Benchmark Reconciliations Revisited

Bettina Aten and Alan Heston[1]

Summary; Producers of benchmark comparisons are naturally interested in whether the results from one benchmark are consistent with a later benchmark using growth rates between the two to produce a comparison in the later year. Our interest is somewhat different. We recognize that there are differences in price samples, weighting and other factors that will always lead to some differences between extrapolated and new benchmark estimates. In the spirit of some strands of research on the statistical discrepancy in national accounts, we focus on the question of whether there is anything to be learned about the world from examining such differences. The technique we employ looks at several pairs of OECD benchmark comparisons for 1990s, first 1990-96 as officially reported and then less intensively at other benchmarks in the 1990s using a common aggregation procedure. The approach is to examine whether there is strong autocorrelation among the residuals of the countries. We then examine some possible influences that might reduce this auto-correlation, including price similarity, distance and a measure of real exchange rates between each pair of countries. The main conclusion of the paper is that our measure of real exchange rate movements has a strong explanatory effect. While this is a suggestive result it would require more research to establish the nature of any causal relationships.

Introduction

The question we ask is how one should think about multiple benchmark results. Suppose for example, that the per capita GDP in country A were 80% of B in 1990 and between 1990 and 1996 A reports 2% per capita GDP growth and B reports 12%. The rough expectation is that a 1996 benchmark would find A at about 73% [(1.02/1.12)*80%] of the B GDP per capita. If the 1996 benchmark put A at 77% of B we would be happy to find the direction is correct but unhappy that the 1996 extrapolation and benchmark are off by such an amount. In this example one could describe the discrepancy in several ways. A benign description would be that the error in the level of GDP of A relative to B was about .04 on an average benchmark level of .785, or 5%. However, one could also express the .04 discrepancy relative to the difference between the two benchmarks, namely .04/.(80-.77), and the difference is much larger. Still another measure of the discrepancy would be the difference in the growth rates (deflators) that is inherent in the two benchmarks. In the above example, the national accounts growth rates of A is 2% and the growth rate inherent in two benchmarks (assuming the growth rate of B is truth), would be 8%. The difference in the two growth rates would be 6%, representing more than a 100% difference between national growth rates and those inherent in two benchmarks.

It is the latter discrepancy that is dealt with in this paper. For brevity we will describe the difference, $\Delta_{0,t}$ as the difference between the national growth rate and that

[1] The authors are from Bridgewater State College and the University of Pennsylvania. Many of the ideas in this paper have been developed in collaboration with Robert Summers, who, of course, is not responsible for remaining errors. Support from the National Science Foundation (Award # 9911377)) is gratefully acknowledged.

implied by two benchmarks; namely, $Ä_{0,t} = r - r^*$ where r is the national accounts growth rate, and r^* is the implied growth rate between benchmarks. An implied growth rate between two benchmarks in current prices can only be calculated with respect to some referent, like the average of the group. The definition of the implied growth rate is: $r^* = [BM_t(1+r_r)/ BM_0 (1 + r_i)] - 1$, where r_r is the growth rate of the reference country or country group between benchmarks and r_i is the national growth rate of any other country in the benchmark. It should be remembered that the absolute value of the Äs will be the same as those obtained working with changes in benchmark PPPs extrapolated by the relative movements of GDP deflators. In our discussion we will sometimes move back and forth between PPPs and growth rates.

In Part A of this paper we briefly indicate some of the obvious reasons associated with data differences that are frequently used to explain Äs. This section is brief in part because Seppo Varjonen (2002) has provided a very thorough discussion of the issues in the context of the OECD countries. In Part B we discuss the way that these Äs have been treated in some past applications. In Part C a different framework for the thinking about the Äs is proposed. Part D provides some illustrative results for the OECD countries in the 1990s.

A. Extrapolations vs Benchmarks: Some Data Issues

Current and Constant Prices

One major data issue relates to current and constant prices. In the illustration of Countries A and B above, the two benchmark results are in current prices of 1990 and 1996 respectively. However, the extrapolations are typically based on a constant price GDP measure for each country. The constant price measures may be constructed from indexes of price changes based upon direct collection or on an indirect estimate obtained by dividing current price aggregates by growth of one or more physical quantity indicators. The sample of items and weights used for direct and indirect price deflation will differ between countries, and that alone would explain some of the discrepancies embodied in the Äs.

Benchmark Price Samples

In general the bundle of items priced in two benchmarks will consist of items common to both, items only in the first benchmark and items only in the second benchmark, the usual situation in temporal comparisons.[2] Further the weights applied to these price relatives will differ between benchmarks and will contribute to the Äs. Whether these differences will be offset or accentuated by the constant price measures is not clear.

Benchmark Country Composition

If the number of countries involved in two multilateral benchmark comparisons differs in two years it will also influence the relationship between any pair of countries; or

[2] This statement needs qualification for the EU and OECD countries because their item selection is built up from matching of items between country pairs where the item is considered representative in one of the two countries and is either representative or commonly available in the other.

the relationship between any country and the average of the group. While only a partial list of data issues in benchmark reconciliations it suggests some of the major problems. One could also treat the issue of trading gains and losses as a data issue, but it is dealt with in Part C of our presentation.

B. Approaches to Benchmark Reconciliation

The Eurostat Approach

Eurostat has dealt with these problems in a straightforward manner by moving to annual benchmarks where many of the data issues are greatly reduced. In the Eurostat framework new price surveys are made every 3 years for a rotating group of expenditure items and headings. In the intervening years Eurostat uses national price indexes at a detailed heading level to update parities. In this way the effect of different expenditure weights between countries is substantially reduced as a source of data error. There still may be differences due to changing the sample of goods priced and the country composition of benchmark comparisons.[3] The question is whether at the end of the day the Eurostat or any other reconciliation method is the right way to model what is going on at levels of aggregation like GDP.[4]

Other Reconciliation Approaches

The OECD comparisons including the EU countries are done at 3 year intervals. In his paper, Varjonen (2002) has examined the Äs for disaggregated expenditure headings to find out where differences arise. He finds that government and investment have larger Äs than consumption. This is clearly an important research direction to explore particularly for organizations that are producing benchmark comparisons.

In earlier work Summers and Heston (1984) pursued a much less detailed but analogous approach decomposing differences between extrapolations and benchmark estimates of domestic absorption into consumption, government and domestic investment. This exercise found a similar pattern of Äs for the 1970 and 1975 benchmarks as Varjonen, in that investment and government differences were more pronounced than consumption. The approach in the 1984 paper was to use a "consistentization" procedure that assumed errors could occur in national growth rates as well as in the benchmark estimates. In terms of Eurostat procedures this would be equivalent to saying that the time to time country price deflators may contain errors as well as the benchmark parities. Because of the reluctance of countries to accept adjustments of their national indexes of growth and price change, we have not pursued this approach in developing PWT 5.6 and 6.0.

---

[3] A way to deal with this problem is to make estimates for the countries common to both benchmarks, and to only use the results for the larger set of countries to link in the new countries. If the results for the original group of countries are preserved, there are a variety of ways to bring in the new countries.

[4] It should further be noted that the way in which Eurostat uses prices in the 'Star' item method to build up multilateral parities, actually works to reduce the number of prices entering into the construction of a heading parity compared what each country would include in their time to time indexes for the heading. While this may introduce some noise into the estimation procedure, there does not appear to be any clear direction in which this would affect results.

However, it is still worth noting that the two components of expenditures for which the Äs are largest, government and investment, are also the areas for which the national accounts deflators are more questionable. Productivity assumptions must be made for government employees and for construction on the basis of very weak data; or it is necessary to come up with some output indicator for government and construction that is independent of expenditures. In addition many items of producers durables are difficult to price because of quality changes and terms of sale. The focus of this paper, however, is not on reconciling benchmark estimates solely on the basis of errors in growth rates or the sample of items priced and their importance, but rather on macro-factors. By macro-factors we mean influences that lead to departures of the national price level of a country from what is typical of countries of their size, level of income, and degree of involvement in international trade. Some of the variables affecting national price levels have been discussed by Kravis and Lipsey (1987)

C. Another Look at the Problem

It is our view that not all differences between extrapolations and benchmarks can be reconciled, either by the Eurostat approach or by the Summers and Heston (1984) approach. In addition to the macro factors mentioned above and to be developed further below, there are also some conceptual issues in the reconciliation process. We begin with one of the conceptual issues, the difference between Gross Domestic Income and Gross Domestic Product.

Trading Gains and Losses

Benchmark comparisons are done in current prices so they incorporate any gains or losses that may have accrued to countries because export or import prices had moved in different directions from each other and the domestic price level. If a reconciliation is being carried out at an aggregate level, then conceptually one would want to include these gains and losses in the constant price series. Put another way, it does not make much sense to seek to reconcile GDP benchmarks and extrapolations at the GDP level, since in principle, the latter take out the effects of trading gains and losses in deflation of exports and imports. If an aggregate is to be reconciled, domestic absorption is clearly preferable to GDP. However, the way in which exports and imports have been valued in benchmark comparisons does not usually correspond to the practice in national income accounts. This means the above comments do not necessarily apply where exports and imports have been converted at exchange rates in benchmark comparisons, as is current practice in the EU and OECD. However it does remain the case that it make more sense to reconcile estimates of domestic absorption rather than GDP.

The 1993 SNA considers three methods of producing Gross Domestic Income estimates, which modify constant price GDP figures to take account of trading gains and losses. The preferred method is to take the difference between constant price exports and imports that enter GDP and subtract from those the current net foreign balance divided by the deflator for domestic absorption. Explicitly the trading gains, TG, are:

$$(1)\quad TG = \frac{\text{Current Exports}}{\text{Export Deflator}} - \frac{\text{Current Imports}}{\text{Import Deflator}} + \frac{(\text{Current Exports} - \text{Current Imports})}{\text{Domestic Absorption Deflator}}$$

Is this any of this likely to be of practical importance? Not surprisingly that depends on the country and the extent of its trade and the pattern of its exports and imports. When trading gains or losses are added to GDP to produce GDI and expressed as a ratio to GDP on a 1996 base most large countries do not experience great departures of the ratio from 1.0 and only moderate variation. For example the US has an average of 1.015 over the period 1950-98 using domestic absorption deflator to estimate trading gains and losses. Though moderate even for the US the coefficient of variation (CV) is 1.3%. For smaller countries with unusual balance of payments patterns, the CV is much higher, 6.6% for Switzerland and 8.5% for Luxembourg; further for Switzerland the average ratio of GDI to GDP of .909 indicates a major upward trend in GDI relative to GDP over the period prior to 1996.

A symmetric treatment of benchmark comparisons would call for deflating the net foreign balance by the deflator for domestic absorption, which independently of these considerations has been the practice in PWT.[5]. When the Äs for 1990 -96 for the OECD countries were compared with the average of trading gains and losses over the period both expressed as a percent of GDP the relationship was not close. Given the benchmark treatment of the net foreign balance in the OECD, this is not surprising. It does mean that the Äs remain to be explained.

Growth Factors

Countries with high growth rates of output typically experience sectoral shifts in output and price structures, which may in turn be in reflected in relative prices of final expenditures. Most commonly these shifts affect the relative prices of traded and non-traded goods. Not only are these changes difficult for national statistical offices to measure, they are also likely to have an impact on the national price level (NPL) of a country. By national price level we mean the PPP over GDP divided by the exchange rate.

Our conjecture is that countries that grow very rapidly may have larger differences in extrapolations and benchmarks, other things equal. In the empirical work, this idea is extended to taking into account the pair-wise growth experience of OECD countries.

Exogenous Exchange Rate Changes and PPPs

A national price level is an indicator of relative prices in one country versus another given the exchange rate and the expenditure distribution over GDP. The national price level is not necessarily an indicator of the level of trade competitiveness of a country, but a decline (rise) in the NPL usually signals a rise (decline) in competitiveness.[6] For countries with high rates of inflation, most changes in the NPL arise from changes in the numerator, the PPP. However, for countries with relatively

[5] The reason the deflator for DA has been used in PWT is that it is a base country independent as compared to using the exchange rate of one country, like the United States. Use of an exchange rate with respect to the group of countries is base country invariant, though the average should be weighted in the same way as the aggregation procedure for consistency. This would not be the case for the Euro in the EU for example.

[6] National price levels will involve only a portion of the goods and services of a country's international trade. In addition NPLs will also be affected by the current exchange rate that in turn will depend on expectations regarding the future exchange rate of a country due to relative interest rates and the like.

stable prices, most changes in the NPL arise from changes in the exchange rate. In the discussion that follows, two points are argued. First, fluctuations in NPLs can be quite large and persist for a number of years. And second, when the NPL rises, it signals a less competitive trade situation that will tend to raise the price of non-traded goods in a country relative to traded goods. The opposite will happen if the NPL falls primarily due to a drop in the value of a country's exchange rate.

A glance at Table 1 suggests that for the United States there has been a lot of variation in its NPL over the past 25 years (Heston and Summers, 1993). Column (1) of Table 1 presents an index of the price level of the US expressed relative to the 4 largest EU countries with 1990 = 100.[7] If there is a correspondence between movements in exchange rates and relative inflation rates in the U.S. and these countries the index in column (1) of Table 1 would remain around 100. This is clearly not the case. As noted national price level fluctuations can occur when relative inflation rates differ, while exchange rates remain the same; or when exchange rates change and relative inflation rates are similar; or some combination of the two.

The other two columns in Table 1 provide measures of the US exchange rate and suggest that the price level movements of the United States are primarily related to exchange rate movements. This is shown in Column 3 where the nominal rate of the US$ to the Special Drawing Rights (SDR) of the IMF is given. Column 2 is the real effective exchange rate of the US$ as estimated by IMF, which is a measure of changes in the dollar prices at exchange rates of goods that the United States buys and sells, weighted by country specific exports and imports. When the effective exchange rate rises, it means the relative cost of importing goods has fallen and the cost of exporting has risen. These relationships are also represented in Figures 1 and 2. (Table 1 and the Figures are from Heston, Summers and Aten (2001).

In both Figures 1 and 2 the index of the US price level relative to the four EU countries is on the vertical axis. In Figure 1, the real exchange rate beginning with the 1975 benchmarks, is on the horizontal axis, and there is a strong positive relationship (r = .98). In Figure 2, the exchange rate or the US$ to the SDR is on the horizontal axis, and we find that there is a strong negative relationship (r = -.97) with the index of the U.S. price level. So the index of the US price level is not flat across benchmarks, but is subject to major shifts that are driven by movements in exchange rates.

[7] When the U.S. is used as numeraire the price level is taken as 100 and variation from benchmark to benchmark shows up in other countries. To take this into account in Table 1 we first express the price level of the U.S. relative to the four largest countries in the European Union that have participated in all of the benchmark comparisons, namely France, Germany, Italy and the U.K. In this method of presentation, the US price level can vary from benchmark to benchmark; and a glance at Column (1) of Table 1 shows that it most certainly does.

Figure 1: Relation of US Price Level To Real Exchange Rate

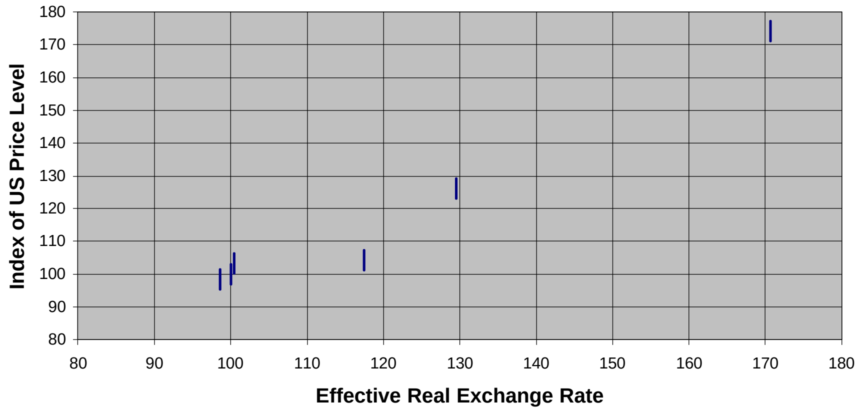


Figure 2: Relation of US Price Level to US$ per SDR Rate

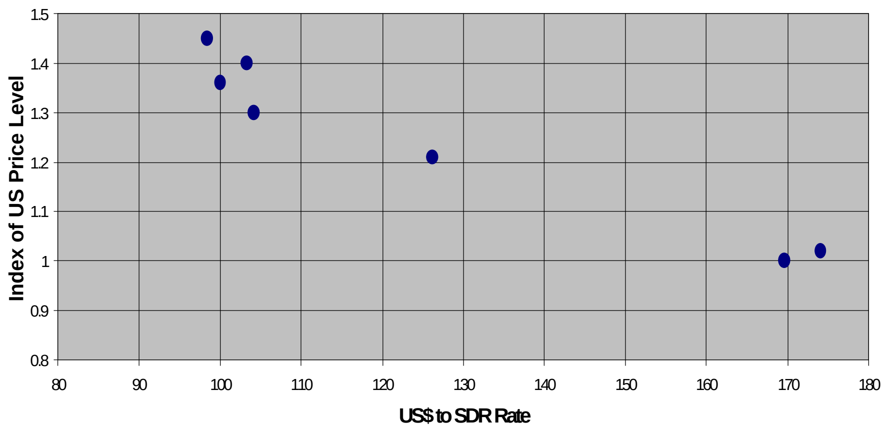


Table 1: Movements of the US Price Level and Measures of the US Exchange Rate

Benchmark Year	Index of US Price Level to EU4 (1990=100) (1)	Effective Real Rate of Exchange (1990 = 100) (2)	$ Rate to SDR (3)
1970	169.6	Not available	1.00
1975	126.2	129.5	1.21
1980	104.2	117.4	1.30
1985	174.1	170.6	1.02
1990	100.0	100.0	1.36
1993	103.3	100.5	1.40
1996	98.4	98.6	1.45

This major divergence from a flat relationship illustrates our first point that NPL can be very large. The most striking instance of course is 1985 when Fed policies supported a very strong dollar. Many would say this was at the cost of a world recession and the creation of a rust belt in the Midwestern states.[8] These consequences of a strong dollar occur because at least in the U.S. the gains that manufacturing firms realize from lower cost of imported inputs appears for the years 1975-98 to be more than offset by the loss of revenue from manufacturing exports (Goldberg and Crockett, 1998).

This brings us to the second point about departures of the NPL of a country from its expected level, namely the structural effects. With regard to the United States one special thing to note is that the dollar has played an asymmetric role in the world financial system over this whole period. Some countries, including individuals and central banks have been willing to use the U.S. dollar as a reserve currency often allowing it to appreciate significantly for a number of years. Other countries, such as Luxembourg and Switzerland, have also played a role in world financial markets well beyond their economic size but with less influence than the United States. However, this point about the U.S. only means that structural effects of persistent national price level fluctuations may require correction less quickly than for countries that do not have a reserve currency role

How is this structural effect related to relatively persistent departures of national price levels from average values typical for a country? During periods when the NPL of a country is rising primarily because of exchange rate appreciation a country is likely to experience a reduction in the prices of traded goods relative to non-traded goods. This may but need not be related to Dutch disease type phenomena. Beyond the shift in relative prices, the consequences for an economy are typically to hurt exporters of

---

[8] This is strikingly revealed in a research report of the OECD (Scarpatta, Bassanini, Pilat and Schreyer, 2000) that shows the maximum and minimum values of the national price levels observed in various benchmarks relative to the US. What this report shows is that for the benchmarks from 1980-1996, the maximum PPP values to the US$ for the 22 OECD countries that participated in the 1985 benchmark are all observed for 1985. Though it is not shown here, this would also be true if the 1975 and 1970 benchmarks comparisons were considered. However, as indicated in Table 1, the index of the US price level in 1970 was very near the 1985 level; for what it is worth, 1970 was in a quasi-fixed exchange rate era.

tradables, including invisibles like tourism. In the empirical section of the paper we will look at whether such shifts appear to explain the Äs.

D. Some Illustrative Results

In the previous section we have listed some factors beyond those considered in most empirical attempts to reconcile differences between extrapolations and benchmarks. This section looks at these effects employing two basic statistical frameworks, one examining the simple relationships treating each country independently, and the second, allowing for possible interactions between countries that have similarities in physical location, price structure, growth experience, or relative price changes. The difference in these two formulations is first discussed with respect to growth rates.

Will countries with similar growth rates have similar Äs? This question can be examined looking at simple measures of association like a least squares equation with the growth (1+r*) in income implicit in two benchmark comparisons on the left hand side and on the right the value of the same ratio where the later benchmark is an estimate using national growth rates.[9] Does adding the growth rate as another variable on the right hand side improve the explanatory value of the regression equation? Probably not. What seems more likely is that any pair of countries with high growth rates will have similar Äs. A way to test the latter relationship is to create a weights matrix, W, that consists of the product of the growth rates of each possible pair of countries between the two benchmarks. One can then test whether the interdependence of country growth effects is associated with the residual variation in the basic relationship. This is done using Moran's I measure of correlation. For a W matrix where the correlation is significant, one can then introduce W into an estimating equation to take account of this type of autocorrelation.

Illustrative Data for the OECD 1990-96

In the future it is planned to carry out more extensive tests, but initially we have used the 1990 and 1996 benchmarks at the GDP level involving 23 OECD countries. Because there are so few observations and because we are only considering the differences between two benchmarks, these results at best will be suggestive. The analysis begins with a simple formulation with $G_{BM96/90,}$ the implied growth rate of GDP per capita based on the 1990 and 1996 benchmarks relative to the average of the OECD as a function of $G_{96/90}$, the extrapolated value of the same variable. Various candidate variables that might be considered, like openness, the ratio of exports plus imports to GDP were then added to this equation.

As one would suspect the simple relationship was typically fairly strong, though far from perfect, with unadjusted R^2 between the two relative growth rates being 0.76.[10]

---

[9] The benchmark and extrapolated variables are of the form (1+ r*) and (1 + r) where in both cases the growth is relative to a reference country or country group. In the empirical work a reference group of countries has been used, namely the average of the OECD countries.

[10] The R^2 between the actual benchmark estimate and the extrapolated estimate is much higher, .979, because the two levels will not differ much in 6 years. We believe it is more informative to work with the present formulation because it is more sensitive to the differences in benchmark estimates and does not introduce the apparent high correlation of level measures.

The candidate variables were then considered as possible additions to this relationship, and many could be quickly eliminated, including openness. Also, we considered the growth rate of domestic absorption and it also did not add to explained variance. However, we did try several measures related to changes in price level that were of the following form. For domestic absorption, government plus investment, and investment, the change of the dollar exchange rate was divided by the change in the national accounts deflator. These variables, DA_{PRX}, GI_{PRX}, and I_{PRX}, are defined in Table 2 below the descriptive statistics.

Table 2. Main Variables

N=23	**Mean**	**Std. Dev.**
$\mathbf{G_{BM\ 96/90}}$	1.032	0.079
$\mathbf{G_{96/90}}$	1.033	0.078
DA_{PRX}	0.875	0.098
GI_{PRX}	0.897	0.098
I_{PRX}	0.941	0.109
XM_{PRX}	0.950	0.055

Domestic Absorption:

$DA_{PRX} = (Xrate_{96} / Xrate_{90}) / ((DA\ cur_{96} / DA\ kon_{96}) / (DA\ cur_{90} / DA\ kon_{90}))$

Government + Investment:

$GI_{PRX} = (Xrate_{96} / Xrate_{90}) / ((GI\ cur_{96} / GI\ kon_{96}) / (GI\ cur_{90} / GI\ kon_{90}))$

Investment:

$I_{PRX} = (Xrate_{96} / Xrate_{90}) / ((I\ cur_{96} / I\ kon_{96}) / (I\ cur_{90} / I\ kon_{90}))$

Exports + Imports

$XM_{PRX} = (Xrate_{96} / Xrate_{90}) / ((XM\ cur_{96} / XM\ kon_{96}) / (XM\ cur_{90} / XM\ kon_{90}))$

The motivation for each of these variables has been suggested in the text. In a world where there was little change in price structure over time, and exchange rates adjusted to relative price changes over time, these variables would be 1. The means of all three variables are much below 1; that is on average the change in dollar exchange rates has been much less than in national prices. Even if there were relative price changes it would be expected that ÄERate/ ÄDA was 1. When each of these variables is introduced into the benchmark and extrapolated growth regressions, their coefficients are in each case negative and statistically significant as can be seen in Table 3.

Table 3. OLS Estimates: Implied growth rate, $G_{BM\ 96/90}$, (based on benchmarks 90-96) as a function of national growth rate of GDP per capita, $G_{96/90}$. and other variables

Dep. Var: $\mathbf{G_{BM\ 96/90}}$	**1**	**2**	**3**	**4**
Constant	0.13 (0.24)	0.30 (0.04)	**0.36 (0.02)**	0.32 (0.03)
$\mathbf{G_{96/90}}$	0.87**	0.84**	**0.81****	0.82**
DA_{PRX}	-	-0.16 (-.06)	-	-
GI_{PRX}	-	-	**-0.18 (-.04)**	-
I_{PRX}	-	-	-	-0.15 (0.05)
R^2	*0.76*	*0.79*	***0.79***	*0.79*
Sig-Sq	*0.00150*	*0.00132*	***0.00127***	*0.00130*
Likelihood	*43.17*	*45.21*	***45.66***	*45.38*

N=23, Pr>|t| in ( ), ** Pr<0.0001, * Pr<0.0005.

How do we interpret this result? For countries where the national price levels of DA, GI or I are rising the implied growth between benchmarks is greater than by extrapolation by national growth rates. One interpretation is that such rapid changes in relative prices lead to an overstatement of national growth rates. It is very difficult to explain the difference by predictable changes in benchmark results

We next examined whether there was any residual autocorrelation with the variables that might help explain the Äs. This test involved various W matrixes that looked at all possible pairs of values for openness, growth rates, the similarity index of prices developed by S. Sergueev (2001) and the other variables in Table 2. Whereas the growth rate of a country was not significant in the simple relationship in Table 3, there was significant autocorrelation between the residuals of the Table 3 equations and the matrix of pair-wise growth rates. This relationship is given in Table 4. In interpreting Table 4 the R^2 statistics cannot be compared in any simple fashion with those in Table 3. The test of the weights matrix is the Lambda (W) coefficient, the error measure, $ó^2$, and the likelihood ratio value. On these measures the addition of the W matrix shown in Table 3 improves the explanatory value of the relationship. The W matrix used was the absolute value of the difference between the national growth rates of per capita DA between each pair of countries.

Table 4. Estimates taking into account residual autocorrelation (Spatial Error Model)

Dep. Var: $G_{BM\ 96/90}$	**1**	**2**
Constant	0.124 (0.052)	**0.389****
$G_{96/90}$	0.881**	**0.790****
GI_{PRX}	-	**-0.194***
Lambda (W)*	-0.9577 (0.024)	**-1.377***
R^2	*0.77*	***0.80***
Sig-Sq	*0.0010*	***0.0006***
Likelihood	*45.32*	***49.68***

N=23, Pr>|t| in ( ). ** Pr<0.0001, * Pr<0.0005.
*with $W_{ij} = |(DA_{Grow})_i - (DA_{j\ Grow})_j|^{-1}$

The weights matrix essentially modifies the values of both the dependent and independent variables in the equation[11]. What equation in column (1) of Table 4 says is that countries with similar growth rates, positive or negative, will have Δs of similar magnitude. In column (2), the W matrix is used with the GI_{PRX} variable, improving the relationship compared to that when either effect is considered separately. These results strongly suggest that benchmark reconciliations can be improved by taking into factors

11

$y = Xb + e$ with $e = lWe + h$ as i.i.d

This implies that $e = (I - lW)^{-1}h$ and $y = Xb + (I - lW)^{-1}h$,

or $(I - lW)y = (I - lW)Xb + h$ (both y and X are spatially filtered variables),

equal to $y = lWy + Xb - lWXb + h$. The latter can also be interpreted as

a spatial common factor constraint, with $y = g_1 Wy + Xg_2 - WXg_3$ and $g_1 g_2 = g_3$

independent of and not typically considered in attempts to reconcile benchmarks. This result is obtained considering only two benchmarks and with experimentation with alternative variables. We are not able at this time to go much further in our analysis, but we are at least able to check whether the included variables in Table 2 have similar explanatory value when more benchmarks are considered.

<u>Four OECD 1990s Benchmarks</u>

In this section we report on an exercise using OECD data for the 1990, 1993, 1996, and 1999 benchmarks. One element of discontinuity exists in that the OECD countries implemented the 1993 SNA during this period. The main change affecting the organization of the results is that instead of using household consumption outlays for goods and services, actual household consumption including goods and services provided to households by governments and non-profit organizations are included. These are mainly in the education and health sectors. Since the comparisons presented here are not disaggregated, this change should not affect our findings.

In generating this comparison we have introduced one element that is new in the sense that the aggregation procedure adopted is stochastic, based on the formulation suggested by Rao (2001) that generalizes the CPD approach. In the original CPD formulation of Summers, coefficients on heading parities and coefficients for each item within the basic heading were estimated. An analogous set of coefficients is generated by applying the method to a set of heading parities, making use of the corresponding set of heading expenditure weights. A clear advantage associated with any simultaneous estimation of PPPs and heading coefficients is that it permits a natural disaggregation of the results. The Rao-CPD approach given in (1) below has been explored by Heston and Aten (2002) for a larger group of countries, and the results are generally similar to the EKS method. In equation (1) p_{ij} is the heading parity, â$_j$ is coefficient on each heading *z* and á$_i$ is the coefficient on each of the country dummies except the numeraire. The estimation of (1) minimized the weighted sum of the squared residuals where the weights are the expenditure shares.

$$(1)\ \ln \mathrm{p}_{\mathrm{ij}} = \sum_{i=1}^{n} b_i z_i + \sum_{j=2}^{m} a_j D_j + e_{ij}$$

Estimation of this equation was carried out for 24 OECD countries in each of the 4 benchmarks. Table 5 presents the data underlying both the earlier results provided in Tables 3 and 4, as well as estimates using the modified CPD approach. In each column the per capita income of each country is expressed relative to the OECD average of 24 countries (=100), except for the 1990 and 1996 numbers from Varjonen that are relative to the EU15 and given in columns (1) and (2). Columns (2) – (6) present the benchmark results, and columns (7) – (9) the extrapolations of 1990 to 1993, 1993 to 1996, and 1996 to 1999. Finally the extrapolation of the benchmark in 1990 to 1996 is presented in column (10).

Table 5. GDP per capita in PPPs and Extrapolated by National Accounts growth rates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ISOcode	**Vbm90**	**Vbm96**	**Bm 90**	**Bm 93**	**Bm 96**	**Bm 99**	**Extrap 93**	**Extrap 96**	**Extrap 99**	**Extrap 9690**
AUS	106	113	103.6	103.8	105.3	104.2	108.0	106.8	108.6	111.1
AUT	106	109	106.9	107.5	105.5	105.2	109.6	104.4	102.4	106.4
BEL	108	112	104.7	109.1	106.3	100.4	103.0	106.3	104.1	100.4
CAN	119	117	114.7	109.6	110.2	105.9	112.0	109.6	112.8	111.9
CHE	137	126	132.2	130.6	119.7	119.0	126.0	120.9	113.7	116.7
DNK	109	121	107.9	110.7	114.7	112.8	106.1	112.4	111.4	107.7
ESP	80	78	76.1	76.7	76.2	76.0	74.2	75.5	77.5	73.1
FIN	104	97	103.5	88.5	91.9	96.3	90.1	90.6	96.0	92.3
FRA	109	102	107.5	103.9	96.9	90.7	103.9	99.6	94.3	99.5
GBR	101	100	101.1	96.2	95.4	95.2	98.9	97.1	93.3	99.8
GER	105	107	100.9	104.5	102.5	100.2	101.5	100.1	96.9	97.2
GRC	60	66	57.9	62.0	61.4	62.9	57.1	60.7	61.1	55.9
IRL	78	97	72.1	78.7	89.3	109.1	75.2	90.0	106.9	86.1
ISL	114	115	109.5	103.8	106.5	113.0	103.6	104.2	109.4	104.0
ITA	104	106	101.0	98.2	98.9	97.7	98.7	95.4	94.0	95.9
JPN	117	123	114.1	117.6	116.1	101.0	115.3	113.8	105.7	111.6
LUX	151	171	149.5	164.2	165.3	179.7	174.0	168.0	181.2	178.0
NLD	106	109	103.0	103.0	104.8	107.5	104.4	102.4	105.7	103.8
NOR	112	127	110.7	120.5	123.2	120.6	117.0	126.7	120.0	123.0
NZL	87	88	83.1	83.6	83.5	78.2	85.2	86.9	80.0	88.6
PRT	64	69	59.3	64.0	66.3	67.8	59.3	64.6	66.1	59.8
SWE	113	103	111.0	96.7	96.8	97.1	101.8	96.1	96.0	101.2
TUR	30	30	30.4	30.7	28.0	24.6	34.2	30.1	26.5	33.6
USA	142	146	139.1	136.0	135.3	134.8	141.0	137.5	136.7	142.5
average	102.58	105.50	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
C.V.	25.7%	26.2%	26.3%	26.8%	26.4%	28.1%	28.1%	26.7%	28.2%	28.2%

There are some striking movements in Table 5, particularly the upward move of Ireland, Portugal, Norway, to a lesser extent Denmark, and of course Luxembourg with its unusual financial interactions with the rest of Europe. On the downside, Switzerland, Japan, France, Canada, Turkey, Sweden and to a lesser extent Great Britain all experience declines relative to the average, even going from above to below the OECD average. The net effect of all of this however, is not to move the OECD countries closer together, at least as indicated by the coefficient of variation, which if anything rises slightly in the 1990s. Let us move to our particular interest in these estimates, and that is the relation of benchmarks and extrapolations.

The way the results in Table 5 have been analyzed thus far to directly examine the extrapolations and benchmarks. First, a ratio has been formed between benchmark values for GDP per capita relative to the group average from the later benchmark as a ratio to the earlier year benchmark. That is, the ratio of the 1993 and 1990 benchmarks becomes the left hand side variable. Values greater than 1.0 would indicate that the relative income of a country has improved between the two benchmarks. The next step is to make an estimate of the same ratio, where the numerator is the extrapolation from 1990 to 1993 and the benchmark value for 1990 is again the denominator. Other combinations of years are possible but we have concentrated on the successive benchmarks, and on 1990 and 1996 because the latter was used in Tables 2-4.[12]

Full analysis of these 4 benchmarks has not been possible[13] but there are several preliminary points to be made. First, consider the following simple descriptive statistics, the standard deviation of the benchmark changes, and the square root of the residual sum of squares from a regression of this change on the extrapolated estimate:

	1993/90	1996/1993	1996/90	1999/96
Standard Deviation	.062	.044	.084	.070
Root Mean Square Error	.045	.028	.061	.030

If the benchmarks are the truth, then the variation of relative positions of countries varies most between 1999 and 1996 and least between 1996 and 1993. Consider now the extent that these variations can be explained by an extrapolation as indicated by the residual sum of squares from a simple regression. What is unexplained by extrapolations is usually more, the more there is to explain, with the notable exception of 1999/96. And it is worth noting that the relative performance of extrapolations is worse over longer periods in the one example given, the1996/90 case.

Another point relates to the variables we have earlier touted. How well do these variables do in explaining these additional benchmarks? The answer is they do well, but not consistently well. Three measures of price movement of the components of GDP, namely DA, I, and I + G, had been expressed as ratios to exchange rate changes. At least one of these three variables was significant when included with the extrapolated value ratios in each of the four benchmark intervals considered. While somewhat encouraging, these results suggest the importance of seeking a better specification of the relationships involved.

Conclusions

This paper has explored the question of whether the differences between two benchmarks and the extrapolated values can provide insights about larger economic issues. We have tried to examine these differences using recent OECD benchmark experience employing aggregative variables not directly involved in the benchmark calculations. That is instead of assuming that inconsistencies between the elements entering benchmarks and national accounts deflators/growth rates can reconcile the results, we have asked whether there may be systematic factors at work on a more aggregative level. Two results emerge from the empirical analysis. First, countries experiencing sharp changes in their national price level relative to a group average tend to have larger differences between extrapolated and benchmark estimates. Further, common country experiences, like similar country income growth, are likely to

[12] In Tables 2-4 we did not include unified Germany because other variables were not available at the time. So there are some non-comparabilities between Tables 2-4 and Table 5 for the 1990-96 benchmarks. Also, the earlier exercise was based on a different aggregation method and probably slightly different 1996-90 growth rates (we have used the most recent OECD national accounts information) so that may be introduce slight differences.

[13] Please see Appendix A for preliminary results and tables.

experience similar differences between extrapolated values from earlier benchmark and a later benchmark.

These results seem strong enough to warrant further research on the question, delving deeper into OECD experience and extending the exercise to a larger sample of countries. And clearly the patterns documented here require a more interpretative explanation than we have been able to provide thus far. The seminal paper of Obstfeld and Rogoff (2000), suggests the importance of border effects that may be an important element in understanding why exchange rate changes occurring independently of current domestic price movements, may themselves not be passed through to domestic prices. Why such situations persist and exactly how they impinge on domestic price structures is not well understood. However, because they are a puzzle to those working in international trade they should also enter into the framework of those concerned with estimating purchasing power parities. First, a major stylized fact with which those in international trade are concerned is the major divergence from the law of one price that is a fundamental finding of the ICP. Second, the explanations sought in international trade may contribute to our understanding of why inconsistencies between successive benchmark comparisons may be an inconvenient fact of life, but are not a fact of life that will necessarily go away even as the quality of benchmark estimates are improved.

References

Goldberg, Linda S. and Keith Crockett (1998), "The Dollar and U.S. Manufacturing," Current Issues in Economics and Finance, Federal Reserve Bank of New York, November.

Heston, A., and R. Summers (1993)"What Can We Learn from Successive ICP Benchmark Estimates?," in A. Szirmai and D. Pilat ed., Explaining Economic Growth; Essays in Honour of Angus Maddison, North Holland, Amsterdam.

Heston, A., with R. Summers and D. Nuxoll "Comparative Country Performance at Own-Prices or Common International Prices," 1994) in Economics, Econometrics and the Link Essays in Honor of Lawrence R Klein, ed, Jan Dutta, et al. Elsevier, New York.

Kravis, Irving and Robert Lipsey (1987), "The Assessment of National Price Levels", in S. W. Arndt and J. D. Richardson eds., Real Financing Linkages Among Open Economies, Cambridge, Mass, MIT Press.

Obstfeld Maurice and Kenneth Rogoff (2000), The Six Major Puzzles in International Macroeconomics: Is There a Common Cause? , NBER Working Paper No.w7777

Rao, D. S. Prasada (2001), "Weighted EKS and Generalized CPD Methods for Aggregation at Basic Heading Level and Above Basic Heading Level," Joint World Bank-OECD Seminar on Purchasing Power Applications, Washington D.C., www.oecd.org/std/ppp/MTG2001/index.htm, January 30-February 2nd.

Scarpatta, Stefano, Andrea Bassanini, Dirk Pilat and Paul Schreyer,(2000) "Economic Growth in the OECD Area: Recent Trends at the Aggregate and Sectoral Level," Economic Department Working Papers No. 248 OECD, 26-June.

Sergueev, Sergei (2001), "Measures of the Similarity of the Country's Price Structures and their Practical Application," Statistical Commission and Economic Commission for Europe, Geneva 12-14 November.

Summers, R. and A. Heston (1984) "Improved International Comparisons of Real Product and its Composition, 1950-80," Review of Income and Wealth, 207-61.

Varjonen, Seppo (2001), "Consistency Between GDP Based on PPPs and National Accounts Time Series," OECD Meeting on National Accounts Oct 9-12, 2001. STD/NA (2001)34.

APPENDIX A

Results for All Benchmark Years

The results for the 1996/90 Varjonen (2001) data are in the text. Below are the results for 1993/90, 1996/93 and 1999/96 using PPP estimates from the Rao-CPD approach. The dependent variable is again the ratio of the benchmark GDP per capita in PPPs, relative to the average of the 24 countries, while the independent variables are the national accounts growth rates, plus variables reflecting exchange rate changes relative to national price changes. The estimates shown are for the best spatial error model (based on the highest maximum likelihood value) and its corresponding OLS estimate.

Table A1. OLS and Best Spatial Error Models for Each Year

Dep. Var: G_{BM}	**1**	**2**	**3**	**4**	**5**	**6**
	1993/90	**1993/90**	**1996/93**	**1996/93**	**1999/96**	**1999/96**
Constant	0.49 (3.1)*	059 (6.5)**	0.22 (1.7)*	0.29 (3.0)*	-0.11 (-.86)	-0.15 (-1.9)
$G_{National}$	0.63 (4.7)**	0.57 (7.1)**	0.59 (3.8)**	0.48 (3.8)**	0.69 (3.2)*	0.69 (4.2)**
I_{PRX}	-0.13(-2.7)	-0.17 (-4.7)**	-	-	-	-
DA_{Grow}			-	-	0.30 (1.4)	0.34 (2.1)
I_{Grow}			0.12 (1.9)	0.15 (2.9)*	-	-
Lambda (W)*	-	-0.70 (-2.8)*	-	-0.93 (-3.8)**	-	-0.83 (-4.9)**
R^2	*0.63*	*0.66*	*0.74*	*0.71*	*0.79*	*0.86*
Sig-Sq	*0.0016*	*0.0010*	*0.0006*	*0.0003*	*0.0011*	*0.0006*
Likelihood	*45.0*	*47.0*	*57.6*	*60.5*	*49.5*	*54.0*

N=24, z-values in ( ). ** Pr |t|<0.001, * Pr |t|<0.005.
*with W_{ij} in 1993/90 $=| (GI_{Grow})_i - (GI_{j\,Grow})_j |^{-1}$,
W_{ij} in 1996/93 $=| (DA_{PRX})_i - (DA_{PRX})_j |^{-1}$
W_{ij} in 1999/96 $=| (GI_{PRX})_i - (GI_j\,PRX)_j |^{-1}$

The estimates for the spatial error model (columns 2,4 and 6) are consistently better than the simple OLS formulation in columns 1,3 and 5, confirming the fact that the residuals are autocorrelated. The spatial error model assumes that the residuals of the regressions in each time period are not independent. The form of this dependence is specified in footnote 11 and repeated below:

$y = Xb + e$ with $e = lWe + h$ as i.i.d

This implies that $e = (I - lW)^{-1}h$ and $y = Xb + (I - lW)^{-1}h$,

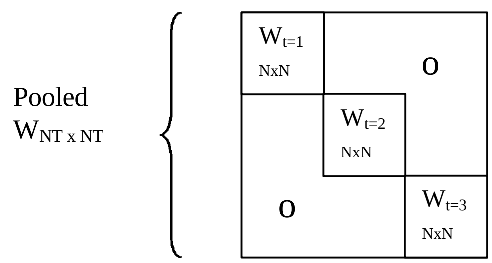
or $(I - lW)y = (I - lW)Xb + h$ (both y and X are spatiall y filtered variables),

eq ual to $y = lWy + Xb - lWXb + h$. The latter can also be interpreted as

a spatial common factor constraint, with $y = g_1 Wy + Xg_2 - WXg_3$ and $g_1 g_2 = g_3$

The lambdas are negative and very significant, ranging from –0.93 in 1996/93 for the weight matrix of differences in price changes for DA, to –0.70 for 1993/90 for matrix of difference of growth rate for G+I. Note that the best model from the 1996/90 Varjonen data was for a different matrix - growth rate differences for DA.

Although these models account for a large variation in benchmark changes, is there a similar pattern over time? One way to look at this is with a simultaneous space-time model, where the lagged values are either in time or in space (as modified by a W matrix). Although the spatial W most commonly refers to geographical space, it can be any binary relationship among all observations, with care taken to create large values for 'close' relationships, and small values for more distant relationships. For example, we take the inverse of the absolute differences in the Ws because we want countries whose growth rates are similar to be 'important', i.e., to have a greater weight relative to each other than countries that have dissimilar growth rates. If we pool all observations, stacking the dependent and independent variables by time periods, we have 72 observations (24x3). The Pooled W matrix consists of three block diagonal matrices expressing the interaction between observations in each time period. In graphic form:



Here, Pooled W has elements equal to zero outside the block diagonal, i.e., the $W_{t's}$ capture the interaction among current neighbors, but not neighbors in a different time period *t*. The estimated model is given below and the results are in Table A2.

$$y_{it} = X_{it}b + e_{it} \text{ where } e_{it} = l\,[We]_{it} + h_{it}$$

$$[1 - lW]y_{it} = [1 - lW]X_{it}b + h_{it} \text{ or}$$

$$y_{it} = l\,[Wy]_{it} + X_{it}b + l\,[WX]_{it}b + h_{it}$$

Table A2. OLS and Best Pooled Spatial Error Model

Dep. Var: G_{BM}	**1**	**2**
Pooled	**OLS**	**Sp. Error**
Constant	0.47	0.28
	(5.0)**	(3.1)*
$\mathbf{G_{National}}$	0.56	0.74
	(7.3)**	(10.1)**
I_{PRX}	-0.08	-0.09
	(-2.7)	(-3.0)*
Lambda (W)*	-	0.65
		(6.4)**
R^2	*0.52*	*0.86*
Sig-Sq	*0.0017*	*0.0010*
Likelihood	*128.7*	*141.4*

N=72, z-values in ( ). ** Pr |t|<0.001, * Pr |t|<0.005.
*with $W_{ij} = |(I_{Grow})_i - (I_{j\,Grow})_j|^{-1}$,

The corresponding results for three separate regressions with the same variables as in the pooled model are shown in Table A3.

	1993/90	**1996/93**	**1999/96**
Constant	0.47	-0.08	0.18
	(3.2)*	(-0.50)	(1.3)
$G_{National}$	0.64	0.89	0.85
	(5.0)**	(8.5)**	(8.2)**
I_{PRX}	-0.13	0.12	-0.11
	(-2.8)*	(1.4)	(-2.5)*
Lambda (W)*	0.098	-0.53	0.44
	(0.3)	(-1.6)	(2.0)
R^2	0.63	0.74	0.86
Sig-Sq	0.0014	0.0005	0.0008
Likelihood	45.0	57.1	51.6

N=24, z-values in ( ). ** Pr |t|<0.001, * Pr |t|<0.005.
*with $W_{ij} = |(I_{Grow})_i - (I_{j\,Grow})_j|^{-1}$,

Note that only the 1999/96 Lambda coefficient is positive, and that in all three cases, these particular variables are not the best, with lower likelihood values than those in Table A1. When pooled, it seems that the 1999/96 relationship dominates that of the other years, that is, the investment price change is negative and significant, and the differences in investment growth between pairs of countries is positive and stronger than other spatial 'filters'.

It is not clear whether separate regressions or a pooled version is 'best', but it is encouraging that we can capture this amount of variation. A weight matrix with time dependencies explicitly included could also improve the model, but it may lead to difficulty in interpreting the lambda. As it is, the lambda is a fairly straightforward measure of interaction across countries, as measured by aggregate indicators of price change or growth.