
Aggregate multi-factor productivity: measurement issues in OECD countries

Document de Travail
Working Paper
2018-20

Balázs Egert



UMR 7235

EconomiX - UMR7235
Université Paris Nanterre
Bâtiment G - Maurice Allais, 200, Avenue de la République
92001 Nanterre cedex

Email : secretariat@economix.fr



AGGREGATE MULTI-FACTOR PRODUCTIVITY: MEASUREMENT ISSUES IN OECD COUNTRIES

Balázs Égert¹

Abstract

This paper analyses for 34 OECD countries the extent to which the calculation of aggregate multi-factor productivity (MFP) is sensitive to alternative parameterisations. The starting point is the definition of MFP used in previous work in the OECD's Economics Department (e.g. Johansson et al. 2013). They include alternative MFP measures, with human capital included or excluded, with different measures of Purchasing Power Parity (PPP) exchange rates, using time-varying capital depreciation rates and different measures of capital stock and labour input (headcount against hours worked). The main result of the paper is that whether or not human capital is included in MFP makes a significant difference for the level and dynamics of MFP. At the same time, MFP measures are less sensitive to other parameters of the calculation.

Keywords: multi-factor productivity, measurement, human capital, OECD

AGGREGATE multi-factor productivity: measurement issues in OECD countries

JEL codes: C23, C51, J2, L43, L51, O4

1. OECD Economics Department; CESifo, and Economix at the University of Paris X-Nanterre. Email: balazs.egert@oecd.org. The paper benefited from very helpful comments and discussion from Alain de Serres, Yvan Gillemette, Alberto González Pandiella, Mikkel Hermansen, Vincent Koen, Jan Strasky, David Turner, Ben Westmore and participants at an internal seminar at the Economics Department. Many thanks go to Isabelle Wanner for excellent statistical assistance.

1. Introduction

1. Measuring multi-factor productivity correctly at the macroeconomic level and understanding its main drivers is key to implementing pro-growth reforms. Progress has been made in measuring multi-factor productivity (MFP) and understanding its determinants at firm and industry level (O'Mahony and Timmer, 2009; Syverson, 2011; Andrews and Cingalo, 2014; Bartelsman et al., 2014). However, measuring MFP at the aggregate/macroeconomic level remains surrounded by measurement problems (Madsen 2006, 2014). Considerable work has been done in recent years to measure MFP but MFP is to some extent a measure of what we do not know. For instance, measures of MFP capturing the contribution of human capital within MFP may distort spatial and temporal comparisons. Indeed, the uncertainty about the precise level, comparability across countries, and the evolution over time has important implications for how precisely one is able to pin down the drivers of MFP in an empirical analysis. Not measuring accurately the MFP frontier masks our ability to reveal how far countries or firms are from the frontier and may lead to flawed conclusions regarding the speed of convergence to it. A wide margin of error may produce dubious results regarding the role of various institutions and policies for the evolution of MFP.

2. Against this background, this paper proposes to carry out an assessment of MFP measurement using the approach adopted in earlier OECD Economics Department work (Johansson, 2013). It is important to note that the objective is not to create new OECD productivity measures but just to carry out an extensive sensitivity analysis.² The paper shows the differences for alternative MFP measures, with human capital included or excluded, with different measures of Purchasing Power Parity (PPP) exchange rates, using time-varying capital depreciation rates and different measures of capital stock and labour input (headcount against hours worked).

3. The main results of this sensitivity analysis can be summarised as follows. First, whether or not human capital is included in MFP makes a significant impact for the level and dynamics of MFP.³ The absolute level and also the time profile of MFP series, from which the effect of human capital are stripped off, are often counterintuitive. This might indicate measurement problems with human capital. Southern European countries such as Spain and Italy have been top performers in the 1980s and early 1990s in terms of MFP levels. The United States was in the middle of the distribution for quite some time. At the same time, Italy and Spain experienced a trend decline in their MFP over three decades. By contrast, MFP series that include the effect of human capital show more plausible patterns. Second, while time-varying Purchasing Power Parity (PPP) has a number of advantages over a constant PPP rate, MFP measures based on constant (2005) PPPs provide a more plausible ranking of OECD countries in terms of relative MFP levels. Whether employment or total hours worked are used for labour input or how exactly the capital stock is calculated do not make too much difference. This finding is not in line with other studies such as OECD (2017).

2. Official OECD data for MFP are reported in the OECD Compendium of Productivity Indicators 2017 and available in the OECD productivity database.

3. Whether MFP is obtained as output minus human and physical capital and labour (MFP excluding human capital) or as output minus physical capital and labour (MFP including human capital)

4. The rest of the paper is organised as follows. Section 2 gives an overview on the calculation of MFP. Section 3 discusses the role of human capital in MFP calculations. Section 4 analyses the use of purchasing power parity (PPP) in the context of cross-country comparison. Sections 5 and 6 present MFP measures based on alternative labour and capital input definitions. Section 7 sketches out a framework for how MFP series should be chosen for empirical analysis from the large number of alternative MFP series produced in this paper. Section 8 finally provides some concluding remarks.

2. Calculating multi-factor productivity: a production function framework

5. MFP can be calculated as a residual of output once all inputs including capital and labour are accounted for (see e.g. Johansson et al., 2013). Typically, the log-level of MFP is derived on the basis of equation (1) below.

$$\ln(\text{MFP}_t) = \ln(Y_t) / \tau - \ln(HK_t) - \ln(L_t) - (1 - \tau) / \tau \times \ln(K_t) \quad (1)$$

where Y_t is real output, HK_t , K_t and L_t are the stock of human and physical capital and labour input, respectively.⁴ τ is the share of labour and $1 - \tau$ the share of capital in the production function.⁵ Y expressed in national currency units is sufficient to study the evolution of MFP over time. Cross-country comparison of MFP levels necessitates the conversion of output into a common currency unit. It is usually done using the purchasing power parity exchange rate (the ratio of absolute price levels in the domestic and foreign (reference) country), which accounts for different levels of economic development. Our benchmark MFP calculation is based on 2005 PPPs.⁶ Labour force surveys typically consider people living in the country and therefore exclude people working but not living in the country.⁷ Most studies and the OECD's Productivity database are unaffected by this as the data on labour input is drawn from business surveys that measure input based on the 'domestic' as opposed to 'national' concept. However, this study uses labour force survey data and as a consequence an adjustment (CLF) is necessary as shown below. Equation (1) is therefore complemented by a labour force adjustment term (CLF):

$$\ln(\text{MFP}_t) = \ln(Y_t) / \tau - \ln(HK_t) - \ln(L_t) - \ln(\text{CLF}_t) - (1 - \tau) / \tau \times \ln(K_t) \quad (2)$$

The inclusion of human capital is related to the standard Jorgensonian approach towards quality-adjusting labour input. Human capital is approximated by the Mincer equation using mean years of schooling of total population (MYS). Throughout this paper, the Mincer equation will be parameterised in line with Johansson et al. (2013) as follows:

-
4. Equation (1) gives MFP which is consistent with Harrod-neutral labour augmenting technical progress. This paper does not address how alternative function forms of the production function would impact on the measure of MFP.
 5. Capital and labour shares can differ across countries and can change over time. Applying different factor shares across countries would make the comparison of productivity levels difficult. Applying factor shares changes over time would change the slope of the series, again making cross-country level comparison difficult. Indeed, factor shares are kept constant over time and the same factor shares are used for specific sectors across countries in firm-level productivity analysis (Gal, 2013).
 6. The absolute price level of a given country in 2005, divided by the absolute price level in the USA in 2005.
 7. For instance many French people work in Luxembourg and Switzerland but live in France.

$$HK_t = 0.1254 \times MYS - 0.002 \times MYS^2 \quad (3)$$

6. Figure 1 plots MFP series obtained from Johansson et al. (2013), *labeffsd*). These series are calculated using equations (2) and (3) with $\tau = 0.67$, 2005 PPPs and the number of total employment. The underlying series are smoothed to filter out short-term disturbances. Figure 1 also shows our reference MFP series, calculated along the lines of equation (2) using raw (unsmoothed) data series for human and physical capital and labour (see Table 1).⁸ The two sets of data have a number of common features:

- Some Southern European countries including Spain, Italy and, to a lesser extent, Portugal were among the highest productivity countries from the mid-1980s to the mid-1990s.⁹
- Luxembourg, Ireland and Norway have been among the top performers over the last decade.
- The United States was in the middle of the pack in the 1980s and became a top performer after the 2007 crisis.
- Japan has been at the bottom of the productivity distribution. Its absolute productivity level is very close to those one can observe in Central and Eastern European countries.
- The absolute MFP level of Germany and Switzerland has been persistently in the low range of advanced OECD countries.
- A number of countries experienced a trend decline in MFP levels, especially for smoothed MFP series. Spain and Italy, which were among the highest-MFP countries in the 1980s, underwent decades of steady decay.

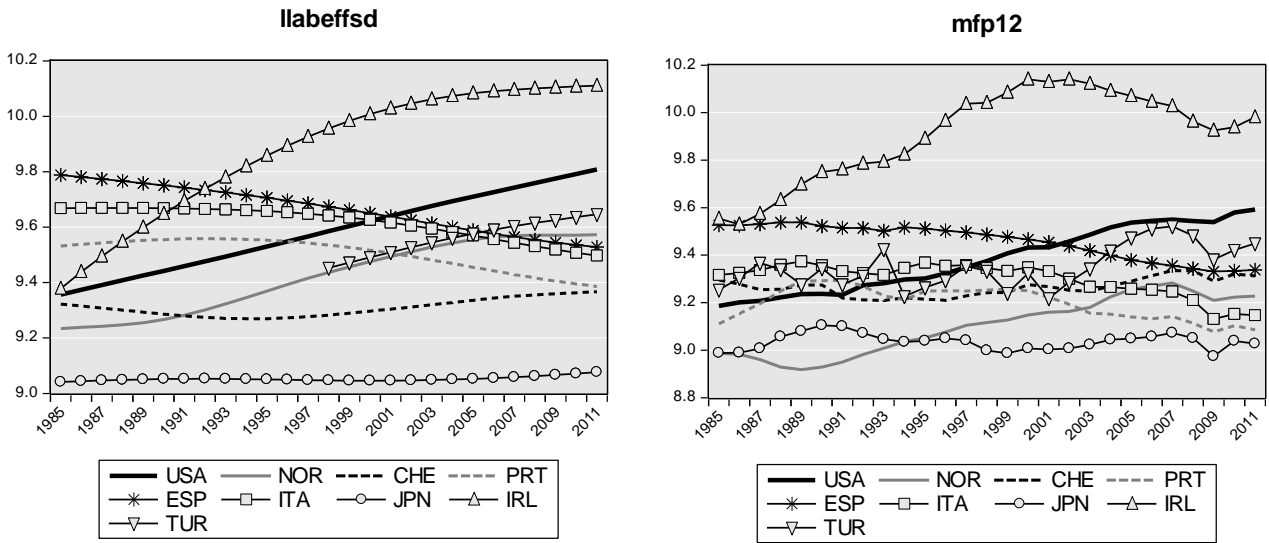
7. These observations, especially the counterintuitive relative levels and the trend decline observed in some of the countries prompt questions about the strength of assumptions used to estimate MFP, and human capital. Against this backdrop, a systematic assessment of factors causing these counterintuitive results is carried out. At the centre of this exercise, a reference measure (MFP12) is constructed, which comprises the following elements:

- Subtracting human capital from output (MFP obtained as output minus human and physical capital and labour)
- Using total employment (rather than hours worked)
- Capital stock with time-varying depreciation rates (drawn from the Penn World Table 8.0)
- 2005 GDP PPP for output and 2005 investment PPP for the capital stock (from PWT 8.0).

8. Recent studies using similar MFP series include Johansson et al., (2013), Westmore (2013) and Andrews and Westmore (2014).

9. Not shown here, but Brazil and Saudi Arabia had very high MFP levels in the late 2000s.

Figure 1. MFP in selected OECD countries, 1981-2013



Note: LABEFFSD is the MFP series obtained from the OECD Economic Outlook database. MFP12 is the reference MFP series for further investigation. It is based on raw (unsmoothed) series (see Table 1).

MFP12 is very similar to the series used in Johansson et al (2013) and earlier vintages of OECD databases. Using MFP12 as a starting point, a number of alternative MFP measures are calculated to investigate the sensitivity of MFP measures to changes in the way they are calculated including i) the treatment of human capital, ii) whether the number of employment or total hours worked is used, iii) how the capital stock is calculated, and iv) the type of PPPs used to obtain MFP levels comparable across countries (see Table 1 for an overview). Overall, the treatment of human capital is what matters most. The type of PPP applied can also make a big difference. By contrast, the way the capital stock is constructed and the measurement of employment matters to a considerably lesser extent (Appendix B presents the alternative MFP measures for all countries).

Table 1. Overview of alternative MFP definitions

	Human capital excluded from MFP	PPP conversion rate	Labour input: total employment	Capital stock: time-varying depreciation rate
LABEFFSD	YES	2005 PPP	YES	Partly
MFP12	YES	2005 PPP	YES	YES, PWT8.0
MFP13	YES	2005 PPP	total hours	YES, PWT8.0
MFP14	YES	2005 PPP	YES	Partly, OECD data
MFP15	YES	2005 PPP, same for GDP and investment	YES	YES, PWT8.0
MFP16	YES	chained PPP	YES	YES, PWT8.0
MFP17	YES	time-varying PPP	YES	YES, PWT8.0
MFP22	NO	2005 PPP	YES	YES, PWT8.0
MFP23	NO	2005 PPP	total hours	YES, PWT8.0
MFP24	NO	2005 PPP	YES	Partly, OECD data
MFP25	NO	2005 PPP, same for GDP and investment	YES	YES, PWT8.0
MFP26	NO	chained PPP	YES	YES, PWT8.0
MFP27	NO	time-varying PPP	YES	YES, PWT8.0

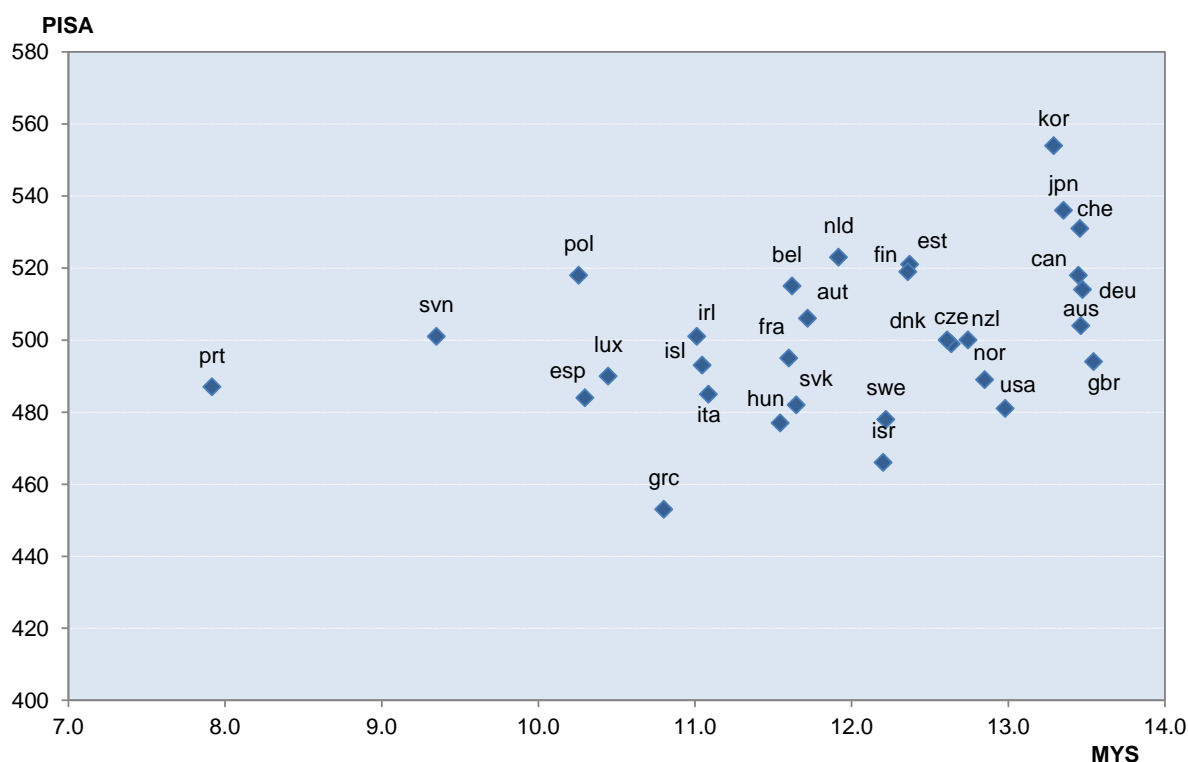
Note: 2005 PPP is the ratio of absolute price levels in 2005. Chained PPP is a variant of a time-varying PPP: the ratios of absolute price levels, observed at a multi-year frequency, are linked via interpolation for the missing years. "Time-varying PPP" is calculated as 2005 PPP extended forwards and backwards GDP and investment deflators.

3. Human capital

8. Whether or not human capital is included in MFP measures makes a significant difference. Human capital is measured on the basis of mean years of schooling. Mean years of schooling may not be strictly comparable across countries if education systems equip pupils with different knowledge and capabilities. Not adjusting for differences in the quality of teaching may seriously bias cross-country comparisons of human capital.¹⁰ The PISA scores published by the OECD can be taken as a rough approximation for the quality of teaching, at a given point in time, up to the years of age 15. The mean years of schooling fairly represents the quantity of schooling. Plotting these two series against each other show very limited correlation (Figure 2). In particular, data are clustered: for instance, the mean years of schooling is about 13.5 years with a 10% variation in PISA scores. A similar effect can be observed when mean years of schooling is about 11.5 years.

10. The proxy proposed by Madsen (2014) to adjust for the quality of teaching, the teacher-pupil ratio modified for weekly hours, appears to be problematic because it most probably do not account for quality differences in tertiary education.

Figure 2. The relationship between the quality and quantity of schooling

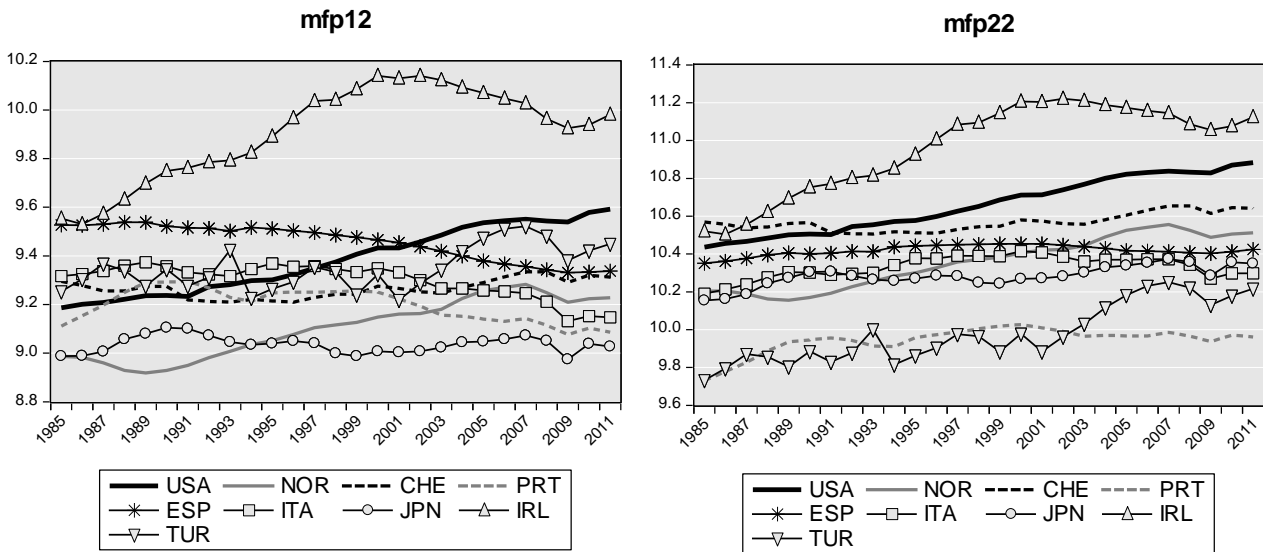


Note: the vertical axis depicts the overall PISA score 2012, the horizontal axis represents mean years of schooling (MYS) in the same year.

Source: OECD.

9. Human capital based only on quantity of inputs and not adjusted for quality can have a substantial impact on the level and the dynamics of the MFP series. If human capital is subtracted from output (as in MFP12), changing all the other parameters does not alter the overall picture. In contrast, keeping human capital in MFP (MFP22) changes the overall profile of the series. The United States is among the top performers from 1985 to 2013. Italy and Spain have more plausible MFP levels and the decline only starts in the mid-1990s for Italy (Figure 3). It is important to recognise that MFP measures (such as MFP22) including human capital are consistent with MFP measures calculated at the sectoral and firm levels: it is indeed very difficult to measure sector- or firm-specific human capital (Gal, 2013).

Figure 3. MFP measures excluding (MFP12) and including (MFP22) human capital



4. PPP and cross-country comparability

10. Constant PPPs are an accurate conversion factor (exchange rate) for temporal comparisons across countries if the real exchange rate is essentially unchanged, that is if the gap between PPPs and actual exchange rates remains stable over time and inflation rates are similar. In such a case, using the PPP rate rather than the market exchange rate would increase the productivity level of the less developed country, and hence lower the gap to the benchmark country.^{11,12} But for emerging market and developing economies in particular, typically, the faster a country is catching up vis-à-vis advanced economies the more likely the use of constant PPPs will give rise to a bias in the measurement of TFP levels. The relative price level (or the level real exchange rate) of catching-up countries tends to increase compared to the benchmark countries (the United States in our case). A major explanation is the Balassa-Samuelson effect,¹³ but other factors can also generate higher prices in catching-up economies.¹⁴

11. The PPP rate is given as the absolute price level of the domestic country expressed in national currency (P) divided by the absolute price level of the benchmark country (P*). That is, $PPP = P/P^*$. The relative price level is the comparison of the two countries' price levels in the same currency. The domestic country's price level expressed in the benchmark country's currency is P/NER where NER is the exchange rate calculated as national currency units (NCU) over one unit of the foreign currency (FCU) NCU/FCU e.g. GBR/USD. The relative price level is $(P/NER)/P^*$. This indicates whether the domestic country's absolute price level is lower or higher compared to the benchmark country. The relative price level can be rewritten as: $(P/P^*)/NER$: the PPP rate divided by the nominal exchange rate. This is indeed the level real exchange rate.

12. The market exchange rate would underestimate the level of MFP in the less developed country, similarly as in the case of per capita income. The reason is that if that the relative price level of less developed countries is lower, $P/P^* < NER$.

13. The dynamic version Balassa-Samuelson effect is at work if productivity gains in the open sector exceeding those in the closed sector are higher in the domestic country compared to the foreign reference country. Wages in the open sector increase in line with productivity gain. If wages equalise across the open and closed sectors, wages will also rise in the closed sector. Wage increases in the closed sector, in turn, will translate into higher prices in the closed sector. Overall, productivity gains in the opens sector will raise services prices, which will

Box 1. The concepts of different PPP measures

Using 2005 PPP would, for the period prior to 2005, underestimate the level and overestimate the change in MFP of catching-up economies, as it implies a higher absolute domestic price level (P) and a higher conversion (exchange rate) rate (hence a smaller adjustment compared to the market exchange rate).

Using time-varying PPPs, which reflect real convergence phenomena, would potentially attenuate these biases. The so-called chain-linked PPP is one avenue. Chain-linked PPP is constructed by connecting different vintages of absolute price levels through some kind of interpolation.¹⁵ Nevertheless, chained PPPs are not without problems. Most importantly, comparing different vintages of PPPs can be very problematic. Generally, it is not recommended the use of GDP numbers converted using time-varying PPPs from different vintages for time series analysis (Eurostat-OECD (2016), and Feenstra et al., 2013, Table 5).¹⁶ Nevertheless, a potential problem is that current PPPs are not consistent with GDP deflators.

Chain-linked PPPs are subject to quality and new goods biases. Ideally, PPPs should measure the price of almost identical goods. The International Comparison Programme, which is the basis of PPPs, notes that goods and services included in a country's price level is mostly based on a broad specification, precisely because the same good may not be available or representative in a country. This means that more developed countries may have higher-quality and hence higher-priced goods in their price basket. As a result, PPPs may overestimate cross-country price differences. But this has also important implications for the time series dimension. With real catching-up, there will be a shift from lower quality to higher quality goods and services and the variety of goods can increase (new goods can enter and old items exit the basket). Higher quality and new goods show up in higher price levels from one price survey to another one. This potentially overstates price level convergence.

Quality and new-goods bias are usually taken care of in consumer price indices, even if imperfectly. This argues for the use of a single PPP vintage (for instance 2005 PPPs) but which are extended backward and forward in time by cumulated inflation differentials (the evolution of the level CPI in the domestic country relative to the one in the benchmark country – the USA in our case). The main caveat is, however, that inflation differentials tend to largely underestimate real convergence compared to price level convergence. Weights used in the CPI basket represent the consumption basket of the average household in any given country. According to Engel's law, poorer households spend a larger proportion of their income on food than on other items. Analogously, they also spend less on services. In fact, the share of services in the CPI basket of developing/emerging market economies tend to be considerably lower than in more advanced countries. The low weight of services will considerably understate the Balassa-Samuelson effect in consumer price indices. A mid-way would be to rely on GDP deflators. The weight of the services sector therein is considerably higher than in the CPI basket. The Balassa-Samuelson effect will not be hidden in the GDP deflators of different countries. Hence, the constant 2005 PPP could be extended back and forth using the GDP deflator differential. Yet the weights on services and other sectors also change in the GDP deflator over time. Overall, time-varying PPPs have a number of appeals over constant PPPs. Nevertheless, they are plagued by various measurement problems. It is not clear which PPP would be best for MFP measurement.

increase the overall price level. At the same time, the Balassa-Samuelson effect is competitiveness neutral: PPP is assumed to hold for tradable goods. In other words, the real exchange rate based on tradable prices is assumed to be stable over time.

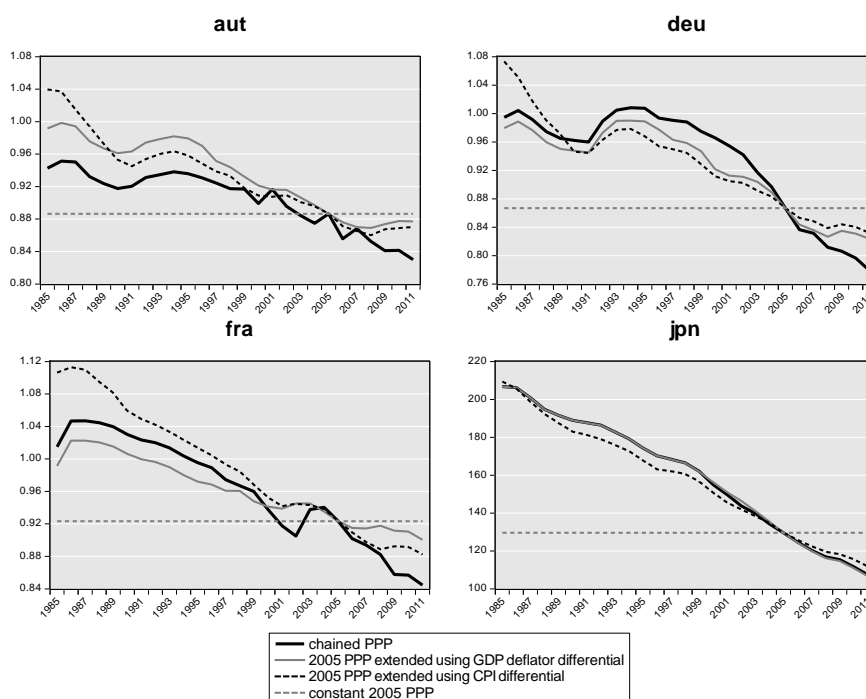
14. A number of other channels can increase services prices (Baumol effect, Penn effect, Baghwati effect) and goods prices. For an overview, see e.g. Égert (2007).
15. They usually come at three- or five-year intervals.
16. http://www.rug.nl/research/ggdc/data/pwt/v80/pwt_80_user_guide.pdf, Table 5.

11. Looking at different PPPs over time, a couple of striking features of the data emerge (Figure 4 below):

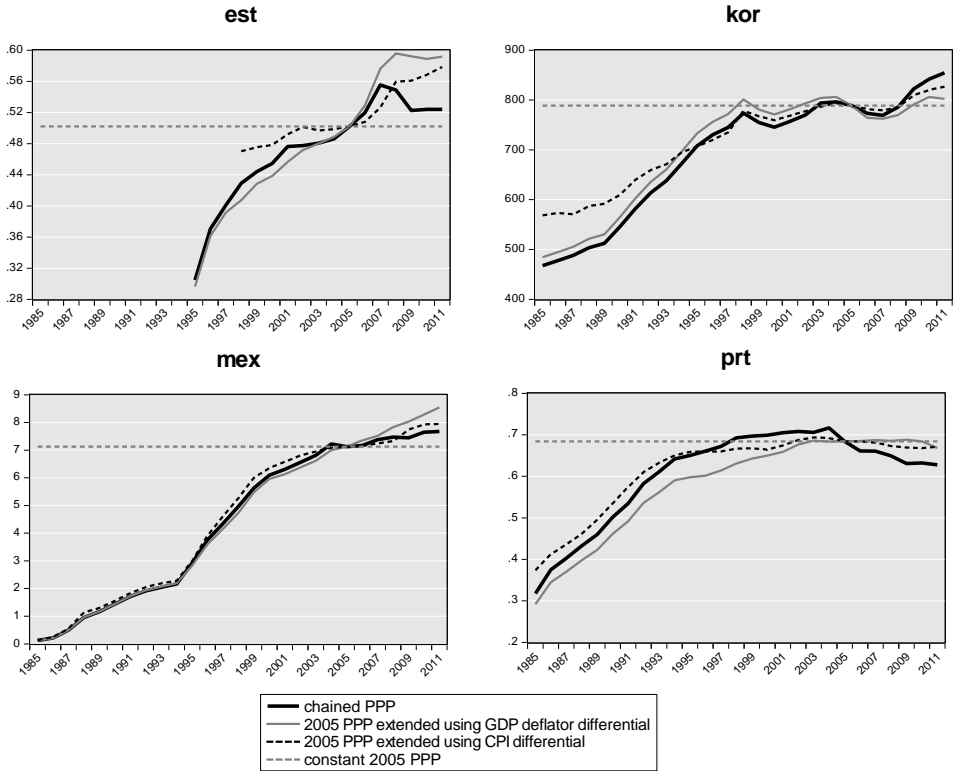
- There is for almost all countries a substantial wedge between constant 2005 PPP and other PPPs (chain-linked, GDP deflator and CPI-based PPPs)
- For some countries including Austria, Germany, France and Japan, the time-varying PPPs were above the 2005 PPP for most of the period under observation but the time-varying PPP measures were steadily converging to the 2005 PPP (Figure 4, panels A).
- For catching-up economies, the initially large negative gap is closing over time (Figure 4, panels B).
- For the last group of countries, there is no systematic gap between time-varying and constant 2005 PPP. The only exception is the CPI differential-extended PPP, which can deviate persistently from 2005 PPP (Figure 4, panels C).

Figure 4. Comparison of different PPPs over time, selected examples

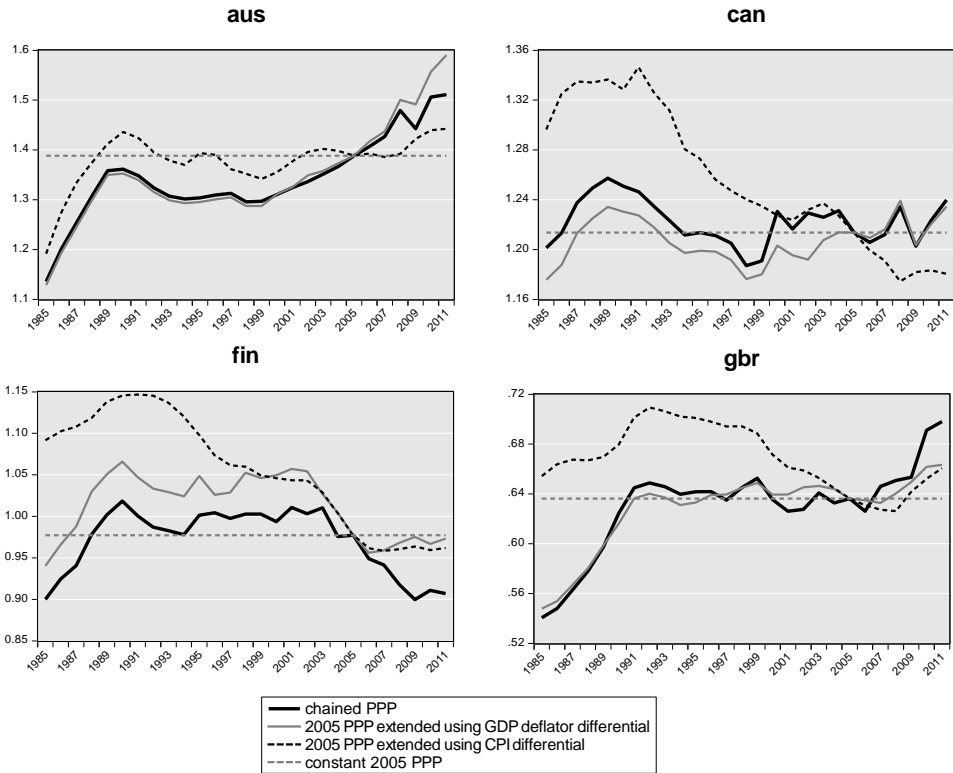
Panel A. Large positive gap narrowing over time between the time-varying PPP measures and the 2005 PPP.



Panel B. Large negative gap narrowing over time between the time-varying PPP measures and the 2005 PPP.



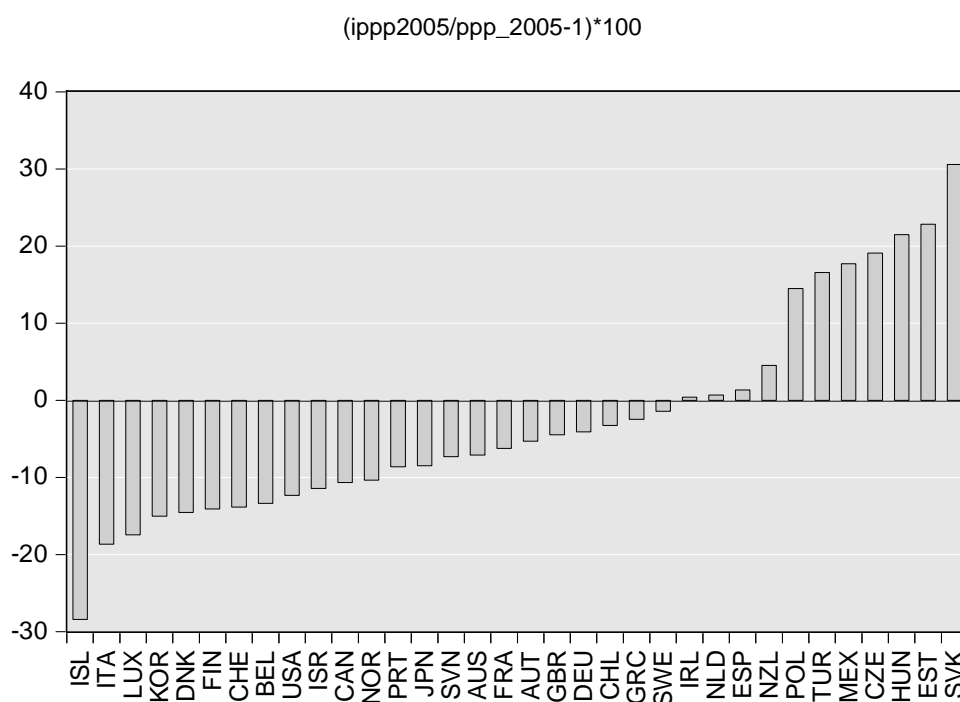
Panel C. No systematic gap between the time varying PPP measures and the constant 2005 PPP



Source: Author's calculations based on data obtained from PWT 8.0 and various OECD databases.

12. A final question related to PPPs is at what rates can the capital stock be converted into the foreign benchmark's currency. The conversion can be done using the rate applied to real GDP. Yet, the capital stock price level may be different from the overall price level. If productive capital is mainly composed of machinery and equipment and to the extent that these goods can be freely traded across countries, the conversion rate may be closer to the market exchange rate than the overall PPP. In such a case, applying the overall PPP rate to the capital stock may overstate the level of MFP (as it lowers the capital stock-to-GDP ratio). On the other hand, if the share of largely nontradable goods in the capital stock exceeds that in GDP, the use of the overall PPP rate will underestimate the true level of MFP. The difference between investment-specific PPP and overall PPP can be substantial. It is striking that investment PPPs in Central and Eastern Europe, Mexico and Turkey are systematically higher than the overall PPP. This implies that the relative price level of investment is closer to other countries', as expected, than the overall price level, largely influenced by the presence of non-tradable goods (Figure 5). Also, over time, the 2005 investment PPP extended by the investment deflator can show very large and persistent deviations from other PPP measures (Appendix B).

Figure 5. Comparison of overall and investment-specific PPPs in 2005

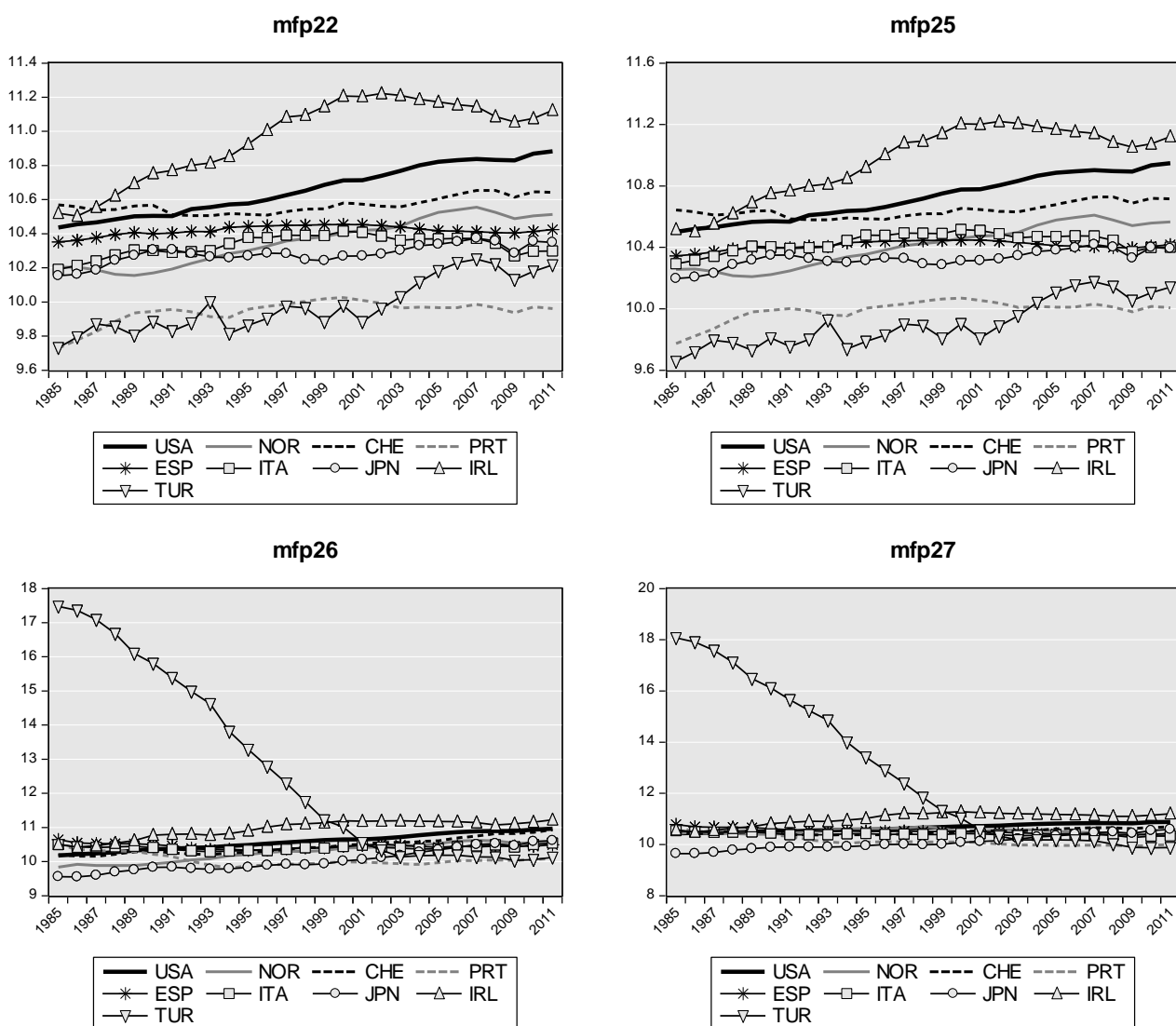


Source: Author's calculations using data obtained from PWT8.0.

13. In the MPF measures presented thus far, GDP was converted using 2005 PPPs (the ratio of the domestic and foreign absolute price levels in 2005) and the capital stock was converted using 2005 investment PPPs drawn from the Penn World Table 8.0. Two issues arise here. The first concerns whether investment-specific or economy-wide PPP conversion rates should be used for the capital stock. MFP25 shows that, as expected in line with Figure 5, applying economy-wide PPP to the capital stock reduces Turkey's MFP level for instance. But the difference is not very large. The second issue relates to the use of constant and time-varying PPPs. Using chained PPPs (MFP26) or 2005 PPPs adjusted for GDP deflator

differences over time (MFP27) result in an unlikely pattern in Turkey's MFP series: very high at the beginning and constantly decreasing. This is the result of the hyperinflation Turkey experienced around 2000. This suggests that time-varying PPP measures should not be used in such circumstances (Figure 6).

Figure 6. Comparison of MFP measures (including human capital) and different PPP conversion rates

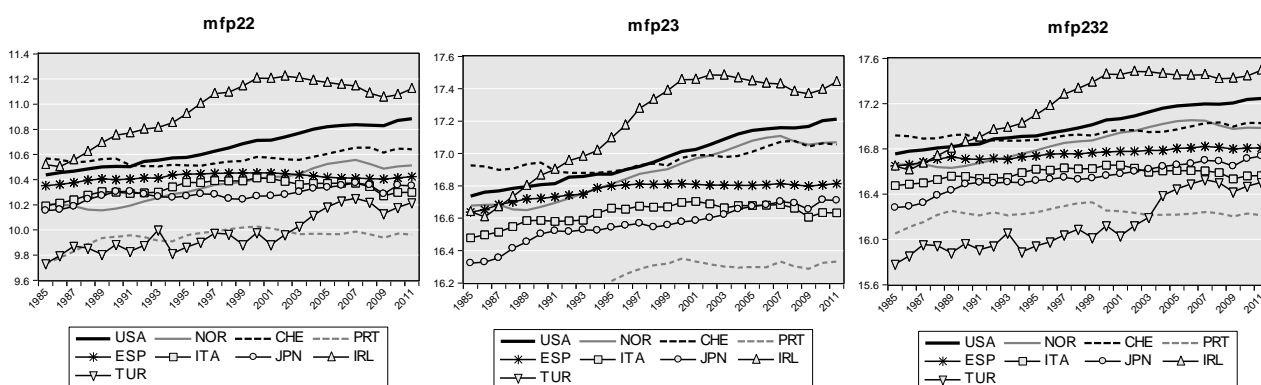


5. Labour input

14. **Whether to use employment or total hours worked makes a smaller difference.** The number of persons in employment (employees and self-employed) are used in the baseline OECD MFP calculations as a measure of labour input. This is something that can be measured with relatively high accuracy. Nevertheless, total employment measures may measure badly total labour input if the average yearly hours worked per employed person differs across countries (see e.g. OECD, 2011). Such a divergence can happen for a number of reasons: i) regulation may put an upper ceiling on weekly hours worked such as France's 35-hours week work, ii) different number of bank holidays and the regulations regarding the level (and the

evolution over a career) of holidays a person can take per year, iii) the share of part-time workers, and iv) the share of self-employed and entrepreneurs in total employment (as they can work as many hours as they wish). Using hours worked is desirable from a methodological and practical viewpoint.¹⁷ However, changing employment to hours worked in MFP23 (including human capital) does not substantially change the overall profile of the series beside some shifts in relative levels (Figure 7). Total hours worked obtained from PWT8 (MFP232) are longer (e.g. Portugal) and available for more countries (e.g. Turkey) than those drawn from OECD countries (MFP23).

Figure 7. Comparison of MFP measures with total employment (MFP21) and total hours worked (MFP23)



Note: MFP23: total hours worked obtained from the OECD Economic Outlook database, MFP232: total hours worked obtained from PWT8.0

6. Capital stock

15. The way physical capital stocks are constructed can matter. There are three broad ways of measuring capital stock: i) using the perpetual inventory method (PIM), relying on investment and depreciation series, ii) using capital stock surveys, or iii) using balance sheet data. PIM is feasible in a wider set of countries, but it has limitations. First, the choice of the initial stock has to be based on some assumptions and it has a large influence on cross-country capital stock levels. Second, whether or not the depreciation rate is allowed to change over time can influence the level of, and the change in, the capital stock. The depreciation can change over time and in a cross-country comparison as well. Different types of capital may depreciate at different speed. If the composition changes over time and if it differs across countries, the depreciation rate will also change. In particular, investment in ICT has a much faster depreciation rate than other types of investment. Countries that started investing more heavily in ICT may have higher depreciation rates. More investment in ICT over time also implies increasing depreciation rates. These countries will also have different depreciation rates than countries that invest less in ICT.

17. There are problems related to measuring hours worked by the self-employed are based on self-reporting. Legally restricted weekly hours can be circumvented by supplementary contracts including a flat-rate remuneration of overtime work. Officially reported hours worked will not be able to capture this overtime work. But measurement problems might be even greater for total employment. Employment numbers for example will include those self-employed who work for example the bare minimum needed for inclusion, i.e. one hour a week.

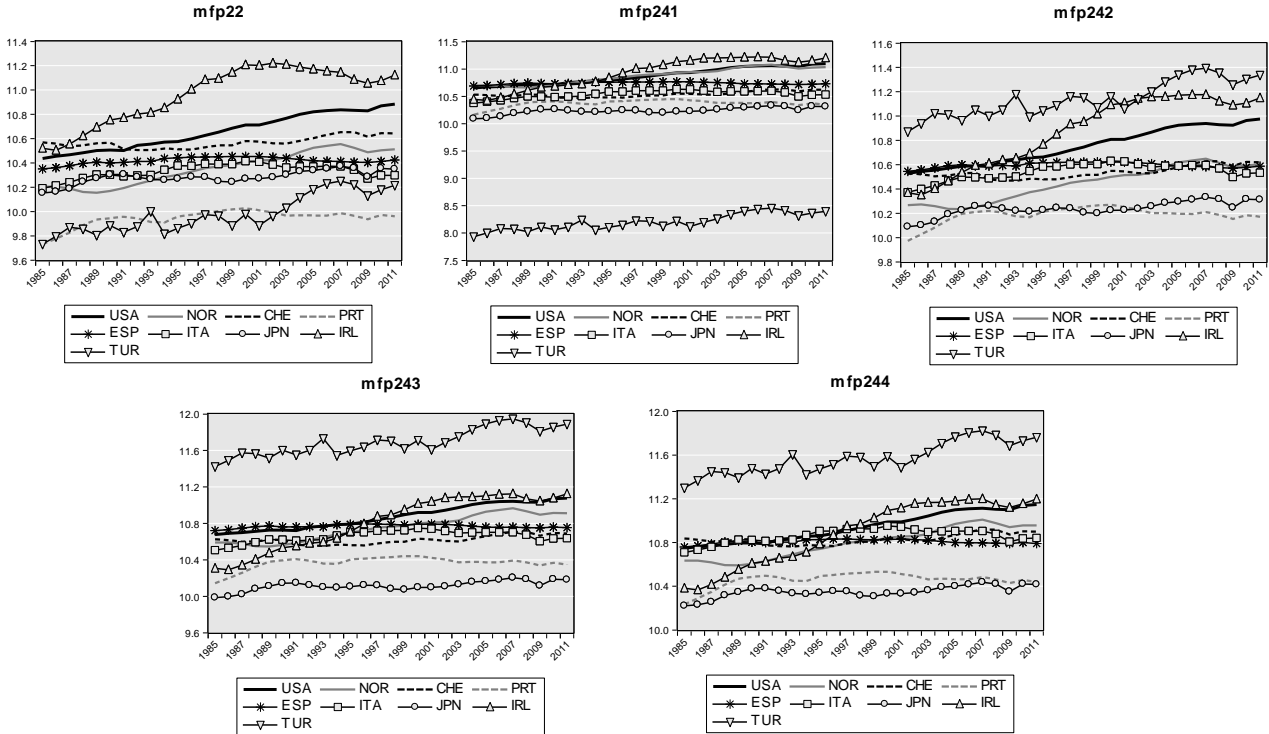
16. MFP12, MFP21, contain capital stocks with time-varying depreciation rates for all countries (from World Penn Table 8.0). MFP24 is calculated on the basis of capital stock series available from the OECD Economic Outlook database. For the core OECD countries, capital stocks are constructed using statistical surveys.¹⁸ For the remaining countries, capital stocks are calculated using the PIM.¹⁹ Using these data for level comparisons may be difficult: the base year used for individual countries can differ. A case in point is Turkey with a base year of 1998. Figure 8 shows how much lower Turkey's MFP series (MFP241) is compared to the other selected OECD countries. Rebasings all series to 2005 seems a solution. However, only real GDP can be rebased, as no capital stock deflators are available for the same series. The result, MFP 242 is depicted in Figure 8. There is little change for most countries (as they mostly have a base year of 2005). Nevertheless, the consequences are substantial for Turkey: rebasing GDP but not the capital stock, given the large gap between 2005 and 1998 raises MFP to a large extent. A more recent vintage of capital stock and real GDP series are based on 2010. The problem with 2010 is that the base year for all PWT series is 2005. Nevertheless, capital stock deflators are available for all but about five countries with different base years. Turkey is one of those, hence the additional increase in its MFP level (MFP 243). The picture remains broadly similar if 2010 PPP is used instead of the 2005 PPP to convert the series into a common currency. While the dynamics of the series may not change too much, it is clear that these series should not be used for cross-country comparisons.

17. An easier way to handle different base years is to obtain country-specific MFP series domestic currency measures and the same base year for all components (output and capital stock). The MFP measure obtained can then be converted into a common currency using constant price PPPs. In this way, only the output (MFP) is converted, not the individual inputs.

18. These countries are as follows: Australia, Austria, Belgium, Canada, Switzerland, Germany, Denmark, Spain, Finland, France, United Kingdom, Ireland, Italy, Japan, Korea, the Netherlands, New Zealand, Portugal, Sweden and the United States.

19. These countries are the following: Chile, Czech Republic, Estonia, Greece, Hungary, Iceland, Luxembourg, Norway, Poland, Slovakia, Slovenia, Turkey, Mexico and Israel.

Figure 8. Comparison of MFP measures with alternative capital stock measures



Note: MFP241: base year such as in Economic Outlook 95, usually 2005; MFP242: real GDP rebased for 2005 for all countries, capital stock not rebased, MFP243: real GDP and capital stock rebased for 2010 if possible, conversion done using 2005 PPP; MF244: real GDP and capital stock rebased for 2010 if possible, conversion done using 2010 PPP

7. Plausibility checks

18. We rely on a set of plausibility checks, which would help to choose from the alternative MFP measures. MFP measures will be considered as realistic if they pass the following filters:

1. The top performers should include countries that plausibly belong to this group. In particular, the United States should be close to the frontier.
2. MFP in the countries close to the frontier should be broadly rising (and at least not declining) over decades. More generally, countries should not have a trend decrease in MFP over the sample period covering three decades.

19. Based on these criteria, the measures MFP22 and MFP23 can be viewed as reasonable measures of MFP. MFP 12 to 17 suffer from the human capital bias: counterintuitive relative levels and trend decline for a number of countries. MFP24 faces the problem of cross-country comparability due to the different base years of the series, a major problem being Turkey. MFP25 appears OK in terms of criteria 1 to 3. Nevertheless, MFP 22 and MFP23 are more convincing conceptually: the capital stock should be converted

using capital stock-specific prices rather than economy-wide prices. Finally, MFP26 and MFP27, based on time-varying PPPs, show weird patterns for catching-up countries.

8. Concluding remarks

20. Measuring multi-factor productivity correctly at the macroeconomic level and understanding its main drivers is key to assessing the impact of pro-growth reforms. Progress has been made in measuring multi-factor productivity (MFP) and understanding its determinants at the firm and industry levels (O'Mahony and Timmer, 2009; Syverson, 2011; Andrews and Cingalo, 2014; Bartelsman et al., 2014). Considerable work has been done in recent years to measure MFP but MFP is to some extent a measure of what we do not know. For instance, measures of MFP capturing the contribution of human capital within MFP may distort spatial and temporal comparisons. Indeed, the uncertainty about the precise level, comparability across countries, and the evolution over time has important implications for how precisely one is able to pin down the drivers of MFP in an empirical analysis. Not measuring accurately the MFP frontier masks our ability to reveal how far countries or firms are from the frontier and may lead to flawed conclusions regarding the speed of convergence to it. A wide margin of error may produce dubious results regarding the role of various institutions and policies for the evolution of MFP

21. This paper carried out a systematic sensitivity analysis of MFP measurement by comparing a set of alternative MFP measures, with human capital included or excluded, with different measures of Purchasing Power Parity (PPP) exchange rates, using time-varying capital depreciation rates and different measures of capital stock and labour input (headcount against hours worked). The main result of the paper is that whether or not human capital is included in MFP makes a significant difference for the level and dynamics of MFP. At the same time, MFP measures are less sensitive to other parameters of the calculation.

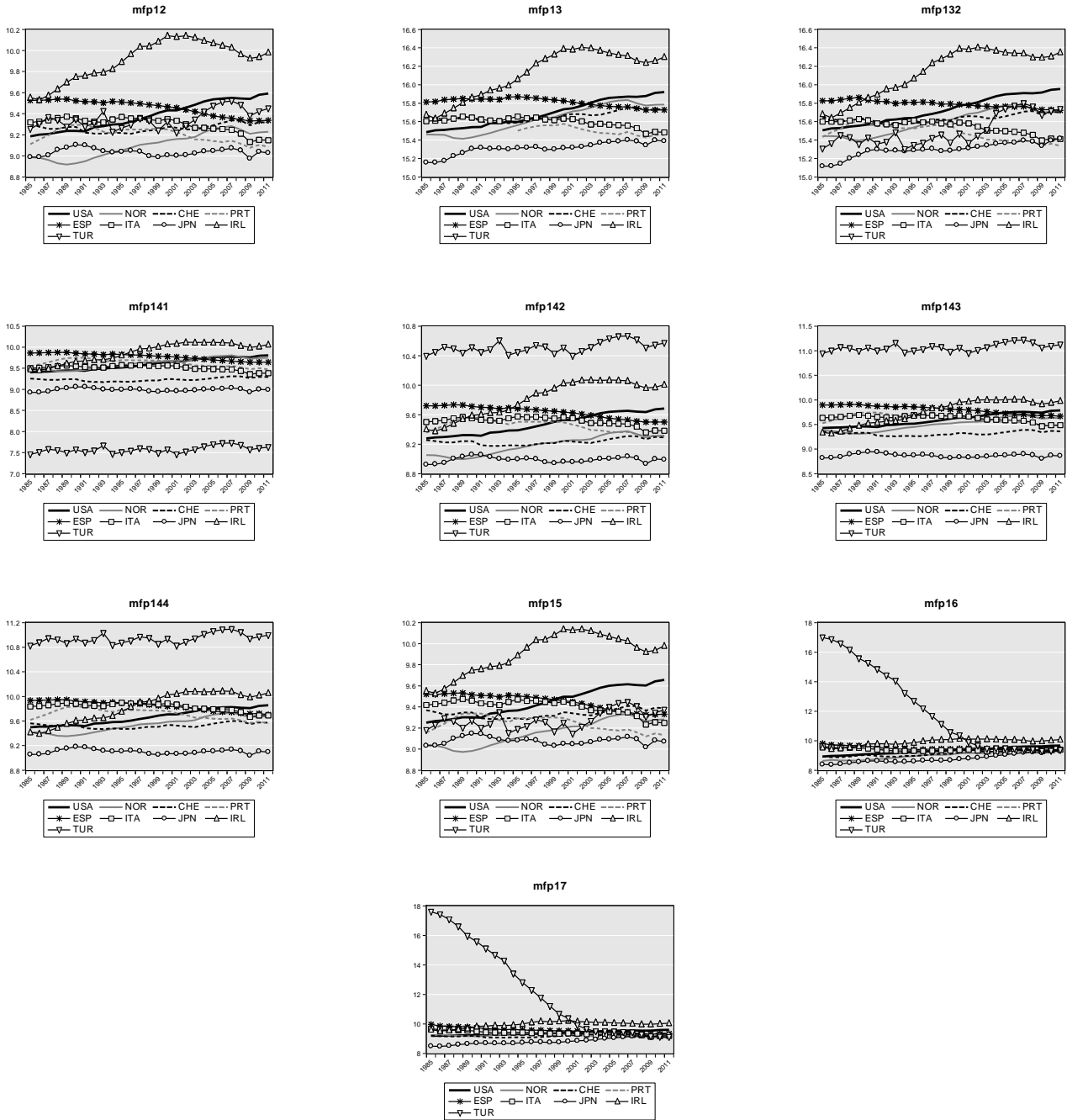
REFERENCES

- Andrews, D. and F. Cingano (2014), Public policy and resource allocation: evidence from firms in OECD countries, *Economic Policy*, 29(78), 253-296.
- Andrews, D. and B. Westmore (2014), Managerial capital and business R&D as enablers of productivity convergence, *OECD Economics Department Working Papers*, No.1137, OECD Publishing.
- Égert, B. (2007), Real convergence, price level convergence and inflation differentials in Europe, CESifo Working Paper No. 2127.
- Eurostat-OECD (2016), Eurostat-OECD methodological manual on purchasing power parities (PPPS), OECD Publishing.

- Feenstra, R. C., R. Inklaar and M. Timmer (2013), PWT 8.0 – A user guide. http://www.rug.nl/research/ggdc/data/pwt/v80/pwt_80_user_guide.pdf
- Gal, Peter (2013), Measuring total factor productivity at the firm level using OECD-ORBIS, *OECD Economics Department Working Papers* No. 1049.
- Johansson, A., Y. Guillemette, F. Murtin, D. Turner, G. Nicoletti, C. de la Maisonneuve, P. Bagnoli, G. Bousquet and F. Spinelli (2013), Long-term growth scenarios, *OECD Economics Department Working Papers* No. 1000
- Madsen, J. (2007), Technology spillovers through trade and TFP convergence: 135 years of evidence for the OECD countries, *Journal of International Economics*, 72(2), 464-480.
- Madsen, J. B. (2014), Human capital and the world technology frontier, *The Review of Economics and Statistics*, 96(4), 676-692.
- OECD (2017), *OECD Compendium of Productivity Indicators 2017*, OECD Publishing.
- O'Mahony M. and M.P. Timmer (2009), Output, input and productivity measures at the industry level: the EU KLEMS database, *The Economic Journal*, 119, 374-403.
- Syverson, C. (2011), What determines productivity, *Journal of Economic Literature*, 49(2), 326-362.
- Westmore, B. (2013), R&D, patenting and growth: The role of public policy, *OECD Economics Department Working Papers*, No. 1047, OECD Publishing.

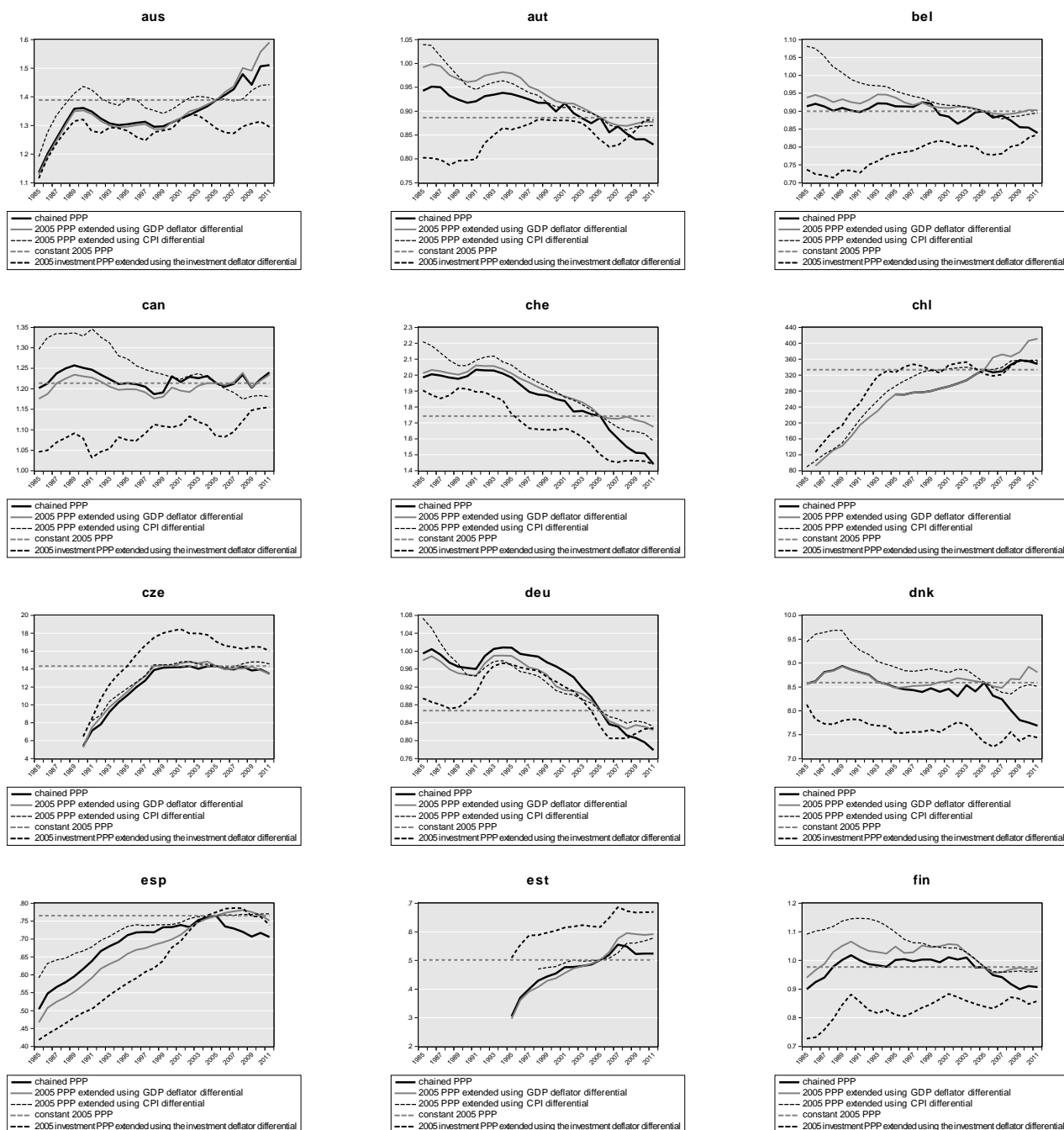
APPENDIX A

Figure A1. Alternative MFP measures including human capital, 1985-2013

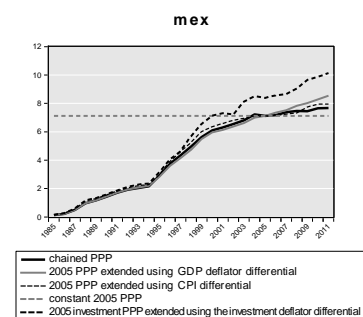
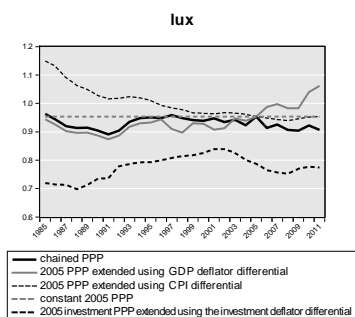
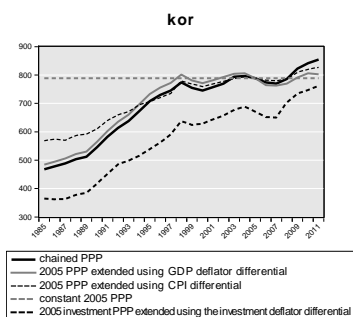
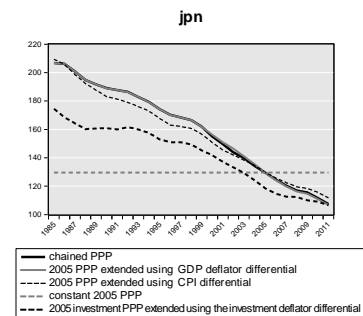
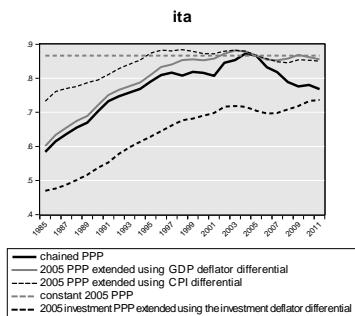
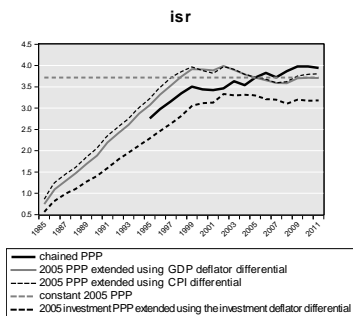
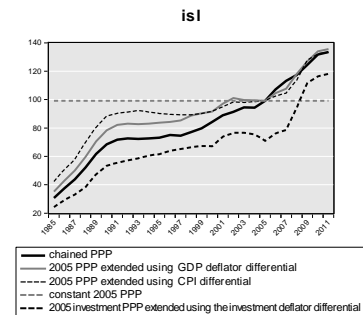
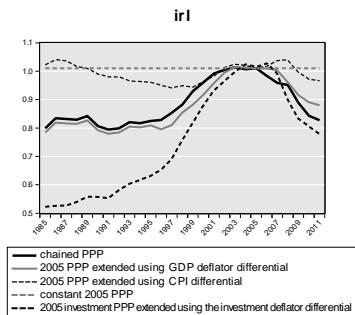
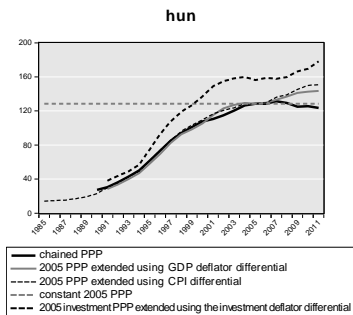
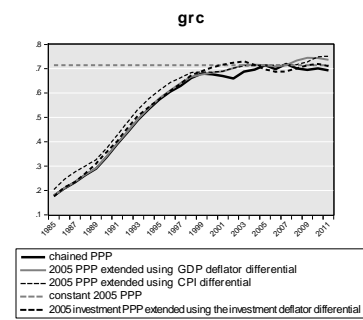
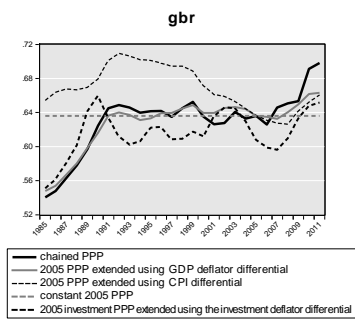
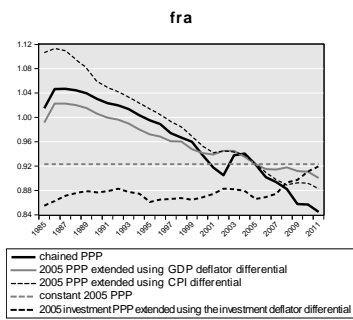


APPENDIX B

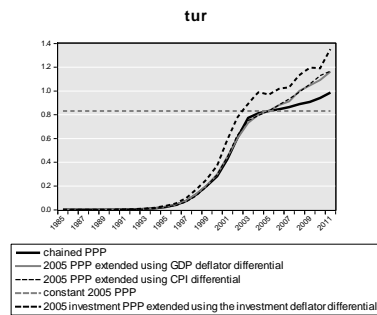
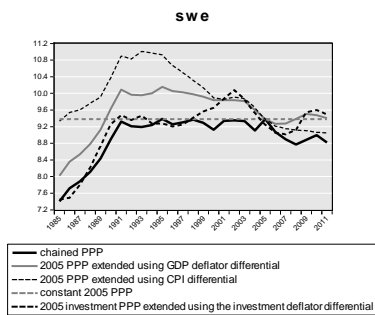
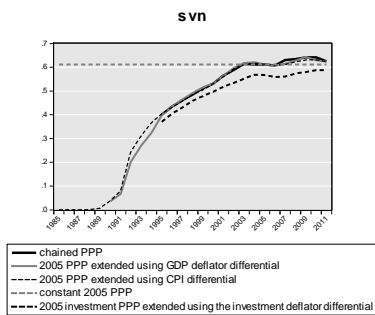
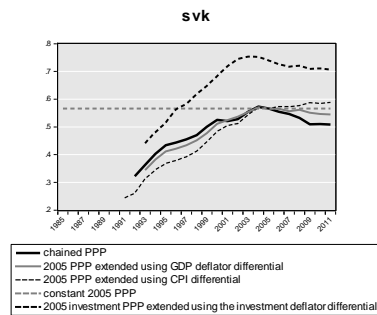
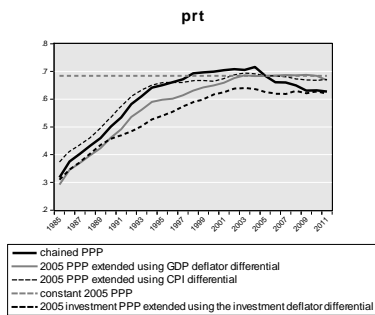
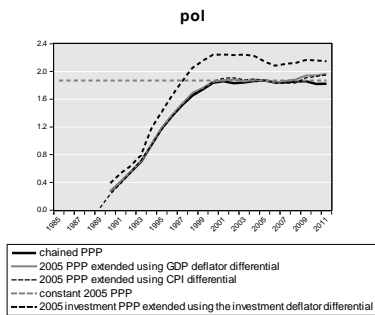
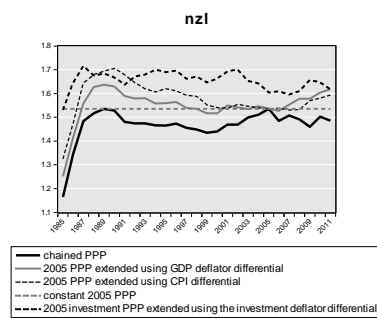
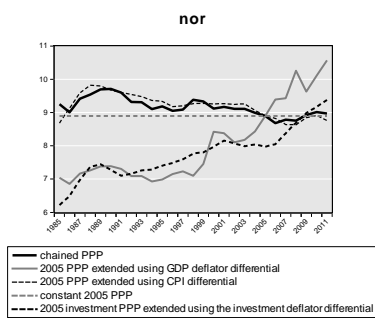
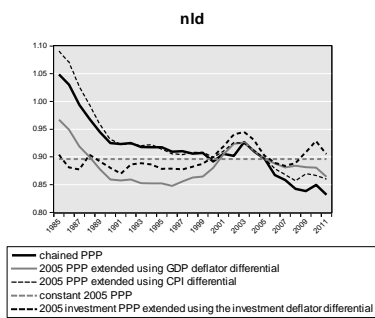
Figure B1. Comparison of the evolution and level of different PPP measures



Note: PPP_X Denotes chain-linked PPP, PPP_PGDP_X Is constant 2005 PPP extended using GDP deflator differentials, PPP_CPI_X Is the CPI-differential adjusted constant 2005 PPP and PPP_2005_X denotes constant 2005 PPP. IPPP_PIT_X is constant 2005 investment-specific PPP adjusted for the investment deflator differential.



Note: PPP_X Denotes chain-linked PPP, PPP_PGDP_X Is constant 2005 PPP extended using GDP deflator differentials, PPP_CPI_X Is the CPI-differential adjusted constant 2005 PPP and PPP_2005_X denotes constant 2005 PPP. IPPP_PIT_X is constant 2005 investment-specific PPP adjusted for the investment deflator differential.



Note: PPP_X Denotes chain-linked PPP, PPP_PGDP_X is constant 2005 PPP extended using GDP deflator differentials, PPP_CPI_X is the CPI-differential adjusted constant 2005 PPP and PPP_2005_X denotes constant 2005 PPP. IPPP_PIT_X is constant 2005 investment-specific PPP adjusted for the investment deflator differential.