July 12, 2009

Is Newer Better?

The Penn World Table Revisions and the Cross-Country Growth Literature

Simon Johnson, William Larson, Chris Papageorgiou, and Arvind Subramanian[1]

Abstract

Versions 6.1 and 6.2 of Penn World Tables (PWT) have essentially the same methodology but report significantly different GDP—growth and level—numbers. The revisions in these numbers in PWT6.2 are in part due to updated national income accounts data, but are in part inherent in the PWT methodology, making it difficult to determine which of the existing versions is "best." The methodology renders data for countries with relatively small total GDP and data distant from the current benchmark year especially variable. We examine 13 leading studies of growth to check for robustness across different versions of PWT. Table invariant results are common in studies examining cross-sectional or very long run data; in contrast, results based on higher frequency data are less likely to be robust in this sense, and annual data are particularly problematic. We also find, surprisingly, that the PWT methodology also leads to GDP estimates that are not truly at PPP prices except for the single year for which international prices are collected. We propose an alternative way to use PWT data that might be less prone to these problems of data variability and inadequately reflecting PPP prices.

[1] Johnson: MIT Sloan School of Management, Peterson Institute for International Economics, and NBER, sjohnson@mit.edu. Larson: George Washington University, larsonwd@gmail.com. Papageorgiou: IMF, cpapageorgiou@imf.org. Subramanian: Peterson Institute for International Economics, Center for Global Development, and Johns Hopkins University, asubramanian@piie.com. We thank Michael Clemens, Robert Feenstra, Aart Kraay, Mick Silver, Antonio Spilimbergo and especially Angus Deaton, Alan Heston, and David Romer for comments, encouragement and suggestions, and seminar participants at Brown University, the Bureau of Economic Analysis, IMF, Oxford University, the University of Pennsylvania (2008 Penn World Table workshop), Princeton University, the Econometric Society Summer Meeting (Boston 2009), the NBER Summer Institute (Small Growth Group, 2009), and the Society for Economic Dynamics Meeting (Istanbul 2009). The views expressed here are those of the authors only.

1. Introduction

The Penn World Table (PWT) is the most widely used source of data for cross-country income comparisons.[2] In particular, version 5.0 of these Tables, which was published in 1991, represented a major step forward in data availability and helped spawn a large empirical literature on growth and its determinants. Without doubt, the economics profession has made substantial progress by paying close attention to the pioneering work of Irving Kravis, Alan Heston, and Robert Summers.

But economists may not have been sufficiently attentive to the caveats and warnings that were clearly attached to both the first version of the Penn World Table and all subsequent versions. In particular, the Table clearly and unequivocally warned that data quality was weak for many countries. But it turns out there is more to guard against than just data quality. This paper draws attention to three consequences for the growth data which is inherent to PWT methodology. First, we document that GDP data are highly variable. Revisions to GDP growth rates are driven to a great extent by changes made by the PWT to the underlying national income accounts (NIA) data. Revisions to GDP levels and the purchasing power parities, in addition to being affected by changes to underlying NIA data, are affected systematically by certain inherent features of the PWT methodology. This variability is especially pronounced: at <u>high frequency</u> (annual data are much more variable than longer averages); for <u>small countries</u> (i.e. countries with small total GDP); and for historical data, so that the <u>further the data are from the benchmark year</u> in the PWT, the more variable they are. To the extent that variability is

[2] Roughly 2/3 of all cross-country empirical work is based on PWT. Second place is held by the World Bank's World Development Indicators, which were originally based on the Penn World Tables, but which have subsequently diverged. The IMF's World Economic Outlook dataset places a distant third.

inherent to the methodology, it is difficult to know whether newer data are better for estimates of GDP levels and growth.

Second, the variability of growth data has implications for the cross-country growth literature. Results based on annual data prove to be less robust across versions of the Tables than are results based on 10-year averages. And results based on long run growth rates or levels turn out to be quite robust in our "Table Invariant" (i.e. robust across PWT versions). The Table's health warnings need to be taken seriously and probably expanded upon, and this implies further robustness checks for all growth studies. If this is done, the Penn World Table would become even more valuable.

Third, the PWT methodology raises a more basic question. The rationale for the PWT is to come up with GDP level and growth data that are at common international (the so-called purchasing power parity) prices so that the data are comparable across countries. The methodology, however, leads to the construction of GDP estimates that are not at common international prices. It turns out that these estimates are based on a mixture of PPP and national prices, making it difficult to really compare across countries.

Is there a way to address problems of variability of data not being at truly PPP prices? We propose an alternative way of using the PWT data that goes some way toward addressing these problems.

This paper is organized as follows. Section 2 briefly reviews the background and history. Section 3 depicts the variability of growth data and the underlying patterns to this variability, focusing on versions 6.2 and 6.1 of the Table. In Section 4 we summarize the results of our robustness/replication studies for leading growth studies. Section 5 illustrates the key aspects of the methodology used in the PWT that go toward

construction of the GDP data and help understand the reasons for data revisions. In Section 6, we explain why not all PWT data fully reflect PPP prices. Section 7 describes our proposal that addresses the problems with PWT methodology relating to data variability and to data not being at PPP prices. Section 8 draws some conclusions, including offering some practical suggestions for researchers. Technical Appendix 1 contains methodological details on the construction of the PWT data. Technical Appendix 2 explains the theoretical reasons for the pattern in data revisions. Technical Appendix 3 illustrates why the current PWT methodology leads to GDP level and growth estimates that are not at international prices.

2. The Evolution of Penn World Table

Background

The Penn World Tables have reported on seven rounds of data, starting in 1970 (Table 1).[3] The latest published version is 6.2, which was released in October 2006.[4] The prior version was 6.1 and that was preceded by 5.6. These are the most commonly used versions of the Table today, and it is on these that we focus our attention. However, it is our understanding that the points made below also apply to other versions.

Conceptually, the approach adopted by the Table is straightforward. The innovation and great contribution of the Table was of course to convert national measures of income into internationally comparable PPP estimates. This is done – in principle – by collecting prices for the same or similar goods in different countries, and deriving price

[3] Appendix Figure 1 shows the distribution of countries by the number of benchmark studies in which they participated.
[4] Version 7.0 will be available in 2009, but the precise release date is not yet known.

indices that can be used to compare what people can actually buy. The price collection operation is known as the International Comparison Programme/Project (ICP).[5] The Table obtains local currency data from the national income accounts of countries. Then, based on international price comparisons, it converts these local currency data into Purchasing Power Parity (PPP)-based figures, which should be comparable across countries. These are more appealing as the basis for comparisons than data based on market exchange rates, which can fluctuate considerably and hide the real progress (or lack thereof) in economic prosperity for a country's inhabitants.

Each "generation" of the Table is based on a different round of the ICP, i.e., so versions 5.6 and 6.1 use different prices. Each generation also uses updated estimates of GDP and its components measured in domestic currency. Within a generation of the Table, the ICP remains the same, e.g., this is the case for 6.2 and 6.1, but there are other revisions – the nature of which varies.

Methodology Changes Over Time

The ICP is a massive undertaking, requiring a vast amount of resources – mostly in terms of people's time, but the computer resources, at least until recently, were also significant. The only organizations that can sustain such an investment are government-funded, and even international organizations, such as the UN, have a hard time coming

[5] The International Comparison Project began in 1968, although its antecedents date back to the 1950s. Irving Kravis was the first director; http://unstats.un.org/unsd/methods/icp/ipco_htm.htm. In 1989 the "P" became Programme, rather than Project. International price comparisons have been completed for 1970, 1975, 1980, 1985, 1990, 1996 and 2005.

up with all the money required.[6] A substantial amount of the needed investment has come directly from governments.

As a result, control over the ICP has shifted over time. "After phase III [1975], the role of the University of Pennsylvania, which had until then been the main engine of the ICP, was gradually transformed into that of adviser on methodological issues. Another notable change in ICP responsibility was the increasing role of the Statistical Office of the European Communities (EUROSTAT). EUROSTAT, in fact, became not only the organizer of the European Community comparison, but also, with its experienced staff, it has provided substantial technical assistance to a number of regional comparisons and to the work on establishing links among the various regions."[7]

And as control has shifted, so have preferences regarding methodology both for particular regions and for how these are aggregated to global estimates. The most important change may have been regionalization: "In phase IV [1980] and onward, countries participated through regions or country groups; first regional (e.g., African, OECD etc.) comparisons were carried out and then the world comparison was built up by linking across these groups."[8] Deaton and Heston (2008) focus on problematic

---

[6] The Ford Foundation did provide critical early financial support through grants to the University of Pennsylvania.

[7] Quotation is from http://unstats.un.org/unsd/methods/icp/ipco_htm.htm

[8] The pattern of support for the Tables has shifted over time. "For the most part, participating countries have provided domestic resources for the data collection for ICP, while the Statistical Division of the Department of Economic and Social Development of the United Nations Secretariat, EUROSTAT, OECD and the Austrian Central Statistical Office, in its capacity as organizer of the Eastern European comparison, have units concerned with ICP-type work. Early sponsors of the ICP, such as the Ford Foundation and a consortium of contributing countries organized by the World Bank, provided resources that allowed coordination of the benchmark comparisons, particularly among developing countries. In phases IV and V, EUROSTAT became the financial supporter of the African and Caribbean comparisons. In phase IV, the Inter-American Development Bank (IADB) provided a major contribution to the Latin American regional comparison; however, this effort was organized through a group of experts visiting

(continued)

methodological points, all of which arise because particular organizations control part of the methods, and their concern suggests that the methodology changes significantly such that it is harder to say if the Table are improving or not in their accuracy. But there are other reasons to worry about this exact same issue.

3. Data Variability and Patterns

We chose to focus on the variability between PWT version 6.1 and 6.2. These versions differ in four relatively small ways.[9] Both versions are within the same generation of the Table, and therefore have almost identical methodologies, and more importantly, between these versions, there is no new ICP and hence no new international price data.[10] Despite these small changes, estimates of GDP levels and growth vary substantially. The one substantive difference is that PWT 6.2 uses more "updated" national income accounts data. Specifically, it is not that the current price national income accounts data change across versions (indeed, the current price data are virtually identical across versions), rather the PWT re-bases the constant price national income

---

countries and little experience was gained by country statistical offices, so that when IADB support was unavailable in phase V, no Latin American regional comparisons took place. The ESCAP regional comparison in phase V was assisted by the Government of Japan, the United Nations Development Programme (UNDP), the World Bank, and the Asian Development Bank." http://unstats.un.org/unsd/methods/icp/ipc6_htm.htm

[9] There are new purchasing power parities for the OECD countries, 20 additional countries (mainly transition economies), coverage was extended from 2000 to 2004, and the reference year for calculation of the purchasing power parities has been moved from 1996 to 2000.

[10] In addition, an important point emerges from a careful examination of the methodological history of the Tables (e.g., Deaton and Heston, 2008). We usually think that a revised data series is better than the original series, due to various kinds of corrections. But in the case of the Tables, the methodology has not necessarily improved over time. There have been innovations, but some of these happened for bureaucratic or even political reasons, as international organizations became involved in the data collection, preparation, and presentation. As a result, it is not reasonable to assume that one version of the Tables (e.g., the latest) is necessarily better than other versions. This is another reason why results that hold in one version of the Tables, but not other versions, should probably be viewed with greater skepticism.

accounts data which yields different level and growth estimates in national currency (we discuss this in greater detail below).

Figure 1 illustrates that the 30-year average annual growth rates (1970-1999; the available period for comparisons; see Table 15 for countries in sample) differ between Penn World Table 6.2 and 6.1.[11] The average difference in the growth rate (for the period 1970-1999) generated by the two versions is close to zero (0.1 percent). But the standard deviation of the differences in growth is about 1.1 percent. This is quite substantial when compared with the fact that the average growth rate (in 6.2) is 1.56 percent. Put differently, growth is more than 1 percent per annum different in more than half the countries and more than 2 percent different nearly a quarter of all cases. (To anticipate our finding on proximate causes below, Figure 2 illustrates that there is also a great deal of variation across versions in the calculated change in prices, measured in PPP terms.)

The first systematic pattern to variability—namely that it increases when the data are at higher frequency—becomes evident when we compute the growth rate over one, ten, and thirty-year periods. These are presented in the three panels of Figure 3 (where the third panel is the same as Figure 1) with the same scale for the growth rates. It is dramatically illustrated that growth computed on an annual horizon is considerably more variable across versions than growth computed over 10 and 30-year horizons. For example, the standard deviation of the growth rates across the two versions is 5.39 percent (relative to the average growth rate of 1.86 percent over that horizon) compared with a standard deviation of 1.08 percent (relative to average growth of 1.51 percent) for

[11] Figure 1 shows that for the sample for which there are data on real per capita PPP GDP in both PWT6.1 and PWT 6.2, the magnitudes are broadly similar when we compare PWT6.2 with WDI or with PWT5.6 (see Table 6) or with the WEO numbers (although the sample varies depending on the comparison).

the 30-year horizon. The striking differences in this figure intuitively explain the results we find later about the robustness of leading growth studies. Results based on annual data prove to be less robust across versions of the Table than are results based on 10-year or 30-year averages. Evidently, errors get averaged out over longer horizons.

Figure 4 shows some prominent cases of revisions in the GDP numbers across the three latest PWT versions and this affects many of the narratives about growth. Consider one such narrative. As an illustration of the severity of the data problem and associated policy discussion for some low income countries, consider the very basic question of who has done well and who has done badly, say over the past 25 years.[12]

Table 2 shows the answer, looking at annual average growth rates between 1975 and 1999 just for African countries. Versions 6.1 and 6.2 agree that Botswana has done best, and they also agree that Egypt, Cape Verde, Tunisia, and Lesotho are in the top 10. But there are seven countries that appear on one list and not the other. While according to version 6.2 Equatorial Guinea is the second best performer with 4% annual growth rate, it is the worst performer with -2.7% annual growths according to version 6.1! The disagreement is even more severe for the "worst performers." Now the lists are almost completely disjoint. There are a total of 10 countries that appear to have the slowest growth according to one version but not according to the other version.[13]

---

[12] This basic and important question was asked, for example, by the World Bank's recent Spence Commission. We reexamine their conclusions in a separate paper, currently under preparation.

[13] More generally, do the leading alternative versions of growth data of the Tables agree on the fastest growing countries around the world? For this we can compare "top 10" and "bottom 10" in terms of growth performance according to versions 6.1 and 6.2 of the Tables, alongside growth rates calculated from the IMF's World Economic Outlook (WEO) and the World Bank's World Development Indicators (WDI). This is the subject of another paper, so we only mention the highlights here. All four sources agree on seven of the fastest growing countries: South Korea, China, Botswana, Thailand, Hong Kong, Ireland and Malaysia. But there are also seven countries

(continued)

A second pattern in the variability of growth estimates is one that the PWT has always highlighted, namely data quality. The Table have always been quite transparent about this data quality issue, and in fact they prominently assigned "Quality Grades" for each country (e.g., to version 5.0) – these grades are subjective assessments made by the authors of the PWT, based on a number of factors described in the Technical Appendix to PWT 6.1, pp.13-18. Figure 5 shows the breakdown of countries by their Grades in version 6.1. Strikingly, only 32 countries received a grade of A or B. Grades of C or D were received by 147 countries.

At first blush, these letter grades turn out to be meaningful, in the sense that they point in the right direction with regard to what data do or do not change across versions of the Table. For example, the left-hand panel in Figure 6 shows differences in 30-year annual average growth rates (1970-1999) for countries with Data Quality Grades of A or B. The right hand panel shows the same for countries with Grades of C or D. All the major variation across versions of the Table occurs in the countries with lower grades. The same comparison also holds across higher frequency growth data.

The quality grades are obviously informative, in the sense of "explaining" the data variability across revisions. The PWT's quality grading is part objective and part subjective and is based on other criteria including the size of the country and the number of benchmark studies that it has participated in. But is quality per se the cause or is quality suggestive of something else?

---

that appear on at least one list but not on all lists. And even between the IMF and the World Bank data there is disagreement on the exact ranking (and on whether Indonesia, Portugal and Luxembourg make the cut). Again, there is more disagreement on the countries that have done worst – probably because poorer performing countries tend to have less reliable data. Only five countries appear on all four lists: Madagascar, Nicaragua, Togo, Venezuela, and Zambia.

Next, we examine the underlying causes of the revisions, by using simple OLS regression analysis. It is worth noting that our aim is to provide some more structure to the correlations presented in the figures above, rather than establishing a tight causal relationship. In this spirit, we run regressions for three different and important variables estimated by the PWT—the purchasing power parity prices, the level of PPP GDP, and the growth rate of PPP GDP.

For reasons explained in greater detail below, we posit that there are four possible factors that could cause revisions to the PWT estimates. First, PWT estimates could change because the underlying NIA data—part of the inputs for the PWT system—change. Second, even apart from changes in the underlying data, revisions could be systematically related to the quality of the data. Third, the PWT grading is in part related to the size of the country, so we explore whether size could play a role in data revisions. Finally, revisions could be systematically related to time and history. We should note that it is not the case that revisions to the Table change only more recent data. Appendix Figure 2 shows that there are large revisions for the 1970s, the 1980s, and the 1990s, looking at 10-year average annual growth rates. The variable we use to capture this is the distance of the estimate from the benchmark year for the ICP.

Regressions for revisions to the PPPs, level of GDP and GDP growth are reported in Tables 3-5 respectively. In Table 3, for example, the dependent variable is the absolute value of the difference in log of international price between version 6.2 and version 6.1.[14] We use two alternative measures for data quality: the first is the grading by PWT itself

[14] All the regressions are done at annual frequency but we have also replicated them at a decadal and 25-year frequencies and obtained similar results with regards to coefficient signs and significance levels. Consistent with Figure 3, when using higher frequency data the coefficient magnitudes obtain considerably larger values for all regressors.

and we convert the 4 letter grades for data quality into an index, from 1 to 4, where 4 is the best ("A"). The second measure is simply the number of ICP studies in which a country has participated. For size, we use the log of a country's gross domestic product. Since the benchmark year for PWT 6.1 is 1996, our measure of time is the number of decades from 1996 (with each year being $1/10^{th}$ of a decade; the observation for 1995 will therefore be 0.1). We use the log of the absolute value of differences in the growth of GDP (at constant domestic prices) as well as the log of the absolute value of differences in the level of GDP (at constant prices) as our measures of changes in national income accounts data.

Column 1 shows that a higher grade is negatively related to differences in the data, which confirms that the PWT team knows their data well. Column 2 shows that if there have been more ICP benchmark studies for a country, this also reduces the extent of data revisions between versions 6.2 and 6.1 – again, no surprise. Log GDP (column 3) has a similar effect, although of course this may be a proxy for data quality or the existence of a benchmark study. These results are consistent with the quality grading as advocated for years by the PWT researchers.

Perhaps more novel is the result in column 6. This shows that the further a datapoint from the benchmark year, measured in absolute value of years from 1996, the more likely is its data to change between versions 6.1 and 6.2, i.e., the coefficient is positively signed. These results on the variability of the PPPs and their relationship to country size (log total GDP) and distance are graphically represented in Figures 2 and 7, respectively. Figure 7 is striking in depicting a funnel relationship between variability and distance from benchmark year, particularly for countries with Grades C and D. Data

variability increases as the datapoint moves away from 1996 both forward and backward in time. A key implication is that historical data in the PWT need to be viewed with particular caution.

Table 4, which reports results for the level of GDP per capita, suggests that the determinants of the revisions for GDP are similar to those for international prices (see figures 7 and 8).[15]

In Table 5, however, we find a startling result. Each of the determinants is significant on its own; but when they are all introduced together (column 7), the variable relating to changes in national income accounts data trumps all other variables. In other words, the revisions to PWT GDP growth rates are almost exclusively driven by the updating of the NIA data. This is graphically illustrated when we plot the two revisions (Figure 9). The simple correlation is 0.94.

We explore this result a little further. We find that there are very few revisions to the current price NIA data. Indeed, when we add this variable to the growth regression in Table 5, it has little explanatory power on its own and it does not affect the power of other variables.

[15] Appendix Tables 1 and 2 and Appendix Figures 3 and 4 confirm the same broad pattern holds across investment and consumption data also. The reason we show this relationship for investment and consumption (specifically the shares of investment and consumption in GDP measured at their respective PPP prices) is that these shares are crucial in the computation of growth rates in the PWT (see the discussion in Technical Appendix 1 under "Step 3."). Distance from benchmark year has a consistently positive effect on the size of the data revision, although it is a bit less marked for consumption than investment, and this effect is larger for countries with lower quality data.

4. Replication Exercises for the Cross-Country Growth Literature

The size of these revisions to data between versions of the Table is surprising and the pattern to the revisions is interesting, but do they matter? To assess this we turn now to examine some of the most prominent studies in the growth that have used the Table.

Most papers in the empirical growth literature use as dependent variable either the level of real per capita GDP, in PPP terms, or the growth rate in the same. Some papers use the level of income or growth rates as a right-hand side, control variable. Given the considerable variation in the data on growth rates (and levels) and the fact that there is no one "best" set of Penn World Table to use, a natural question is which results in the literature are robust to changing versions of the Table.

Criteria for considering papers

We have examined many of the leading papers in the growth literature based on Penn World Table 5.6 or 6.1.[16] In each case, we attempted to run exactly the same specifications and samples, but using version 6.2 of the Table instead. This approach cannot prove that a particular set of results is right or wrong, but it may illustrate patterns in terms of what kind of results are more or less robust.

We had four criteria for inclusion. First, the list had to include papers that studied the level of GDP, the growth rate of GDP, and the volatility of GDP. Second, since data revisions can be thought of as measurement error, and since estimation depends upon

[16] We focused our attention on "high impact" papers, measured either in terms of citations (using Google Scholar) or based on discussion with active researchers or papers that we think will prove influential. We sought to examine a range of papers, in terms of the frequency of data and methodology. However, our sample is not comprehensive, and the results are intended as illustrations only.

whether measurement error is on the left or right hand side, we included papers in which the various GDP measures (level, growth, and volatility) featured as regressors and as regressands. Third, since our preliminary analysis of the data showed that measurement error varies significantly between high and low frequency data, we chose papers that adopted a pure cross-sectional approach (very low frequency); a panel approach, with data averaged over 4-5 years (high frequency); a panel approach with annual data (very high frequency). Table 13 lists the various papers in the appropriate analytical category. After having decided the universe of potential papers based on these criteria, we narrowed the list to papers that we considered influential. Finally, for inclusion here, we had to be able to obtain the original data and to be able to replicate the paper's main results with those data.

Papers with table invariant results

In all, we tested the robustness of 13 papers in the growth literature. Note that we did not check all specifications in all papers. Rather we concentrated on what appear to us – or to others citing the work – as the "main" results. The lower part of Table 13 lists 9 papers for which we found basically no or small changes in results. In addition, there were more substantial changes for 4 papers: Ramey and Ramey (1995), Jones and Olken (2005), Hausmann et al. (2005), and Aghion et al. (2005). We go through these in more detail below.

Before that, we should emphasize that many prominent papers survived the robustness checks. Whatever else is right or wrong with the growth literature, the bulk of it is not afflicted by the problem of sensitivity to changes in GDP data. This list includes

work such as Acemoglu, Johnson, Robinson and Thaicharoen (2003), Barro (1999), Burnside and Dollar (2000), Delong and Summers (1991), Demirguc-Kunt, Levine and Laeven (2004), Easterly, Kremer, Pritchett and Summers (1993), and Sachs and Warner (1995). The minor changes that we found in results from these papers are mentioned briefly in Table 13. Miguel, Satyanath, and Sergenti (2004) proved robust in a particularly interesting way, which we expand on below. And we also discuss Mankiw, Romer and Weil (1992) in more detail below, because their results come close to being affected.

Ramey and Ramey (1995)

This paper, published in the *American Economic Review*, 1995, tests the link between growth and volatility. The specification involves running an annual panel with growth rate on the left hand side, a set of country- and time-varying controls on the right hand side, and a time fixed effect. There is no country fixed effect; instead there is a country-specific and time-invariant measure of volatility of growth, proxied by the standard deviation of the growth rate over the period covered by the growth data.

The paper used growth data from the PWT 5.6. In the original paper, there are 2208 observations from 92 countries, while in the balanced sample (i.e., where there are data from both PWT 5.6 and 6.2), there are 1776 observations from 74 countries.

In the paper, the coefficient on volatility is negative and significant at the 1 percent level. The magnitude of the coefficient is -0.177. We replicate this result in the first column of Table 7.

When we re-estimated their core specification on the balanced sample (column 2), the coefficient on the volatility term is smaller and barely significant. When we switch to using PWT 6.2, the volatility coefficient becomes even smaller and quite far from being significant (column 3). The same is true if we use the original dataset but drop countries with a Data Quality Grade of D (column 4) or if we drop the same countries while using version 6.2 (column 5).

Jones and Olken (2005)

This paper, published in the *Quarterly Journal of Economics*, 2005, relates growth to leadership. Specifically, it estimates the effect of random leader deaths on a country's growth rate. The two key findings are: such deaths have a significant impact on growth on average; and second, the death of leaders has a significant impact in autocracies but not in democracies. This paper used annual growth data from PWT 6.1 and their results are replicated in column 1 of Table 8.

When we re-estimated the core specification using data from PWT 6.2, in Table 7, we obtained the following differences compared with the original.

1. In the original paper, the coefficients of the random leader death are significant for the year of the leader death and for the two subsequent years. When we re-estimated it, using new data, the contemporaneous effect remained significant, but the effect for the two subsequent years ceased to be significant (p-value of 0.12 and 0.13, respectively); see column 3.
2. The really striking differences are related to the disaggregation of the results by type of political regime. In Jones and Olken (2005), random leader deaths had a

significant impact in autocracies in the year of the leader death and in the two subsequent years (column 1 in the 3rd panel, labeled Autocrats). In the robustness check of column 3 in the same panel, they were not significant for any of these three time horizons.

3. Even more interestingly, Jones and Olken find no significant effects in democracies. In the re-estimation using version 6.2 of the Table, leader deaths are significant in the year of the leader death and the following year (column 3 of the 4th panel, labeled Democrats).

Thus, not only the average effect but the pattern across political regimes seems to vary depending on whether data from PWT 6.1 or 6.2 are used.

The same pattern held when we dropped from the sample those countries categorized as very poor quality (grade D) from a data point of view (column 5 of Table 8), i.e., their original results hold (although the coefficients are smaller) with PWT 6.1 but the reversal of results is still the case with version 6.2.

Hausmann, Rodrik and Pritchett (2005)

This paper, published in the *Journal of Economic Growth*, 2006, identifies a set of countries that are deemed to have sustained growth over a long period of time. The paper used data from PWT 6.1. The criteria used to define sustained growth are that countries must have experienced: an *improvement* in growth rates of at least 2 percentage points per capita (this captures the idea of acceleration); *sustained* growth of at least 3.5 percent per capita for seven years; and a higher post-acceleration income level than the pre-acceleration peak (this is to ensure that accelerations are not simply a rebound from a

prior period of very bad performance, for example, due to wars or conflict or other shocks). In addition, *growth per capita* must remain *above 3 percent* after seven years, which captures the sense that good performance is sustained.

On this basis, Hausmann et al. identified 82 country-periods (see their Table 3) that met these criteria.[17] Our exercise was simply to see how the list of countries changes when the data from PWT 6.2 are used; the results are in Table 8. Between PWT 6.1 and 6.2, 100 country-periods were identified as sustained growers on this definition. However, only 65 cases were common to both data sets. In other words, there is a discrepancy of 35 percent between the two versions. There were 17 cases which Hausmann et al., using PWT 6.1 (for example, India in 1982 and Sri Lanka in 1979), identified as sustained growers but which did not show up as such using PWT 6.2. Conversely, there were 18 instances (for example, Indonesia in 1987) which were identified by PWT 6.2 as sustained growers but which were not so categorized by Hausmann et al. using PWT 6.1. One way of illustrating this discrepancy is to note that only in about 65 percent of the cases was there agreement between the two datasets; in other words, in 4 out of 10 instances, the answer depends on which data set is used (Table 9).

Miguel, Satyanath, and Sergenti (2004)

This paper, published in the *Journal of Political Economy*, 2004, attempts to identify and quantify the impact of economic factors on civil wars. The estimation is

[17] They actually identified 83 growth transitions but one of them has been excluded because data for Botswana does not go farther back than 1970 in PWT 6.2.

restricted to 41 countries in Africa and uses an annual panel framework using GDP data from PWT 6.1.

The key results of the paper are that in an OLS or probit framework, contemporaneous and one-period lagged growth does not have a statistically significant effect on civil conflict in Africa. However, when economic growth is instrumented with rainfall, one-period lagged growth has a statistically and economically significant effect on conflict – see column 1 (OLS) and column 4 in their Table 9 (IV) from the original.

When we re-ran these regressions using data from PWT 6.2 (our Table 10), we found that the IV results of Miguel et al. remained unchanged. However, the OLS results changed. Specifically, in the OLS framework, contemporaneous economic growth, which did not have a significant effect in the OLS becomes borderline statistically significant (in column 3, the coefficient on contemporaneous growth has a t-statistic of 1.67).

Our explanation of these results is that the OLS and probit estimations are more prone to measurement error (in this case on the right hand side) which is substantial in the annual panel framework, especially for Africa. This results in substantial differences in results. On the other hand, a good instrument addresses measurement error, resulting in a more robust estimation.

Aghion, Howitt, and Mayer-Foulkes (2005)

This paper, published in the Quarterly Journal of Economics, 2005, examined the effect of financial development on income (or technological) convergence. The paper's two core findings are that income convergence will depend on the level of financial

development; and that there will be a threshold level of financial development above which countries will converge and below which they will diverge.

Aghion et al. run a simple cross-country growth regression for a sample of 63 countries, where the key innovation is an interaction term between the level of financial development and the initial level of income divergence with the frontier country, the U.S. The key prediction of the model for which the estimation provides confirmation is that the coefficient on this interaction is negative (see the second row in column 1 of Table 11). When we re-run the Aghion et al. core regression using PWT 6.2 data – shown in Table 11 – we find that while the magnitude of this coefficient drops by about 40 percent, it remains statistically significant (second row of column 3).

However, the second implication of the Aghion et al. paper is not robust to changing the GDP data. Aghion et al. derive a threshold value of financial development, which is the ratio of the coefficient on the income convergence term and the coefficient on the interaction between income convergence and financial development. In the Aghion et al. specifications, this value is about 25 percent.

In the revised estimations, using PWT 6.2 data, the coefficient on the income convergence term is economically close to zero and statistically insignificant from zero (in one of the core specifications, the coefficient on this term switches signs relative to that in the paper). This yields a critical value for financial development of zero, suggesting that all countries will converge, which runs counter to the spirit of the main results implied by Aghion et al. (see column 5 of Table 11).

Mankiw, Romer and Weil (1992)

This paper, published in the *Quarterly Journal of Economics* in 1992, tests an augmented version of the Solow growth model, with the augmentation consisting of adding human capital as an additional input (apart from capital and labor) into production. The test comprises 2 elements: first, checking for the significance of the coefficients on savings (proxied by investment), human capital, and population growth (with the latter augmented to reflect technical progress and adjusted for depreciation); and second, importantly, checking for the magnitudes of these coefficients. In particular, the Solow model yields the result that the sum of the coefficients on savings, population and human capital should add to zero; or, in practice, the coefficient on the savings term adjusted for population growth should be equal in sign and magnitude to the coefficient on the human capital term also adjusted for population growth.

Mankiw et al. estimate a simple cross-section regression in which the left hand side is the level of per capita GDP (PPP, log). They estimate it for three samples, one consisting of 98 countries for which data are available and another for a sample of 75 countries in which countries classified as D in terms of data quality by the PWT are dropped, and a third where non-oil countries are dropped from the sample.

When we re-estimate their equations for the Solow model using PWT 6.2 data (and the non-oil sample), we find that the coefficients on savings, population growth and human capital are all correctly signed and statistically significant (Table 12). However, we find that the key test of the Solow model that the coefficients on savings and human capital (both adjusted for population) should be equal in magnitude comes close to being rejected by the data (in column 5, the p-value for the test of the Solow restriction is

0.166). This result is interesting because it contrasts with the general finding that long-horizon cross-sectional growth studies as well as those that use the level of GDP do not see significant changes in the results. In this study, the results change because the regressions use the investment-GDP ratio, which as the constant term in the specifications in Table 12 shows, is particularly prone to variability.

Discussion

Is there an explanation for the pattern observed? The least robust papers are those that use annual panel data, for which measurement error is large. Conversely, all the papers that survive the robustness are those that use low frequency data (i.e. cross-sectional estimations) or 5-year panel data. Evidently, averaging reduces "data revision error" and makes the estimations more robust.

But this raises another question. As long as this "data revision error" – which is in principle a particular kind of measurement error – is random why isn't it captured by the random error term in the growth regression? The answer is twofold: First, when the PPP-adjusted GDP appears as the independent variable (in level or growth) in regression estimation then the question posed above is indeed very valid. That is, *if* the "data revision error" were random then it should have been absorbed by the random error term. Unfortunately, the error is not random and suffers from the biases imposed by the PWT methodology as discussed previously. For example, whether the countries used in the estimation sample are small or large or whether the data used extends to the early 1960s biases the PPP-adjusted GDP measures and therefore cannot be appropriately captured by the random error term. Second, a typical growth regression usually incorporates

regressors that are either directly related to PPP-adjusted GDP (i.e. initial GDP), or components thereof (such as consumption, investment or government spending; directly measured by PWT as well). This introduces a measurement error in the dependent variable problem which although challenging could potentially be corrected econometrically. That an error-in-variable is a probable cause for papers we could not replicate, is supported by the fact that the 2SLS results of the Miguel et al. (2003) survive the robustness checks – despite using annual panel data. In Miguel et al., the core result is not robust to data revision in the OLS version but it is robust in the IV version. Having a good instrument helps overcome measurement error.

5. Understanding the Revisions to GDP Data: PWT Methodology

Thus far we have shown that large revisions to GDP data occur across PWT versions; that they are systematic; and that they matter for the cross-country growth literature. We now turn to explaining this variability. There is a large literature assessing the basic methodology employed by the PWT for determining the international prices and PPPs which go into the internationally comparable estimates of the level of per capita GDP. In addition to the series of papers by Kravis, Summers and Heston (1978) and Summers and Heston (1980, 1991 and 1996), notable contributions include Ciccone and Jarocinksi (2008), Deaton (2006), Dowrick (2005), Dowrick and Quiggin (1997), Feenstra et al. (2009), Heston (1994), Neary (2004), Nuxoll (1994), Rao and Selvanathan (1992), Samuelson (1994), Srinivasan (1994), and van Veelen (2002).

However, most of the discussion and criticism of the PWT methodology focuses on and relates to the *level* of international prices and *level* of PPP-adjusted GDP

especially as estimated for the benchmark year. Surprisingly, little attention has been given to the methodology and estimates for the *growth* rates of PPP-adjusted GDP, which then has implications for PPP-adjusted GDP levels for non-benchmark years. How are these growth rates and GDP levels for non-benchmark years calculated by the PWT?

Description of PWT methodology for growth estimates

Here we focus on the PWT growth estimates. Once the PWT calculates the level of PPP-adjusted GDP (and the associated domestic absorption, DA) for the benchmark year (say 1996), it calculates the levels for non-benchmark years according to the following equations. GDP for a non-benchmark year, say 1995, is calculated as:

$$Y_{95} = DA_{95} + \text{Net Foreign Balance}_{95} \qquad (1)$$

where Y is PPP-adjusted GDP and DA is domestic absorption.

$$DA_{95} = DA_{96}/(1 + \text{DAgrowth}_{95\text{-}96}) \qquad (2)$$

Now DA_{96} is domestic absorption for the benchmark year which is, by definition, in PPP terms. Consider next how $\text{DAgrowth}_{95\text{-}96}$ is computed:

$$\begin{aligned}\text{DAgrowth}_{95\text{-}96} = {} & [(\text{share of consumption in } DA_{95})*(\text{real consumption growth}_{95\text{-}96})] + \\ & + [(\text{share of investment in } DA_{95})*(\text{real investment growth}_{95\text{-}96})] + \\ & + [(\text{share of government in } DA_{95})*(\text{real government growth}_{95\text{-}96})] \qquad (3)\end{aligned}$$

Domestic absorption growth is calculated as the weighted average growth of its three components, **C**, **I** and **G**. The weights assigned to each of these components are the shares of each component in domestic absorption measured in 1995 and measured in international prices. These shares are obtained from "Step 2" described in Technical Appendix 1. But the growth rates of real **C**, **I**, and **G**, are from the national income accounts (constant price series) and as such these growth rates are at domestic not PPP prices. The chain series in the PWT uses contemporaneous shares while the Laspeyre's series uses fixed shares obtained for the benchmark year.

Equation (3) above offers clues to the striking result depicted in Chart 9. If the updating of the national income accounts leads to revisions in the growth of **C**, **I** and **G** (which are on the right hand side of this equation), correspondingly domestic absorption growth and hence GDP growth calculated by the PWT (on the left hand side) will also change. Note that national income accounts revisions could also change the shares of consumption, investment and government in domestic absorption in equation (3). However, these shares are at international prices, which are likely to be less sensitive to changes in national income accounts, especially to changes in data for the smaller countries.

A final point is worth noting. PWT 6.2 used updated national income accounts data. But in what sense is this newer or better data? The answer is: not clear. The current price series NIA data are virtually identical between PWT versions 6.1 and 6.2 with a GDP growth rate correlation of 0.992, pre-1996. It is the constant price series that changes, with a GDP growth rate correlation of 0.665 in the same period and sample. Essentially, the PWT re-based all national income account data from a 1990 base year to

2000. But the growth rates that this procedure yields do not always correspond to the growth rates reported say in the World Development Indicators by national authorities, raising the question whether the PWT's re-basing is appropriate and whether their estimate is superior to that of the national authorities?

Variability of estimate and size

In thinking about the variability of the estimates and trying to locate its causes, it is useful to recall the sequence for constructing the estimates. When a revision to the PWT occurs, the PPPs, and hence the GDP figures, change for the benchmark year due to the GK aggregation procedure. Next, the PPPs are changed for the non-benchmark years. This leads to a change in the share of real **C**, real **I**, and real **G** for non-benchmark years. Since these shares are used to construct the growth rate of domestic absorption, this is the channel by which GDP growth rates change across revisions.

The regressions in Tables 3-5 showed that per capita GDP growth but especially levels and the PPPs were significantly related to the size of a country as proxied by its level of GDP; more specifically, the smaller the country the larger the resulting revisions. Why is this the case? Recall that growth revisions happen because the shares of the three components of absorption change across PWT versions. In turn, these shares change because the PPP estimates for the non-benchmark years change. So, the question becomes why do PPP estimates change more for small countries than for the large ones?

A formal answer to this question is provided by Rao and Selvanathan (1992). They show that the PPPs and international prices can be seen as weighted averages, which makes it possible to interpret them as estimators of parameters from appropriately

specified regression models. Once these are seen as such, it then also becomes possible to assess the reliability of the estimates of these parameters. They show—equation (5) in Rao and Selvanathan (1992)—that the standard error of estimates of the PPPs are inversely related to a country's size (more specifically, its total consumption expenditure). The smaller the country, the less reliable is the estimate of its PPP, a proposition supported in the data (Table 3, columns 3, and 7).

Variability of estimate over time

One of the less recognized aspects of the PWTs is that the GDP estimates for the non-benchmark years are based on *extrapolated* PPPs. It is striking for example, that in PWT 6.2 and 6.1, the calculated GDPs for 1985 do not use actual international prices that were available from the benchmark studies from PWT 5.6. Instead, the calculations for 1985 start with the 1996 benchmark data and extrapolate backward using national income accounts data and the aggregation methods outlined in the Technical Appendix. "Step 2".

So, the question arises, does a new PWT version provide more reliable data for years other than the benchmark year? In Tables 3-5, we saw that the variability of the data between versions may actually increases the farther away one gets from the benchmark year. A heuristic explanation for this phenomenon is the following. Take 1995, the first non-benchmark year in PWT 6.1. The calculations of the international prices for 1995 are estimates based on the actual price data for 1996. So, some error is added to the 1995 estimates. If we take this back one more year to 1994, we know that 1994 international prices are in turn estimated from 1995 numbers. This adds one more layer of error. Relative to the benchmark year 1996, there are two sources of error for the

1994 estimates, and so on. This error structure going back in time is analogous to the error structure going forward imposed by many forecasting exercises. It would therefore not seem surprising that estimates are more variable the farther away from the benchmark year (see the pronounced funnel shape of Figure 7).

6. Are PWT GDP Estimates Really PPP Estimates?

At first blush, this is an odd question to ask. The raison d'être of the PWT is to come up with PPP prices and use them for computing GDP estimates. For the benchmark year, PPPs are calculated from the price data that are collected and these PPPs are used to compute GDP. So, in the benchmark year, the GDP estimates are at international prices. But matters turn out to be much murkier for years other than the benchmark year.

Recall equation (3):

$$\begin{aligned} DAgrowth_{95\text{-}96} = & [(\text{share of consumption in } DA_{95})*(\text{real consumption growth }_{95\text{-}96})] + \\ & + [(\text{share of investment in } DA_{95})*(\text{real investment growth }_{95\text{-}96})] + \\ & + [(\text{share of government in } DA_{95})*(\text{real government growth }_{95\text{-}96})] \end{aligned}$$

In this equation, the shares of the different components in DA are at international prices while the growth rates of the components are at domestic prices.

It is possible to further decompose equation (3) to see how the additions to domestic absorption are being valued. In Technical Appendix 3, we illustrate with a simple example that when there is more than one good, equation (3) results in valuing the quantity additions to each of the components of domestic absorption (between 1995 and 96) at some hybrid of domestic and international prices (see equation (5) in Technical Appendix 3). In general, this will be different from valuing these quantity changes at

international prices and this difference is likely to be systematically larger for smaller countries. Note that since GDP levels for non-benchmark years are calculated by applying growth rates to benchmark year level GDP estimates, the farther away we move from the benchmark year the more the level GDP estimates will be comprised of quantities that are not valued at international prices.

This leads to the striking conclusion: not only are growth rates of GDP not at international prices, but because of this fact, level GDP estimates are also not fully at international prices for non-benchmark years.[18]

7. Some Thoughts for an Alternative PPP-GDP Chained Growth Series

The existing PWT PPP-GDP series suffer from three shortcomings. First, as documented previously, the methodology leads to large and systematically biased variations across versions of the Table. Second, estimates of PPP-adjusted GDP growth rates and PPP-adjusted level of GDP for the non-benchmark years are not at international but rather some "hybrid" prices, which goes against the raison d'etre of PWT. Third, each new generation of PWT leads to discarding useful information from the disaggregated price data in all previous ICP benchmark studies.

Is there a way of calculating GDP estimates that overcome these problems? In principle, yes. First and foremost, the alternative approach could use the data and the estimates of the level of GDP compiled in the benchmark years (i.e., 1980, 1985, 1996,

[18] This is even more true for the WDI estimates of the level and growth rate of GDP because, not only are the growth rates of the components of domestic absorption not at international prices, but neither are the weights assigned to the three components.

and 2005). All data for the intervening years are extrapolated and not measured at international prices and should not be the basis for estimates.

The current methodology starts with the level of GDP in the benchmark year, then calculates the growth rates, which are in turn used to calculate the level of GDP in the benchmark years. We would propose doing it the other way around: to use the level estimates for the benchmark years and calculate the growth rates from these level estimates. Take the PPP-adjusted GDPs calculated for 1985 and 1996 from disaggregated ICP data. While cross-sectional comparisons can be made for each of the two years 1985 and 1996, there cannot be intertemporal comparisons based on these estimates because they are not at common prices. The aim, therefore, is to come up with GDP level estimates that are at common prices.

Here is a sketch of our proposed method. Assume that the level of GDP, for a country, at international prices in the benchmark year 1985 is:

$$Y_{85} = c_{85}^{1} P_{85}^{1} + c_{85}^{2} P_{85}^{2} + i_{85} P_{85}^{i}, \qquad (7)$$

where the P's are all international prices calculated from the Geary-Khamis aggregation, and superscripts refer to the type of consumption good (1 or 2) or to the investment good i (we assume that GDP comprises only C and I, and that C comprises only c^1, c^2).
The level of PPP-adjusted GDP in the second benchmark year is,

$$Y_{96} = c_{96}^{1} P_{96}^{1} + c_{96}^{2} P_{96}^{2} + i_{96} P_{96}^{i}. \qquad (8)$$

To calculate growth rates, we would ideally need to value either period 1985 quantities at 1996 prices or 1996 quantities at 1985 prices. Thus we can derive two new GDP estimates that are, respectively:

$$Y_{85}^{2} = c_{85}^{1} P_{96}^{1} + c_{85}^{2} P_{96}^{2} + i_{85} P_{96}^{i}\,, \qquad (9)$$

$$Y_{96}^{1} = c_{96}^{1} P_{85}^{1} + c_{96}^{2} P_{85}^{2} + i_{96} P_{85}^{i}\,. \qquad (10)$$

The point is that the data for computing Y according to equations (9-10) are generated by the PWT. For each benchmark year, the quantities of the different consumption and investment goods are available at the disaggregated level (recall that these are the inputs for the Geary-Khamis aggregation procedure, and these quantities are derived simply by deflating expenditures on each good at domestic prices by domestic prices). And for each benchmark year, the Geary-Khamis procedure yields disaggregated international prices.

Equations (7) and (10) will yield a GDP series that is intertemporally comparable (at period 1985 prices); Equations (8-9) will also yield a GDP series that is internationally comparable but at period 1996 prices. It is now simple to calculate growth rates.

We can then derive the growth rates of GDP either at base-year prices, which is

$$\dot{Y}_{85} = \frac{\Delta c^{1} P_{85}^{1} + \Delta c P_{85}^{2} + \Delta i P_{85}^{i}}{Y_{85}}\,. \qquad (11)$$

Or at current year prices which is,

$$\dot{Y}_{96} = \frac{\Delta c^1 P_{96}^1 + \Delta c^2 P_{96}^2 + \Delta i P_{96}^i}{Y_{85}^2}. \qquad (12)$$

Our preferred approach, however, would be to calculate a new chained series that can be used for international comparisons. This would involve averaging the two growth rates emerging from equations (11-12) to yield a chained growth estimate between these two years. The way we would envisage is that the chaining would be based on successive ICPs. Thus, there would be a growth estimate between 1985 and 1996 based on the ICPs in 1985 and 1996. Similarly, there would be an estimate between 1996 and 2005 based on the ICPs in these two years. And so on. The advantage of this approach would be that every time a new ICP is conducted, it would not change historical growth rates of GDP. Only growth rates going forward until the next ICP would be estimated.

Of course, a number of practical complications will have to be addressed. First, the basket of disaggregated goods varies across benchmark years, so some way of deriving a common basket will have to be found (i.e. equations (7-10) will need to have a common basket of goods).

A second problem with our proposal is that it would only provide a growth rate for the horizon between successive ICPs (either ten years, or over five years, in the future if ICPs are conducted every five years). There will not be a high frequency (i.e. annual) growth rate. But in assessing whether this is a major loss, two points should be kept in mind: annual data seem particularly problematic because of the current extrapolation methodology. Moreover, the motivation of the PWT is to make meaningful comparisons

of standards of living across countries at a given point in time. Taking long averages seems to preserve this scope but high frequency data (especially annual frequency comparisons) becomes unusable.

A third problem relates to samples. For 1985 and 1996, but especially for previous benchmark years, the sample of countries for which ICPs have been done remains limited. How to derive the chained growth numbers for a larger enough sample of countries going back before 1985 or 1980 needs further thought.

Once we have a chained growth estimate of the type we have proposed, this is how we would envisage the use of different PWT GDP series. First, for pure cross-country comparisons of the level of per capita income, researchers should use the estimates from the different benchmark years. Thus, if a researcher wants to compare incomes across countries in 2005, the benchmark estimate for 2005 should be used. If the time period is 1996, the benchmark estimate from the ICP estimate of PWT 6.1 (not 6.2) should be used. And if the time period is 1985, the estimate from PWT 5.6 for 1985 should be used. The key message here is that researchers should use that version of the PWT that is closest to the timing of inquiry and should NOT use the most recent version. That is, if the year of inquiry is 1985, the most recent PWT series (PWT 6.1 or 6.2) should NOT be used because the level estimate for 1985 in PWT 6.1 and 6.2 is NOT at international prices. Of course, if the year of inquiry is say 1990, researchers could use either the PWT 6.1 estimate for 1996 or the PWT 5.6 estimate for 1985.

Second, if the object is intertemporal growth comparisons, especially growth over the medium and long run, researchers should use the new chained growth estimates that we have proposed. Of course, these chained growth estimates could in principle lead to a

level GDP series (at least for the benchmark years) that is at a common international price. Thus, based on a chained growth estimate based on 1985 and 1996 ICPs, we could have a chained level estimate for 1985 that applies the new growth estimate to the level estimate in 1996, and extrapolates backwards. Or, we could have a new chained level estimate for 1996 that applies the new growth estimate to the level GDP numbers for 1985. But for pure cross-country comparisons, we would urge that researchers not use the level estimates derived from this chained growth procedure. Rather, they should follow the suggestion in point 1 above.

8. Conclusion

There is considerable variation in the level and growth of PPP-adjusted GDP estimates across alternative versions of the Penn World Table. This variation stems in part from changes to underlying national income accounts data but is also systematically related (inversely) to the size of a country and to the distance of the data from the benchmark data. Because this variability is intrinsic to the PWT methodology there is little basis for knowing whether newer versions are better than older versions. In fact, the distance from benchmark finding suggests that in every new version, historic data tends to become more variable.

This variation is enough so that some standard results in the growth literature are not robust across alternative versions of the Table. However, it does appear safe (from a Table Invariant point of view) to use PPP GDP level data, i.e., looking at cross-sections. Long-run changes, e.g., over 30 years, also appear to be robust. Medium-run growth

rates, e.g., 10 year panels, may also be safe, although we flag this as a point for further investigation.

It is not safe, in a Table Invariant sense, generally to use annual data from the Penn World Table. The exception would be for countries with quality grades of A and B or if there is a good instrumentation strategy. In general, annual data from non-OECD countries should be treated with a very large grain of salt.

Thus, growth studies should demonstrate robustness to different versions of the PWT. Where possible, robustness should also be demonstrated for samples of countries for which benchmark data are available and for samples excluding the small countries or those with a quality grading of C or D. If results do not survive one or both of these checks, a much bigger health warning should attach to any policy implications.

We also found, surprisingly, that GDP data from the PWT are not at PPP prices. To overcome this problem as well as the systematic variability (with respect to country size and distance from the benchmark year) of the GDP numbers, we suggested an alternative approach to calculating GDPs which has conceptual limitations but which may be easier to implement in practice. It is one of the remarkable features of the PWT exercise that each new version of the data leads to an almost complete discarding of international price data from previous versions and previous ICP studies. Our suggestion would avoid this problem. Moreover, by using estimates based on actual ICP prices, the proposal would avoid using extrapolations for all years other than the benchmark year, extrapolations which do not fully preserve the spirit that all comparisons of the standard of living should be at international prices. There are drawbacks to our proposed approach – particularly as it would reduce sample sizes and limit the period under consideration –

but given the substantial disparity in growth estimates between our suggested approach and the current PWT approach, we strongly suggest that researchers consider adding it to their robustness-checking toolkit. Indeed, a promising step in this direction is in Feenstra and Rao (2009).

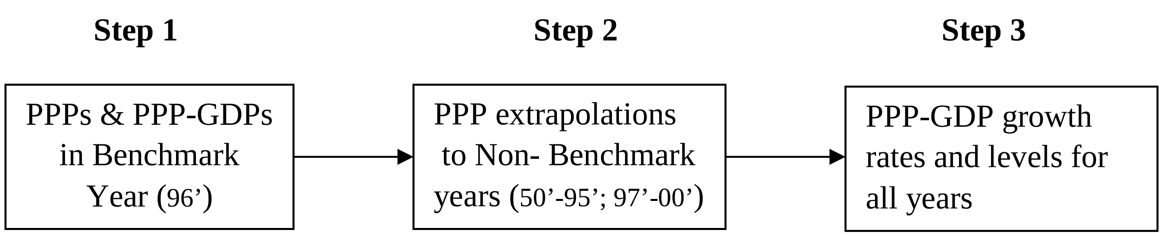
It remains to be seen how version 7.0 of the Penn World Table will confirm or change any of this assessment. It is also unclear whether that version will definitively supersede all previous versions (and all alternatives, including the World Development Indicators.)

Technical Appendix 1
The Penn World Table Calculation of the Level and Growth of PPP-Adjusted GDP

We outline below, the construction of the purchase power parity (PPP)-adjusted level of GDP and its growth rate. In the process, we highlight some key and relatively unknown facts about PWT data construction. The description below refers to version 6.1 of PWT.[19]

The procedure used by PWT to construct PPP-adjusted GDPs can be stripped down to three main steps as shown in Figure A1. In the first step, international price levels and PPP-adjusted GDPs are calculated for a benchmark year (for PWT 6.1 in 1996). In the second step these benchmark international prices are extrapolated backward and forward to produce PPP time series for each country considered. Finally, in step 3 PPP-adjusted GDP levels and growth rates are calculated for all countries and years using the relevant PPPs from step 2.

Figure A1: Constructing PPP-Adjusted GDPs



Next, we look into each of the three steps focusing on what is essential in the final construction of the PWT series.

Step 1: Obtaining PPPs and PPP-adjusted GDPs for the benchmark year
All calculations start for a benchmark year (the year for which detailed price data are collected; for PWT 6.1 in 1996) and for benchmark countries (the countries which participate in the International Comparison of Prices Program (ICP)); that is, benchmark countries are those for which actual and detailed price data are collected.[20]

The raw data for calculating PPPs for the benchmark year are:

[19] Special procedures are used for certain countries (i.e. China and India) at various stages. These are not described here and the interested reader is advised to consult the technical appendix to the PWT 6.1.
[20] See Table 1 for a list of benchmark years and benchmark countries used for different versions of PWT.

1. Expenditures in local currencies $(pq)_{ij}$ (obtained from the national income accounts), where *i* refers to categories of goods and services and *j* to the country; In PWT 6.1 there were 31 basic categories of goods and services.[21]
2. The individual prices p_{ij} for these basic goods and services categories in each country covered in the ICP exercise.

Dividing these expenditures by the individual prices yield notional quantities for each of the categories (i.e. $q_{ij} = (pq)_{ij} / p_{ij}$).

These data are used to derive two sets of <u>international</u> prices:

1. A set of international prices for each of the (31) categories of goods and services (π_i); that is these prices are the same for each of category for all countries.[22]
2. A purchasing power parity for each country (pl_j); PPPs for each country.

How are the (π_i)s and (pl_j)s derived? They are obtained from a system of two simultaneous equations which is also known as the Geary (1958) and Khamis (1972) (GK thereafter) aggregation procedure. These equations follow from the definition of the two prices as follows:

$$\pi_i = \sum_j \left[\left(\frac{p_{ij}}{pl_j} \right) \left(\frac{q_{ij}}{\sum_j q_{ij}} \right) \right] \qquad (1)$$

Equation (1) defines the international price for each commodity. The first expression on the right hand side ($\frac{p_{ij}}{pl_j}$) is the price of a category of goods *i* in country *j* deflated by that country's purchasing power parity, and the second ($\frac{q_{ij}}{\sum_j q_{ij}}$) is the share of *i* in country *j* in the world consumption of that good. The equation makes clear that the international price for each the 31 commodities is the weighted average of their domestic prices, where the weights are the share of each country in total world consumption of that commodity. Table A1, at the end of this appendix, provides a simple 2-good, 3-country example that illustrates the GK procedure.[23]

---

[21] Expenditures are multiplied by 'super–country weights' used by PWT to minimize the differences in results from adding or subtracting countries in the aggregation. For details on these weights see the technical appendix to the PWT 6.1 (pp. 6-7 and Table 1).

[22] International prices for each category are expressed relative to the U.S.; $p_{i,us} = 1$ for every *i*.

[23] In the example, in Table A1, take the equation for food:

$$10/PPP_A + 10/PPP_B + 12/PPP_C = 7*P_{food}$$

Here the expenditures on food in domestic prices by the three countries are 10, 10, and 12 respectively. From the example, information on quantities and prices in each country can be used to express the equation in the following manner:

$$1*10/PPP_A + 2*5/PPP_B + 4*3/PPP_C = 7*P_{food}$$

$$1*(P_{A,food}/PPP_A) + 2*(P_{B,food}/PPP_B) + 4*(P_{C,food}/PPP_C) = 7*P_{food}$$

(continued)

$$pl_j = \frac{\sum_i (pq)_{ij}}{\sum_i \pi_i q_{ij}} \qquad (2)$$

Equation (2) defines the price level of a country or its purchasing power parity (PPP). These relative country prices are the heart of the PWT, subsequently used to produce cross-country comparable GDP measures. For each country, the total expenditure in domestic prices divided by the sum of expenditures on each commodity at international prices yields the PPP for that country.[24,25]

As a result of the GK procedure, in PWT 6.1 there are international prices for 31 individual commodities, denoted as (π_i)s, and a set of country specific PPPs, denoted as (pl_j)s.

A final point is worth making here. The 31 categories of goods and services are also placed into 3 more aggregate categories corresponding to the three basic categories of Domestic Absorption (DA)—Consumption (*C*), Investment (*I*), and Government (*G*). For each country *j*, the country specific purchasing power parities for *C*, *I*, and *G* (pl_{Cj}, pl_{Ij}, pl_{Gj}) are also calculated. These will vary across countries simply because the composition of the 31 categories of goods and services (for which there are common international prices) will vary across countries.[26]

$$P_{food} = 1/7*(P_{A,food}/PPP_A) + 2/7*(P_{B,food}/PPP_B) + 4/7*(P_{C,food}/PPP_C)$$

In other words, the international price of food (P_{food}) is the weighted average of the individual domestic prices, where the weights are the share of each country in total world consumption of that commodity. Evidently, the international price will be closer to the domestic prices of those countries that have a larger weight in consumption and be different from the prices in the smaller countries. This is called the Gerschenkron effect (Gerschenkron, 1947).

[24] Once the PPPs are calculated for the benchmark countries for the benchmark year, the PWT proceeds to calculate the same for non-benchmark countries. This is done essentially through a two-step regression procedure, which yields coefficients which can be used to estimate the PPPs for non-benchmark countries. For the purpose of this paper, the estimation procedure is not relevant and is therefore skipped.

[25] There are also historical data on PPPs obtained from previous ICP rounds. For example, while for PWT 6.1 the ICP was done for 115 countries for the year 1996, a similar exercise was done in 1985 for 64 countries for version PWT 5.6. These 1985 PPPs can be extrapolated using national income accounts deflators for consumption, investment and government, to yield "PPPs" for 1996. So, for a number of countries, there are multiple data on PPPs for the benchmark year (1996). The PWT then utilizes a weighting method for these multiple data sources to arrive at a final set of international prices or PPPs for the year 1996 for all countries (see pages 9-11 of technical appendix to the PWT 6.1 for details).

[26] To demonstrate this, take the example in Table A1 and assume that food is 50% consumption and 50% investment and that haircuts are 100% investment. The PPPs for consumption and investment for country A will be calculated as follows:

Real consumption expenditure (at international prices) = quantity * price = 0.5 * 3.76 = 1.88.
Hence PPP for consumption, according to equation (2), is the expenditure at domestic prices divided by the expenditure at international prices which is (0.5*10)/1.88 = 2.66.

(continued)

These prices are used to calculate the PPP-adjusted GDPs for the benchmark year. GDP is calculated as the sum of real domestic absorption (measured at international prices) plus the net foreign balance. The constituent elements of domestic absorption—*C*, *I*, and *G*—are also measured at international prices by multiplying the quantities of C, *I*, and *G* by their respective international price obtained above.

$$\text{Real}(C_j) = \pi_c * q_c, \quad (3)$$

where Real(C_j) is consumption at international prices, π_c is international price of consumption and q_c is the notional quantity of consumption. By a similar procedure real *I* and real *G* are produced. Summing these three yields real domestic absorption as follows:

$$\text{Real}(DA_j) = \text{Real}(C_j) + \text{Real}(I_j) + \text{Real}(G_j) \quad (4)$$
$$\text{Real}(GDP_j) = \text{Real}(DA_j) + NFB_j, \quad (5)$$

where Real(DA_j) is domestic absorption in international prices for country *j*, Real(*GDP*)*j* is PPP-adjusted GDP and NFB is Net Foreign Balance. Note that the net foreign balance does not require the calculation of PPPs because they involve all traded goods.

Step 2: Calculating international prices for the non-benchmark years

Once PPPs for the benchmark year are calculated, the PWT extrapolates these PPPs going forward and backward in time. How does it do so? Take PWT 6.1 and the year 1995, which is one year before the benchmark year. For 1995, the PWT obtains nominal expenditures for the three components of domestic absorption—consumption, investment and government—from the national income accounts. These can be denoted as nomC_j, nomI_j and nomG_j,where *j* refers again to the country and the values are all for 1995. Each country *j* will also have a price level associated with for *C*, *I*, and *G* (pl_{Cj}, pl_{Ij}, pl_{Gj}, respectively) for 1995. These are just the 1996 PPPs for each of these categories, calculated as shown above, deflated by the price change between 1995 and 1996 for *C, I*, and *G*, where the price changes are obtained from national income accounts.

These extrapolated domestic prices for the three components of domestic absorption obtained (pl_C, pl_I, pl_G) are the equivalent of the detailed price data collected for the benchmark year. However, unlike in the case of the benchmark year where detailed prices collected for a large set of commodities for the non-benchmark years, each country has only three sets of international prices.

With these prices and with national income data for 1995 on consumption, investment and government expenditures, the PWT uses the GK aggregation procedure to calculate international prices (π)s for *C*, *I*, and *G*, as well as the PPPs (*pl*)s for all countries exactly as in step 1.

Similarly, PPP for investment = (0.5*10+ 10*1)/(0.5*3.6+10*0.8) = 1.52,
By a similar process, PPPs can be calculated for countries B and C.

These international prices are then used to convert consumption, investment, and government expenditures at domestic prices into expenditures at international prices as in equation (3) in step 1.

Step 3: Calculating the Level of PPP-adjusted GDP and its Growth rate for non-benchmark years

Step 2 yields, real *C*, real *I* and real *G* for all countries for non-benchmark years. How is the level of PPP-adjusted GDP then calculated, say for 1995? Essentially via a circuitous procedure that first calculates the growth rate of domestic absorption between 1995 and 1996 at international prices and then applies this growth rate to the 1996 level of Domestic Absorption (*DA*) to derive the level of DA in 1995. To this, the net foreign balance for 1995 is added to obtain the level of PPP-adjusted GDP for 1995.

The Chain Series:

$$\begin{aligned}\text{Real}(DA\text{growth}_{95\text{-}96})_j = & (\text{share}C \text{ in } DA_{95})_j * (\text{nom}C \text{ growth}_{95\text{-}96})_j \\ & + (\text{share}I \text{ in } DA_{95})_j * (\text{nom}I \text{ growth}_{95\text{-}96})_j \\ & + (\text{share}G \text{ in } DA_{95})_j * (\text{nom}G \text{ growth}_{95\text{-}96})_j \qquad (6)\end{aligned}$$

$$\text{Real}(DA_{95})_j = \text{Real}(DA_{96})_j \,/\, \text{Real}(DA\text{growth}_{95\text{-}96})_j \qquad (7)$$

$$\text{Real}(GDP_{95})_j = (DA_{95})_j + (NFB_{95})_j \qquad (8)$$

where [Real(DAgrowth$_{95\text{-}96}$)$_j$] is the growth of domestic absorption in international prices between 1995 and 1996 in country *j*, [nom*C* growth] is the growth of consumption expenditure in domestic price values, and [share*C* in *DA*] is the share of consumption in domestic absorption in international prices.

Three points are worth emphasizing about equation (6). First, the growth rates of real *C*, *I*, and *G* (nom*X* growth$_{95\text{-}96}$), are from the national income accounts and mostly do not change across PWT versions. Because they are from national income accounts, these growth rates are at domestic not PPP prices. However, the weights assigned to each of these components, the shares of each component in domestic absorption (share*X* in DA_{95}), are measured at international prices in 1995 which are obtained from steps 1 and 2; - i.e. (share*C* in DA_{95})$_j$ = Real(C_j) / Real(DA_j) = $\pi_c * q_c$ / (Real(C_j) + Real(I_j) + Real(G_j)).

Second, the PWT computes two PPP-adjusted GDP series,[27] the chained series (RGDPCH) which is the most commonly used and the one recommended by the authors of the PWT, and an RGDPL (or Laspeyre's series). The difference between the two is simply that in the RGDPCH, the shares change for every year the growth rate is

[27] There is a third series called RGDPTT which is not discussed here.

calculated; hence note that the shares are time-sensitive. In contrast, in the RGDPL series, the shares remain the same for all years and are the shares (calculated at international prices) for the benchmark year. Therefore, when the PWT is revised, the shares are revised for the benchmark year and all subsequent years. But for the RGDPL series, the changes in non-benchmark years are not relevant.

Third, this leads to understanding why the growth rates differ between the two series in the PWT and in turn how these differ from the PPP-adjusted growth rate calculation in the World Development Indicators. In the WDI, the growth rates will typically NOT change across revisions (other than to reflect revesions of national accounts) because the growth rates are calculated from the national income accounts. In terms of equation (4), PWT and WDI use the same numbers for real consumption, investment, and government growth, all obtained from the national income accounts. It is the shares that are different: in the WDI, these shares are from the national income accounts themselves, and therefore measured at domestic prices, and change with time; for the RGDPL series, the shares are at international prices but fixed at the levels of the benchmark year; for the RGDPCH series, the shares are at international prices and change every year based on changing international prices every year.

Steps 2-3 are then repeated for each of the years before and after the benchmark year to yield PPP-adjusted GDP growth rates and levels for these years.

Table A1: A Simple Example for Calculating PPPs and PPP-Adjusted GDPs

Basic data

Commodities	Country A			Country B			Country C			Total
	Quantity	Price	Expenditure	Quantity	Price	Expenditure	Quantity	Price	Expenditure	Quantity
Food	1	10	10	2	5	10	4	3	12	7
Haircuts	10	1	10	12	1	12	15	1	15	37
Total			20			22			27	
Currency exchange rates	Rupees 3 per $			Pesos 2 per $			Dollars			

Geary-Khamis procedure (Need to Find PPP for each country and International price for Food and Haircuts)

For each commodity: Each country's expenditure at national prices divided by PPP for that country yields total expenditure on that commodity at international prices

For Food $10/PPP_A + 10/PPP_B + 12/PPP_C = 7*P_{food}$-----------(1)

For Haircuts $10/PPP_A + 12/PPP_B + 15/PPP_C = 37*P_{hcuts}$----------(2)

For each country: The sum of expenditures on each commodity at international prices is the real expenditure divided by the national PPPs

For country A $1*P_{food}^{int} + 10*P_{hcuts}^{int} = 20/PPP_A$------------(3)

For country B $2*P_{food}^{int} + 12*P_{hcuts}^{int} = 22/PPP_B$------------(4)

For country C $4*P_{food}^{int} + 15*P_{hcuts}^{int} = 27/PPP_C$------------(5)

Since expenditures on all commodities equals that by all countries, there are 4 equations with four unknown, with one price PPP_C being the numeraire and set at 1.

The system yields the following solution:

$P_{food}^{int} = 3.76$ $PPP_A = 1.69$

$P_{hcut}^{int} = 0.8$ $PPP_B = 1.28$

	Real GDP at PPP exchange rates	**Real GDP at Market exchange rates**
Country A	20/1.69 = 11.83	6.7
Country B	22/1.28 = 17.19	11
Country C	27	27

Technical Appendix 2
Variability of PPP-Estimates with respect to Size and Distance from Benchmark Year

Variability of estimate and size

In thinking about the variability of the estimates and trying to locate its causes, it is useful to recall the sequence for constructing the estimates. When a revision to the PWT occurs, the PPPs, and hence the GDP figures, change for the benchmark year due to the GK aggregation procedure. Next, the PPPs are changed for the non-benchmark years. This leads to a change in the share of real **C**, real **I**, and real **G** for non-benchmark years. Since these shares are used to construct the growth rate of domestic absorption, this is the channel by which GDP growth rates change across revisions.

The regressions in Tables 4 and 5 showed that GDP level and growth revisions were significantly related to the size of a country as proxied by its level of GDP; more specifically, the smaller the country the larger the resulting revisions. Why is this the case? Recall that growth revisions happen because the shares of the three components of absorption change across PWT versions. In turn, these shares change because the PPP estimates for the non-benchmark years change. So, the question becomes why do PPP estimates change more for small countries than for the large ones?

A formal answer to this question is provided by Rao and Selvanathan (1992). They show that the PPPs and international prices can be seen as weighted averages, which makes it possible to interpret them as estimators of parameters from appropriately specified regression models.

The Geary-Khamis system can be represented in the following regression model:

$$\frac{p_i}{p_{ij}} = PPP_j + u_{ij} \ , i = 1,2,\ldots,N \qquad (1)$$

where p_i is the international price of good i, p_{ij} the price of good i in country j, PPP_j the purchasing power parity for country j, and u_{ij} is a random variable with zero mean and variance σ_{ij}^2.

If the following assumptions are made about the error structure:

$$E(u_{ij}) = 0$$
$$V(u_{ij}) = \sigma_u^2/(p_{ij}q_{ij})$$
$$E(u_{ij}u_{lk}) = 0$$

for all *i, j, l, k,* and $i \neq l, j \neq k$.

The generalized least squares estimator (GLS) of PPP_j is given by:

$$P\hat{P}P_j = \left(\sum_{i=1}^{N} P_i q_{ij}\right) \Big/ \left(\sum_{i=1}^{N} p_{ij} q_{ij}\right) \qquad (2)$$

By re-arranging equation (2)

$$P\hat{P}P_j = \sum_{i=1}^{N} w_{ij} \left(P_i / p_{ij}\right) \qquad (3)$$

the $P\hat{P}P_j$ can be interpreted as a weighted average of the international prices where the weights are the share of commodity i in country j in total expenditures in that country

$$w_{ij} = p_{ij} q_{ij} \Big/ \sum_{i} p_{ii} q_{ij}$$

With this system, it then also becomes possible to assess the reliability of the estimate of the PPPs. Thus, the standard error associated with $P\hat{P}P_j$ is given by:

$$SE\left(P\hat{P}P_j\right) = \left[\hat{\sigma}_u^2 \Big/ \left(\sum_{i=1}^{N} p_{ij} q_{ij}\right)\right]^{1/2} \qquad (4)$$

Equation (4) (equation (5) in Rao and Selvanathan (1992)) shows that the standard error is inversely related to the total consumption expenditure in country j, $\sum p_{ij} q_{ij}$. The smaller the country, the less reliable is the estimate of its PPP, a proposition supported in the data (Table 3, columns 3, and 7)

Variability of estimate over time

One of the less recognized aspects of the PWTs is that the GDP estimates for the non-benchmark years are based on *extrapolated* PPPs. It is striking for example, that in PWT 6.2 and 6.1, the calculated GDPs for 1985 do not use actual international prices that were available from the benchmark studies from PWT 5.6. Instead, the calculations for 1985, start with the 1996 benchmark data and extrapolated backward using national income accounts data and the aggregation methods outlined in steps 2 and 3.

So, the question arises, does a new PWT version provide more reliable data for years other than the benchmark year? In Tables 3-5, we saw that the variability of the data between versions may actually increases the farther away one gets from the benchmark year. A heuristic explanation for this phenomenon is the following. Take 1995, the first non-benchmark year in PWT 6.1. The calculations of the international prices for 1995 are estimates based on the actual price data for 1996. So, some error is

added to the 1995 estimates. If we take this back one more year to 1994, we know that 1994 international prices are in turn estimated from 1995 numbers. This adds one more layer of error. Relative to the benchmark year 1996, there are two sources of error for the 1994 estimates, and so on. This error structure going back in time is analogous to the error structure going forward imposed by many forecasting exercises. It would therefore not seem surprising that estimates are more variable the farther away from the benchmark year (see the pronounced funnel shape of Figure 7).

A slightly more formal version of this argument can be made based on the findings of Rao and Selvanathan (1992). Going back to equation (3) above, we see that the standard error of the GLS estimator for the PPPs is not only related to the size of a country (noted above) but also $\hat{\sigma}_u^2$, where this is an estimate of the σ_u^2 based on the GLS residuals in equation (1) above. The greater is $\hat{\sigma}_u^2$, the greater will be the variability of the estimates of the PPPs.

Note that under the PWT procedure equation (3) is estimated every year sequentially, starting with the benchmark year and going backward and forward. So, every year as random error is added in equation (1), the variance of $\hat{\sigma}_u^2$ increases. That is, the variance of $\hat{\sigma}_u^2$ for 1995 (the first non-benchmark year) is less than for 1994 which is less for 1993 and so on. Correspondingly, the standard errors of the estimates of the PPPs increase as we move away from the benchmark year which results in the funnel shape in Figure 7.

Technical Appendix 3: Growth Rate Measurement in the PWT

How are growth rates computed in the Penn World Table? To answer this question and focus on the key points, we assume that GDP comprises consumption (**C**) and investment (**I**) and is hence equal to domestic absorption (**DA**).

One consumption and one investment good
We start with the case where there is only one consumption and one investment good. First, some notation. The quantities of these goods are denoted by small case letters, c and i, respectively. Subscripts, which apply only to value and price variables, refer to whether they are measured at domestic or international (i.e. purchasing power parity) prices. Thus P_{dom} refers to domestic price and P_{int} to international prices. Superscripts refer to the year of measurement, and can either be the current year (T) or the base year (B). For example, C^T denotes real consumption for the year T. For the price variables, there will be two superscripts, the first referring to the year of measurement and the second to the good (consumption, c, or investment, i). Thus P_{dom}^{BC} refers to the price of the consumption good measured at international prices for the base year.

Thus GDP = DA = C+ I = Y (1)

$$\dot{Y}^T = \frac{C_{\text{int}}^T}{Y_{\text{int}}^T} \cdot \frac{\Delta C_{dom}^T}{C_{dom}^T} + \frac{I_{\text{int}}^T}{Y_{\text{int}}^T} \cdot \frac{\Delta I_{dom}^T}{I_{dom}^T} \qquad (2)$$

Equation (2) simply says that GDP growth in year T is a weighted average of growth of consumption and investment. In the PWT's chain series (RGDPCH), the weights are the shares of C and I in domestic absorption measured at current international prices. So, if the growth rate is calculated for the period 1993 to 1994, the weights are at 1993 PPP prices obtained from the Geary-Khamis aggregation procedure described in step xx in Appendix. Note that the growth rates of C ($\frac{\Delta C_{dom}^T}{C_{dom}^T}$) and I ($\frac{\Delta I_{dom}^T}{I_{dom}^T}$) are obtained from the national income accounts and are therefore at domestic base-year prices, where this base-year could vary across countries.

Equation (2) can be re-written as:

$$\dot{Y}^T = \frac{c^T P_{\text{int}}^{TC}}{Y_{\text{int}}^T} \cdot \frac{\Delta c^T}{c^T} \cdot \frac{P_{dom}^{BC}}{P_{dom}^{BC}} + \frac{i^T P_{\text{int}}^{Ti}}{Y_{\text{int}}^T} \cdot \frac{\Delta i^T}{i^T} \cdot \frac{P_{dom}^{Bi}}{P_{dom}^{Bi}}$$

which in turn simplifies to:

$$\dot{Y}^T = \frac{\left(\Delta c^T P_{\text{int}}^{TC} + \Delta i^T P_{\text{int}}^{Ti}\right)}{Y_{\text{int}}^T} \qquad (3)$$

Equation (3) shows that the PWT chain series for the growth rate essentially involves valuing the additions to (the quantities) of consumption and investment at current year international prices. This is a chain-weighted index because the prices used for valuing these additions change every year.

As a result, the RGDPCH series does not use benchmark year international prices; it uses current year international prices that are obtained in the PWT by extrapolating from the benchmark years (see Appendix) and applying the G-K aggregation procedure. These prices change with every revision as shown below and the revisions are not random but systematically related to country attributes.

A second feature of these current year international prices is that they are calculated at a highly aggregated level, namely at the level of aggregate consumption and investment. For the benchmark year, aggregate consumption is obtained by adding up consumption of the different goods for which disaggregated price data are available. For non-benchmark years, it is as if there is only one consumption and one investment good.

Does this aggregation affect the calculation of growth rates? Suppose that there are two consumption goods and one investment good. The spirit of the PWT suggests that all three goods should be valued at international prices. But how are they actually valued? If we had disaggregated international prices, we can write down how growth of DA should be computed. Essentially, equation (2) should be re-written to take account of the extra consumption good. Thus,

$$\dot{Y}^T = \frac{\left(\Delta c^{T1} P_{\text{int}}^{T1} + \Delta c^{T2} P_{\text{int}}^{T2} + \Delta i^T P_{\text{int}}^{Ti}\right)}{Y_{\text{int}}^T} \qquad (4)$$

Now, the second superscript is 1 or 2 for the two consumption goods and i for the investment good. Equation (4) is just an extension of equation (3) and says that GDP growth is obtained by valuing each of the quantity changes (to the consumption goods and investment good) at their respective current international prices.

But for non-benchmark years, we do not have disaggregated international price data. GDP growth is measured thus:

$$\hat{\dot{Y}}^T = \frac{\overline{C}_{\text{int}}^T}{Y_{\text{int}}^T}\left(\frac{\Delta C_{dom}^{T1} + \Delta C_{dom}^{T2}}{\overline{C}_{dom}^T}\right) + \frac{I_{\text{int}}^T}{Y_{\text{int}}^T} \cdot \frac{\Delta I_{dom}^T}{I_{dom}^T} \quad \text{which can be simplified to:}$$

$$\hat{\dot{Y}}^T = \frac{\left(c^{T1} + c^{T2}\right)\overline{P}_{\text{int}}^{TC}}{Y_{\text{int}}^T}\left\{\frac{\left(\Delta c_{dom}^{T1} P_{dom}^{B1} + \Delta c_{dom}^{T2} P_{dom}^{B2}\right)}{\left(c^{T1} + c^{T2}\right)\overline{P}_{dom}^{TC}}\right\} + \frac{\Delta i^T}{Y_{\text{int}}^T} \cdot \frac{P_{dom}^{Bi}}{P_{dom}^{Bi}} \quad \text{and further to:}$$

$$\hat{\dot{Y}}^T = \frac{1}{Y_{\text{int}}^T}\left\{\Delta c^{T1}\left(\frac{P_{dom}^{B1}}{\overline{P}_{dom}^{BC}}\cdot \overline{P}_{\text{int}}^{TC}\right) + \Delta c^{T2}\left(\frac{P_{dom}^{B2}}{\overline{P}_{dom}^{BC}}\cdot \overline{P}_{\text{int}}^{TC}\right)\right\} + \frac{\Delta i^T P_{\text{int}}^{Ti}}{Y_{\text{int}}^T} \quad (5)$$

Equation (5) shows that each of the quantity changes is valued not at current year international prices (as in equation (4)) but at some hybrid of domestic and international prices. The bar sign over a variable denotes that it refers to the composite good as does the second superscript C. The price term in brackets for the two consumption goods consists of ($\left\{\left(\frac{P_{dom}^{B1}}{\overline{P}_{dom}^{BC}}\right)\right\}$ which is the domestic relative price of that good (that is, it the domestic price of good 1 relative to the average price of the composite of goods 1 and 2); and $\overline{P}_{\text{int}}^{TC}$ is the average international price of the composite of goods 1 and 2.

To more clearly identify the difference between how GDP growth ought to be measured and how it is, we can take the difference between equations (4-5), which yields:

$$\dot{Y}^T - \hat{\dot{Y}}^T = \frac{\overline{P}_{\text{int}}^{TC}}{Y_{\text{int}}^T}\left\{\Delta c^{T1}\left(\frac{P_{\text{int}}^{T1}}{\overline{P}_{\text{int}}^{TC}} - \frac{P_{dom}^{B1}}{\overline{P}_{dom}^{BC}}\right) + \Delta c^{T2}\left(\frac{P_{\text{int}}^{T2}}{\overline{P}_{\text{int}}^{TC}} - \frac{P_{dom}^{B2}}{\overline{P}_{dom}^{BC}}\right)\right\} \quad (6)$$

This equation shows that the difference between the two depends on how different is the relative price of a consumption good at domestic prices ($\frac{P_{dom}^{B1}}{\overline{P}_{dom}^{BC}}$) from its relative price at international prices ($\frac{P_{\text{int}}^{T1}}{\overline{P}_{\text{int}}^{TC}}$). Note that this difference will vary across time because domestic prices are computed for a fixed base period (which can be different across countries) while the international prices are current prices.

We know that this difference in the relative prices will vary systematically across countries. It will be greater for smaller countries because under the G-K procedure domestic prices of larger countries have a greater weight when computing international prices. This is the Gerschenkron effect. Thus, GDP growth rates are likely to be measured with greater error (relative to the true growth rate that is consistent with the spirit of the PWT as represented in equation (4).) for smaller countries.

A second problem is that the farther away T is from the base year, the greater the discrepancy. Hence growth rate calculations for years farther away from the benchmark year are likely to have greater measurement error.

References

Acemoglu, Daron, Simon Johnson, James Robinson and Yunyong Thaicharoen. (2003). "Institutional Causes, Macroeconomic Symptoms: Volatility, Crises and Growth," *Journal of Monetary Economics* 50, 49–123.

Aghion, Philippe, Peter Howitt and David Mayer-Foulkes. (2005)."The Effect of Financial Development on Convergence: Theory and Evidence," *Quarterly Journal of Economics* 120, 173–222.

Barro, Robert, J. (1999). "Determinants of Democracy," *Journal of Political Economy* 107, S158–29.

Burnside, Craig and David Dollar. (2000). "Aid, Policies, and Growth," *American Economic Review* 90, 847–868.

Ciccone, Antonio and Marek Jarocinksi. (2008). "Determinants of Economic Growth: *Will Data Tell?" Working Paper Series No. 852, European Central Bank.*

Deaton, Angus. (2006). "Purchasing Power Parity Exchange Rates for the Poor: Using Household Surveys to Construct PPPs" working paper, Research Program in Developing Studies, Princeton University.

Deaton, Angus and Alan Heston. (2008). "Understanding PPPs and PPP-Based National Accounts," *American Economic Journal – Macroeconomics*, forthcoming.

DeLong, Bradford J. and Lawrence H. Summers. (1991). "Equipment Investment and Economic Growth," *Quarterly Journal of Economics* 106, 445–502.

Demirguc-Kunt, Asli, Luc Laeven and Ross Levine. (2004). "Market Structure, Institutions, and the Cost of Financial Intermediation," *Journal of Money Credit and Banking* 36, 593–622.

Dowrick, Steve. (2005). "The Penn World Table: A Review," *Australian Economic Review* 38, 223–228.

Dowrick, Steve and John Quiggin. (1997). "True Measures of GDP and Convergence," *American Economic Review* 87, 41–64.

Easterly, William, Michael Kremer, Lant Pritchett and Lawerence Summers. (1993). "Good Policy or Good Luck?: Country Growth Performance and Temporary Shocks," *Journal of Monetary Economics* 32, 459–483.

Feenstra, Robert, Alan Heston, Marcel P. Timmer and Haiyan Deng. (2009). "Estimating Real Production *and* Expenditures Across Nations: A Proposal for Improving the Penn World Tables," *Review of Economics and Statistics* 91, 201-212.

Feenstra, Robert, and D.S. Prasada Rao, "Consistent Comparions of Real Incomes across Space and Time," 2009, mimeo.

Geary, Roy C. (1958). "A Note On the Comparison of Exchange Rates and Purchasing Power Between Countries," *Journal of the Royal Statistical Society, Series A (General)* 121, 97–99.

Gershenkron, Alexander. (1947). "The Soviet Indices of Industrial Production," *Review of Economic Studies* 29, 217–26.

Handbook of the International Comparison Programme. http://unstats.un.org/unsd/methods/icp/ipc6_htm.htm

Hausmann, Ricardo, Lant Prichett and Dani Rodrik. (2005). "Growth Accelerations," *Journal of Economic Growth* 10, 303–29.

Heston, Alan. (1994). "National Accounts: A Brief Review of Some Problems in Using National Accounts Data in Level of Output Comparisons and Growth Studies," *Journal of Development Economics* 44, 29–52.

International Monetary Fund. (2008). *World Economic Outlook*, Washington, DC.

Jones, Ben and Benjamin Olken. (2005). "Do Leaders Matter? National Leadership and Growth Since World War II," *Quarterly Journal of Economics* 120, 835–864.

Khamis, Salem H. (1972). "A New System of Index Numbers for National and International Purposes," *Journal of the Royal Statistical Society, Series A (General)* 135, 96–121.

Kravis, Irving, B., Alan Heston and Robert Summers. (1978). "Real GDP Per Capita for More Than One Hundred Countries," *Economic Journal* 88, 215–242.

Mankiw, Gregory N., David Romer, and David N. Weil. (1992). "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics* 107, 407–437.

Miguel, Edward, Shanker Satyanath and Ernest Sergenti. (2004). "Economic Shocks and Civil Conflict: An Instrumental Variables Approach," *Journal of Political Economy*, 112(4), 725-753.

Neary, Peter, J. (2004). “Rationalizing the Penn World Table: True Multilateral Indices for International Comparisons of Real Income,” *American Economic Review* 94, 1411–1428.

Nuxoll, Daniel, A. (1994). “Differences in Relative Prices and International Differences in Growth Rates,” *American Economic Review* 84, 1423–1436.

Penn World Table, versions 1.0, 4.0, 5.0, 5.6, 6.1, 6.2. Constructed by Alan Heston, Irving B. Kravis and Robert Summers (and Bettina Aten, versions 6.1 and 6.2). http://pwt.econ.upenn.edu/.

Ramey, Valerie A. and Garey Ramey. (1995). “Cross-Country Evidence on the Link between Volatility and Growth,” *American Economic Review* 85, 1138–1151.

Rao, Prasada, D.S. and Selvanathan E.A. (1992). “Computation of Standard Errors for Greary-Khamis Parities and International Prices: A Stochastic Approach,” *Journal of Business and Economics Statistics* 10, 109–115.

Sachs, Jeffrey. D. and Warner, Andrew. M. (1995). “Economic Reform and the Process of Global Integration,” *Brookings Papers on Economic Activity*, 1–118.

Samuelson, Paul, A. (1994). “Facets of Balassa-Samuelson Thirty Years Later,” *Review of International Economics* 3, 201–226.

Srinivasan, T. N. (1994). “Data Base for Development Analysis: An Overview,” *Journal of Development Economics* 44, 3–27.

Summers, Robert and Alan Heston. (1980). “A New Set of International Comparisons of Real Product and Price Levels Estimates For 130 Countries, 1950-1985,” *Review of Income and Wealth* 26, 19–66.

Summers, Robert and Alan Heston. (1991). “The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988,” *Quarterly Journal of Economics* 106, 327–368.

Summers, Robert and Alan Heston. (1996). “International Price and Quantity Comparisons: Potentials and Pitfalls,” *American Economic Review* 86, 20–24.

van Veelen, Matthijs. (2002). “An Impossibility Theorem Concerning Multilateral International Comparisons of Volumes,” *Econometrica* 70, 369–75.

World Bank. (2007). *World Development Indicators,* Washington, DC.

World Bank. (2008a). *Global Purchasing Power Parities and Real Expenditures: 2005*

International Comparison Program, Washington, DC.

World Bank. (2008b). *ICP 2003–2006 Handbook*, online at http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/ICPEXT/0,contentMDK:20126612~pagePK:60002244~piPK:62002388~theSitePK:270065,00.html.

Table 1. The Evolution of the Penn World Table

PWT Version	Benchmark Year	Year Released	Countries in the PWT
1.0	1970	1980	119
2.0	-	-	-
3.0	1975	1984	115
4.0	1980	1988	128
5.0	1985	1991	134
5.6	1985	1995	151
6.1	1996	2002	168
6.2	1996	2006	188
7.0	2005	TBD	TBD

ICP Phases	Year(s)	Number of Countries
Phase 1	1970	10
Phase 2	1973	16
Phase 3	1975	35
Phase 4	1980	61
Phase 5	1985	62
Phase 6	1996	115
Phase 7	2005	146

Notes:

Penn World Table, Version 2, was never released. There are some discrepancies in these numbers across different PWT sources. This table uses information from the datasets themselves and the PWT 6.1 appendix. Other information can be found in the various papers published along with releases of versions of the Penn World Table.

Table 2: Average Growth Rates between 1975 and 1999: African Countries

PWT 6.2		PWT 6.1		
Top 10 Countries				
PWT 6.2		PWT 6.1		Countries Not Appearing on Both Lists
Country	Growth	Country	Growth	
Botswana	4.5%	Botswana	5.1%	Congo, Republic of
Equ.Guinea	4.0%	Cape Verde	4.7%	Equatorial Guinea
Cape Verde	3.7%	Mauritius	4.3%	Ethiopia
Egypt	3.7%	Egypt	3.7%	Malawi
Mauritius	3.7%	Tunisia	2.5%	Mali
Lesotho	3.5%	Uganda	1.7%	Mauritius
Tunisia	2.7%	Morocco	1.7%	Uganda
Mali	2.0%	Lesotho	1.5%	
Ethiopia	1.6%	Congo, Rep of	1.5%	
Morocco	1.6%	Malawi	1.2%	
Bottom 10 Countries				
PWT 6.2		PWT 6.1		Countries Not Appearing on Both Lists
Country	Growth	Country	Growth	
Gabon	-2.6%	Equ. Guinea	-2.7%	Chad
Zambia	-2.1%	Mozambique	-2.4%	Cote d'Ivoire
Madagascar	-1.9%	Zambia	-1.8%	Gabon
Togo	-1.7%	Comoros	-1.6%	Guinea-Bissau
Guinea-Bissau	-1.4%	Madagascar	-1.4%	Mauritania
Comoros	-1.2%	Cote d'Ivoire	-1.4%	Mozambique
Niger	-0.7%	Niger	-1.3%	Namibia
Nigeria	-0.5%	Mauritania	-1.3%	Niger
Chad	-0.5%	Togo	-1.0%	Nigeria
Mozambique	-0.4%	Namibia	-0.9%	Togo
				Country Switching Lists: Equ. Guinea

Notes:
This table presents the countries with the top 10 and bottom 10 average growth rates between 1975 and 1999 calculated using PWT 6.2 and PWT 6.1 GDP data (RGDPCH). Seven countries appear on the top 10 list for one dataset but not the other. Ten countries appear on the Bottom 10 list for one dataset but not the other. One country, Equatorial Guinea, switches lists. According to PWT 6.1, it is the worst-performing country; according to PWT 6.2, it is the second highest performing country.

Table 3. Explaining Differences in Price Levels Across Datasets

Dependent variable →	**Abs(difference in log P)**						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Estimation	**OLS**						
Dataset Comparison →	**PWT6.2 vs. PWT6.1**						
Sample	**1970 to 1999**						
PWT Grade (D=1…A=4)	-5.657***						
	[-15.0]						
Total ICP Studies		-2.409***					0.964***
		[-13.0]					[3.60]
Log(GDP)			-2.091***				-0.916***
			[-10.1]				[-4.65]
Abs(Difference in NA GDP)				-0.00220*			0.00846***
				[-1.69]			[6.61]
Abs(Difference in NA GDP Growth)					0.799***		0.496***
					[6.87]		[4.91]
Distance from Benchmark Year, in Decades (1996)						5.579***	10.96***
						[11.3]	[9.98]
Distance * Total ICPs							-2.319***
							[-8.02]
Constant	23.46***	16.11***	61.86***	10.45***	8.073***	3.925***	22.46***
	[20.5]	[22.6]	[11.7]	[31.3]	[22.2]	[9.92]	[5.07]
N	3016	3016	3016	3016	3016	3016	3016
R-squared	0.085	0.062	0.058	0.000	0.057	0.062	0.198

Notes:

The dependent variable is the absolute value of the difference in the log-level of prices (P) across PWT 6.2 and PWT 6.1. The variable PWT grade can take on four values, 1-4, with 1 representing the worst PWT rating of D, and 4 representing the best PWT rating of A. Total ICP studies refers to the number of ICP studies a country has participated in. GDP is total GDP and is measured at purchasing power parity. The distance variable is calculated as the absolute difference between the year of the observation and 1996 (i.e., abs(t-1996)), and is divided by 10 to express the variable in decades. 1996 is the "Benchmark Year" used in constructing both PWT 6.2 and PWT 6.1. The sample consists of the 104 countries in the “Long Run Sample” used in other tables and figures. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Robust t-statistics in brackets.

Table 4. Explaining Differences in Levels of GDP Per Capita Across Datasets

Dependent variable →	Abs(difference in log per capita PPP-adjusted GDP)						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Estimation	**OLS**						
Dataset Comparison →	**PWT6.2 vs. PWT6.1**						
Sample	**1970 to 1999**						
PWT Grade (D=1…A=4)	-4.036***						
	[-18.0]						
Total ICP Studies		-2.319***					-0.733***
		[-18.1]					[-3.39]
Log(GDP)			-1.758***				-0.452***
			[-13.6]				[-3.25]
Abs(Difference in NA GDP)				0.00118			0.0121***
				[1.05]			[11.1]
Abs(Difference in NA GDP Growth)					0.752***		0.534***
					[7.35]		[5.21]
Distance from Benchmark Year, in Decades (1996)						2.456***	4.813***
						[6.99]	[7.05]
Distance * Total ICPs							-1.074***
							[-6.10]
Constant	21.69***	17.90***	55.67***	12.21***	10.21***	9.495***	20.32***
	[30.7]	[35.4]	[16.8]	[45.7]	[33.5]	[23.7]	[6.01]
N	3016	3016	3016	3016	3016	3016	3016
R-squared	0.070	0.091	0.066	0.000	0.081	0.019	0.18

Notes:

The dependent variable is the absolute value of the difference in the log-level of GDP across PWT 6.2 and PWT 6.1. The variable PWT grade can take on four values, 1-4, with 1 representing the worst PWT rating of D, and 4 representing the best PWT rating of A. Total ICP studies refers to the number of ICP studies a country has participated in. GDP is total GDP and is measured at purchasing power parity. The distance variable is calculated as the absolute difference between the year of the observation and 1996 (i.e., abs(t-1996)), and is divided by 10 to express the variable in decades. 1996 is the "Benchmark Year" used in constructing both PWT 6.2 and PWT 6.1. The sample consists of th 104 countries in the “Long Run Sample” used in other tables and figures. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Robust t-statistics in brackets.

Table 5. Explaining Differences in Growth of GDP Per Capita Across Datasets

Dependent variable →	Abs(difference in growth rate of per capita GDP)						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Estimation	**OLS**						
Dataset Comparison →	**PWT6.2 vs. PWT6.1**						
Sample	**1970 to 1999**						
PWT Grade (D=1…A=4)	-1.716***						
	[-17.8]						
Total ICP Studies		-0.678***					-0.0598
		[-14.0]					[-1.50]
Log(GDP)			-0.784***				-0.0990***
			[-16.1]				[-3.65]
Abs(Difference in NA GDP)				-0.00262***			0.000377**
				[-7.39]			[2.02]
Abs(Difference in NA GDP Growth)					0.842***		0.825***
					[26.6]		[24.6]
Distance from Benchmark						0.366***	0.0416
Year, in Decades (1996)						[3.21]	[0.37]
Distance * Total ICPs							-0.0106
							[-0.37]
Constant	6.767***	4.409***	22.12***	2.962***	0.432***	2.355***	3.020***
	[22.6]	[23.6]	[17.5]	[30.2]	[6.32]	[15.4]	[4.25]
N	3016	3016	3016	3016	3016	3016	3016
R-squared	0.093	0.058	0.096	0.005	0.747	0.003	0.750

Notes:

The dependent variable is the absolute value of the difference in the growth rate of GDP across PWT 6.2 and PWT 6.1. The variable PWT grade can take on four values, 1-4, with 1 representing the worst PWT rating of D, and 4 representing the best PWT rating of A. Total ICP studies refers to the number of ICP studies a country has participated in. GDP is total GDP and is measured at purchasing power parity. The distance variable is calculated as the absolute difference between the year of the observation and 1996 (i.e., abs(t-1996)), and is divided by 10 to express the variable in decades. 1996 is the "Benchmark Year" used in constructing both PWT 6.2 and PWT 6.1. The sample consists of the 104 countries in the “Long Run Sample” used in other tables and figures. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Robust t-statistics in brackets.

Table 6. Differences Between PWTs 6.2, 6.1, and 5.6

	Annual Growth Rates	Decadal Growth Rates		Long-Run Growth Rate
	1970 to 1990	1970s	1980s	1970-1990
GDP (RGDPCH)				
PWT 6.2 - PWT 6.1				
Avg GR (6.2)	1.6%	2.2%	1.0%	1.6%
Mean Difference	0.0%	0.0%	0.0%	0.0%
SD Difference	5.3%	1.7%	1.4%	1.0%
PWT 6.2 - PWT 5.6				
Avg GR (6.2)	1.6%	2.2%	1.0%	1.6%
Mean Difference	0.1%	-0.2%	0.5%	0.1%
SD Difference	5.9%	1.9%	1.3%	1.1%
PRICES (P)				
PWT 6.2 - PWT 6.1				
Avg (6.2)	62.5	70.4	67.4	65.0
Mean Difference	2.7%	3.3%	2.4%	2.9%
SD Difference	20.7%	23.6%	14.5%	18.6%
PWT 6.2 - PWT 5.6				
Avg (6.2)	62.5	70.4	67.4	65.0
Mean Difference	-10.2%	-9.8%	-5.1%	-7.5%
SD Difference	27.2%	31.5%	28.9%	29.5%
Obs	2000	100	100	100

Table 7. Ramey and Ramey (1995) Replication Results

Dependent variable →	**growth of per capita GDP**				
	[1]	[2]	[3]	[4]	[5]
Dataset Substitution	None	None	PWT6.2	None	PWT6.2
Sample	Original	Balanced	Balanced	Balanced	Balanced
PWT Grades Included	All	All	All	A,B,C	A,B,C
Standard Deviation of Growth Rates	-0.177***	-0.151*	-0.074	-0.107	-0.132
	[2.426]	[1.821]	[0.94]	[1.254]	[1.460]
N	2208	1776	1776	1608	1608

Notes:
This table presents a replication exercise of Ramey and Ramey (1995) by updating their original PWT 5.6 GDP data with PWT 6.2 data. All specifications replicate Table 1, equation (1), in Ramey and Ramey (1995). For presentation purposes we focus only on the key parameter - variation of growth rate - and omit other parameter estimates. Specification [1] replicates the main result in Ramey and Ramey (1995) using their original data. Specification [2] presents the same result using PWT 5.6 after dropping observations to balance the data with available observations in PWT 6.2. Specification [3] presents our main result when we replace the balanced sample using PWT 5.6 with PWT 6.2. Specification [4] replicates [2] when countries with quality grading "D" are dropped. Specification [5] replicates [3] when countries with quality grading "D" are dropped. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. t-statistics in brackets.

Table 8. Jones and Olken (2005) Replication Results

Dependent variable	→ **Annual Growth rate**				
Hypothesis	→ **Leader deaths affect growth (Wald p-value)**				
	[1]	[2]	[3]	[4]	[5]
Dataset Substitution	None	None	PWT6.2	None	PWT6.2
Sample	Original	Balanced	Balanced	Balanced	Balanced
PWT Grades Included	All	All	All	A,B,C	A,B,C
Years after leader's death	All Leaders				
0	0.057*	0.054*	0.093*	0.012**	0.110
1	0.085*	0.122	0.218	0.028**	0.130
2	0.067*	0.133	0.203	0.039**	0.141
Number of leader deaths	57	52	52	45	45
Years after leader's death	All Leaders, Tenure >=2 Years				
0	0.039**	0.039**	0.04**	0.011**	0.049**
1	0.054*	0.087*	0.140	0.021**	0.068*
2	0.031**	0.076*	0.102	0.023**	0.049**
Number of leader deaths	47	42	42	36	36
Years after leader's death	Autocrats				
0	0.019**	0.032**	0.199	0.012**	0.304
1	0.016**	0.049**	0.356	0.011**	0.227
2	0.028**	0.100*	0.314	0.026**	0.183
Number of leader deaths	29	26	26	21	21
Years after leader's death	Democrats				
0	0.460	0.326	0.044**	0.236	0.039**
1	0.552	0.415	0.092*	0.291	0.069*
2	0.432	0.370	0.134	0.359	0.113
Number of leader deaths	22	20	20	19	19

Notes:
This table presents a replication exercise of Jones and Olken (2005) by updating their original PWT 6.1 GDP data with PWT 6.2 data. All specifications replicate results from Tables III and V in Jones and Olken (2005). Specification [1] replicates the main result in Jones and Olken (2005) using their original data. Specification [2] presents the same result using PWT 6.1 after dropping observations to balance the data with available observations in PWT 6.2. Specification [3] presents our main result when we replace the balanced sample using PWT 6.1 with PWT 6.2. Specification [4] replicates [2] when countries with quality grading "D" are dropped. Specification [5] replicates [3] when countries with quality grading "D" are dropped. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 9. Haussman, Pritchett and Rodrik (2005) Replication Results

PWT 6.1 (Haussman et al.)		PWT 6.2		PWT 6.1 (Haussman et al.)		PWT 6.2		PWT 6.1 (Haussman et al.)		PWT 6.2	
ARG	1963	ARG	1963	IDN	1967	IDN	1967	PAK	1979	PAK	1977
ARG	1990	ARG	1990	IDN	1987	IDN	1985	PAN	1959	PAN	1959
AUS	1961	**No break**		IND	1982	**No break**		PAN	1975	PAN	1975
BEL	1959	BEL	1959	IRL	1958	IRL	1958	PER	1959	PER	1959
BRA	1967	BRA	1968	IRL	1985	IRL	1986	**No break**		PHL	1970
No break		BWA	1984	**No break**		IRN	1966	PNG	1987	**No break**	
CAN	1962	CAN	1961	ISR	1957	ISR	1957	POL	1992	POL	1992
CHL	1986	CHL	1986	**No break**		ISR	1989	**No break**		PRT	1959
CHN	1978	CHN	1977	ISR	1967	ISR	1967	PRT	1985	PRT	1985
CHN	1990	CHN	1991	JOR	1973	**No break**		PRY	1974	PRY	1974
CMR	1972	**No break**		JPN	1958	JPN	1958	ROM	1979	ROM	1971
No break		CMR	1978	**No break**		JPN	1984	RWA	1975	RWA	1975
COG	1969	COG	1968	KOR	1962	KOR	1963	SGP	1969	SGP	1967
COG	1978	COG	1976	KOR	1984	KOR	1984	**No break**		SGP	1987
COL	1967	COL	1967	LKA	1979	**No break**		SYR	1969	SYR	1969
DNK	1957	DNK	1957	**No break**		LKA	1958	SYR	1974	**No break**	
DOM	1969	DOM	1969	LSO	1971	LSO	1971	SYR	1989	SYR	1991
DOM	1992	DOM	1991	**No break**		LSO	1992	TCD	1973	**No break**	
DZA	1975	**No break**		MAR	1958	MAR	1958	THA	1957	THA	1957
ECU	1970	ECU	1970	MLI	1972	MLI	1973	THA	1986	THA	1986
EGY	1976	EGY	1975	MUS	1971	MUS	1970	TTO	1975	**No break**	
ESP	1959	ESP	1959	MUS	1983	MUS	1983	TUN	1968	TUN	1968
ESP	1984	ESP	1984	MWI	1970	MWI	1970	TWN	1961	TWN	1961
No break		ETH	1988	MWI	1992	MWI	1990	**No break**		TWN	1985
FIN	1958	FIN	1958	MYS	1970	MYS	1967	**No break**		TZA	1992
FIN	1967	**No break**		MYS	1988	MYS	1988	UGA	1977	**No break**	
No break		GAB	1969	**No break**		NER	1974	UGA	1989	**No break**	
FIN	1992	FIN	1992	NGA	1957	**No break**		URY	1974	URY	1974
GBR	1982	GBR	1982	NGA	1967	NGA	1968	URY	1989	URY	1988
GHA	1965	GHA	1965	NIC	1960	NIC	1960	USA	1961	USA	1961
GNB	1969	**No break**		NOR	1991	NOR	1992	**No break**		VEN	1971
GNB	1988	**No break**		NZL	1957	NZL	1957	**No break**		ZMB	1963
No break		HND	1974	PAK	1962	PAK	1961	ZWE	1964	ZWE	1967
HTI	1990	**No break**									

Notes:
This table compares growth breaks in per capita GDP (RGDPCH series) obtained from using the original PWT 6.1 in Haussman et al. (2005) with those obtained using PWT 6.2. Growth breaks are reproduced using Haussman et al. (2005) original Gauss code. **No break** indicates an inconsistency between PWT 6.1 and PWT 6.2 in obtaining growth break dates that are more than 3 years apart for each country considered. Haussman et al. (2005) identified an additional break for BWA in 1969 using PWT 6.1, but data for Botswana does not go farther back than 1970 in PWT 6.2. There are 35 cases (out of 100; 35%) where such inconsistencies are detected.

Table 10. Miguel, Satyanath and Sergenti (2004) Replication Results

Dependent variable	→		log (Growth Gap)			
	[1]	[2]	[3]	[4]	[5]	[6]
Estimation	OLS	OLS	OLS	2SLS	2SLS	2SLS
Dataset Substitution	None	None	PWT	None	None	PWT
Sample	Original	Balanced	Balanced	Original	Balanced	Balanced
PWT Grades Included	All	All	All	All	All	All
Economic Growth Rate, t	-0.145	-0.144	-0.327*	-0.383	-0.284	-1.985
	[0.767]	[0.727]	[1.668]	[0.276]	[0.217]	[1.351]
Economic Growth Rate, t-1	0.071	0.079	0.006	-2.139**	-2.078**	-2.307**
	[-0.368]	[-0.395]	[-0.029]	[2.078]	[2.096]	[2.327]
N	743	724	724	743	724	724
R-squared	0.52	0.50	0.49	0.39	0.38	0.29

Notes:
This table presents a replication exercise of Miguel, Satyanath and Sergenti (2004) by updating their original PWT 5.6 GDP data with PWT 6.2 data. All specifications replicate Table 4, equations (3-5), in Miguel, Satyanath and Sergenti (2004). For presentation purposes we focus only on the key parameter - current and lagged growth rates - and omit other parameter estimates. Specification [1] replicates the main OLS result in Miguel, Satyanath and Sergenti (2004) using their original data. Specification [2] presents the same result using PWT 5.6 after dropping observations to balance the data with available observations in PWT 6.2. Specification [3] presents our main result when we replace the balanced sample using PWT 5.6 with PWT 6.2. Specifications [4], [5], [6] replicate specifications [1], [2], [3], respectively, using 2SLS, using rainfall as an instrument for current and lagged economic growth. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. t-statistics in brackets.

Table 11. Aghion, Howitt and Mayer-Foulkes (2005) Replication Results

Dependent variable →	log (Growth Gap)				
	[1]	[2]	[3]	[4]	[5]
Dataset Substitution	None	None	PWT6.2	None	PWT6.2
Sample	Original	Balanced	Balanced	Balanced	Balanced
PWT Grades Included	All	All	All	A,B,C	A,B,C
Financial Development	-0.015	-0.015	0.011	-0.009	0.013
	[0.93]	[0.94]	[0.93]	[0.60]	[1.21]
Financial Development	-0.061***	-0.061***	-0.041***	-0.048***	-0.031**
* Initial GDP Gap (1960)	[5.35]	[4.84]	[4.16]	[3.87]	[3.25]
Initial GDP Gap (1960)	1.507***	1.505***	0.402	1.031*	0.090
	[3.14]	[2.83]	[1.02]	[1.95]	[0.25]
N	71	60	60	57	57
R-squared	0.51	0.34	0.37	0.31	0.37
Implied Convergence Threshold	24.70	24.83	9.76	21.61	2.92
Number of Countries Above Threshold	37 out of 71	37 out of 71	65 out of 71	48 out of 71	71 out of 71

Instruments: legal origins, legal origins * initial GDP gap
Conditioning set: EMPTY

Notes:
This table presents a replication exercise of Aghion, Howitt, and Mayer-Foulkes (2005) by updating their original GDP data obtained from Levine, Loayza and Beck (2000) dataset with PWT 6.2 data. All specifications replicate Table 1, equation (1), in Aghion, Howitt, and Mayer-Foulkes (2005). Specification [1] replicates the main result in Aghion, Howitt, and Mayer-Foulkes (2005) using their original data. Specification [2] presents the same result using Levine, Loayza and Beck's data after dropping observations to balance the data with available observations in PWT 6.2. Specification [3] presents our new result when we replace the balanced sample using Levine, Loayza and Beck's data with PWT 6.2 data. Specification [4] replicates [2] when countries with quality grading "D" are dropped. Specification [5] replicates [3] when countries with quality grading "D" are dropped. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. t-statistics in brackets.

Table 12. Mankiw, Romer, and Weil (1992)

Dependent variable →	**log of GDP Per Capita in 1985**				
	[1]	[2]	[3]	[4]	[5]
Dataset Substitution	None	None	PWT6.2	None	PWT6.2
MRW Sample	Non-Oil	Non-Oil	Non-Oil	Non-Oil	Non-Oil
Sample	Original	Balanced	Balanced	Balanced	Balanced
PWT Grades Included	All	All	All	A,B,C	A,B,C
log(I/GDP)-log(n+g+d)	0.738***	0.713***	0.431***	0.754***	0.430***
	[5.96]	[5.21]	[3.22]	[5.28]	[2.99]
log(school) –log(n+g+d)	0.657***	0.622***	0.711***	0.639***	0.761***
	[9.07]	[7.92]	[6.98]	[7.59]	[6.86]
N	98	83	83	77	77
R-squared	0.78	0.79	0.73	0.78	0.71
Solow Restriction p-value	0.654	0.653	0.210	0.584	0.166

Notes:
This table presents a replication exercise of Mankiw, Romer, and Weil (1992) by updating their original PWT 4.0 GDP and I/GDP data with PWT 6.2 data. All specifications replicate the Table 2 restricted regressions in Mankiw, Romer, and Weil (1992). The "Solow Restriction" is the restriction that the parameters for log(I/GDP)-log(n+g+d) and log(school)- log(n+g+d) are of equal magnitude. Specification [1] replicates the main result in Mankiw, Romer, and Weil (1995) using their original data. Specification [2] presents the same result using PWT 4.0 for the sample of non-oil countries after dropping observations to balance the data with available observations in PWT 6.2. Specification [3] presents our main result when we replace the balanced sample using PWT 4.0 with PWT 6.2. Specification [4] replicates [2] when countries with quality grading "D" are dropped. Specification [5] replicates [3] when countries with quality grading "D" are dropped. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. t-statistics are in brackets.

Table 13. Replication Exercise Results Summary

Author	Journal	Year	Tables Replicated	Replication Exercise Results Summary
Results not robust to dataset changes:				
Aghion et. al.	QJE	2005	Table 1, Cols 1-4	Initial GDP gap becomes insignificant in many exercises. Implied convergence thresholds become implausibly low.
Hausmann et. al	JEG	2005	Table 1	Discrepancies identified in 35 out of 100 growth acceleration cases between PWT 6.2 and PWT 6.1
Jones and Olken	QJE	2005	Tables 3, 5	Many differences across exercises. Most pronounced are changes in significance of Democrats' and Autocrats' deaths.
Ramey and Ramey	AER	1995	Table 1	Standard deviation of growth rates coefficient becomes weakly significant in one exercise and insignificant in three.
Results robust to dataset changes:				
Acemoglu, Johnson, et. al	JME	2003	Table 2, Rows 1-4	Minor changes to parameter magnitudes and significance levels.
Barro	JPE	1999	Table 1, Col 2	Minor changes to parameter magnitudes and significance levels.
Burnside and Dollar	AER	2000	Tables 1-3	Some changes in significance levels, but parameter magnitudes remain largely unchanged.
DeLong and Summers	QJE	1991	Table 1	Some changes in parameter magnitudes, but significance levels remain largely unchanged.
Demirguc-Kunt et. al.	JMCB	2004	Tables 5, 7	Minor changes to parameter magnitudes and significance levels.
Easterly et. al.	JME	1993	Table 5, Cols 2, 4	Minor changes to parameter magnitudes and significance levels.
Mankiw et. al.	QJE	1992	Table 2, Cols 1, 2	I/GDP coefficient halves in magnitude, nearly violating CRS Solow restriction in some exercises. Otherwise estimates are robust.
Miguel et. al	JPE	2004	Table 4, Cols 1-7	Instruments seem to reduce or eliminate parameter estimate differences when using PWT 6.2 data vs. their data.
Sachs and Warner		1995	Table 11, Col 7	Some changes in significance levels, but parameter magnitudes remain largely unchanged.

Table 14. Alternative Estimates of Annual Real Per Capita GDP (PPP) Growth Rate, 1985-1996

COUNTRY	Combining Benchmark Data from PWTs 5.6 and 6.2 [1]	Using Data from PWT 6.2 [2]	Difference [1]-[2]	Absolute value of difference
Zimbabwe	5.15%	0.03%	-5.11%	5.11%
Swaziland	7.89%	2.89%	-5.00%	5.00%
Cameroon	0.06%	-4.54%	-4.60%	4.60%
Grenada	5.25%	0.94%	-4.31%	4.31%
Nigeria	-3.43%	0.85%	4.28%	4.28%
Barbados	4.37%	0.51%	-3.85%	3.85%
St. Lucia	6.60%	2.99%	-3.61%	3.61%
Portugal	6.65%	3.73%	-2.92%	2.92%
Congo, Republic of	-5.19%	-2.44%	2.75%	2.75%
Tanzania	-2.92%	-0.18%	2.74%	2.74%
Benin	-2.81%	-0.23%	2.58%	2.58%
Trinidad &Tobago	-2.37%	0.21%	2.58%	2.58%
Nepal	-0.29%	2.24%	2.52%	2.52%
Philippines	3.74%	1.25%	-2.49%	2.49%
Greece	2.96%	0.66%	-2.30%	2.30%
Hungary	2.13%	0.05%	-2.08%	2.08%
Bangladesh	-0.18%	1.89%	2.07%	2.07%
Sri Lanka	1.58%	3.57%	1.99%	1.99%
Mauritius	6.86%	5.11%	-1.75%	1.75%
Austria	3.70%	1.98%	-1.72%	1.72%
Luxembourg	6.08%	4.41%	-1.67%	1.67%
Tunisia	3.82%	2.20%	-1.63%	1.63%
Morocco	2.82%	1.36%	-1.46%	1.46%
Spain	3.92%	2.59%	-1.33%	1.33%
Senegal	-1.95%	-0.75%	1.20%	1.20%
Jamaica	3.38%	2.24%	-1.14%	1.14%
Botswana	5.42%	4.31%	-1.12%	1.12%
Denmark	2.71%	1.66%	-1.05%	1.05%
Madagascar	-2.43%	-1.49%	0.94%	0.94%
Canada	0.15%	1.09%	0.94%	0.94%
Netherlands	2.98%	2.06%	-0.92%	0.92%
Iran	-0.70%	0.20%	0.90%	0.90%
Norway	3.46%	2.60%	-0.86%	0.86%
Turkey	1.77%	2.64%	0.86%	0.86%

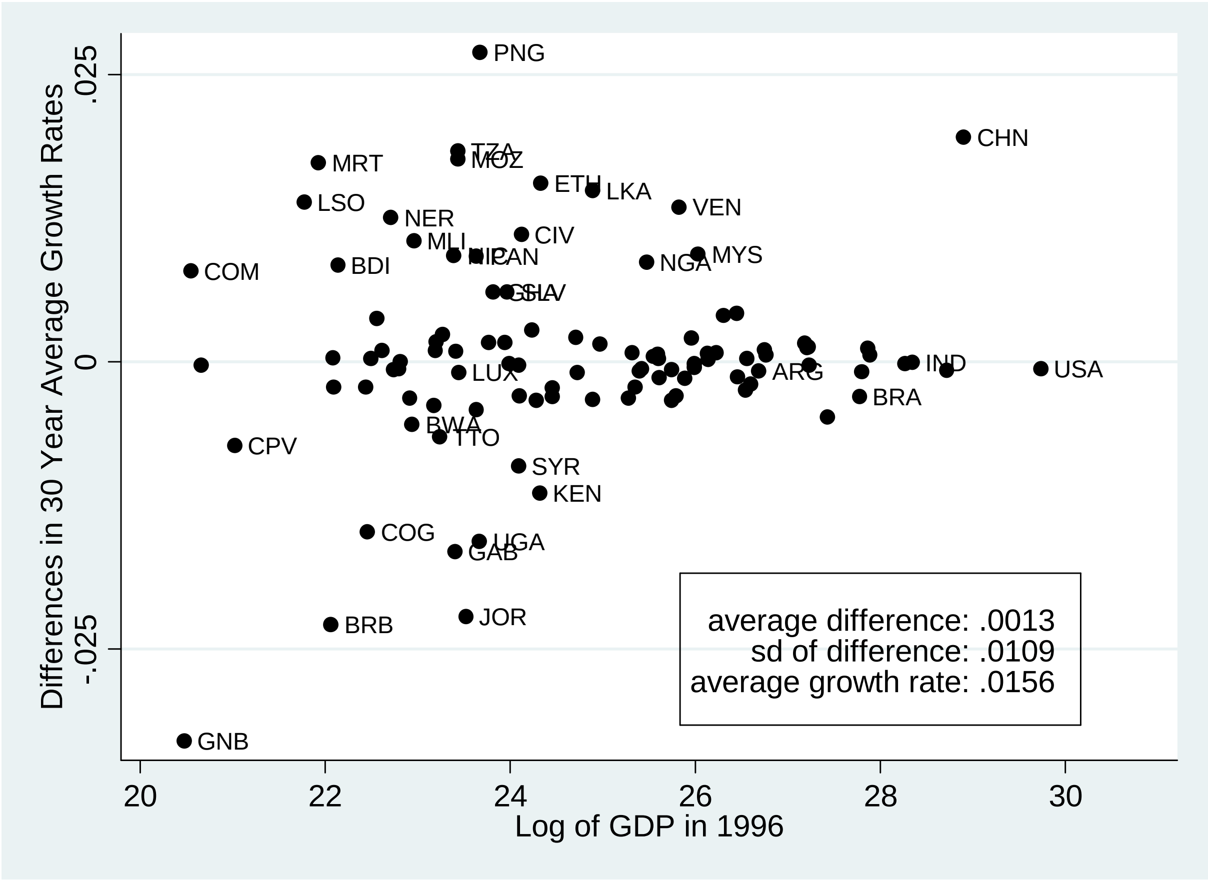
Italy	2.62%	1.80%	-0.82%	0.82%
Belgium	2.81%	2.05%	-0.76%	0.76%
Rwanda	-1.67%	-2.42%	-0.75%	0.75%
Sierra Leone	-2.71%	-3.44%	-0.73%	0.73%
Ethiopia	3.71%	3.05%	-0.66%	0.66%
Suriname	-2.10%	-2.74%	-0.64%	0.64%
Bahamas	-0.58%	-1.17%	-0.58%	0.58%
France	2.11%	1.58%	-0.53%	0.53%
Zambia	-2.99%	-3.48%	-0.49%	0.49%
Thailand	6.16%	6.63%	0.47%	0.47%
Mali	2.37%	2.80%	0.43%	0.43%
United Kingdom	2.73%	2.30%	-0.43%	0.43%
India	3.35%	2.93%	-0.43%	0.43%
Pakistan	2.49%	2.07%	-0.42%	0.42%
Ireland	4.82%	4.41%	-0.40%	0.40%
Poland	1.62%	1.89%	0.27%	0.27%
Japan	3.04%	2.79%	-0.26%	0.26%
Egypt	3.46%	3.22%	-0.24%	0.24%
USA	2.14%	1.91%	-0.22%	0.22%
Australia	1.60%	1.77%	0.17%	0.17%
Korea, Republic of	7.72%	7.60%	-0.12%	0.12%
Kenya	1.20%	1.08%	-0.12%	0.12%
Sweden	1.12%	1.02%	-0.10%	0.10%
Finland	0.64%	0.57%	-0.07%	0.07%
Cote d'Ivoire	-0.65%	-0.66%	-0.01%	0.01%
Malawi	0.78%	0.78%	0.00%	0.00%

Table 15. Countries in the Long-Run Sample

ISO	Country	log GDP	Differences between PWT 6.2 and PWT 6.1					
			Prices	GDP Growth	I/GDP	C/GDP	Benchmarks	PWT Grade
ARG	Argentina	26.69	2%	-0.08%	-9%	-1%	2	B
AUS	Australia	26.77	8%	0.06%	-1%	-5%	3	A
AUT	Austria	25.99	-6%	-0.02%	-7%	-3%	5	A
BDI	Burundi	22.14	5%	0.84%	-30%	2%	0	C
BEL	Belgium	26.14	2%	0.02%	-5%	2%	6	A
BEN	Benin	22.61	8%	0.10%	16%	-9%	2	C
BFA	Burkina Faso	22.92	-2%	-0.32%	3%	13%	0	C
BOL	Bolivia	23.77	4%	0.17%	6%	0%	2	C
BRA	Brazil	27.77	30%	-0.30%	-7%	3%	3	C
BRB	Barbados	22.07	10%	-2.29%	-68%	-15%	2	C
BWA	Botswana	22.94	14%	-0.54%	2%	-7%	3	C
CAN	Canada	27.24	6%	-0.03%	0%	1%	4	A
CHE	Switzerland	25.96	-1%	0.20%	4%	2%	2	A
CHL	Chile	25.74	-5%	-0.34%	11%	-5%	2	B
CHN	China	28.90	25%	1.95%	49%	-16%	0	C
CIV	Cote d'Ivoire	24.12	12%	1.11%	5%	-4%	3	C
CMR	Cameroon	24.09	19%	-0.03%	-33%	-7%	3	C
COG	Congo, Republic of	22.46	-8%	-1.48%	16%	9%	2	C
COL	Colombia	26.13	3%	0.07%	7%	1%	3	C
COM	Comoros	20.55	14%	0.79%	56%	-9%	0	D
CPV	Cape Verde	21.02	-11%	-0.73%	-11%	-7%	0	D
CRI	Costa Rica	23.94	0%	0.17%	-42%	17%	1	C
DNK	Denmark	25.60	4%	0.03%	-6%	-8%	5	A
DOM	Dominican Republic	24.45	-2%	-0.31%	-18%	-3%	1	C
DZA	Algeria	25.80	4%	-0.30%	-9%	-12%	0	D
ECU	Ecuador	24.71	-14%	0.21%	17%	6%	2	C
EGY	Egypt	26.30	4%	0.41%	17%	2%	2	C
ESP	Spain	27.23	1%	0.13%	-7%	-11%	5	B
ETH	Ethiopia	24.33	12%	1.55%	-16%	5%	2	C
FIN	Finland	25.28	4%	-0.32%	10%	-21%	4	A
FJI	Fiji	22.09	3%	-0.22%	-6%	10%	1	C
FRA	France	27.89	0%	0.06%	-7%	-12%	6	A
GAB	Gabon	23.40	4%	-1.65%	-61%	-43%	1	C

NER / GBR	Niger / United Kingdom	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
NGA / GER	Nigeria / Germany	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
NIC / GHA	Nicaragua / Ghana	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	C
NLD / GIN	Netherlands / Guinea	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
NOR / GMB	Norway / Gambia, The	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
NPL / GNB	Nepal / Guinea-Bissau	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
NZL / GNQ	New Zealand / Equatorial Guinea	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
PAK / GRC	Pakistan / Greece	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	4	[illegible]
PAN / GTM	Panama / Guatemala	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	C
PER / HKG	Peru / Hong Kong	[illegible]	[illegible]	[illegible]	[illegible]	-1%	[illegible]	[illegible]
PHL / HND	Philippines / Honduras	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	C
PNG / HUN	Papua New Guinea / Hungary	[illegible]	[illegible]	[illegible]	-38%	[illegible]	[illegible]	[illegible]
PRT / IDN	Portugal / Indonesia	[illegible]	-5%	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
PRY / IND	Paraguay / India	[illegible]	-10%	[illegible]	[illegible]	[illegible]	[illegible]	C
ROM / IRL	Romania / Ireland	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
RWA / IRN	Rwanda / Iran	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	C
SEN / ISL	Senegal / Iceland	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
SLV / ISR	El Salvador / Israel	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
SWE / ITA	Sweden / Italy	[illegible]	1%	[illegible]	[illegible]	[illegible]	[illegible]	A
SYR / JAM	Syria / Jamaica	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	C
TCD / JOR	Chad / Jordan	[illegible]	[illegible]	[illegible]	-8%	[illegible]	[illegible]	[illegible]
TGO / JPN	Togo / Japan	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
THA / KEN	Thailand / Kenya	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	C
TTO / KOR	Trinidad & Tobago / Korea, Republic of	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
TUN / LKA	Tunisia / Sri Lanka	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	C
TUR / LSO	Turkey / Lesotho	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
TZA / LUX	Tanzania / Luxembourg	23.44	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
UGA / MAR	Uganda / Morocco	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
URY / MDG	Uruguay / Madagascar	[illegible]	-3%	[illegible]	[illegible]	[illegible]	3	[illegible]
USA / MEX	USA / Mexico	[illegible]	[illegible]	[illegible]	-2%	[illegible]	[illegible]	[illegible]
VEN / MLI	Venezuela / Mali	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	C
ZAF / MOZ	South Africa / Mozambique	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0	[illegible]
ZMB / MRT	Zambia / Mauritania	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	C
ZWE / MUS	Zimbabwe / Mauritius	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	C
MWI	Malawi	22.80	2%	-0.06%	-23%	7%	4	C
MYS	Malaysia	26.03	2%	0.94%	-2%	-2%	2	C
NAM	Namibia	22.81	-26%	0.00%	-36%	1%	0	D

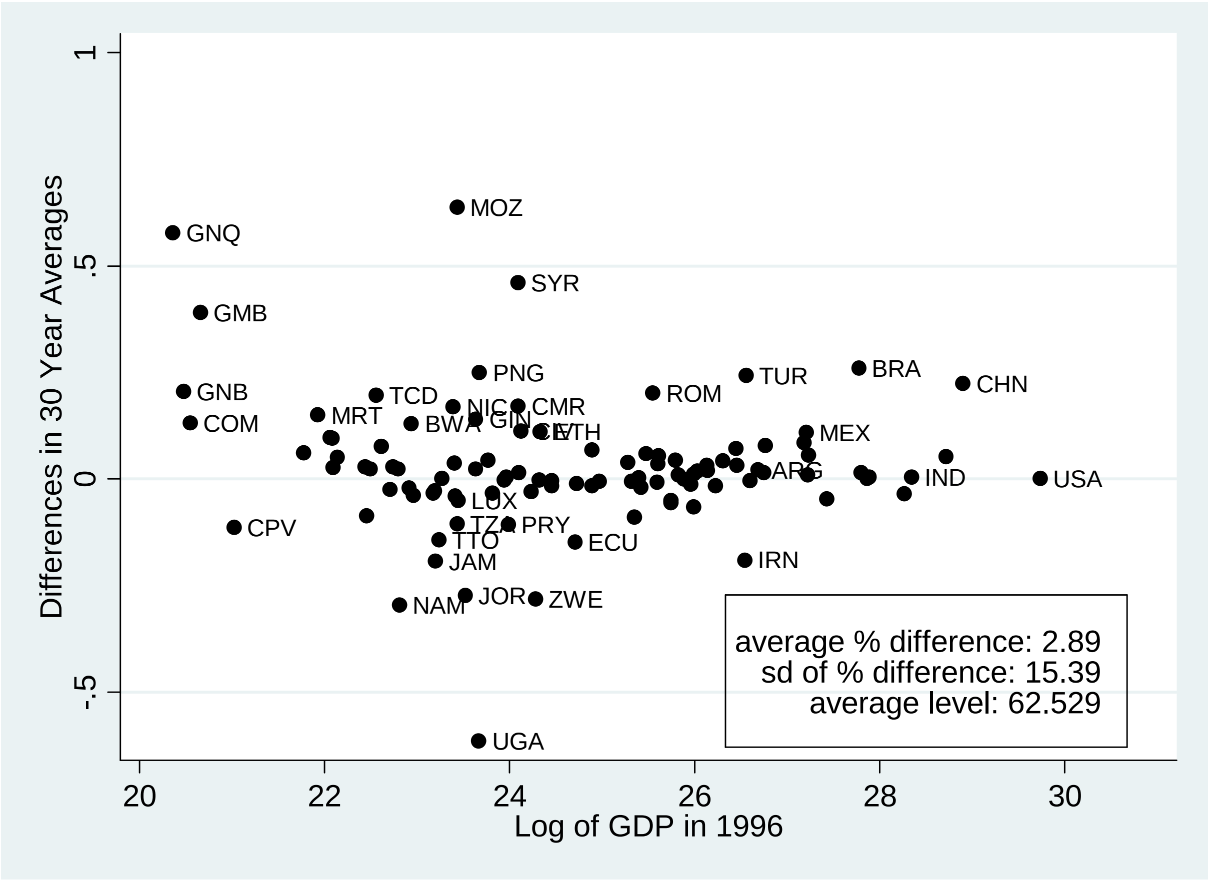
Figure 1: Differences in 30-Year Average Per Capita GDP Growth Rates between PWT 6.2 and PWT 6.1



Notes

1. 30-year average annual per capita GDP growth rates are computed using the RGDPCH series for the period 1970-1999.
2. Differences in growth rates between the two versions of PWT are calculated as GDP growth from PWT 6.2 minus GDP growth from PWT 6.1.
3. Average differences for 30-year average annual per capita GDP growth rates are very close to zero (.0013) whereas the standard deviation is .0109. Average growth rate is about 1.56 percent.
4. The sample consists of the 104 countries in the “Long Run Sample” used in other tables and figures.

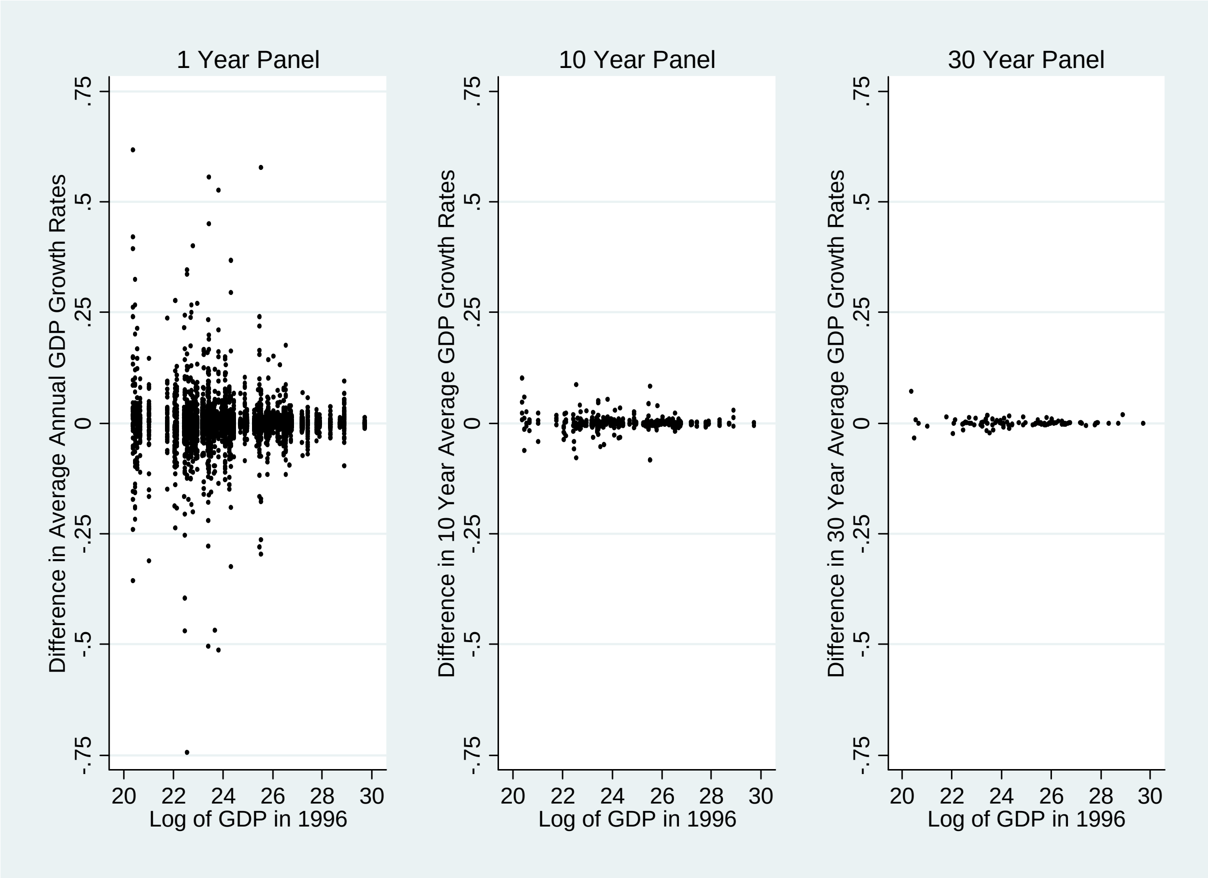
Figure 2: Differences in 30-Year Average Prices between PWT 6.2 and PWT 6.1



Notes

1. 30-year average prices are computed using the average of the P series for the period 1970-1999.
2. Differences in averages between the two versions of PWT are calculated as the log of average prices from PWT 6.2 minus the log of average prices from PWT 6.1.
3. Average differences for 30-year average prices are very close to zero (2.89) whereas the standard deviation is 15.39. The average price level is 62.93, relative to a price level of 100 for the United States.
4. The sample consists of the 104 countries in the “Long Run Sample” used in other tables and figures.

Figure 3: Differences in Annual Per Capita GDP Growth between PWT 6.2 and PWT 6.1



Notes

1. Differences in growth rates between the two versions of PWT are calculated as GDP growth from PWT 6.2 minus GDP growth from PWT 6.1.

2. 1 Yr Panel: # of obs: 3840; avg. difference = .0008; sd of difference = .0539; avg. growth rate = .0186,
10 Yr Panel: # of obs: 384; difference = .0008; sd of difference = .0163; avg. growth rate = .0186,
30 Yr Panel: # of obs: 96; difference = .0010; sd of difference = .0108; avg. growth rate =.0151.

4. The sample of countries is constant across the three figures above, and consists of 96 of the 104 countries in the "Long Run Sample" used in other tables and figures. 8 countries were dropped because they did not have data for all four decades. This sample is called the "Decades Sample."

Figure 4: GDP Per Capita ($2000): PWT 6.2, PWT 6.1, and PWT 5.6

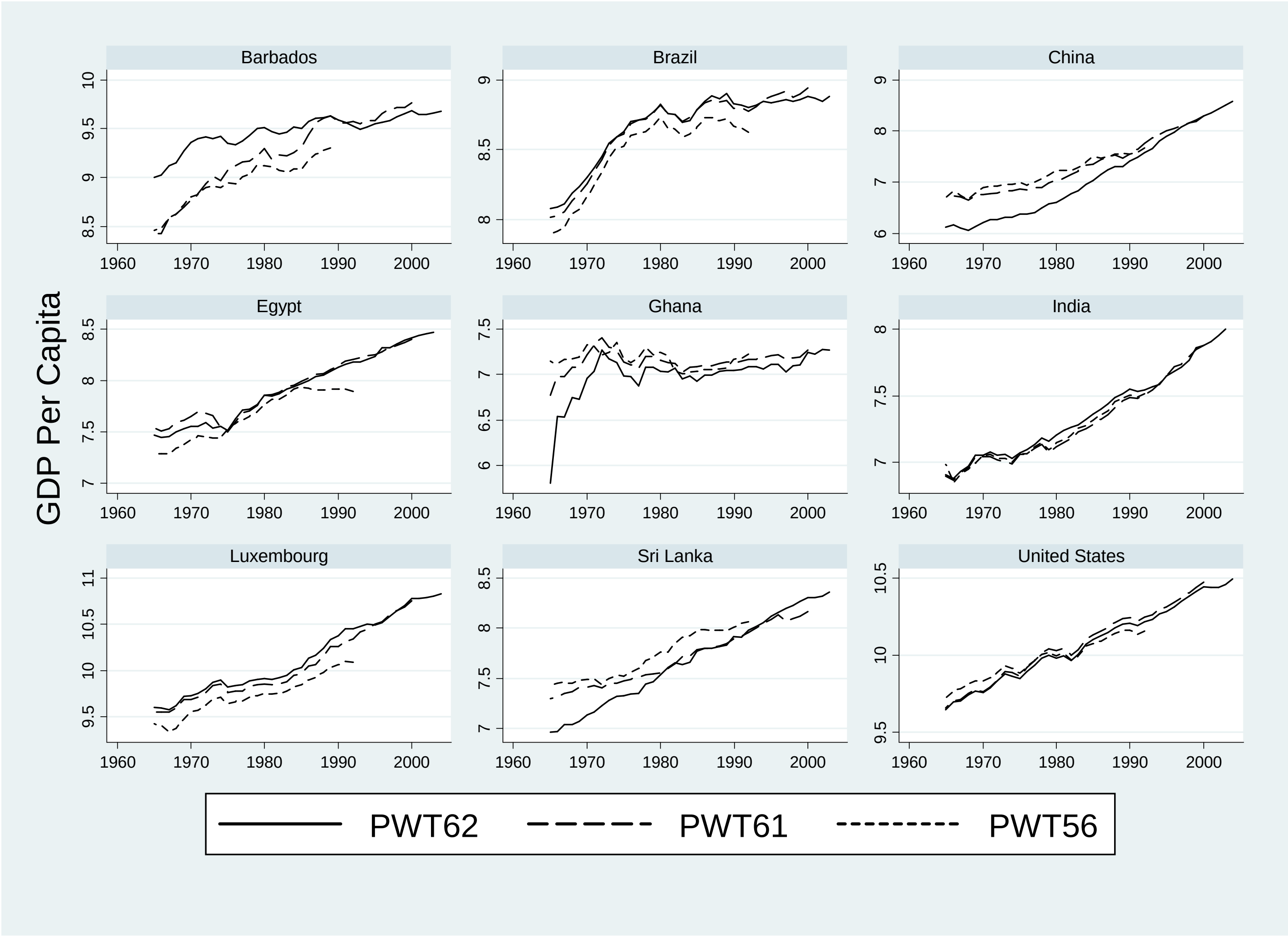
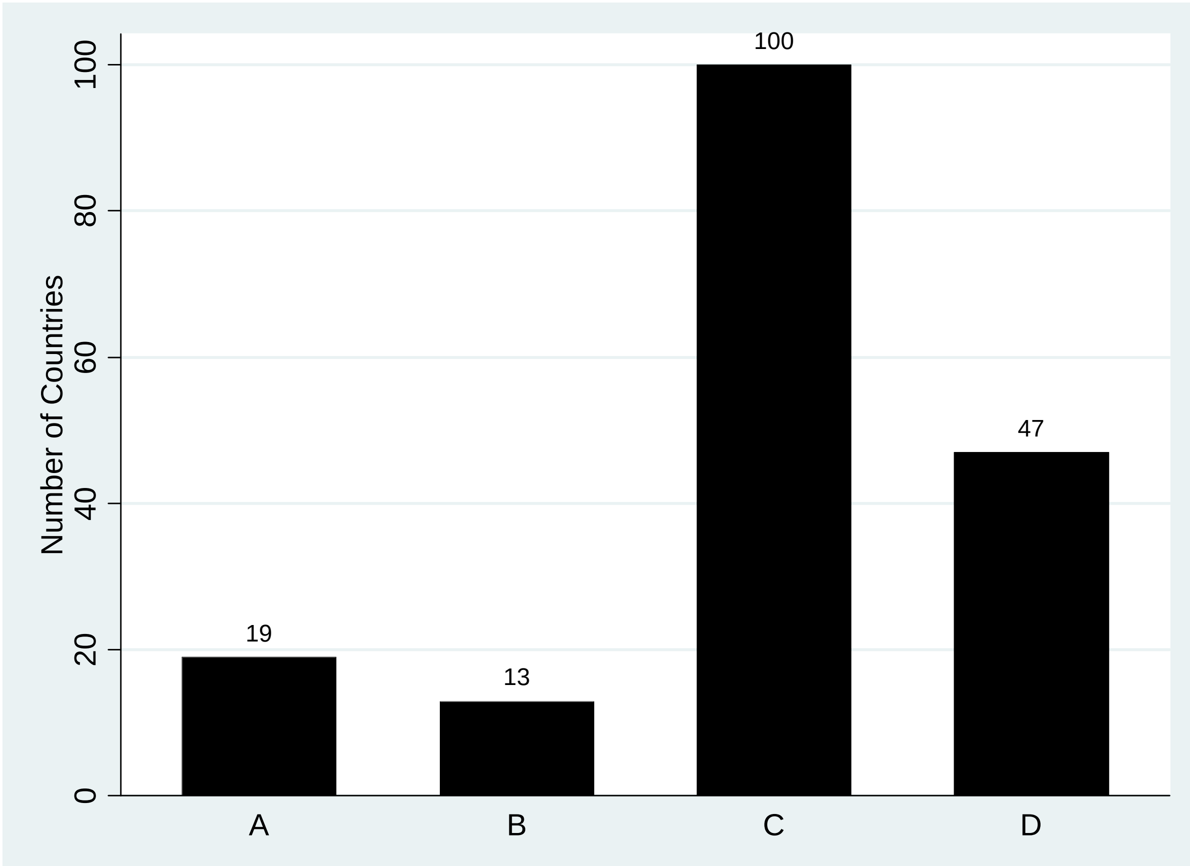


Figure 5: Countries with Respective PWT Data Quality Grades



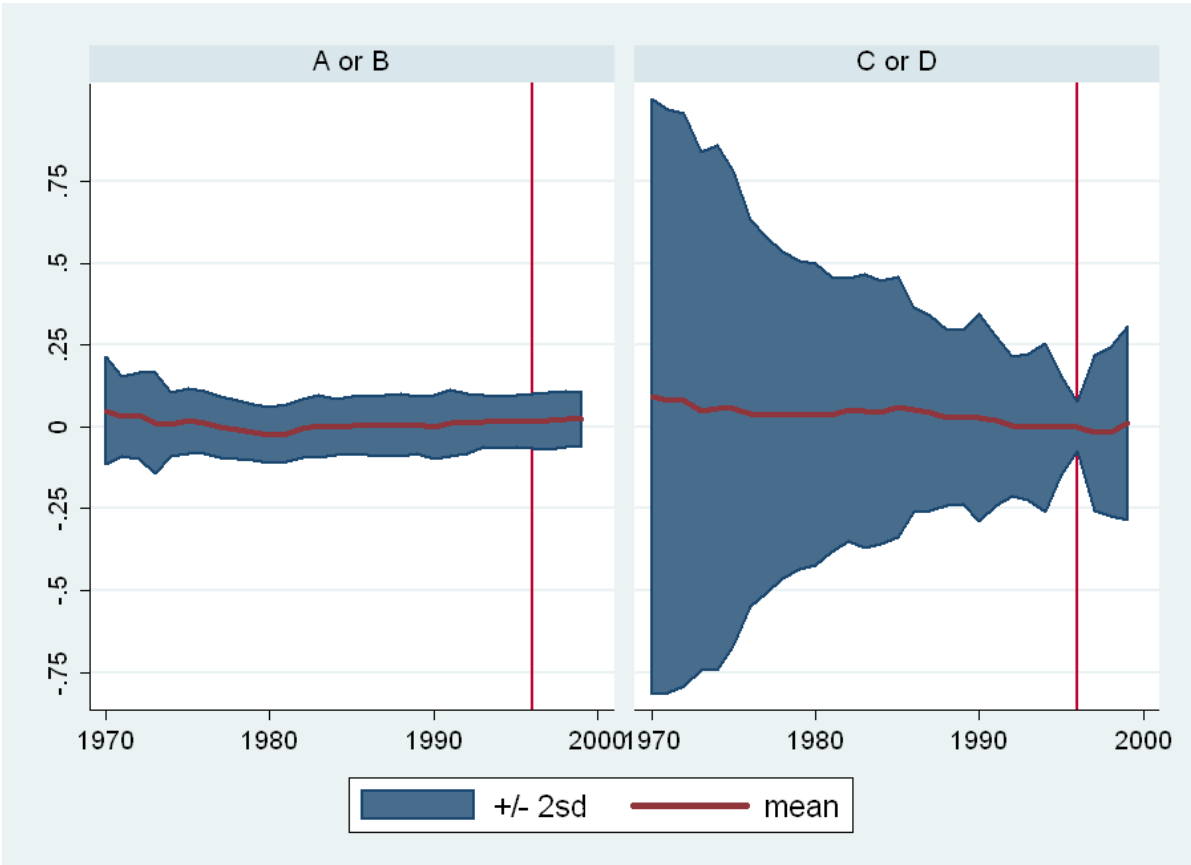
Notes
1. ICP benchmark studies and quality grades used in PWT 6.1.

Figure 6: Differences in 30-Year Average GDP Growth by PWT Quality Grade between PWT 6.2 and PWT 6.1

Notes

1. 30-year average annual per capita GDP growth rates are computed using the RGDPCH series for the period 1970-1999.
2. Differences in growth rates between the two versions of PWT are calculated as GDP growth from PWT 6.2 minus GDP growth from PWT 6.1.
3. The sample consists of the 104 countries in the "Long Run Sample" used in other tables and figures.
4. A or B: average difference: -.0002; sd of difference: .0015; average growth rate: .0233; countries: 29.
   C or D: average difference: .0019; sd of difference: .0127; average growth rate: .0126; countries: 75.

Figure 7: Evolution of Differences in Levels of Prices Across Time between PWT 6.2 and PWT 6.1

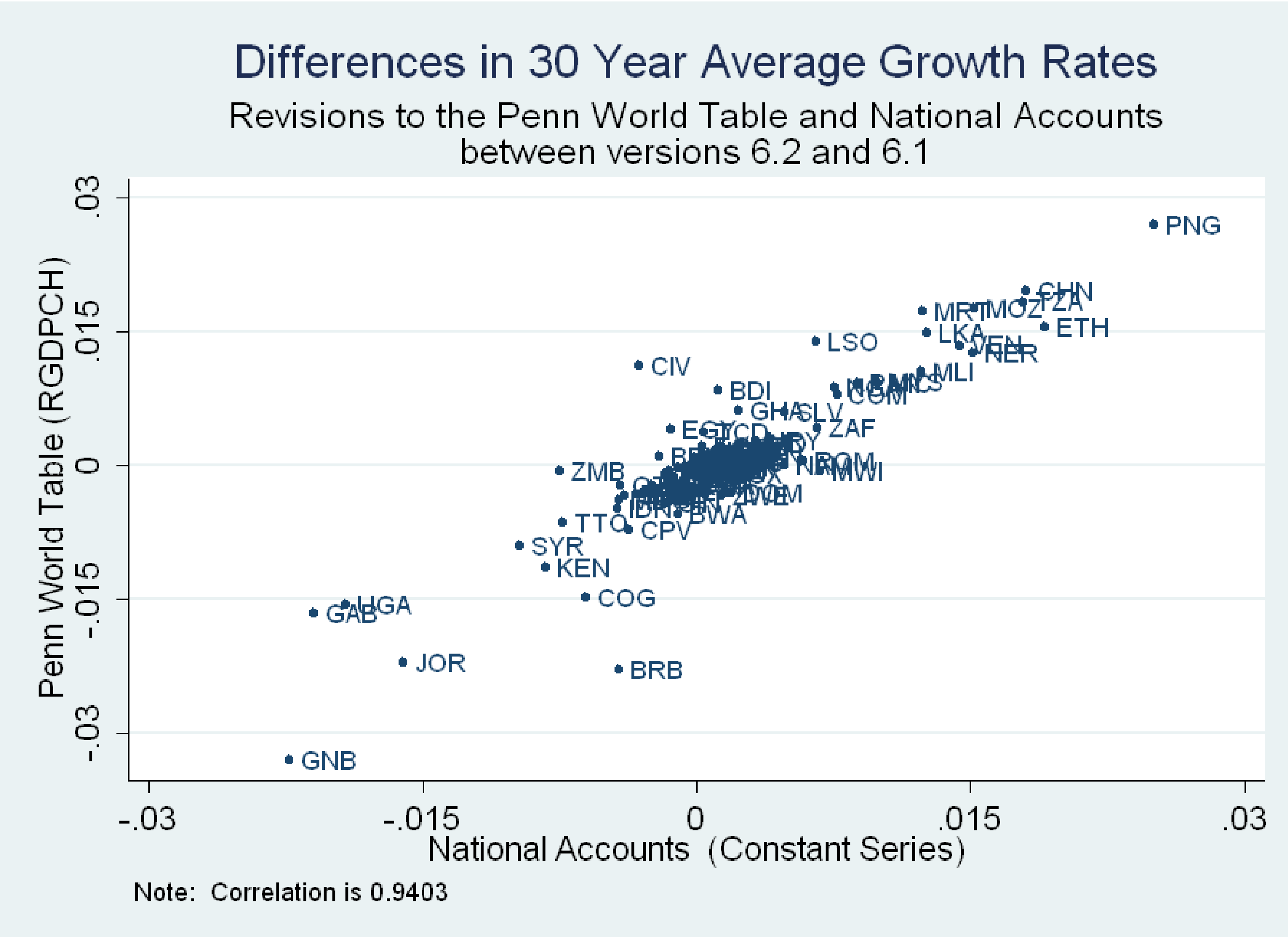


Notes
1. Prices is the Price (P) series.
2. Differences in levels rates between the two versions of PWT are calculated as the log-level of P from PWT 6.2 minus the log-level of P from PWT 6.1.
3. Each mean and standard deviation is computed across countries for a given year. Sample of countries for each year includes countries for which there is data for every year between 1970 and 1999.
4. Vertical line denotes benchmark year, 1996.
5. The sample consists of the 104 countries in the "Long Run Sample" used in other tables and figures.

Figure 8: Evolution of Differences in Levels of Per Capita GDP Across Time between PWT 6.2 and PWT 6.1

Notes
1. GDP is Real GDP per capita, chained series (RGDPCH).
2. Differences in levels between the two versions of PWT are calculated as the log of GDP per capita from PWT 6.2 minus the log of GDP per capita from PWT 6.1.
3.Each mean and standard deviation is computed across countries for a given year. Sample of countries for each year includes countries for which there is data for every year between 1970 and 1999.
4.Vertical line denotes benchmark year, 1996.
5. The sample consists of the 104 countries in the “Long Run Sample” used in other tables and figures.

Figure 9: Correlation Between Revisions to PWT Growth Rates and National Income Accounts Growth Rates (30 Year Averages)

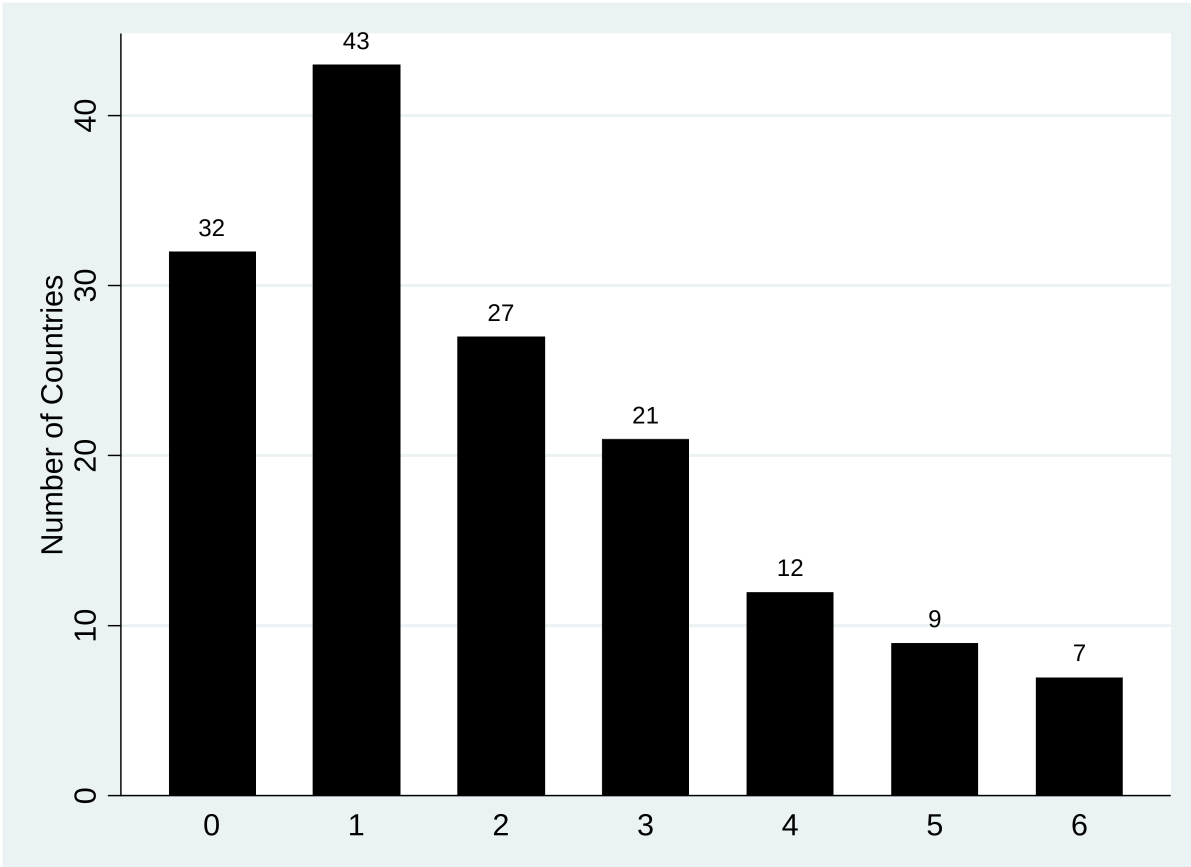


Notes

1. GDP from PWT is Real GDP per capita, chained series (RGDPCH).

2. GDP from Nationa Accounts, Constant Series as modified by PWT.

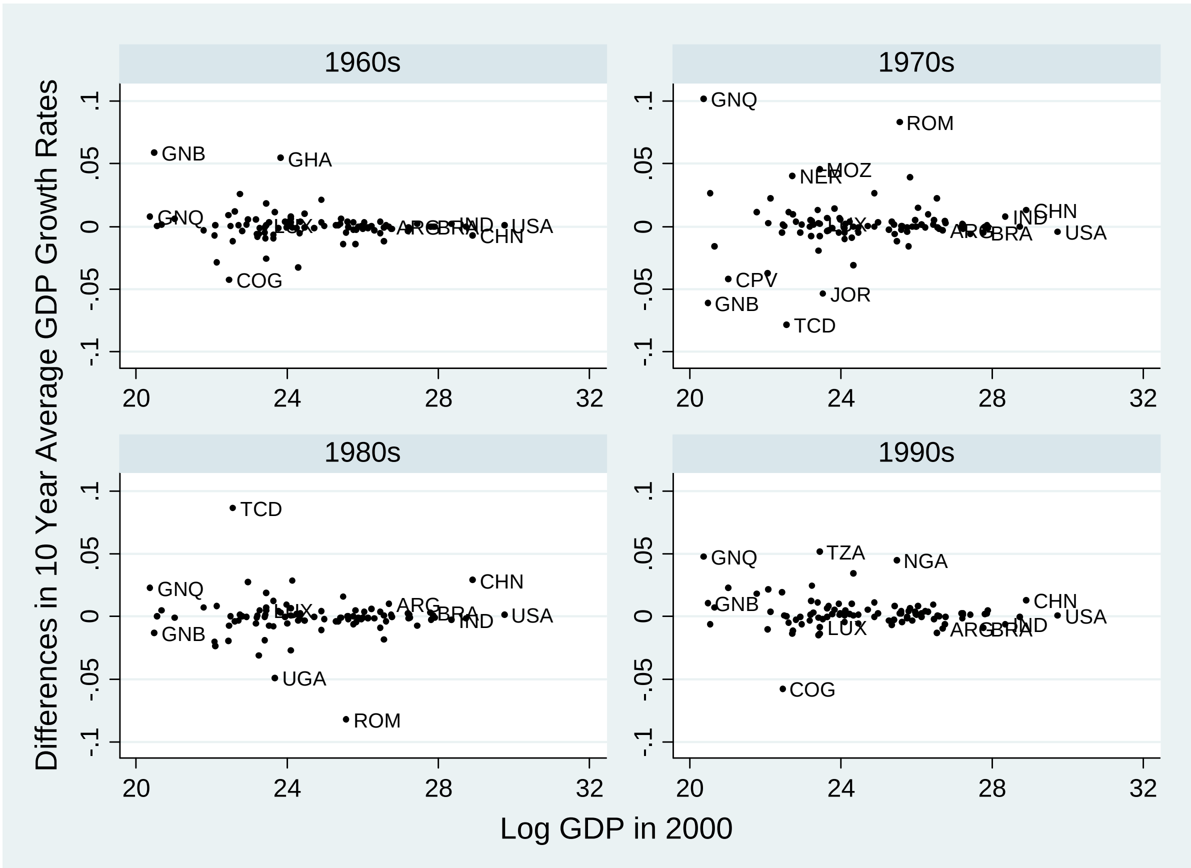
Appendix Figure 1: Countries with Respective Number of ICP Benchmark Studies



Notes

1. ICP benchmark studies used in PWT 6.1.

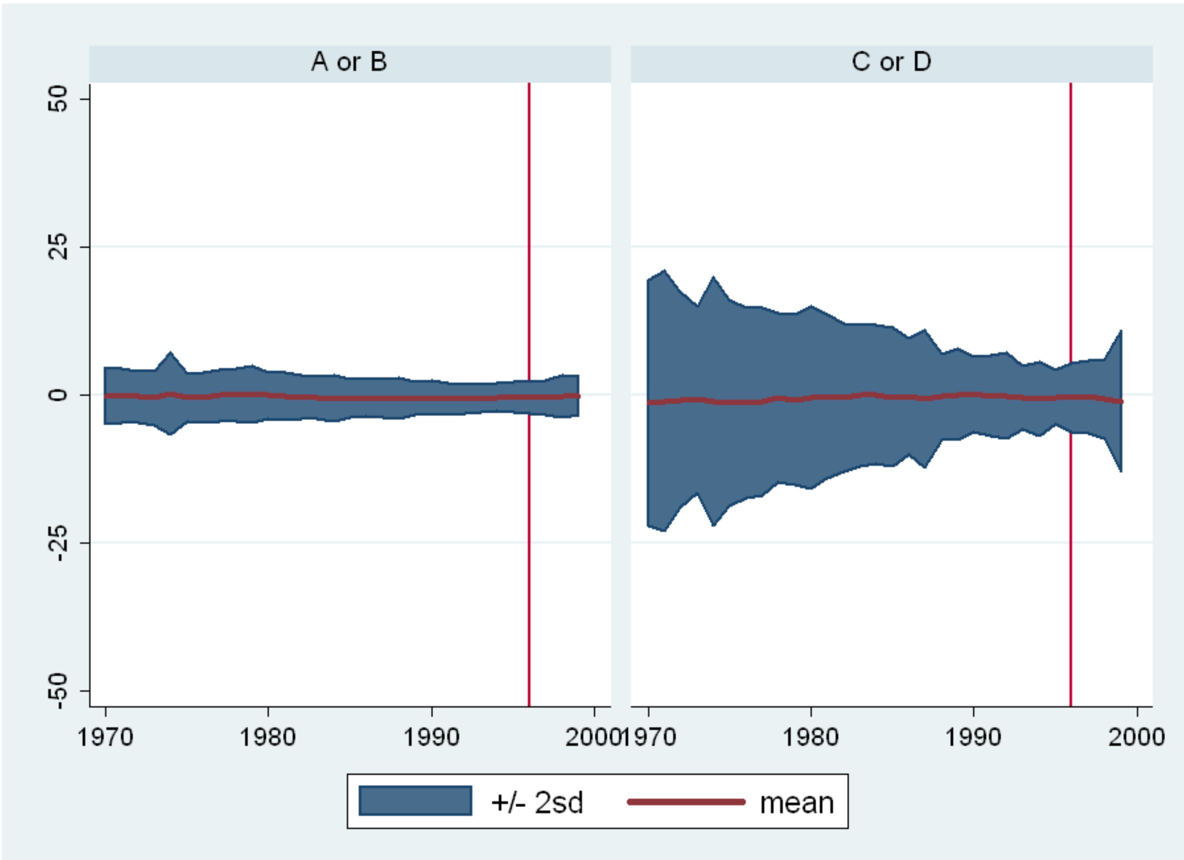
Appendix Figure 2: 10-Year Average Per Capita GDP Growth by Decade between PWT 6.2 and PWT 6.1



Notes

1. 10-year average annual per capita GDP growth rates are computed using the RGDPCH series for the period 1960-2000. Average annual growth rates are calculated as $[log(RGDPCH_{t+10})- log(RGDPCH_t)]/10$.
2. Differences in growth rates between the two versions of PWT are calculated as GDP growth from PWT 6.2 minus GDP growth from PWT 6.1.
3. 1960s: avg. difference = .0001; sd of difference = .0123; avg. growth rate = .0279,
   1970s: avg. difference = .0012; sd of difference = .0217; avg. growth rate = .0212,
   1980s: avg. difference = -.0008; sd of difference = .0163; avg. growth rate = .0084,
   1990s: avg. difference = .0028; sd of difference = .0130; avg. growth rate = .0168.
4. The sample of countries is constant across decades, and consists of 96 of the 104 countries in the "Long Run Sample" used in other tables and figures. 8 countries were dropped because they did not have data for all four decades. This sample is called the "Decades Sample."

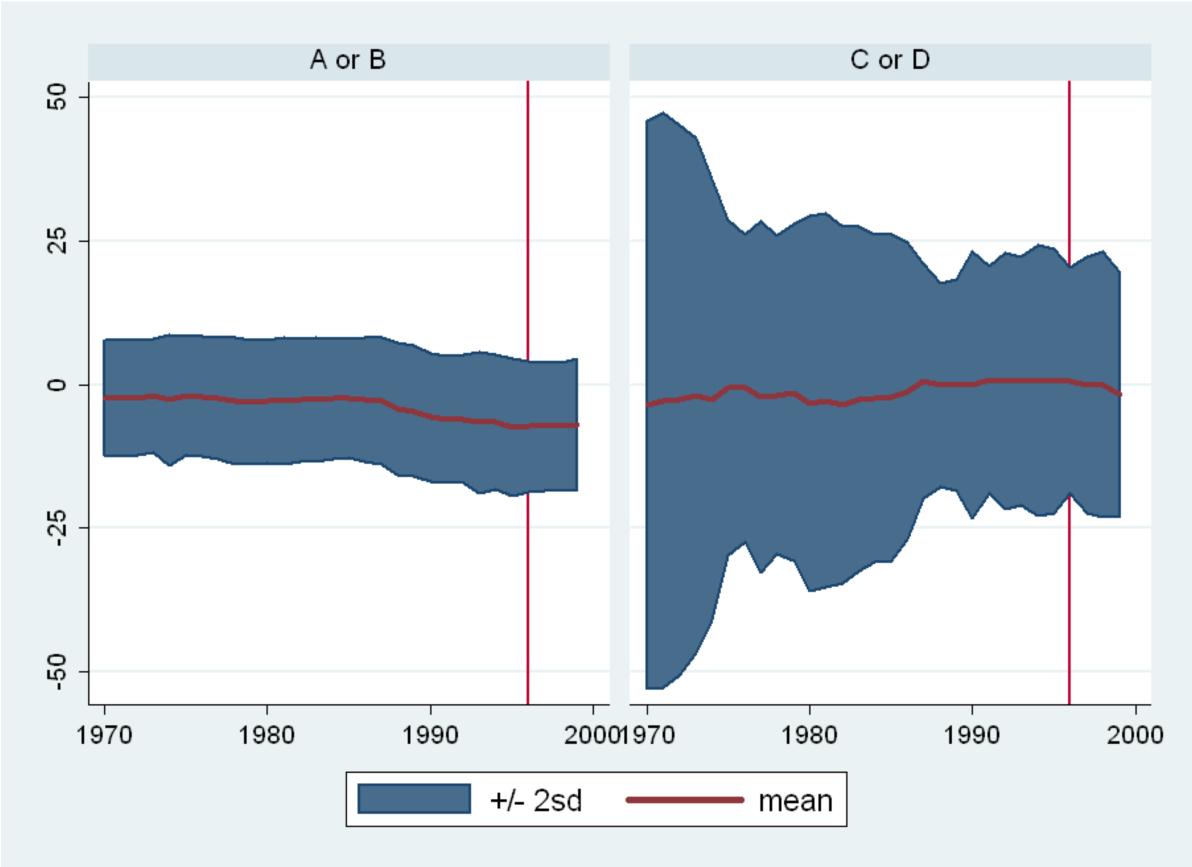
Appendix Figure 3: The Evolution of Differences in Levels of Investment Across Time between PWT 6.2 and PWT 6.1



Notes

1. Investment is the Investment/GDP (KI) series.
2. Differences in levels between the two versions of PWT are calculated as the log-level of investment from PWT 6.2 minus the log-level of investment from PWT 6.1.
3. Each mean and standard deviation is computed across countries for a given year. Sample of countries for each year includes countries for which there is data for every year between 1970 and 1999.
4. Vertical line denotes benchmark year, 1996.
5. The sample consists of the 104 countries in the “Long Run Sample” used in other tables and figures.

Appendix Figure 4: The Evolution of Differences in Levels of Consumption Across Time between PWT 6.2 and PWT 6.1



Notes

1. Consumption is the Consumption/GDP (KC) series.
2. Differences in levels between the two versions of PWT are calculated as the log-level of consumption from PWT 6.2 minus the log-level of consumption from PWT 6.1.
3. Each mean and standard deviation is computed across countries for a given year. Sample of countries for each year includes countries for which there is data for every year between 1970 and 1999.
4. Vertical line denotes benchmark year, 1996.
5. The sample consists of the 104 countries in the "Long Run Sample" used in other tables and figures.

Appendix Table 1. Explaining Differences in Levels of Investment Across Datasets

Dependent variable →	**Abs(difference in Investment/GDP ratio)**						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Estimation	**OLS**						
Dataset Comparison →	**PWT6.2 vs. PWT6.1**						
Sample	**1970 to 1999**						
PWT Grade (D=1…A=4)	-0.954***						
	[-14.6]						
Total ICP Studies		-0.308***					0.051
		[-7.98]					[0.84]
Log(GDP)			-0.443***				-0.323***
			[-10.7]				[-5.02]
Abs(Difference in NA GDP)				-0.000701**			0.000596*
				[-2.02]			[1.76]
Abs(Difference in NA GDP Growth)					0.206***		0.158***
					[5.33]		[3.94]
Distance from Benchmark						1.185***	1.238***
Year, in Decades (1996)						[9.79]	[6.70]
Distance * Total ICPs							-0.0496
							[-0.85]
Constant	5.028***	3.552***	13.74***	2.859***	2.235***	1.456***	8.886***
	[23.1]	[27.4]	[12.9]	[31.7]	[20.4]	[13.7]	[5.79]
N	3016	3016	3016	3016	3016	3016	3016
R-squared	0.035	0.014	0.037	0.000	0.054	0.040	0.106

Notes:

The dependent variable is the absolute value of the difference in the ratio of Investment/GDP across PWT 6.2 and PWT 6.1. The variable PWT grade can take on four values, 1-4, with 1 representing the worst PWT rating of D, and 4 representing the best PWT rating of A. Total ICP studies refers to the number of ICP studies a country has participated in. GDP is total GDP and is measured at purchasing power parity. The distance variable is calculated as the absolute difference between the year of the observation and 1996 (i.e., abs(t-1996)), and is divided by 10 to express the variable in decades. 1996 is the "Benchmark Year" used in constructing both PWT 6.2 and PWT 6.1. The sample consists of the 104 countries in the “Long Run Sample” used in other tables and figures. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Robust t-statistics in brackets.

Appendix Table 2. Explaining Differences in Levels of Consumption Across Datasets

Dependent variable →	**Abs(difference in Consumption/GDP ratio)**						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]
Estimation	**OLS**						
Dataset Comparison →	**PWT6.2 vs. PWT6.1**						
Sample	**1970 to 1999**						
PWT Grade (D=1…A=4)	-2.400***						
	[-10.3]						
Total ICP Studies		-1.054***					1.019***
		[-9.31]					[4.92]
Log(GDP)			-1.601***				-1.386***
			[-11.8]				[-9.90]
Abs(Difference in NA GDP)				-0.00435***			-0.000197
				[-6.60]			[-0.23]
Abs(Difference in NA GDP Growth)					0.492***		0.316***
					[6.30]		[4.33]
Distance from Benchmark Year, in Decades (1996)						1.038***	2.911***
						[3.65]	[4.39]
Distance * Total ICPs							-0.831***
							[-4.62]
Constant	13.38***	10.33***	47.29***	8.105***	6.422***	6.606***	37.62***
	[19.7]	[24.9]	[13.7]	[41.5]	[29.0]	[23.2]	[11.8]
N	3016	3016	3016	3016	3016	3016	3016
R-squared	0.044	0.034	0.098	0.003	0.062	0.006	0.141

Notes:
The dependent variable is the absolute value of the difference in the ratio of Consumption/GDP across PWT 6.2 and PWT 6.1. The variable PWT grade can take on four values, 1-4, with 1 representing the worst PWT rating of D, and 4 representing the best PWT rating of A. Total ICP studies refers to the number of ICP studies a country has participated in. GDP is total GDP and is measured at purchasing power parity. The distance variable is calculated as the absolute difference between the year of the observation and 1996 (i.e., abs(t-1996)), and is divided by 10 to express the variable in decades. 1996 is the "Benchmark Year" used in constructing both PWT 6.2 and PWT 6.1. The sample consists of the 104 countries in the “Long Run Sample” used in other tables and figures. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Robust t-statistics in brackets.