[03.01]

PPPs for Government Services

Robert Inklaar and Marcel Timmer

7th Technical Advisory Group Meeting
September 17-18, 2012

Washington DC

PPPs for government services

Robert Inklaar and Marcel Timmer[1]

Groningen Growth and Development Centre, University of Groningen

September 2012

Abstract

For a substantial number of countries the relative price of government services in ICP 2005 was initially approximated using the wages of public servants, education and health workers. This is based on the implicit assumption that productivity of workers does not differ across countries which is implausible. Various adjustments have been suggested in the past. In this paper, we built upon previous efforts and argue for a modification to the ICP2005 productivity adjustment. We draw on a broader basis of data and our procedure is more straightforward to implement than the existing ones. This leads to adjustment factors that are similar in scale to those implemented in ICP2005, but with differences in distribution across countries.

Introduction

Measuring the volumes of government services is one of the more challenging areas of the ICP program because output prices of those services cannot be directly observed. This affects a major part of the economy, namely the general government, education and health sectors. One solution to this challenge would be to measure outcomes of these activities, such as the number of pupils taught. Given such information on relative quantities, relative output prices can be indirectly inferred. As argued by, for example, Schreyer (2012) this would be the first-best approach, but it requires much detailed information (see also Blades, 2012 and Heston, 2012) which is only avaible for a small set of countries.

In the long tradition of ICP various alternatives have been suggested (see e.g. discussion in Sergueev 1998). In the most recent round in some regions of the ICP, input-proxies have been used, assuming that output volumes are proportional to input volumes. This approach is followed and relative wages of civil servants, teachers, nurses, doctors, etc. are used to convert labor costs to relative labor input volumes. However, this approach is only correct if workers in government are equally productive across countries.[2] World Bank (2008) and Heston (2012) have both suggested alternative approaches to adjust relative

[1] We thank Wen Chen and Juan Ricardo Perilla Jimenez for excellent research assistance and Sergey Sergueev and Alan Heston for helpful cooperation and comments.

[2] For simplicity. throughout the paper we assume all education, health and government services are carried out by the public sector, no output prices or volumes are available and expenditure is measured by wages only .

wages for differences in productivity and these have been applied for Africa, Asia-Pacific and West Asia in the 2005 round.

In this paper, we outline a new procedure to estimate these productivity differences based on differences in capital input and present the results, building upon this work. We compare both the procedure and results to the existing approaches by World Bank (2008) and Heston (2012). The main contribution of this research is that we develop new measures of capital input based on investment broken down by up to six different assets for all 145 countries in the ICP 2005 comparison. In addition, we use newly developed data on capital shares in GDP for 40 countries from across the development spectrum to estimate capital shares for all 145 countries.

Our approach

The aim of the approach is to correct for labour productivity differences in government activities across countries. This requires direct output measures of these activities and this data is obviously not available. One alternative would be to proxy this by means of labour productivity in market activities, which would require employment estimates of the market sector and a market volume output. This has been tried in the past with limited success as proper employment statistics were difficult to derive (see Sergueev 1998 for a discussion of the experience) and relative differences in market and non-market activities are known to differ (giving rise to the famous Harrod-Balassa-Samuelson effect).

More recently, World Bank (2008) and Heston (2012) suggested an alternative approach to proxy differences in labour productivity through differences in the use of capital per worker. Here we follow this line of thinking. To estimate the extent to which differences in wages reflect differences in prices and productivity, we assume that output of government services Y is produced using capital K and labor L with efficiency level A:

(1) $$Y = f(K, L, A)$$

If we assume that the production function has constant returns to scale and exhibits Hicks-neutral efficiency, we can rewrite (1) as:

(2) $$\frac{Y}{L} = A\left(\frac{K}{L}\right)^{\alpha} \Leftrightarrow y = Ak^{\alpha}$$

where α is the output elasticity of capital. The aim is to compare labor productivity y between any given pair of countries i and j and each country could be at a different point on the production function and thus have a different output elasticity of capital. The standard approach to this problem is to define a hypothetical 'average' country, with variables denoted by an upper bar, and compare each country to this average; see Caves, Christensen and Diewert (CCD, 1982). This procedure is akin to the EKS index number approach, but is based on the Törnqvist index instead of the Fisher index. Relative labor productivity between country i and the average country is then equal to:

(3) $$ln\left(\frac{y_i}{\bar{y}}\right) = ln\left(\frac{A_i}{\bar{A}}\right) + \frac{1}{2}(\alpha_i + \bar{\alpha})ln\left(\frac{k_i}{\bar{k}}\right)$$

If we were to know all elements in the right-hand side in equation (3), we would now be able to adjust government relative wages for productivity differences, but we have no direct information on any of the variables. Efficiency differences A are the hardest to pin down and, following earlier work (Inklaar and Timmer 2008), we assume that efficiency in the use of labour and capital inputs is the same across countries. This is not likely to be true, but we have no basis for a better-founded assumption. It is unlikely, though, that efficiency differences are as large in this sector as in the economy as a whole. According to the arguments of Balassa (1964) and Samuelson (1964), efficiency differences are likely to be smaller in non-traded goods, such as government services, because these are labor intensive, making it harder to achieve returns to scale.

To approximate capital per worker, we will assume that relative capital per worker in government is the same as for the economy as a whole due to a lack of data on capital use in the public sector. Results for 20 advanced economies, presented in Inklaar and Timmer (2008), show that this is not an unreasonable assumption. Though the correlation between relative capital input in government and the total economy is not perfect, it is positive and significant at 0.45.

The output elasticity of capital, α, is not directly observable, but a common approach is to assume perfect competition in the factor input and product market so that the revenue share of capital can be used instead. Again, we have no information about the revenue share of capital for government services, so we will use the capital share in GDP instead.

Given these data (more on which below), we compute adjustment factors (F) for relative wages. These are based on capital input (relative to the average country) for country *i* compared with capital input for the base country:

(4) $$F_i = ln\left(\frac{y_{US}}{y_i}\right) = \frac{1}{2}(\alpha_{US} + \bar{\alpha})ln\left(\frac{k_{US}}{\bar{k}}\right) - \frac{1}{2}(\alpha_i + \bar{\alpha})ln\left(\frac{k_i}{\bar{k}}\right)$$

Here we take the USA as the base country, but the CCD method is base country independent. The adjustment factors, as defined in equation (4), are used to adjust relative wages in country i (w_i) for labour productivity differences:

(5) $$\widetilde{w_i} = w_i e^{F_i}$$

This adjustment factor will generally be higher than 1, since most countries have lower levels of capital per worker than the US, the base country in our comparison. Since our model implies that a government worker is less productive in a country with less capital, his productivity-adjusted wage should be higher. This leads to lower relative input volumes when applied to nominal input values.

Another element that would normally lead to cross-country differences in labor productivity, and hence wages, are differences in levels of schooling. Since the ICP wages are collected for precisely specified categories of workers, distinguished also by their educational qualifications (Blades, 2012), we assume no further adjustments are required. The productivity adjustment needs to be applied across all categories of workers.

We will derive two sets of adjustment factors: one based on capital stock estimates and another based on capital services. In the next section, we discuss the difference in theory and practice. By providing the two alternatives, a robustness analysis can be made.

Comparison to alternative approaches

Our approach is most similar to that of World Bank (2008). The main difference is that their approach was based on sparser data on capital input and capital income shares. In the implementation, they applied estimates of the capital-output ratio and capital from a limited set of countries by income level. Specifically, low-income countries would be assumed to have a capital-output ratio of 2.5 and a capital share of 50%, while high-income countries would have a capital-output ratio of 3.5 and a capital share of 30%. A practical complication of their approach was that to move from capital-output ratios to capital per worker ratios, initial estimates of GDP per worker were needed. However, the wage adjustment would affect GDP estimates, so an iterative procedure was used. Our approach improves on this by generating capital and labor input estimates for each of the ICP 2005 countries and providing capital share estimates based on broader country-level data. This also allows us to implement our approach without any need for iteration.

Heston (2012) proposes an approach that relies on an econometric estimate of the contribution of capital per worker to differences in GDP per worker, using data on capital input for 106 countries. As a result, no data on capital cost shares are needed, since the output elasticity is estimated econometrically. The output per worker level predicted from capital input levels provides a continuous set of adjustment factors. Rather than using these directly, he groups countries into broad bins, where the wages of the countries in the highest-productivity bin are not adjusted, and subsequently higher adjustments are made as countries fall into lower-productivity bins. A likely advantage of this discretization is that small differences in capital input, which could be caused by measurement error, do not feed one-for-one into the results While this approach has the benefit of being less data-intensive, econometric estimates of capital's output elasticity also have their problems. The most obvious drawback is a lack of cross-country variations in the elasticity, which is an assumption that seems at odds with the observed variation in capital cost shares. Other econometric challenges, such as omitted variables, simultaneity, etc. should also be dealt with.

Data and methodology

Investment at current and constant prices

The measurement of productive capital inputs is a data-intensive exercise; see for instance OECD (2009). There are two main challenges, namely the collection of investment flows and prices, and the great variation in productive asset lives and thus, presumably, in their contribution to output. For example, an office building may well be used for several decades, while computers are typically replaced after five years. A computer should therefore generate much larger returns for every dollar invested than an office building since the investment has to be recouped in a much shorter period of time. A common shortcut method is to ignore this heterogeneity and estimate capital input based on a common and constant assumed asset life. This ignores important changes in investment composition over time and differences across countries. It is well documented that in the course of economic development the share of equipment and machinery investments increases. Also, richer countries typically devote a higher share on short-lived ICT assets than poorer countries at any point in time.

In this research, we have developed a new dataset of investment by assets for all countries that have ever participated in an ICP benchmark comparison, so including the 145 countries in ICP 2005. We distinguish up to six assets, shown in Table 1 with their geometric depreciation rates. These rates are assumed to be common across countries and constant over time. As these data are not readily available for all countries, we use a variety of sources in compiling the investment data. To begin with, we ensure that detailed investment sums to total gross fixed capital formation for each country and year from the UN National Accounts Main Aggregates Database. This control total is extrapolated back to 1950 whenever available using data from the UN National Accounts Official Country Data; national sources and PWT7.0.

Table 1, Assets covered and geometric depreciation rates

Asset	Depreciation rate
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 Fraumeni (1997)

We follow a two-stage procedure for estimating the asset composition of investment. In the first stage, we distinguish structures, transport equipment and equipment and software. 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 additionally use 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. Combined with asset investment shares from ICP benchmarks we apply a so-called commodity flow method, whereby the trend in investment is approximated by the supply of investment goods in the economy.[3] The combination of data sources and the application of the commodity flow method is described in detail in Appendix B.

In the second stage, we use data compiled by The Conference Board from EU KLEMS on information and communication technology (ICT) investment and WITSA on ICT expenditure to split up investment in equipment and software into the last four assets shown in Table 1. This provides us with a dataset showing investment at current national prices.

We also need deflators, and here 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 detailed enough data, such as the US. For ICT assets, we thus 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. The result is still though that for many countries, only the total investment deflator is used for non-ICT assets.[4]

Initial capital stocks

Some assets, in particular structures, have very low depreciation rates, corresponding to long asset lives. As a result, an assumption has to be made on the initial capital stock to use in building up capital stocks. This is an important decision, particularly for formerly Communist countries for which National Accounts data start in 1990. A common assumption is to use the steady-state relationship from the Solow growth model:

$$(6) \qquad K_0 = \frac{I_0}{g+\delta}$$

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. Experiments with this approach showed that many formerly Communist countries would have implausibly high capital levels.

An alternative is to make an assumption about the initial capital-output ratio. This is in the same vein as World Bank (2008), who applied a capital-output ratio between 2.5 and 3.5 depending on a country's income level. Based on experiments with a steady-state initial capital stock described below, we found that the capital-output ratio does not actually vary systematically by income level, neither for the sum of assets, nor for most individual assets. The exceptions

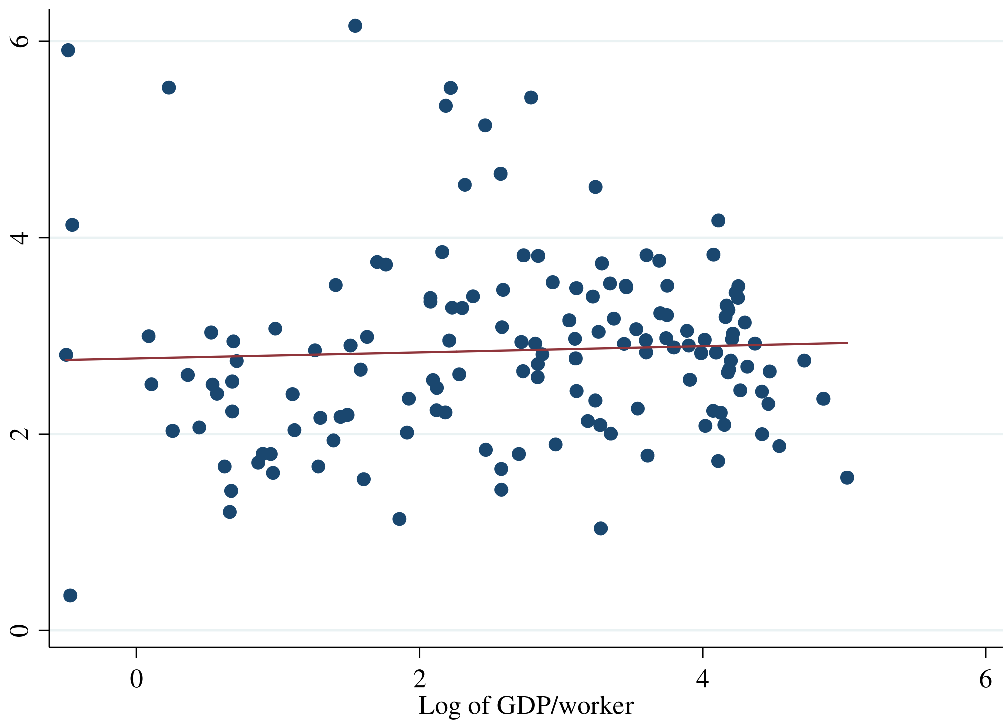
[3] This approach has also been used by Caselli and Wilson (2004), though without the constraint that investment had to add up to gross fixed capital formation in the National Accounts.
[4] An alternative that can be explored in the future is to use quality-adjusted relative import prices for capital goods and transport equipment, newly developed by Feenstra and Romalis (2012).

are ICT assets, but given their short asset lives, we set the initial capital/output ratios to zero without much effect on the capital stocks in 2005. Table 2 shows the initial capital/output ratios that we assume for all countries.

Table 2, Assumptions for nominal capital stock/GDP ratios

Asset	Capital/output ratio
Structures (residential and non-residential)	2.25
Transport equipment	0.1
Other machinery and assets	0.25
ICT assets	0
Total	2.6

Figure 1, Overall Capital stock/GDP ratio, in 2005 versus log GDP/worker



Note: capital output ratios are computed based on data at current national prices and capital is summed over all assets.

Figure 1 plots capital/output ratios in 2005, summed over all assets, against GDP per worker; the figure is almost identical for GDP/capita. The figure shows considerable variation in capital/output ratios around an average of about 2.6. The least squares regression line illustrates that there is no systematic relationship between GDP/worker and capital/output ratios. This holds also when restricting the sample to countries for which at least 45 years of investment data is available, i.e. the group of countries with the most reliable estimates. This suggests the Kaldor (1957) fact fit the data fairly well.

Capital stocks and services

Given an initial capital stock, nominal investment flows and prices and depreciation rates, it is straightforward to compute capital stocks for asset *a* at time *t* using the Perpetual Inventory Method (PIM):

(7) $$K_{at} = (1-\delta) K_{at-1} + I_{at}$$

Where K and I are at constant national prices. Multiplying this capital stock by the asset deflator then gives capital stocks at current national prices. For the comparison of capital stocks across countries, one needs to convert capital stocks to a common currency and aggregate over assets. Given the results from equation (7), the computation of aggregate capital stocks at common international prices requires only investment PPPs. From ICP 2005, we aggregate the investment basic headings to one for transport equipment, one for structures and one for the four equipment and software assets. For countries in the OECD/Eurostat group, we have more detailed investment basic headings that allow us to compute a specific PPP for each asset. Aggregation to three assets (ICP 2005) or 6 assets (OECD/Eurostat) is done using an EKS approach with investment as weights.

Capital stocks thus derived can be used directly in equation (4). However, relative capital stocks across countries do not properly reflect the contribution of each asset to production. As discussed above, computers should receive a relatively larger weight than structures because of more rapid depreciation and asset price declines. This is taken into account in the so-called capital services approach. Following OECD (2009), we define the user cost of capital *uc* for asset *a* at time *t* as:

(8) $$uc_{at} = i_t + \delta_a - dp_{at}$$

where *i* is the nominal required rate of return on capital and *dp* the asset-specific price change. Following Inklaar and Timmer (2008), we set *i* equal to a financial market interest rate. Specifically, we use data on lending rates, government bond yields and government bill yields from the IMF's International Financial Statistics. The lending rate is closest to the interest rate likely faced by borrowers, so we use the government yields to extend the data for lending rates, assuming a constant margin of lending rates over government yields. To extend the data further, we assume constant real interest rates and extrapolate using the GDP deflator. Finally, we assume real interest rates should be at least 2%. For countries without any interest rate data, we assume a constant real interest of 4%, in line with Diewert (2001).

In a more data-rich environment, a more sophisticated approach could be taken, such as estimating the weighted average cost of capital based on equity and bond yields (Inklaar, 2010), but given the sample of countries, this would not, in general, be feasible. Another alternative would have been to set the required rate of return so that the sum of capital compensation (user cost times capital stock at current prices) equals GDP minus labor compensation; a so-called internal rate of return. However, any measurement error in capital stock construction or official data would have a major impact on this internal rate. Even for advanced economies in EU KLEMS, this can easily lead to large differences and swings in the rate of return.

As the asset-specific price change, we use the average annual investment price change over the previous five years, to smooth out potentially severe year-to-year fluctuations in asset inflation. Even with all these adjustments to avoid implausibly high or low user costs, there are still extreme values from applying equation (8). We therefore set a minimum of 2% and a maximum of 100% for the user costs and replace more extreme values by this minimum or maximum. This affects only a small fraction of observations, around 1% of observations showed user costs at less than 2% and only in Zimbabwe did user costs exceed 100%.

As for capital stocks, one needs to convert capital services into a common currency. Since capital compensation is defined as the user cost times the capital stock at current prices, the relative price of capital compensation is given by the relative user cost times the relative price of capital, i.e. the asset's investment PPP (omitting the time subscript):

$$(9) \qquad PPP_{ija}^{Kser} = \frac{uc_{ia}}{uc_{ja}} PPP_{ija}^{I}$$

Based on the capital services PPPs, we derive relative levels of capital service volumes across countries, which can be used as an alternative for capital stocks in calculating relative capital intensities in equation (5).

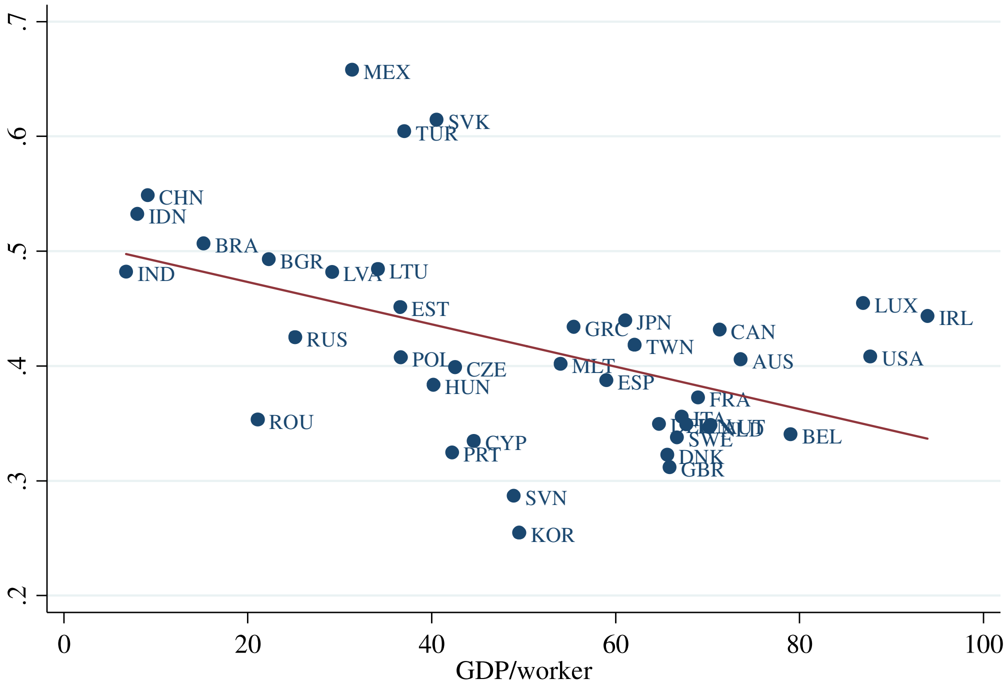
The preceding discussion suggests there is a trade-off between using capital stocks or capital services. On the one hand, a measure of capital services is preferable from a conceptual point of view: assets that generate more output should be weighted more heavily. On the other hand, computing capital services requires more data, which may increase measurement error. The trade-off between these factors of course also depends on whether capital services provide a noticeably different perspective than capital stocks. This will be most likely where investment in ICT assets are important, because the user costs for those assets are much higher than for other assets, so they are weighted much more heavily in capital services than in capital stocks measures. Data for the limited number of countries in the EU KLEMS database suggests that in government services, ICT is used more intensively in government services than in the economy as a whole (the average intensity is 50% larger), which suggests we may even be underestimating capital services inputs in government services by relying on total economy estimates. This underestimation would be even larger with capital stock estimates.

Capital shares

Estimating the share of GDP earned by capital requires not just information about the labor compensation of employees, a figure that is generally published as part of the National Accounts, but also an estimate of the labor income of the self-employed. Detailed work on this has recently been done in the World Input-Output Database (see Timmer, 2012) for 40 countries spanning a broad range of economic development.

Figure 2 plots the capital shares from this database against GDP per worker. This shows that the majority of countries have a capital share that is considerably higher than the 1/3 that is often used by growth economists. It is more in line with the 0.3-0.5 that was used in World Bank (2008). There is also a clear and highly significant negative relationship between capital shares and GDP per worker. This also confirms the approach by the World Bank (2008), but provides a more fine-grained scale along which to put countries that were not covered in WIOD. We thus use the least-squares regression line from Figure 2 to predict the capital share for countries for which we have no data from WIOD. This leads to a range of labor shares between 0.23 and 0.51, a range that is well spanned by the WIOD data.[5]

Figure 2, Share of capital in GDP in 2005 versus GDP per worker (in US$)



Source: WIOD Socio-Economic Accounts, February 2012

Labor

Finally, a measure of labor input is also required to implement equation (5). As illustrated by Heston (2012), using the population or even working-age population is likely to be misleading as labor force participation rates vary considerably. We build on this by relying, first, on data for the number of persons engaged from The Conference Board's Total Economy Database (January 2012, henceforth TED). This database covers 111 countries of the countries in ICP 2005. For 10 additional countries, we use employment data from ILO; for

[5] The relationship between the capital share and the log of GDP per worker gives a slightly better fit than the GDP per worker level. The main drawback is that the least-productive countries would have a capital share of about 70 percent rather than 51 percent, which is highly implausible.

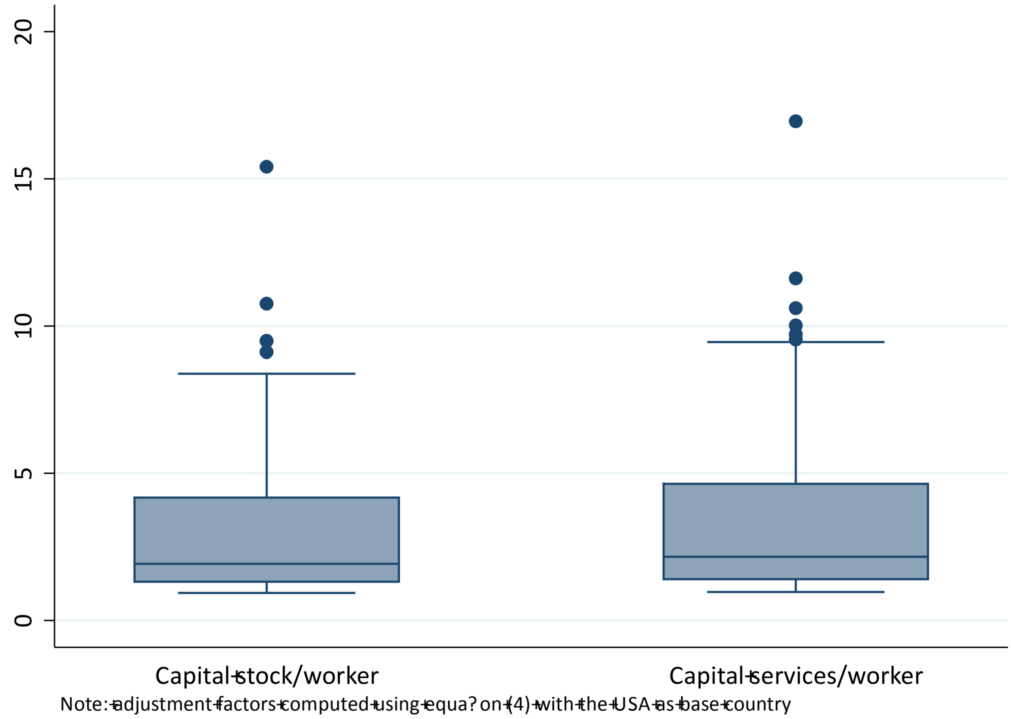
another 24 we use employment data from WDI and for one country, we use data on the labor force from WDI.[6]

Results

Wage adjustment factors

The data described in the previous section provide the necessary ingredients to implement equation (4) for all 145 ICP 2005 countries and compute wage adjustment factors for these countries. Figure 3 shows box plots for the wage adjustment factors based on the estimates of capital stock per worker and capital services per worker, using the US as the base. The adjustment factors for capital services are generally higher than for capital stocks, reflecting the fact that the US has a relatively large share of ICT assets, which have a larger weight in capital services than in capital stocks. This mirrors the findings for the smaller set of countries in Inklaar and Timmer (2008).

Figure 3, Box plot of country adjustment factors for government wages in 2005, based on capital stock and capital services estimates



Note: Box plot ranges 5, 25, 50, 75 and 95%.

The two sets of adjustment factors are overall very similar. The capital services adjustment factors are somewhat larger, 2.2 versus 1.9 for capital stocks at the

---

[6] There are cases where data from, for instance ILO, is available for an earlier or later year than 2005. In that case, we use the data from the more preferred source and extrapolate it to 2005 using the less preferred source. Our preference ranking is 1) TED (constructed for maximum international comparability); 2) ILO (same concept as TED and closest to TED where both are available); 3) WDI employment (same concept as TED); and 4) WDI labor force (different concept).

median. The dispersion of capital services adjustment factors is also larger, which could be due to the added computational steps in the construction of capital services and corresponding measurement error. The similarities dominate, though, with a correlation between the two sets of adjustment factors of 0.997. We find that most of the adjustment factors are between 1 and 4, which is the range used by Heston (2012). However, about 40, mostly African, countries shows higher adjustment factors, with a maximum of 15 to 17 for Zimbabwe. In only a few countries do we find capital levels higher than in the USA and the lowest adjustment factor is 0.93 for Luxembourg based on capital stocks.

GDP per capita

For the final part of the analysis, we gauge the impact on real GDP per capita, comparing three sets of estimates. The first uses basic headings for health, education and government services for which PPPs are based on relative wages unadjusted for productivity differences.[7] The first alternative uses the capital stock adjustment factors for health, education and government sevices PPPs and the second uses the capital services adjustment factors. For each of the three cases, we compute a GEKS price index across all basic headings and countries and use this GDP PPP to compute GDP per capita. The overall impact of our adjustments on GDP will depend on the adjustment factor and the share of health, education and government services in overall GDP.

Figure 4, GDP per capita in 2005 across productivity adjustments

Figure 4 plots the ratio of adjusted to unadjusted GDP per capita versus the log of unadjusted GDP per capita for both adjustment factors. Both sets of adjustment

[7] This is based on a set of basic heading PPPs with unadjusted wage PPPs for for health, education and government sevices underlying Heston (2012). They do not correspond for all countries to the original source material and are used in this paper for illustrative purposes only.

factors have very large effects on GDP per capita, with downward adjustments of up to 30 percent, though for all but 10 percent of the countries the adjustment is less than 20 percent. The adjustment also becomes smaller with increasing income levels, reflecting how capital levels increase with increasing output levels. The two adjustment factors yield very similar GDP per capita numbers, with differences in the order of 0-3.5 percent.

Table 3, Adjustment factors and GDP per capita before and after adjustment in 2005,

Country	ISO code	Adjustment factor		Adjusted GDP relative to unadjusted	
		Stock	Services	Stock	Services
Top 10 lowest adjustment factors					
Luxembourg	LUX	0.93	0.96	1.00	0.99
Singapore	SGP	0.95	1.00	0.92	0.92
United States	USA	1.00	1.00	1.00	1.00
Japan	JPN	0.96	1.01	0.95	0.94
Australia	AUS	1.07	1.07	0.95	0.95
Switzerland	CHE	1.14	1.08	0.95	0.95
Hong Kong	HKG	1.04	1.11	0.98	0.97
Bahrain	BHR	0.96	1.11	0.97	0.95
Belgium	BEL	1.09	1.11	0.97	0.96
Italy	ITA	1.08	1.12	0.96	0.96
Top 10 highest adjustment factors					
Guinea	GIN	8.19	8.90	0.80	0.79
Togo	TGO	7.93	8.91	0.80	0.79
Chad	TCD	8.30	9.07	0.75	0.74
Rwanda	RWA	8.38	9.46	0.69	0.68
Ethiopia	ETH	8.04	9.54	0.79	0.77
Gambia, The	GMB	9.11	9.71	0.71	0.71
Burundi	BDI	8.28	10.02	0.75	0.74
Mozambique	MOZ	9.49	10.61	0.75	0.74
Liberia	LBR	10.76	11.62	0.73	0.72
Zimbabwe	ZWE	15.41	16.96	0.70	0.69

A full set of country results is shown in Appendix A. Here we show some results for the top-10 countries with lowest and highest adjustment factors in Table 3. The ranking of countries is based on their average adjustment factor. The two adjustment factors sometimes show noticeable differences but the labor cost of government workers represents a modest share of GDP, so the difference between the stock-adjusted and services-adjusted GDP per capita results are much more modest. The biggest adjustment is for Rwanda for which GDP is 32%

lower when government wages are adjusted for differences in capital services. For 29 countries out of 146 countries the adjustment is between 20 and 32%.[8]

Concluding remarks

Comparing like-for-like is a crucial element of the ICP program and this is particularly challenging when it comes to government services, which relies on comparing inputs volumes rather than output volumes. To put countries on a more comparable basis, we should account for differences in capital input across countries and in this paper, we outlined and implemented a methodology for this. By and large, the productivity adjustment is similar in approach and outcomes to earlier exercises. However, we have developed capital input per worker measures for all ICP 2005 countries and estimated capital shares based on data for a broader range of countries. This allows for a more fine-grained adjustment and the now-established estimation framework will facilitate a similar exercise based on ICP 2011 results when these are available. But before that, there are number of issues for discussion and further steps are needed before this could be completed.

Questions for discussion

1. *Capital stocks or capital services*? As discussed above, relative capital services measures are conceptually appealing but involve more data and may be prone to more measurement error. The results show somewhat larger adjustment factors based on capital services than capital stocks, but the final GDP per capita numbers differ by no more than about 3 percent.
2. *Continuous versus discrete distribution of adjustment factors?* We use the adjustment factors directly from our capital computations while Heston (2012) made this distribution discrete by assigning the same adjustment factor to all countries in a group defined by their capital contribution. On the one hand, small differences between adjustment factors may well be due to measurement error in the capital data. On the other hand, a discrete distribution involves an arbitrary number of groups and possibly severe jumps in adjustment factors when moving from one group to the next.
3. *Capital input or labor productivity in the private sector?* In this paper, we have followed earlier approaches to labour productivity adjustment by estimating capital input. This has two downsides. First, capital stock or capital services estimation is data-intensive; and second, we cannot make any adjustments for differences in efficiency levels in the use of capital and labour. An alternative would be to measure labor productivity in the private sector and adjust government services wages using relative labor productivity (see e.g. Sergueev 1998). This would involve computing a private sector PPP based on the basic headings for household consumption and investment and collecting information about employment in the private sector.[9] An advantage would be

[8] One noticeable feature in Table 3 is that Japan's GDP per capita is adjusted downwards despite a capital stock adjustment factor lower than one. This seems an artifact of the GEKS index number calculations: if the results are computed based only on data for Japan and the US, Japan's GDP per capita would be adjusted upwards.

[9] Ideally this should be corrected for by import and export PPPs to arrive at the sectoral output measure needed.

that this is computationally less intensive than capital stock or services measurement, though of course suitable employment figures would need to be collected which is far from straightforward (Sergueev, 1998). A disadvantage could be that we would be over-adjusting wages, because efficiency differences in the (mostly tradable) private sector are likely to be larger than in the (mostly non-tradable) government sector.

Way forwards to ICP 2011

Based on the outcomes for each of the discussion questions, there would be a clear method for making a productivity adjustment to the wages in ICP 2011.

- The work on productivity adjustment cannot properly start without the basic heading PPPs and expenditure levels for investment goods. The PPPs are needed for comparing capital input across countries and the expenditure levels serve as a benchmark for the investment series.
- For the countries currently covered, we would need to extend the capital and labor data from 2005 to 2011 based on the ICP 2011 investment data and National Accounts up to 2011.
- In addition, country coverage would need to be extended to ensure that all 200 countries covered in ICP 2011 have capital data. This would typically involve implementing the commodity flow method, since detailed National Accounts data would typically not be available for this of countries. It is unsure whether the detailed type of data needed for this is available for these "new" countries and ways around it have to be tried.
- The adjustment factors would need to be shifted to a regional basis, since the global adjustment as described in this paper would not be applicable within ICP.

Appendix A Table of results

Appendix B Data Documentation

References

Balassa, B. (1964), "The purchasing-power parity doctrine: a reappraisal" *Journal of Political Economy* 72: 584–596.

Blades, D. (2012), "Health and Education", in World Bank (ed.) *Measuring the Size of the World Economy*, Washington, DC: World Bank, chapter 11.

Caselli, Francisco and Daniel J. Wilson (2004), "Importing technology" *Journal of Monetary Economics* 51: 1-32.

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

Diewert, W. Erwin (2001), "Measuring the Price and Quantity of Capital Services under Alternative Assumptions," *Department of Economics Working Paper* No 01-24, University of British Columbia.

Feenstra, Robert C. and John Romalis (2012), "International Prices and Endogenous Quality" *mimeo*.

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

Heston, Alan (2012), "Government Services Productivity Adjustments" in World Bank (ed.) *Measuring the Size of the World Economy*, Washington, DC: World Bank, chapter 16.

Inklaar, Robert and Marcel P. Timmer (2008), "GGDC Productivity Level Database: International Comparisons of Output, Inputs and Productivity at the Industry Level" *GGDC Research Memorandum*, GD-104.

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.

OECD (2009), *OECD Manual: Measuring Capital*, Paris: OECD.

Samuelson, P. A. (1964), "Theoretical notes on trade problems" *Review of Economics and Statistics* 46(2): 145-54.

Schreyer, Paul (2012), "Output, outcome, and quality adjustment in measuring health and education services" *Review of Income and Wealth* 58(2): 257-278.

Sergueev, S. (1998), "Comparison of Non-Market Services at Cross Roads (experience, considerations, proposals)", backgruond paper for SCECE, June 1998.

World Bank. 2008. *Global Purchasing Power Parities and Real Expenditures: 2005 International Comparison Program.* Washington, DC: World Bank. http://siteresources.worldbank.org/ICPINT/Resources/icp-final.pdf.

Appendix A

ISO code	Adjustment factor		Adjusted GDP relative to unadjusted	
	Stock	Services	Stock	Services
ALB	1.82	2.04	0.89	0.87
AGO	3.24	3.67	0.85	0.83
ARG	1.40	1.64	0.92	0.90
ARM	2.53	2.97	0.86	0.84
AUS	1.07	1.07	0.95	0.95
AUT	1.11	1.13	0.95	0.95
AZE	2.74	3.04	0.84	0.83
BHR	0.96	1.11	0.97	0.95
BGD	4.18	4.87	0.85	0.83
BLR	2.07	2.38	0.88	0.86
BEL	1.09	1.11	0.97	0.96
BEN	6.13	6.91	0.81	0.80
BTN	1.86	2.22	0.88	0.86
BOL	2.98	3.41	0.84	0.83
BIH	2.22	2.38	0.88	0.87
BWA	1.39	1.37	0.93	0.93
BRA	1.87	2.10	0.89	0.88
BRN	1.05	1.18	0.94	0.92
BGR	1.97	2.15	0.88	0.87
BFA	5.90	6.67	0.80	0.79
BDI	8.28	10.02	0.75	0.74
KHM	5.67	6.63	0.78	0.76
CMR	4.39	4.64	0.83	0.82
CAN	1.16	1.17	0.95	0.95
CPV	2.18	2.34	0.87	0.86
CAF	5.82	6.86	0.82	0.80
TCD	8.30	9.07	0.75	0.74
CHL	1.69	1.84	0.90	0.89
CHN	2.57	3.04	0.86	0.84
COL	1.89	2.25	0.89	0.87
COM	4.55	5.44	0.80	0.78
COD	5.59	6.55	0.83	0.82
COG	4.79	5.11	0.82	0.81
CIV	7.23	7.51	0.78	0.78
HRV	1.31	1.45	0.93	0.91
CYP	1.28	1.46	0.94	0.93
CZE	1.32	1.41	0.92	0.91
DNK	1.18	1.16	0.95	0.95
DJI	3.31	3.84	0.81	0.80

ECU	1.92	2.15	0.88	0.87
EGY	2.99	3.22	0.84	0.83
GNQ	3.37	3.43	0.86	0.85
EST	1.55	1.64	0.90	0.90
ETH	8.04	9.54	0.79	0.77
FJI	2.76	2.83	0.84	0.84
FIN	1.14	1.19	0.95	0.94
FRA	1.09	1.14	0.96	0.96
GAB	1.66	1.76	0.89	0.88
GMB	9.11	9.71	0.71	0.71
GEO	2.24	2.58	0.87	0.86
DEU	1.17	1.23	0.95	0.94
GHA	4.68	4.71	0.83	0.82
GRC	1.16	1.23	0.95	0.94
GIN	8.19	8.90	0.80	0.79
GNB	4.39	5.49	0.82	0.80
HKG	1.04	1.11	0.98	0.97
HUN	1.39	1.50	0.92	0.91
ISL	1.05	1.13	0.96	0.96
IND	3.75	3.66	0.83	0.83
IDN	2.54	3.05	0.86	0.85
IRN	1.33	1.55	0.91	0.89
IRQ	2.63	2.89	0.85	0.84
IRL	1.28	1.30	0.94	0.94
ISR	1.24	1.30	0.94	0.94
ITA	1.08	1.12	0.96	0.96
JPN	0.96	1.01	0.95	0.94
JOR	1.60	1.86	0.91	0.89
KAZ	1.95	2.20	0.88	0.87
KEN	6.45	6.78	0.78	0.77
KOR	1.33	1.41	0.92	0.92
KWT	1.16	1.35	0.97	0.95
KGZ	4.21	4.87	0.80	0.78
LAO	4.39	4.96	0.74	0.73
LVA	1.70	1.82	0.90	0.89
LBN	1.03	1.19	0.95	0.94
LSO	2.91	3.46	0.86	0.84
LBR	10.76	11.62	0.73	0.72
LTU	1.69	1.87	0.90	0.88
LUX	0.93	0.96	1.00	0.99
MAC	1.24	1.40	0.92	0.91
MKD	1.39	1.61	0.92	0.90
MDG	6.77	7.53	0.81	0.80
MWI	5.22	6.00	0.84	0.82

MYS	1.52	1.58	0.91	0.90
MDV	2.14	2.27	0.87	0.86
MLI	5.47	5.76	0.79	0.78
MLT	1.25	1.37	0.94	0.93
MRT	2.71	3.05	0.86	0.84
MUS	1.65	1.77	0.90	0.89
MEX	1.54	1.67	0.91	0.90
MDA	2.78	3.34	0.85	0.84
MNG	1.75	2.11	0.90	0.87
MNE	1.62	1.87	0.91	0.89
MAR	2.83	2.79	0.85	0.86
MOZ	9.49	10.61	0.75	0.74
NAM	1.74	1.84	0.91	0.90
NPL	4.17	5.58	0.82	0.80
NLD	1.19	1.19	0.94	0.94
NZL	1.43	1.43	0.92	0.92
NER	5.68	6.54	0.82	0.80
NGA	4.89	5.10	0.82	0.82
NOR	1.13	1.19	0.95	0.94
OMN	1.26	1.29	0.94	0.94
PAK	3.09	3.61	0.86	0.84
PRY	2.68	3.08	0.86	0.85
PER	1.73	1.98	0.90	0.88
PHL	2.63	3.03	0.87	0.85
POL	1.51	1.65	0.91	0.90
PRT	1.33	1.42	0.94	0.94
QAT	1.11	1.19	0.92	0.91
ROU	1.91	2.08	0.88	0.87
RUS	1.75	1.98	0.89	0.88
RWA	8.38	9.46	0.69	0.68
STP	3.65	3.93	0.84	0.83
SAU	1.13	1.27	0.96	0.95
SEN	3.96	4.38	0.83	0.82
SRB	1.83	2.17	0.89	0.87
SLE	7.91	8.90	0.78	0.77
SGP	0.95	1.00	0.92	0.92
SVK	1.29	1.32	0.91	0.91
SVN	1.29	1.37	0.94	0.93
ZAF	2.18	2.15	0.89	0.89
ESP	1.11	1.20	0.95	0.94
LKA	2.54	2.90	0.87	0.85
SDN	5.61	5.92	0.83	0.82
SWZ	1.99	2.20	0.89	0.88
SWE	1.34	1.24	0.93	0.94

CHE	1.14	1.08	0.95	0.95
SYR	2.56	2.83	0.86	0.85
TWN	1.22	1.22	0.92	0.91
TJK	4.51	5.64	0.76	0.74
TZA	6.47	7.07	0.83	0.83
THA	1.94	2.03	0.88	0.87
TGO	7.93	8.91	0.80	0.79
TUN	1.83	1.90	0.90	0.89
TUR	1.76	1.77	0.90	0.89
UGA	7.29	8.34	0.76	0.75
UKR	2.11	2.39	0.87	0.86
GBR	1.31	1.32	0.94	0.94
USA	1.00	1.00	1.00	1.00
URY	1.46	1.62	0.91	0.90
VEN	1.57	1.75	0.90	0.89
VNM	3.39	4.06	0.81	0.79
YEM	1.93	2.37	0.88	0.86
ZMB	4.97	5.25	0.83	0.82
ZWE	15.41	16.96	0.70	0.69

Appendix B Data documentation

This appendix provides information on the methods and sources of data used in deriving the investment numbers of structures, machinery and transport equipment at current prices. It begins with a brief description of the key features of the data and the major goal of this project, followed by in-depth discussion on how the data are combined and further processed for investment estimations.

Table B1 Types of data, origins and availability

Name	Provider	Availability*
ECLAC National Accounts	Economic Commission for Latin America and the Caribbean	32 countries; 1950-2008
EU KLEMS	Groningen Growth and Development Center	16 countries; 1970-2007
International Comparison Program (ICP)**	World Bank	168 countries; 1970-2005
UNIDO (INDSTAT rev.3)	United Nations Industrial Development Organization	180 countries; 1963-2003
OECD National Accounts	Organization for Economic Cooperation and Development	40 countries; 1950-2011
UN Comtrade	United Nations	196 countries; 1962-2011
UN National Accounts	United Nations	211 countries; 1950-2010
World Trade Flows	Center for International Data at	201 countries; 1963-2000
World Trade Flows Updated	UC Davis	countries; 1994-2008

* The availability of the years is not applicable to all countries. It only indicates that for one or more countries the data cover the entire time-span. For instance, in ECLAC national accounts only half of the countries have data available in 1950; while the other half starts to have data in much later years (e.g. 1977, 1978 etc.).

** For ICP data it is available quintennially between 1970 and 1985 and the last two available years are 1996 and 2005, of which the former year (i.e. 1996) do not disentangle machinery and transport equipment.

Goal of the project and the main method used

Since the goal of this project is to have time series of the investment estimates that go back as far in time as possible and have the coverage of country equal to that of the international comparison program (ICP), we complement the existing sources (i.e. ECLAC national accounts, EU KLEMS, and OECD national accounts) that already have data on investment by asset type using the commodity-flow method. In case of structures, we first use the actual investment number in structures that is covered in at least one ICP benchmark year; while for other years we use data on value added in the construction industry provided by UN national accounts to extrapolate.

As for machinery and transport equipment, the following estimation equation is applied since for many countries most of these assets are imported:

$$\hat{I}_{it} = Y_{it} - X_{it} + M_{it}$$

where Y is gross output, X are exports and M are imports. Gross output is available from UNIDO INDSTAT and exports and imports can be obtained from either United Nations Comtrade database or Feenstra's World Trade Flows. We opt for the former source and use Feenstra's data only in case of absence of a certain country in UN Comtrade. The main reason for this decision is because there is a wide disparity between the old Feenstra's data and the updated one. Take Australia and Panama as two random examples. The numbers are of comparable size for Australia during the overlapping time period with the only exception of year 2000;[10] while for Panama exports from the old data source are consistently much larger than that in the updated source. We do not know how to interpret and reconcile this dramatic discrepancy between the two data sources. Thus, for consistency UN Comtrade is preferred.

Table B2 Comparison between WTFOld and WTFNew

	AUS		PAN	
	Exports-old	Exports-new	Exports-old	Exports-new
1994	716,263	705,036	917,485	632
1995	838,285	987,792	874,157	499
1996	1,123,345	1,348,689	11,661,824	231
1997	1,719,278	1,954,317	937,277	5
1998	1,255,998	1,490,084	386,594	217
1999	1,671,955	1,972,337	323,950	8
2000	3,137,684	1,847,332	659,214	83

Now we turn to detailed discussion on how trade and output data are further processed to enable investment estimations.

Trade data process

As argued earlier, trade data is obtained from UN Comtrade database for the time period of 1962 to 2011 for all the countries available.[11] The commodity classification codes used in downloading the data is Standard International Trade Classification (SITC) revision 1 since this classification goes furthest back in time (i.e. 1962). With SITC rev.1 at two-digit level we are interested in the following four codes: 71, 72, 73 and TOTAL, of which we label the sum of the first two codes as trade in machinery, 73 as trade in transport equipment. TOTAL indicates the total amount of trade across all industries.

In order to have a balanced panel we first fill in the gaps that exist in the data. That is to say, for each and every country-industry pair we have time series covering the entire period of 1962 to 2011. One immediate consequence of this, however, is that there are 2120 missing values (or about 10.8% of total number

[10] A further check shows that there is a large difference between the two data sources in 2000 for nearly all countries covered in world trade flows. Thus, the observed data inconsistency is not an isolated event unique to Australia in 2000.

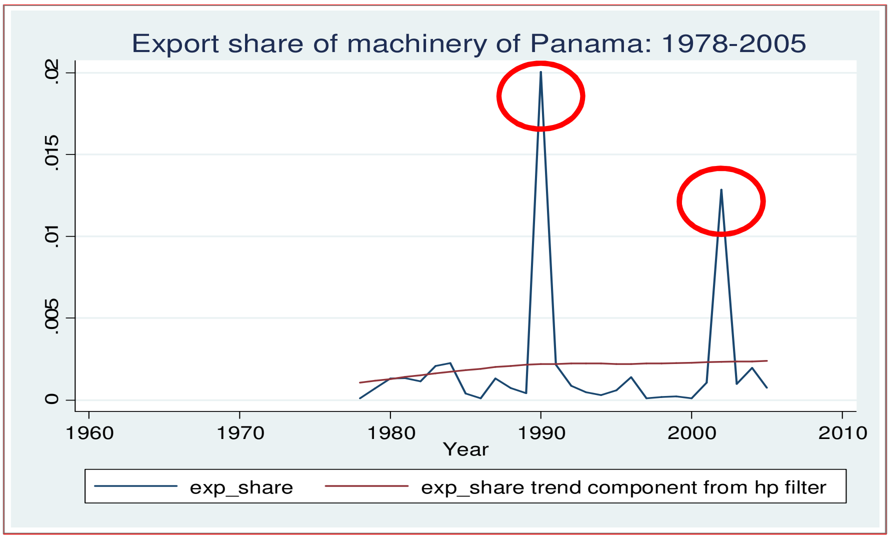
[11] Due to the lack of data on Equatorial Guinea (GNQ) and Sao Tome and Principe (STP), we obtain the trade data of these two countries from the World Trade Flows compiled by Feenstra (2005).

of observation) generated. As a quick fix, we apply the linear interpolation technique, which takes the following general form:

$$X = X_a + (X_b - X_a)\frac{(t-t_a)}{(t_b-t_a)}; \quad M = M_a + (M_b - M_a)\frac{(t-t_a)}{(t_b-t_a)}$$

where X indicates exports, M indicates imports and the subscripts (i.e. a and b) indicate the years at which the trade data are available; t denotes year. This interpolation is easily done in Stata (the main statistical software employed in this project) and the same calculating procedure applies whenever interpolation is used again in this project.

Figure B1



After filling in the missing values we proceed to compute the share of trade in machinery and transport equipment. Surprisingly, we find abnormally large jumps of these two computed shares. For illustration purposes, again take Panama as an example. There are two outstanding spikes in the export share of machinery between 1978 and 2005 (circled in red in Figure B1).

By contrast, its corresponding import share appeared to be much less volatile. To be precise, the largest exports share of machinery is about 240 times larger than its smallest share; whereas for imports share, this factor is around merely 2.12. In order to remove those undesirable spikes, Hodrick-Prescott (HP) time-series filtering technique is applied. Although the rule of thumb is to apply the smoothing parameter λ=6.25 for annual data, we used λ=1600 instead to obtain a more sensitive detector of outliers by allowing for a much smoother trend. The criterion used to identify the outliers is ad-hoc. We consider observations that have the cyclical components in the top and bottom five percentiles as outliers. Thus, as shown in Table B3 whenever the cyclical components exceed |5%| for imports and |3%| for exports they are considered as outliers whose values are then replaced by the product of their (HP generated) trend shares and total -imports or -exports.

Table B3, Distribution of the Cyclical Components after HP Filtering

cyclical components of imports share					cyclical components of exports share				
	Percentiles	Smallest				Percentiles	Smallest		
1%	-.06728	-.19661			1%	-.06184	-.20362		
5%	**-.03841**	-.18751			5%	**-.02443**	-.19049		
10%	-.02742	-.1855	Obs	11928	10%	-.01474	-.1871	Obs	11543
25%	-.01333	-.18342	Sum of Wgt.	11928	25%	-.00514	-.18465	Sum of Wgt.	11543
50%	**-.00073**		Mean	-4.2e-12	**50%**	**-.00051**		Mean	-9.5e-13
		Largest	Std. Dev.	.02705			Largest	Std. Dev.	.02223
75%	.01149	.26974			75%	.00335	.28523		
90%	.02854	.3725	Variance	.00073	90%	.0141	.32598	Variance	.000494
95%	**.04189**	.44743	Skewness	1.5021	95%	**.02645**	.34909	Skewness	2.945
99%	.07618	.45242	Kurtosis	25.047	99%	.07485	.42308	Kurtosis	51.0454

In addition, in order to accomplish the goal of having time series that go as far back as possible we extrapolate the trade data all the way back to 1950 using data on gross capital formation (GCF) provided by UN national accounts.[12] That is:

$$X_{it-1} = X_{it} * \frac{(GCF^{NA})_{t-1}}{(GCF^{NA})_t}; \qquad M_{it-1} = M_{it} * \frac{(GCF^{NA})_{t-1}}{(GCF^{NA})_t}$$

where *I* indexes asset type, namely *machinery* and *transport equipment*; *t* denotes year

Output data process

For production we rely on the data provided by United Nations Industrial Development Organization (UNIDO) compiled in 2006. In this data set the commodity classification codes used is International Standard Industry Classification (ISIC) revision 2 at three-digit level. The following five codes are of interest to us: 300, 382, 383, 384, and 385. The first code (i.e. 300) denotes the total output value of the manufacturing industry, 384 denotes the output value of transport equipment, and the sum of the rest three (i.e. 382+383+385) shows the output of machinery. Similar to the trade data, we first obtain a balanced panel by filling in the gaps. Since missing values can be found virtually in all countries[13], in addition to using interpolation technique to fix 'in-between' missing values we also use the data on value added in manufacturing from UN national accounts (i.e. D^{NA}) to fill in the missing data from the 'two ends' by extrapolation. In equation terms, we extrapolate the total output value of manufacturing (denoted by D^{UNIDO}) as follows:

[12] Not all countries have data on gross capital formation back to 1950, thus trade data is extrapolated backwards only for those that have such data available in UN national accounts.
[13] With a balanced panel there are, in fact, more observations with missing values than otherwise (i.e. 3451 missing versus 3068 observations that have data available).

Moving backwards: $(D^{UNIDO})_{t-1} = (D^{UNIDO})_t * \frac{(D^{NA})_{t-1}}{(D^{NA})_t}$

Moving forwards: $(D^{UNIDO})_{t+1} = (D^{UNIDO})_t * \frac{(D^{NA})_{t+1}}{(D^{NA})_t}$

After fixing D^{UNIDO}, we fill in missing values for output in machinery and transport equipment, respectively, by keeping the last observable share constant and multiply it with their year-specific output. That is:

Moving backwards: $Y_{it-1} = (Y_{share})_{it}(D^{UNIDO})_{t-1}$

Moving forwards: $Y_{it+1} = (Y_{share})_{it}(D^{UNIDO})_{t+1}$

where subscript i denotes the asset type (i.e. machinery and transport equipment); t denotes year.

Similar to the attempt of removing outliers in trade data, we applied the same filtering technique here to correct for spikes in output. We first merge the value of GDP from UN national accounts and define a share as $share = (Y^{UNIDO})/GDP^{NA}$. Whenever the cyclical components of this share and the share itself exceeds one we identify it as a outlier, which is then replaced by the product of the (HP generated) trend share and the value of GDP.[14] In addition, we also use data on GDP^{NA} to extend the time series back to 1950 by extrapolation:

$$Y_{t-1} = Y_t * \frac{(GDP^{NA})_{t-1}}{(GDP^{NA})_t}$$

Calculating investments for machinery and transport equipment

After processing the trade and output data we are now ready to apply the commodity flow method to compute investment series by asset type (i.e. machinery and transport equipment) with the following equation:

$$\hat{I}_{it} = Y_{it} - X_{it} + M_{it}$$

Since machinery and transport equipment are generally considered high-tech products, it is very likely that the least-developed countries do not produce but mostly import them from other countries. Thus, output Y is set to zero whenever the country-industry pair under concern has missing data for the entire time period. That is to say, for those countries with zero output their investment in either machinery or transport equipment are simply the net difference between imports and exports. One exception to note is Belarus, which has its exports volume consistently larger than that of imports. After the failure of finding the output data for Belarus from alternative sources (e.g. national bureau of statistics), we resort to using the shares of the Russian data combined with

[14] As the rule of thumb used in this project we would normally identify observations that are in the top and bottom five percentiles as outliers but this criterion is not applied here, since in retrospect we find there is only one observation has an extreme jump in output. As we aim to stay as close as possible to the original data, the criterion used here only identifies this particular observation as the outlier and correct it with the trend share multiplied by GDP^{NA}, while leaving the rest of the output data intact.

Belarus's own number of value added in manufacturing (i.e. D^{NA}) as a crude measure to derive output for Belarus.[15]

In less than 0.2 percent of the cases we find negative investment estimations and these negative values are either due to a small import share or large export share. In the former case, we applied the mean share of their imports to correct for the small values of imports; while in the latter case, we replace the exports value to missing and then interpolate the values in between. Though these correction measures are rather crude, it helps to get rid of the negative investment values and yield somewhat more plausible numbers.

Calculating shares of structures, machinery and transport equipment using ICP data

As already mentioned, we use data on value added in the construction industry to proxy for investment in structures. More specifically, we first apply the structure share from the international comparison program (ICP) as the benchmark share and then with six waves in ICP (i.e. 1970, 1975, 1980, 1985, 1996 and 2005) we obtain six benchmark values of investment in structures:

$$I_S = SS^{ICP} * GFCF^{NA}$$

where I_S denotes investment in structures, SS^{ICP} indicates the actual share of structures provided by ICP and $GFCF^{NA}$ denotes total gross fixed capital formation in national accounts.[16]

Using data on value added in construction, we define a ratio (i.e. $I_S/V^{NA}_{Construc}$) which is interpolated and then multiplied by $GFCF^{NA}$ to obtain a complete time series of investment in structures running from 1970 to 2005. For years before 1970 and after 2005 we extrapolate it as follows:

$$\text{Moving backwards: } (I_S)_{t-1} = (I_S)_t * \frac{(V^{NA}_{Construc})_{t-1}}{(V^{NA}_{Construc})_t}$$

$$\text{Moving forwards: } (I_S)_{t+1} = (I_S)_t * \frac{(V^{NA}_{Construc})_{t+1}}{(V^{NA}_{Construc})_t}$$

After obtaining the investment values of structure, we continue to compute its share by dividing total GFCF. Doing so, implausible numbers arise. Some of the structure shares exceeded unity, which is clearly insensible. To correct for this, HP filtering technique is again applied to detect the abnormal jumps of value added in construction (i.e. $V^{NA}_{Construc}$). We consider observations that are in the top and bottom five percentiles as outliers. Thus, as shown in Table B4 whenever the cyclical components exceed |10%| the values are replaced by the product of their (HP generated) trend shares and $GFCF^{NA}$.

Table B4 Distribution of the Cyclical Components for Output Shares

HP generated cyclical components

[15] One could also argue to use the shares of another country (e.g. Ukraine) rather than Russia, since this is after all an arbitrary choice.

[16] For ease of exposition, SM^{ICP} indicates share of machinery, ST^{ICP} indicates share of transport equipment and SMT^{ICP} represents share of machinery and transport equipment combined. The superscript denotes the source of data.

	Percentiles	Smallest		
1%	-.20326	-.88514		
5%	-**.09068**	-.83112		
10%	-.05658	-.82781	Obs	8027
25%	-.02377	-.77994	Sum of Wgt.	8027
50%	**-.00256**		Mean	1.51e-11
		Largest	Std. Dev.	.08744
75%	.01854	1.1041		
90%	.05083	1.2050	Variance	.00765
95%	**.08477**	1.1632	Skewness	3.8189
99%	.26225	1.6709	Kurtosis	70.499

However, although this measure helps to mitigate the problem there is still a large number of cases with structure shares larger than one. Thus, a second attempt is made by setting $V_{Construc}^{NA}$ values to missing whenever structure shares exceed unity, which are then interpolated or extrapolated. This correction measure does not help to fully solve the problem but the number of cases with structure share exceeding unity reduced sizably by more than two-thirds. In order to stay as close as possible to the original data we content ourselves with these two correction attempts and leave the other unsolved large structure shares as they are.

To derive the shares of machinery and transport equipment using ICP data we follow the exact same approach as above-described but with $V_{Construc}^{NA}$ substituted by $\hat{I}_M$ for machinery and $\hat{I}_T$ for transport equipment; I_S replaced by I_M and I_T; and SS^{ICP} replaced by SM^{ICP} and ST^{ICP}.

Combing structure, machinery and transport equipment shares

With the structure, machinery and transport equipment shares computed, we combine them and sum them up. Ideally, we would want the sum being equal to one but this is unsurprisingly not the case. Therefore, in order to rescale the sum of these three shares back to one we divide each of the shares by the sum. For illustration purposes, lets consider Azerbaijan. In 1991, Azerbaijan has the most implausible estimations. Its structure share is estimated at 1.166, machinery share at 1.737 and transport equipment share at 0.134. Thus, the sum of these three shares is equal to 3.037. To rescale it back to unity, the structure share is calculated as 1.166 divide by 3.037 (38.4%). Analogously, machinery share is 57.2%, and the residual is the transport equipment share (4.4%). The following table compares the estimated shares of Azerbaijan in its worst year (i.e. 1991) with the benchmark numbers from ICP (i.e. 2005). The values undoubtedly differ, but such a rescaling already does a very decent job as the proportion of each share is in the 'right' order (i.e. machinery accounts for the largest share, followed by structure and transport equipment).

Table B5 Comparison of the distribution of the asset type shares of Azerbaijan

	1991	2005
Structure share	38.4%	43.2%
Machinery share	57.2%	46.3%
Transport equipment share	4.4%	10.5%

4.2 Integrate data from different sources

To obtain a complete list of investment shares by asset type we integrate the data from different sources. As shown in table 5, we start with OECD national accounts data and complement that with EU KLEMS, which is further complemented by ECLAC and commodity-flow data.

Table B 6 Integration of the data

Order of integration	Source	
1	OECD NA	*extrapolate using*
2	EU KLEMS	*extrapolate using*
3	ECLAC	*extrapolate using*
4	ICP/CFM	

For illustration, we consider how OECD national accounts data are complemented by data from EU KELMS:

Moving backwards:

$$(S_i^{OECD\ NA})_{t-1} = (S_i^{OECD\ NA})_t * \frac{(S_i^{EU\ KLEMS})_{t-1}}{(S_i^{EU\ KLEMS})_t};\quad i{\sim}(Mach, TraEq, Struc)$$

Moving forwards:

$$(S_i^{OECD\ NA})_{t+1} = (S_i^{OECD\ NA})_t * \frac{(S_i^{EU\ KLEMS})_{t+1}}{(S_i^{EU\ KLEMS})_t};\quad i{\sim}(Mach, TraEq, Struc)$$

The same extrapolation procedure carries through when the data are integrated from the other sources. In addition, there are two issues warrant further explanation. First, asset types are distinguished at a much more disaggregated level in EU KLEMS. Thus, data on machinery from EU KLEMS are aggregated by summing up *IT*, *CT*, *Other machinery* and *Software*. On the contrary, ECLAC only identifies assets dichotomously, namely structures and machinery and transport equipment combined. In order to isolate machinery from transport equipment we use the following extrapolation technique[17]:

Step 1: compute the combined machinery and transport equipment share using CFM data

$$SMT^{CFM} = SM^{CFM} + ST^{CFM}$$

[17] This splitting technique is also applied to OECD national accounts data when needed. For instance, the value of transport equipment is embedded in other machinery for Australia. We first rely on EU KLEMS data to spilt transport equipment from other machinery and for years that are not covered in EU KLEMS we apply the shares from commodity-flow estimation.

Step 2: compute the respective share of machinery and transport equipment using CFM data

$$\widehat{SM}^{CFM} = \frac{SM^{CFM}}{SMT^{CFM}}; \qquad \widehat{ST}^{CFM} = \frac{ST^{CFM}}{SMT^{CFM}}$$

Step 3: apply the share to isolate machinery and transport equipment in ECLAC

$$SM^{ECLAC} = \widehat{SM}^{CFM} * SMT^{ECLAC}; \qquad ST^{ECLAC} = \widehat{ST}^{CFM} * SMT^{ECLAC}$$

Investment numbers of structures, machinery and transport equipment at current prices

Having computed the shares of three assets, we multiple them with total gross fixed capital formation denominated in current prices. That is:

$$I_{it} = S_{it} * GFCF^{NA}$$

where I denotes investment and S denotes share; subscripts *i* denotes asset type (i.e. structure, machinery, and transport equipment), and *t* denotes year. These are the investment series that are then used in estimation capital stocks.