries, the production function is combined with necessary conditions for producer equilibrium. These conditions state the equality between the share of each input in the value of output (v_L and v_K) and the elasticity of output with respect to that input:

$$v_K = \frac{\partial \ln Y}{\partial \ln K}(K, L, M, D) = \alpha_K + \beta_{KK} \ln K + \beta_{KL} \ln L + \beta_{KM} \ln M + \beta_{KD} D \qquad (A.4)$$

$$v_L = \frac{\partial \ln Y}{\partial \ln L}(K, L, M, D) = \alpha_L + \beta_{KL} \ln K + \beta_{LL} \ln L + \beta_{LM} \ln M + \beta_{LD} D \qquad (A.5)$$

and

$$v_M = \frac{\partial \ln Y}{\partial \ln M}(K, L, M, D) = \alpha_M + \beta_{KM} \ln K + \beta_{LM} \ln L + \beta_{MM} \ln M + \beta_{MD} D \qquad (A.6)$$

with $v_K = \frac{p_K K}{p_Y Y}$, $v_L = \frac{p_L L}{p_Y Y}$ and $v_M = \frac{p_M M}{p_Y Y}$ and p_K, p_L, p_M and p_Y denoting prices of capital, labour, intermediate inputs and output respectively. Assuming constant returns to scale, the sum of the elasticities with respect to all inputs is equal to unity, so that the value shares also sum to unity.

12. Subscripts for industry are suppressed for expositional ease.

46. Given these assumptions, the difference in technology level v_D between countries A and B can be expressed as the relative output quantity levels less a weighted average of relative input quantities known as the Törnqvist index:

$$\hat{v}_D^{AB} = \ln \frac{Y^B}{Y^A} - \hat{v}_K \ln \frac{K^B}{K^A} - \hat{v}_L \ln \frac{L^B}{L^A} - \hat{v}_M \ln \frac{M^B}{M^A} \qquad \text{(A.7)}$$

with $\hat{v}_K$, $\hat{v}_L$ and $\hat{v}_M$ denoting simple average of shares in total input compensation, $\hat{v}_K = \frac{1}{2}[v_K^A + v_K^B]$, $\hat{v}_L = \frac{1}{2}[v_L^A + v_L^B]$ and $\hat{v}_M = \frac{1}{2}[v_M^A + v_M^B]$. The expression for the average difference in technology level, $\hat{v}_D^{AB}$, is referred to as the translog index of difference in technology. The Törnqvist index is a superlative index number formulae which means that it is exact for a functional form, in this case the translog, which provides a second order approximation to an arbitrary aggregator function (Diewert 1976). Hence it can handle all kinds of different formulation of technical change such as different combinations of input-augmenting technical change.

CHAPTER III: DATA SOURCES AND COMPARABILITY ISSUES (NOT INCLUDED)

3.1 Comparing output and intermediate inputs

3.1.1 National accounts and input-output tables

3.1.2 Industrial surveys and censuses

3.2 Comparing labour input

3.2.1 Employment

3.2.2 Adjusting for hours worked

3.2.3 Adjusting for the composition of the labour force

3.3 Comparing capital input

3.4 Compensation for labour, capital and intermediate inputs

CHAPTER IV: PURCHASING POWER PARITIES ACCORDING TO EXPENDITURE AND PRODUCTION APPROACH

7. For comparisons of productivity levels across countries (or regions) one will mostly be interested in comparisons of physical (or volume) productivity. In practice, however, the basic data that can be relied upon for comparisons of productivity levels are not physical quantities of output, but value measures of output and inputs, such as the value of gross output, value added and intermediate inputs. There will be a difference between the measure of volume and value productivity when the relative price level between countries differ. For example, while the average value productivity of a hairdresser in Mexico (for example, turnover per working day in US$ using the peso/$ currency exchange rate) may be 25 per cent of that of his colleague in the U.S., his physical productivity (for example, the number of haircuts per day) may be as much as 75 per cent compared to the U.S. This is because the price of a haircut in Mexico is only one third of that in the U.S. For comparisons of productivity levels across space, value measures need to be corrected for these differences in relative prices between countries. This correction can be made by using purchasing power parities (PPPs).

56. A purchasing power parity (PPP) is defined as the ratio of the price of a product or a bundle of products between two countries, with prices expressed in each country's own currency.[13] The relative price level is defined as the (average) price of one country relative to the (average) price of the other country, with prices expressed in a common currency. When countries have different currencies, the relative price level is obtained as the ratio of the PPP to the currency exchange rate. So the relative price level of a haircut in Mexico compared to the U.S. is obtained by comparing the PPP of the haircut (for example, 33 pesos in Mexico to 10 US$ in the U.S.) to the currency exchange rate (for example, 10 pesos to one US$). The relative price level of Mexico relative to the U.S. is then (3.3/10 =) 33 per cent. When two countries have the same currency, for example, the euro, the relative price level can be directly derived from the PPP. For example, when the ex-factory price of a ton of flat steel of identical quality is 2,000 euro in Portugal against 2,500 euro in Germany, the Portuguese price level is 80 per cent of that in Germany. As a result the output and productivity in flat steel manufacturing in Portugal relative to the Germany would be understated by (2,500/2,000 –1 =) 25 per cent if directly compared without the PPP adjustment.

57. The previous discussion implies that when using the currency exchange rate for the purpose of output and productivity comparisons, one does not adjust for international price differences between countries. The comparison is then a value comparison, not a volume comparison. Only when the PPP equals the currency exchange rate, value productivity equals volume productivity. Similarly, for a comparison between countries or regions using the same currency, an adjustment for relative price

---

[13] A few terminological issues are important to keep in mind for the remainder of this chapter. 'Product' may refer to a good or a service. 'Country' may also refer to a region (for example, an U.S. state) or a group of countries (for example, the European Union) as the PPP adjustment will be necessary for any comparison of geographical units when relative price levels differ. Finally, the term 'purchasing power parity' essentially refers to comparisons of expenditure prices, indicating that with a unit of purchasing power, say one U.S. dollar, one can purchase the same bundle of products in one or more countries. Although this terminology may be less appropriate for comparisons of output or income, it is in practice used in more general terms for any comparison of relative prices across space.

differences is required to obtain a volume comparison. Refraining from an adjustment for relative prices (for example, by using the currency exchange rate instead of the PPP) will be particularly a problem when comparisons are made at the level of industries or sectors of the economy. Even in sectors which are open to international trade, and for which it may be assumed that relative price levels converge to one in the long run (that is, the PPP equals the exchange rate), different degrees of monopoly power, lags in response to exchange rate movements, etc., make this assumption unlikely to be fulfilled in practice. For non-traded sectors, there is in fact no reason at all to suppose price ratios will equal the exchange rate. Finally, it is well recognised that the exchange rate is also influenced by short-term capital movements which should not be reflected in comparative volume measures of output and productivity (Taylor and Taylor 2004).

58. Parallel to the distinction made in the national accounts framework for deflating current expenditure, output and income over time, there are essentially three methods to construct PPPs for deflation across space.[14] The first method, the expenditure approach, consists of constructing expenditure-based purchasing power parities, which provide PPPs for expenditure categories, such as private consumption, gross fixed capital formation and government consumption. Expenditure PPPs refer to purchase prices for final expenditure categories (food, clothing, rent, machinery, education, etc.). Alternatively one can develop purchasing power parities on the basis of the production (or industry-of-origin) approach, hereafter to be named production PPPs. These PPPs are based on prices faced by the producers and provide PPPs for gross output and intermediate inputs for individual industries (see table 4.1). Finally, one can develop PPPs for comparisons of factor income, that is, for income from labour and capital. Income PPPs are not extensively used in international comparisons and will not be discussed further in this reader.

Table 4.1 PPP concepts for output and productivity comparisons

Expenditure approach	***Production approach***
PPP of private consumption (C)	PPP of gross output (GO)
PPP of gross fixed capital formation (I)	PPP of intermediate inputs (II)
PPP of government consumption (G)	PPP of value added (VA)
PPP for trade balance (X-M)	

59. Aggregate comparisons of productivity levels for the total economy are usually based on gross domestic product (GDP) and deflated on the basis of expenditure-based purchasing power parities. Expenditure PPPs have originally been compiled within the framework of the International Comparisons Project (ICP) for the purpose of comparing per capita expenditure patterns on private and government consumption and fixed capital formation. Nowadays expenditure PPPs are based on a regular survey which is specifically designed for this purpose, and carried out under the auspices of Eurostat and the OECD. Expenditure PPPs are primarily used for comparative studies of GDP, per capita income and expenditure, but OECD and Eurostat also publish measures of labour productivity based on expenditure PPPs. This is also done by several academic outlets such as the Penn World Tables (Summers and Heston 1991) and the

---

[27] In the remainder of this chapter we apply, for the sake of convenience, the term 'deflation' also to comparisons across space, instead of something like 'adjustment for differences in relative prices', even though the former of course primarily relates to comparisons over time. The use of the term 'deflation' indicates that any comparison of output and productivity, either over time or across space, requires an adjustment for changes or differences in prices.

Total Economy Database of the Groningen Growth and Development Centre (GGDC) and The Conference Board.[15]

60. For comparisons of productivity levels by industry, one ideally requires production PPPs by industry for industry level comparisons. Production PPPs usually relate to the level of gross output. When making comparisons of levels of value added, which is a common output concept for productivity studies, one must either develop separate PPPs for gross output and intermediate inputs from which the PPP for value added can be implicitly derived, or one must assume that the average PPP for gross output at industry level equals the PPP for intermediate inputs. In the latter case value added is based on, what is called, single deflation, whereas in the former case one obtains a double-deflated value added-measure.[16]

61. As production PPPs are not as readily available as expenditure PPPs, some scholars have chosen to proxy the former by using the latter. Expenditure PPPs for aggregate GDP can only be used at industry level if one applies the very strict assumption of identical comparative price levels across industries. Other studies have resorted to using selective expenditure PPPs, which are based on an allocation of PPPs for specific expenditure items to individual industries. For example, the expenditure PPP for clothing would then be applied to output in the wearing apparel industry. What is required here, however, is an adjustment of the PPPs from expenditure to producer price level using information on taxes and transport and distribution margins. Just as for PPPs for gross output, intermediate input PPPs can also be based on expenditure or production PPPs.

62. The present chapter proceeds as follows. Section 4.1 briefly describes how expenditure PPPs have been used for comparisons of output and productivity at the aggregate level. Section 4.2 discusses how expenditure PPPs can be adjusted to producer price level by taking account of differences in margins and taxes. Section 4.3 goes into the methods used to directly develop production PPPs. As both the expenditure and the production approach have their strengths and weaknesses, an optimal approach may be to combine the best of both worlds, and use a mix of expenditure and production PPPs for comparison at industry level. The mix approach is discussed in Section 4.4, and its methodological underpinning, providing a reconciliation of expenditure and productivity prices using a supply-use framework is introduced in the Appendix to this chapter. Section 4.5 goes in some more detail into issues of multilaterisation of production PPPs, and section 4.6 deals with the issue of the consistency of benchmark and time series of PPPs. Chapter 5 discusses sources and methods to obtain industry level PPP for gross output on a sector-by-sector basis using the mix approach described in section 4.4. Chapter 6 goes more deeply into the derivation of PPPs at the level of intermediate inputs, capital and labour input and value added.

4.1 The use of expenditure PPPs for productivity comparisons at the aggregate level

63. Expenditure PPPs are well known for their use in comparisons of GDP per capita (Maddison 1995, 2001). They can also be applied to comparisons of productivity, notably labour productivity (van Ark and McGuckin 1999; Scarpetta et al. 2000; Pilat 2004; van Ark and McGuckin 2004). Expenditure comparisons were pioneered by Gilbert and Kravis (1954) and Gilbert and Associates (1958). Since then, the International Comparisons Program (ICP) has developed the systematic construction of expenditure PPPs under the auspices of the United Nations and the World Bank. A major standard work on ICP was provided by Kravis, Heston and Summers (1982). Since the early 1980s OECD has regularly published

[15] For Penn World Tables, see http://datacentre2.chass.utoronto.ca/pwt/. For GGDC, see http://www.ggdc.net/dseries.

[16] See Section 2.2 for a more detailed explanation of single and double deflation.

estimates of expenditure PPPs, derived from its joint programme with Eurostat. Eurostat publishes PPPs on an annual (current) basis, using a rolling sample of which one-third is refreshed every year, using extrapolations for the other two years, throughout each three-year cycle. Benchmark expenditure PPPs (constant PPPs) for OECD countries are currently available for 1980, 1985, 1990, 1993, 1996, 1999 and 2002. OECD uses these constant PPPs for their measures of annual GDP in US dollars in Volume I of the *Annual National Accounts*. OECD's *Main Economic Indicators* shows per capita volume indices using current PPPs. The *OECD Compendium on Productivity Indicators* uses constant PPPs for 2004 only.

64. Expenditure PPPs are based on prices mostly collected through specific surveys in the various countries participating in the comparisons. These are purchase prices for a common basket of products, which cover the whole range of goods and services included in the final expenditure on GDP. The products in the basket are deemed comparable between countries through detailed product specifications. In some cases the specifications are brand and model specific but in many instances generic specifications were used which only describe the relevant characteristics of the product without mentioning of a particular brand and model.[17] The product prices are used to calculate price relatives between countries, which are aggregated to the so-called basic heading level. A basic heading consists of a group of similar well-defined products from which a sample can be selected that are both representative of their type and of the purchases made in the participating countries. In the OECD PPP round of 1999 there were 221 of such basic headings. It is at the level of the basic heading that expenditure categories are defined and PPPs are calculated.[18]

65. Expenditure PPPs reflect the relative price ratios for private consumption, investment and government expenditure. They can be directly used for comparisons of levels of expenditure (gross national expenditure), but some caution is required when using them for comparing levels of output and productivity. To move from PPPs for final expenditure to PPPs for gross domestic product, a correction is required for the net foreign trade balance (X-M), that is, for relative prices of exports (X) and imports (M). Currently, the E-PPPs from Eurostat and OECD make an adjustment for the net balance of imports and exports on the basis of the actual exchange rate. In the Penn World Tables the adjustment is made with the overall expenditure PPP. Ideally a correction for differences in terms of trade is required. For this purpose one needs to develop separate export and import PPPs for internationally traded goods. Such PPPs are not yet available on a wide scale. Feenstra et al. (2004) provide a first attempt and show that terms of trade effects can be quite substantial for small open economies.

4.2 Adjusted expenditure PPPs at the industry level

66. Expenditure-based PPP comparisons are based on purchase prices of final goods and services with a detailed product specification. Hence to apply them to output and productivity comparisons by industry, the expenditure PPPs need to be re-allocated from expenditure categories to industry groups (see, for example, Hooper and Vrankovich 1995; Kuroda and Jorgenson 2001; Lee and Tang 2001; van

---

17 Brand and model specific PPPs are mainly for consumer goods and primarily for European countries.

18 During the past decades various adjustments and improvements have been made to the E-PPP method. In 1997, an evaluation of the PPP program was prepared by the former chief statistician of the Australian Bureau of Statistics, Ian Castles (OECD 1997), which has led to a range of improvements in the construction of PPPs. The Internet site of the OECD Statistics Department provides a good overview of some of the key issues related to the construction of purchasing power parities, see http://www.oecd.org/std/ppps.htm. The OECD benchmark study for 1999 also contains an extensive discussion of many of the issues related to the OECD/Eurostat work on PPPs (OECD 2002).

Biesebroeck 2004, Rao et al., 2004).[19] The method is best applied by using E-PPPs from the detailed basic heading level and apply those to 3- or 4 digit level industries (using the NACE, NAICS or ISIC industry classifications). To obtain an expenditure PPP for industry output, thePPPs can be aggregated in translog forms by using nominal output shares of the commodity mix as weights (see, for example, Kuroda and Jorgenson 2001):

$$\ln E\text{–}PPP_j^{BA} = \sum_{i \varepsilon j} \bar{v}_{ij} \ln E\text{–}PPP_{ij}^{BA} \quad (4.1)$$

with $E\text{–}PPP_j^{BA}$ the expenditure PPP for industry *j* between countries A and B, $E\text{–}PPP_{ij}^{BA}$ the expenditure PPP for expenditure category *i* and $\bar{v}_{ij}$ the weight of expenditure category *i* in industry j output, averaged over the two countries.[20] Ideally detailed industry-by-product (supply) tables are needed to indicate the output value of products originating from each industry (see Chapter 5). In the absence of a detailed supply table, business statistics (for example, the annual business survey or a production census) can be used to obtain output weights. It is also possible to use the expenditure shares from ICP as a proxy for output weights. These are relative easy to acquire but do not necessarily reflect output shares as they include imported goods, and exclude production for export.

67. The use of expenditure PPPs at industry level has several significant drawbacks (see also section 4.4). Firstly, a complete correspondence between expenditure categories and industries is not possible since, by definition, prices of goods produced as intermediate input for other industries are not covered by expenditure PPPs. In an industry like chemicals, production for intermediate consumption can make up more than 70 percent of output. To bridge the gap, some studies have included close substitutes for a given industry's intermediate deliveries based on its deliveries to final demand. Secondly, expenditure PPPs are based on a purchaser price concept whereas the output value to which it is applied is based on a producer of basic price concept. For example, a comparison of output in the shoe manufacturing industry should be based on the prices received by the manufacturer of shoes and should not include the transport margin of the transportation company and the retail margin gained by the shoe retailer. Expenditure prices also include indirect taxes and subsidies, which often vary across countries. Therefore the use of E-PPPs at industry level requires an adjustment from purchaser price level to basic price level.[21]

68. The adjusted expenditure PPP approach has been pioneered by Jorgenson and associates through 'peeling off' indirect taxes and transport and distribution margins from the expenditure prices (see, for example, Jorgenson, Kuroda and Nishimizu 1987):

---

19 Aggregate expenditure PPPs are sometimes used at industry level (see, for example, Dollar and Wolff 1993; Bernard and Jones 1996). This implies the strict very assumption of identical relative price levels across industries. This method is not recommended to be used for industry-level comparisons.

20 Of course, other weighting schemes like Laspeyres or Fisher type can also be used. The latter type is often used for production PPPs (see section 4.3).

21 Prices received by manufacturers are sometimes measured as producer prices, i.e. including net taxes (taxes minus subsidies) on products. It is desirable, however, to exclude those and use the basic price. In any case it is crucial that the output and the PPP use the same price concept be it producer or basic prices.

$$E\text{–}PPP^{BA}_{adjusted} = \left[\frac{1+T^{A}}{1+T^{B}}\right] E\text{–}PPP^{BA} \qquad (4.2)$$

Hence the PPP between country A and B for output in a particular industry is proxied by the adjusted expenditure PPP, calculated as the specific PPP for this industry multiplied by the ratio of tax, transportation and distribution margins (T) in country A over those in country B. Input-output tables are best to be used to calculate indirect taxes and transport and distribution margins even though not many countries make this kind of data publicly available at a sufficiently detailed level (see Pilat 1996).

69. As the expenditure PPPs are based on final consumption rather than on domestic output a third step is needed, namely the adjustment for the foreign trade effect. Relative prices for import goods need to be excluded from the expenditure PPPs, whereas those for export goods need to be added back in. This adjustment is important insofar as domestic, import and export price relatives for the same product differ and a significant amount of international trade takes place. This is particularly the case for commodities, such as agricultural, mining and manufacturing goods. Hooper (1996) proposes a solution to this problem for deriving manufacturing sector expenditure PPPs, adjusting domestic prices for import and export price effects. He assumes that import and export prices in manufacturing are set according to 'world' prices (P_{world}). The expenditure PPP for an industry adjusted for imports and export price effects, indicated by a '~', is obtained by multiplying the domestic output prices (which are the expenditure prices adjusted for margins, $P_{adjusted}$) by the difference between the world prices and domestic output prices weighted at the share of net export in total output (*(X-M)/Y*).

$$E\text{–}P\tilde{P}P^{BA}_{adjusted} = \frac{P^{B}_{adjusted} + \left[\frac{(X^{B} - M^{B})}{Y^{B}}\right](P_{world} - P^{B}_{adjusted})}{P^{A}_{adjusted} + \left[\frac{(X^{A} - M^{A})}{Y^{A}}\right](P_{world} - P^{A}_{adjusted})} \qquad (4.3)$$

The world price, P_{world}, for each item is obtained as an output-weighted average of each country's item price.

70. Making the adjustments for component E-PPPs at industry level to take account of taxes, margins and the foreign trade effect is empirically important. Table 4.2 shows for a number of countries the adjustments when moving from a component expenditure PPP for manufacturing to an adjusted E-PPP by first adjusting for wholesale and retail distribution margins, then for net taxes and subsidies, and finally by an adjustment for import and export prices. The results are only shown for aggregate manufacturing. At a lower level of aggregation, differences due to the various adjustments become much more pronounced (see Hooper and Vrankovich 1995, Table 2).

71. The method to adjust for world trade prices proposed by Hooper and Vrankovich (1995) relies heavily on a number of assumptions, in particular that for each country exports and imports are priced at the world price level for the goods in question. It also does not take account of tariffs and non-tariff barriers. Moreover, the average world price level is obviously based only on the prices of the countries which are included in the comparison.[22] Since the pioneering work of Hooper (1996) and Hooper and

[22] See Hooper and Vrankovich (1995) and Hooper (1996) for details. A practical problem of this method is the ad-hoc definition of the world price, for which Hooper has chosen the average price level for the G7 countries

Vrankovich (1995), this method has not been pursued further. A study of PPPs for exports and imports in combination with weights derived from trade data by industry is needed to continue this work (see, for example, Feenstra et al, 2004).

Table 4.2 - Bilateral PPPs for gross output in manufacturing derived from expenditure PPPs (units of local currency per dollar), 1990

	US	Japan	West Germany	France	Italy	UK	Canada
E-PPP	1.00	198.32	2.25	7.78	1794	0.71	1.41
plus adjustment for distribution margins	1.00	221.61	2.36	7.75	1888	0.77	1.41
plus adjustment for taxes	1.00	208.51	2.33	7.70	1957	0.77	1.42
plus adjustment for international trade	1.00	217.86	2.36	8.07	2005	0.79	1.43

Source: Hooper and Vrankovich (1995, Table 2).

72. In conclusion, the main drawbacks of the expenditure-based PPP method for comparisons of output and productivity at industry level, is that the adjustment of the PPPs from a purchaser price to basic price concept and the correction for foreign trade are complicated. Furthermore by definition expenditure PPPs only cover prices for final expenditure and do not reflect relative prices of intermediate goods. Finally, this approach still leaves many questions concerning the validity of the original expenditure PPPs for some of the basic headings in ICP. For example, in health services, education and parts of intermediate consumption the basic heading PPPs are not based on comparisons of expenditure prices, but on proxies derived from other basic headings or based on input prices (OECD 2002, Table I, p. 30) and as such are not useful for output and productivity comparisons (see also section 4.4 and Chapter 5).

4.3 The production approach

73. The theoretically most appropriate approach for international comparisons of output and productivity levels is to apply the production or industry-of-origin approach. The industry-of-origin approach was pioneered by Paige and Bombach (1959) in a comparison of the United Kingdom and the United States. The earlier work was conveniently summarised by Kravis (1976). In practice, two methods are available within this approach. The first method does not require the calculation of PPPs as it makes only use of direct comparisons of physical quantities of output (tons, litres, units). The second method converts output by industry to a common currency with an industry-specific production PPP.

74. In the past, international productivity studies often applied the physical quantity method (Rostas, 1948) and in many historical studies this is still common practice (Broadberry, 1997). Over the past four decades most studies have switched to the use of production PPPs instead of physical quantity comparisons. The switch is primarily caused by the rapid increase in the number of products and product varieties that need to be compared. If the total output produced would be matched by comparisons of physical quantities or comparisons of prices, and if these quantities and prices equal the value of output

normalised to US dollar price level by using the ratio of the adjusted component PPP relative to the exchange rate. Also, the adjustment is made at the industry, rather than the product level.

being compared, the physical quantitiy and production PPP methods provide exactly the same result. However with less than full coverage, there are differences between the two methods due to the underlying assumptions concerning the measured for non-matched output. The physical quantity method assumes that the quantity relatives of matched output are representative for the unknown quantity relatives of non-matched output. In case of the production PPP method the price relatives of matched output are considered representative for the unknown price relatives of non-matched output. It may be assumed that the representativity of price level comparisons of matched output for non-matched output is greater than for quantity ratios, in particular when the representativity assumption is limited to items in the same product group or industry. Hence nowadays there is a preference for using production PPPs rather than physical quantities.[23]

75. In contrast to the expenditure PPP method, there is no internationally co-ordinated survey for collecting information on specified output or producer prices from which detailed production PPPs can be calculated. There are various alternative ways to obtain production PPPs for gross output, partly depending on the data availability for individual industries. One way is to make use of producer or basic prices for specified products which are collected for alternative purposes.[24] For example, the Food and Agriculture Organization (FAO) collects prices of agricultural products for a large set of countries, which has been mostly applied for comparisons of agricultural output and productivity (see Chapter 5). The most widely used approach to obtain production PPPs is the unit-value-ratio (UVR) method. This method makes use of production statistics (censuses or business statistics surveys) that record the output values and quantities for product items. By dividing the output value by the corresponding quantities, one obtains unit values, which can then be used for calculating unit value ratios (UVRs) for matched items between countries.

76. The UVR-based method was pioneered by Paige and Bombach (1959) in a comparison of the United Kingdom and the United States. It was later applied by, among others, the United Nations Economic Committee for Europe (ECE) for a comparison of some East and West European countries (Conference of European Statisticians, 1972; Drechsler and Kux, 1972), and between Germany, United Kingdom and the United States (Smith, Hitchens and Davies, 1982). In recent decades, the UVR method has been explored and used in the ICOP project (International Comparisons of Output and Productivity) at the University of Groningen (Maddison and van Ark, 1989, 2002; van Ark, 1993, 1996; Timmer, 2000, van Ark and Timmer, 2001), the National Institute of Economic and Social Research (Rostas, 1948; Smith, Hitchens and Davies, 1982; O'Mahony, 1992, 1999, 2002) and the Centre d'études prospectives et d'informations internationales (Freudenberg and Ünal-Kesenci, 1994; Nayman and Ünal-Kesenci, 2000). As the UVR method has been mostly applied to manufacturing industries, the methodological presentation below focuses on that sector. However, the unit value ratio can also be used for several other sectors of the

---

23 Nowadays, physical quantities are still used to a limited extent, for example, for breaking down output of larger product categories. For example, in the case of passenger cars, the total value and volume of passenger cars produced can be broken down in a number of categories (SUV, luxury, medium, small types) using quantity information from industry specific sources (McKinsey Global Institute, 1992; van Ark, 1993). The physical quantity method is still in use for comparisons at plant level – in which case quality differences can be better accounted for – and for comparisons of output in some industries, for example, in mining and part of services, such as transport and communication (see Chapter 5 and van Ark, Monnikhof and Mulder (1999) and Mulder (1999)). See van Ark and Maddison (1994) for a review of the two approaches for manufacturing and for a description of the conditions under which these yield the same results.

24 In principle producer price data can be obtained from the national statistical programmes on producer price indices, but in practice it is very difficult to obtain comparable data due to a lack of harmonisation across countries. Moreover, such producer price measures would require an adjustment to basic price level, in case the output concept used is at basic prices.

economy. In Chapter 5 practical issues in comparisons of agriculture, services and manufacturing are discussed.

77. Unit value ratios are derived on the basis of unit values for matched product items. As a first step, the unit value (*uv*), which is a proxy for the precise prices of each product *i*, is derived by dividing the ex-factory output value (*o*) by the produced quantities (*q*) of product *i*:

$$P_i = uv_i = \frac{o_i}{q_i} \qquad (4.4)$$

The unit value can be considered as an average price, averaged throughout the year for all producers and across a group of nearly similar products. Subsequently, in a bilateral comparison, broadly defined products with similar characteristics are matched. For each matched product, the ratio of the unit values in both countries is taken. This unit value ratio (UVR) is given by:

$$\mathrm{UVR}_i^{BA} = \frac{\mathrm{uv}_i^A}{\mathrm{uv}_i^B} \qquad (4.5)$$

with A and B the countries being compared, B being the base country. The product UVR indicates the relative producer price of the matched product in the two countries. Table 4.3 gives an example of how UVRs are derived. The example is drawn from a comparison of output prices for the meat processing industry between the United States and Germany in 1997. In this industry 17 product matches have been made. Their unit values in national currency are given columns 5 and 6. The product unit value ratios are given in the last column.

78. Product UVRs need to be aggregated to derive PPPs for gross output for individual industries or for the aggregate sector. This can be done in one step from product to aggregate manufacturing, but also in various steps. Given the fact that mostly only a selected number of products are matched the latter is recommended. The UVRs are then reweighted several times, first according to their output share in the industry, then according to the industry share in the branch and finally according to the branch share in aggregate manufacturing. For example, following the latest version of the International Standard Industrial Classification (ISIC rev 3) product UVRs can be aggregated up 101 three-digit manufacturing industries (following ISIC rev 3), then to 22 two-digit manufacturing branches and finally to aggregate manufacturing. As a result the aggregate production PPP better reflects the actual share of each underlying product item for which UVRs are available in total output. The aggregation from product level to the 3-digit industry level can be compared to the practice of aggregating expenditure prices of individual goods first to basic headings. The difference is that product UVRs have weights attached, while expenditure PPPs below the basic heading level have not.

79. The production PPP for industry j based on the industry-of-origin approach is given by:

$$O\text{–}PPP_j^{BA} = \sum_{i=1}^{I_j} W_{ij}\ \mathrm{UVR}_{ij}^{BA} \qquad (4.6)$$

with i=1,.., I_j the matched products in industry *j*; $w_{ij} = o_{ij} / o_j$ the output share of the i^{th} commodity in industry j; and $o_j = \sum_{i=1}^{I_j} o_{ij}$ the total matched value of output in industry *j*. In bilateral comparisons the weights of either the base country (*B*) or the other country (*A*) can be used, which provide a Laspeyres and a Paasche type PPP respectively. As the output shares are consistent with the quantities that are used to derive the unit values, the weights and unit value ratios in (14.6) are consistent. This is an important advantage over the use of specified item prices as for the E-PPP approach. The bottom lines in Table 14.4 show the calculation of industry-level O-PPPs for gross output. The geometric average of the Laspeyres and Paasche indices, the Fisher index, is often used when a single currency conversion factor is required.[25]

80. The next aggregation step can be made by using either the gross output or the value added of each 3-digit industry to obtain an industry-weighted mean of all industry PPPs in a 2-digit branch. Again weights from base country *B* or the other country *A* can be used to arrive at Laspeyres and Paasche indices of the branch PPPs respectively. The latter step is repeated for the final aggregation step from branch level to the level of total manufacturing. Columns 4 to 6 in Table 4.4 provides an example of the various branch level production PPPs for gross output (weighted at gross output) for the Germany-US comparisons for 1997.[26]

81. As mentioned above, because of the increased variety and complexity of products not all products in an industry *j* can be matched, even not in a comparison between two countries. This is because of missing value or quantity data, difficulties in matching corresponding products between countries, and because of products that are unique for a specific country. Indeed the composition of output tends to differ much more across countries than the composition of expenditure, which is a complicating factor for industry-of-origin comparisons. Information on the number of product matches made and the output covered by these matched products can be used to assess the reliability and possible biases of the UVRs (see last columns of Table 4.4). The reliability of the PPP of gross output for a given industry or sector can also be evaluated on the basis of the coefficient of variation of the UVRs. On the assumption that relative price levels should not vary too much within homogeneous industries, large variations in unit value ratios within an industry or branch signal a lower reliability of the measures. Other indicators of reliability include a measure of the Paasche/Laspeyres spread between PPPs, which indicate differences in production structure between countries.

---

[25] Alternatively a Törnqvist aggregation can be made which is computationally almost identical. So far Fisher indices have been mostly been used for industry-of-origin comparisons.

[26] The example comes from the ICOP project at the University of Groningen. For a more detailed exposition of the UVR-methodology as applied in the ICOP project, the reader is referred to chapter 5.

Table 4.3: Example of Product Matching in The meat processing industry (ISIC 151), United States and Germany, 1997

	Quantity		Value (own currency)		Unit value		Value (other currency)		Unit
	U.S.	Germany	U.S.	Germany	U.S.	Germany	U.S.	Germany	Value Ratio
	(1000 ton)	(1000 ton)	(mil. USD)	(mil. DEM)	USD	DEM	(mil. DEM)	(mil. USD)	DEM/USD
Product name	(1)	(2)	(3)	(4)	(5)=(3)/(1)	(6)=(4)/(2)	(7)=(1)*(6)	(8)=(2)*(5)	(9)=(4)/(3)
Fresh and frozen beef	4757	498	15621	2069	3.28	4.16	19779	1634	1.27
Fresh and frozen boneless beef,	1448	222	3429	1076	2.37	4.84	7011	526	2.04
Fresh and frozen variety meats	650	263	951	342	1.46	1.30	846	384	0.89
Fresh and frozen veal	85	19	334	146	3.92	7.59	647	75	1.94
Fresh and frozen lamb and mutton	30	3	140	28	4.72	9.79	290	13	2.07
Fresh and frozen pork	647	1615	1276	5163	1.97	3.20	2068	3186	1.62
Cattle hides, skins, and pelts	933	42	1961	125	2.10	2.96	2764	88	1.41
Edible tallow and stearin	710	201	492	162	0.69	0.81	572	139	1.16
Lard	79	122	43	102	0.55	0.83	65	67	1.51
Boiled ham, barbecue pork, and bacon	1530	382	5788	3047	3.78	7.97	12191	1447	2.11
Sausages	2105	1256	7547	10485	3.59	8.35	17576	4502	2.33
Jellied goods	1635	336	6391	1954	3.91	5.82	9523	1312	1.49
Broilers and fryers	9607	398	12373	1186	1.29	2.98	28658	512	2.32
Fryer roaster turkeys	1868	222	3768	858	2.02	3.86	7216	448	1.92
Cooked or smoked turkey	385	38	1403	279	3.64	7.26	2797	140	1.99
Cooked or smoked chicken	1343	91	4125	661	3.07	7.26	9751	279	2.36
Cooked or smoked poultry hams	454	103	1838	1049	4.05	10.22	4646	415	2.53
Total Value			67480	28731			126400	15169	
Paasche PPP for gross output (4)/(8)									1.89
Laspeyres PPP for gross output (7)/(3)									1.87
Fisher PPP for gross output									1.88

Source: ICOP, University of Groningen, based on U.S. Dept. of Commerce, *U.S. Census of Manufactures 1997*; Statistisches Bundesamt, *Produktion im Produzierendes Gewerbe*, 1997.

Table 4.4: Example of Aggregation of PPPs for Gross Output from Industry, to Branch, to Sectoral Level, United States and Germany, 1997

	ISIC	No. of Product Matches	Coverage ratio (%)		Production PPP for Gross Output			Coefficient of variation	
			United States	Germany	Laspeyres	Paasche	Fisher	Laspeyres	Paasche
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Food products, beverages and tobacco	15-16	132	61.9	65.4	1.83	1.59	1.71	0.003	0.019
of which, Meat processing	*151*	*17*	*60.6*	*72.3*	*1.87*	*1.89*	*1.88*	*0.039*	*0.030*
Textiles	17	25	53.8	49.5	2.84	2.35	2.58	0.055	0.069
Wearing apparel	18	39	73.4	40.5	3.59	3.28	3.43	0.030	0.040
Leather products and footwear	19	12	61.7	31.2	2.21	2.12	2.16	0.089	0.123
Wood products	20	13	28.5	51.8	2.11	1.81	1.95	0.083	0.105
Paper products, printing & publishing	21-22	19	16.4	25.7	1.70	1.60	1.65	0.072	0.085
Petroleum and coal products	23	0	0.0	0.0	n/a	n/a	n/a	n/a	n/a
Chemicals and chemical products	24	59	17.5	12.9	1.77	1.88	1.83	0.052	0.061
Rubbert and plastic products	25	4	22.9	7.4	2.00	1.75	1.87	0.047	0.151
Non-metallic mineral products	26	23	28.5	22.0	1.50	1.06	1.26	0.052	0.053
Basic metals	27	43	69.6	71.3	1.75	1.56	1.66	0.028	0.027
Fabricated metal products	28	11	3.7	6.5	1.58	1.45	1.51	0.110	0.105
Machinery and equipment	29	39	10.2	9.9	2.03	1.93	1.98	0.075	0.069
Transport equipment	30-33	6	21.9	35.9	2.21	2.16	2.18	0.071	0.054
Office, accounting and computing machinery	30	6	42.9	38.3	1.79	1.58	1.68	0.217	0.131
Electrical and optical equipment	34-35	49	15.3	20.7	2.37	1.70	2.01	0.111	0.123
Furniture and miscellaneous manufacturing	36-37	36	19.0	30.1	1.66	1.47	1.56	0.064	0.062
Total manufacturing		516	27.0	28.5	1.89	1.82	1.86	0.024	0.021

Source: ICOP, University of Groningen, based on U.S. Dept. of Commerce, *U.S. Census of Manufactures 1997*; Statistisches Bundesamt, *Kostenstrukturerhebung*, 1997.

80. The choice for gross output or value added weights in the aggregation procedure depends on the actual purpose of the comparison. If productivity is ultimately going to be compared on a value added basis, thePPP for gross output may be assumed representative of the value added PPP (the single deflation procedure), and value added weights are then the proper weighting system at industry and branch level.[27] However, single deflation has a number of disadvantages.[28] From the perspective of a comparison of multifactor productivity levels, it is desirable to compute separate PPPs for gross output and for intermediate inputs, and to use gross output and intermediate input values as weights respectively (see Chapter 6 for a more extensive treatment of the double deflation issue in an MFP level comparison).

82. An important point question concerning the UVR method to obtain production PPPs is how well the unit value approaches the exact price received by the producer. The reliability of this type of production PPPs is therefore likely to be greatest in industries with homogeneous products, for which values and quantities are most readily available in the production statistics. For example, in basic goods industries, such as pulp and paper, wood products, metallic and non-metallic mineral products, output coverage is usually quite large as there are relatively few product differences between countries. Product matching is more difficult in manufacturing industries that produce durable consumer goods and investment goods, and the coverage percentages are therefore often lower than for basic goods and the UVRs less representative of all products in an industry.

83. The ‘product matching’ problem may be divided into two categories, namely the ‘product mix’ problem and the ‘product quality’ problem. The ‘product mix’ problem is caused by the fact that production statistics report quantity and values for product groups rather than for specified products. This problem is aggravated in international comparisons because of the lack of a harmonised product coding system for industry statistics, so that items need to be combined in order to obtain a correct match between countries.[29] The direction of the bias due to differences in product mix is undetermined.

84. The ‘product quality’ problem relates to differences in characteristics of a similar product. In this case it is not possible to find exactly the same model of a particular product, even if there was value and quantity information on specified products. In practice detailed product characteristics are difficult or impossible to observe directly from production statistics. Some studies make use of information from secondary, often private, industry sources that may include information on specified product prices. Such sources may prove particularly useful for products from typical high-tech industries, such as information technology equipment (computers and telecommunication equipment), passenger cars and pharmaceutical products.

---

27 Value added weights can only be used for aggregation of PPPs from industry levels up to higher levels as no value added weights are available at the product level.

28 Firstly, double deflation puts larger requirements on data, as besides PPPs for gross output, PPPs for intermediate inputs are also needed. Secondly, double deflation introduces a new possible source of error not present in single deflation measures, i.e. errors associated with the measurement of input prices. In many sectors, material input prices are either unavailable or crudely measured. Hill (1971) suggests that the use of double deflation when material input prices are measured with error may be more misleading than using single deflation. This problem is aggravated by the fact that double deflated value added is defined as relative output quantity minus relative intermediate input quantity. This subtraction gives room for errors. A small percentage error in the index of gross output appears as a much larger percentage error in double deflated real value added than in the case of single deflation as value added is smaller than gross output. Double deflation can also lead to measures of output that are unstable and erratic, especially at lower levels of aggregation when value added is a small share of gross output. This explains the ongoing popularity of single deflated value added in international comparisons of productivity levels, despite its theoretical inferiority.

29 The recent availability of harmonised product codes in the PRODCOM system for the European Union is a major step forward for international comparisons based on unit value ratios in mining and manufacturing. But comparisons with the United States are still hampered by the use of different product classifications. See Chapter 5 for more details

85. Another fruitful, though data demanding, way forward is to make greater use of hedonic price measurement for international price comparisons. Instead of observing the prices of products themselves, the hedonic method obtains the price of a bundle of characteristics of a product through regression analysis. The international hedonic function hypothesis assumes that the coefficients of the product characteristics in the hedonic function should be identical across countries (Triplett, 2004). The increased availability of scanner data makes such an approach possible, but so far it has mainly been experimented with for expenditure PPPs and hardly for production PPPs.[30] The estimation of hedonic functions with scanner data can be costly, however, and should therefore typically confined to areas where quality problems are particularly important and where these problems can be addressed by hedonic methods.

4.4 Assessment of PPP alternatives

86. It may be clear from the discussion in the previous sections that in practical terms neither the expenditure PPP approach or the production PPP approach is clearly superior when it come international comparisons of levels of output and productivity comparisons. At the aggregate level expenditure PPPs have been most widely applied mainly because it is readily available from the Eurostat and OECD price survey for the International Comparisons Project. Expenditure PPPs also have the advantage of being based on a specific survey of prices for specified items, and the problem of having to double deflate output and intermediate inputs as is necessary when using production PPP is avoided. However, a specific problem of using expenditure PPP for comparisons of GDP and aggregate productivity is the need to adjust for a terms of trade-effect, which according to the ICP approach is not standard done.

86. At the industry level, the production PPP approach is theoretically most preferable as it deals with the correct price that is required at output level. But production PPPs are mostly based on unit value ratios as basic prices for specified items at producer level are often not available. These unit value ratios have the disadvantage of introducing ‘product matching’ problems in international comparisons. Expenditure PPPs can also be used for output and productivity comparisons, but their use is theoretically disadvantaged because no price data are available for intermediate product items. The practical disadvantage of expenditure PPPs is that it requires detailed adjustments for margins, taxes and terms of trade effects. The choice on whether to use component expenditure PPPs or production PPPs is an empirical one, and will differ by industry. Below we discuss the criteria for choosing between expenditure PPPs and production PPPs. A methodological exposition, which sets a framework to reconcile expenditure and production using a supply-use framework, is provided in the appendix to this chapter.

87. It should be emphasised that the methodology and sources to obtain expenditure PPPs and production PPPs are fundamentally different. Expenditure PPPs are derived from specified prices that are obtained from a specific survey set up for the purpose. Production PPPs are mostly based on ratios of unit values derived from (national) production censuses and business statistics surveys. The strengths and weaknesses of the two PPP concepts can be classified in the following four broad areas: price concept, representativeness of sampled items, sampling error and weighting systems.[31]

88. As far as the price concept is concerned, PPPs at industry level should reflect relative prices for domestic production. Production PPPs directly reflect prices at industry level. Expenditure PPPs reflect prices of final expenditure on products, which either have been produced domestically or have been

---

[30] For example, the use of a country-product dummy (CPD) method, which relates through a regression analysis the price of a product to its characteristics and to a dummy variable for the country of origin, makes it possible to calculate the quality adjusted PPP from the coefficients of the dummy. Even when hedonic functions are not stable across countries, quality adjustments can be made on the basis of country-specific coefficients to derive PPPs. See Heravi, Heston and Silver (2003) for an application to television set in France, Netherlands and the U.K. See Moch and Triplett (2001) for an application to personal computer prices in France and German. See Van Mulligen (2002) who provides first results of a comparison of unit value ratios for cars between selected European countries.

[31] See also the discussion in O’Mahony (1996).

imported. In addition, not all domestically produced goods are destined for final expenditure, as part is being used for intermediate demand or exported. Expenditure PPPs therefore only approach the industry level prices in case most of the expenditure is on domestically produced goods and the industry output being produced goes mostly to domestic final expenditure and not to intermediate consumption or exports. This can be assessed using an input-output framework. In this framework total use of goods is given by the summation of final expenditure (consumer, investment and government demand), demand for intermediate consumption and exports. Total supply is given by domestic production (output) plus imports. Appendix 4A provides a detailed reconciliation of expenditure and production prices, which can be used for an assessment of strengths and weaknesses of both price concepts in different industries.

88. A second issue in the choice for expenditure PPPs or production PPPs for comparisons at industry level is representativity of the underlying prices. Representativeness is important in two areas: (1) representativeness of the matched products for the industry output of the countries being compared (also called 'characteristicity'), and (2) the representativiness of the relative prices for sampled products for the relative prices of unsampled products. In the ICP methodology, the expenditure PPP for a comparison of two countries is constructed from a sub-set of varieties that exist in both countries. In some cases, particularly for comparisons between countries with great differences in income levels or expenditure patterns, this overlapping subset concerns a small number of varieties which may be characteristic of neither country. Characteristicity is less of a problem with production PPPs, because the measures are not based on specification pricing but on unit value for more aggregate product groups that are being matched. This implies that the output coverage of matched products is substantially larger than the percentage of total expenditure that is covered by expenditure PPPs. Also, the unit value underlying production PPPs reflect prices averaged throughout the year, instead of prices at one point of time as in the ICP survey methods. Another important advantage of production PPPs is that they cover not only final expenditure goods, but also sales of goods for intermediate consumption.

89. Although O-PPPs may be more characteristic of the output in an industry then E-PPPs, there can be a problem of representativity of matched items for non-matched items in the same industry. Due to the 'product matching' problem O-PPPs are often based on samples of products which are biased towards relatively homogeneous, less sophisticated goods. The underlying unit value ratios may not always be representative of the more upgraded, high-quality varities in the same industry.

89. A third issue, linked to the previous one, concerns the sampling error of expenditure PPPs vis-à-vis production PPPs. In principle one might expect a bigger sampling error for unit value ratios compared to specified expenditure PPPs. As more aggregate items are being matched, product mix problems and the quality problem may raise the sampling error of production PPPs. Lichtenberg and Griliches (1989) showed that in an intertemporal context producer price indices in the US based on specified product prices are superior to those based on census unit value ratios as price dispersions are much higher for the latter. However, Balk (1999) shows that the sampling error is not only dependent on the extent of price dispersion for a product variety but also on the number of observations. In cases where production PPPs are based on more observations than the expenditure PPPs, the final sampling error might well be less than the error introduced by the quality problems.

89. In summary, there is no final judgement on the superiority of expenditure PPPs or production PPPs in adjusting output for relative price differences at the industry level. In practice a mixture of production PPPs and expenditure PPPs should be used for comparisons at industry level. The importance attached to the advantages and disadvantages of each method differs between industries. In Table 4.5 an assessment is made of the usefulness of expenditure and output PPPs for 19 different sectors in the economy. This is based on an assessment of expenditure PPPs from OECD for the 1999 ICP-round and production PPPs for 1997 from the ICOP project at the University of Groningen. PPPs are ranked from 0 (not useful) to 5 (very useful) on the basis of the following criteria.

Table 4.5: Assessment of usefulness of expenditure PPPs and productioin PPPs by sector for OECD countries

		Grade		Remark		
Industry	ISIC rev. 3 code	Adjusted expenditure PPP	Production PPP	Adjusted expenditure PPP	Production PPP	Final choice in this database
Agriculture	01-05	0	5	Not available	-	production
Mining and quarrying	10-14	0	4	Not available	-	production
Manufacturing	15-37					
Food, drink & tobacco	15,16	3	4	Trade intensive	-	Mainly production
Basic goods	17,20,21,23-28	1	4	Small expenditure share	-	Mainly production
Non-durable	18,19,22,36,37	2	4	Trade intensive	-	Mainly production
Durable	29-35	2	2	Trade intensive	Quality problem	production/expenditure
Electricity, gas and water supply	40,41	3	2	Low expenditure share	Quantity problem	expenditure/production
Construction	45	4	0	-	Not available	expenditure
Trade	50-52	0	2	Not available	Quality problem	production
Hotels & catering	55	4	0	-	Not available	expenditure
Transport	60-63	1	3	Small expenditure share	Quality problem	production
Communications	64	2	3	Small expenditure share	Quality problem	production
Finance	65-67	1	0	Reference PPP	Not available	expenditure
Real estate activities	70	4	0	-	Not available	expenditure
Business services	71-74	1	0	Small expenditure share	Not available	expenditure
Public administration and defence, education and health	75, 80,85	1	0	Mainly based on input PPPs	Not available	expenditure
Other services	90-95	2	0	Small expenditure share	Not available	expenditure

Note: ranking indicates 0 (not useful), 1 (very poor), 2 (poor), 3 (acceptable), 4 (useful) and 5 (very useful).
Source: assessment based on expenditure PPPs for OECD from 1999 round and production PPPs for 1997 from Groningen Growth and Development Centre.

90. The usefulness of expenditure PPPs at industry level depends on the share of final expenditure in total use, the share of import in total supply, and the quality of the expenditure PPPs.[32] The first two issues can be evaluated on the bases of supply-use tables as described in the appendix. As discussed above, the main weaknesses of production PPPs are the product mix and quality problems. Especially for high-tech goods, or heterogeneous services, production PPPs can be affected. In addition, for many services there is no data on unit values due to a lack of appropriate value data and the difficulty of defining quantities. Production PPPs are therefore particular useful for industries for which products are relatively homogeneous and for which differences in product quality problems are small. A more detailed assessment of the practical usefulness of adjusted expenditure PPPs and production PPPs is provided on an industry-by-industry basis in Chapter 5.

4.5 PPPs for inputs

4.5.1 PPPs for intermediate input

91. When double deflated value added measures are used in a productivity comparison, or a full KLEMS multifactor productivity measure is estimated, intermediate input PPPs are necessary in addition to the gross output PPPs discussed above. The data problems to obtain input PPPs for individual industries are larger than for output. There is often no input price parallel to the production PPPs. Business statistics surveys and productivity censuses provide little or no information on quantities and values of inputs in manufacturing, and for non-manufacturing industries the information is largely absent. Moreover, expenditure PPPs by definition do not reflect prices of intermediate inputs as they cover only final expenditure categories. However, it is possible to use a combination of production PPPs and expenditure PPPs as proxies for relative input prices.

92. From the producer's point of view, intermediate input PPPs should reflect the relative costs of acquiring intermediate deliveries, hence they need to reflect relative purchaser prices of intermediate inputs. But as for the case of output, the most important issue is to have consistency between the input valuation concept and the input price relatives. So the actual choice for a particular input price concept depends on the input value data used in the productivity comparison. For example, transactions in most of the input-output tables in the latest OECD I/O-database (OECD 2001d) are valued at basic prices. Trade and transportation margins are allocated to the retail and wholesale trade, transport, warehousing and insurance industries. Consequently, intermediate input values in these tables need to be converted using relative basic prices of inputs, rather than purchaser prices.[33]

93. Expenditure PPPs for final goods can be used as reasonable proxies for intermediate input PPPs when the same type of final goods are also used as intermediate inputs. One has to assume that the purchase prices are the same, irrespective of the industry of use of final consumption destination. The lower the level of aggregation for which this assumption is made, the higher confidence one can have in its validity. For example, the assumption that the expenditure PPP for a particular telecommunication service, say long-distance calling, is the same for the final consumer as for the intermediate user is less stringent than assuming that the average expenditure PPP for transport and communication services is the same for final users and intermediate users.

94. Production PPPs have the advantage that they actually directly cover intermediate deliveries. Hence it is straightforward to use the production PPPs as a proxy for relative intermediate input prices. For example, the production PPP for a particular type of paper, say kraft paper, can be used as a proxy for the

32. See, for example, the Castles report (OECD, 1997) for an assessment of E-PPPs.

33. When output is valued at basic prices and intermediate consumption at purchasers' prices, the residual value added does not represent final output of separate products valued at either of the two price concepts. With common prices separation of final output is possible, but the treatment of net taxes on products requires attention (Aulin-Ahmavaara (2003).

input PPP for those industries which use kraft paper as an input. Especially for manufactured intermediate inputs, production PPPs provide a valuable set of proxy intermediate input PPPs. However, a crucial assumption i is that import prices do not differ from prices of domestically produced intermediate inputs. This might not hold, especially when imports make up a large share of nitermediate consumption. In that case a possible alternative is to deflate domestically produced and imported intermediate inputs separately. Normally, imported intermediate inputs are separately identified in most I/O tables (but not for the U.S.). Import PPPs based on, for example, trade data may then be used in addition to domestic production PPPs for the domestically produced intermediate inputs. Exchange rates may be used as a rough proxy for intermediate inputs that are heavily traded in international markets such as crude oil.

95. The aggregation method to obtain PPPs for intermediate inputs is not fundamentally different from that for PPPs for gross output, *i.e.* it requires an average PPP for each input category, which is then weighted at the value share of each input category in the total value of intermediate inputs, using either base country or other country weights. The Use tables (which are part of the input-output framework) are the key source to obtain the weights of intermediate input categories as these detailed weights are normally not provided in the national accounts nor in production censuses or business statistics surveys. Assume that input prices can be expressed as translog functions of the prices of their components. Hence the differences between the logarithms of aggregate input prices for the two countries (*PPP of II*) can be expressed as a weighted average of differences between logarithms of the component intermediate input prices *m*, each weighted by its share in total input value v_m which is averaged over the two countries A and B:

$$\ln PPP\,of\;II = \sum_{m\,\varepsilon\,II} \bar{v}_m \ln PPP_m^{BA} \quad (4.7)$$

4.5.2 PPPs for labour input

96. Comparisons which use a homogenous (or 'raw') labour concept in the denominator of the productivity equation, such as number of workers or total hours worked do not need currency converters for labour input, as the comparison is already given in terms of volume. In the case of a heterogenous labour concept, for example workers of different skill types, relative labour input PPPs are needed to correct total labour compensation for differences in relative prices of skill categories of workers. Ideally, this labour input PPP should be based on labour costs including all costs incurred by the producers in employment of labour such as taxes levied, health cost payments, other types of insurance and contributions to retirement paid by the employer, and the value of payments in kind and allowances (such as housing and rent).[34]

97. The PPP for labour input should represent the relative price of one unit of domestic currency's worth of labour input. It is a weighted average of the PPPs for detailed labour categories with shares of each labour type in total labour costs as weights. Types of labour might include age, gender and skills. As discussed in Chapter 3, for large scale international comparisons it is currently not feasible to distinguish more than three detailed skill categories at industry level: low skilled (pre-primary, primary and lower secondary education, ISCED 0-2), medium skilled (upper secondary education, 3) and high skilled labour (total tertiary education, 5-7).[35] O'Mahony (1999) argues for a refinement of skill categories to also reflect levels of vocational education. Alternatively labour categories might be defined on the basis of occupation. In that case relative wage levels by occupation may be used to obtain labour input PPPs.

98. For many countries, data on compensation per working hour by skill category and industry cannot be obtained from the same source as the numbers on quantities of hours by labour type. The latter is normally collected by the labour force survey. Compensation data by labour type can sometimes be derived from a wage statistics, which may provide monthly earnings by industry and skill category. Such statistics on

34. Ideally, this should also include the value of financial benefits such as stocks options, see discussion in OECD (2001, p. 45)

35. See O'Mahony, Inklaar and Timmer (2004) for an attempt for four EU countries and the U.S.

earnings often need to be scaled up to total labour costs. The weights for each labour type need to be adjusted to the labour compensation measures as obtained from the national accounts, and adjusted for compensation of self-employed persons. In sum, data constraints limit what can be done in this context.

4.5.3 PPPs for capital input

99. To convert capital input measured in national prices into common prices, capital input PPPs need to be developed. Capital PPPs give the relative price of the use of a unit of capital in two countries from the purchasers' perspective. The calculation of the capital input PPP is less straightforward than for output, intermediate input or labour input PPPs. This is due to the conceptualisation of capital input as capital services rather than capital stocks (see Chapter 2). Expenditure PPPs for new investment goods are available from the ICT PPP programme. In rare cases, an income PPP for capital may be obtained from capital goods surveys, but the underlying price may represent either represent the book value or the perceived market value of the capital goods. In all cases the PPPs for capital service input will differ from the PPPs for new investment goods or capital goods if there are differences between countries in terms of the rate of return, the rate of depreciation, and/or the composition of the capital stock (Jorgenson, 1995, p.191 and 340).

100. The expenditure PPPs for investment have been criticised in recent evaluations of the PPP programme (OECD, 1997), and it remains unclear how reliable they are. Work on production PPPs for investment goods is one way to improve the measures for investment itself. Hedonic pricing techniques as discussed in section 4.3 can also help to improve measure of PPPs for information and communication equipment.

101. Once PPPs for investment or capital are available, they need to be transformed into PPPs for capital service input with the help of so-called annualisation factors. In other words, the PPP for investments or capital is adjusted for differences in user cost value between two countries. Under the assumption that the relative efficiency of new capital goods is the same in both countries, the relative capital rental price is calculated as (the asset subscript k is dropped for convenience):

$$\frac{p_{t,C}^K}{p_{t,US}^K} = \frac{p_{t,C}^I}{p_{t,US}^I} \frac{\left(\frac{p_{t-1,C}^I}{p_{t,C}^I} \tilde{r}_{t,C} + \delta \right)}{\left(\frac{p_{t-1,US}^I}{p_{t,US}^I} \tilde{r}_{t,US} + \delta \right)} \qquad (4.8)$$

with $\tilde{r}$ the real rate of return. This can be simplified to

$$PPP_t^K = PPP_t^I \frac{u_{t,C}^K}{u_{t,US}^K} \qquad (4.9)$$

where u indicates the user cost of one currency unit's worth of capital stock, PPP^I is purchasing power parity for investment, and PPP^K the PPP for capital services. Capital PPPs represent the price of a unit of domestic currency's worth of capital input in terms of a common currency. Ideally, differences in the effective rate of taxation on capital should also be taken into account in the calculation of relative user cost (Jorgenson 1995).

102. There are some key issues in the calculations of relative user cost across countries, which are parallel to the issues concerning the construction of capital services series (OECD, 2001*a*, 2001*b*). For the calculation of the rates of return the key question is whether these should be computed using an ex-ante approach, which assumes a given nominal rate of return, or an ex-post approach. The latter approach derives the rate of return on the basis of the capital revenue relative to the current price value of the total gross fixed capital stock. The capital revenue is based on the gross operating surplus as reported in the national accounts,

from which the imputed income of self-employed persons is deducted. The advantage of the ex-post approach is that the estimates are consistent with the overall accounting framework from which the productivity level estimates are derived. Less demanding is the use of an ex ante rate of return such as yields on government bonds.[36]

4.6 Multilateralisation of PPPs

103. Because of the complications in deriving industry PPPs described above, comparisons on a two-country basis are often the most practical approach for industry comparisons of productivity.[37] But if various bilateral comparisons are made and the same country is used as the base for this set of binaries, multilateral comparisons can be made as well. This variant is known as a 'star comparison' with the base country being the centre of the star. This type of comparison has the advantage that it complies with the requirement of country characteristicity. However, since each comparison only involves one pair of countries, the totality of the comparisons lacks internal consistency, *i.e.* they are not transitive. Direct comparisons of A and B will generate different results than indirect comparions through a third country (a with base and base with B). Multilateralisation of the results therefore seems desirable. Multilateralisation of the productivity measures are discussed in section 2.3.2. Here we discuss possible approaches to the multilateralisation of output and input price relatives.

104. Multilateralisation has been the corner stone of the ICP expenditure PPP programme, and an important source of debate over the past three decades. Essentially multilateralisation depends on the choice of aggregation method. Aggregation takes place after price ratios for individual goods and services have been averaged to obtain unweighted parities between countries for small groups of homogeneous commodities (basic headings). It involves weighting and aggregating these basic heading parities to arrive at PPPs and real values for each category of expenditure up to the level of total GDP. The seminal study on ICP by Kravis, Heston and Summers (1982) provided a wide range of aggregation methods, updated more recently by Diewert (1999) and Balk (2001). At the centre stage among the various methods discussed are the Geary-Khamis (GK) method, and those based on the Elteto-Koves-Szulc (EKS) method.[38] In the latest benchmark comparisons by OECD and Eurostat only these two alternatives are offered.

105. The EKS method, proposed by Elteto and Koves (1964) and Szulc (1964)[39], is designed to construct transitive multilateral comparisons from a matrix of binary/pairwise comparisons derived using a formula which does not satisfy the transitivity property. The EKS method in its original form uses the binary Fisher PPPs (F_{jk}: j,k=1,..M) as the starting point. The computational form for the EKS index is given by

$$EKS_{jk} = \prod_{l=1}^{M} \left[F_{jl} \cdot F_{lk} \right]^{1/M} \quad (4.10)$$

The formula defines the EKS index as an unweighted geometric average of the linked (or chained) comparisons between countries j and k using each of the countries in the comparisons as a link.

106. The EKS method shown in (4.10) produces comparisons that are transitive. In addition these indices also satisfy the important property that they deviate the least from the pairwise Fisher binary comparisons. This property is in line with the property of characteristicity discussed by Drechsler (1973). Since the Fisher index is considered to be ideal and possesses a number of desirable properties, the EKS

36. Van Ark *et al.* (2002) present annualisation factors for a large set of OECD countries at the aggregate level.

37. One complicating factor here is that measures from the ICP expenditure-PPP program are essentially multilateral, and additional tabulations are required to show those PPPs on a binary basis.

38. See Elteto and Koves (1964); Szulc (1964); Geary (1958); and Khamis (1970).

39. It is now well recognised that Gini proposed this method in 1924. We will continue to refer to this as the EKS-method as this isin line with most publications of international organisations.

method has a certain appeal since it preserves the Fisher indices to the extent possible, while constructing multilateral index numbers. However, a major problem with the EKS formula is that it gives equal weights to all linked comparisons [F_{jl} . F_{lk}], effectively assuming that they are of equal reliability. It can be argued that in practice it is possible to show that some link comparisons are intrinsically more reliable than others. For example, it might be that some pairwise Fisher indices are based on price data for many commodities while in other cases comparisons are based on prices for only one or two items. It is desirable to take this information into account when constructing the EKS multilateral indices (Rao, 2001; Rao and Timmer 2003).

107. The Geary-Khamis (G-K) method, unlike the standard index numbers, defines the purchasing power parities of currencies PPP_j (j=1,..M), and also a set of international average prices P_i (i=1,..N), one for each commodity, in terms of observed price and quantity data. Equations that define the PPP_j's and P_i's can be written as below. An international price, P_i, of *i*-th commodity is defined as:

$$P_i = \frac{\sum_{j=1}^{M} \frac{p_{ij} q_{ij}}{PPP_j}}{\sum_{j=1}^{M} q_{ij}} \qquad (4.11)$$

108. Thus the international price of *i*-th commodity is defined by first calculating the total value of output of *i*-th commodity across all countries which are in national currency units, converted into a common currency unit using the purchasing power parities. This total value, now expressed in a common currency unit, is then divided by the total output of this commodity across all countries. This definition of international average price is consistent with standard national accounts and statistical practices used in defining national average price from regional price data. To implement equation (4.10), it is necessary to define the parities, PPP_j. The G-K method defines these parities as below. For each j:

$$PPP_j = \frac{\sum_{i=1}^{N} p_{ij} q_{ij}}{\sum_{i=1}^{N} P_i \, q_{ij}} \qquad (4.12)$$

109. Equation (4.12) is in the form of a Paasche index where PPP_j is defined as the ratio of total value of production derived using national prices (p_{ij}) and international prices (P_i). Essentially, PPP_j in (4.12) measures the level of prices in country j relative to the international average price. Equations (4.11) and (4.12) are a system of equations which have to be solved simultaneously, see Kravis, Heston and Summers (1982) for more on this approach.

110. The Geary-Khamis and EKS aggregation methods differ substantially. The EKS method treats countries as a set of independent units with each country being assigned equal weight at the basic heading level. The EKS PPPs are obtained by minimising the differences between multilateral binary PPPs and bilateral binary PPPs. The EKS PPPs are thus close to the PPPs that would have been obtained if each pair of countries had been compared individually. The GK method treats countries as members of a group. Each country is weighted according to its share in total group GDP and the prices that are calculated are characteristic of the group overall. Both methods have advantages and disadvantages. For countries with price structures that are very different from the average, the GK approach generally leads to higher estimates of volumes than if more characteristic prices would have been used. This effect is particularly important when comparing countries with large differences in income levels. The EKS method leads to results that are more characteristic of each country's own prices, and thus leads to estimates of GDP per capita that are relatively similar to those resulting from the use of characteristic prices. Economically seen, this method allows for substitution behaviour and hence its results are not additive: the real value of aggregates is not equal to the sum of the real value of its components. This is preferable from the economic viewpoint, but might be a practical disadvantage if additivity is desired. For OECD countries, the empirical differences between the two methods are relatively small, since national price structures of final expenditure are quite

close. Most comparisons of income and productivity utilise the EKS results since these give equal weight to individual countries and are not biased against overstating income levels for relatively low-income economies. The EKS method is also more closely aligned with index number and economic theory.[40]

112. One problem with constructing multilateral price ratios is that they change with the addition of each new country in the comparisons. To overcome this, Eurostat has developed block fixity within the European countries ensuring that the addition of new countries outside the EU does not alter rankings within the EU. This is achieved by constructing EKS multilateral price ratios for the block of EU countries and then combining this with non-EU countries.

113. For industry-of-origin comparisons of productivity several attempts have been undertaken to multilateralise the binary results. Pilat and Rao (1996) constructed a first set of multilateral production PPPs by applying various multilateral indices (including GK and EKS) to original binary ICOP results at aggregate levels of manufacturing branches. Rao and Timmer (2003) applied multilateral aggregation procedures at the product level. First, they constructed a consistent list of over 250 manufacturing products for which price and quantity data were available for at least three countries. Aggregation took place in two stages. First, products were aggregated to branches; second, branches were aggregated to total manufacturing. The results suggest that for industry-of-origin comparisons, the choice of an aggregation formula at the product level is quantitatively much more important than at the branch level. Rao and Timmer also implemented weighted EKS formulae. The weighted EKS index includes information on the reliability of the various underlying binary comparisons. Such information may relate to the number of product matches in a group, the coefficient of variation of PPPs in such a group, or the spread between the Paasche and Laspeyres PPPs which is an indication of differences in production structure between countries. The results suggest that weighted EKS are superior to unweighted EKS and to the Geary-Khamis as long as additivity is not required.

114. Recently, an alternative approach to construct transitive price and quantity indices has been proposed besides Geary-Khamis and EKS. It is based on the procedure of chaining. A key problem in any spatial comparison is that there is no logical sequence of countries to be compared, as there is for comparisons between points in time. In the spatial context, the idea of data-independence of the ordering must be given up. Hill (1999) provides a structure of spatial chain indices based on the Paasche-Laspeyres (PL) spread of each underlying binary index. Using the graph theoretic concept of Minimum Spanning Trees (MST), a method is proposed of deriving a system of transitive multilateral comparisons from a matrix of binary comparisons. The approach is based on the fact that direct binary comparisons may not always be best. This method provides a spanning tree that links all binaries with the smallest PL spread. Multilateral indices are then obtained by linking together bilateral indices as specified by the spanning tree. An apparent problem in this approach is the unstability of the underlying spanning tree when new data (countries or time periods) is added although they tend to generate similar clusters of countries. In practice, it also appears that the multilateral index numbers are not very sensitive to the underlying spanning tree (Hill 2002). Hill and Timmer (2004) show that is possible toe generalize the Minimum Spanning Tree method by including various measures of reliability at the same time using standard errors on bilateral price indices as weights.

4.7 Time series of PPPs

115. PPPs are normally not made on an annual basis, but only for a given benchmark year.For example, expenditure PPPs are currently available for OECD countries for the years 1970, 1975, 1980, 1985, 1990, 1993, 1996, 1999 and 2002. The Penn World Tables have provided expenditure PPPs for 1967 (ICP I), 1970 (ICP 2), 1975 (ICP 3), 1980 (ICP 5), 1985 (ICP 5), 1990 (ICP 6), 1993 (PWT mark 6) and 1996 (PWT mark 6.1). Production PPPs are not systemically collected. The most regular effort is done within the framework of the ICOP project at the University of Groningen, which has produced production PPPs for manufacturing for the years 1975, 1987, 1992 and 1997 for different sets of countries. Benchmark year PPPs provide a

40. The EKS method is closely related to superlative index numbers, such as the Theil-Tornqvist index. See Neary (2005) for a recent attempt to provide a theoretical foundation for the Geary-Khamis system.

snapshot of relative prices in that particular year. PPPs for other years can be estimated by extrapolations with national price series. For example, the expenditure PPP in a country for year t can be obtained on the basis of extrapolation of the E-PPP from benchmark year 0 with the movement in the relevant consumer price index (CPI) in the country relative to the CPI in the reference country during [0, T]. Similarly, O-PPPs can be updated using producer price indices.

116. Let PPP_T^{BA} be the benchmark PPP for year T defined by:[41]

$$PPP_T^{BA} = \sum_i w_{T,i}^{BA} PPP_{T,i}^{BA} \quad (4.13)$$

with $PPP_{T,i}^{BA}$ the PPP for category *i* in year *T* and $w_{T,i}^{BA}$ the weight of category i in year T based on a common weighting scheme (industry output or expenditure at basic heading level) which will not be defined any further. The benchmark for T can also be proxied by updating the PPP for year 0, using national price indices

$$P\tilde{P}P_T^{BA} = PPP_0^{BA} \frac{P_T^A / P_0^A}{P_T^B / P_0^B} \quad (4.14)$$

with *P* the national price level. The PPP for T is estimated by extrapolating the benchmark PPP for 0 on the basis of relative changes in price levels in countries A and B. $P\tilde{P}P_T^{BA}$ and PPP_T^{BA} are normally not the same. Equation (15) can be rewritten in terms of prices of individual categories which are aggregated with weights for the year 0, assuming Laspeyres type national price indices for simplicity:

$$P\tilde{P}P_T^{BA} = \left[\sum_i w_{0,i}^{BA} PPP_{0,i}^{BA} \right] \left[\frac{\sum_i w_{0,i}^A (P_{T,i}^A / P_{0,i}^A)}{\sum_i w_{0,i}^B (P_{T,i}^B / P_{0,i}^B)} \right] \quad (4.15)$$

117. The procedure above may be described as the ‘constant’ PPP method. Although PPPs change over time but the weights are kept constant, in this example, to those of year 0.[42] In reality, an extrapolated PPP to year T is rarely equal to the benchmark year T PPP (when available). This is both for theoretical and practical reasons. The impact of practical implementation is probably more important but we start with the former.

118. When comparing (4.15) and (4.13) it can be easily seen that two index number problems plague the extrapolation procedure (Szilagyi, 1984). The first element is that the weights for the base year are preserved as the weighting system for the time series (fixed weight bias) in the extrapolated benchmark. The second element is that each of the the time series are based on national weights of each individual country, whereas benchmark estimates are based on a common weighting system for both countries (weight inconsistency). Both weighting problems are well-known in the price-index number literature. They are related and, in an internatioal context, they are referred to as the tableau effect.[43] The “tableau effect” can be significant in case relative price and quantity structures between countries and over time differ considerably, for example in

41. Here we make no distinction between E-PPPs and O-PPPs as the issues are the same for both PPP concepts.

42. In essence the procedure is the same as extrapolating a comparison of output or productivity between two countries from year 0 (using that year’s PPP) by using time series of output or productivity growth. In principle these two methods would give the same result if PPPs and GDP levels use are consistent with GDP price and volume series applied.

43. As called by Summers and Heston (1991, p. 340). Various smoothing methods can be used to straighten out these differences (see e.g. Krijnse Locker and Faerber 1984), but this implies that in the process benchmark PPPs and/or national CPIs are modified compared to the original estimates.

case there is a wide gap between successive benchmark comparisons. However, when the period considered is relatively short and the levels of development across countries not too different, the procedure would itself only lead to relatively minor inconsistencies, as price and quantity structures may be assumed constant over time and across countries.

119. The inconsistency due to the tableau effect can be reduced to a minimum when the extrapolation is not done at the aggregate level but at the lowest level of detail possible. In that case the updated PPP for category *i* ($P\tilde{P}P_{T,i}^{BA}$) is defined by

$$P\tilde{P}P_{T,i}^{BA} = PPP_{0,i}^{BA} \frac{P_{T,i}^{A} / P_{0,i}^{A}}{P_{T,i}^{B} / P_{0,i}^{B}} \quad (4.16)$$

The updated PPP for category *i* in year *T* ($P\tilde{P}P_{T,i}^{BA}$) is equal to the PPP for category i in year 0 ($PPP_{0,i}^{BA}$), using relative changes in national prices of i (P_i) in country A and B over the period [0,T]. In case that the basic price and quantity data for the benchmark PPPs and time series are consistent, the updated benchmark PPP for T will be equal to the original benchmark PPP for T when weights for T are used in the aggregation procedure as in (14). This method approaches the alternative to a constant PPP method, namely redoing the PPP calculation every year and weighting the categories at each year's expenditure. This is called the 'current' PPP method. However, as the weights differ for every year compared to the previous year, the current PPP method cannot be used for comparisons of relative price levels over time and it therefore only provides a 'snapshot' for each individual year. Eurostat uses a mix of the current PPP method and the annual extrapolation of PPPs at basic heading level.[44]

120. In practice consistency between basic price data and weights for the benchmark PPPs vis-à-vis the time series of national prices is never achieved. There are a host of problems which can be categorised into two groups: 1. differences in methodology and procedures between the national accounts and PPP programmes, and 2. inconsistency between PPP benchmarks and national price indices due to changes in methods and procedures.

121. The price methodology for PPPs differs from that for national price indices because of the different aims of both exercises. These include differences in the price index formulae used (for example, a multilateral index or Fisher index for the PPP and a Paasche index for the price series) but also a difference in the actual weights entering into the country's intertemporal price indexes. In addition, there are differences between the specifications and pricing of products in both pricing exercises. For a PPP comparison, items are priced which are comparable between countries. For domestic time series, items are priced which are representative in both periods 0 and T. These item baskets are by no means identical. The product selection and definition for spatial comparisons often has to take into account large differences in output or expenditure structure whereas structures only change gradually over time for a single country.

122. Other methodological differences between PPPs and price series over time are related to the treatment of foreign trade which differs between ICP and national accounting practice. Also pricing procedures differ between ICP and national accounts, as for ICP one compares products and service prices irrespective of location and type of outlet is not taking into account.[45] The practical problems may be

---

44 Eurostat estimates PPPs on an annual basis using the rolling benchmark method which relies on the assumption that basic price and quantity data for the benchmark PPPs and time series are consistent. The starting point of the rolling benchmark approach is the complete matrix of basic heading PPPs from the latest benchmark year calculation. In each year following the benchmark year, about a third of the basic heading PPPs are replaced by new PPPs calculated using prices collected during the year, while the basic heading PPPs that are not replaced are extrapolated using price indices specific to these basic headings. Expenditure weights are updated every year for each basic heading. Eurostat constructs current PPPs for EU Member States, Iceland, Norway, Poland and Switzerland on an annual basis.

45. This is the so-called "a potato is a potato" rule.

somewhat different between production PPPs and expenditure PPPs, but they are of a similar nature. For example, whereas the O-PPPs for manufacturing products are based on unit values from production statistics, the national price indices are either survey-based producer price indices or national accounts-based deflators by industry.

123. In practice the product specifications, the expenditure and industry classifications, and the data collection and pricing procedures evolve over time, in both price index and PPP work. For example, price methodologies, especially for rents, high-tech goods and (non-market) services might differ between countries and affect the intertemporal consistency of CPIs. Initiatives such as the Harmonised Consumer Price Index by Eurostat are aimed at minimizing these inconsistencies between countries. Revisions such as those which came with the introduction of the 1993 System of National Accounts and the introduction of NACE rev. 1 and NAICS brought with it changes in product, industry and expenditure classifications, which affect the weights for price index and PPP computations. Especially the PPP methodologies and quality of the underlying price data and weights are changing over time. For example, multilateral PPP results are sensitive to the number of countries which participate in the ICP and ICOP programmes.[46] Finally, there is the problem that PPPs cannot be extrapolated at the item level as they should ideally, but at best at the basic heading or industry level which introduces a fixed weighting problem below the basic heading.

124. The discussion above suggests that there is no unique way of obtaining a set of PPPs and national price indices which are internally consistent both in the interspatial and intertemporal dimension. This problem can only be resolved by further integration of national CPI programs and the ICP work on expenditure PPPs (Prasada Rao, 2001) as well as PPI programs and ICOP-type work on production PPPs. Meanwhile pragmatic choices must be made, depending on the type of comparisons the user is interested in. For example, if one wants a snapshot view of comparative performance in year t one is left with two choices. First, to use the set of current PPPs for year T, when it is available and of sufficient quality. Or, second, to extrapolate a benchmark set of PPPs, say 0, with the help of national price series at the most detailed level and using year t expenditure or industry weights for aggregation to minimise the "tableau effect". For example, in their recent *Main Economic Indicators* OECD uses benchmark PPPs for GDP whenever available. Estimates on intermediate years are one-year extrapolations backwards or forward to benchmark years.[47] In contrast, in previous versions of the Penn-World Tables (PWT, see Summers and Heston, 1991) a simple errors-in-measurement has been used that provides adjustment factors to both ICP benchmark PPPs and national accounts price series to reconcile them. In the most recent version of PWT (version 6.1), this use of consistentization involved only modifying different benchmark estimates, not country growth rates (see documentation on PWT at http://pwt.econ.upenn.edu).

125. Ideally, what is needed is a fuller integration of prices collected for various PPP rounds and national price indices, and proper aggregation procedures that can generate consistent comparisons over space and time. This field is evolving, see *e.g.* Heston, Summers and Aten (2001), Prasada Rao (2001) and Hill (2002). One of the obstacles here is that comparisons for one benchmark year will be altered, sometimes significantly, when data for a new benchmark is introduced and consistency or transitivity is imposed on the combined data set. Also greater attention for problems of quality differences of PPP data for various countries and benchmarks is necessary.

---

46. See Heston, Summers and Aten (2001). This can be remedied by imposing "bloc fixity" for a particular group of countries (for example, EU member states) as done by Eurostat and OECD.

47. Strikingly OECD has chosen not to link the three year PPP series which leaves (small) breaks between years based on different benchmarks. OECD extrapolates the PPP only at the level of aggregate GDP, so that changes in expenditure weights are not reflected as is the case with the Eurostat estimates.

Box 4.1. **Conclusions and recommendations**

PPPs for productivity comparisons at the aggregate GDP level

Comparisons of aggregate productivity levels can best be made using expenditure-based GDP PPPs derived from the International Comparisons Project. Genuine PPPs are only available for benchmark years. By extrapolating with relative movements in national price indices at a detailed level time series of PPPs can be generated. Consistency between PPP benchmarks and national time series, however, is not to be expected until a further integration of national and international price collection efforts takes place.

Output and input PPPs for industries

For comparisons of productivity levels by industry it can be recommended to combine information from the expenditure approach (E-PPPs) and the industry-of-origin approach (O-PPPs). O-PPPs are mostly based on unit value ratios, and are best used in those areas where product quality differences across countries are limited, and where output is relatively homogeneous and/or mainly concerns intermediate inputs. These sectors include for example agriculture, mining, many manufacturing industries, public utilities and transport and communication. Unit value ratios may be complemented by 'producer price' information from secondary sources. Expenditure PPPs are useful for industries which product is primarily destined for final use, where the effect of import and export prices is small, and where the expenditure PPPs themselves can be assumed to be of sufficient quality. These industries include manufacturing industries producing, for example, investment goods and durable consumer goods, the construction sector and services such as hotels & restaurants, finance & business services. E-PPPs must be adjusted for differecnes in trade and transport margins and taxes, and international trade, when applicable. Intermediate input PPPs can be derived by transforming output PPPs by means of a use-table. In addition to PPPs for output and intermediate input, relative factor input prices are also needed. Labour PPPs can be based on information of relative wages, depending on the number of labour types being distinguished. Capital input PPPs can be derived from E-PPPs for investment goods, which need to be adjusted for differences in user cost of capital across countries.

Multilateralisation of PPPs

For comparisons involving more than two countries, multilateral PPPs that are transitive are to be preferred. When additivity is also required, the Geary-Khamis procedure is the best choice. This implies that the real value of aggregates is the sum of the real value of its components. However, the GK procedure has a well-known bias towards the price structures of large countries in terms of GDP. When the requirement of additivity is dropped, (a weighted variant of) the EKS procedure is to be preferred, which maintains country characteristicity.

Appendix 4.A The Link between Expenditure and Production Prices within a Supply-Use Framework

See Section 2 from the paper by Timmer, Ypma and van Ark (item 25 on the agenda)

CHAPTER 5: PRACTICAL ISSUES IN DEVELOPING PPPS BY INDUSTRY (NOT INCLUDED)

See Section 4 from the paper by Timmer, Ypma and van Ark (item 25 on the agenda)

5.1 The aggregate level

5.2 PPPs for agriculture

5.3 PPPs for manufacturing

5.4 Productivity comparisons for services

5.4.1 Productivity comparisons for construction

5.4.2 Productivity comparisons for transport and communication

5.4.3 Productivity comparisons for wholesale and retail trade

CHAPTER 6: MEASURING PRODUCTIVITY LEVELS (NOT INCLUDED)

See sections 4 and 5 of the paper by Inklaar and Timmer (item 26 on the agenda)

CHAPTER 7: CONCLUSIONS AND MAIN RECOMMENDATIONS (NOT INCLUDED)

REFERENCES

AHMAD, N. (2002) "Measuring Investment in Software", *STI Working Paper*, OECD, Paris, forthcoming.

AHMAD, A. F. LEQUILLER, P. MARIANNA, D. PILAT, P. SCHREYER and A, WÖLFL (2003), "Comparing Labour Productivity Growth In The OECD Area: The Role Of Measurement", *Statistics Directorate Working Paper 2003/5*, December 2003, OECD, Paris

ARK, B. VAN (1993), *International Comparisons of Output and Productivity*, Monograph series no. 1, Groningen: Groningen Growth and Development Centre (downloadable from http://www.eco.rug.nl/GGDC/ThesisArk.html)

ARK, B. VAN (1996), "Issues in Measurement and International Comparison of Productivity – An Overview", in OECD, *Industry Productivity. International Comparisons and Measurement Issues*, Paris: OECD, pp. 19-47.

ARK, B. VAN and R.C. INKLAAR (2002), *Purchasing Power Parities for Industry Comparisons*, Final Report for the International Benchmark Club, BAK Basel (Switzerland), Groningen Growth and Development Centre and The Conference Board.

ARK, B. VAN, and A. MADDISON (1994), "The International Comparison of Real Product and Productivity", *Research Memorandum*, no. 567 (GD-6), Groningen Growth and Development Centre and The Conference Board.

ARK, B. VAN and P.H. VAN MULLIGEN (1998), *Purchasing Power Parities for Industry Comparisons. Final Report of a project for BAK-Basel, Switzerland*, University of Groningen.

ARK, B. VAN, and D. PILAT (1993), "Productivity Levels in Germany, Japan and the United States", *Brookings Papers on Economic Activity, Microeconomics*, 2, pp.1-48.

ARK, B. VAN, and M.P. TIMMER (2001), "PPPs and International Productivity Comparisons: Bottlenecks and New Directions", Joint World Bank–OECD Seminar on PPPs, 30 January – 2 February, Washington D.C.

ARK, B. VAN, and M.P. TIMMER (2002), "Notes and Communications: Industry Productivity Comparisons", *De Economist* (Quarterly Review of the Royal Netherlands Economic Association), Volume 140, No. 1, pp. 95-109.

ARK, B. VAN, E. MONNIKHOF, and N. MULDER (1999), "Productivity in Services: An international Comparative Perspective", *Canadian Journal of Economics*, 32(2), pp.471-99.

ARK, B. VAN and R.H. McGUCKIN (1999), "International Comparisons of Labor Productivity and per Capita Income", *Monthly Labor Review*, July 1999, pp. 33-41.

ARK, B. VAN (2006), *Performance 2006: Productivity, Employment and Income in the World's Economies*, Research Report, The Conference Board, New York.

ARK, B. VAN, J. MELKA, N. MULDER, M. TIMMER and G. YPMA (2002), "ICT Investment and Growth Accounts for the European Union, 1980-2000", Final Report on "ICT and Growth Accounting", Prepared for DG ECFIN, European Commission, Brussels, June 2002.

AULIN-AHMAVAARA, P. (2003), "The SNA93 Values as a Consistent Framework for Productivity Measurement: Unsolved Issues", *The Review of Income and Wealth*, March 2003, pp. 117-133.

BAILY, M.N. (1986) "Productivity Growth and Materials Use in U.S. Manufacturing", *Quarterly Journal of Economics*, February 1986, v. 101, iss. 1, pp. 185-95

BAILY, M.N. and R.M. SOLOW (2001), "International Productivity Comparisons Built from the Firm Level", *Journal of Economic Perspectives*, 15(3), pp. 151-172.

BAILY, M.N and J.F. KIRKEGAARD (2004, *Transforming the European Economy*. Institute for International Economics, Washington D.C.

BALK, B. (1999), "On the use of unit value indices as consumer price indices", *Proceedings of the Fourth International Working Group on Price Indices*, Washington DC: BLS.

BALK, B. (2001), "Aggregation Methods in International Comparisons. What Have We Learned?", Paper presented to Joint World Bank–OECD Seminar on PPPs, 30 January – 2 February, Washington D.C.

BALL, V.E., J-C BUREAU, J-P BUTAULT and R. NEHRING (2000), "Levels of Farm Sector Productivity: An International Comparison", draft manuscript.

BAUMOL, W.J. and E.N. WOLFF (1984), "On Interindustry Differences in Absolute Productivity", *Journal of Political Economy*, 92(6), pp. 1017-1034.

BERNARD, A.B. and C.I. JONES (1996), "Comparing Apples to Oranges: Productivity Convergence and Measurement Across Industries and Countries," *American Economic Review*, 86(5), pp. 1216-1283.

BIESEBROECK, J. VAN (2004), "Cross-country conversion factors for sectoral productivity comparisons", *NBER Working Paper Series* No. 10279.

BLADES, D.W. (1993), "Comparing Capital Stocks", in A.E. Szirmai, B. van Ark and D. Pilat (eds.), *Explaining Economic Growth. Essays in Honour of Angus Maddison*, North Holland.

BOSKIN, M.J. and L.J. LAU (1990), "Post-war Economic Growth of the Group-of-Five Countries: A New Analysis", Technical Paper 217, Center for Economic Policy Research, Stanford University, Stanford, Ca.

BROADBERRY, S.N. (1997), *The Productivity Race. British Manufacturing in International Perspective*, 1850-1990, Cambridge University Press.

BUREAU OF LABOR STATISTICS (1983), *Trends in Multifactor Productivity, 1948-81*, Bulletin 2178.

CAPDEVIELLE, P. and M.K. SHERWOOD (2002), "Providing comparable international labor statistics", *Monthly Labor Review*, June, pp. 3-14.

CAVES, R.E., ed. (1992), *Industrial Efficiency in Six Nations*, MIT Press, Cambridge MA.

CAVES, R.E., L.R. CHRISTENSEN and E. DIEWERT (1982), "Multilateral Comparisons of Output, Input and Productivity Using Superlative Index Numbers", *Economic Journal*, 92, pp. 73-86.

CHRISTENSEN, L.R , CUMMINGS and D.W. JORGENSON (1981), "Relative Productivity Levels, 1947-1973: An International Comparison", European Economic Review, May 1981, v. 16, iss. 1, pp. 61-94.

CHRISTENSEN, L.R and D.W. JORGENSON (1973), "Measuring Economic Performance in the Private Sector", in M. Moss, ed., *The Measurement of Economic and Social Performance*, Studies in Income and Wealth, New York.

COELLI, T, D.S. PRASADA RAO and G.E. BATTESE (1998), *An Introduction to Efficiency and Productivity Analysis*, Kluwer Academic Publishers.

CONRAD, K. and D.W. JORGENSON (1985), "Sectoral Productivity Gaps between the United States, Japan and Germany, 1960-1979", in H. Giersch (ed.), *Probleme und Perspektiven der Weltwirtschaftlichen Entwicklung*, Duncker and Humblot.

CONFERENCE OF EUROPEAN STATISTICIANS (1972), Comparisons of Labour Productivity in Industry in Austria, Czechoslovakia, France and Hungary, *Statistical Standards and Studies no. 24,* United Nations, New York.

DANZON, P.M. and L-W CHAO (2000), "Cross-national price differences for pharmaceuticals: how large, and why?", *Journal of Health Economics*, vol. 19, pp. 159-195.

DENISON, E.F. (1985), *Trends in American Economic Growth 1929-82*, Brookings Institution, Washington, D.C.

DIEWERT, W. ERWIN (1986), "Micro-economic Approaches to the Theory of International Comparisons", NBER Technical Working Paper No. 53, National Bureau of Economic Research, Cambridge MA.

DIEWERT, E.W. (1999), "Axiomatic and Economic Approaches to International Comparisons", in A. Heston and R.E. Lipsey (eds), *International and Interarea Comparisons of Income, Output and Prices* pp.13-108.

DOLLAR, D. and E.N. WOLFF (1993), *Competitiveness, Convergence, and International Specialization,* Cambridge MA: MIT Press.

DRECHSLER, L. (1973), "Weighting of Index Numbers in International Comparisons", *The Review of Income and Wealth*, March.

DRECHSLER, L. and J. KUX (1972), *International Comparisons of Labour Productivity* (in Czech), SEVT, Prague.

DREW, C., C. RICHARDSON and P. VAZE (2001), "International Comparisons of Productivity – An Update on Developments", *Economic Trends*, May 2001.

ENGLANDER, S. and A. GURNEY (1994), "Medium-Term Determinants of OECD Productivity", *OECD Economic Studies*, No. 22, Spring, pp. 49-109.

ELDRIDGE, L.P. and M.K. SHERWOOD (2001), "A Perspective on the U.S. – Canada Manufacturing Productivity Gap", *Monthly Labor Review*, February, pp. 31- 48.

ELTETO, O. and P. KOVES (1964), "On an Index Computation Problem in International Comparisons" (in Hungarian), *Statistztikai Szemle*, Vol. 42, pp. 507-518.

EUROSTAT (2002), Agricultural Productivity Indicators. State of Development, Eurostat F1 – Agricultural Accounts and Prices (AAP), Luxembourg.

FÄRE, R., S. GROSSKOPF, M. NORRIS and Z. ZHANG (1994), "Productivity Growth, Technical Progress, and Efficiency Changes in Industrialized Countries", *American Economic Review*, vol. 84, pp. 66-83.

FAO (1986), *Intercountry Comparisons of Agricultural Production Aggregates*, FAO Economic and Social Development paper 61, Rome.

FAO (1993), *Intercountry Comparisons of Agricultural Output and Productivity*, FAO Economic and Social Development paper 112, Rome.

FEENSTRA, R., A. HESTON, M.P. TIMMER and H. DENG (2004), 'Estimating real production and expenditures across nations: a proposal for improving the Penn World Tables', *NBER Working Paper Series No. 10866*, Cambridge Mass.

FREUDENBERG, M. and D. ÜNAL-KESENCI (1994), "French and German Productivity Levels in Manufacturing", in K. WAGNER and B. VAN ARK (eds.), International Productivity Differences. Measurement and Explanations, North Holland, pp. 53-84.

FUJIKAWA, KIYOSHI and CARLO MILANA (1996), "Bilateral and Multilateral Comparisons of Productivity in Input-Output Analysis Using Alternative Index Numbers", in OECD, *Industry Productivity . International Comparisons and Measurement Issues*, OECD Proceedings, Paris 1996, pp. 217-241.

GEARY (1958), "A Note on Comparisons of Exchange Rates and Purchasing Power between Countries", *Journal of the Royal Statistical Society*, Vol. 121, 97-99, pt. 1.

GERSBACH, H. and B. VAN ARK (1994), "Micro Foundations for International Productivity Comparisons", *Research Memorandum* GD-11, Groningen Growth and Development Centre.

GILBERT, M. and I.B. KRAVIS (1954), An International Comparison of National Products and the Purchasing Power of Currencies, OEEC, Paris.

GILBERT, M. and Associates (1958), Comparative National Products and Price Levels, OECD, Paris.

GRILICHES, Z., ed. (1992), *Output Measurement in the Service Sectors*, NBER, The University of Chicago Press.

HALL, R. and C. JONES (1997), "Levels of Economic Activity Across Countries", *American Economic Review*, 87, pp. 173-177.

HALL, R.E. and D.W. JORGENSON (1967), "Tax Policy and Investment Behaviour", *American Economic Review*, Vol. 57.

HAYAMI, Y. and V. RUTTAN (1970), *Agricultural Development: an International Perspective,* John Hopkins University Press, Baltimore.

HESTON, A., R. SUMMERS and B. ATEN (2001), *Price Structures, the Quality Factor and Chaining,* paper presented at Joint World Bank–OECD Seminar on PPPs, 30 January – 2 February, Washington D.C.

HILL, T.P. (1977), "On Goods and Services", *Review of Income and Wealth*, vol. 123, no. 4, pp. 315-338.

HILL, R. (1999), "Comparing Price Levels Across Countries using Minimum Spanning Trees", *Review of Economics and Statistics*, 81(1), pp. 135-142.

HILL, R. (2002), "Measuring price differences across countries and time in the European Union", Paper to be presented at the 27th General Conference of the International Association for Research in Income and Wealth, Djurhamn, Sweden, 18-24 August 2002.

HOOPER, P. (1996), "Comparing Manufacturing Output Levels among the Major Industrial Countries", in OECD, *Productivity. International Comparison and Measurement Issues*, OECD Proceedings, Paris.

HOOPER, P. and E. VRANKOVICH (1995), "International Comparisons of the Level of Unit Labor Costs in Manufacturing", in *International Finance Papers*, No. 527, Federal Reserve Board, Washington D.C.

HULTEN, C.R. and F.C. WYKOFF (1981), "The Estimation of Economic Depreciation Using Vintage Asset Prices", *Journal of Econometrics*, vol. 15, pp. 367-96.

INKLAAR, R., H. WU and B. VAN ARK (2003), ""Losing Ground" Japanese Labour Productivity and Unit Labour Cost in Manufacturing in Comparison to the U.S." *Research Memorandum* GD-64, Groningen Growth and Development Centre, July.

INTERNATIONAL LABOUR OFFICE (ILO) (1999), *Key Indicators of the Labour Market 1999*, Geneva.

INTERNATIONAL LABOUR OFFICE (ILO) (2001), *Key Indicators of the Labour Market 2001-2002*, Geneva.

ISLAM, N. (1995), "Growth Empirics: A Panel Data Approach", *Quarterly Journal of Economics*, CX, pp. 1127-70.

ISLAM, N. (1999), "International Comparison of Total Factor Productivity: A Review", The Review of Income and Wealth, 45(4), pp. 493-518.

JORGENSON, D.W. (1995), *Productivity,* MIT Press, Cambridge Mass., 2 Volumes.

JORGENSON, D.W., F.M. GOLLOP and B.M. FRAUMENI (1987), *Productivity and US Economic Growth*, Cambridge MA: Harvard University Press.

JORGENSON, D.W. and M. KURODA (1990), "Productivity and International Competitiveness in Japan and the United States, 1960-1985", in C.R. HULTEN (ed.), *Productivity in the U.S. and Japan,* Studies in Income and Wealth, 51, Chicago: University of Chicago Press.

JORGENSON, D.W., M. KURODA and M. NISHIMIZU (1987), "Japan-U.S. Industry-level Productivity Comparisons, 1960-1979", *Journal of the Japanese and International Economics*, vol. 1, 1-30.

JORGENSON, D.W. and M. NISHIMIZU (1978), "U.S. and Japanese Economic Growth, 1952-1974", *Economic Journal*, vol. 88, No. 352, pp. 707-726.

JORGENSON, D.W., M. KURODA and M. NISHIMIZU (1981), "Japan-US Industry-Level Productivity Comparisons, 1960-1979", *Journal of Japanese and International Economies*, 1(1), pp.1-30.

JORGENSON, D.W. and K.J. STIROH (2000), "Raising the Speed Limit: U.S. Economic Growth in the Information Age", *Brookings Papers on Economic Activity*, Vol. 2.

JORGENSON, D.W. and E. YIP (1999), "Whatever Happened to Productivity Growth?", Manuscript Harvard Universoty, revised June 28, 1999.

JORGENSON, D.W. and F.C. LEE, eds. (2001), Industry-level Productivity and International Competitiveness Between Canada and the United States", Industry Canada Research Monograph, Ottawa.

JORGENSON, D.W., M.S. HO AND K.J. STIROH (2005), *Productivity. Volume 3: Information Technology and the American Growth Resurgence*, MIT Press, Cambridge, Mass.

KHAMIS, S.H. (1970), "Properties and Conditions for the Existence of a New Type of Index Number", *Sankhya*, Series B, Vol. 32, Parts 1 & 2.

KRAVIS, I.B. (1976), "A Survey of International Comparisons of Productivity", *The Economic Journal*, Vol. 86, March, pp. 1-44.

KRAVIS, I.B., A. HESTON and R. SUMMERS (1982), *World Product and Income*, Johns Hopkins University Press, Baltimore.

KRIJNSE LOCKER, H. and H.D. FAERBER (1984), "Space and Time Comparisons of Purchasing Power Parities and Real Values", *Review of Income and Wealth*, March 1984, 30(1), pp. 53-83.

KURODA, M., K. MOTOHASHI and K. SHIMPO (1996), "Issues on the International Comparison of Productivity: Theory and Measurement", in OECD (1996*a*), *Industry Productivity – International Comparison and Measurement Issues*, Paris, pp. 49-95.

KURODA, M and D.W. JORGENSON (2001), *Economic Growth and Structural Change in Japan*, Draft Manuscript Keio University 2001.

LANDAU R. (1989), "Technology and Capital Formation" in Jorgenson, Dale W.; Landau, Ralph, (ed.) *Technology and capital formation*, pp. 485-505, Cambridge, Mass. and London: MIT Press.

LEE, F.C. and Y. TANG (2001), "Productivity Levels and International Competitiviness between Canada and the US", in Jorgenson and Lee (ed.) (2001).

LEWIS, W.W. (2004), *The Power of Productivity: Wealth, Poverty, and the Threat to Global Stability*, The University of Chicago Press, Chicago.

LICHTENBERG, F. and Z. GRILICHES (1989), "Errors of Measurement in Output Deflators", *Journal of Business and Economic Statistics*, January 1989, 7(1).

MADDISON, A. (1952), "Productivity in Canada, the United Kingdom and the United States", *Oxford Economic Papers*, October.

MADDISON, A. (1980), “Monitoring the Labour Market: A Proposal for a Comprehensive in Official Statistics”, *Review of Income and Wealth*, June, pp. 175-217.

MADDISON, A. (1983), “A Comparison of Levels of GDP Per Capita in Developed and Developing Countries, 1700-l980”,*Journal of Economic History*, March.

MADDISON, A. (1987), “Growth and Slowdown in Advanced Capitalist Countries: Techniques of Quantitative Assessment”, *Journal of Economic Literature*, vol. 25, no. 2.

MADDISON, A. (1995), *Monitoring the World Economy, 1820-1992*, OECD, Paris.

MADDISON, A. (2001), *The World Economy: A Millennial Perspective*, OECD, Paris.

MADDISON, A. and B. VAN ARK (2002), “The International Comparison of Real Product and Productivity,” in A. Maddison, D.S. Prasada Rao and W.F. Shepherd (eds.), *The Asian Economies in the Twentieth Century*, Edwar Elgar, Cheltenham, pp. 5-26.

MADDISON, A. and H. VAN OOSTSTROOM (1993), “The International Comparison of Value Added, Productivity, and Purchasing Power Parities in Agriculture”, *Research Memorandum GD1*, Groningen Growth and Development Centre.

MADDISON, A. and D.S. PRASADA RAO (1996), “A Generalized Approach to International Comparisons of Agricultural Output and Productivity”, *Research Memorandum* GD-27), Groningen: Groningen Growth and Development Centre.

MASON, G., and M. O’MAHONY (1997), “Capital Accumulation and Manufacturing Productivity Performance: US-European Comparisons”, Discussion Paper No. 124, National Institute of Economic and Social Research, London.

McGUCKIN, R.H. and B. VAN ARK (2005), *Performance 2005: Productivity, Employment and Income in the World's Economies*, Research Report R-1364-05-RR, The Conference Board, New York.

McKINSEY GLOBAL INSTITUTE (1992), *Services Productivity*, Washington D.C.

McKINSEY GLOBAL INSTITUTE (1993), *Manufacturing Productivity*, Washington D.C.

McKINSEY GLOBAL INSTITUTE (2000), *Why the Japanese Economy is not Growing: Micro Barriers to Productivity Growth*, Washington D.C.

MOCH, D. and J. TRIPLETT (2001), “International Comparisons of Hedonic Price Indexes for Computers: A Preliminary Examination”, Prepared for NBER/CRIW Summer Institute meeting, July 30-31.

MORRISON CATHERINE and W. ERWIN DIEWERT (1990), “Productivity Growth and Changes in the Terms of Trade in Japan and the United States”, in C. Hulten (ed) *Productivity Growth in Japan and the US*, NBER Studies in Income and Wealth Vol 53.

MULDER, N. (1999), *The Economic Performance of the Service Sector in Brazil, Mexico and the USA, A Comparative Historical Perspective*, Monograph series no. 4, Groningen: Groningen Growth and Development Centre.

MULLIGEN, P.H. van (2002), *Quality Differences and Hedonic Pricing in International Comparisons*, Ph.D. thesis, University of Groningen, forthcoming.

NAYMAN, L. and D. ÜNAL-KESENCI (2000), "Comparaison France-Allemagne des niveaus de production et de productivité dans le secteur manufacturier", CEPII, Paris, mimeographed.

OECD (1996), *Industry Productivity: International Comparison and Measurement Issues*, OECD Proceedings, Paris.

OECD (1997), "Review of the OECD-Eurostat PPP Program", Paris, available at http://www.oecd.org/std/ppp/pps.htm.

OECD (1998a), "OECD Data on skills: Employment by industry and occupation", *STI Working Papers 1998/4*, OECD, Paris.

OECD (1998b), "Employment measures in the OECD national accounts: Comparison between national accounts and labour force statistics", OECD, Paris.

OECD (1999*a*), *Employment Outlook*, Paris.

OECD (1999*b*), "Purchasing Power Parities and Real Expenditures – EKS Results, 1996", Paris.

OECD (2000), "Underground, informal and illegal activities (non-observed economy): creation and content of a measurement handbook", STD/NA(2000)21, OECD, Paris.

OECD (2001*a*), *OECD Productivity Manual: A Guide to the Measurement of Industry-Level and Aggregate Productivity Growth*, OECD, Paris.

OECD (2001*b*), *Measuring Capital*, OECD, Paris.

OECD (2001*c*), *The New Economy: Beyond the Hype*, OECD, Paris.

OECD (2001*d*), *The Input-Output Database*, OECD, Paris (http://www1.oecd.org/dsti/sti/stat-ana/stats/io/inout.pdf)

OECD (2001*e*), "Estimates of Annual Hours Actually Worked in OECD Countries", mimeo.

OECD (2001*f*), *Education at a Glance*, OECD, Paris.

OECD (2001*g*), "Investment in Human Capital Through Post-Compulsory Education and Training", *OECD Economic Outlook*, No. 70, Paris, December.

OECD (2001*h*), "Productivity and firm dynamics", *OECD Economic Outlook*, No. 69, Paris, June.

OECD (2001*i*), "Problem areas in calculating times series of PPPs", paper presented at joint World Bank-OECD seminar on PPPs Recent Advances in Methods and Applications, Washington DC 30 January – 2 February 2001.

OECD (2001*j*), "Consistency between GDP based on PPPs and National Accounts Time Series", *OECD memo*, STD/NA(2001)34, October 2001.

OECD (2002), "Purchasing Power Parities – Measurement and Uses" , *OECD Statistics in Brief*, March 2002, No. 3.

OECD (2002*a*), "ICT and Business Performance - Preliminary Findings", mimeo.

OECD (2002*b*), *Measuring the Non-Observed Economy. A Handbook*, OECD, Paris.

OECD (2002*c*), *Employment Outlook*, June.

OECD (2002*d*), "Measures of Skill from Labour Force Surveys – An Assessment", Working Party of National Experts on Science and Technology, OECD, Paris

OECD (2002*e*), *Purchasing Power Parities and Real Expenditures, 1999 Benchmark Year*, OECD, Paris.

OECD (2002f), "Purchasing Power Parities – Measurement and Uses" , *OECD Statistics in Brief*, March 2002, No. 3.

OECD (2004a), *The Sources of Economic Growth in OECD Countries*, OECD, Paris.

OECD (2004b), *Purchasing Power Parities and Real Expenditures, 2002 Benchmark Year*, OECD, Paris.

OECD (2005), *OECD Compendium of Productivity Indicators*, OECD, Paris.

OECD (2006), *Economic Policy Reforms: Going for Growth*, OECD Paris.

O'MAHONY (1992), "Productivity Levels in British and German Manufacturing Industry", *National Institute Economic Review*, No. 139

O'MAHONY (1996*a*), "Measures of Fixed Capital Stock in the Post-war Period: A Five-Country Study", in: B. VAN ARK and N.F.R. CRAFTS (eds.), *Quantitative Aspects of Post-War European Economic Growth*, CEPR/Cambridge University Press, pp. 165-214.

O'MAHONY, M. (1996*b*), "Conversion Factors in Relative Productivity Calculations: Theory and Practice", in OECD, *Industry Productivity. International Comparisons and Measurement Issues*, Paris: OECD, pp. 245-262.

O'MAHONY (1999), *Britain's Productivity Performance 1950-1996 – An International Perspective*, National Institute of Economic and Social Research, London.

O'MAHONY (2002), "Productivity and Convergence in the EU", *National Institute Economic Review*, April.

O'MAHONY, M. and W. DE BOER (2002), "Britain's relative productivity performance: Updates to 1999", National Institute for Economic and Social Research, mimeographed.

O'MAHONY, M., N. OULTON and J. VASS. (1998), "Market Services: Productivity Benchmarks for the UK", *Oxford Bulletin of Economics and Statistics*, November, p. 529-551.

O'MAHONY and N. OULTON (2000), "International Comparisons of Labour Productivity in Transport and Communications: The US, the UK and Germany", *Journal of Productivity Analysis*, 14, 7–30.

OOSTSTROOM, H. VAN and A. MADDISON (1984), "An International Comparison of Levels of Real Output and Productivity in Agriculture in l975", *Research Memorandum 162*, Institute of Economic Research, Groningen.

PAIGE, D. and G. BOMBACH (1959), *A Comparison of National Output and Productivity*, OEEC, Paris.

PILAT, D. (1994), *The Economics of Rapid Growth. The Experience of Japan and Korea*, Aldershot UK and Brookfield US: Edward Elgar.

PILAT, D. (1996), *Labour productivity levels in OECD countries: estimates for manufacturing and selected service sectors*, OECD Economics Department Working Papers no. 169, Paris.

PILAT, D. (1997), "Competition, Productivity and Efficiency", *OECD Economic Studies*, No. 27, 1996/II, pp. 107-145, OECD, Paris.

PILAT, D. and D.S. PRASADA RAO (1996), "Multilateral Comparisons of Output, Productivity and Purchasing Power Parities in Manufacturing", *Review of Income and Wealth*, 42(4), pp.1-18.

PILAT, D. and P. SCHREYER (2004), "The OECD Productivity Database – An Overview", *International Productivity Monitor*, Number 8, Center for the Study of Living Standards, Spring, pp. 59-65.

PRAIS, S.J. (1995), *Productivity, Education and Training. An International Comparative Perspective*, Cambridge University Press.

PRC BOUWCENTRUM (1995), *Een vergelijking van internationale bouwkostenvergelijkingen*, Bodegraven.

RAO, D.S. PRASADA (2001) "Weighted EKS and Generalized CPD Methods for Aggregation at Basic Heading Level and above Basic Heading Level", Paper presented at the Joint World Bank-OECD Seminar on Purchasing Power Parities: Real Advances in Methods and Applications, 30 January-1 February, 2001, Washington DC.

RAO, D.S. PRASADA and M.P. TIMMER (2000), "Purchasing Power Parities for Industry Comparisons Using Weighted Elteto-Koves-Szulc (EKS) Methods", *The Review of Income and Wealth*, 49(4), pp. 491-512

RAO, S., J. TANG and W. WANG (2004), "Measuring the Canada-U.S. Productivity Gap: Industry Dimensions", *International Productivity Monitor*, no. 9, Fall 2004, pp. 3-14.

RICHARDSON, C. (2001*a*), "International Comparisons of Productivity – A Detailed Discussion", Office of National Statistics, http://www.statistics.gov.uk

RICHARDSON, C. (2001*b*), "International Comparisons of Productivity – New Methodology, New Data", Office of National Statistics, http://www.statistics.gov.uk

ROSTAS, L. (1948), *Comparative Productivity in British and American Industry*, National Institute of Economic and Social Research, Cambridge University Press, London.

SCARPETTA, S., A. BASSANINI, D. PILAT and P. SCHREYER (2000), "Economic Growth in the OECD Area: Recent Trends at the Aggregate and Sectoral Level", *OECD Economics Department Working Papers No. 248*, OECD, Paris.

SCHREYER, P. and D. PILAT (2001), “Measuring Productivity”, *OECD Economic Studies*, No. 33, 2001/2, pp. 127-170.

SØRENSON, A. (2001), “Comparing Apples to Oranges: Productivity Convergence and Measurement Across Industries and Countries: Comment”, *American Economic Review*, 91(4), pp. 1160-1167.

SUMMERS, R. and A. HESTON (1991), ‘The Penn World Table (Mark 5): an Expanded Set of International Comparisons, 1950-1988’, *Quarterly Journal of Economics*, 106, pp. 327-68.

SZULC, B. (1964), “Index Numbers of Multilateral Regional Comparisons” (in Polish), *Przeglad Statysticzny*, Vol.3, 239-254.

SMITH, A.D., D.M.W.N. HITCHENS and S.W. DAVIES (1982), *International Industrial Productivity: A Comparison of Britain, America and Germany*, National Institute of Economic and Social Research, Cambridge University Press, London.

SMITH, A.D. and D.M.W.N. HITCHENS (1985), *Productivity in the Distributive Trades – A Comparisons of Britain, America and Germany*, National Institute of Economic and Social Research, Cambridge University Press, London.

SUMMERS, R. (1973), “International Price Comparisons Based Upon Incomplete Data”, *Review of Income and Wealth*, March.

SZILAGYI, G. (1984), “Procedures for Updating the Results of International Comparisons”, *Review of Income and Wealth*, June, Vol. 30(2), p 153-168.

SZIRMAI, A. and D. PILAT (1990), “Comparisons of Purchasing Power, Real Output and Labour Productivity in Manufacturing in Japan, South Korea and the USA, 1975-1985”, *Review of Income and Wealth*, 36(1), pp.1-31.

TAYLOR, A.M. and M.P. TAYLOR (2004), “The purchasing power parity debate”, *Journal of Economic Perspectives*, 18(4), pp. 135-158.

TIMMER, M.P. (2000), *The Dynamics of Asian Manufacturing. A Comparative Perspective in the late Twentieth Century*, Edward Elgar, Cheltenham.

TIMMER, M.P. and B. VAN ARK (2002), “Capital Formation and Foreign Direct Investment in Korea and Japan: Coping with Diminishing Returns,” in J. Th. Lindblad, ed., *Asian Growth and Foreign Capital: Case Studies from Eastern Asia*, Aksant, Amsterdam, pp. 59-98.

TIMMER, M.P. and B. VAN ARK (2005), “IT in the European Union: A driver of productivity divergence?”, *Oxford Economic Papers*, vol. 57 no. 4, pp. 693-716.

TRIPLETT, J. (2004), *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products*, Paris: OECD.

VERMANDE, H. and P.H. VAN MULLIGEN (1999), “Construction Costs in the Netherlands in an International Context”, *Construction Management and Economics*, Vol. 17, pp. 269-283.