Total Economy Database

Sources & Methods

This version : August 2015

Klaas de Vries
Abdul Azeez Erumban

Contents

List of tables

Introduction

This document describes the sources and methods used to construct the variables presented in The Conference Board Total Economy Database™. As such, it provides a general overview. It discusses the construction of individual variables in the database covering the period 1950-2014. Estimates for the ongoing year 2015 are forecasts based on insights from our in-house experts, or based on forecasts from other organizations such as the European Commission, the Organization for Economic Cooperation and Development (OECD) and the International Monetary Fund (IMF).

The Total Economy Database™ (TED) is a comprehensive database with annual data covering Gross Domestic Product (GDP), population, employment, hours, labor quality, capital services, labor productivity, and total factor productivity for 128 countries in the world, including 5 former economies. TED was originally developed by the Groningen Growth and Development Centre (University of Groningen, The Netherlands) in the early 1990s. Beginning in the late 1990s, it was produced in partnership with The Conference Board. In 2007, the database was transferred from the University of Groningen to The Conference Board, which has maintained and extended the database since then. In January 2010, The Conference Board expanded the database by adding a module on sources of growth, including labor quantity and quality, capital services (non-ICT and ICT), and total factor productivity (TFP). The extended module aims to integrate two previous data sets, namely the world economy productivity data set created by Dale Jorgenson and Khuong Vu of Harvard University and the Total Economy Growth Accounting Database of the Groningen Growth and Development Centre.

The first section of this document covers the sources and methods used to construct the basic productivity and income variables. These variables are GDP, employment, hours worked and population. Time series data for these variables are available from the file 'Output, Labor and labor Productivity', on the TED webpage. The second section focuses on the growth accounting variables, which are the factors that contribute to GDP growth as identified by the growth accounting method. The data, which cover 120 countries, are available from the file 'Growth Accounting and Total Factor Productivity'. The third and last section outlines the method of aggregation used to obtain regional growth rates, as available from the file 'Regional Aggregates'.

1 – Basic productivity and income variables

1.1 Gross Domestic Product

1.1.1 Basic sources used

The Total Economy Database (TED) uses GDP, valued at market prices whenever possible, as a measure of output in a country. GDP statistics are available for most countries in the world and are collected and constructed by national statistics agencies, using international guidelines set by the United Nations (UN) in their *System of National Accounts* handbook (UN, 2009). This assures a certain level of consistency, even though the quality of the data may differ between countries.

The historical GDP series in TED for the period before 1991 are obtained from Maddison Historical Statistics (Maddison, 2010). These estimates are extrapolated forward until 2014 using official estimates of GDP, which are freely available from various national (country-specific) or international sources. Table 1 provides a list of the general sources used to construct the time series of GDP. Estimates for the ongoing year 2015 are forecasts by The Conference Board, using country-specific insights from our experts.

Table 1 – Sources used to construct GDP time series (1990-2014)

Region	*Source*
Europe	Eurostat – Database
	UN Economic Commission for Europe – Statistical Database
Other high income countries	Organization for Economic Co-Operation and Development – StatExtracts
Asia	Asian Development Bank – Statistical Database System
Latin America	UN Economic Commission for Latin America – CEPALSTAT
Africa, Middle East	UN National Accounts Main Aggregates Database

1.1.2 Converting GDP data to a common denominator

National Statistics Offices prepare GDP estimates on a national currency basis. As such they are not internationally comparable, making it necessary to convert them into a common currency unit. Exchange rate conversion is an easy method of translating national currency GDP and other economic indicators into a common currency. However, exchange rates are not reflective of purchasing power differences across countries, and therefore, a more pertinent approach is to use purchasing power parities (PPP). PPPs are seen as better capturing the 'true' value of what a dollar can buy in goods and services in any country, while exchange rates are set in the marketplace. In general, products and services are consumed and produced at lower prices in low-income countries than in high-income countries. Exchange rates fail to capture this, as they do not consider price differences across countries.

PPPs and exchange rates – an example

Suppose a tin of tuna fish of identical quality is purchased at U.S.$3 in the United States and at half a dinar in Bahrain, then one obtains a direct "purchasing power parity" (PPP) for canned tuna of 0.167 BHD to the U.S. dollar (0.5 BHD/3 U.S.$). Suppose the official exchange is 0.33 BHD per dollar, then one would understate the expenditure in Bahrain relative to the United States by 50 percent when making the comparison on the basis of the exchange rate rather than the PPP. In other words, the price level of canned tuna fish in Bahrain is 50 percent lower than in the United States. If this would be true for all expenditure, real income in Bahrain measured at PPP, would be twice as high as nominal income measured at the exchange rate. The computation of purchasing power parities (PPPs) at the aggregate level of GDP is done on the basis of a large sample of relative price comparisons for expenditure on representative products and services. See World Bank Group (2015), for a detailed discussion on PPPs and its construction.

Therefore, PPPs are preferred to exchange rates as a method of converting national currency values into a common denominator in productivity analysis. From a business point of view, PPPs are vital to understanding the market size or real standards of living of a country, while from a revenue perspective, exchange rate converted data are more useful.

In the TED, three GDP series – GDP GK, GDP EKS and GDP XR – are available. All three are measured in constant market prices and cover the period from 1950 to 2015. The first two are converted using PPPs while the latter is based on exchange rates. The ***GDP GK*** series, which is converted at 'Geary-Khamis' PPPs[1], is expressed in 1990 U.S. dollars and is available for all of the 128 countries present in TED, including former economies such as East and West Germany. Both ***GDP EKS*** and ***GDP XR*** series are given for 123 countries and are expressed in 2014 US dollars. Whereas GDP XR is presented in market exchange rates of 2014, obtained from the IMF World Economic Outlook (April 2015), GDP EKS is presented in 2014 PPP $.[2] The 2014 PPPs are based on the World Bank-ICP 2011 round, updated to express 2014 prices using the change in national GDP deflators. In order to construct current price GDP series in US$, we derive such PPPs for each year using changes in relative prices:

$$PPP_t^i = PPP_{t-1}^i * e^{\Delta ln\left(\frac{p^{GDP_t^i}}{p^{GDP_t^{US}}}\right)} \quad \text{For } t > 2011 \qquad (1\text{-A})$$

$$PPP_t^i = PPP_{t+1}^i / e^{\Delta ln\left(\frac{p^{GDP_{t+1}^i}}{p^{GDP_{t+1}^{US}}}\right)} \quad \text{For } t < 2011 \qquad (1\text{-B})$$

[1] The Geary-Khamis method is an aggregation method in which category "international prices" (reflecting relative category values) and country PPPs (depicting relative country price levels) are estimated simultaneously from a system of linear equations. It has the property of base-country invariance, matrix consistency and transitivity. For more details, see chap. V and annex II in UN (1992).

[2] The EKS method is a multilateral method developed by Eltoto, Kovacs and Szulc, that computes the nth root of the product of all possible Fisher indexes between n countries. It has been used at the detailed heading level to obtain heading parities, and also at the GDP level. EKS has the properties of base-country invariance and transitivity. For more details, see chap. V and annex II in UN (1992).

where PPP_t^i is the PPP between country i and the United States in year t and $\left(\frac{p^{GDP_t^i}}{p^{GDP_t^{US}}}\right)$ is the ratio of the GDP deflator in country i and in the United States and the subscripts. GDP deflators are obtained from the national accounts as the ratio of current and constant GDP in a given country.

1.1.3 Adjustments to ensure comparability

Some countries report GDP based on the (country specific) fiscal year, as opposed to using the calendar year which is common for most countries. In order to ensure comparability across countries, GDP data is adjusted – when necessary and possible – to refer to the calendar year for all countries in TED. There are two ways of adjusting fiscal year reported data. The most straightforward way is by using quarterly national accounts data, so that calendar year based data is constructed using data from the relevant quarters. When quarterly data is not available, we use the average of data for two fiscal years to represent the calendar year based value, so that

$$GDP_t^{CY} = \frac{GDP_t^{FY} + GDP_{t-1}^{FY}}{2} \qquad (2)$$

where CY denotes calendar year and FY fiscal year. Obviously, this method only provides relevant results when the country specific fiscal year starts at the middle of the calendar year. Table 2 presents an overview of the countries for which an adjustment is made.

Table 2 – Adjustments to fiscal year based GDP data

Country	*Start of Fiscal Year*	*Adjustment method*
Iran	March 21	Quarterly National Accounts data
India	April 1	Quarterly National Accounts data
New Zealand	April 1	Quarterly National Accounts data
Australia	July 1	Quarterly National Accounts data
Bangladesh	July 1	Average of two years annual data
Egypt	July 1	Average of two years annual data
Pakistan	July 1	Average of two years annual data
Ethiopia	July 8	Average of two years annual data

The following adjustment method ensures intertemporal consistency--in other words, it makes certain that individual time series are consistent over time in what they cover. Since the Total Economy Database is essentially an update and extension of Maddison Historical Statistics (Maddison, 2010) it still uses the same list of countries as in that database. However, ever since the Maddison database was constructed, some countries have undergone significant changes to their borders. These countries are Ethiopia, from which Eritrea seceded in 1993, Serbia & Montenegro, which separated into two independent states in 2006 and Sudan, which broke up into the Republic of Sudan and South Sudan in 2011. Data in the TED refer to the old borders of these countries, so that Ethiopia includes Eritrea, Serbia & Montenegro are one country, as is the Republic of Sudan and South Sudan. Real GDP growth rates for the period for which these countries no longer form a union are weighted using nominal GDP data.

1.1.4 GDP estimates for China and Argentina

In the case of two countries, China and Argentina, there has been concern over the reliability of the officially produced GDP data. Following arguments by many scholars that Chinese GDP growth rates are often overstated, we use an adjusted series from Maddison for the period 1950-2003.[3] However, the PPPs used in Maddison series were significantly lower than more up to date PPPs, and therefore, the GDP series from Maddison have been further adjusted downwards by a factor of 0.774. For the period after 2003 we use official GDP growth rates.

The data for Argentina is obtained from the ARKLEMS + LAND project, a research project that focuses on the measurement and analysis of the source of economic growth, productivity and competitiveness of the Argentinean economy. The TED relies on this source, because the quality of the Argentinian official data is often questioned.[4] In particular the government inflation figures are said to understate actual price growth.

Country specific sources may apply in some cases, the details of which are available from source notes provided on the Total Economy Database website:

http://www.conference-board.org/data/economydatabase/index.cfm?id=27770

1.2 Employment

1.2.1 Concept of employment used

From the perspective of productivity, the measure of employment should be consistent with the measure of output. Therefore, the employment figures should cover all persons engaged in some activity that falls within the production boundary of the system of national accounts. It should include employees, self-employed as well as unpaid family members that are economically engaged, apprentices and the military. In TED, whenever possible the domestic concept of employment is used, as it includes all workers employed domestically, but excludes any nationals working abroad, and hence is in line with the production boundary for GDP.[5]

[3] See: Maddison, Angus and Wu, Harry (2008). Measuring China's Economic Performance. *World Economics*, 9(2).
[4] See: Coremberg, A. (2014). Measuring Argentina's GDP Growth. World Economics, 15(1); http://arklems.org/
[5] Alternatively, one can use the national concept which counts all nationals working domestically and abroad, but excludes foreign workers employed domestically.

1.2.2 Basic sources used

The United Nations *System of National Accounts 2008* (UN, 2008, chapter 19) prescribes that national accounts statistics must also include compatible measures of employment, as described above. Both Eurostat, the European Union's statistical agency, and OECD report National Accounts data on employment submitted by member countries.[6] For several advanced economies in TED we use the employment figures reported under the National Accounts (domestic concept) in these databases. One notable exception is the United States. Employment estimates provided by the Bureau of Economic Analysis (BEA) in the framework of the National Accounts report number of jobs, and are therefore not comparable to statistics for other countries where employment is reported in number of persons. For the U.S. we utilize data on the employed (civilian) population from the Current Population Survey, available from the Bureau of Labor Statistics, and data on military personnel provided by the BEA.

For most other countries, the national accounts do not provide employment statistics, or at least do not publish employment data based on all persons engaged. In those cases we rely on Labor Force Survey (LFS) data. The concept of employment used in the LFS may differ from one country to the other, but in general it includes every person (usually aged 15 years and over) who worked for one hour or more during the reference period (usually last week), thus including own-account and unpaid family workers. Employment estimates derived from LFS may differ from national accounts based estimates, since the latter integrates information from many sources, of which the LFS is only one. Furthermore, the two sources also differ conceptually. While the national accounts make a distinction between the national and domestic concept of employment –and as described above we opt for the domestic concept– the LFS only covers resident households and thus the national concept. The difference between the national and domestic concept are cross-border workers, which is small for most countries, though small countries such as Luxembourg form notable exceptions. Another point of difference are the recording thresholds. National accounts do not exclude individuals from employment because of age, while the LFS does. In developed economies the difference is generally very small.

For a number of African and Latin American countries we use employment data provided by the Groningen Growth and Development Center (Timmer, de Vries and de Vries, 2014), and for some Asian countries we rely on the Asian Productivity Organization's *Asian Productivity Database*. The employment estimates from these databases are carefully constructed using (historical) data from population censuses, labor force surveys and establishment surveys, and pertain to all persons engaged in production and thus correspond closely with GDP estimates for these countries.

With regard to employment data for the remaining countries, which are mainly in the Middle East and Africa, we rely on ILO estimates of the employed population (ILO Key Indicators of the Labor Market 8th edition (KILM), table 2A). The procedure followed by the ILO is roughly as follows. Employment data from the KILM database is constructed using ILO estimates of the labor force (ILO, 2013), subtracted by

[6] These data are compiled via a joint Eurostat/OECD questionnaire. See 'Situation of Annual National Accounts in the OECD Database and New Features of the Joint OECD-EUROSTAT Questionnaire', available from www.oecd.org/dataoecd/9/29/24336184.doc.

the number of unemployed to arrive at total employment. Data on the unemployed are collected from LFS data whenever possible, and are otherwise estimated using econometric techniques (ILO, 2010).

Most countries do not provide employment series for the complete period as covered in the TED (1950-2015). We use benchmark estimates for 1950 from Maddison (1995) for 27 countries, and extend it for the missing years in between, and for the countries for which none of the above sources are available using ILO estimates of the Labor Force (ILO, 2013, 2011, 1996), which provide time series on the labor force covering all countries for the period 1950-2015.

Table 3 provides a list of the general sources used to construct the employment time series. Note that for a number of countries we use country-specific sources. For more details please refer to detailed source notes available from the Total Economy Database website.

Table 3 – Sources used to construct employment time series

Type of data	*Source*
National accounts based	Eurostat - Database
	Organization for Economic Co-Operation and Development (OECD)– StatExtracts
Labor Force Survey based	UN Economic Commission for Europe – Statistical Database
	Asian Development Bank – Statistical Database System
	UN International Labor Organization – ILOstat
	Haver analytics
Third-party constructed estimates	Groningen Growth and Development Center – 10 Sector Database
	Asian Productivity Organization – Asian Productivity Database
	ILO Key Indicators of the Labor Market 8th edition, Table 2A Estimates of the employed population

1.2.3 Adjustments to ensure comparability

The employment data for Australia and New Zealand are obtained from their national accounts, and are available for the fiscal year. Therefore, we adjust the fiscal year series using equation (2) – i.e. the average of two fiscal year series – to obtain calendar year data. As is evident from Table 2, calendar year GDP in these two countries are obtained by aggregating quarterly data, an approach which is not feasible in the case of employment due to lack of quarterly data. The second adjustment noted under section 1.1.3, referring to the intertemporal consistency, also holds for the employment estimates in TED, so that Ethiopia includes Eritrea, Serbia & Montenegro are one country, as is the Republic of Sudan and South Sudan.

1.3 Total hours worked

1.3.1 Concept of hours worked used

Estimates of working hours involve serious measurement problems and international comparability is difficult. Even individual countries often provide differing estimates, which are variously based on labor force surveys or establishment surveys. An advantage of estimates based on labor force surveys is that they are usually quite comprehensive. These estimates include adjustments for overtime, sickness, etc. A disadvantage, however, is that often they slightly overestimate actual hours worked. Figures based on establishment surveys usually only cover hours paid (which may include overtime), and require further adjustments to account for various types of absence.[7]

Our approach to hours worked data is similar to our employment data, as described in section 1.2.1. Our first choice is the national accounts based data on hours worked, which is generally available only for advanced countries in the TED. For a number of Asian and Latin American countries we rely on estimates from the Asian Productivity Organization's *Asian Productivity Database*, and publications from Maddison (1982, 1995, 2001), Crafts (1997) and Hoffman (1998). The measure used in TED is hours actually worked, so it includes paid overtime and excludes paid hours that are not worked due to sickness, vacation and holidays, etc.

1.3.2 Basic sources used

Series on hours actually worked per person are available for 66 countries. Table 4 displays a list of the general sources used to construct the time series of total hours worked.

Table 4 – Sources used to construct hours worked time series

Type of data	*Source*
National accounts based	Eurostat - Database
	Organization for Economic Co-Operation and Development – StatExtracts
Third-party constructed estimates	Groningen Growth and Development Center – 10 Sector Database
	Asian Productivity Organization – Asian Productivity Database
	Maddison (1982, 1995, 2001)
	Crafts (1997)
	Hoffman (1998)

1.2.3 Adjustments to ensure comparability

As in the case of GDP and employment, hours data for Australia and New Zealand are adjusted to reflect the calendar year, by taking the average of two fiscal year series (equation 2).

[7] The method for making such adjustments is described in A. Maddison (1980), "Monitoring the Labour Market: A Proposal for a Comprehensive Approach in Official Statistics", Review of Income and Wealth, June, pp. 175-217.

1.4 Population

The population figures reported in the TED refer to mid-year estimates, and are sourced from Maddison Historical Statistics (Maddison, 2010) for the period 1950-1990. For the period 1991-2015, we use the trends from the US Census Bureau *International Data Base (IDB)*. Table 5 displays a list of the general sources used to construct the population time series.

Table 5 – Sources used to construct population time series

Period	*Source*
1950-1990	Maddison Historical Statistics (Maddison, 2010)
1990-2015	US Census Bureau – International Data Base

1.5 Per capita income and labor productivity

So far we have discussed the concepts, sources and methods used to construct the basic variables in the Total Economy Database. These basic variables are used in turn to analyze per capita income and labor productivity, the two variables that are discussed in this subsection.

One of the most widely used measures to compare living standards between countries is GDP per capita or per capita income (PCI). It is calculated by dividing the total income earned in a country (as measured by GDP) by the total population i.e.:

$$PCI = \frac{GDP}{POP} \quad (3)$$

where GDP is Gross Domestic Product measured in PPP terms, either in 1990 GK or in 2014 EKS, and POP is total population. The use of PPP converted GDP is of vital importance for international comparisons of PCI, as it adjusts for the difference in price levels between countries, as outlined in section 1.1.2. Note, however, that by definition, PCI does not account for the inequality in the distribution of income, as it is a measure of average income per person as if total income is equally distributed.

The most important source of rising living standards, as measured for example by PCI, is productivity. As countries become more productive, they are essentially producing more value with the same effort. As such, labor productivity, or y, is a measure of efficiency, and can be measured either as output per hour or output per person, so that,

$$y = \frac{GDP}{L} \quad (4)$$

where L is labor input, measured either as persons employed or total hours worked.

Output per hour is the preferred measure of labor productivity since it measures labor intensity more effectively. For example, it adjusts for differences in average working time per person employed between countries. However, as stated in section 1.3.2, the availability of reliable and comparable data on total hours worked is limited, which explains why the TED presents hours worked data for roughly half of the countries in the database. The TED presents full coverage of data on persons employed, so that estimates of labor productivity based on GDP per person employed are available for all countries in the database.

2 – Growth accounting variables

2.1 Total Factor Productivity

The measure of labor productivity discussed in the previous section is a partial measure of productivity. A more sophisticated measure is Total Factor Productivity (TFPG), which takes into account not only labor as an input but also the contributions of physical, human and other intangible capital to the production of goods and services. As will become apparent later in this section, TFP is in essence a subcomponent of Labor Productivity. TFPG is not measured directly; rather, it is obtained as a residual after accounting for the contributions of all other factors of production to growth in output.

Assuming a neo-classical aggregate production function, aggregate value added (or GDP) growth can be decomposed into contributions from aggregate capital input (K), aggregate labor input (L) and aggregate total factor productivity (A) growth as:[8]

$$\Delta\text{lnGDP} = \bar{v}_K \Delta\text{lnK} + \bar{v}_L \Delta\text{lnL} + \Delta\text{lnTFP} \qquad (5)$$

where v_K and v_L are respectively the share of capital compensation and labor compensation in nominal GDP, both averaged over the current and previous years. Under constant returns to scale $v_K + v_L = 1$, so that the capital compensation share can be obtained by subtracting labor compensation from nominal value added. ΔlnK is the capital services growth rate and ΔlnL is the labor input growth rate. ΔlnTFP is the growth of total factor productivity.

Following Jorgenson (1963), capital services and labor input are measured as translog aggregates of heterogeneous types of capital and labor.

$$\Delta\text{lnK} = \sum_k \bar{s}_k \Delta\text{lnK}_k\,; \;\; \text{and } \Delta\text{lnL} = \sum_l \bar{s}_l \Delta\text{lnL}_l \qquad (6)$$

where s_k is the share of each type of capital k in total capital compensation, and s_l is the share of each type of labor l in total labor compensation, defined as:

$$s_k = \frac{P_{K,k} K_k}{\sum_K P_{K,k} K_k} \quad \text{and} \quad s_l = \frac{P_{L,l} L_l}{\sum_l P_{L,l} L_l} \qquad (7)$$

where $P_{K,k}$ is the rental price of capital type k, and $P_{L,l}$ is the price (wage rate) of labor type l. As before $\bar{s}$ in (6) is the two-period averages of these shares. In TED we distinguish between three types of labor (low-skilled; medium-skilled; and high-skilled), and six types of capital assets (computer hardware;

[8] This section draws on Jorgenson et al (2007) and Erumban and Das (2015).

computer software; telecommunications equipment; non-residential buildings and structures[9]; transport equipment; and non-ICT machinery). Assets hardware, software and communication equipment are considered as ICT assets, and all other assets are non-ICT assets.

In equation (6) the growth rate of each type of employee is weighted by their compensation share, so that it reflects both the quantity and the quality or composition effect (differences between different worker groups). In TED, we distinguish between the quantity effect (H) and the quality effect. Adding across different worker groups and obtaining their growth rates, we can obtain:

$$\Delta \ln H = \Delta \ln \sum_{l} L_l \qquad (8)$$

where H is the labor quantity measured either by employment or hours obtained by summing across different types of workers. The difference between labor input growth rates aggregated using equation (8) and the aggregation across worker types in equation (6) is labor quality (or the labor composition effect), as it captures the heterogeneity within labor input (see Jorgenson, 2001).[10] In TED, we present the contribution of labor, split into the contribution of pure employment quantity (H) and labor quality (LQ), and the contribution of capital services split into ICT capital services (ICT) and non-ICT capital services (non-ICT). Then, equation (5) can be re-written as

$$\Delta \ln GDP = \bar{s}_{K,it} \Delta \ln K_{it} + \bar{s}_{K,nit} \Delta \ln K_{nit} + \bar{s}_L \Delta \ln H + \bar{s}_L \Delta \ln LQ + \Delta \ln A \qquad (9)$$

Subtracting the growth rate of labor quantity, H, we can decompose the growth rate of labor productivity into capital deepening and TFPG as:

$$\Delta \ln v = \Delta \ln GDP - \Delta \ln H = \bar{s}_{K,it} \Delta \ln k_{it} + \bar{s}_{K,nit} \Delta \ln k_{nit} + \bar{s}_L \Delta \ln LQ + \Delta \ln A \qquad (10)$$

where y is the total economy labor productivity growth, measured as the difference between GDP growth and labor quantity growth and k is capital deepening, measured as the difference between capital service growth and labor quantity growth. When available, total hours worked is used as a measure of labor quantity L, otherwise it refers to persons employed.

Equation (9) illustrates that growth in output is the result of growth in labor and capital services (weighted by their respective shares in total output) and a residual that is labeled TFP growth. The

[9] Note that we try to exclude residential structures from our capital series whenever possible, which is true for most OECD economies. However, for most emerging economies this asset group is inclusive of residential structures as well.
[10] Similar approach can be followed in the case of capital also, by aggregating investment across asset types, which however, undermines the differences in marginal productivities across asset types. Therefore, the difference between such simply aggregated capital and the capital services obtained using (3) is often considered as capital compensation effect (see Erumban, 2008).

interplay between labor productivity, and TFP becomes clear from equation (10), as does the fact that the latter is a more refined measure of productivity. The equation shows that growth in labor productivity really depends on three proximate factors, which are capital deepening, labor quality and TFP. Growth in capital deepening arises from increases in the amount of capital available per worker. Improvements in labor quality can be accomplished by increasing the skill level of workers, as well as through education, training or experience. TFP growth captures what is left unexplained, which is sometimes referred to as technical change or the overall efficiency.

The accurate measurement of the variables in the productivity and growth accounting equations is the key to comparing and evaluating the sources of GDP growth. In section 1 of this document (see subsections 1.1, 1.2 and 1.3), we have discussed in detail the construction of GDP and labor quantity (both in terms of hours worked and number of persons employed). In what follows, we discuss which sources and methods are used to construct the remaining variables in equation (9), which are capital services (ICT and non-ICT), labor quality, and labor compensation as a share of GDP.

2.2 Capital Services

2.2.1 Basic sources used

The growth in capital services is estimated by creating series on capital stock by different asset types. As mentioned before, we distinguish between information and communication technology assets (ICT) and non-ICT capital assets. Non-ICT capital services are based on three assets, namely non-residential construction, machinery and transport equipment. ICT capital services are based on investment in software, hardware and telecommunications equipment. Table 6 displays a list of the general sources used to construct the investment time series.

Table 6 – Sources used to construct investment time series

Investment type	*Source*
Total gross fixed capital formation	UN National Accounts Main Aggregates Database
Gross fixed capital formation by asset type	Eurostat - Database Organization for Economic Co-Operation and Development – StatExtracts Erumban, Abdul. A, 2008
ICT investment data	EUKLEMS Jorgens/Vu Kuznets dataset WITSA – Digital Planet Report 2010

2.2.2 Consolidating data from different sources

As investment data and price indices are from various sources, two adjustments were made to consolidate the data. Firstly, all price indices, constant investment and constant GDP data (GDP EKS as described in section 1.1) are benchmarked to 2000 prices. Secondly, as constant investment data from various data sources have different measurement units (for example, OECD data are in PPP adjusted USD, EU KLEMS data are in national currency), we harmonize them with GDPEKS (in 2000 international dollars) by multiplying the ratio of the GDPEKS over constant GDP from the same source as the investment data, i.e.:

$$I_i = I_t^j \times GDP^{EKS}/GDP^j \qquad (11)$$

Where I_i is real investment in asset i used in our capital services estimation, I_t^j denotes real investment data for asset i as obtained from source j (e.g., PWT) and GDP^j is GDP data as obtained from source j.

Machinery investment, as reported in the source data, include hardware and telecommunication data that we must exclude to obtain non-ICT machinery in our classification. To do that, we deduct the latter two asset types from machinery, all in current prices, to obtain an adjusted machinery investment series. Then, we replace the machinery price indices by other machinery (machinery excluding hardware and telecom) price indices reported in EU KLEMS.[11] For countries not covered by EU KLEMS, the original machinery price indices are maintained. The adjusted machinery investment series in current prices terms are deflated using the adjusted machinery price to obtain the adjusted constant machinery investment, i.e., machinery excluding hardware and telecom.

The difference in ICT investment price relative to non-ICT investment price will affect the contribution of these two assets. For example, in the early 1990s, Brazil experienced hyper-inflation which made its non-ICT price surge relative to its ICT price, leading to a sudden drop in the contribution of non-ICT capital growth to total output growth. ICT assets are subject to substantial quality changes, which are often not fully captured by reported investment price deflators. Therefore, it is advisable to use quality-adjusted prices to deflate ICT goods, so that their true effect will be captured properly. However, for most countries, such quality adjusted prices are not available. For EU KLEMS countries, we use the reported ICT price deflators from EU KLEMS. For other countries we adopt the price harmonization approach pioneered by Schreyer.[12] We harmonize the ICT price deflator for a given country using the U.S. ICT hedonic price deflator, adjusted for domestic inflation, and assuming the difference in growth rates between ICT and non-ICT capital in country i to be equivalent to that in the U.S.:

$$\Delta lnp_{ICT}^i = \Delta lnp_{non-ICT}^i + (\Delta lnp_{ICT}^{US} - \Delta lnp_{non-ICT}^{US}) \qquad (12)$$

where *p* denotes the investment price. ICT and non-ICT price indices are each calculated as weighted averages over three asset types (ICT denotes hardware, telecom and software; non-ICT denotes construction, transport and machinery) using their share of current asset investment in total ICT or non-ICT asset investment.

[11] Among our data sources, only EU KLEMS reports price indices for machinery excluding hardware and telecom. We use such adjusted machinery indices whenever possible and extend it to earlier and later years using the growth rates of the original (unadjusted) machinery price indices.

[12] See Schreyer (2002) for a discussion of the approach

2.2.3 Constructing capital input series and calculating its contribution to growth

For each type of asset, capital stock series, $\mathrm{K}_{\mathrm{i,t}}$, is constructed using the Perpetual Inventory Method (PIM) with an assumed geometric depreciation rate

$$K_{i,t} = (1-\delta_i)K_{i,t-1} + I_{i,t} \qquad (13)$$

in which $I_{i,t}$ is the constant investment data. The asset type specific geometric depreciation rates δ_i,, which are assumed to be constant across countries, are provided in Table 7 . PIM requires an initial capital stock, to the depreciated value of which the stream of investment is added. The initial capital stock for each asset, $K_{i,0}$is calculated using a steady state capital output ratio, as suggested by Harberger (1978)

$$K_{i,0} = \frac{I_{i,0}}{\delta_i + g} \qquad (14)$$

where $I_{i,0}$is the level of constant investment in the initial period and g is the growth rate of constant GDP (calculated as a 20 year average, starting from the 10 years before the capital data start).

As mentioned in equation (6), growth in capital services is measured as the weighted sum of capital stock growth in different asset types

$$\Delta\ln\mathrm{K}_t = \sum_{\mathrm{k}} \bar{\mathrm{s}}_{\mathrm{k,t}}\Delta\ln\mathrm{K}_{\mathrm{k,t}} \qquad (15)$$

where weights in the above equation are two-period average shares of each asset type in the value of total capital compensation. As is evident from (7), to obtain the compensation share of each asset type, we need data on rental price of each asset, which is obtained as:

$$p_{k,t}^K = p_{k,t}^I r_t + \delta_k p_{k,t}^I - (p_{k,t}^I - p_{k,t-1}^I) \qquad (16)$$

Where p_k^K is the rental price of capital type k, r_t is the internal rate of return, δ_k is the depreciation rate for asset k, and p_k^I is the investment price of asset k. The last term, which is the difference between current and previous year investment prices, is a measure of capital gain. Thus the rental fee is determined by the nominal rate of return, the rate of depreciation and the asset-specific capital gains. It can also be rewritten as $p_{k,t}^K = p_{k,t}^I r_{k,t} + \delta_k p_{k,t}^I$ with $r_{k,t}$ being the real rate of return, defined as the nominal rate of return r_t adjusted for asset-specific capital gains ($r_{k,t} = k_t - (\frac{p_{k,t}^I}{p_{k,t-1}^I} - 1)$). In equation 16 above, investment price $p_{k,t}^I$ can be derived from the ratio of current and constant investment price indices. (Note: The rate of depreciation is the same as the rate used in the construction of capital stock in equation 11.)

There are two different methods of estimating the nominal rate of return r_t in equation 14 above (see Erumban, 2008). The first is the *ex-post* or internal rate of return. It is calculated as:

$$r_t = \frac{M_t - \sum_k [\delta_k p^I_{k,t} K_{k,t} - (p^I_{k,t} - p^I_{k,t-1}) K_{k,t}]}{\sum_k p^I_{k,t} K_{k,t}} \qquad (17)$$

where M_t is the total nominal capital compensation (which can be derived as value added minus labor compensation using labor share as discussed in the next section). This approach ensures complete consistency between income and production accounts and is thus is the preferred method of calculating the nominal rate of return. It is assumed that the total value of capital services equals its compensation for all assets. For each country, this nominal rate of return is invariant across different asset types, but varies across time. When the internal rate of return cannot be calculated using equation (18) due to data deficiency, an *ex-ante* rate of return approach is used. This method is based on the opportunity cost concept using an exogenous value for the rate of return. Specifically, we use the maximum value from among a country's Central Bank Discount Rate, Government Bond Yield, and Lending Rate, all of which are from IMF International Financial Statistics database, as the external rate of return.

It is often the case that measured internal rate of return and rental prices are negative. However, theoretically, the rental price of capital services cannot be negative. In case of negative values, we use the average of the two adjacent years' positive internal rates of return/rental price of capital services to replace those years. If the negative value appears in the start or end year, then we use the average of all non-negative value to replace it.

Table 7 – Asset-specific depreciation rates used

Asset	*Depreciation rate*
Construction	3%
Transportation	20%
Machinery	13%
IT Hardware	30%
Telecom Equipment	12%
Software	46%

2.3 Labor Quality

2.3.1 Basic sources used

Data on hours worked or total employment as described in section 1.2 and 1.3 represents a series of labor quantity. In order to measure labor's contribution to output growth, an adjustment for changes in the quality or composition of labor is needed. This adjustment is made by calculating the difference between labor input growth rates and labor quantity growth rates, i.e.

$$\Delta \ln LQ = \Delta \ln \sum_l L_l - \sum_l \bar{s}_l \Delta \ln L_l \qquad (18)$$

Alternatively, labor composition growth rates can be constructed based on shares of employment or hours of different worker categories and their respective wage shares, using a Törnqvist index:

$$\Delta lnLQ_t = \sum_i \bar{s}_{l,t} \, \Delta lnh_{l,t} \qquad (19)$$

in which $\bar{s}_{l,t}$ is the two-year average of share of total labor compensation by labor type *l* and $\Delta lnh_{l,t}$ is the growth rate of share of hours worked by labor type *l*.

Table 8 displays the various sources used to divide the total workforce into different skill types, as well as the coverage of those sources.

Table 8 – Sources of data on labor by skill type

Database	*Period*	*Coverage*
Cohen & Soto (2007, 2014)	1960-2020	Average years of schooling
Barro & Lee (2012)	1950-2010	Population 15 years of age and older having received either no schooling, primary, secondary or tertiary education
EUKLEMS	1970-2005	Percentage of total hours worked by low, medium or high skilled workers
Wittgenstein Centre for Demography and Global Human Capital (2015)	1970-2100	Population 15 years of age and older having received either no schooling, primary, secondary or tertiary education

2.3.2 Methodology

Though there are discrepancies between the three datasets in terms of coverage (i.e., population vs. hours worked) and definitions used, we lack the information to harmonize these datasets. Instead we have tried to find a statistical relationship among them. The EU KLEMS dataset provides detailed information on skill shares of employment and compensation for 16 European economies, United States, South Korea and Japan. We consider the EUKLEMS dataset as the most accurate measure among the ones listed in Table 8, and is used as our benchmark for econometric estimation.

We use a seemingly unrelated regression (SUR) model to estimate the relationships between these different databases:[13]

Cohen & Soto and Barro & Lee
Cohen & Soto and Wittgenstein Centre for Demography and Global Human Capital (WC)
EUKLEMS and Cohen & Soto[14]

We obtain a consolidated combined data set by filling in the missing countries in Cohen & Soto by the projected values of the Barro & Lee data. Then we extrapolate the combined data set into the future using the growth rates of the WC, for countries missing Cohen & Soto and Barro & Lee data we use the

[13] Data set mentioned first is used as dependent variable
[14] Ignoring the separation in gender and age groups for EUKLEMS

projected values of the WC. Finally we transform the combined data set of Cohen & Soto, Barro & Lee and WC using the estimated relationship between Cohen & Soto and EUKLEMS. As we consider EUKLEMS to be the most reliable measure of skill levels, we keep the original EUKLEMS data if available in the transformed combined dataset. This way, we obtain a dataset with EUKLEMS data together with the projected data of the combined dataset based on EUKLEMS between 1960 and 2015, divided into three skill levels. Thirdly, to smooth out the data for those EUKLEMS countries between the years covered by EUKLEMS and those not covered by EUKLEMS (years that are covered by the projected values of the combined data set), we use the growth rate of the projected non-EUKLEMS years to extend the EUKLEMS year levels. Finally, an exponential interpolation using constant growth rates is applied to obtain the skill levels in between the available 5 or 10 year intervals.

The compensation data by skill type is constructed using the data set described above and compensation data from EUKLEMS. The shares in total labor compensation by skill type differ widely between countries and years within EUKLEMS. However, since wage ratios are fairly stable over time, we can reconstruct wage ratios using

$$s_{j,l,t} = h_{j,l,t}\frac{w_{j,l,t}}{\overline{w}_{j,t}} \qquad (20)$$

in which $w_{j,l,t}$ is the wage earned by labor type *l* at time *t* in country *j,* and $\overline{w}_{j,t}$ is the average wage at time *t* in country *j*. For each country in EUKLEMS we obtain $s_{j,l,t}$, $h_{j,l,t}$ and $\overline{w}_{j,t}$ and use that to calculate $w_{j,l,t}$ for each country. For each year we then calculate:

$$\frac{\overline{w}_{l,t}}{\overline{\overline{w}}_t} = \frac{\sum_j w_{j,l,t}}{\sum_j \overline{w}_{j,t}} \qquad (21)$$

For missing years we estimate a trend to extrapolate the wage ratio backward and forward in time.[15] Multiplying the wage ratios by the share in hours worked by skill type gives an estimate of the share in labor compensation for all countries. Again, for EUKLEMS countries the actual share of labor compensation is used and extrapolated backward and forward using the estimated values of labor compensation for missing years. For non-EUKLEMS countries the estimated values of labor compensation are used. The change in labor composition, as described in equation (19), is then calculated by using the data on the share in labor compensation and hours worked.

2.4 Labor income's share of total GDP

2.4.1 Basic sources used

In order to calculate the contributions of labor and capital to GDP growth, as described in equation (5), weights need to be assigned to the different factor inputs. These weights are determined by the share of

[15] For years outside the EUKLEMS range we estimate the trend using the Arellano and Bond instrumental variable approach.

total income accrued by either labor or capital. Since the share of capital and labor in total income equal 1, as explained in section 2.1, we construct data on the labor income share directly, while deriving the capital income share indirectly. Table 9 displays a list of the general sources used to construct time series data on the labor income share.

Table 9 – Sources used to construct the labor share time series

Type of data	*# of countries*	*Source*
National accounts based	40	Eurostat - Database Organization for Economic Co-Operation and Development – StatExtracts
Third-party constructed estimates[16]	6 45	WIOD PWT 8.0
Country specific	4	Various
Assuming 0.5	25	-

2.4.2 Methodology

Whenever available, we construct time series on the labor share using national accounts based data. The national accounts only identify the income accrued by employees separately, and hence excludes income earned by self-employed workers. This means that, in order to include all income accrued by labor, the income of the self employed must be estimated. Under the assumption that labor compensation for the self-employed can be imputed from the average compensation for employees, we use the following equation,

$$s_L = \frac{w_e . L_e}{Y} . \frac{L}{L_e} \qquad (22)$$

where s_L is the nominal income share of labor in GDP, w_e is the wage rate earned by employees, L_e is the total hours worked by employees, L is the total hours worked by all employed including self-employed, and Y is total nominal GDP at market prices. As is evident from equation (22) a common wage rate is assumed for both self-employed and employees so that we can re-write the equation as,

$$s_L = w_e . L/Y \qquad (23)$$

When there is no data on hours worked broken down by all persons employed (L) and employees (L_e), L and L_e can be replaced by E and E_e respectively, where E denotes all employed persons and E_e all employees.

For the remaining countries we either rely on data from World Input Output Database (WIOD), which uses the method as explained in equation (22), Penn World Tables (PWT), which uses a number of different estimation methods, or we assume a plain 50 percent labor share.

---

16 These databases use GDP at basic prices, which is GDP at market price – the measure used in the TED – less of taxes and subsidies on products.

3 – Aggregation of growth rates

This section outlines the method of aggregation used to obtain regional growth rates, as available from the file 'Regional Aggregates'. The following equations are used for GDP, labor input and labor productivity growth respectively,

$$\Delta \ln Y_{region} = \sum_{i} \overline{w_i}\, \Delta \ln GDP_i \tag{24}$$

$$\Delta \ln L_{region} = \Delta \ln \sum_{i} L_i \tag{25}$$

$$\Delta \ln y_{region} = \sum_{i} \overline{w_i}\, \Delta \ln y_i + R \tag{26}$$

with w_i denoting the country share in PPP adjusted nominal GDP of the region for each year and a bar denoting the two-period average. Thus, aggregate GDP growth is the weighted sum of individual countries' GDP growth, with relative GDP size of each country being the weight. Growth in labor quantity (employment or hours) is simply the log difference of summed total labor quantity of all the countries in one region. The aggregate labor productivity growth is the sum of country contributions to regional productivity growth (within country productivity gain, obtained as a weighted sum of country productivity growth) and a reallocation term R. The reallocation term is the difference between regional employment growth as measured in (25) and the weighted growth rate of country employment growth rates, i.e. $R = (\sum_i \overline{w_i}\, \Delta \ln L_i - \Delta \ln \sum_i L_i)$ The reallocation term is positive if employment shifts from low productivity countries towards high productivity countries.

Aggregate Total Factor Productivity (TFP) growth rates for various country groups are calculated using the following steps. First, aggregate labor input growth rates and aggregate capital services growth rates are calculated by taking the weighted average of individual country growth rates. The weights used are two-period averages of the country shares in PPP-adjusted nominal GDP of the group for each year. Second, the aggregate labor compensation share for each year is obtained by summing the labor compensation (PPP adjusted) of individual countries and then dividing this sum by total nominal GDP (PPP adjusted) of the group. Finally, TFP growth rates are calculated using,

$$\Delta \ln TFP_{region} = \Delta \ln GDP_{region} - \bar{v}_{L\_region}\, \Delta \ln L_{region} - \bar{v}_{K\_region}\, \Delta \ln K_{region} \tag{27}$$

where $\bar{v}_{L\_region}$ is the two-period average of the regional labor compensation shares, and $\bar{v}_{K\_region}$ is the regional capital compensation share, obtained as 1- regional labor share. $\Delta \ln Y_{region}$ is defined in equation (24).

References

Barro, R. J., and Lee, J. W. (2012). “A new data set of educational attainment in the world, 1950–2010”, Journal of Development Economics

Cohen, D., and M. Soto. (2007). International Educational Attainment Database, updated by Cohen-Leker in 2014

Crafts, N. (1997). Economic Growth in East Asia and Western Europe Since 1950: Implications for Living Standards, National Institute Economic Review 4/97.

Erumban, Abdul. A. (2008). “Measurement and analysis of capital, productivity and economic growth,” University of Groningen, available from: http://dissertations.ub.rug.nl/faculties/feb/2008/a.e.erumban/

Erumban, A.A., Das, D.K. (2015). Information and Communication Technology and Economic Growth in India, Telecommunication Policy, forthcoming.

Feenstra, R., R. Inklaar, and M.P. Timmer (2013). "The Next Generation of the Penn World Table" available for download at www.ggdc.net/pwt

Harberger, A. C., 1978. Perspectives on Capital and Technology in Less Developed Countries. In: M. J. Artis and A. R. Nobay (Eds.): Contemporary Economic Analysis, London, 42–72.

Hoffman, A.A. (1998). Latin American economic development, a causal analysis in historical perspective, GGDC monograph series (no.3).

International Labor Organization (2013). ILO Estimates and Projections of the Economically Active Population, 1990-2030 (7th edition), available from: http://www.ilo.org/ilostat/content/conn/ILOSTATContentServer/path/Contribution%20Folders/statistics/web_pages/static_pages/EAPEP/EAPEP%20Methodological%20paper%202013.pdf

International Labor Organization (2010). ILO Employment Trends Unit, Trends Econometric Models – A Review of the Methodology, available from: http://www.ilo.org/wcmsp5/groups/public/---ed_emp/---emp_elm/---trends/documents/publication/wcms_120382.pdf

International Labor Organization (2011). Economically Active Population, Estimates and Projections, 1980-2020 (6th edition), available from: http://laborsta.ilo.org/applv8/data/EAPEP/eapep_E.html

International Labor Organization (1996). Economically Active Population Estimates and projections, 1950-2010 (4th edition).

Jorgenson, D.W., (1963). Capital theory and investment behavior, American Economic Review, 53 (2), 247-259.

Jorgenson, D.W., Ho, M.S., Samuels, J.D., Stiroh, K.J. (2007). Industry origins of the American productivity resurgence, Economic Systems Research, 19(3), 229-252.

Maddison A., (1982.) Phases of Capitalist Development.

Maddison, A. (1995). Monitoring the World Economy, 1820-1992.

Maddison, A. (2001). The World Economy, a Millennial Perspective.

Maddison, A. (2010). Historical Statistics of the World Economy, 1-2008 AD, available from www.ggdc.net/maddison/Historical_Statistics/horizontal-file_02-2010.xls.

Schreyer, P. (2002). "Computer Price Indices and International Growth and Productivity Comparisons," *Review of Income and Wealth*, 48 (1), pp. 15-31, March.

Timmer, MP., de Vries, G.J. and K. de Vries (2014). "Patterns of Structural Change in Developing Countries." GGDC research memorandum 149 .

United Nations (2009). System of National Accounts 2008 – 2008 SNA, available from http://unstats.un.org/unsd/nationalaccount/sna2008.asp

United Nations (1992), Handbook of the International Comparison Programme. Washington, 1992.

Wittgenstein Centre for Demography and Global Human Capital, (2015). *Wittgenstein Centre Data Explorer Version 1.2.* Available at: http://www.wittgensteincentre.org/dataexplorer

World Bank Group (2015), Purchasing Power Parities and the Real Size of World Economies. A Comprehensive Report of the 2011 International Comparison Program. Washington, 2015.