American Economic Review 2015, 105(10): 3150–3182
http://dx.doi.org/10.1257/aer.20130954

The Next Generation of the Penn World Table†

By Robert C. Feenstra, Robert Inklaar, and Marcel P. Timmer*

We describe the theory and practice of real GDP comparisons across countries and over time. Version 8 of the Penn World Table expands on previous versions in three respects. First, in addition to comparisons of living standards using components of real GDP on the expenditure side, we provide a measure of productive capacity, called real GDP on the output side. Second, growth rates are benchmarked to multiple years of cross-country price data so they are less sensitive to new benchmark data. Third, data on capital stocks and productivity are (re)introduced. Applications including the Balassa-Samuelson effect and development accounting are discussed. (*JEL* C43, C82, E01, E23, I31, O47)

For over four decades, the Penn World Table (PWT) has been a standard source of data on real GDP across countries. Making use of prices collected across countries in benchmark years by the International Comparisons Program (ICP), and using these prices to construct purchasing-power-parity (PPP) exchange rates, PWT converts gross domestic product (GDP) at national prices to a common currency—US dollars—making them comparable across countries. Previous versions of PWT, each based on a newer ICP benchmark, were described extensively by their originators (Summers and Heston 1988, 1991; Heston and Summers 1996). From version 8 onward, development has moved to the University of California, Davis and the University of Groningen, while retaining the PWT initials and with continued input from Alan Heston at the University of Pennsylvania.[1] In this paper we describe the main changes to the measurement of real GDP that have been introduced in this "next generation" of PWT.

* Feenstra: Department of Economics, University of California-Davis, 1 Shields Avenue, Davis, CA 95616, and NBER (e-mail: rcfeenstra@ucdavis.edu); Inklaar: Faculty of Economics and Business, University of Groningen, PO Box 800, 9700 AV Groningen, Netherlands (e-mail: R.C.Inklaar@rug.nl); Timmer: Faculty of Economics and Business, University of Groningen, PO Box 800, 9700 AV Groningen, Netherlands (e-mail: M.P.Timmer@rug.nl). This research received support through grants from the National Science Foundation and the Sloan Foundation. We thank Bettina Aten, Andy Atkeson, Susanto Basu, Angus Deaton, Erwin Diewert, Alan Heston, Chad Jones, Prasada Rao, Andrés Rodríguez-Clare, and Carlos Végh for helpful comments as well as other seminar participants at the University of Groningen, University of Pennsylvania, Norwegian School of Economics and Business, and the Economic Growth group of the National Bureau of Economic Research. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.
† Go to http://dx.doi.org/10.1257/aer.20130954 to visit the article page for additional materials and author disclosure statement(s).

[1] PWT version 7 is based on the 2005 ICP prices. PWT version 8.1 is still based on the 2005 benchmark but has new features described in this paper, and is available online at: http://www.rug.nl/research/ggdc/data/pwt/. Version 9 will be based on the new ICP 2011 benchmark that became available in 2014.

Most importantly, we distinguish whether real GDP is intended to measure the standard of living across countries or to measure productive capacity. As argued by Feenstra et al. (2009), real GDP in previous versions of PWT, or its components such as consumption or domestic absorption, was intended to measure the standard of living across countries. They refer to this concept as "real GDP on the expenditure side," or real GDP^e. This variable was close to what is called "command-basis GDP" in the United States. We contrast this concept with "real GDP on the output-side," or real GDP^o, which is intended to measure the productive capacity of an economy. Countries that have strong terms of trade—meaning higher than average prices for exports or lower than average prices for imports—will have higher real GDP^e than real GDP^o as a result. We have incorporated a new dataset of quality-adjusted prices of exports and imports so that both real GDP variables are now reported in PWT8.

Second, to hold prices constant over time, past versions of PWT relied upon real GDP growth from the national accounts for each country. That is, the level of real GDP across countries was constructed for the most recent ICP benchmark and then projected backward and forward in time by using national accounts growth rates for each country. That approach meant that past years of ICP data were discarded. In PWT8 we likewise include a variable that uses real GDP growth from the national accounts, but we further introduce measures of real GDP that correct for changing prices over time and use ICP benchmarks from *multiple years*. All of these measures of real GDP in PWT8 resolve the problem noted by Johnson et al. (2013) that, in past versions, growth rates were dependent on the benchmark year of ICP data used in PWT.

Third, we reintroduce measures of the capital stock across countries based on data of investment by type. They are used in conjunction with measures of human capital to provide, for the first time, measures of total factor productivity across countries. New data on labor income shares in GDP allow factor substitution elasticities to differ across countries and over time. This opens the possibility of analyzing the proximate sources of differences in productivity and living standards across countries.

The paper is organized as follows. In Section I we provide a guided tour to the new PWT, highlighting the main sets of variables, briefly discussing their construction and indicating areas of research where they can be useful. Compared to previous PWT versions, PWT8 allows us to forge a much closer link between the variables in PWT and the theoretical concepts of welfare and production in the literature. In Sections II–IV we describe this theory behind real GDP comparisons. We use a familiar model with traded and nontraded goods, whereby more technologically developed countries have higher prices for nontraded goods: this is the Balassa-Samuelson effect (Balassa 1964, Samuelson 1964), or "Penn effect" (Samuelson 1994). In this context, we argue that it is highly misleading to use a single good—even a traded good—as numeraire to measure "real" GDP. If the law of one price holds, then that approach is equivalent to deflating GDP across countries using their *nominal* exchange rates, and will give a biased measure of the standard of living or productive capacity across countries. Instead, real GDP must be measured by holding the entire vector of prices constant across countries and over time, and we discuss practical ways to achieve that end. In Section II we discuss comparisons of real expenditure, while Section III covers measurement of real output across countries and over time. In Section IV we outline the measurement of total factor productivity.

In Section V we discuss computational issues within PWT8 and show how the concepts discussed in the theoretical sections are empirically implemented. It reviews our approach of dealing with multiple ICP benchmarks and how we incorporate quality-adjusted prices of exports and imports from Feenstra and Romalis (2014), needed to compute real GDP^{o}. Core details on the construction of capital stock and productivity measures are provided. To illustrate potential uses of the new PWT data, three applications are presented in Section VI. We show the differences between real GDP^{e} and real GDP^{o} and explain this gap based on familiar relationships in the literature. We also document how the new measures of factor inputs and productivity can explain more of the cross-country variation in real GDP per capita than standard approaches in the literature. Finally, we show that our use of multiple ICP benchmarks has important implications for estimating the Balassa-Samuelson effect, the positive relationship between a country's relative price level and its income per capita. Section VII concludes and the (online) Appendix contains the proofs of our theorems and further details on the calculation of variables in PWT8.

I. A Guided Tour of PWT8

What is "real" GDP? In macroeconomics, this concept means GDP evaluated at constant prices over time. Likewise, for international comparisons research, real GDP means GDP that is evaluated at constant prices across countries. It is not enough to hold just one price constant across countries (i.e., to have a numeraire such as a traded good), but it is essential to hold *all* prices for goods and services constant across countries when evaluating real GDP. This is the basic approach taken in the PWT since its inception. Up to version 7, PWT used information on relative prices of consumption and investment from ICP that allowed for the measurement of relative standards-of-living across countries. For PWT8, we have developed new data that allow us to also provide measures on relative productive capacity across countries. Combined with new data on capital and labor input, cross-country comparisons of productivity can be made as well. In this section we outline the main variables in PWT and their uses, and provide pointers to more detailed discussions in the remainder of the paper.

An important distinction is between GDP measured from the expenditure side and the production side. Traditionally, PWT measured GDP from the expenditure side, and in earlier versions this was the *only* measure of real GDP. It was constructed as nominal GDP, deflated by the relative price level for domestic absorption.[2] To achieve this, the ICP would collect detailed data on consumer expenditures as well as the prices for those expenditure categories, and by dividing expenditures by prices it obtained the consumption quantities relevant to the standard of living. In conjunction with the prices and quantities of investment goods and government expenditures, also collected by the ICP, real GDP from the expenditure side was

[2] In the US National Income and Product Accounts, a comparable measure is referred to as "command-basis GDP." Command-basis GDP is obtained by deflating nominal GDP by the price index for gross domestic purchases. See equation (16) for the comparable definition of expenditure-side real GDP in PWT.

computed. In PWT8 we refer to this real GDP concept as GDP^e to distinguish it from real GDP measured from the production side. This emphasis is important because PWT8, for the first time, includes output-based real GDP, or real GDP^o.

Output-based real GDP was previously not feasible because its computation requires not only relative prices of consumption and investment, but also of export and imports. Incorporating such data is challenging as there is no cross-country survey that collects prices for traded goods of comparable quality across countries, as the ICP does for consumption and investment products. Instead, we are forced to start with the unit values of traded goods. A recent body of research in international trade shows how to correct these unit values for quality, thereby obtaining quality-adjusted prices across countries, as in Feenstra and Romalis (2014).[3] Dividing export and import values by these prices we obtain quality-adjusted quantities, which are treated as outputs and inputs, respectively, to production and to the construction of output-based real GDP. Real GDP^o can be used to compare the productive capacity across countries in a given year and will typically be different from real GDP^e as countries face differing terms of trade; see Feenstra et al. (2009) and Section V for more.

A second important distinction between various sets of variables in PWT is whether they are constructed holding prices for goods and services constant across countries *as well as over time*, or not. This distinction leads to the following two definitions of real GDP as appear in PWT8: real GDP using prices that are constant across countries but depend on the *current year* (variables $CGDP^e$ and $CGDP^o$); and real GDP using prices that are constant across countries and are also *constant over time* ($RGDP^e$ and $RGDP^o$). We prefix the first concept by *C* because it uses prices in the *current year*: this concept is sometimes called "current-price" real GDP in the literature on international comparisons. It is straightforward to correct this concept for inflation in the United States, but it is not purely real since the vector of (reference) prices at which GDP is evaluated can change over time. Accordingly, the *C* variables are best-suited for comparisons across countries in a particular year. We prefix the second concept by *R* because it also holds prices constant over time and therefore corresponds to what economists normally think of as "real": this concept is sometimes called "constant-price" real GDP. The *R* variables are well-suited for comparisons across countries and over time, e.g., the productive capacity of China's economy today as compared to the US economy at some point in the past. By construction these two sets of variables are equal in the benchmark year 2005 ($RGDP^o = CGDP^o$ and $RGDP^e = CGDP^e$), but otherwise differ because the *C* variables are evaluated at different prices in other years. Sections II, III, and V provide more detailed discussions.

The key variables in PWT8.1 are shown in Table 1, where part A lists the "current-price" or *C* variables and part B lists the "constant-price" or *R* variables.[4] Focusing first on part A, the variable $CGDP^e$ and its components (consumption, investment, and government expenditures) play an important role in measures of

[3] The starting point of this literature is that a good that is imported in high quantity but without having a low unit value must be of high quality: see Khandelwal (2010) and Hallak and Schott (2011). Feenstra and Romalis (2014) extend this demand-side measurement by also building in a supply side, as discussed in Section V.

[4] The variables *CCON*, *CDA*, and *CWTF* were not included in PWT8.0, but are newly added in PWT8.1.

Table 1—Key Variables in PWT Version 8.1 and Their Uses

Acronym	Name	Units	Useful for comparing	See also
Panel A. Based on prices that are constant across countries in a given year				
$CGDP^e$	Expenditure-side real GDP, using prices for final goods that are constant across countries	Millions of 2005 US $	Living standards across countries in each year	Section V
$CGDP^o$	Output-side real GDP, using prices for final goods, exports, and imports that are constant across countries	Millions of 2005 US $	Productive capacity across countries in each year	Section III, V
CCON	Real consumption of households and government, using prices that are constant across countries	Millions of 2005 US $	Living standards across countries in each year	Section II, V
CDA	Real domestic absorption, computed as real consumption (*CCON*) plus real investment	Millions of 2005 US $	Living standards across countries in each year	Section II, V
CK	Capital stock using prices for structures and equipment that are constant across countries	Millions of 2005 US $	Capital stock across countries in each year	Section IV, V
CTFP	TFP level, computed with $CGDP^o$, *CK*, labor input data, and *LABSH*	USA value = 1 in all years	Productivity level across countries in each year	Section IV, V
CWTFP	Welfare-relevant TFP level, computed with *CDA*, *CK*, labor input data, and *LABSH*	USA value = 1 in all years	Living standards across countries in each year	Section IV, V
Panel B. Based on prices that are constant across countries and over time				
$RGDP^e$	Expenditure-side real GDP, using prices for final goods that are constant across countries and over time	Millions of 2005 US $, $RGDP^e = CGDP^e$ in 2005	Living standards across countries and across years	Section III, V
$RGDP^o$	Output-side real GDP, using prices for final goods exports and imports that are constant across countries and over time	Millions of 2005 US $, $RGDP^o = CGDP^o$ in 2005	Productive capacity across countries and across years	Section III, V
Panel C. Based on national prices that are constant over time				
$RGDP^{NA}$	Real GDP at constant national prices, obtained from national accounts data for each country	Millions of 2005 US $, $RGDP^{NA} = CGDP^o$ in 2005	Growth of GDP over time in each country	
$RCON^{NA}$	Real household and government consumption at constant national prices	Millions of 2005 US $, $RCON^{NA} = CCON$ in 2005	Growth of consumption over time in one country	
RDA^{NA}	Real domestic absorption at constant national prices	Millions of 2005 US $, $RDA^{NA} = CDA$ in 2005	Growth of domestic absorption over time in each country	
RK^{NA}	Capital stock at constant national prices, based on investment and prices of structures and equipment	Millions of 2005 US $, $RK^{NA} = CK$ in 2005	Growth of the capital stock over time in each country	Section V
$RTFP^{NA}$	TFP index, computed with $RGDP^{NA}$, RK^{NA}, labor input data, and *LABSH*	2005 value = 1 for all countries	Growth of productivity over time in each country	Section V
$RWTFP^{NA}$	Welfare-relevant TFP index, computed with $RGDP^{NA}$, RK^{NA}, labor input data, and *LABSH*	2005 value = 1 for all countries	Growth of welfare-relevant productivity over time in each country	Section V

(*Continued*)

TABLE 1—KEY VARIABLES IN PWT VERSION 8.1 AND THEIR USES (*Continued*)

Acronym	Name	Units	Useful for comparing	See also
Panel D. Other variables				
PL_CON	Price level of *CCON*, equal to the PPP (ratio of nominal *CON* to *CCON*) divided by the nominal exchange rate	USA value = 1 in 2005	How consumption price levels differ across countries	Section V
PL_DA	Price level of *CDA* and $CGDP^e$, equal to the PPP (ratio of nominal *DA* to *CDA*) divided by the nominal exchange rate	USA value = 1 in 2005	How expenditure price levels differ across countries	Section V
PL_GDP^o	Price level of $CGDP^o$, equal to the PPP (ratio of nominal *GDP* to $CGDP^o$) divided by the nominal exchange rate	USA value = 1 in 2005	How output price levels differ across countries	Section V
LABSH	The share of labor income of employees and self-employed workers in GDP	Fraction of nominal GDP	Total inputs across countries or over time	Section V

comparative living standards. PWT8 provides a number of alternatives. Jones and Klenow (2011) ask by how much consumption of a random person in the United States would have to be adjusted to make this person indifferent between living for a year in the United States or in another country. This involves taking into account differences between the two countries in the real level of consumption, but also in life expectancy, leisure, and income inequality. The relevant building block for such a "consumption-equivalent" welfare measure from PWT8 is real consumption, the sum of real household and government consumption, denoted by *CCON*.[5] Starting with this variable, we can add real investment to obtain *CDA*, and likewise adding the real trade balance we get back to $CGDP^e$ (the details of this calculation are in Section V). From the point of view of the representative consumer, $CGDP^e$ essentially treats the trade balance as an income transfer that is then deflated by the local prices, including prices for nontraded goods. $CGDP^e$ can be viewed as a measure of the standard of living, but extended to incorporate the real trade balance.

The new measure of productive capacity of an economy (variable $CGDP^o$) is particularly relevant in studies that account for the proximate determinants of GDP levels, also known as development accounting, as in Hall and Jones (1999); Caselli (2005); and Hsieh and Klenow (2010). Its construction is discussed in detail in Sections III and V. PWT8 also provides new information on real inputs that enables one to compare total factor productivity (TFP) across countries. Measures of the capital stock are cumulated from series on investment in buildings and different types of machinery and converted with relative prices for structures and equipment that are constant across countries (variable *CK*).[6] New measures of labor input are provided as well, corrected for differences in schooling. In addition, we expand upon the work of Gollin (2002) and estimate the share of labor income in GDP that varies over time and across countries (variable *LABSH*, in part D of Table 1). Combining this with (more standard) measures of human capital, one can compare the *level* of

[5] As also argued in Jones and Klenow (2011), the dividing line between household and government consumption is very country-specific and based on the institutional details of how the education and healthcare systems are organized. A total consumption measure is thus the most relevant.

[6] Some earlier versions of PWT had also included capital stock information, but the current data have been newly developed for PWT8; see Section V and online Appendix C.

productivity across countries at a point in time (variable *CTFP*, with $CTFP = 1$ for the United States). The new data on real inputs are relevant in accounting for productivity differences, as in Caselli (2005), but can also be used in constructing welfare-relevant TFP measures along the lines of Basu et al. (2014). They show that the welfare of a country's infinitely lived representative consumer is summarized, to a first order, by total factor productivity and by the capital stock per capita. To calculate this welfare-relevant TFP, they argue that a measure of real domestic absorption is needed which includes consumption as well as investment. This measure is called *CDA* in PWT8 and the TFP measure based on this is called *CWTFP.* Details are provided in Sections IV, V, and online Appendix C.

In past versions of PWT the growth rate of RGDP was computed solely based on the growth rate of real GDP—or its components—obtained from national accounts (NA) data.[7] In the PWT8 the measures of $RGDP^e$ and $RGDP^o$, listed in part B of Table 1, are based on growth rates that are tied to *multiple* ICP benchmarks and correct for changing prices between these benchmarks. Because we interpolate between multiple ICP benchmarks, there is no guarantee that the growth rate of real GDP so obtained will necessarily be close to the NA growth rate.[8] We now indicate the real series with national-accounts growth rates with the superscript *NA*, so that $RGDP^{NA}$ in PWT8 is based on those growth rates. We normalize it such that $RGDP^{NA} = CGDP^o$ in the benchmark year 2005.[9] In all of our measures of real GDP, the growth rates will not change in between existing benchmark years as new benchmarks become available, unless the underlying nominal GDP data from the national accounts are revised.[10] This "invariance of growth rates between benchmarks" was not previously a feature of PWT—as discussed by Johnson et al. (2013)—which meant that ICP benchmarks often led to considerable changes in real GDP growth rates for all prior years. That deficiency is no longer the case in PWT8.

In addition we provide two new variables also based on national accounts growth rates. To measure capital stocks *over time* we include RK^{NA}, which is also computed based on cumulated investment in structures and equipment, but deflated with national prices that allow for a comparison over time. It is set equal to *CK* in 2005. The corresponding measure of productivity, $RTFP^{NA}$, is computed using the growth rate of real GDP from national-accounts data, $RGDP^{NA}$, in conjunction with the growth rates of RK^{NA} and the labor force, to obtain productivity growth rates for each country. $RTFP^{NA}$ is normalized to 1 in 2005 for all countries; see Section V.

Finally, PWT8 provides various relative price levels, which equal the PPP exchange rate divided by the nominal exchange rate. These variables show how

[7] Up to version 6.1, the variable "rgdpl" in PWT relied upon a weighted average of the NA growth rates of the *components* of GDP, i.e., *C*, *I*, *G*, *X*, and *M*. The weights used depended on the ICP benchmark being used, leading to the criticism of Johnson et al. (2013). Beginning in version 6.2, a second real GDP variable "rgdpl2" was introduced that relied instead on the NA growth rate of total absorption, and therefore was not subject to that criticism.

[8] India, for example, is found to have a higher standard of living in its 1975 ICP benchmark than predicted from the 1985 benchmark and back-casting using the growth of national accounts prices. It follows that the change in real GDP from 1975 onward is correspondingly reduced.

[9] $RGDP^{NA}$ is similar to the series "rgdpl2" that was used in PWT6.2 and v7 except that: (i) "rgdpl2" used the real growth of absorption from the national accounts of each country rather than the real growth of GDP; (ii) "rgdpl2" was normalized to equal expenditure-side $CGDP^e$ in the relevant ICP benchmark year, whereas $RGDP^{NA}$ is normalized to equal the output-side measure $CGDP^o$ in 2005.

[10] These changes can be large. For example, Jerven (2013) discusses Ghana's upward revision of nominal GDP by 60 percent in 2012. More recently, Nigeria announced an upward revision of almost 100 percent.

prices differ across countries when converted at the nominal exchange rate. The ratio of nominal GDP in local currency to $CGDP^o$ equals that country's PPP exchange rate relative to the US dollar (PL_GDP^o). The price levels of *CON* and *DA* in a country are given by *PL_CCON* and *PL_DA*. Price level concepts are discussed in Section V.

To summarize, PWT8 includes a range of measures useful for comparing living standards and productive capacity across countries and over time, including five different measures of real GDP. Many of these measures, with the *C*-prefix, are best-suited when comparing levels across countries in the current year. The variables with the *R*-prefix are best-suited for comparisons over time, though only $RGDP^e$ and $RGDP^o$ are simultaneously suitable for over time and cross-country comparisons. The *CGDP* and *RGDP* series, on both the expenditure and on the output sides, are tied to *multiple* ICP benchmarks whenever price data for a country have been collected multiple times. If the sole object is to compare the growth performance of economies, we would recommend using the $RGDP^{NA}$ series (and this is closest to earlier versions of PWT). In the remainder of this paper, we provide a more detailed discussion of the concepts, definitions, and measurement of the PWT variables.

II. Measurement of Real Expenditure

To illustrate the challenges to constructing "real" GDP, we use a familiar model with traded and nontraded goods. Let $\mathbf{q}_{Nj}$ be a vector of consumption of nontraded goods in country j, with prices $\mathbf{p}_{Nj}$, and $\mathbf{q}_{Tj}$ be a vector of consumption of traded goods in country j, with prices $\mathbf{p}_{Nj}$. We suppose that there is a representative consumer in each country with expenditure function denoted by $E_j(\mathbf{p}_{Nj}, \mathbf{p}_{Tj}, u_j)$, where u_j is utility in country j. Consider a simplified version of this model as discussed in Obstfeld and Rogoff (1996, ch. 4) and Végh (2014, ch. 4 and 6), with a single traded and a single nontraded good. In the monetary version of the model with prices quoted in national currencies (Végh 2014, ch. 6), we might initially assume that the law of one price holds,

$$p_{Tj} = \mathcal{E}_j p_{T0}, \tag{1}$$

where $\mathcal{E}_j$ is the nominal exchange rate in units of country j currency per unit of country 0 currency. Then this model can readily yield the prediction that the relative price of the nontraded good is higher in a country that is more productive in the traded good sector. The reason, of course, is that increased productivity of the traded good leads to higher wages, which in turn increases the relative price of the nontraded good, p_{Nj}/p_{Tj}; this is the celebrated Balassa-Samuelson hypothesis.

The problem that international comparisons seek to solve is how to compare real GDP across countries when their prices differ, as nontraded prices surely do. The "solution" to this problem will depend on what we want real GDP to measure. Throughout this section we maintain that real GDP should measure the standard of living across countries, to be contrasted with real GDP as a measure of productive capacity as outlined in the next section. In order to measure the standard of living—or the cost of obtaining the actual level of utility—it is not enough to just choose a common numeraire: comparing GDP across countries with a common numeraire will give a misleading idea of how the standard of living differs across countries. To show this, let us choose the single traded good as the numeraire and suppose that (1) holds. Then allowing for a vector of nontraded goods, "real" expenditure in each country is measured as

$$\frac{E_j(\mathbf{p}_{Nj}, p_{Tj}, u_j)}{p_{Tj}} = E_j(\mathbf{p}_{Nj}/p_{Tj}, 1, u_j), \tag{2}$$

where the equality follows because the expenditure function is homogeneous of degree 1 in prices. Compare this to *nominal* expenditure measured in terms of the currency of country 0:

$$\frac{E_j(\mathbf{p}_{Nj}, p_{Tj}, u_j)}{\mathcal{E}_j} = \frac{p_{Tj} E_j(\mathbf{p}_{Nj}/p_{Tj}, 1, u_j)}{\mathcal{E}_j} = p_{T0} E_j(\mathbf{p}_{Nj}/p_{Tj}, 1, u_j), \tag{3}$$

where we again make use of homogeneity of degree 1 of the expenditure function, and (1). It is evident that nominal expenditure in a common currency in (3) differs from "real" expenditure in (2) by just the traded good price, p_{T0}. So the ratio of (2) across countries will be identical to the ratio of equation (3). But it is well known that expenditure converted at the nominal exchange rate—which is what we are measuring in equation (3)—gives a highly misleading measure of the standard of living. The reason is that in (3) we are still using the high prices of nontraded goods in more productive countries, leading to higher nominal expenditure and also higher "real" expenditure in (2) when measured in terms of the traded goods price. Conversely, the poor countries will look even poorer when their expenditure is converted to the currency of a rich country, as in (3), if we do not also recognize that their nontraded prices are low. To demonstrate this point in our model, choose country 0 as the United States or a European country with high relative nontraded prices, so that $\mathbf{p}_{Nj}/p_{Tj} < \mathbf{p}_{N0}/p_{T0}$. Then because the expenditure function is increasing in prices it follows that $E_j(\mathbf{p}_{Nj}/p_{Tj}, 1, u) < E_j(\mathbf{p}_{N0}/p_{T0}, 1, u)$, so we obtain

$$\frac{E_j(\mathbf{p}_{Nj}/p_{Tj}, 1, u_j)}{E_0(\mathbf{p}_{N0}/p_{T0}, 1, u_0)} < \frac{E_j(\mathbf{p}_{N0}/p_{T0}, 1, u_j)}{E_0(\mathbf{p}_{N0}/p_{T0}, 1, u_0)} \text{ and} \tag{4}$$

$$\frac{E_j(\mathbf{p}_{Nj}/p_{Tj}, 1, u_j)}{E_0(\mathbf{p}_{N0}/p_{T0}, 1, u_0)} < \frac{E_j(\mathbf{p}_{Nj}/p_{Tj}, 1, u_j)}{E_0(\mathbf{p}_{Nj}/p_{Tj}, 1, u_0)}.$$

The expressions appearing on the right of the inequality signs in (4) both measure the cost of obtaining the utility levels in each country at *common* relative prices $\mathbf{p}_{N0}/p_{T0}$ or $\mathbf{p}_{Nj}/p_{Tj}$. Regardless of which prices are chosen, the relative standard of living on the right of (4) is *higher* than the ratio of "real" or nominal expenditure from (2) or (3), respectively, that appear on the left of (4). This finding demonstrates that low-income countries (with lower relative prices of nontraded goods) will look poorer if we simply convert their expenditures at the nominal exchange rate. To give just one example from PWT8.1, the GDP of China in 2011 when converted at its nominal exchange rate is \$5,439 per capita. That is 11.3 percent of nominal GDP per capita in the United States. We will later measure real GDP per capita in China at 20.5 percent of that in the US in 2011, so that converting at the nominal exchange rate understates its value by nearly one-half.[11] Part of this understatement could come from an undervalued exchange rate, so that the law of one price in (1) does not hold for traded goods, but the deeper problem is that the *nontraded* goods are cheaper in China than in the United States when converted at the official exchange rate.

To resolve this problem and obtain an accurate measure of the standard of living or real GDP, one approach would be to collect the price data across countries and estimate expenditure functions as on the right of the inequality signs in (4). The collection of data for comparable goods across countries is undertaken by the International Comparisons Program (ICP)—a joint project of the United Nations, the World Bank, and other international agencies. But these statistical agencies do not like to rely on econometrically estimated expenditure functions to obtain the standard of living, preferring index-number methods that we discuss below. Of course, researchers can estimate expenditure functions and a leading example is Neary (2004), who estimated an AIDS expenditure function across countries to measure the standard of living. Neary pooled data across countries so that there is a single representative consumer with nonhomothetic tastes. Likewise, we shall drop the country subscript from the expenditure function, and now use $E(\mathbf{p}_{Nj}, \mathbf{p}_{Tj}, u_j)$. Note that if tastes are homothetic then the expenditure function is written as $E(\mathbf{p}_{Nj}, \mathbf{p}_{Tj}, u_j) = e(\mathbf{p}_{Nj}, \mathbf{p}_{Tj})u_j$, in which case the right-hand side of (4) simply becomes the ratio of utilities, u_j/u_0.

Short of estimating the expenditure function, the approach that is taken by statistical agencies and PWT is to evaluate the expenditures that appear on the right of the inequality signs in (4) using the *observed* consumption vectors in each country. Let $\mathbf{q}_j = (\mathbf{q}_{Nj}, \mathbf{q}_{Tj})$ be the vector of consumption goods (traded and nontraded) in country j, with $\mathbf{p}_j = (\mathbf{p}_{Nj}, \mathbf{p}_{Tj})$ denote the country j prices. Then we consider evaluating the two ratios

$$\frac{\mathbf{p}_0'\mathbf{q}_j}{\mathbf{p}_0'\mathbf{q}_0} \quad \text{and} \quad \frac{\mathbf{p}_j'\mathbf{q}_j}{\mathbf{p}_j'\mathbf{q}_0}. \tag{5}$$

Let us return to the case of a single nontraded good and a single traded good. If country 0 is a rich, productive country then it will have a higher relative price of the nontraded good $p_{N0}/p_{T0} > p_{Nj}/p_{Tj}$. With substitution in consumption we would then expect that $q_{N0}/q_{T0} < q_{Nj}/q_{Tj}$. Using these inequalities in (5) and dividing both expressions by (q_{Tj}/q_{T0}), we obtain

$$\frac{\mathbf{p}_0'\mathbf{q}_j/q_{Tj}}{\mathbf{p}_0'\mathbf{q}_0/q_{T0}} = \frac{(p_{N0}/p_{T0})(q_{Nj}/q_{Tj}) + 1}{(p_{N0}/p_{T0})(q_{N0}/q_{T0}) + 1} > \frac{(p_{Nj}/p_{Tj})(q_{Nj}/q_{Tj}) + 1}{(p_{Nj}/p_{Tj})(q_{N0}/q_{T0}) + 1} = \frac{\mathbf{p}_j'\mathbf{q}_j/q_{Tj}}{\mathbf{p}_j'\mathbf{q}_0/q_{T0}}.$$

[11] The difference between nominal and real GDP per capita is even greater for lower income countries: Cambodia, for example, has nominal (real) GDP per capita that is 1.9 percent (5.9 percent) of the United States in 2011.

The inequality above is obtained because the higher relative price $p_{N0}/p_{T0} > p_{Nj}/p_{Tj}$ is applied on the left-hand side to relative quantities $q_{Nj}/q_{Tj} > q_{N0}/q_{T0}$ that are higher in the numerator than in the denominator. In words, this expression says that real consumption in one country relative to another is higher when evaluated at the prices of the *other* country, or to put it most simply, "the grass is greener on the other side." This result shows that the assessing the standard of living by evaluating the consumption quantities at a particular country's prices will be quite sensitive to *which* country's prices are used.

We stress that the above inequality does not depend on having just two goods, and also does not depend on having higher prices of nontraded goods in richer countries, but holds quite generally for any price differences across countries that are consistent with demand-side substitution. Since the country 0 quantity is in the denominator in (5), the first ratio is a Laspeyres quantity index and the second is a Paasche quantity index, and the former exceeds the latter provided that there is negative correlation between the price and quantity differences between countries.[12]

These indexes differ from the ratio of expenditures on the right of (4) because in $E_j(\mathbf{p}_{N0}/p_{T0}, 1, u_j)$, for example, we use country 0 prices but would allow the consumption quantities in country j to be *optimal* at those prices; in contrast, in (5) we hold the consumption quantities fixed at their *observed* levels and are not allowing for substitution in response to prices. Under certain conditions, this limitation can be corrected by taking the geometric mean of the Laspeyres and Paasche indexes in (5), obtaining the Fisher ideal quantity index:

$$(6) \qquad Q_{j0}^F \equiv \left[\left(\frac{\mathbf{p}_0'\mathbf{q}_j}{\mathbf{p}_0'\mathbf{q}_0}\right)\left(\frac{\mathbf{p}_j'\mathbf{q}_j}{\mathbf{p}_j'\mathbf{q}_0}\right)\right]^{\frac{1}{2}}.$$

For a bilateral comparison with only two countries, it is known that if the representative consumer's utility function has a homothetic, quadratic functional form, then the Fisher ideal quantity index in (6) will exactly measure the ratio of utilities u_j/u_0 (Diewert 1976). So in that case, the Fisher ideal quantity index is the "right" way to measure the standard of living across countries. When there are many countries, however, then the comparison is more difficult. Computing (6) for two countries j compared with h, and then again for h compared with k, and multiplying these, we do not necessarily get the same result as directly comparing real expenditure in j with k. To overcome this lack of transitivity, we compare country j with k by indirectly comparing them via all other countries $h = 1, \ldots, C$:

$$(7) \qquad Q_{jk}^{GEKS} = \prod_{h=1}^{C}\left(Q_{jh}^F Q_{hk}^F\right)^{\frac{1}{C}}, \text{ with } Q_{hh}^F \equiv 1.$$

[12] More precisely if the price and quantity differences between countries, weighted by values, are negatively correlated, then the Laspeyres index exceeds the Paasche index. See Balk (2008, p. 64).

This so-called GEKS index is transitive by construction and is an accepted method for making multilateral comparisons.[13, 14]

We have introduced the reader to these index number comparisons of real expenditure because they play a role in PWT8. Specifically, we shall use a two-stage aggregation procedure that first aggregates the prices of items collected by the ICP *within* the categories of consumption C, investment I, and government expenditures G. Prices within these categories are collected by the ICP in each benchmark year and are aggregated using a GEKS approach, i.e., using Fisher-Ideal price indexes that are made transitive across countries using a formula like (7). Besides the desirable property of transitivity, there is a very practical reason for aggregating the categories of C, I, and G: in this way, prices *outside* the benchmark years of the ICP can be interpolated or extrapolated using the time-series data on consumption, investment and government price indexes for each country from their national accounts, as described in Section V.

Having thus obtained a complete time-series and cross-country dataset on the prices of C, I, and G relative to a base country (the US), the second stage is to aggregate to total expenditure. In this second stage we *do not* again use a GEKS procedure to aggregate the prices of C, I, and G in each year, and in this respect we differ from the World Bank who construct the ICP purchasing-power-parity (PPP) price deflators (or real exchange rates) in this way: real GDP is then obtained as nominal GDP divided by the PPPs. As we shall explain in the next section, such an approach severely limits the ability to compare real GDP both across countries and also *over time*. In order to obtain a time-series and cross-section comparison, we believe that it is essential to adopt another approach to the measurement of real GDP, which will involve using *reference prices*.

In general, the reference-price approach to measuring real expenditure means that a vector $\boldsymbol{\pi}$ of reference price is used to evaluate real expenditure across countries as

$$\frac{\boldsymbol{\pi}'\mathbf{q}_j}{\boldsymbol{\pi}'\mathbf{q}_0}.$$

In the specific application to PWT8, we are starting with price indexes and hence relative quantities of C, I, and G obtained from the first-stage GEKS aggregation, so these three components of GDP are multiplied by reference prices and summed in the second stage of aggregation (which is extended to include exports and imports, as discussed below). The question is: what reference prices are used? The most common procedure to use is the *quantity-weighted average over countries of the prices of each good*. This particular choice of reference prices is called the Geary-Khamis (GK) approach.[15] The GK approach satisfies the desirable axiomatic property that

[13] After Gini, Eltetö, Köves, and Szulc. A modern treatment and references are provided by Balk (2008); see also online Appendix B. An alternative approach based on "minimum spanning trees" is presented in Hill (1999). In this method, pairs of countries are compared, either directly or indirectly through a sequence of chained bilateral comparisons involving other countries, with the sequence of countries chosen so that the resulting multilateral indices are least sensitive to the bilateral formula that is used.

[14] Neary (2004) questions whether the GEKS index can accurately reflect the standard of living across countries when preferences are nonhomothetic, however, so this research area is far from resolved. Feenstra, Ma, and Prasada Rao (2009) discuss transitive comparisons with AIDS and nonhomothetic translog expenditure functions.

[15] Due to Geary and Khamis. A modern treatment is provided by Balk (2008) and is described in Section V.

it maintains additivity, so that the components of GDP at reference prices sum to overall real GDP.[16] We now justify the reference-price approach more carefully in the context of measuring real output across countries.

III. Measurement of Real Output across Countries and over Time

GDP measured from the expenditure side (GDP^e) and its components such as consumption and investment play an important role in measures of comparative living standards. We contrast this concept with "real GDP on the output-side," or real GDP^o, which is intended to measure the productive capacity of an economy. In order to measure real output we need to hold the entire vector of prices constant across countries and use those prices to evaluate the *production* quantities rather than the consumption quantities. If there were only final goods, one could simply compute production as the difference between consumption and net exports. However, with intermediate goods, the mapping from consumption to production is not straightforward and one approach would be to calculate the value-added components of consumption categories (Herrendorf, Rogerson, and Valentinyi 2013). The data to do so is not widely available. So we take another, indirect approach of specifying the entire production vector for the economy as $\mathbf{y}_j \equiv (\mathbf{q}_j, \mathbf{x}_j, -\mathbf{m}_j)$, where $\mathbf{q}_j$ is the quantity of final goods as before, $\mathbf{x}_j$ is the quantity of exports and $-\mathbf{m}_j$ is minus the quantity of imports. Domestic prices for the exports and imports are denoted by $\mathbf{p}_j^x$ and $\mathbf{p}_j^m$, and the vector of prices is $\mathbf{P}_j = (\mathbf{p}_j, \mathbf{p}_j^x, \mathbf{p}_j^x)$. We are treating all final goods as nontraded in the sense that some retailed services at least have been added, whereas all imports are intermediate inputs into the production process, possibly only into retailing.

To evaluate output we use the revenue or GDP function for the economy,

$$r_j(\mathbf{P}_j, \mathbf{v}_j) = \max_{\mathbf{q}_j, \mathbf{x}_j, \mathbf{m}_j \geq 0} \left\{ \mathbf{P}_j' \mathbf{y}_j \,|\, F_j(\mathbf{y}_j, \mathbf{v}_j) = 1 \right\}, \tag{8}$$

where $F_j(\mathbf{y}_j, \mathbf{v}_j)$ is a transformation function for each country, which depends on the vector $\mathbf{v}_j$ of primary factor endowments and has an index for country j due to technological differences across countries. Let us denote a vector of reference prices by $\boldsymbol{\Pi} = (\boldsymbol{\pi}, \boldsymbol{\pi}^x, \boldsymbol{\pi}^m)$. Then real output can be compared across countries using the ratio of revenue functions evaluated at these reference prices:

$$\frac{r_j(\boldsymbol{\Pi}, \mathbf{v}_j)}{r_0(\boldsymbol{\Pi}, \mathbf{v}_0)}. \tag{9}$$

One approach to measuring real output would be to estimate the revenue functions in (9). But estimating revenue functions across all countries is even harder than estimating the expenditure function—as Neary (2004) does—because the revenue functions are indexed by country j, indicating technological differences between

[16] This additivity property does not hold, however, when the GEKS approach alone is used to measure real GDP.

them. For this reason, we must rely on indexes that can be used to approximate the ratio of revenue functions in (9).

As in the previous section, the most obvious choice of prices for evaluating the output vectors of two countries is the prices in either country. We have already discussed the inequality that arises from substitution in demand, with the real consumption of one country versus another being higher when evaluated at the other country's prices. The same inequality holds when evaluating the real output of two countries, despite the fact that this comparison is being made using production data rather than consumption data:

$$\left(\frac{\mathbf{P}_j'\,\mathbf{y}_j}{\mathbf{P}_j'\,\mathbf{y}_0}\right) < \left(\frac{\mathbf{P}_0'\mathbf{y}_j}{\mathbf{P}_0'\mathbf{y}_0}\right). \tag{10}$$

This inequality can be interpreted by noting that the right-hand side of (10) is the Laspeyres quantity index, which exceeds the Paasche quantity index on the left due to substitution in demand. According to production theory, however, the inequality should be reversed, since those goods whose prices have raised the most will have the greatest quantity increase. Nevertheless, various studies confirm that the "demand-side bias" in (10) holds in empirical work, and this inequality is known as the Gerschenkron effect. Gerschenkron (1951) was the first to provide evidence that the relative GDP of a country was higher when evaluated at another country's prices. Indeed, for the 146 countries in the 2005 ICP comparison, we find that this inequality holds for more than 98 percent of country pairs.

By taking a geometric mean of the Paasche and Laspeyres indexes, we obtain the Fisher quantity index of real output. The question is how this index-number approach will compare to a reference-price approach as in (9). We can establish a rather tight relationship between these two approaches with the following result, proved in online Appendix A:

THEOREM 1: *Suppose that the outputs are revenue-maximizing and that the inequality in* (10) *holds. Then there exists a reference price vector* $\mathbf{\Pi}$ *between* $\mathbf{P}_j$ *and* $\mathbf{P}_k$ *such that*

$$\frac{r_j(\mathbf{\Pi}, \mathbf{v}_j)}{r_k(\mathbf{\Pi}, \mathbf{v}_k)} = \left[\left(\frac{\mathbf{P}_j'\mathbf{y}_j}{\mathbf{P}_j'\mathbf{y}_k}\right)\left(\frac{\mathbf{P}_k'\mathbf{y}_j}{\mathbf{P}_k'\mathbf{y}_k}\right)\right]^{\frac{1}{2}}.$$

This new result says that computing a Fisher ideal quantity index of production between the countries is a valid comparison of real output between them, in the sense that it is equivalent to using some reference price vector. Remarkably, it does not depend on the functional form of the revenue function but only on optimizing behavior. This theoretical result suggests that there may not be a substantial difference between using the Fisher ideal index of real output—or its generalization, the GEKS approach in (7)—as compared to a reference price approach. We have confirmed that this result holds in PWT8 in a *single* year: whether we are measuring real output or real expenditure, the results from using a GEKS approach do not differ that much from using reference prices constructed as the weighted average of prices across countries.

But this similarity between the index number (GEKS) and reference price (GK) approaches breaks down when we also make comparisons *across time*. In that case we need to recognize that the reference price vector $\mathbf{\Pi}$ established by Theorem 1 is only implicit, and it depends on the level of prices $\mathbf{P}_j$ and $\mathbf{P}_k$. While this enables us to obtain a valid comparison of real output between two countries in each year, we would not be able to compare those real outputs across time because we do not know how the implicit reference price vector is changing over time, and therefore cannot make a "constant price" comparison that we normally expect in "real" variables.

It turns out, however, that we can readily extend Theorem 1 to obtain a consistent comparison of real GDP across countries and simultaneously over time (such variables in PWT8 are indicated by a prefix R). Let the subscript t on all variables indicate time. Suppose that we start in a situation where we have two reference price vectors at two points in time, $\mathbf{\Pi}_\tau = (\boldsymbol{\pi}_\tau, \boldsymbol{\pi}_\tau^x, \boldsymbol{\pi}_\tau^m)$, $\tau = t-1, t$, using the reference prices for all final goods plus exports and imports. In order to also compare real output over time, it would be desirable to use a single vector $\mathbf{\Pi}$ and compute the ratios

$$\frac{r_{jt}(\mathbf{\Pi}, \mathbf{v}_{jt})}{r_{jt-1}(\mathbf{\Pi}, \mathbf{v}_{jt-1})}, j = 1, \cdots, C,$$

for each country. Notice that the endowments in this comparison can change over time, as well as the revenue function itself due to technological change, but the reference prices are held constant.

We can apply Theorem 1 by treating the bilateral comparison there as between country j using reference prices $\mathbf{\Pi}_{t-1}$ and $\mathbf{\Pi}_t$ in the two periods. The optimal outputs at these prices are denoted by $y_{j\tau}^* \equiv \partial r_{j\tau}(\mathbf{\Pi}_\tau, \mathbf{v}_{j\tau}) / \partial \mathbf{\Pi}_\tau$, $\tau = t-1, t$. We assume that the time-series analogue of (10) holds, which states that for country j

$$\left(\frac{\mathbf{\Pi}_t' \mathbf{y}_{jt}^*}{\mathbf{\Pi}_t' \mathbf{y}_{jt-1}^*}\right) < \left(\frac{\mathbf{\Pi}_{t-1}' \mathbf{y}_{jt}^*}{\mathbf{\Pi}_{t-1}' \mathbf{y}_{jt-1}^*}\right). \tag{11}$$

Again, we interpret (11) as stating that the Laspeyres quantity index (on the right) exceeds the Paasche quantity index (on the left). This inequality is another illustration of the Gerschenkron effect.[17] Then an immediate corollary of the earlier theorem is obtained by changing the notation to compare time periods rather than countries, as follows:

COROLLARY 1: *Suppose that the outputs are revenue-maximizing and the Gerschenkron effect in* (11) *holds. Then there exists a reference price vector* $\mathbf{\Pi}$ *between* $\mathbf{\Pi}_{t-1}$ *and* $\mathbf{\Pi}_t$ *such that*

$$\frac{r_{jt}(\mathbf{\Pi}, \mathbf{v}_{jt})}{r_{jt-1}(\mathbf{\Pi}, \mathbf{v}_{jt-1})} = \left[\left(\frac{\mathbf{\Pi}_t' \mathbf{y}_{jt}^*}{\mathbf{\Pi}_t' \mathbf{y}_{jt-1}^*}\right)\left(\frac{\mathbf{\Pi}_{t-1}' \mathbf{y}_{jt}^*}{\mathbf{\Pi}_{t-1}' \mathbf{y}_{jt-1}^*}\right)\right]^{\frac{1}{2}}. \tag{12}$$

[17] Evidence for US exports and imports comes from Alterman, Diewert, and Feenstra (1999). They find that the Laspeyres price or quantity indexes for imported goods over time exceed the Paasche price or quantity indexes, consistent with demand-side substitution in the United States. The same inequality holds for many exported goods, too, which must reflect foreign demand-side substitution rather than US supply-side substitution.

To understand how this result is applied in PWT8, recall that we start with a set of prices for C, I, and G, constructed across countries (relative to a base country, the United States) and over time, constructed from the GEKS method described in (7). To these we add relative prices for exports X and imports M, as described in Section V. That is the first stage of aggregation. In the second stage, we use the GK method to construct reference price for each of C, I, G, X, and M as the weighted average of these prices (relative to the US) across countries: those are the reference prices $\mathbf{\Pi}_t$ in each year. Then the right-hand side of formula (12) can be used to obtain a *constant reference-price growth rate* of real output. In practice, instead of using the optimal quantities as on the right of (12) we instead use observed quantities (see Section V). In this way, we obtain data for real GDP across countries that are consistent with the reference prices established for each year and also correct for *changing* reference prices when making comparisons across time. These variables are denoted in PWT8 by $RGDP^e$ (using only prices for C, I, and G) and $RGDP^o$ (also using prices for X and M). We believe that they offer the best cross-country and time-series comparisons of real GDP. As we mentioned at the end of Section I, however, for research questions that can be answered with the growth rate of real GDP from the national accounts, that growth rate is used to construct RGDP^{NA} and this variable is the closest to real GDP as reported in past versions of PWT.[18]

IV. Total Factor Productivity

Having obtained the comparison of real GDP across countries and over time, we now show how total factor productivity can be computed. We rely heavily on our earlier results and on Caves, Christensen, and Diewert (1982a,b)—henceforth, CCD—and Diewert and Morrison (1986)—henceforth, DM. We drop the time subscript and return to the ratio of revenue functions given in (9), $r_j(\mathbf{\Pi}, \mathbf{v}_j)/r_k(\mathbf{\Pi}, \mathbf{v}_k)$, which measures real output in country j relative to k. Real output can vary due to differing factor endowments, as indicated by v_{lj} and v_{lk} for factors $l = 1, \ldots, L$, or due to differing technologies, as indicated by the country subscript j and k on the revenue function. We can isolate the effect of productivity differences by considering two alternative ratios

$$A_j \equiv \frac{r_j(\mathbf{\Pi}, \mathbf{v}_j)}{r_k(\mathbf{\Pi}, \mathbf{v}_j)}, \quad \text{and} \quad A_k \equiv \frac{r_j(\mathbf{\Pi}, \mathbf{v}_k)}{r_k(\mathbf{\Pi}, \mathbf{v}_k)}.$$

Both of these ratios measure the overall productivity of country j to country k, holding fixed the level of factor endowments. Neither ratio can be measured directly from the data, however, because the numerator or the denominator involves a revenue function that is evaluated with the productivity of one country but the endowments of the other. But the results of CCD and DM tell us that if the revenue function has a translog functional form, then we can precisely measure the geometric mean of these two ratios:

[18] See footnotes 7 and 9.

THEOREM 2: *Assume that the revenue functions* $r_j(\boldsymbol{\Pi}, \mathbf{v}_j)$ *and* $r_k(\boldsymbol{\Pi}, \mathbf{v}_k)$ *are both translog functions that are homogeneous of degree* 1 *in* $\mathbf{v}$ *and have the same second-order parameters on factor endowments, but may have different parameters on prices and on interaction terms due to technological differences between countries. Then the overall productivity of country j relative to k can be measured by*

$$(13) \qquad (A_j A_k)^{\frac{1}{2}} = \frac{r_j(\boldsymbol{\Pi}, \mathbf{v}_j)}{r_k(\boldsymbol{\Pi}, \mathbf{v}_k)} / Q_T(\mathbf{v}_j, \mathbf{v}_k, \mathbf{w}_j^*, \mathbf{w}_k^*),$$

where $Q_T(\mathbf{v}_j, \mathbf{v}_k, \mathbf{w}_j^*, \mathbf{w}_k^*)$ *is the Törnqvist quantity index of factor endowments, defined by*

$$(14) \qquad \ln Q_T(\mathbf{v}_j, \mathbf{v}_k, \mathbf{w}_j^*, \mathbf{w}_k^*) = \sum_{l=1}^{L} \frac{1}{2}\left(\frac{w_{lj}^* v_{lj}}{\sum_m w_{mj}^* v_{mj}} + \frac{w_{lk}^* v_{lk}}{\sum_m w_{mk}^* v_{mk}}\right) \ln\left(\frac{v_{lj}}{v_{lk}}\right),$$

and where $w_{lj}^* = \partial r_j(\boldsymbol{\Pi}, \mathbf{v}_j) / \partial v_{lj}$, $w_{lk}^* = \partial r_k(\boldsymbol{\Pi}, \mathbf{v}_k) / \partial v_{lk}$ *are the factor prices using reference prices* $\boldsymbol{\Pi}$.

CCD establish a result like Theorem 2 using the translog distance and transformation functions, whereas DM establish an analogous result using a time-series rather than cross-country comparison. For completeness, we include a proof in online Appendix A, where we explain that the restriction that the second-order parameters of the factor endowments restricts the technology differences across countries to be of the Harrod-neutral type on factors, or to apply to sectors. The GDP ratio $r_j(\boldsymbol{\Pi}, \mathbf{v}_j) / r_k(\boldsymbol{\Pi}, \mathbf{v}_k)$ in (13) is measured as in Theorem 1, while the Törnqvist quantity index is measured as in (14) but using *observed* factor prices (and therefore observed factor shares) rather than factor prices evaluated at the reference prices, as discussed in the next section.

Theorem 2 tells us that by dividing the observed difference in real GDP^o by the Törnqvist quantity index of factor endowments, we obtain a meaningful measure of the productivity difference between the countries. This result, like the GDP function in (8) and Theorem 1, relies on strict neoclassical assumptions and in particular on perfect competition in product and factor markets. Then with the added assumptions on the translog function described in Theorem 2, the productivity measure in (13)–(14) reflects cross-country differences in aggregate technology.

We recognize that the requirement of perfect competition in product and factor markets, needed for Theorems 1 and 2, is strong. Recent literature has incorporated imperfect competition into the measurement of productivity: e.g., de Loecker (2009) for a single firm or industry, and Basu et al. (2014) for the entire economy. While we expect that our results could be extended to incorporate imperfect competition, such an extension is beyond the scope of the present paper. Burstein and Cravino (2015) relate empirical productivity measures (using procedures of statistical agencies that are similar to ours) to aggregate productivity and welfare changes in international trade models featuring monopolistic competition, and find that those empirical productivity measures are well-grounded. Likewise, Basu et al. (2014) argue that even with imperfect competition in product markets, TFP calculations based on aggregate consumption (rather than output) still provide valid welfare comparisons across countries. Specifically, they show that welfare can be measured through the present value of future relative TFP and the relative current capital stock (per capita). Most important, this result does not rely on assumptions regarding market structure and technology, but follows only from assuming a price-taking, optimizing representative consumer. Furthermore, they show that in an open economy, this welfare-relevant measure of TFP should be computed based on real domestic absorption.[19] For these various reasons, we expect that the methods used to construct PWT8, as outlined in the next section, while derived from perfectly competitive behavior as in Theorems 1 and 2, may well apply more generally.

V. Implementation in PWT

Measures of real GDP in PWT8 are built up from detailed price data on consumption, C and G; investment, I; exports and imports, X and M; as well as nominal expenditures and trade. This is done in a two-step aggregation procedure: using the GEKS price indexes (7) to compute aggregates *within* the major categories of GDP; and then using reference prices *for each* of these major categories computed as the world average prices with the Geary-Khamis (GK) approach. We first outline the measurement of GDP from the expenditure side, then from the output side, and finally discuss productivity.

Within each category C, I, and G, we first aggregate the ICP prices using GEKS price indexes.[20] ICP prices are available for the benchmark years 1970, 1975, 1980, 1985, 1996, and 2005. There is an expanded set of countries available from the ICP in each benchmark, and in total 167 countries are used in one benchmark or another. That is the set of countries included in PWT8 (this set will expand as more countries are included in future benchmarks).[21] For each country, we keep track of which benchmarks were used; years in-between benchmarks will have the prices for final goods *interpolated* using the corresponding price trends from countries' national accounts data; and for years before the first or after the last benchmark for each country the prices of final goods are *extrapolated* using national account data (see online Appendix B).

In a second step the GEKS price indexes are used to obtain a (3×1) vector of reference prices for C, I, G (and later, exports and imports).[22] The quantity of domestic final goods C, I, and G are included within the (3×1) vector $\mathbf{q}_j$.[23] Given the (3×1) vector of reference prices for domestic final goods, $\boldsymbol{\pi}$, the PPP exchange rate can be defined as

$$PPP_j^q = \mathbf{p}_j'\mathbf{q}_j / \boldsymbol{\pi}_j'\mathbf{q}_j. \tag{15}$$

[19] As discussed in Section I and in the next section, PWT8.1 includes the TFP measure *CWTP* that is based on domestic absorption rather than output.

[20] Since output prices for government consumption, G, are typically unobservable, ICP provides information on relative input prices, notably relative wages. For PWT, we modify the ICP numbers by implementing a common productivity adjustment approach described in Chapter 4 of World Bank (2014); see also Heston (2013). This leads to results that are more comparable between countries and to what is implemented in ICP 2011.

[21] The new PWT9 will be based on the 2011 ICP and cover nearly 180 countries.

[22] Below, we outline how these reference prices are estimated from the GK procedure; see also online Appendix B.

[23] The relative quantity of these variables is obtained by dividing their relative value by the GEKS price index.

This equation shows that the PPP exchange rate is just the ratio of expenditure at local prices to that at reference prices measured in the currency of the base country, in our case the US. Because the PPP is in units of the currency of country *j* per unit of the currency of the base country, it is common to divide it by the nominal exchange rate to obtain what is called the "price level" of country *j*:

$$PL_j \equiv \frac{PPP_j}{\mathcal{E}_j}.$$

The ratio of price levels is typically known as the *real exchange rate* between countries. These price levels are given in PWT for each country relative to the United States. Denoting nominal GDP in national currency by GDP_j, and the trade balance by $(X_j - M_j)$, *real GDP on the expenditure side* is then computed as

$$\text{(16)} \qquad CGDP_j^e \equiv \boldsymbol{\pi}'\mathbf{q}_j + \left(X_j - M_j\right)/PPP_j^q = GDP_j/PPP_j^q.$$

The expression $\boldsymbol{\pi}'\mathbf{q}_j$ on the left is just real expenditure on final goods, which is obtained by deflating nominal expenditure $\mathbf{p}_j'\mathbf{q}_j$ by the PPP exchange rate in equation (15). In the second term, we *also* deflate the trade balance by the same PPP exchange rate that is constructed over final goods. From the point of view of the representative consumer, we are essentially treating the trade balance as an income transfer that is then deflated by the local prices, including prices for nontraded goods. By this logic, one can view (16) as a measure of the standard of living for country *j*, but now extended to incorporate the trade balance.

In addition to $CGDP^e$, PWT8 also includes a measure of real consumption and a measure of real domestic absorption. The measure of real domestic absorption is equal to $CGDP^e$ except for the trade balance, so $CDA_j = \boldsymbol{\pi}'\mathbf{q}_j$. Real consumption includes both private (C) and public consumption (G), but in contrast to real domestic absorption excludes investment, so $CCON_j = \pi_C q_{Cj} + \pi_G q_{Cj}$. In PWT8, we provide these real consumption and real GDP^e variables and also, for the first time, we provide estimates of real GDP on the output side (GDP^o) for the full set of PWT countries and all years. This requires relative price data for imports and exports, as discussed in Feenstra et al. (2009). Compared with their experimental estimates, the real GDP^{o} results in PWT8 are much more reliable due to the use of new relative prices of exports and imports that correct for quality, as constructed by Feenstra and Romalis (2014). This quality correction is crucial as the prices of traded goods are computed as unit values of export and imports products, rather than the precisely specified prices collected for consumption and investment goods in the ICP.

To correct the unit values for quality, recent literature such as Khandelwal (2010) and Hallak and Schott (2011) presume that a good that is imported in high quantity but without having a low price must be of high quality. One shortcoming of this approach is that a good might be imported in high quantity because there are many varieties of it (e.g., many models of cars from Japan).[24] So Feenstra and Romalis

[24] We have also computed the quality-adjusted export prices using the technique of Khandelwal (2010), who uses country population as a proxy for export variety. As shown in Feenstra and Romalis (2014, Figure XIII), there

(2014) refine this demand-side measurement by adding a supply side with monopolistically competitive firms. Using the assumption of free entry they solve for the variety of each good produced, so that differences in the range of varieties sold from each exporter to each importing country are accounted for. Dividing the unit-values of exports and imports by the quality estimates, quality-adjusted prices are obtained. This procedure is implemented at the level of four-digit Standard International Trade categories between each pair of countries, and then aggregated to 6 one-digit Broad Economic categories, such as consumer goods or fuel. The quality-adjusted price indexes in these broad categories show much less variation across countries than do the raw unit-values, since most of the variation in the unit values is due to quality.

The quality-adjusted trade prices are an important ingredient for real GDP^{o}. They are averaged across countries to obtain reference prices for exports and imports. These are included within $\boldsymbol{\Pi} = (\boldsymbol{\pi}, \boldsymbol{\pi}^{x}, \boldsymbol{\pi}^{m})$ and applied to the revenue function[25] to measure real GDP on the output side as

$$(17) \quad CGDP_j^o \equiv \boldsymbol{\pi}'\mathbf{q}_j + \boldsymbol{\pi}'^{x}\mathbf{x}_j - \boldsymbol{\pi}'^{m}\mathbf{m}_j = \frac{C_j + I_j + G_j}{PPP_j^q} + \frac{X_j}{PPP_j^x} - \frac{M_j}{PPP_j^m} = \frac{GDP_j}{PPP_j^o},$$

where the equalities follow by defining the PPPs of final goods, exports, imports, and GDP as

$$(18) \quad PPP_j^q \equiv \frac{\mathbf{p}_j'\mathbf{q}_j}{\boldsymbol{\pi}'\mathbf{q}_j}, \quad PPP_j^x \equiv \frac{\mathbf{p}_j'^{x}\mathbf{x}_j}{\boldsymbol{\pi}'^{x}\mathbf{x}_j}, \quad PPP_j^m \equiv \frac{\mathbf{p}_j'^{m}\mathbf{m}_j}{\boldsymbol{\pi}'^{m}\mathbf{m}_j},$$

$$PPP_j^o \equiv \frac{\mathbf{p}_j'\mathbf{q}_j + \mathbf{p}_j'^{x}\mathbf{x}_j - \mathbf{p}_j'^{m}\mathbf{m}_j}{\boldsymbol{\pi}'\mathbf{q}_j + \boldsymbol{\pi}'^{x}\mathbf{x}_j - \boldsymbol{\pi}'^{m}\mathbf{m}_j}.$$

It is apparent that nominal exports and imports in (17) are not deflated by a PPP computed over final goods, as in (16), but are deflated by PPPs that are specific to exports and imports. The use of reference prices for all goods, including exports and imports as in (17), makes real GDP^{o} an appropriate measure of the productive

is then a strong negative correlation between export quality and population, and so a strong positive correlation between the quality-adjusted terms of trade and population. As a result, the Khandelwal procedure leads to countries with large populations, such as India, having $CGDP^e$ noticeably higher than $CGDP^o$. We believe that this tendency is artificial (i.e., India does not have such a strong terms of trade) and it does not occur using our own methods.

[25] The revenue function presumes perfect competition, whereas the quality-adjusted export and import prices have been obtained from a model of monopolistic competition. This does not create any inconsistency for import prices, because the quality-adjusted demands are still a standard function of the quality-adjusted prices. But on the export side, monopolistically competitive firms are charging a fixed CES markup over marginal costs, contrary to the standard revenue function. Still, Feenstra and Kee (2008) show that in the monopolistic competition model with CES preferences, a well-specified GDP function is being maximized. Further, Burstein and Cravino (2015) allow for monopolistic competition in an international trade model, and find that conventional measures of GDP construction are still adequate to a first order. For these reasons and because there is no practical alternative, we are willing to use the quality-adjusted export prices even with the perfectly competitive revenue function.

capacity of countries. If we divide the PPPs in (18) by the nominal exchange rate, then we obtain the *price levels* of these components of GDP.

The reference prices used in computing real GDP^{o} and GDP^{e} have not been defined up to this point; in PWT8 we compute these based on the Geary-Khamis (GK) approach. The first equation is the definition of the PPP for GDP^{o}, PPP_j^o, in (18). Given this PPP, the reference price for each product is computed as the (quantity-weighted) average of the country prices relative to their PPP:

$$(19) \qquad \pi_i = \frac{\sum_{j=1}^{C} (p_{ij}/PPP_j^o) q_{ij}}{\sum_{j=1}^{C} q_{ij}}, \quad \pi_i^x = \frac{\sum_{j=1}^{C} (p_{ij}^x/PPP_j^o) x_{ij}}{\sum_{j=1}^{C} x_{ij}},$$

$$\pi_i^m = \frac{\sum_{j=1}^{C} (p_{ij}^m/PPP_j^o) m_{ij}}{\sum_{j=1}^{C} m_{ij}},$$

where the index i in the reference prices for final goods, π_i, runs over C, I, G, and in the reference prices for exports and imports, π_i^x and π_i^m, runs over the one-digit Broad Economic categories. Then PPP_j^o in (18) together with (19) are a system of equations that can be solved up to a normalization.

Real GDP on the expenditure side and output side will differ due to the terms of trade faced by countries. This is apparent by taking the difference between (16) and (17):

$$CGDP_j^e - CGDP_j^o = \left(\frac{PPP_j^x}{PPP_j^q} - 1\right)\frac{X_j}{PPP_j^x} - \left(\frac{PPP_j^m}{PPP_j^q} - 1\right)\frac{M_j}{PPP_j^m}.$$

To simplify this expression, we can divide by $CGDP_j^o$ and rearrange terms to obtain

$$(20) \quad \underbrace{\frac{CGDP_j^e - CGDP_j^o}{CGDP_j^o}}_{Gap} = \frac{1}{2}\underbrace{\left(\frac{PPP_j^x}{PPP_j^q} - \frac{PPP_j^m}{PPP_j^q}\right)}_{Terms\ of\ trade}\underbrace{\left(\frac{X_j/PPP_j^x}{CGDP_j^o} + \frac{M_j/PPP_j^m}{CGDP_j^o}\right)}_{Real\ openness}$$

$$+ \underbrace{\left[\frac{1}{2}\left(\frac{PPP_j^x + PPP_j^m}{PPP_j^q}\right) - 1\right]}_{Traded/nontraded\ price}\underbrace{\left(\frac{X_j/PPP_j^x}{CGDP_j^o} - \frac{M_j/PPP_j^m}{CGDP_j^o}\right)}_{Real\ balance\ of\ trade\ share}.$$

We see that the gap between real GDP^{e} and real GDP^{o} can be expressed as the sum of two terms: the first is the terms of trade (expressed as a difference rather than a ratio) times real openness; and the second is the relative prices of traded goods (again expressed as a difference) times the real balance of trade. The influence of both these terms on the gap between real GDP from the expenditure and output sides has also been shown by Kohli (2004, 2006) and Reinsdorf (2010), and we will illustrate this relation with some examples from PWT8.1 in Section VI.

The above formulas are computed for each year, obtaining the measures of real GDP that are based on current-year reference prices, i.e., $CGDP_j^e$ and $CGDP_j^o$. To correct for changing reference prices over time, we use Corollary 1 to define the growth rate of real GDP^{o} as

$$(21)\quad \frac{RGDP_{jt}^{o}}{RGDP_{jt-1}^{o}} \equiv \left[\left(\frac{\boldsymbol{\Pi}'_{t-1}\mathbf{y}_{jt}}{\boldsymbol{\Pi}'_{t-1}\mathbf{y}_{jt-1}}\right)\left(\frac{\boldsymbol{\Pi}'_{t}\mathbf{y}_{jt}}{\boldsymbol{\Pi}'_{t}\mathbf{y}_{jt-1}}\right)\right]^{\frac{1}{2}}$$

$$= \left[\left(\frac{\boldsymbol{\pi}'_{t-1}\mathbf{q}_{jt} + \boldsymbol{\pi}'^{x}_{t-1}\mathbf{x}_{jt} - \boldsymbol{\pi}'^{m}_{t-1}\mathbf{m}_{jt}}{\boldsymbol{\pi}'_{t-1}\mathbf{q}_{jt-1} + \boldsymbol{\pi}'^{x}_{t-1}\mathbf{x}_{jt-1} - \boldsymbol{\pi}'^{m}_{t-1}\mathbf{m}_{jt-1}}\right)\left(\frac{\boldsymbol{\pi}'_{t}\mathbf{q}_{jt} + \boldsymbol{\pi}'^{x}_{t}\mathbf{x}_{jt} - \boldsymbol{\pi}'^{m}_{t}\mathbf{m}_{jt}}{\boldsymbol{\pi}'_{t}\mathbf{q}_{jt-1} + \boldsymbol{\pi}'^{x}_{t}\mathbf{x}_{jt-1} - \boldsymbol{\pi}'^{m}_{t}\mathbf{m}_{jt-1}}\right)\right]^{\frac{1}{2}}.$$

Thus, the quantities of final goods, exports, and imports change from $t-1$ to t in both ratios, and are evaluated using the reference prices from one period or the other, and then taking the geometric mean. PWT8 uses the growth rates from this formula to compute real GDP^{o} in all years other than the 2005 benchmark, for which $RGDP^{o} = CGDP^{o}$.

In addition, the constant-price growth rates of real GDP^{e} are obtained by using only the reference prices $\boldsymbol{\pi}'_{t-1}$ and $\boldsymbol{\pi}'_{t}$ of the final consumption goods. $RGDP^{e} = CGDP^{e}$ is defined by (16) in the benchmark year 2005, and its growth rate to other years is obtained as

$$(22)\quad \frac{RGDP_{jt}^{e}}{RGDP_{jt-1}^{e}} \equiv \left[\left(\frac{\boldsymbol{\pi}'_{t-1}\mathbf{q}_{jt} + (X_{jt} - M_{jt})/PPP_{jt}^{q}}{\boldsymbol{\pi}'_{t-1}\mathbf{q}_{jt-1} + (X_{jt-1} - M_{jt-1})/PPP_{jt-1}^{q}}\right)\left(\frac{\boldsymbol{\pi}'_{t}\mathbf{q}_{jt} + (X_{jt} - M_{jt})/PPP_{jt}^{q}}{\boldsymbol{\pi}'_{t}\mathbf{q}_{jt-1} + (X_{jt-1} - M_{jt-1})/PPP_{jt-1}^{q}}\right)\right]^{\frac{1}{2}}.$$

Notice that in (22) we deflate nominal exports and imports by the PPPs for final goods, PPP_{jt}^{q} and PPP_{jt-1}^{q}, computed from the reference prices for those goods. This is in contrast to (21) where the actual reference prices of exports and imports are used.

Theorem 2 tells us that by deflating the observed difference in real GDP^{o} by the Törnqvist quantity index of factor endowments, we obtain a meaningful measure of the productivity difference between the countries. The Törnqvist quantity index is constructed using the factor prices that are implied by the reference prices for goods, $\boldsymbol{\Pi}$. In practice we do not observe these factor prices, and so we replace the theoretical expressions in (13)–(14) with versions that we can measure from the data:

$$(23)\quad CTFP_{jk} \equiv \frac{CGDP_j^{o}}{CGDP_k^{o}} \Big/ Q_T(\mathbf{v}_j, \mathbf{v}_k, \mathbf{w}_j, \mathbf{w}_k),$$

where we use $CTFP_{jk}$ to denote the (current-year price) productivity of country j relative to k, and the Törnqvist quantity index of factor endowments Q_T is evaluated with observed factor prices and shares. PWT8 includes $CTFP_{jk}$ computed with current year prices for each country j relative to the United States. In addition to the production-side measure of $CTFP_{jk}$, PWT8 also includes a welfare-relevant measure of TFP based on the work of Basu et al. (2014). This measure is based not on relative $CGDP^o$ levels, but instead on relative domestic absorption, CDA:

$$CWTFP_{jk} \equiv \frac{CDA_j}{CDA_k} / Q_T\left(\mathbf{v}_j, \mathbf{v}_k, \mathbf{w}_j, \mathbf{w}_k\right). \tag{24}$$

An analogous expression is used for *productivity growth* in each country, which is defined by reintroducing time subscripts and using real GDP and factor input growth rates obtained from national accounts data:

$$RTFP^{NA}_{jt,t-1} \equiv \frac{RGDP^{NA}_{jt}}{RGDP^{NA}_{jt-1}} / Q_T\left(\mathbf{v}_{jt}, \mathbf{v}_{jt-1}, \mathbf{w}_{jt}, \mathbf{w}_{jt-1}\right). \tag{25}$$

For this purpose, we have developed new data on factor inputs—capital and labor—and factor income shares.[26] Specifically, PWT8 (re)introduces a measure of the physical capital stock, based on long time-series of investment by asset. For each country, we distinguish investment in structures, transport equipment, and machinery, and for a range of countries, we also separately distinguish investment in computers, communication equipment, and software. Investments are cumulated into capital stocks using asset-specific geometric depreciation rates using the perpetual inventory method. The relative factor price of the capital stock is computed by aggregating asset-specific investment prices using shares of each asset in the total (current cost) capital stock. PWT has long included data on the number of workers in an economy, but a more accurate measure of relative labor input should account for the large differences in schooling across countries. To that end, PWT8 includes an index of human capital per worker based on the average years of schooling, linearly interpolated from Barro and Lee (2013), and an assumed rate of return for primary, secondary, and tertiary education, as in Caselli (2005).[27]

We have also developed new information about the share of labor income in GDP. An important measurement challenge, well known since Gollin (2002), is that self-employed workers earn income for both the labor and capital they supply. We follow Gollin (2002) in splitting this mixed income between capital and labor income using the same shares as found for nonmixed income. Where mixed income is not available as a separate data item in a country's national accounts, we impute the labor income of the self-employed either by assuming that self-employed earn the same average income as employees or based on the share of agriculture in value added. In online Appendix C, we go into greater detail on these measurement choices and their implications. One important result, though, is that the global decline in the corporate labor income share that was documented by Karabarbounis and Neiman (2014) is also found for our economy-wide labor shares. In computing

[26] See online Appendix C for more details on the data sources and measurement methodology.

[27] Though we note that this is an imperfect measure of human capital as differences in the returns to experience (Lagakos et al. 2014) and the quality of schooling (Hanushek and Woessman 2012) are not accounted for.

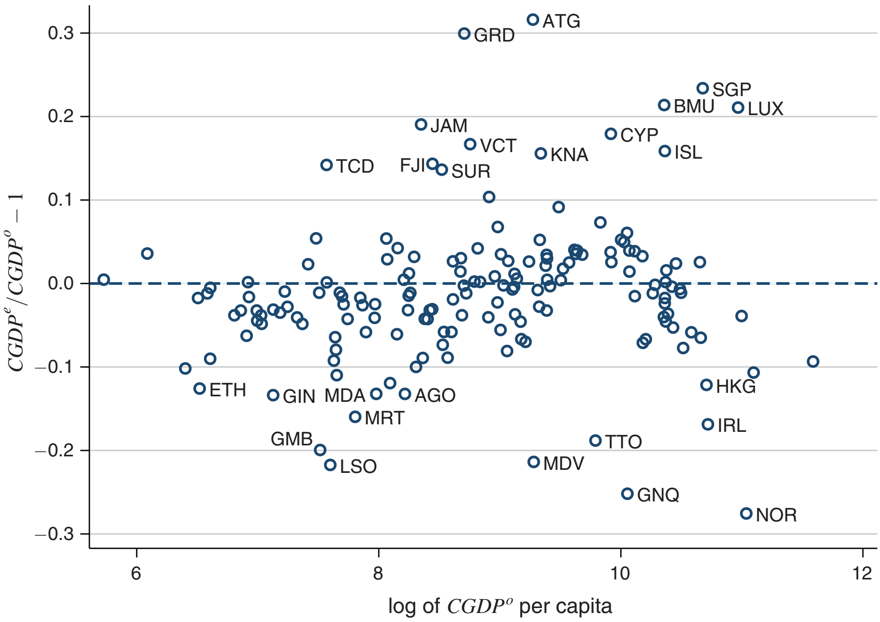

FIGURE 1. REAL GDP PER CAPITA IN 2005 AND THE GAP BETWEEN $CGDP^e$ AND $CGDP^o$

Notes: Includes 166 of the countries in PWT. Excluded is El Salvador with a gap of 1.5.

Source: Computations based on PWT8.1.

the overall quantity index of factor endowments, we assume that the income share of physical capital equals 1 minus the labor income share, though future work on distinguishing natural from physical capital, as done for one year in Caselli and Feyrer (2007), would be an important improvement.

VI. Applications

Based on the next generation of PWT, version 8.1, we provide three applications to illustrate its usefulness: an analysis of the difference between real GDP from the expenditure and output sides; an analysis of the Balassa-Samuelson effect; and a decomposition of the variance of real GDP per capita into the variance of factor inputs and productivity (known as development accounting).

A. GDP^e versus GDP^o

Figure 1 illustrates how real GDP^o differs from real GDP^e in 2005. For some countries the differences are clearly notable, with several absolute differences near 30 percent. At the same time, many differences are not so large: 153 of the 166 countries have a GDP^e level within 10 percent of their GDP^o level. Note also that the gap between $CGDP^e$ and $CGDP^o$ does not vary systematically with the level of $CGDP^o$ per capita. To better understand what is driving these gaps, we use the decomposition introduced in equation (20). According to this, consumption possibilities can exceed productive capacity when a country faces favorable terms of

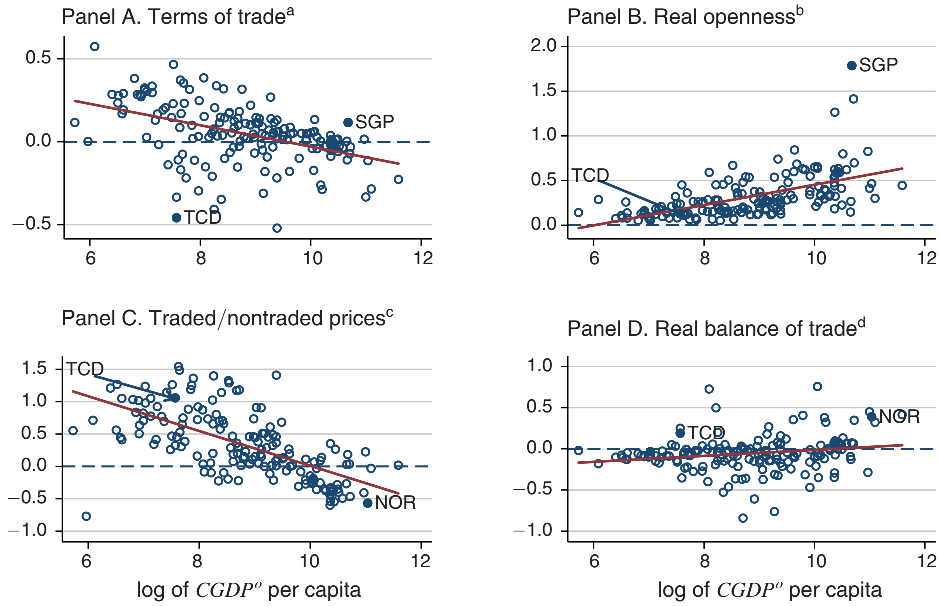

Figure 2. Decomposing the Gap between $CGDP^e$ and $CGDP^o$ in 2005

Notes: See equation (19) for the definition of the concepts plotted in the four panels.
[a]Omitted is Zimbabwe (1.0)
[b]Omitted is El Salvador (4.2)
[c]Omitted are Tajikistan (2.4), Uzbekistan (3.9), and Zimbabwe (3.0)
[d]Omitted is El Salvador (−2.2)

Source: Computations based on PWT8.1.

trade and the gains are larger whenever real openness is larger. In contrast, a country only gains excess consumption possibilities from a positive real balance of trade when traded goods are expensive compared to nontraded goods. Following the basic argument from Section II this will mostly be the case in poorer countries while in rich countries, traded goods are relatively cheap. Figure 2 illustrates the decomposition of the gaps from Figure 1 in 2005, according to equation (19).

Panel A of Figure 2 shows that terms of trade are negatively related to $CGDP^o$ per capita, which follows from the results of Feenstra and Romalis (2014).[28] There is a positive relationship between real openness and income levels as shown in panel B; a finding that is consistent with Alcalá and Ciccone (2004). Panel C shows how the ratio of traded to nontraded prices declines with income, consistent with the Balassa-Samuelson effect, while the real balance of trade in panel D shows a positive relationship with income, reflective of the Lucas (1990) paradox where capital flows from poor to richer countries. The overall gap between $CGDP^e$ and $CGDP^o$ is not systematically related with income levels, however, because the various positive and negative relationships with income levels cancel out when combined.

For illustration purposes, we have highlighted the observations for Chad (TCD), Singapore (SGP, panels A and B), and Norway (NOR, panels C and D). Figure 1

[28]Feenstra and Romalis (2014) find that quality-adjusted import prices are lower for poor countries, leading to the negative relationship between the terms of trade and country income. This relationship is weak before the mid-1990s, while there is a consistently significant negative relationship since 1996.

shows considerable positive gaps between $CGDP^e$ and $CGDP^o$ for both Singapore (+23 percent) and Chad (+13 percent), while the gap in Norway is strongly negative (−28 percent). The reason for Singapore's large gap is a small but positive terms of trade combined with the largest observed real openness—its real balance of trade and traded/nontraded price ratio contribute little. In contrast, Chad had very negative terms of trade but combined with low real openness, the negative contribution to the overall gap is limited. Because of a high traded/nontraded price ratio and a positive real balance of trade, Chad's overall gap ends up strongly positive. Norway, finally, also has a positive real balance of trade but because nontraded prices are relatively high in 2005, the overall result is a negative gap between $CGDP^e$ and $CGDP^o$.

While the above discussion has highlighted some countries with large differences between $CGDP^e$ and $CGDP^o$, for many important countries the difference is quite small. China, for example, has $CGDP^e$ ($CGDP^o$) per capita that is 13.7 percent (13.8 percent) of that in the United States in 2005 and 20.4 percent (20.5 percent) in 2011. The similar values for the expenditure- and output-side measures of real GDP follow from variables in (20) that are relatively small and offsetting in sign: China's terms of trade is slightly negative, but its ratio of traded to nontraded prices is greater than unity with a positive but modest real balance of trade.[29]

B. *The Balassa-Samuelson Effect in PWT8.1*

In panel C of Figure 2 we illustrated the Balassa-Samuelson or Penn effect for 2005: the observation that the relative price of nontraded goods to traded goods increases with the income level of a country. Surprisingly, Bergin, Glick, and Taylor (2006) found that there was no evidence of a Penn effect in the early 1950s, and that the effect gradually became significant and strengthened over time. Their analysis was based on PWT version 6 and we revisit it using version 8. Consider the regression

$$\ln\left(\frac{PPP_{it}^o}{\mathcal{E}_{it}}\right) = \beta_0 + \beta_1 \ln\left(\frac{CGDP_{it}^o}{POP_{it}}\right) + \varepsilon_{it}, \tag{26}$$

where $\mathcal{E}_{it}$ is the exchange rate and POP is the population of country i at time t.[30] The dependent variable in (26) is the price level of each country relative to the United States.

The finding of Bergin, Glick, and Taylor (2006) is puzzling, as the Balassa-Samuelson effect was already identified in data for the 1950s and 1960s

[29] We note that the 2005 prices for final consumption goods for China used in PWT8 have been adjusted downward by 20 percent as compared to the ICP values, which is the same adjustment that was made in PWT7 and reflects the fact that the ICP prices were collected in large part from urban areas in China. This correction is discussed further in Feenstra et al. (2013). Further adjustments for biases in ICP 2005 are also made in PWT8.1, based on Inklaar and Rao (2014) and described in Feenstra, Inklaar, and Timmer (2015b).

[30] Bergin, Glick, and Taylor (2006) divide the country's GDP per capita level by the US level in every year, but this only affects the estimate of β_0. Also note that sometimes the exchange-rate-converted GDP per capita level is used as the explanatory variable instead of the PPP-converted GDP per capita level. We follow the approach of Bergin, Glick, and Taylor (2006), which was also advocated by Officer (1982). Officer argued that a productivity measure would be preferable to a GDP per capita level. Results using $CGDP^o$ per capita or $CTFP$ are very similar.

in, e.g., Balassa (1964), therefore raising the question why the effect shows up so much later in PWT8. One possibility could be that this effect is an emergent property and Bergin, Glick, and Taylor (2006) propose a model that yields this outcome. An alternative possibility is that the estimation of the PPPs used in the regression (26) in early years is problematic. As explained in Section V, PWT8 includes two types of observations: those based directly on ICP benchmark price survey results or interpolated between benchmarks; and those based on extrapolations from the oldest or most recent benchmarks using relative inflation rates from countries' national accounts data. The benchmark observations (and in effect the interpolated observations) make no assumptions regarding the evolutions of PPPs over time; they are simply based on the observed benchmark survey.

In contrast, extrapolating assumes that the change in PPPs is well-approximated by relative inflation. Deaton (2012) argues, however, that relative inflation will be a systematically biased estimate of the change in PPPs across countries. Specifically, he argues that under plausible conditions, the PPP of a poor country relative to a rich country will increase at a faster rate than implied by the difference in overall inflation. Figure 3 provides evidence in support of this contention. Panel A plots relative prices and income levels for benchmark or interpolated observations, while panel B plots observations based on extrapolations from earlier or later benchmarks. For both sets of observations, there is a significant positive relationship between price and income levels, but the regression line in panel A is significantly steeper than the line in panel B.[31] In online Appendix D we show more systematically that the extrapolation procedure used for nonbenchmark observations indeed lead to the supposed disappearance of the Balassa-Samuelson effect in early years.

The study of Bergin, Glick, and Taylor (2006) is based on PWT6 which relied exclusively on extrapolation of PPPs from a recent benchmark year, which would explain their findings. This illustrates the usefulness of including historical benchmark material and clearly distinguishing between benchmark/interpolated and extrapolated observations, as done in PWT8.

C. Capital and Productivity

Traditionally, the main strength of PWT has been its information on GDP per capita, useful for comparing the standard of living across countries. Yet to gain an understanding of the (proximate) sources of the differences in living standards, we should analyze differences in the level of output, inputs, and productivity: see, e.g., Klenow and Rodríguez-Clare (1997); Hall and Jones (1999); and Caselli (2005). In PWT8, the introduction of relative prices of exports and imports and the resulting real GDP^o variable means that there is now a true measure of relative output. PWT has also long provided information on labor input, i.e., the number of workers, but information on physical capital has been absent since PWT version 6 and there has never been information in PWT on human capital or productivity. This has left researchers to their own devices in compiling productivity estimates. As a

[31] Note that we only include observations from ICP benchmark years (1970, 1975, 1980, 1985, 1996, and 2005) for a more balanced set of observations across the income spectrum in both panels. Online Appendix D shows results for the full range of years.

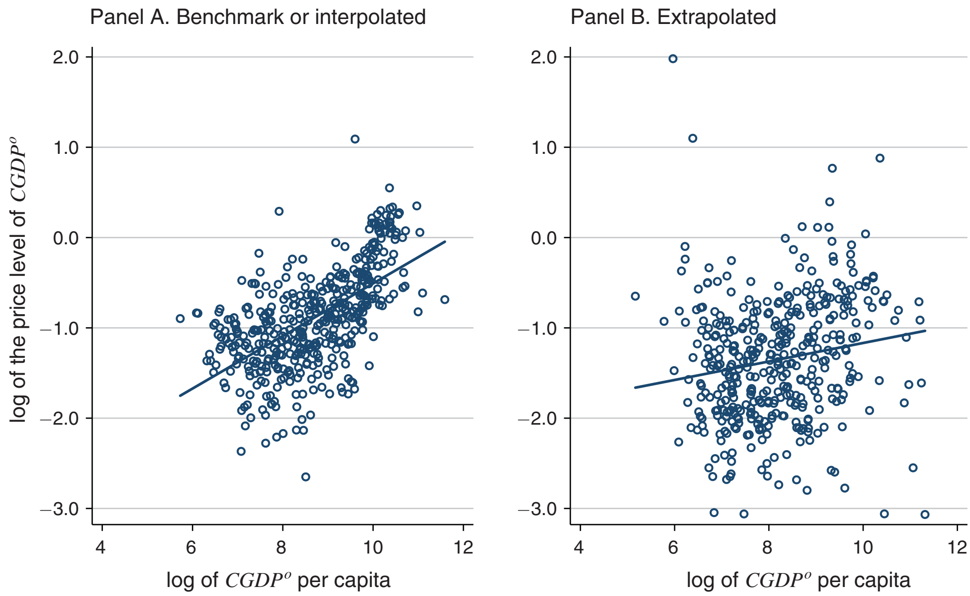

FIGURE 3. THE BALASSA-SAMUELSON EFFECT IN PWT8.1

Notes: Included are only observations in ICP benchmark years: 1970, 1975, 1980, 1985, 1996, and 2005. Observations are distinguished by whether the price level for a country is from that ICP benchmark, interpolated between ICP benchmarks, or extrapolated from an earlier or later benchmark. See online Appendix D for a comprehensive analysis. The solid line is the least-squares regression line; the slope of the regression line in panel B is significantly smaller than in panel A.

Source: Computations based on PWT8.1.

result these estimates tend to rely on numerous simplifying assumptions, such as a Cobb-Douglas production function and homogeneous physical capital.

The new version of PWT represents an important step forward by including measures of physical and human capital and estimates of productivity based on the translog production function (see Theorem 2) which allows for substitution elasticities to differ across countries and over time. The first novelty is to estimate physical capital stocks for all countries in PWT based on data of investment by asset. The second novelty is to estimate the share of labor income in GDP for a large majority of PWT countries. These are combined with (more standard) measures of human capital to arrive at measures of total factor productivity (TFP). A detailed description of the data is included as online Appendix C and here we provide an outline of the approach and show its implications for development accounting results as in Caselli (2005).

Physical capital stocks are computed by cumulating the depreciated past investments, but we distinguish investments by type of asset.[32] This has two important implications when contrasted with the study by Caselli (2005), which is representative for the broader literature. First, the average depreciation rate now varies across countries and over time, as countries differ in the asset composition

[32] Investment by type data is partly from National Accounts statistics, partly from estimates using the commodity-flow method along similar lines as Caselli and Wilson (2004).

of their capital stock and depreciation differs across assets. Second, while existing studies (implicitly) use the relative price of investment to compare the level of the capital stock across countries, we use information on the asset composition of the capital stock to compute a relative price of the capital stock. This relative price of the capital stock gives much larger weight to the price of buildings—which are comparatively cheap in poorer countries, than to the price of machinery, which is relatively more expensive as buildings are the longest-lived assets: see Hsieh and Klenow (2007). So while the price of investment goods relative to consumption goods declines rapidly with income—as in Hsieh and Klenow (2007)—there is a weaker relation with income levels when comparing the price of the capital stock to the price of consumption goods.

The typical approach in development accounting is to assume that the output elasticity of labor is identical across countries and constant over time at 0.7, yet the evidence in support of this assumption is modest at best. The oft-cited work of Gollin (2002) shows substantial cross-country variation in the income share of labor in GDP, as well as Bernanke and Gürkaynak (2002). More recently, Karabarbounis and Neiman (2014) shows that, for many countries, the labor share has been declining in the last two decades. In PWT8, we therefore estimate the share of labor income in GDP for as many countries and years as possible. Information on the labor compensation of employees is broadly available, but the labor compensation of self-employed workers needs to be separately estimated. Here we broadly follow the existing approaches in the literature, but on a substantially larger scale.[33] This yields some key findings, namely: (i) an average labor share of 0.52, which is much lower than the 0.7 that is typically assumed; (ii) no systematic variation of labor shares with income levels; (iii) declining labor shares over time in 89 of the 127 countries.

Combining the new information on labor shares, physical capital, and estimates of human capital (based on the average years of schooling of Barro and Lee 2013) yields data on overall factor inputs (denoted Q) and relative productivity levels ($CTFP$). These can be used to provide a variance analysis of real GDP per capita across countries as in Caselli (2005). Combining (20) with (23), we obtain a decomposition of current-price real GDP on the expenditure side:

$$\frac{CGDP_j^e}{CGDP_k^e} = CTFP_{jk} \times Q_T(\mathbf{v}_j, \mathbf{v}_k, \mathbf{w}_j, \mathbf{w}_k) \times \left(\frac{1 + Gap_j}{1 + Gap_k}\right),$$

where the Gap between real GDP^e and real GDP^o is defined by the various terms-of-trade and balance-of-payments expressions on the right of (20). We report summary statistics from this decomposition in Table 2 and illustrate how these new measures compare to the earlier measures based on simplifying assumptions, namely a Cobb-Douglas production function with labor share of 0.7 and homogeneous capital stock (baseline).

[33] This involves using data on total income (from capital and labor) of self-employed, assuming similar wages for self-employed as for employees and information on the importance in the economy of agriculture, the dominant sector for self-employed workers. See online Appendix C for details.

TABLE 2—COUNTING FOR CROSS-COUNTRY VARIATION OF $CGDP^E$ PER CAPITA IN 2005

	Baseline	(1)	(2)
Variance (ln($CGDP^e$))	1.433	1.433	1.433
Variance (ln(1+*Gap*))	0.006	0.006	0.006
Variance (ln(Q))	0.351	0.332	0.485
Variance (ln(*CTFP*))	0.412	0.452	0.345
Fraction of variance accounted for by factor inputs	0.253	0.232	0.338

Notes: $CGDP^e$ is expenditure-based real GDP; *Gap* is the difference between expenditure-based and output-based real GDP; Q is inputs of physical and human capital per capita, and *CTFP* is total factor productivity. The baseline decomposition is based on a constant labor share of 0.7 and homogeneous capital. Variant (1) allows for asset heterogeneity and variant (2) also allows for variation in the labor share.

Source: Computations based on PWT8.1.

This exercise aims to account for differences in $CGDP^e$ per capita by variation in the gap between $CGDP^e$ and $CGDP^o$—the effect of the terms of trade on standards of living—variation in factor inputs and variation in TFP. The first column mimics the approach of Caselli (2005), the second column accounts for the heterogeneity of physical capital, and the third column also accounts for the heterogeneity of labor shares and this is our preferred measure. Accounting for labor share heterogeneity has an important impact on the variation in $CGDP^e$ per capita that is accounted for by variation in factor inputs, which increases from 25.3 percent to 33.8 percent. The very small share of variation accounted for by the gap between $CGDP^e$ and $CGDP^o$ was already implicit in Figure 1, as there was no systematic relationship between this gap and income levels.

Important to note, though, is that physical capital in PWT only covers *produced* capital, such as buildings and machinery, not natural capital such as land or subsoil resources. This natural capital is particularly important in developing economies, as shown by Caselli and Feyrer (2007). Augmenting capital inputs with natural capital using data from World Bank (2011) leads to a decline in the variance of $CGDP^e$ per capita accounted for by variation in factor inputs (to 27.6 percent), suggesting that efficiency differences in the use of natural capital might be bigger than that from produced capital. As yet, data are only available for a few recent years and are not yet ideally suited for cross-country comparisons of inputs levels. However, including natural capital would be an important future improvement for PWT.

VII. Conclusions

From its inception, the International Comparisons Program (ICP), on which PWT is built, only collects the prices of final products—for consumption, investment, and the government—across countries. It was prohibitively expensive to further collect comparable prices for the whole range of industrial and intermediate inputs used in economies, many of which are also traded. This limitation means that calculations based on ICP prices only are best thought of as representing the standard of living of countries rather than their production possibilities. Feenstra et al. (2009) argued that a measure of the productive capacity of countries could be obtained by combining the ICP data with prices for exports and imports. These two approaches lead to measures of real GDP on the expenditure-side and real GDP on the output-side, respectively, both of which are included in the new PWT version 8.1.

The second contribution of PWT8 is to improve upon the growth of real GDP previously reported in PWT, which was based on national accounts data. Johnson et al. (2013) criticized growth rates as being dependent on the benchmark year of ICP data, and thereby dependent on the version of PWT being used. That problem is resolved in PWT8 by using *multiple* ICP benchmarks: for all of our measures of real GDP, the growth rate will not change in between existing benchmark years as new benchmarks become available, unless the underlying nominal GDP data from the national accounts are revised. We have shown that incorporating multiple ICP benchmarks also ensures that relationships such as the Balassa-Samuelson effect remain apparent in the dataset, rather than disappearing when going back further in time.

The final contribution of PWT8 is to reintroduce a measure of the capital stock and, for the first time, include a measure of relative TFP across countries. We have shown that, compared to standard findings in the literature, cross-country variation in factor inputs can account for more of the cross-country variation in $CGDP^e$ per capita. This is mostly because PWT8.1 incorporates new estimates of the labor share in GDP that vary in a meaningful fashion across countries and over time.

Taken together, these contributions show that PWT version 8 breaks new ground in providing a cross-country dataset that is closer linked to the theoretical concepts of welfare and production, more consistent over time and more transparent in its methods. It should be noted however, that revisions will remain part of future versions of PWT. There can be substantial changes to nominal and real national accounts data over time, which will be the principal source of changes in interpolated values as new versions of PWT become available. The release of the 2011 ICP provides new prices for final expenditure which, in conjunction with updated, quality-adjusted prices for exports and imports, will be used to compute real GDP on the expenditure side and output side in PWT version 9. The results for the period 2005–2011, which were *extrapolated* in PWT8, will then be *interpolated* between the 2005 and 2011 benchmarks in PWT version 9. Early analysis on the 2011 ICP prices suggests that they differ quite substantially from extrapolated prices using the 2005 benchmark (Deaton and Aten 2014; Inklaar and Rao 2014). Therefore, we can expect some discussion and analysis of the 2011 benchmark prices before they are used to revise the recent years and are incorporated into PWT version 9.

REFERENCES

Alcalá, Francisco, and Antonio Ciccone. 2004. “Trade and Productivity.” *Quarterly Journal of Economics* 119 (2): 613–46.

Alterman, William F., W. Erwin Diewert, and Robert C. Feenstra. 1999. *International Trade Price Indexes and Seasonal Commodities*. Washington, DC: US Department of Labor, Bureau of Labor Statistics.

Balassa, Bela. 1964. “The Purchasing-power Parity Doctrine: A Reappraisal.” *Journal of Political Economy* 72 (6): 584–96.

Balk, Bert M. 2008. *Price and Quantity Index Numbers: Models for Measuring Aggregate Change and Difference.* Cambridge, UK: Cambridge University Press.

Barro, Robert J., and Jong Wha Lee. 2013. “A New Data Set of Educational Attainment in the World, 1950–2010.” *Journal of Development Economics* 104: 184–98.

Basu, Susanto, Luigi Pascali, Fabio Schiantarelli, and Luis Serven. 2014. "Productivity and the Welfare of Nations." National Bureau of Economic Research Working Paper 17971.

Bergin, Paul R., Reuven Glick, and Alan M. Taylor. 2006. "Productivity, Tradability, and the Long-Run Price Puzzle." *Journal of Monetary Economics* 53 (8): 2041–66.

Bernanke, Ben S., and Refet S. Gürkaynak. 2002. "Is Growth Exogenous? Taking Mankiw, Romer, and Weil Seriously." In *NBER Macroeconomics Annual 2001*, edited by Ben S. Bernanke and Kenneth Rogoff, 11–72. Cambridge, MA: MIT Press.

Burstein, Ariel, and Javier Cravino. 2015. "Measured Aggregate Gains from International Trade." *American Economic Journal: Macroeconomics* 7 (2): 181–218.

Caselli, Francesco. 2005. "Accounting for Cross-Country Income Differences." In *Handbook of Economic Growth*, Vol. 1A, edited by Philippe Aghion and Steven N. Durlauf, 679–741. Amsterdam: Elsevier.

Caselli, Francesco, and James Feyrer. 2007. "The Marginal Product of Capital." *Quarterly Journal of Economics* 122 (2): 535–68.

Caselli, Francesco, and Daniel J. Wilson. 2004. "Importing Technology." *Journal of Monetary Economics* 51 (1): 1–32.

Caves, Douglas W., Laurits R. Christensen, and W. Erwin Diewert. 1982a. "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity." *Econometrica* 50 (6): 1393–1414.

Caves, Douglas W., Laurits R. Christensen, and W. Erwin Diewert. 1982b. "Multilateral Comparisons of Output, Input, and Productivity Using Superlative Index Numbers." *Economic Journal* 92 (365): 73–86.

Deaton, Angus. 2012. "Consumer Price Indexes, Purchasing Power Parity Exchange Rates, and Updating." http://goo.gl/58c9Rw.

Deaton, Angus, and Bettina Aten. 2014. "Trying to Understand the PPPs in ICP2011: Why Are the Results so Different?" National Bureau of Economic Research Working Paper 20244.

De Loecker, Jan. 2009. "Recovering Markups from Production Data." *International Journal of Industrial Organization* 29 (3): 350–55.

Diewert, W. Erwin. 1976. "Exact and Superlative Index Numbers." *Journal of Econometrics* 4 (2): 115–45.

Diewert, W. Erwin, and Catherine J. Morrison. 1986. "Adjusting Outputs and Productivity Indexes for Changes in the Terms of Trade." *Economic Journal* 96 (383): 659–79.

Feenstra, Robert C., 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 (1): 201–12.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015a. "The Next Generation of the Penn World Table: Dataset." *American Economic Review.* http://dx.doi.org/10.1257/aer20130954.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015b. "What is New in PWT 8.1?" www.rug.nl/research/ggdc/data/pwt/v81/what_is_new_in_pwt_81.pdf.

Feenstra, Robert C., and Hiau Looi Kee. 2008. "Export Variety and Country Productivity: Estimating the Monopolistic Competition Model with Endogenous Productivity." *Journal of International Economics* 74 (2): 500–18.

Feenstra, Robert C., Hong Ma, J. Peter Neary, and D. S. Prasada Rao. 2013. "Who Shrunk China? Puzzles in the Measurement of Real GDP." *Economic Journal* 123 (573): 1100–129.

Feenstra, Robert C., Hong Ma, and D. S. Prasada Rao. 2009. "Consistent Comparisons of Real Incomes across Time and Space." *Macroeconomic Dynamics* 13 (S2): 169–93.

Feenstra, Robert C., and John Romalis. 2014. "International Prices and Endogenous Quality." *Quarterly Journal of Economics* 129 (2): 477–527.

Gerschenkron, Alexander. 1951. *A Dollar Index of Soviet Machinery Output, 1927–28 to 1937*. Report R-197. Santa Monica: Rand Corporation.

Gollin, Douglas. 2002. "Getting Income Shares Right." *Journal of Political Economy* 110 (2): 458–74.

Hall, Robert E., and Charles I. Jones. 1999. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *Quarterly Journal of Economics* 114 (1): 83–116.

Hallak, Juan Carlos, and Peter K. Schott. 2011. "Estimating Cross-country Differences in Product Quality." *Quarterly Journal of Economics* 126 (1): 417–74.

Hanushek, Eric A., and Ludger Woessmann. 2012. "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation." *Journal of Economic Growth* 17 (4): 267–321.

Herrendorf, Berthold, Richard Rogerson, and Akos Valentinyi. 2013. "Two Perspectives on Preferences and Structural Transformation." *American Economic Review* 103 (7): 2752–89.

Heston, Alan. 2013. "Government Services: Productivity Adjustments." In *Measuring the Real Size of the World Economy*, 413–40. Washington, DC: World Bank.

Heston, Alan, and Robert Summers. 1996. "International Price and Quantity Comparisons: Potentials and Pitfalls." *American Economic Review* 86 (2): 20–24.

Hill, Robert J. 1999. "Comparing Price Levels across Countries Using Minimum-Spanning Trees." *Review of Economics and Statistics* 81 (1): 135–42.

Hsieh, Chang-Tai, and Peter J. Klenow. 2007. "Relative Prices and Relative Prosperity." *American Economic Review* 97 (3): 562–85.

Hsieh, Chang-Tai, and Peter J. Klenow. 2010. "Development Accounting." *American Economic Journal: Macroeconomics* 2 (1): 207–23.

Inklaar, Robert, and D. S. Prasada Rao. 2014. "Cross-Country Income Levels over Time: Did the Developing World Suddenly Become Much Richer?" GGDC Research Memorandum 151. Groningen: GGDC.

Jerven, Morten. 2013. "Comparability of GDP Estimates in Sub-Saharan Africa: The Effect of Revisions in Source and Methods since Structural Adjustment." *Review of Income and Wealth* 59 (1): S16–36.

Johnson, Simon, William Larson, Chris Papageorgiou, and Arvind Subramanian. 2013. "Is Newer Better? Penn World Table Revisions and Their Impact on Growth Estimates." *Journal of Monetary Economics* 60 (2): 255–74.

Jones, Charles I., and Peter J. Klenow. 2011. "Beyond GDP? Welfare across Countries and Time." National Bureau of Economic Research Working Paper 16352.

Karabarbounis, Loukas, and Brent Neiman. 2014. "The Global Decline of the Labor Share." *Quarterly Journal of Economics* 129 (1): 61–103.

Khandelwal, Amit. 2010. "The Long and Short (of) Quality Ladders." *Review of Economic Studies* 77 (4): 1450–76.

Klenow, Peter, and Andrés Rodríguez-Clare. 1997. "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" In *NBER Macroeconomics Annual 1997*, edited by Ben S. Bernanke and Julio J. Rotemberg, 73–114. Cambridge, MA: MIT Press.

Kohli, Ulrich. 2004. "Real GDP, Real Domestic Income, and Terms-of-Trade Changes." *Journal of International Economics* 62 (1): 83–106.

Kohli, Ulrich. 2006. "Real GDP, Real GDI, and Trading Gains: Canada, 1981–2005." *International Productivity Monitor* 13: 46–56.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman. 2014. "Experience Matters: Human Capital and Development Accounting." National Bureau of Economic Research Working Paper 18602.

Lucas, Robert E., Jr. 1990. "Why Doesn't Capital Flow from Rich to Poor Countries?" *American Economic Review* 80 (2): 92–96.

Neary, Peter J. 2004. "Rationalizing the Penn World Table: True Multilateral Indices for International Comparisons of Real Income." *American Economic Review* 94 (5): 1411–28.

Obstfeld, Maurice, and Ken Rogoff. 1996. *Foundations of International Macroeconomics*. Cambridge, MA: MIT Press.

Officer, Lawrence H. 1982. *Purchasing Power Parity and Exchange Rates: Theory, Evidence and Relevance*. Greenwich, CT: JAI Press.

Reinsdorf, Marshall B. 2010. "Terms of Trade Effects: Theory And Measurement." *Review of Income and Wealth* 56 (S1): S177–S205.

Samuelson, Paul A. 1964. "Theoretical Notes on Trade Problems." *Review of Economics and Statistics* 46 (2): 145–54.

Samuelson, Paul A. 1994. "Facets of Balassa-Samuelson Thirty Years Later." *Review of International Economics* 2 (3): 201–26.

Summers, Robert, and Alan Heston. 1988. "A New Set of International Comparisons of Real Product and Price Levels: Estimates for 130 Countries, 1950–1985." *Review of Income and Wealth* 34 (1): 1–25.

Summers, Robert, and Alan Heston. 1991. "The Penn World Table (Mark 5): An Expended Set of International Comparisons, 1950–1988." *Quarterly Journal of Economics* 106 (2): 327–68.

Végh, Carlos A. 2014. *Open Economy Macroeconomics in Developing Countries*. Cambridge, MA: MIT Press.

World Bank. 2011. *The Changing Wealth of Nations: Measuring Sustainable Development in the New Millennium.* Washington, DC: World Bank.

World Bank. 2014. *Purchasing Power Parities and the Real Size of World Economies: A Comprehensive Report of the 2011 International Comparison Program*. Washington, DC: World Bank.

The Next Generation of the Penn World Table
Online Appendix

By Robert C. Feenstra, Robert Inklaar and Marcel P. Timmer

Appendix A: Proof of Theorems

Proof of Theorem 1

Because the outputs $\mathbf{y}_k$ are feasible for country *k,* but not optimal at the prices $\mathbf{P}_j$, it follows that $r_k(\mathbf{P}_j, \mathbf{v}_k) \geq \mathbf{P}_j'\mathbf{y}_k$. This establishes the first inequality below and the second is established similarly:[1]

$$\frac{r_j(\mathbf{P}_j, \mathbf{v}_j)}{r_k(\mathbf{P}_j, \mathbf{v}_k)} \leq \left(\frac{\mathbf{P}_j'\mathbf{y}_j}{\mathbf{P}_j'\mathbf{y}_k}\right) \text{ and } \frac{r_j(\mathbf{P}_k, \mathbf{v}_j)}{r_k(\mathbf{P}_k, \mathbf{v}_k)} \geq \left(\frac{\mathbf{P}_k'\mathbf{y}_j}{\mathbf{P}_k'\mathbf{y}_k}\right).$$

Using the Gerschenkron effect in (10) it follows that:

$$\frac{r_j(\mathbf{P}_j, \mathbf{v}_j)}{r_k(\mathbf{P}_j, \mathbf{v}_k)} \leq \left(\frac{\mathbf{P}_j'\mathbf{y}_j}{\mathbf{P}_j'\mathbf{y}_k}\right) < \left[\left(\frac{\mathbf{P}_j'\mathbf{y}_j}{\mathbf{P}_j'\mathbf{y}_k}\right)\left(\frac{\mathbf{P}_k'\mathbf{y}_j}{\mathbf{P}_k'\mathbf{y}_k}\right)\right]^{\frac{1}{2}} < \left(\frac{\mathbf{P}_k'\mathbf{y}_j}{\mathbf{P}_k'\mathbf{y}_k}\right) \leq \frac{r_j(\mathbf{P}_k, \mathbf{v}_j)}{r_k(\mathbf{P}_k, \mathbf{v}_k)}.$$

Now define $\mathbf{\Pi} = \lambda \mathbf{P}_j + (1-\lambda)\mathbf{P}_k$, and consider forming the ratio $r_j(\mathbf{\Pi}, \mathbf{v}_j)/r_k(\mathbf{\Pi}, \mathbf{v}_k)$. By construction, this ratio equals the right-hand side of the above expression when $\lambda = 0$ and the left-hand side when $\lambda = 1$. Therefore, there exists a value of $\lambda \in (0,1)$ at which this ratio equals the Fisher ideal price index in the middle of the above expression. QED

Proof of Corollary

This follows immediately from Theorem 1 by indexing country *j* with the time subscript *t*, and then treating country *k* as identical to country *j* in year *t*–1.

[1] These inequalities are due to Malmquist (1953), as noted by Diewert (2008).

Proof of Theorem 2

Let us write the vector of reference prices $\mathbf{\Pi}$ as $\mathbf{\Pi} = (\pi_1, \dots, \pi_N)'$ and the vector of endowments as $\mathbf{v}_j = (v_{1j}, \dots, v_{Lj})'$ for country *j*. Since the reference prices are equal in the revenue functions $r_j(\mathbf{\Pi}, \mathbf{v}_j)$ and $r_k(\mathbf{\Pi}, \mathbf{v}_k)$, we can be quite flexible about how they appear. In particular, suppose that $r_j(\mathbf{\Pi}, \mathbf{v}_j)$ is of the form:

$$\ln r_j(\mathbf{\Pi}, \mathbf{v}_j) = \ln h_j(\mathbf{\Pi}) + \sum_{l=1}^{L} \alpha_{lj} \ln v_{lj} + \sum_{l=1}^{L}\sum_{i=1}^{N} \beta_{lij} \ln v_{lj} \ln \pi_i$$

$$+ \frac{1}{2}\sum_{l=1}^{L}\sum_{m=1}^{L} \gamma_{lm} \ln v_{lj} \ln v_{mj}\,, \text{(A1)}$$

where $h_j(\mathbf{\Pi})$ is homogenous of degree one. The function $\ln r_k(\mathbf{\Pi}, \mathbf{v}_k)$ is specified similarly but with *k* replacing *j*.

Without loss of generality we can assume that $\gamma_{lm} = \gamma_{ml}$, and notice that only these second-order parameters do not depend on countries *j* or *k*. To interpret this restriction, note that we can multiply the endowments v_{lj} in (A1) by country-specific Harrod-neutral productivity terms λ_{lj}, which after simplification are absorbed into the parameters α_{lj} and also into $h_j(\mathbf{\Pi})$, but the parameters γ_{lm} on $\ln v_{lj} \ln v_{mj}$ remain unchanged. Very general types of country-specific sectorial productivity terms can be introduced through the parameters β_{lij} and through the function $h_j(\mathbf{\Pi})$. But the condition that the second-order parameters γ_{lm} do not differ across countries restricts the generality of factor-augmenting productivity differences between countries (beyond Harrod-neutral) that can be introduced.

Using the definition of the factor prices $w_{jl}^* = \partial r_j(\mathbf{\Pi}, \mathbf{v}_j)/\partial v_{jl}$ and the assumption that $r_j(\mathbf{\Pi}, \mathbf{v}_j)$ is homogeneous of degree one in endowments, we have $w_{lj}^* v_{lj} / \sum_m w_{mj}^* v_{mj} = \partial \ln r_j(\mathbf{\Pi}, \mathbf{v}_j)/\partial \ln v_{jl}$. It follows from (A1) that:

$$\frac{w_{lj}^* v_{lj}}{\sum_m w_{mj}^* v_{mj}} = \alpha_{lj} + \sum_{i=1}^{N} \beta_{lij} \ln \pi_i + \sum_{m=1}^{L} \gamma_{lm} \ln v_{mj}\,. \text{(A2)}$$

Then using the definition of A_j and A_k along with (A1) and (A2), we can compute:

$$
\begin{aligned}
\ln\left(A_j A_k\right)^{\frac{1}{2}} &= \frac{1}{2}\left[r_j\left(\boldsymbol{\Pi},\mathbf{v}_j\right) - r_k\left(\boldsymbol{\Pi},\mathbf{v}_j\right) + r_j\left(\boldsymbol{\Pi},\mathbf{v}_k\right) - r_k\left(\boldsymbol{\Pi},\mathbf{v}_k\right)\right] \\
&= \left[r_j\left(\boldsymbol{\Pi},\mathbf{v}_j\right) - r_k\left(\boldsymbol{\Pi},\mathbf{v}_k\right)\right] - \frac{1}{2}\left[r_j\left(\boldsymbol{\Pi},\mathbf{v}_j\right) + r_k\left(\boldsymbol{\Pi},\mathbf{v}_j\right) - r_j\left(\boldsymbol{\Pi},\mathbf{v}_k\right) - r_k\left(\boldsymbol{\Pi},\mathbf{v}_k\right)\right] \\
&= \left[r_j\left(\boldsymbol{\Pi},\mathbf{v}_j\right) - r_k\left(\boldsymbol{\Pi},\mathbf{v}_k\right)\right] \\
&\quad - \frac{1}{2}\left[\sum_{l=1}^{L}\left(\alpha_{lj} + \alpha_{lk} + \sum_{i=1}^{N}\left(\beta_{lij} + \beta_{lik}\right)\ln\pi_i\right)\ln\left(\frac{v_{lj}}{v_{lk}}\right) + \sum_{l=1}^{L}\sum_{m=1}^{L}\gamma_{lm}\ln v_{lj}\ln v_{mj}\right. \\
&\quad \left. - \sum_{l=1}^{L}\sum_{m=1}^{L}\gamma_{lm}\ln v_{lk}\ln v_{mk}\right] \\
&= \left[r_j\left(\boldsymbol{\Pi},\mathbf{v}_j\right) - r_k\left(\boldsymbol{\Pi},\mathbf{v}_k\right)\right] \\
&\quad - \frac{1}{2}\left[\sum_{l=1}^{L}\left(\alpha_{lj} + \alpha_{lk} + \sum_{i=1}^{N}\left(\beta_{lij} + \beta_{lik}\right)\ln\pi_i\right)\ln\left(\frac{v_{lj}}{v_{lk}}\right)\right. \\
&\quad \left. + \sum_{l=1}^{L}\sum_{m=1}^{L}\gamma_{lm}\left(\ln v_{mj} + \ln v_{mk}\right)\left(\ln v_{lj} - \ln v_{lk}\right)\right] \\
&= \left[r_j\left(\boldsymbol{\Pi},\mathbf{v}_j\right) - r_k\left(\boldsymbol{\Pi},\mathbf{v}_k\right)\right] - \frac{1}{2}\left[\sum_{l=1}^{L}\left(\frac{w_{lj}^{*}v_{lj}}{\sum_m w_{mj}^{*}v_{mj}} + \frac{w_{lk}^{*}v_{lk}}{\sum_m w_{mk}^{*}v_{mk}}\right)\ln\left(\frac{v_{lj}}{v_{lk}}\right)\right],
\end{aligned}
$$

where the first equality follows directly from the definition of A_j and A_k; the second equality follows from simple algebra; the third equality follows from the translog formula in (A1); the fourth equality follows from algebra on the double-summations that $\gamma_{lm} = \gamma_{ml}$; and the final equality follows from the share formula in (A2). QED

Appendix B: PWT Variables and Formulas

Detailed PPP estimation

We start from the following three datasets:

(i) Benchmark data from ICP, OECD and Eurostat, providing data for detailed consumption and investment products[2] on prices p_{ijt}^{k} and quantities[3] q_{ijt}^{k} for product *i*, country *j*, year *t* and with *k* equal to household consumption *C*,

[2] What we refer to as 'products' are typically referred to as 'basic headings' in the benchmark data.

[3] The relative quantity of any variable is obtained by dividing its relative expenditure by its relative price.

government consumption G or investment. Prices are given relative to those in the United States and expenditures are in current national prices.

(ii) Data for Broad Economic categories (BEC) of exports and imports from Feenstra and Romalis (2014). Prices p_{ijt}^{k} and quantities k_{ijt} have been aggregated from 4-digit Standard Industrial Trade Classification (SITC) categories to six, one-digit BEC categories i, country j, year t and with k equal to exports x or imports m. Prices are expressed relative to the United States and expenditures are in current national prices.

(iii) National Accounts data of expenditure and price deflators $P_{jt}^{NA,k}$. Data are only available for the spending categories $k = C, I, G, X, M$ in country j at time t. Expenditures are in current national prices and the price deflators are relative to a benchmark year with 2005 = 1.

In the first step, we aggregate the benchmark price data as to obtain price indexes for overall household consumption, government consumption and investment using a GEKS procedure. For a given year t, define the Fisher price index for spending category k between country j and country h as:

$$P_{jh}^{F,k} \equiv \left[\left(\frac{\mathbf{p}_j'^k \mathbf{q}_j^k}{\mathbf{p}_h'^k \mathbf{q}_j^k}\right)\left(\frac{\mathbf{p}_j'^k \mathbf{q}_h^k}{\mathbf{p}_h'^k \mathbf{q}_h^k}\right)\right]^{\frac{1}{2}} . \text{(B1)}$$

Then the GEKS price index for spending category k between country j and country h, computed over a set of countries $c = 1, \dots, C$ to obtain a transitive comparison, is:

$$P_{jh}^{G,k} \equiv \prod_{c=1}^{C} \left(P_{jc}^{F,k} P_{ch}^{F,k}\right)^{1/C} . \text{(B2)}$$

For every benchmark comparison, we apply equation (B2) to get price indexes for C, I and G. These are all initially scaled to with US=1, but for easier comparison of magnitudes over time, we multiply these by the US deflator $P_{USA,t}^{NA,k}$. Denote the resulting price indexes, with USA in 2005=1, as $P_{jt}^{G,k}$. For exports and imports, we use the price indexes from Feenstra and Romalis (2014) defined at the level of six, one-digit Broad Economic Categories (BEC) i, such as fuel or consumer goods. These have also been aggregated from the 4-digit SITC data using a GEKS procedure. Multiplying these by US price deflators gives the corresponding $P_{jt}^{G,k}$.[4]

Since the ICP benchmark comparisons have been done in only six years – 1970, 1975, 1980, 1985, 1996 and 2005 – and not all countries participated in every comparison, we have a dataset for 167 countries[5] and comparatively few observations per country. To get a complete set of price indexes, we interpolate between benchmarks and extrapolate for years before the first and after the last available benchmark. This procedure is illustrated schematically in Figure C1.

For a hypothetical country that participated only in the 1996 and 2005 benchmarks, we use interpolated price indexes for the 1997-2004 period, while the years up to 1995 and from 2006 onwards are based on extrapolated price indexes. Extrapolation to year *t*–1 if price index data for year *t* are available is done using the price deflators from the National Accounts:

$$P_{jt-1}^{G,k} = P_{jt}^{G,k} \times P_{jt-1}^{NA,k} / P_{jt}^{NA,k} \,. \text{(B3)}$$

Interpolation takes the benchmark price indexes as given but uses the pattern of inflation in intervening years to determine how much of the overall change in price indexes between benchmarks should fall in any given year. So for the hypothetical country of Figure B1, this would be:

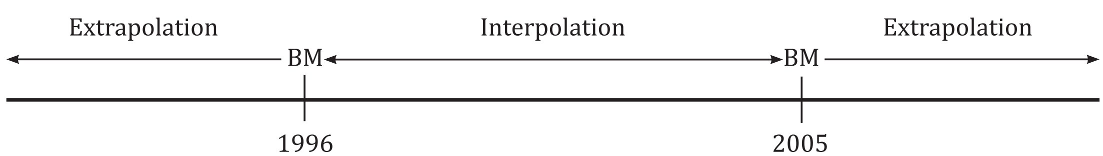

FIGURE B1: SCHEMATIC ILLUSTRATION OF A HYPOTHETICAL PRICE INDEX COMPUTATION

Note: BM stands for 'benchmark'. This hypothetical country participated only in the 1996 and 2005 benchmarks.

$$P_{jt}^{G,k} = P_{j1996}^{G,k} \times \frac{P_{jt}^{NA,k}}{P_{j1996}^{NA,k}} \left(\frac{2005 - t}{2005 - 1996} \right) + P_{j2005}^{G,k} \times \frac{P_{jt}^{NA,k}}{P_{j2005}^{NA,k}} \left(\frac{t - 1996}{2005 - 1996} \right) . \text{(B4)}$$

This approach to interpolation and extrapolation is similar in spirit to the approach of Rao *et al* (2010), who also discuss a method for estimating price indexes for a full set of years and countries using benchmark price indexes and National Accounts deflators. The key distinction is that we

[4] These still have subscript *i* for each BEC.

[5] This is the full set of (currently existing) countries that participated in at least one ICP comparison.

always force the price series to be equal to the benchmark estimates, while this is a special case of Rao *et al* (2010); see also Hill (2004).

Real GDP computation

Given price indexes for the major expenditure categories and expenditures on these categories from the National Accounts, we estimate real GDP. To simplify notation compared to the previous discussion, let us denote for each country *j* and time period *t* the GEKS price index $p_{ijt} \equiv P_{jt}^{G,i}$ for $i = C, I, G$, while the GEKS prices indexes for export and imports are denoted by p_{ijt}^{x} and p_{ijt}^{m}, respectively, across the $i = 1, \ldots, 6$ BEC categories. We therefore have a set of 15 'products' for every country and year: three for domestic final expenditure *C, I* and *G*, six for exports, *X* and six for imports, *M*. Using the Geary-Khamis system, we define the following set of reference prices:

$$\pi_{it} = \sum_{j=1}^{C} \left(p_{ijt}/PPP_{jt}^{o}\right) q_{ijt} \Big/ \sum_{j=1}^{C} q_{ijt}\text{, for } i = C, I, G,$$

$$\pi_{it}^{x} = \sum_{j=1}^{C} \left(p_{ijt}^{x}/PPP_{jt}^{o}\right) x_{ijt} \Big/ \sum_{j=1}^{C} x_{ijt}\text{, for } i = 1, \ldots, 6, \text{(B5)}$$

$$\pi_{it}^{x} = \sum_{j=1}^{C} \left(p_{ijt}^{x}/PPP_{jt}^{o}\right) x_{ijt} \Big/ \sum_{j=1}^{C} x_{ijt}\text{, for } i = 1, \ldots, 6,$$

where the PPP for each country is defined using nominal GDP_{jt} as:

$$PPP_{jt}^{o} = \frac{GDP_{jt}}{\sum_i \pi_{it} q_{ijt} + \sum_i \pi_{it}^{x} x_{ijt} - \sum_i \pi_{it}^{m} m_{ijt}}. \text{(B6)}$$

The system of equations (B5) and (B6) can be solved recursively and has a unique solution up to a normalization, as shown in Feenstra *et al* (2009). The normalization that we choose is $PPP_{USA,t}^{o} = P_{USA,t}^{NA,GDP}$, so that the PPP for the USA is equal to the GDP deflator (and 2005=1). Whether the solution to the system is also positive is less clear *a priori*, because imports (products 10-15) are subtracted in the numerator and denominator. Only if trade shares are not too large is it possible to rule out negative solutions (see Feenstra *et al*, 2009).

One complication is that the 15 products for which we have prices and quantities do not make up total GDP. Our prices of exports and imports only cover merchandise trade, but net trade in services also contributes to total GDP.[6] Furthermore, in some countries there is a statistical discrepancy between total expenditure $C+I+G+X-M$ and total GDP. For example, India defines total GDP from the income side and total expenditure may be larger or smaller than GDP. We therefore define R_{jt} as residual component of GDP.

Given the set of reference prices and PPPs in every year, most of the variables in PWT can be computed. Table B1 shows these variables and corresponding formulas in logical order, so that previously defined variables can be used in subsequent formulas. The remaining computed variables, relating to factor inputs and productivity, are discussed in more detail in Appendix C.

---

[6] Note the in some countries, merchandise exports or imports are larger than National Accounts total exports or imports. In those cases, we use National Accounts trade numbers and assuming net trade in services is zero.

TABLE B1: PWT VARIABLES AND FORMULAS

PWT variable	Formula
$CGDP^o_{jt}$	$\sum_{i=C,I,G}\pi_{it}q_{ijt}+\sum_{i=1}^{6}\pi^x_{it}x_{ijt}-\sum_{i=1}^{6}\pi^m_{it}m_{ijt}+\frac{R_{jt}}{PPP^o_{jt}}\equiv\frac{GDP_{jt}}{PPP^o_{jt}}$
$PL\_GDP^o_{jt}$	$PPP^o_{jt}/\varepsilon_{jt}$
$PL\_GDP^e_{jt}$	$\left(\sum_{i=C,I,G}p_{ijt}q_{ijt}\Big/\sum_{i=C,I,G}\pi_{it}q_{ijt}\right)\Big/\varepsilon_{jt}\equiv PPP^q_{jt}/\varepsilon_{jt}$
$CCON_{jt}$	$\sum_{i=C,G}p_{ijt}q_{ijt}\Big/PPP^q_{jt}$
CDA_{jt}	$\sum_{i=C,I,G}p_{ijt}q_{ijt}\Big/PPP^q_{jt}$
$CGDP^e_{jt}$	GDP_{jt}/PPP^q_{jt}
$\frac{RGDP^o_{jt}}{RGDP^o_{jt-1}}$	$\left[\left(\frac{\sum_{i=C,I,G}\pi_{it-1}q_{ijt}+\sum_{i=1}^{6}\pi^x_{it-1}x_{ijt}-\sum_{i=1}^{6}\pi^m_{it-1}m_{ijt}}{\sum_{i=C,I,G}\pi_{it-1}q_{ijt-1}+\sum_{i=1}^{6}\pi^x_{it-1}x_{ijt-1}-\sum_{i=1}^{6}\pi^m_{it-1}m_{ijt-1}}\right)\times\left(\frac{\sum_{i=C,I,G}\pi_{it}q_{ijt}+\sum_{i=1}^{6}\pi^x_{it}x_{ijt}-\sum_{i=1}^{6}\pi^m_{it}m_{ijt}}{\sum_{i=C,I,G}\pi_{it}q_{ijt-1}+\sum_{i=1}^{6}\pi^x_{it}x_{ijt-1}-\sum_{i=1}^{6}\pi^m_{it}m_{ijt-1}}\right)\right]^{1/2}$
$\frac{RGDP^e_{jt}}{RGDP^e_{jt-1}}$	$\left[\left(\frac{\sum_{i=C,I,G}\pi_{it-1}q_{ijt}+(X_{jt}-M_{jt})/PPP^q_{jt}}{\sum_{i=C,I,G}\pi_{it-1}q_{ijt-1}+(X_{jt-1}-M_{jt-1})/PPP^q_{jt-1}}\right)\times\left(\frac{\sum_{i=C,I,G}\pi_{it}q_{ijt}+(X_{jt}-M_{jt})/PPP^q_{jt}}{\sum_{i=C,I,G}\pi_{it}q_{ijt-1}+(X_{jt}-M_{jt})/PPP^q_{jt}}\right)\right]^{1/2}$
CSH_i_{jt}	$\pi_{it}q_{ijt}/CGDP^o_{jt}$, $i=C,I,G$
CSH_X_{jt}	$\sum_{i=1}^{6}\pi^x_{it}x_{ijt}\Big/CGDP^o_{jt}$
CSH_M_{jt}	$-\sum_{i=1}^{6}\pi^m_{it}m_{ijt}\Big/CGDP^o_{jt}$
CSH_R_{jt}	$\left(R_{jt}/PPP^o_{jt}\right)/CGDP^o_{jt}$
PL_i_{jt}	$\left(p_{ijt}q_{ijt}/\pi_{it}q_{ijt}\right)/\varepsilon_{jt}$, $i=C,I,G$
PL_X_{jt}	$\left(\sum_{i=1}^{6}p^x_{ijt}x_{ijt}\Big/\sum_{i=1}^{6}\pi^x_{it}x_{ijt}\right)/\varepsilon_{jt}$
PL_M_{jt}	$\left(\sum_{i=1}^{6}p^m_{ijt}m_{ijt}\Big/\sum_{i=1}^{6}\pi^m_{it}m_{ijt}\right)/\varepsilon_{jt}$

Appendix C: Capital, labor and productivity

The productivity level in country j relative to country k was defined prior to Theorem 2:

$$(A_j A_k)^{\frac{1}{2}} = \frac{r_j(\mathbf{\Pi},\mathbf{v}_j)}{r_k(\mathbf{\Pi},\mathbf{v}_k)} / Q_T(\mathbf{v}_j,\mathbf{v}_k,\mathbf{w}_j^*,\mathbf{w}_k^*), \tag{C1}$$

where $Q_T(\mathbf{v}_j,\mathbf{v}_k,\mathbf{w}_j^*,\mathbf{w}_k^*)$ is the Törnqvist index of factor endowments. In this appendix we describe the measurement of the factor endowments and factor shares. For that, it is helpful to write equation (C1) in terms of PWT variables:

$$\frac{CTFP_{jt}}{CTFP_{USAt}} = \frac{CGDP^o_{jt}}{CGDP^o_{USAt}} \Big/ Q_{j,USA,t}\,, \tag{C2}$$

where,

$$\begin{aligned} Q_{j,USA,t} &= \tfrac{1}{2}\left(LABSH_{jt} + LABSH_{USAt}\right)\left(\frac{EMP_{jt}}{EMP_{USAt}}\frac{HC_{jt}}{HC_{USAt}}\right) \\ &\quad + \left[1-\tfrac{1}{2}\left(LABSH_{jt} + LABSH_{USAt}\right)\right]\left(\frac{CK_{jt}}{CK_{USAt}}\right) \end{aligned}, \tag{C3}$$

and where $LABSH_{jt}$ denotes the share of labor in total factor income in country j and year t, EMP_{jt} denotes employment, and HC_{jt} denotes a measure of human capital based on the average years of schooling from Barro and Lee (2012). Similarly, we define productivity growth over time as:

$$\frac{RTFP^{NA}_{jt}}{RTFP^{NA}_{jt-1}} = \frac{RGDP^{NA}_{jt}}{RGDP^{NA}_{jt-1}} \Big/ Q_{jt,t-1}\,, \tag{C4}$$

where,

$$\begin{aligned} Q_{j,t,t-1} &= \tfrac{1}{2}\left(LABSH_{jt} + LABSH_{jt-1}\right)\left(\frac{EMP_{jt}}{EMP_{jt-1}}\frac{HC_{jt}}{HC_{jt-1}}\right) \\ &\quad + \left[1-\tfrac{1}{2}\left(LABSH_{jt} + LABSH_{jt-1}\right)\right]\left(\frac{RK^{NA}_{jt}}{RK^{NA}_{jt-1}}\right) \end{aligned}. \tag{C5}$$

The remainder of this appendix describes how the variables CK_{jt}, RK^{NA}_{jt}, and $LABSH_{jt}$ are computed, drawn from Inklaar and Timmer (2013).

Capital stocks

Capital stocks are estimated based on cumulating and depreciation past investments using the perpetual inventory method (PIM). This section first discusses how investment by asset is estimated. Given the long-lived nature of many assets, it is important to start the PIM from an initial capital stock and the method used to estimate these is discussed next. Finally, we show the implications of the more detailed investment data for cross-country depreciation patterns and the relative capital stock levels.

Investment at current and constant prices

There is a wide range of assets in which firms (and governments) can invest in and these tend to have widely varying asset life spans. A common shortcut method is to ignore this heterogeneity and estimate capital input based on a common and constant assumed asset life. But this ignores important changes in investment composition over time and differences across countries. The work of Caselli and Wilson (2004), for instance, shows there are considerable differences in the composition of investment across countries: for example, richer countries tend to invest more in computers.

For PWT8, we develop a dataset with investment in up to six assets, shown in Table C1 with their geometric depreciation rates. These rates are assumed to be common across countries and constant over time. As the breakdown by asset is not readily available for all countries, we use a variety of sources in compiling the investment data.

TABLE C1, ASSETS COVERED AND GEOMETRIC DEPRECIATION RATES

Asset	Depreciation rate (in percent)
Structures (residential and non-residential)	2
Transport equipment	18.9
Computers	31.5
Communication equipment	11.5
Software	31.5
Other machinery and assets	12.6

Notes: depreciation rates are based on official BEA deprecation rates of Fraumeni (1997).

We first distinguish structures, transport equipment and machinery. We do this based on OECD National Accounts, country National Accounts, EU KLEMS (www.euklems.org) and ECLAC National Accounts (Economic Commission for Latin America and the Caribbean). That still leaves many countries with incomplete data, so we use the commodity-flow method (CFM) whereby investment in an asset is assumed to vary with the economy-wide supply (production + imports - exports) of that asset. This approach has also been used by Caselli and Wilson (2004), though without the constraint that investment had to add up to gross fixed capital formation in the National Accounts and without covering investment in structures. The CFM method uses data on value added in the construction industry from the UN National Accounts Main Aggregates Database; imports and exports of equipment from UN Comtrade and Feenstra's World Trade Flows database; and industrial production from UNIDO. The detailed expenditure data underlying the ICP PPP benchmarks is used to fix investment shares at different points in time and the CFM is used to measure the trends over time. In a second step, we use data from EU KLEMS and from The Conference Board on information and communication technology (ICT) investment and WITSA on ICT expenditure to split up investment in machinery into investment in computers, communication equipment, software and other machinery. This second step can only be done for a subset of countries as ICT investment data is not available for all countries. This provides us with a dataset on investment at current national prices.

For data on investment prices over time, we use EU KLEMS, OECD National Accounts, ECLAC or UN National Accounts. This last source only provides a deflator for overall investment, which is most obviously problematic for ICT assets that have shown rapidly declining prices in countries with enough data, such as the US. For ICT assets, we therefore assume that the US price trend also applies to countries for which we have no specific data from other sources, with an adjustment made for overall inflation using the GDP deflator. For many countries, though, only the total investment deflator is used for non-ICT assets.

Initial capital stocks

We have very long time series of investment, back to 1950 for numerous countries, but to also provide good estimates in earlier years, we have to start from an initial capital stock. We have chosen to apply a harmonized procedure for all countries. Based on our choice of an initial capital stock, we then estimate capital stocks using the perpetual inventory method. The choice for an initial capital stock procedure is a consequential one, in particular because structures have such long asset lives, and thus low depreciation rates. With a 2 percent annual depreciation rate and investment data since 1950, almost 30 percent of the 1950 capital stock is still in use in 2011.[7] For countries with data since 1990, such as the former Soviet republics, the surviving fraction is almost two-thirds, so the procedure used for estimating the initial capital stock is important.

[7] Calculated as $(0.98)^{61}=0.29$.

Nehru and Dhareshwar (1993) discuss a number of alternatives for estimating this initial capital stock, including production function estimates and choosing an initial capital/output ratio. Their preferred approach, originally proposed by Harberger (1978), is to use the steady-state relationship from the Solow growth model:

$$K_0 = \frac{I_0}{g+\delta} \ . \quad \text{(C6)}$$

The initial capital stock K_0 for an asset is related to investment in the initial year, the (steady-state) growth rate of investment g and the depreciation rate δ. This requires the strong assumption that all economies were in a steady state in the first year for which data is available and that a reasonable steady-state growth rate of investment can be identified. Harberger (1978) initially chose a three-year average, while Nehru and Dhareshwar (1993) (effectively) use the average growth between 1950 and 1973 and Caselli (2005) uses the average growth until 1970 (which means a 10 to 20-year average growth rate given his selection of countries with data since at least 1960).

Nehru and Dhareshwar (1993) were dissatisfied with assuming an initial capital/output ratio, for reasons they do not spell out in detail. However, we argue that this method actually leads to superior results, in particular in early years of the sample and in transition economies, where the data is available for a limited period of time and where the early years were particularly turbulent. Under this approach, the initial capital stock is estimated as:

$$K_0 = Y_0 \times k, \quad \text{(C7)}$$

where Y_0 is GDP in the initial year and k the assumed capital/output ratio K/Y. To motivate the choice for k, Figure C1 plots capital/output ratios in 2005, where capital is summed over all assets, against GDP per capita. The figure includes the 142 countries with investment data since at least 1970. The figure shows considerable variation in capital/output ratios around a median value of 2.7. The least squares regression line indicates that there is no systematic relationship between GDP per capita and capital/output ratios.

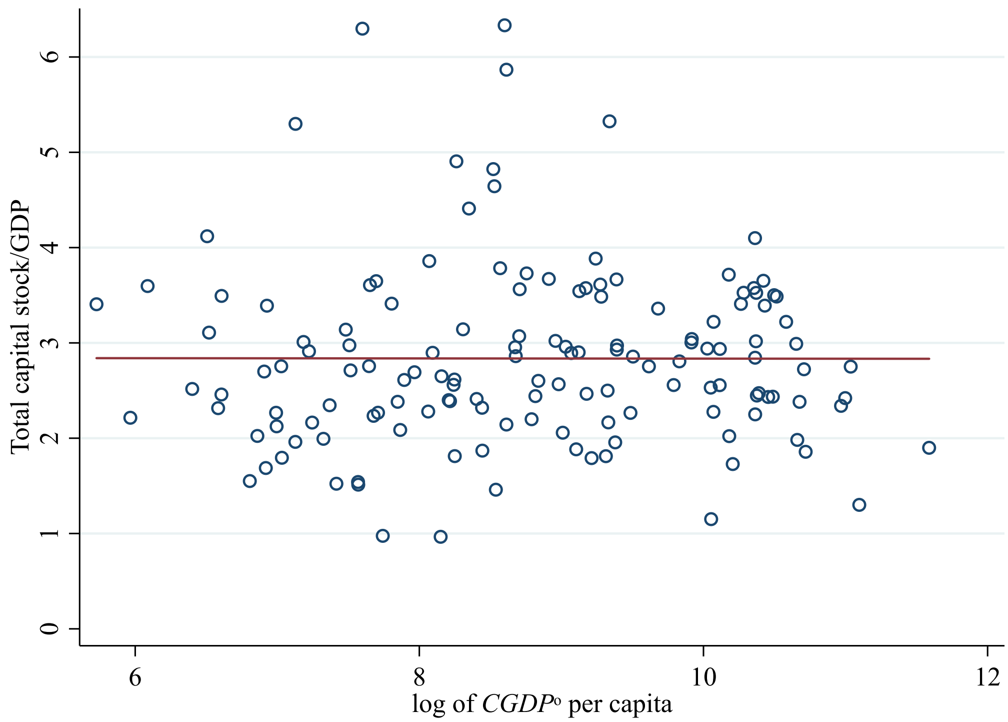

FIGURE C1: CAPITAL/OUTPUT RATIO VERSUS LOG GDP PER CAPITA, IN 2005

Source: PWT8.1

Note: Capital/output ratios are measured using data at current national prices. The plot excludes countries with investment data that start in a year later than 1970.

To solidify this finding, Table C2 shows a number of regressions, aiming to explain variation in capital/output ratios with GDP per capita. Note that the capital/output ratio is at current national prices, so does not reflect differences in inflation or relative prices of the capital stock versus the price of GDP. Column (1) shows the regression on the data in Figure C1, so with data for 2005 and including only countries with investment data since at least 1970. Column (2) includes all countries and years and shows a similarly insignificant relationship between the GDP per capita level and the capital/output ratio.

Columns (3) through (8) analyze the separate components of the total capital of setting an initial capital stock, these relationships are less relevant to account for because the asset life of machinery and communication equipment is much shorter and because the use of computers and software only became widespread since the 1960s in the US and later in other countries. Furthermore, structures account for, on average, 80 percent of the value of the capital stock, so its initial stock will have the most impact on the overall results.

TABLE C2: THE RELATIONSHIP BETWEEN CAPITAL/OUTPUT RATIOS AND GDP PER CAPITA

	(1) 2005 since 1970	(2) Total stock Full sample	(3) Structures	(4) Machinery	(5) Transport equipment	(6) Computers	(7) Communication equipment	(8) Software
log *CGDP^o*/capita	-0.00111	0.0455	0.00373	0.0288***	-0.000838	0.00476***	-0.0136***	0.00673***
	(0.0516)	(0.0439)	(0.0403)	(0.00942)	(0.00485)	(0.000720)	(0.00462)	(0.000724)
Constant	2.846***	2.521***	2.335***	0.113	0.149***	-0.0292***	0.196***	-0.0500***
	(0.479)	(0.399)	(0.368)	(0.0820)	(0.0430)	(0.00644)	(0.0427)	(0.00638)
Observations	142	8268	8268	8268	8268	3293	3293	3293
R-squared	0.000	0.002	0.000	0.029	0.000	0.160	0.029	0.329

Notes: dependent variable is the capital/output ratio, measured using data at current national prices. Robust standard errors are in parentheses; in columns (2)-(8), errors are clustered by country. Column (1) only includes data for 2005 for countries with investment data since at least 1970. The remaining columns include all countries and years in PWT8.1. The number of observations in column (6)-(8) is lower than in (2)-(5) because there is no data on investment in these assets for a range of countries. *** p<0.01, ** p<0.05, * p<0.1

Table C3 shows the initial capital/output ratios (*k* in equation C7) that we assume for all countries for the non-ICT assets, based on the cross-country medians that we observe in the data over the full period. Given their short asset lives and relatively small share in total assets we set initial ICT stocks equal to zero. There is a circularity in setting initial capital stocks based on capital/output ratios that are computed based on those initial stocks. However, if we use equation (C7) to estimate initial stocks, the same median ratios are obtained.

TABLE C3: INITIAL CAPITAL/OUTPUT RATIOS FOR NON-ICT ASSETS

Asset	Capital/output ratio *k*
Structures (residential and non-residential)	2.2
Transport equipment	0.1
Other machinery and assets	0.3
Total	2.6

Note: initial capital/output ratios for ICT assets are set at zero.

Based on these choices, we can contrast the results based on assuming an initial capital/output ratio to results based on the more commonly-used 'steady-state' method described in equation (C6). The average growth of investment for the first ten years of the sample period is used in that equation, but the results are similar if the first five years of data are used.[8] Table C4 shows descriptive statistics for the year 2005, comparing capital/output ratios based on assuming an initial capital/output ratio (K/Y) and based on assuming a steady-state capital level (SS). The first row compares the results for all countries and shows that the median capital/output ratio is very similar

8 The main difference is in the cross-country variation, which is much higher if only the first five years of investment data is used.

across the two approaches, but the variation is much larger using the steady-state approach. Finally, the correlation is high, but at 0.73 far from perfect.

TABLE C4: COMPARING CAPITAL/OUTPUT RATIOS IN 2005: INITIAL RATIO VS. STEADY-STATE ASSUMPTION

	Median		Standard deviation		Correlation
	K/Y	Steady State	K/Y	Steady State	(K/Y, SS)
All countries (167)	2.84	2.78	1.28	2.09	0.74
Investment data starting in:					
1950-1959 (73)	2.65	2.65	0.66	0.68	0.99
1960-1970 (69)	2.74	2.65	1.17	1.23	0.97
1988-1990 (25)	3.64	5.33	2.17	3.73	0.55

Notes: K/Y indicates that an initial capital/output ratio is assumed for the first year in which data is available; SS indicates the steady-state capital stock based on equation (C6) is used.

The next rows split the full sample of 167 countries by the length of the investment series. More than 70 countries have investment time series since before 1950, and another 69 countries have time series since 1970, while the final group of 25 countries has investment data for less than 25 years. In the first two groups of countries the correlation is nearly perfect and the median and standard deviation are very similar. It is in this final group that the largest differences can be seen: the steady-state approach leads to a median capital/output ratio that is much higher, variation that is much higher, and the correlation between the two approaches is quite low at 0.53. The countries in this last group are nearly all countries that emerged from the former Soviet Union, Yugoslavia and Czechoslovakia. In those newly-formed countries, the early 1990s were anything but a steady-state, involving a transition to market-based economies. Using the steady-state approach implies very high capital/output ratios, especially early in the transition period. This is because these countries saw rapidly falling GDP in the first years of their transition and thus a large increase in their capital/output ratio. To illustrate, in 1995 the steady-state approach implies a capital/output ratio in the Czech Republic of 6.5 and in Slovakia of 8.9, while in Poland and Hungary (transition countries with longer time series), the ratios are 3.4 and 3.7. In contrast, if the assumption of initial capital/output ratio is used, the 1995 capital/output ratio in the four countries varies between 3.7 and 3.9.

Figure C2 shows that differences between the initial K/Y ratio and its steady state value are much larger in earlier years. The figure shows the median capital/output ratio for the 142 countries with investment data since 1970 according to the two approaches. As the figure illustrates, from the mid-1980s onwards, the median capital/output ratio fluctuates between 2.5 and 3.1 according to both approaches, which suggests this is the standard range for the capital/output ratio. Assuming an initial capital/output ratio ensures that the data for these countries are continuously within this range, but applying the steady-state assumption implies that capital levels are much too low for a long period of time, starting at a median level of 1.9 and reaching 2.5 only in 1980. We therefore conclude that assuming the initial capital/output ratios from Table C3 ensures more plausible capital stocks across all countries and years.

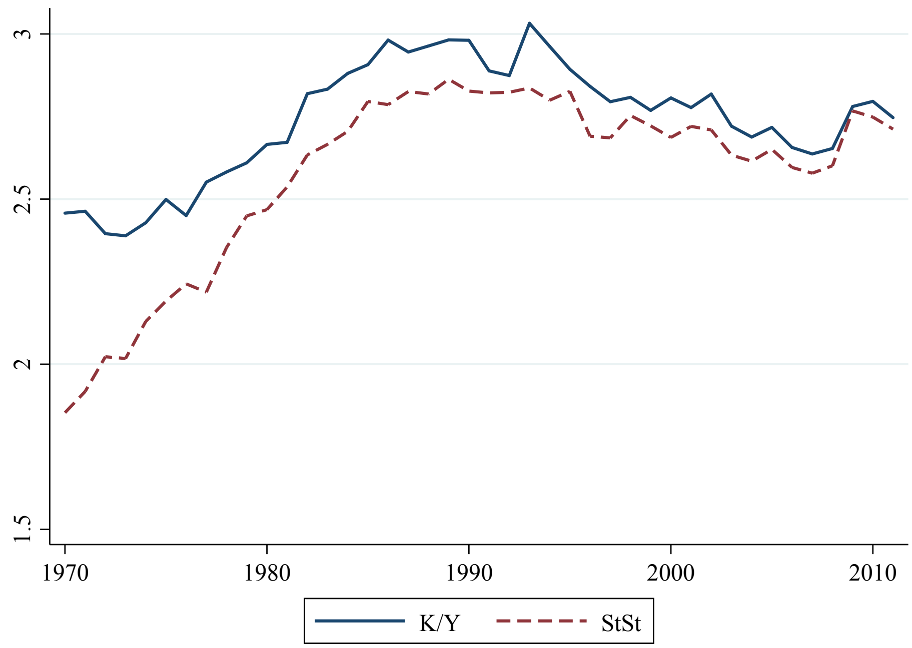

FIGURE C2: COMPARING CAPITAL/OUTPUT RATIOS OVER TIME: INITIAL RATIO VS. STEADY-STATE ASSUMPTION, 1970-2011

Source: PWT8.1

Notes: K/Y indicates that an initial capital/output ratio is assumed for the first year in which data is available; StSt indicates the steady-state capital stock based on equation (C6) is used. For each year, the median, standard deviation and correlation is computed for the cross-section of 142 countries with investment data since 1970 or earlier.

Capital stocks and depreciation

Given an initial capital stock K_0, investment at constant prices I and depreciation rates δ, it is straightforward to compute capital stocks for asset a in country i at time t using the Perpetual Inventory Method (PIM): $K_{ait} = (1 - \delta_a)K_{ait-1} + I_{ait}$. Multiplying this capital stock by the asset deflator P_{ait} then gives capital stocks at current national prices. Compared with the typical approach in the literature, the main benefit is that the assumption of a single depreciation rate for all countries and years is no longer necessary, since the asset composition of investment varies across countries and years.

Figure C3 illustrates the depreciation rates of the total capital stock that we get as a result of this approach.[9] The figure plots the depreciation rate in 2005 against GDP per capita and the least-squares regression line. The slope is not significantly different from zero in this year, but across all years, higher income countries tend to have higher depreciation rates. This fits with the finding from Table C2 that richer countries have higher capital/output ratios for assets with high depreciation rates: machinery, computers and software. So, for example, the U.S. had a depreciation rate in 2011 of 4.1 percent, while China had a rate of 3.1 percent. Since the capital stocks of richer countries are depreciating at a more rapid rate, the capital stock levels we estimate here will be relatively lower compared to capital stocks estimated based on a common depreciation rate across countries.

Capital stock at constant national prices

With capital stocks constructed for each of the assets, we construct a total capital stock at constant national prices, RK^{NA}. Ideally, this would be a measure of capital services, not capital stocks. A capital services measure would reflect that shorter-lived assets have a larger return in production, as indicated by the user cost of capital of each asset. As a result, buildings, which represent on average 80 percent of the capital stock at current prices, would represent a much lower share of capital services. However, the data requirements for estimating capital services are higher than for a capital stock measure. In particular, the user cost of capital of an asset should include, alongside the depreciation rate, a required rate of return on capital and the rate of asset price inflation. Asset-specific inflation rates are not available for many countries, as mentioned above, and the required rate of return on capital is generally hard to measure well (see e.g. Inklaar, 2010). Furthermore, in countries that have experienced periods of extreme inflation in the past, any measurement error in either inflation or the rate of return would lead to substantial swings in the user cost of capital. Finally, the user cost would be needed for comparisons over time but also across countries. This implies that mismeasurement of user costs in one country would affect capital input estimates for other countries as well.

We therefore use the total capital stock as our measure of capital input. The RK^{NA} variable is constructed as a Törnqvist aggregate of the individual asset growth rates:

[9] The depreciation rate of the total capital stock is computed as $\delta_{it} = \sum_a P_{ait}\delta_a K_{ait-1} / \sum_a P_{ait} K_{ait-1}$.

$$\Delta \log RK_{it}^{NA} = \sum_{a} \bar{v}_{ait} \Delta \log K_{ait} \,, \quad \text{(C8)}$$

with $\bar{v}_{ait} = \frac{1}{2}(v_{ait-1} + v_{ait})$ and $v_{ait} = P_{ait}K_{ait} / \sum_a P_{ait}K_{ait}$. So the growth of the capital stock at constant national prices for each assets is weighted by its two-period average share in the capital stock at current national prices. Equation (C8) defines the growth rate of RK^{NA}, with 2005=1.

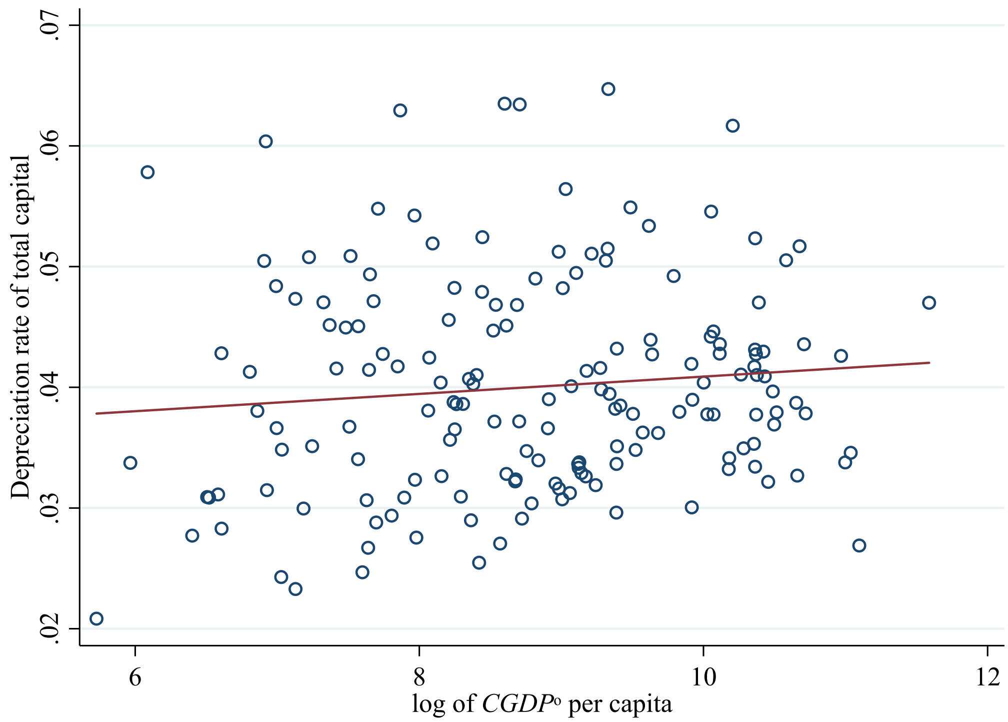

FIGURE C3: DEPRECIATION RATE OF THE TOTAL CAPITAL STOCK AND GDP PER CAPITA IN 2005

Capital stock at current reference prices

The computation of the capital stock at current (reference) prices involves converting the capital stock at current national prices using a price level for the capital stock. Specifically, in terms of PWT variables, the current-price capital stock is defined as:

$$CK_{it} = \frac{\sum_a P_{ait}K_{ait}}{\mathcal{E}_{it} PL_{K_{it}}}, \text{(C9)}$$

where K_{ait} is the capital stock of asset a at national prices and $\mathcal{E}_{it}$ is the nominal exchange rate relative to the US dollar. The relative price of the capital stock PL_K_{it} is computed from investment prices for each of the assets – denoted by the vector $\mathbf{P}_{it}^{I}$ – while using the capital stocks at current national prices as weights and using the U.S. as the comparison country 0:

$$PL_{K_{it}} = \prod_{j=1}^{C} \left(P_{ijt}^{F} P_{j0t}^{F}\right)^{\frac{1}{C}} \Big/ \mathcal{E}_{it}\,, \text{where } P_{ijt}^{F} = \left[\left(\frac{\mathbf{P}_{it}'^{I}\mathbf{K}_{it}}{\mathbf{P}_{jt}'^{I}\mathbf{K}_{it}}\right)\left(\frac{\mathbf{P}_{it}'^{I}\mathbf{K}_{jt}}{P_{jt}'^{I}\mathbf{K}_{jt}}\right)\right]^{\frac{1}{2}}. \text{(C10)}$$

The prices for each asset are based on PPP benchmark surveys: the six ICP surveys since 1970 and the more frequent surveys by the OECD and Eurostat since 1995. The prices from these surveys do not directly map into the six assets we use. In some surveys, there would not be sufficiently detailed data to separately distinguish each of the six assets. At a minimum, it is always possible to distinguish investment in structures from investment in machinery and equipment and usually this latter category is also split between transport equipment and other machinery. In case of missing detail, the more aggregate price is applied to each of the detailed assets, such as the machinery price for ICT assets and for other machinery. There would also sometimes be more detailed PPPs than the six we need and in those cases, we use a GEKS procedure with investments from the surveys as weights to arrive at the required six asset prices. We follow the same procedure as discussed in Appendix C to estimate a full panel of investment prices: using benchmark prices where available, interpolating between benchmarks and extrapolating if no earlier or later benchmark is available.

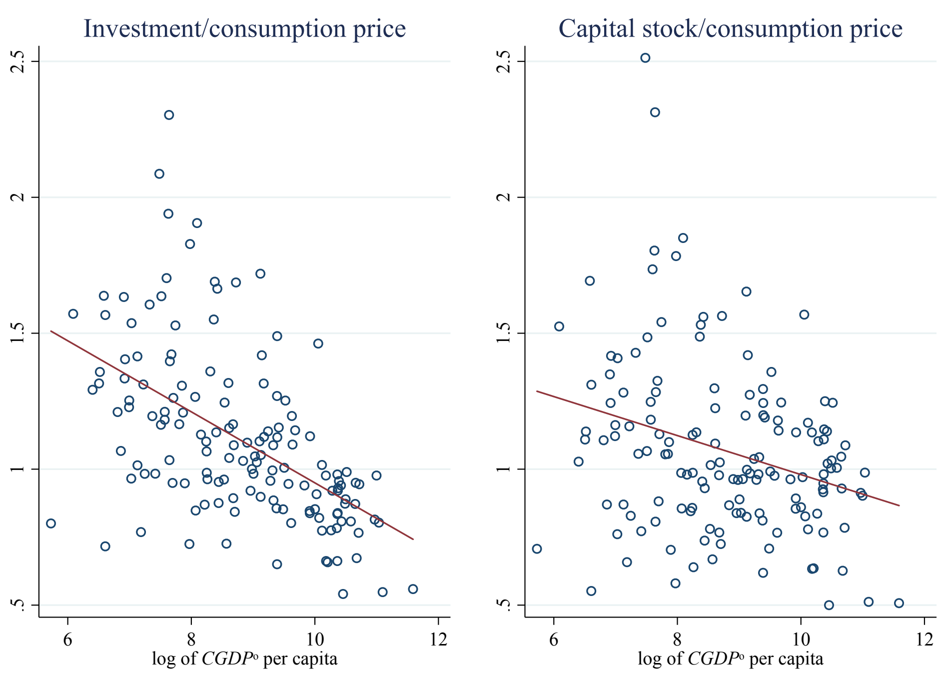

FIGURE C4: THE INVESTMENT AND CAPITAL STOCK PRICE RELATIVE TO THE HOUSEHOLD CONSUMPTION PRICE LEVEL IN 2005

Notes: only countries that participated in the 2005 ICP benchmark survey are included. The price level of investment goods (left panel) and the price level of the capital stock (right panel) is divided by the price level of household consumption.

Figure C4 provides a cross-country view for 2005 of the resulting relative price level of the capital stock. As Hsieh and Klenow (2007) found, the prices of investment goods in poorer countries are high relative to the price of consumption and this is confirmed for 2005 in the left panel of Figure C4. Since many investment goods are traded, their prices are relative close to the exchange rate. In contrast, a considerable part of consumption is non-traded and prices in the non-traded sector tend to be lower in poorer countries. However, structures are non-traded, so their prices will be more similar to consumption prices. Since the depreciation rate of structures is lower than of other assets, the weight of structures in the capital stock will be larger than in total investment. As a result, the price level of the capital stock gives greater weight to the non-traded part of investment than the price level of investment. As a result, the right panel of Figure C4 shows a relationship between the relative price of the capital stock and CGDP^o per capita that is significantly smaller than between the relative investment price and CGDP^o per capita. It follows that the capital stock levels in PWT8 will typically be higher in poorer countries than the single-asset capital stock estimates that have been predominantly used in the literature.

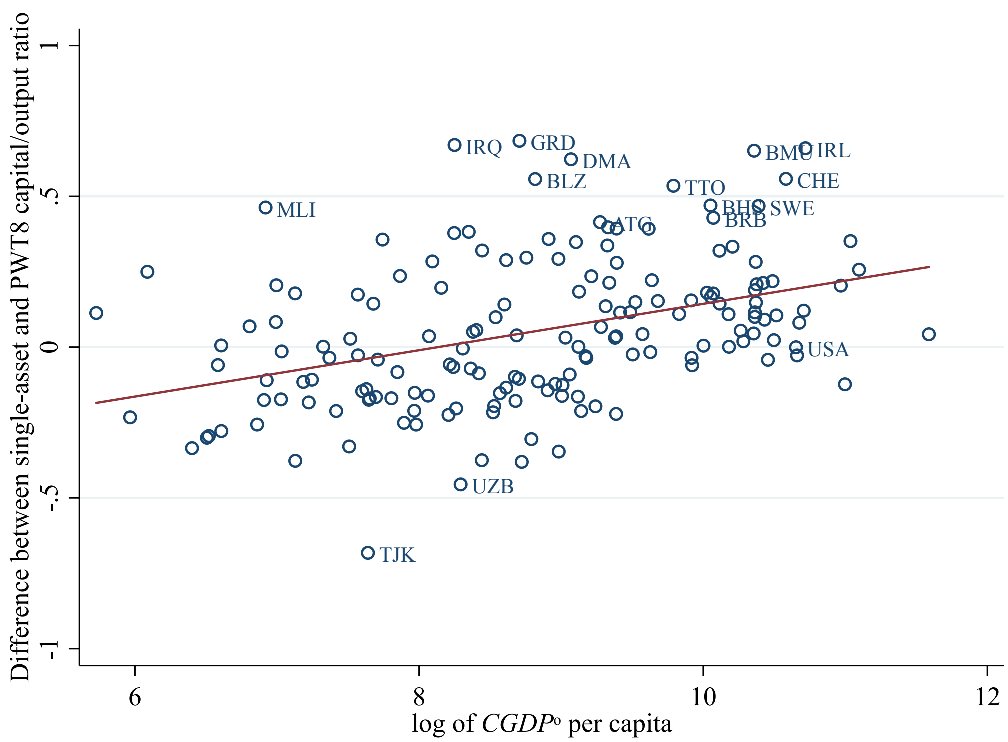

FIGURE C5: THE DIFFERENCE BETWEEN THE SINGLE-ASSET AND PWT8 CAPITAL/OUTPUT RATIO IN 2005

Note: single-asset capital stock cumulates overall real investment and depreciates it at a common 4 percent. The PWT8 capital stock uses asset-specific investment and depreciation rates and converts to a common currency using a capital stock PPP.

The reason is that those single-asset capital stocks were constructed using PWT investment series, without taking the larger weight of structures prices in the relative capital stock into account.

We constructed a single-asset capital stock using the same procedure as in Caselli (2005), so cumulating PPP-converted investment assuming a common depreciation rate (4 percent, based on Figure C3). The differences in 2005 between this single-asset capital stock and the PWT8.1 capital stock are shown in Figure C5. As the figure shows, the differences are quite large and, as predicted from Figure C4, the single-asset capital stock significantly underestimates capital input in poorer countries. The PWT8 capital stocks thus provide a different view than earlier approaches and a view that is more closely linked to the concept it represents.

Labor shares

This section is devoted to estimating the share of labor income in GDP, the *LABSH* variable in PWT8. This is a challenge because, in contrast to the labor income of employees, the labor income of self-employed workers is not directly observable as their income consists of compensation for both their labor supply and any capital they may own. This issue was taken up by Gollin (2002), who discusses different methods for estimating the labor compensation of self-employed workers. He showed for a modest set of countries that suitably adjusted labor shares are much more similar than naïve shares that ignore the labor compensation of the self-employed. This is because in poorer countries, more people are self-employed, which compensates for the lower naïve labor share in poorer countries.

For PWT8 we build upon these efforts by adding an adjustment method, increasing the range of countries, and, most importantly, extending the time period covered. Based on these, we come up with a 'best estimate' labor share based that is subsequently used in our TFP calculations. The end result is labor share estimates for up to 127 out of 167 countries in PWT8, covering the period since 1950.

Basic data and adjustment methods

The starting point is National Accounts data on compensation of employees, GDP at basic prices[10] and mixed income. Mixed income is the total income earned by self-employed workers, so it is a combination of capital and labor income. Given the aim of dividing the income of the self-employed between labor and capital, data on mixed income is very helpful by providing an upper bound to the amount of labor income earned by the self-employed. Two of Gollin's (2002)

[10] Net taxes on products are excluded since this is not income accruing to any of the factor inputs but a direct transfer to the government.

adjusted labor shares rely on mixed income information: the first allocates all mixed income to labor, assuming that self-employed workers only use labor input; while the second assumes that self-employed workers use labor and capital in the same proportion as the rest of the economy.

Mixed income data is available for 60 of the countries in PWT, so additional information is required. Gollin's (2002) third approach additionally uses data on the number of employees and the number of self-employed and assumes that self-employed earn the same average wage as employees. These data are drawn from the ILO LABORSTA database.[11] The 'same-wage' assumption may not be too far off the mark in advanced economies where the share of employees in the total number of persons engaged (employees + self-employed) is 85-95 percent. However, in many emerging economies this share is below 50 percent and as low as 4 percent.[12] In those countries, using information on the wages of employees will overstate the labor income of the self-employed.

We therefore propose an alternative adjustment method. Most self-employed workers are active in agriculture. According to the Socio-Economic Accounts (SEA) of the World Input-Output Database (see Timmer, 2012), agriculture employs about half of the self-employed in poorer countries. The agricultural sector also uses very few fixed assets in these countries as, according to the SEA, the agricultural labor share (accounting for the labor income of self-employed) is over 90 percent of value added, on average. We therefore have an adjusted labor share that adds all of value added in agriculture to labor compensation of employees. This adjustment could be too large as it ignores all income from capital and land and the labor income of employees in this sector is double-counted. On the other hand, the labor income of self-employed outside agriculture is ignored.

This provides us with data on the share of labor compensation of employees in GDP at basic prices and four adjusted shares, namely: two based on mixed income, one based on the share of employees in the number of persons engaged and one based on the share of agriculture in GDP. Table C5 shows descriptive statistics for these shares in 2005. Showing statistics for a single year makes it easier to illustrate the cross-country variation since the basic data for some countries is much more extensive than for others and 2005 gives the largest country coverage.[13]

---

[11] These data are supplemented by data from the Socio-Economic Accounts of the World Input-Output Database, see www.wiod.org.

[12] The share of employees in persons engaged is strongly positively related to GDP per capita.

[13] The sample of countries is different in each row. Descriptive statistics for the common sample of 42 countries shows very similar averages but with a smaller range.

Table C5 shows that, as in Gollin (2002), the naïve approach of using labor compensation of employees leads to very low labor shares of 42 percent on average and as low as 5 percent (in Nigeria). However, there are also some very high labor shares, up to 89 percent (in Bhutan). Adjustments 1 and 2 – based on mixed income data– show notably higher labor shares, though these can only be computed for 53 of the 108 countries. Especially some of the main oil- producing countries (Qatar, Oman, Venezuela) also show quite low labor shares (20-45 percent) based on these adjustments.

TABLE C5: DESCRIPTIVE STATISTICS FOR THE DIFFERENT LABOR SHARE ALTERNATIVES, 2005

Share	Mean	StDev	Min	Max	Obs
Naïve share	0.42	0.14	0.05	0.89	103
Adjustment 1, mixed income	0.60	0.12	0.20	0.90	53
Adjustment 2, part mixed income	0.53	0.13	0.18	0.73	53
Adjustment 3, average wage	0.67	0.28	0.20	2.27	89
Adjustment 4, agriculture	0.51	0.14	0.17	1.13	103

Notes: StDev: standard deviation, Obs: number of observations. Naïve share is the share of labor compensation of employees (COMP) in GDP. Adjustment 1 adds mixed income (MIX): (COMP+MIX)/GDP. Adjustment 2 assumes the same labor share for mixed income as for the rest of the economy: COMP/(GDP-MIX). Adjustment 3 assumes the same average wage for self-employed (SEMP) as for employees (EMPE): (COMP/GDP) *(EMPE+SEMP)/EMPE. Adjustment 4 adds value added in agriculture (AGRI): (COMP+AGRI)/GDP.

Using information on the number of self-employed and assuming they earn the same average wage as employees (adjustment 3) leads to average labor shares that are close to the commonly assumed labor share of 0.7 in Caselli (2005) and many others. Here too, though, there are countries with very low shares (Kuwait: 0.22), and some with unrealistically high shares, such as Bhutan (2.27). Bhutan already had a very high labor share according to the naïve share, so this overestimation is to be expected. This could indicate that, in contrast to National Accounting rules, the statisticians in Bhutan have already imputed the labor income of self-employed in their 'employee compensation' numbers. Lesotho's labor share under adjustment 3 of 2.05 indicates a similar problem.

The fourth adjustment adds the value added share of agriculture to the naïve share. The average share is somewhat lower than the commonly assumed two-thirds labor share but there are no countries with labor shares as extreme as under adjustment 3, though Bhutan is again the country with the highest labor share. In this broad cross-country setting, it would seem that any of the four adjustments would count as an improvement over the naïve share, but also that the mechanical application of one of these adjustments would not fit all countries equally well. This is even more apparent when comparing the cross-country pattern of adjustments 3 and 4, as done in Figure C6.

This shows that adjustment 3 leads to some very high labor shares for the poorer countries. Indonesia (IDN), for instance, shows a labor share of 92 percent under adjustment 3 but a share of only 44 percent under adjustment 4. The 92 percent adjustment is almost certainly too high since more than 80 percent of Indonesia's self-employed work in agriculture or distributive trade (based on SEA data). Adding the full value added earned in those sectors as labor compensation would lead to a labor share of 60 percent. Indonesia's remaining self-employed are unlikely to earn another 30 percent of GDP.

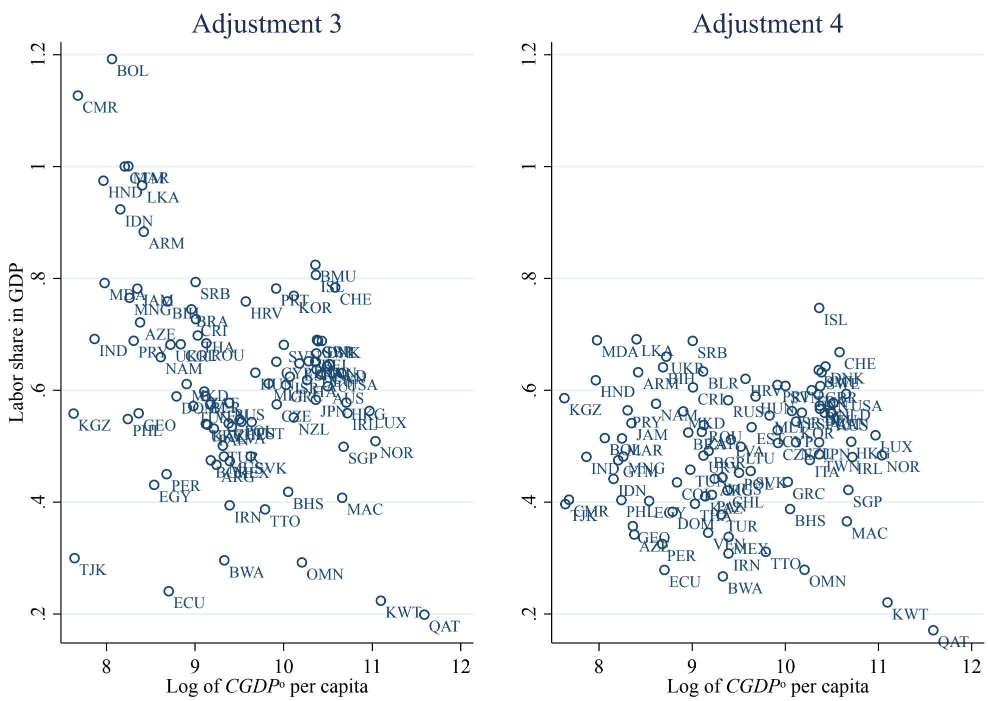

FIGURE C6: LABOR SHARES VS. GDP PER CAPITA IN 2005, ADJUSTMENTS 3 AND 4

Note: see notes to Table C5 for a description of the adjustments. Both figures include only countries for which data on adjustment 3 is available and excludes Bhutan and Lesotho (both with labor shares under adjustment 3 exceeding 2).

Best estimate labor share

These results suggest that a single adjustment approach is not appropriate for all countries. We therefore construct a 'best estimate' labor share based on the following three rules:

(i) Where available, adjustments based on mixed income seem preferable as this income directly relates to the income of self-employed, giving an upper-bound to the labor share. Adjustment 1 seems fairly extreme by assuming that self-employed use no capital at all, so we consider adjustment 2 to be the more plausible approach and use those labor shares if available.

(ii) Whenever mixed income data is available in National Accounts statistics, the naïve labor share never exceeds 0.66. So if the naïve labor share is larger than 0.7 in a particular country, it seems reasonable, like in Bhutan, that this share already includes an imputation for self-employed labor income. In those cases, the naïve share is used directly.

(iii) Given the patterns shown in Figure C6, there seems to be a greater chance of overestimating the labor share than underestimating the labor share. A conservative estimate would thus be the smaller of adjustments 3 and 4. So this is what we use if there is no mixed income data and the naïve share is below 0.7.

To ensure complete coverage over the years, we assume labor shares remain constant or we linearly interpolate if there are missing years in the middle of the sample. After these interpolations and extrapolations, we apply the three rules. Table C6 summarizes the resulting 'best estimate' labor shares for 2005.

TABLE C6: SUMMARY STATISTICS OF THE 'BEST ESTIMATE' LABOR SHARE IN 2005, BY TYPE OF ADJUSTMENT

Share	# of countries	Mean	StDev	Min	Max
Overall	127	0.52	0.14	0.18	0.89
of which:					
Adjustment 2, part mixed income	60	0.52	0.13	0.18	0.73
Adjustment 3, average wage	4	0.40	0.16	0.24	0.58
Adjustment 4, agriculture	62	0.52	0.14	0.22	0.85
Naïve share	1	0.89			

Notes: see notes to Table C5 for details on construction of the naïve and adjusted shares.

By interpolating and assuming shares constant over time, country coverage increases to 127 countries (out of 167 in PWT8). The resulting cross-country average of 0.52 is lower than Gollin's (2002) preferred 0.7 estimate, but it shows only a somewhat larger range than his of 0.34-0.91.[14] The average is lower, which is partly related to revisions of the underlying data for mixed income. Of the 15 countries with mixed income in both Gollin's data and here, the average 'adjustment 2' labor share is 68 percent in Gollin's (2002) data and 60 percent based on the current vintage of National Accounts data.[15] In contrast, the naïve share and 'adjustment 3' share are very similar for

[14] This is the range of his Table 2, not the 0.65-0.80 he mentions in his abstract. Only about half of his labor shares fall within the narrow 0.65-0.80 range.

[15] This compares shares for the same year as in Gollin (2002).

the overlapping set of countries. In addition, there seems to be a downward trend in labor shares over time, see below for more discussion.

Table C6 also illustrates that for almost half the countries, information on mixed income is available and therefore used. Adjustment 4 is used for most other labor share estimates and Adjustment 3 and the naïve share are only used for a few countries. The overall pattern is very similar if labor shares that are interpolated or assumed constant are dropped from the sample. The patterns shown in Table C6 for labor shares in 2005 also hold for the full sample, though there are many fewer labor shares based on observed data before the 1990s.

TABLE C7: LABOR SHARES AND VARIATION ACROSS INCOME LEVELS AND TIME

	(1) Naïve share	(2) Best estimate	(3) Best estimate	(4) Best estimate	(5) Best estimate
Log of CGDP^o per capita	0.0760***	0.00558	0.00594	0.0105	-0.0275**
	(0.00835)	(0.00940)	(0.00797)	(0.00747)	(0.0113)
Oil country			-0.153***	-0.146***	
			(0.0290)	(0.0264)	
Time trend				-0.00336***	-0.00159***
				(0.000495)	(0.000460)
Country dummies	no	no	no	no	yes
Observations	2237	2237	2237	2237	2237
R-squared	0.368	0.003	0.183	0.273	0.204

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors, clustered by country, in parentheses. Dependent variables are the naïve share (column (1)) or the best estimate labor share. Sample excludes labor shares that are interpolated or assumed constant from other years. Oil countries are OPEC countries and those countries for which the share of energy exports exceeds one-third. This share was chosen as all OPEC countries have an energy export share that is at least this large.

Table C7 analyses the cross-country patterns in the labor share data by relating labor shares to (the natural log of) CGDP^o per capita levels, excluding any labor shares that are assumed constant or interpolated. The naïve share in column (1) shows a strong positive relationship with GDP per capita, as Gollin (2002) also observed in his smaller sample. When using the best estimate labor share, no significant relationship with income levels is found in column (2). What does matter substantially is whether it is an oil country – the OPEC members and any other country in which energy exports accounts for at least one-third of total exports. Column (3) shows that those countries have labor shares that are on average 15 percentage points lower. Column (4) adds a linear time trend, and this shows a significant decline in labor shares over time. This remains the case when including country dummies in column (5). In that specification, increases in income levels are even associated with *lower* labor shares, compared with the higher labor shares from column (1). The time trend is less steep in column (5), but still highly significant. Moreover, the pattern of declining labor shares is found across the whole sample of countries as there is a decline in the labor share in 89 of the 127 countries and the trend is there for rich and poor countries alike.[16] Finally, if year dummies rather than a linear time trend is used, the assumed linear relationship over time of Table C7 is confirmed.

These results pose the important question how to explain this variation. Karabarbounis and Neiman (2014) propose cheaper investment goods to explain declines in labor shares over time, but other channels have been highlighted as well: capital-biased technological change, international trade (Harrison, 2005) and increased financialisation. This is fruitful area for additional research. Regardless of the underlying cause, though, this analysis illustrates quite clearly that the standard 'one-size-fits-all' labor share of 70 percent that is commonly used in the literature is a simplification that is not supported by the facts.

Appendix D: Extrapolating PPPs and Biased Balassa-Samuelson estimates

To see why extrapolating PPPs using relative inflation could lead to systematically biased PPP estimates, we use the following stylized example from Deaton (2012). We use the expenditure function to measure real GDP on the expenditure side and consider a two-country setting where consumers in both countries have common but non-homothetic tastes. The rate of inflation in each country can be measured by totally differentiating the expenditure function (obtaining a continuous-time Divisia index):

$$d\ln P_{jt} = \left(\frac{\partial \ln E\left(p_{jt}, u_j\right)}{\partial p_{jt}} \right)' d\ln p_{jt} = \ln s'_{jt} d\ln p_{jt}, \quad j = 1,2, \qquad \text{(D1)}$$

where p_{jt} is the vector of prices in country j, s_{jt} is the vector of the budget shares in country j, E is the expenditure function, and P_{jt} is the overall price level. The shares in (D1) must be evaluated at current prices, so that in-between two prices p_{jt-1} and p_{jt} these shares will be changing. Suppose that we approximate the discrete change in (D1) between these two price vectors by the Törnqvist index:

$$\Delta \ln P_{jt} \approx \tfrac{1}{2}(\ln s_{jt-1} + \ln s_{jt}) \Delta \ln p_{jt}, \quad j = 1,2.$$

Then the inflation differential between the two countries can then be written as:

[16] Though the trend is less steep and less significant in richer countries.

$$\Delta \ln P_{2t} - \Delta \ln P_{1t} \approx \tfrac{1}{2}(\ln s_{2t-1} + \ln s_{2t})'\Delta \ln p_{2t} - \tfrac{1}{2}(\ln s_{1t-1} + \ln s_{1t})'\Delta \ln p_{1t} \qquad \text{(D2)}$$

Next, suppose that we approximate the PPP index of country 2 relative to country 1 by a Törnqvist price index, computed using the shares at time t:

$$\ln PPP_{21,t} \approx \tfrac{1}{2}\left(s_{2t} + s_{1t}\right)'\left(\ln p_{2t} - \ln p_{1t}\right). \qquad \text{(D3)}$$

This PPP index relies on budget shares for the two countries, just like the Fisher indexes from section 3 were based on the quantity vectors of both countries; in both cases this allows for substitution in response to price differences. (D3) is an approximation because, with non-homothetic tastes, the Törnqvist price index does not precisely measure the ratio of expenditures in the two countries needed to obtain the same level of utility. Combining (D2) and (D3), we get the following expression for the change in the PPP index over time:

$$\begin{aligned}
\Delta \ln PPP_{21,t} &\approx \tfrac{1}{2}\left(s_{2t} + s_{1t}\right)'\left(\ln p_{2t} - \ln p_{1t}\right) - \tfrac{1}{2}\left(s_{2t-1} + s_{1t-1}\right)'\left(\ln p_{2t-1} - \ln p_{1t-1}\right) \\
&= \tfrac{1}{2}(s_{2t-1} + s_{2t})'\left(\ln p_{2t} - \ln p_{2t-1}\right) - \tfrac{1}{2}(s_{1t-1} + s_{1t})'\left(\ln p_{1t} - \ln p_{1t-1}\right) \\
&\quad - \tfrac{1}{2}\left(s_{2t} - s_{1t}\right)'\left(\Delta \ln p_{2t} + \Delta \ln p_{1t}\right) + \tfrac{1}{2}\Delta s_{2t}'\Delta \ln p_{2t} - \tfrac{1}{2}\Delta s_{1t}'\Delta \ln p_{1t} \\
&\approx \underbrace{\left(\Delta \ln P_{2t} - \Delta \ln P_{1t}\right)}_{\text{Inflation differential}} - \underbrace{\tfrac{1}{2}\left(s_{2t} - s_{1t}\right)'\left(\Delta \ln p_{2t} + \Delta \ln p_{1t}\right)}_{\text{Bias}}
\end{aligned} \qquad \text{(D4)}$$

The first approximation in (D4) is the same as that in (D3); the following equality is obtained with detailed algebra; and the final approximation follows from (D2) and by treating the final term $(\tfrac{1}{2}\Delta s_{2t}'\Delta \ln p_{2t} - \tfrac{1}{2}\Delta s_{1t}'\Delta \ln p_{1t})$ as small. Then (D4) states that the inflation differential between the two countries is a potentially biased estimator for the change in the PPP index, with the bias indicated by the final term on the right-hand side.

We can infer the direction of the bias when comparing rich and poor countries by considering that inflation of traded products (with higher productivity growth) tends to be lower than inflation of non-traded products. Suppose country 2 is poorer than country 1. In poorer countries traded products – such as food – generally have larger budget shares. As a result, there is a negative correlation between the components of the vectors of budget share differences $\left(s_{2t} - s_{1t}\right)$ in (D4) and average inflation $\tfrac{1}{2}\left(\Delta \ln p_{2t} + \Delta \ln p_{1t}\right)$. Therefore, the final term on the right-hand side of equation (D4) is negative, and including the minus sign will make it positive, so that the overall bias is positive. It follows that the PPP of country 2 relative to country 1 will increase at a *faster rate* than implied by the difference in overall inflation.

This result has important implications for the Balassa-Samuelson effect based on PPPs extrapolated using relative inflation rates. Since PPPs are typically extrapolated backwards in time – as recent ICP surveys cover more countries than the earlier ones – using the difference in relative inflation rates would overstate the 'true' PPP for poor countries if the previous argument holds. This overstatement would raise the price level of the poorer countries, and therefore bias the Balassa-Samuelson regression coefficient towards zero. The further back in time, the larger the bias would be, which explains why the Balassa-Samuelson effect is not found in our estimates when using extrapolated price levels.

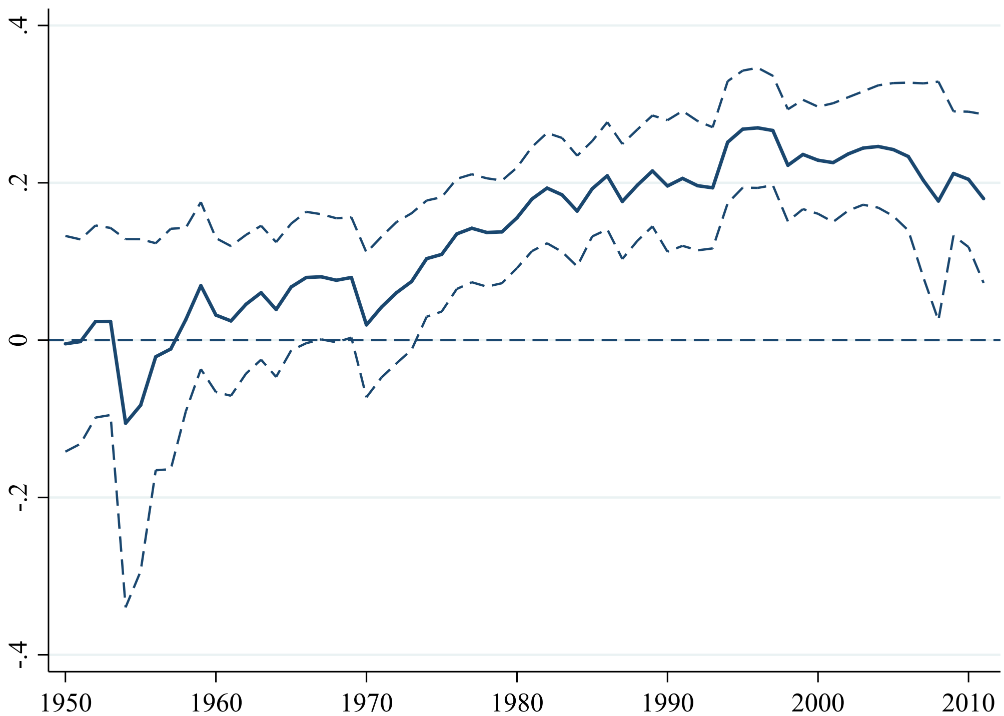

FIGURE D1: THE BALASSA-SAMUELSON EFFECT IN PWT8.1, 1950-2011

Notes: The figure plots β_{1t} from equation (D5) and its 95-percent confidence interval for all observations in PWT8.1, excluding those for which, due to extreme swings, no market exchange rate is available. Each β_{1t} is from a cross-sectional regression for all observations in year *t*.

To more firmly establish this finding that was already suggested in Figure 3 in the main text, consider Figure D1, which plots the annual coefficient β_{1t} from the Balassa-Samuelson regression:

$$\ln\left(\frac{PPP_{it}^{o}}{\mathcal{E}_{it}}\right) = \beta_0 + \beta_{1t}\ln\left(\frac{CGDP_{it}^{o}}{POP_{it}}\right) + \varepsilon_{it}. \quad \text{(D5)}$$

The figure shows a qualitatively similar pattern as Bergin *et al* (2006), with a low and insignificant coefficient in the early years, which rises steadily and becomes highly significant in later years. The main substantive difference is that it takes until 1974 before the slope coefficient in (23) turns significantly positive, versus the mid-1950s in Bergin *et al* (2006).[17] This is a puzzling finding, as the Balassa-Samuelson effect was already identified in data for the 1950s and 1960s in e.g. Balassa (1964), therefore raising the question why the effect shows up so much later in PWT8.

To determine whether the non-benchmark observations are in fact leading to the disappearance of the Balassa-Samuelson effect in early years, we distinguish between PPP benchmarks and observations interpolated between those benchmarks on the one hand, and observations that are extrapolated on the other hand. Figure D2 shows the result from the same cross-sectional regressions as in Figure D1, but run separately on the two subsamples from 1970-2005; note that there are no benchmark observations before 1970 and that the last global benchmark comparison was in 2005. Figure D2 shows that in the *benchmark/interpolated* sample there is a consistently positive and significant coefficient β_{1t} of 0.3 on average, confirming the Balassa-Samuelson effect, while in the *extrapolated* sample, the coefficient is rarely significantly different from zero. This confirms the hypothesis that the extrapolation procedure biases the Balassa-Samuelson effect towards zero.[18] The pattern in Figure D2 can thus be best understood as a weighted average between a positive, significant coefficient in the benchmark/interpolated sample and a zero coefficient in the extrapolated sample. As the weight of the benchmark/ interpolated sample increases over time, the overall coefficient increases and turns significant. The study of Bergin *et al* (2006) is based on PWT 6 which relied exclusively on extrapolation of PPPs from a recent benchmark year which would explain their findings. Our findings illustrates the usefulness of including historical benchmark material and clearly distinguishing between benchmark/interpolated and extrapolated observations, as done in PWT8.

[17] The same analysis based on PWT7.1 shows a qualitatively similar pattern, but the coefficient does not turn significantly positive until 1988.

[18] If we restrict ourselves to the countries and years included in the benchmark-or-interpolated sample and compute extrapolated PPPs, we still find that the Balassa-Samuelson coefficient β_{1t} is biased towards zero.

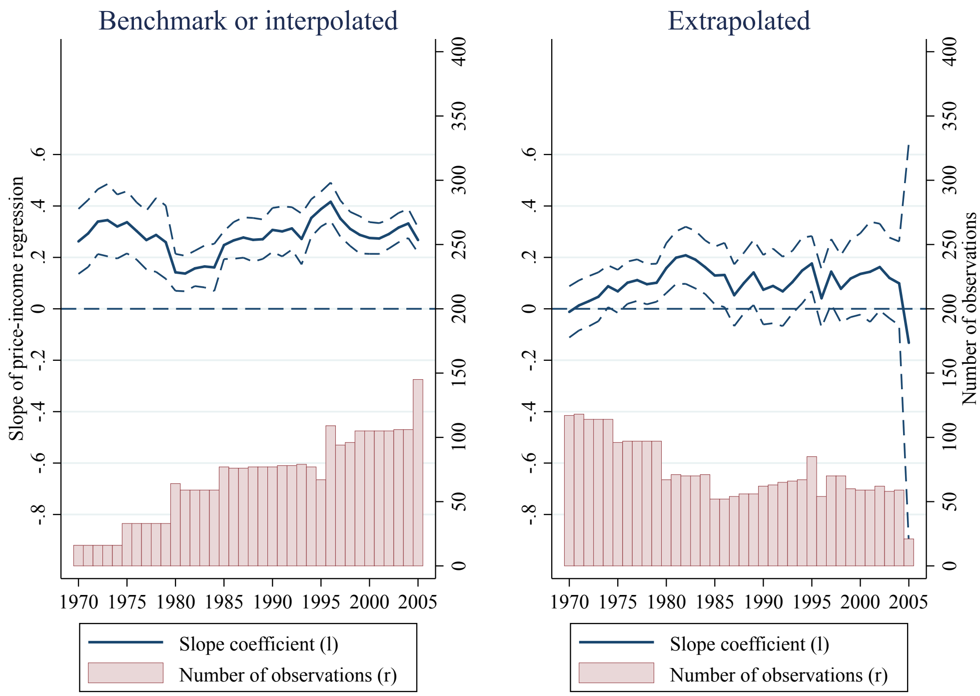

FIGURE D2: THE BALASSA-SAMUELSON EFFECT IN PWT8.1 FOR BENCHMARK OR INTERPOLATED AND EXTRAPOLATED OBSERVATIONS, 1970-2005

Notes: The figure plots β_{1t} from (D5) and its 95-percent confidence interval for all observations in PWT8.1. Each β_{1t} is from a cross-country regression in year *t*. The left-hand panel includes only observations from PPP benchmarks or interpolated between PPP benchmarks. The right-hand panel includes only observations extrapolated from PPP benchmarks using inflation rates, i.e. all other observations. Excluded are those observations for which, due to extreme swings, the market exchange rate is replaced in PWT8.1 by an estimated exchange rate based on a relative PPP assumption.

REFERENCES:

Balassa, Bela. 1964. "The Purchasing-Power Parity Doctrine: A Reappraisal." *Journal of Political Economy* 72 (6): 584–96.

Barro, Robert and Jong-Wha Lee. 2013. "A New Data Set of Educational Attainment in the World, 1950-2010." *Journal of Development Economics* 104: 184–98.

Bergin, Paul R., Reuven Glick and Alan M. Taylor. 2006. "Productivity, Tradability, and the Long-Run Price Puzzle." *Journal of Monetary Economics* 53 (8): 2041–66.

Caselli, Francesco. 2005. "Accounting for Cross-Country Income Differences." In Phillipe Aghion and Steven N. Durlauf, editors, *Handbook of Economic Growth, Volume 1A,* 679-741. Amsterdam: Elsevier,.

Caselli, Francesco and Daniel J. Wilson. 2004. "Importing Technology." *Journal of Monetary Economics* 51 (1): 1–32.

Deaton, Angus. 2012. "Consumer Price Indexes, Purchasing Power Parity Exchange Rates, and Updating." Unpublished paper for the PWT Workshop May 2012, Philadelphia.

Diewert, W. Erwin. 2008. "The Economic Theory of Index Numbers: A Survey." In Angus Deaton, editor, *Essays in the Theory and Measurement of Consumer Behaviour in Honour of Sir Richard Stone*, 163–208. Cambridge: Cambridge University Press.

Feenstra, Robert C., Alan Heston, Marcel P. Timmer, and Haiyan Deng. 2009. "Estimating Real Production and Expenditures Across Countries: A Proposal for Improving the Penn World Tables." *Review of Economics and Statistics* 91 (1): 201–12.

Feenstra, Robert C. and John Romalis. 2014. "International Prices and Endogenous Quality." *Quarterly Journal of Economics* 129 (2): 477–527.

Fraumeni, Barbara M. 1997. "The Measurement of Depreciation in the U.S. National Income and Product Accounts." Survey of Current Business. July: 7-23.

Gollin, Douglas. 2002. "Getting Income Shares Right." *Journal of Political Economy* 110 (2): 458–74.

Harberger, Arnold. 1978. "Perspectives on Capital and Technology in Less Developed Countries." In Michael J. Artis and Avelino R. Nobay, editors, *Contemporary Economic Analysis*, London: Croom Helm.

Harrison, Ann. 2005. "Has Globalization Eroded Labor's Share? Some Cross-Country Evidence." MPRA Paper No. 39649.

Hill, Robert J. 2004. "Constructing Price Indexes across Space and Time: The Case of the European Union." *American Economic Review* 94 (5): 1379–410.

Hsieh, Chang-Tai and Peter J. Klenow. 2007. "Relative Prices and Relative Prosperity." *American Economic Review* 97 (3): 562–85.

Inklaar, Robert. 2010. "The Sensitivity of Capital Services Measurement: Measure All Assets and the Cost of Capital." *Review of Income and Wealth* 56 (2): 389–412.

Inklaar, Robert and Marcel P. Timmer. 2013. "Capital and TFP in PWT8.0." mimeo, available at: www.ggdc.net/pwt.

Karabarbounis, Loukas and Brent Neiman. 2014. "The Global Decline of the Labor Share." *Quarterly Journal of Economics* 129 (1): 61–103.

Malmquist, Sten. 1953. "Index Numbers and Indifference Surfaces." *Trabajos de Estatistica* 4: 209–42.

Nehru, Vikram, and Ashok Dhareshwar. 1993. "A New Database on Physical Capital Stock: Sources, Methodology and Results." *Revista de Analisis Economica* 8 (1): 37–59.

Rao, D. S. Prasada, Alicia Rambaldi and Howard Doran. 2010. "Extrapolation of Purchasing Power Parities using Multiple Benchmarks and Auxiliary information: A New Approach." *Review of Income and Wealth* 56 (S1): S59–98.

Timmer, Marcel. 2012. "The World Input-Output Database (WIOD): Contents, Sources and Methods." WIOD Working Paper Number 10, downloadable at (May 22, 2014): http://www.wiod.org/publications/papers/wiod10.pdf .