Measuring the labour income share of developing countries

Learning from social accounting matrices

Katharina Trapp*

April 2015
Abstract: This paper is the first to address the challenges of measuring the labour income share of developing countries. The poor availability and reliability of national accounts data, and the fact that self-employed people, whose labour income is hard to capture, account for a major share of the workforce and often work in the informal sector, render its computation difficult. I consult social accounting matrices as an additional source of information to construct a labour share dataset backed up with microeconomic evidence. First descriptive results show a significant downward trend in labour shares of developing countries since the early 1990s.

Keywords: income distribution, labour share

JEL classification: E25, O15

Acknowledgements: My thanks go to Konstantin Wacker, who encouraged me to continue working on this topic, and to Jann Lay and Stephan Klasen for their support. I also would like to acknowledge the useful comments from the participants in the poster session at the UNU-WIDER conference ‘Inequality – Measurement, Trends, Impacts, and Policies’, 5-6 September 2014, Helsinki.
1 Introduction

In the recent past, inequality has increased within most countries, not only in high-income regions but also across the developing world (Atkinson and Bourguignon 2015; Goldberg and Pavcnik 2007; Niño-Zarazúa et al. 2014; UNDP 2013 and others). Despite this worrying and serious development, we know far too little about inequality, its magnitude, causes, and consequences. For a long time, the distribution of income between individuals and even more so between the factors of production has been largely unattended by the economic discipline. But studies like the recent Oxfam publication predicting that ‘by 2016, the top 1% will have more than 50% of total global wealth’ (Hardoon 2015: 22) enhances the interest of economists more and more. Not least, the distribution question of capital and labour has been put back on the international agenda because of the widely known book by Thomas Piketty (2014).

The labour income share reflects how much of national value added accrues to labour (as opposed to capital). It is a highly informative macroeconomic variable to explore when analysing the roots of inequality. As has been emphasized by Atkinson (2009), shifting the focus from the personal to the factor level can deliver a more integrated understanding of the determinants of inequality, such as imbalances between different sources of income. By decomposing inequality into factor shares and their concentration, the underlying causes of income disparities can be more easily grasped. For example, profits, rents, and other income from capital are much more concentrated than labour and inclined to go to rich households, as revealed by Piketty (2014) and others. A declining labour share is thus likely to indicate an increase in inequality among individuals (ILO 2014c). This illustrates that personal income distribution can be directly determined from factor income shares once the owners of capital, land, and labour are known (Ray 1998).

Dynamics in the factor income distribution are of particular relevance for developing countries, especially in their effort to fight poverty. Regressive redistribution of factors and their remuneration will be felt strongly in these countries due to weak social safety nets and limited access to capital of the poor. The main asset of the poor certainly is labour, usually in form of agricultural self-employment (Fields 2014). As such, the labour share can serve as an indicator in designing policies for social protection and tax systems as these usually target the factor income distribution (minimum wage policies, tax concessions for investments, etc.).

Measurement of the labour share, however, is notoriously difficult for low- and lower middle-income countries. Most studies rely on the relation of compensation of employees to gross domestic product (GDP) from national accounts statistics when measuring the labour share. A key problem of this simple definition is the fact that compensation of employees does not include the labour income of the self-employed, which accounts for the major fraction of the labour force in developing countries (Gollin 2002). An additional difficulty arises from the fact that most of the self-employed in developing countries are located in the informal sector. Finally, and most importantly, there is reason for concern about the scope, detail, and quality of the national accounts of developing countries (UN 2012). The adjustment of the labour share hence requires more prudent handling in the case of less developed countries. Furthermore, the fact that the economic structure of developing countries fundamentally differs from those of high-income economies makes separate assumptions and estimates indispensable. For example, self-employed workers in Organisation for Economic Co-operation and Development (OECD) countries are more likely to have consciously decided to enter self-employment while it may be a matter of necessity for workers in the developing world.

To date, there is no adjusted panel dataset available that addresses developing countries and their particularities when measuring labour share. As a consequence, research findings on the
development of the labour share in lower-income countries remain incomplete, which is a clear gap in the economic literature. For instance, Piketty’s (2014) highly illuminating results remain limited to OECD countries and a few major emerging economies since the lower-income countries do not dispose of a (well-documented) tax base that could be used as data source.

By exploring the challenges associated with the measurement of the labour share in developing countries, identifying social accounting matrices (SAMs) as favourable approach, and drawing on them to construct a dataset of labour shares of poor countries, this paper intends to take a first step towards closing this gap. SAMs are micro-funded representations of an economy that provide detailed data on all the economic transactions that take place within a country. Insights are used to counter-check the reliability of macroeconomic data and to formulate assumptions required for measuring the labour share. By this means, the paper provides the first macro-level labour share dataset for developing countries that is based on and backed up by microeconomic evidence.

The final dataset covers about 100 developing countries from 1990 to 2011. Descriptive statistics show that labour’s relative income in developing countries has been declining since the early 1990s by about 10 percentage points, indicating that global labour income has been lagging behind overall productivity increases. This effect holds true for most of the developing regions and seems to be independent of a country’s development stage. The hope that globalization will not only lead to economic growth in developing countries but also reduce worldwide and within-country inequality through the promotion of the factor of labour hence seems not to have materialized.

The paper is organized as follows: Section 2 presents the concept of the labour income share and elaborates on its measurement using national accounts. Existing datasets are reviewed in Section 3. Section 4 discusses the challenges associated with measuring the labour share of developing countries, whereupon Section 5 introduces SAMs as an approach that serves both as an information source and a check on robustness. Insights are applied to construct and validate the labour share in section 6. Section 7 presents some properties of the dataset and Section 8 concludes.

2 The labour income share: Concepts and measurements

The labour income share reflects how much of national income is earned by labour and hence measures the factor income distribution of a country. Assuming that value added, or production output, is given by \( Y = f(K, L) \), where \( K \) is capital (including land capital) and \( L \) labour used in production, the income distribution between production factors is given by:

\[
Y = \frac{w}{P} \times L + \frac{i}{P} \times K,
\]

where \( w \) is wage, \( i \) the interest rate, \( r \) rent and \( P \) the price level. The labour share (LS) then can be expressed as:

\[
LS = \frac{w \times L}{P \times Y}
\]

The standard approach to describe the relationship between factors of production and their production output continues to be the Cobb-Douglas production function (Cobb and Douglas 1928). Due to its assumption of constant output elasticities and factor remuneration according to productivity, the Cobb-Douglas production technology predicts factor shares to be consistent over time. Changes in the labour share may only originate from shifts to sectors that use one of the
production factors with relative intensity, non-neutral technological progress, or a deviation of a factor's remuneration from its marginal product caused by labour adjustment costs or bargaining between capital and labour (Bentolila and Saint-Paul 2003). Early empirical estimations of the factor income distribution using the Cobb-Douglas production function found the labour share to range between 0.6 and 0.75 in the United States and other high-income countries, giving rise to the standard value of two-thirds for labour's share in production output (Douglas 1967, 1976). The same studies supported the long-term stability of factor shares, which also became one of the 'stylized facts' of economic growth put forward prominently by Kaldor (1957).

The labour share can be computed from national accounts statistics. The empirical literature usually starts from the relation of compensation of employees, taken from the primary distribution of income accounts, to total value added produced in the respective country (GDP):

\[ \text{LS} = \frac{\text{Compensation of employees}}{\text{GDP}} \]

Data is provided by the UN System of National Accounts (SNA) and is accessible through the National Accounts Official Country Data. This simple measurement, however, tells only half the story and is often referred to as the 'naïve labour share'. As pointed out by Krueger (1998) and Gollin (2002), compensation of employees merely covers wage earners in the corporate sector and ignores self-employment. The challenge with self-employed income is that it is composed of income from labour as well as from capital so that its labour component needs to be filtered out in a first step. Labour income, however, is not easy to identify, especially at the national level. If it is not corrected for self-employment, income of the self-employed would be mistakenly treated as only consisting of non-labour income and would be added to the denominator of the equation but not to the numerator, resulting in a downward bias of the labour share. Furthermore, in a dynamic perspective, ceteris paribus shifts in the composition of employment would automatically change the labour share. For example, when formerly self-employed enter wage employment in developing countries, this typically is the movement away from subsistence agriculture – their labour income suddenly appears in employee compensation statistics, raising the labour share, even though labour income has effectively not changed (or only very little). It is therefore essential to adjust this measure so that it not only reflects the share of national income that is earned by employees but the entire share that accrues to labour input, regardless of how it was earned. Gollin (2002) presented three possible approaches in this regard. His article soon became the standard reference.

Gollin's first two adjustments make use of the item 'mixed income' listed in the UN SNA. Mixed income refers to the remuneration of the self-employed and – as the term already suggests – includes income from labour and capital (EC et al. 2008). By using this item and filtering out its labour income component, which is then added to employee compensation, a meaningful measure of the labour share can be obtained.

In his first adjustment, mixed income of the self-employed is completely added to compensation of employees, assuming income of the self-employed to be only composed of labour income:

1 According to the International Labour Organization (ILO), the self-employed comprise employers, own-account workers, members of producers' co-operatives, and contributing family workers (ILO 2014a).
As this procedure ignores income from factors of production other than labour, it is likely to overestimate the labour share.

The second adjustment assumes mixed income to consist of the same mix of labour and non-labour income as the rest of the economy:

\[
LS_{G2} = \frac{\text{Compensation of employees}}{\text{GDP} - \text{Mixed income}}
\]

This approach is more straightforward but disregards that capital and labour shares might vary substantially across sectors and with the size and structure of businesses.

Gollin’s third adjustment requires only data on the employment structure of a country, which is for example available in the ILO’s Key Indicators of the Labour Market database. Relying on the assumption that the self-employed earn the same labour income as employees, it imputes the average wage bill of employees to the self-employed. Only income of the self-employed that exceeds the mean wage sum is counted as income from capital:

\[
LS_{G3} = \frac{\text{Compensation of employees}}{\text{Employees}} \times \frac{\text{Total employment}}{\text{GDP}}
\]

This adjustment does not take into consideration that self-employed and wage earners might work in different sectors, realizing different labour productivities. If self-employment mainly occurs in subsistence farming and other low-productivity activities, as is typically the case in developing countries, this equation systematically overestimates the labour share. On the other hand, it underestimates the labour share in countries where the self-employed belong to the high-income earners of the economy. To account for such systematic differences, some studies (Arpaia et al. 2009; OECD 2012) impute the income of the self-employed on the basis of the average wage of employees in each sector, using EU KLEMS data. While the naïve labour share reflects the share of the total wage bill of employees in GDP, the labour share computed using this third adjustment relates the wage per employee to the value added per employed person. It follows that the adjusted labour share rises (falls) if the wage per employee grows faster (slower) than the productivity per employed person. It is hence closely linked to the unit labour costs which relate average wages to overall performance, that is, to productivity (Marterbauer and Walterskirchen 2003).

When applying these adjustments, Gollin (2002) finds labour shares to be more or less constant across time and space and therefore suggests adhering to models using a Cobb-Douglas production technology. The size of the labour share is shown to range between 0.6 and 0.85. His results are, however, based on a small sample of 31 high- and low-income countries observed at only one point in time and are heavily challenged by other recent studies.

3 Data review

Aside from Gollin (2002), various international organizations as well as researchers have taken up the analysis of trends in labour shares. The majority of the empirical literature (Arpaia et al. 2009; Bentolila and Saint-Paul 2003; Blanchard 1997; Checchi and García-Peñalosa 2010; EC 2007; Ellis
and Smith 2007; Guscina 2006; Hutchinson and Persyn 2012; ILO 2011; IMF 2007; Jaumotte and Tytell 2007; OECD 2012; Richardson and Khripounova 1998; Slaughter 2001; and others) is restricted to OECD countries, where data quality as well as coverage is high and data on employment structure and mixed income available. These works rely on the naïve labour share or, additionally, Gollin’s third adjustment, as it is for example provided in the EC’s AMECO (Annual Macro-Economic) database.

Piketty (2014) adopts an entirely different approach and uses tax data to study the distribution of top incomes and capital gains.2 His study is restricted to around 30 countries, mainly from the Western hemisphere, that provide income tax data. When he splits mixed income between capital and labour income, he relies on Gollin’s second adjustment.

There are some broader studies (e.g. Bernanke and Gürkaynak 2001; Diwan 2001; Guerriero and Sen 2012; Harrison 2005; Jayadev 2007; Karabarbounis and Neiman 2013; Rodriguez and Jayadev 2010) that conduct worldwide analyses including a number of developing countries. Like the recent study of Karabarbounis and Neiman (2013), they mainly base their analysis on unadjusted labour shares. In fact, about two-thirds of low- and middle-income countries report the item ‘compensation of employees’ to UN SNA on a regular basis since 1990 but a systematic recording of unincorporated businesses remains missing. Just a handful of countries report the necessary data on the number of self-employed or their income, with the result that applying one of Gollin’s adjustments (primarily the first or third) comes with the consequence that only a few developing countries remain in the sample. For example, when Rodriguez and Jayadev (2010) adjust their broad labour share dataset of 135 countries for mixed income, their sample size sharply reduces to 59 (mainly high-income) countries. Some other studies (Decreuse and Mareek 2013; Ortega and Rodriguez 2001) use the UN Industrial Development Organization (UNIDO) dataset INDSTAT3 as an alternative to the UN SNA data but it is even more restrictive as it is limited to the corporate manufacturing sector.

With adjusted labour share data for 127 low-, middle-, and high-income countries for at least 20 years, the recently published Penn World Table (PWT) version 8.0 is the broadest analysis of labour share trends. By using data on total agricultural value added (taken from the Socio-Economic Accounts of the World Input-Output Database) as a proxy for mixed income, Inklaar and Timmer (2013) are able to increase the number of sample countries by about 60. Time and country coverage is further extended by interpolating and keeping labour shares constant over time. They develop a ‘best estimate’ labour share that chooses the most appropriate from the three Gollin adjustments in light of the country-specific data situation.3 Their approach is primarily mechanical, meaning that it is based on a country’s data availability and plausibility rather than on economic conditions.4

In contrast to Gollin’s results, all OECD studies find the labour share to be decreasing in most of the high-income countries over time, regardless of how labour’s share in national income is measured. The global studies confirm this decline and further provide evidence of a significant

2 Available in the World Top Income Database.

3 That is (1) Gollin’s first adjustment assuming all of agricultural value added to be labour income, (2) his second adjustment assuming UN mixed income to consist partly of labour and capital income, (3) his third adjustment using data on the number of self-employed from ILO’s LABORSTA, or (4) the naïve share.

4 For example, an important guideline is that the lowest labour share from either (1) or (3) is chosen, given the high risk of overestimation.
negative trend for the developing world. The labour share is also found to be less than the ubiquitous ‘two-thirds’, both in rich and poor countries.\textsuperscript{5}

To my knowledge, there is no study explicitly exploring the development of the labour share in low- and middle-income countries. This is a clear gap in empirical research as the special economic structures and endowments of developing countries require a distinct model as well as a separate empirical investigation. Instead, the developing world is exclusively covered in global studies, where the same assumptions are applied for industrialized as well as low-income economies. There remains some risk that this approach results in incorrect labour share data. Self-employment systematically differs between the two economic worlds: while it may mean business management and entrepreneurial spirit in the Western hemisphere, it is likely to refer to subsistence farming and precarious employment in developing nations. Similarly, cross-country studies omit verifying the assumptions on which their measurement relies with national or regional data, which might be another source of trouble. My dataset improves upon this literature by being the first to consider region-specific peculiarities as well as differences resulting from a country’s development status when constructing the labour share. It also considers the specific data situation in these countries. To do so, this study investigates self-employment sectors and their typical characteristics in low- and lower middle-income countries consulting national SAMs.

\section{Measurement challenges}

Data review has shown that constructing a broad dataset on the labour share of developing countries is strongly hampered by the limited availability of national accounts data in these countries. Data on self-employed income is lacking for the majority of lower-income economies and even basic figures, such as the number of self-employed, are scarce. This is not surprising in view of the fact that the most prevailing forms of self-employment in developing countries are micro and small enterprises, and subsistence farmers whose economic activities are difficult to capture. Hence, the naïve labour share cannot easily be adjusted for self-employment, as suggested by Gollin (2002). This is only possible against the backdrop of losing the bulk of observations, most probably the least developed countries, which do not have the institutional capacities to collect the data.

Data constraints therefore require proxy variables to be selected either for self-employed income or for their share in the workforce when compiling a comprehensive dataset. This, in turn, involves making assumptions about the composition of the self-employed, their characteristics or value added. To impute the labour income of the self-employed and decide on using one of Gollin’s adjustments, it is necessary, in addition, to formulate assumptions with regard to the intensities and productivities of the self-employed factor. The PWT database uses total value added in agriculture as proxy for self-employed labour income, assuming that most self-employed income stems from agricultural production, with labour being by far the most important input factor (Inklaar and Timmer 2013). This proxy is plausible, yet it disregards capital, especially land, as an agricultural factor and leaves aside labour income from other forms of self-employment. One should also bear in mind that this variable is not able to reflect industrialization of agriculture, meaning a constant or increasing agricultural value added accompanied by a decreasing number of farmers.

Choosing a proxy in the context of developing countries takes place against the background of – in contrast to advanced economies – an average of about two-thirds and up to 90 per cent of the

\textsuperscript{5} For example, it averages 0.52 and 0.46 in Inklaar and Timmer (2013) and Harrison (2005), respectively.
working population being self-employed, with most of them belonging to the informal sector (ILO 2014b). Hence, self-employment is the rule in lower-income countries and systematically differs from self-employment in developed economies: it primarily describes workers who are forced into self-employment as they do not find employment on the regular job market, whereas self-employment in OECD countries rather refers to (formerly employed) people who wish to set up their own business and become their own boss (Fields 2014; World Bank 2013). It is essential to differentiate between these two main circumstances and determine whether it is a conscious decision or rather the lack of alternatives that drives people into self-employment. In fact, the formal labour market in poor countries is relatively underdeveloped offering only limited opportunities. Furthermore, a high share of the workforce is poorly educated, causing difficulties in matching job profiles. By implication, self-employed people in developing countries are likely to have received little or no education at all which, in turn, also suggests low labour productivities. Another particularity of self-employment in developing countries lies in its vulnerability and informality. Mostly, a formal contract and support from a social security system is missing and remuneration is entirely dependent on self-made profits (ILO 2014b). The risk of being active in the informal sector is especially high in Africa, Asia, and Latin America and the Caribbean (ILO 2014b). In developing nations, self-employment thus usually coincides with informality and hence often remains statistically unobserved. UN SNA standards demand that the so-called shadow economy be recorded, but, due to its very nature, national accounting often fails to do so as it can only be estimated (OECD 2004).

The processing of data is challenged, furthermore, by the unreliable scope, detail and quality of national accounting in developing economies (UN 2012). As demonstrated by Jerven (2012), national income accounting in Sub-Saharan Africa in particular is fraught with inaccuracies and inconsistencies, which impairs cross-country comparability. In this case, for example, it is not immediately obvious what is covered by the items ‘compensation of employees’ and ‘mixed income’ of the self-employed, as they are not always standardized with international guidelines (see for example EC et al. 2008). Dealing with international income data is thus likely to be subject to a quality–coverage trade-off, as is common among broad cross-country data analyses: the more countries (especially developing countries) are covered, the greater the chance of poorer quality data and lack of comparability.

To meet these particular demands and challenges, the following section more thoroughly investigates labour income in general, and self-employed income in particular, before turning to the formulation of assumptions. This is done by consulting SAMs as additional sources of information. This prevents the exclusive reliance on macro-level national accounts data of previous studies.

5 Social accounting matrices

A SAM is a square matrix that represents flows of all transactions that take place in an economy. Rows display the income of an account while columns denote its expenditures. SAMs draw a comprehensive picture of the economy and thus can reveal a country’s economic structure much better than national accounts can. They are constructed by matching national accounts, input-output tables, labour force surveys, household surveys, and so on, and using them to complement each other. They bridge income data collected at the household level and macroeconomic data, and therefore imply a higher reliability than national accounts. Already Pyatt and Round (1985) stressed the role SAMs could play in improving the quality of national accounts.

As SAMs disaggregate by sector and type of employment, they make the sectoral composition of labour visible. By this means, SAMs can be used to identify the sectors of the self-employed as
well as the corresponding productivities and factor intensities. This information can be used as the basis for a suitable estimation of the labour income of the self-employed. Unfortunately, SAMs are not available at a large scale, but the International Food Policy Research Institute (IFPRI) and the UN Development and Analysis Division (UN DESA) freely provide data for several developing countries. In addition, a number of country case studies are available from various sources. Usually, the cross-entropy approach is applied to develop SAMs. This method exploits scattered and inconsistent data in a highly flexible and efficient way and thus can deal with the poor data situation of developing countries (Robinson et al. 2001). Since SAMs are usually only constructed at large time intervals, the data mostly covers one observation per country.

SAMs can help in understanding the substance of self-employment in two ways, namely by providing quantitative as well as qualitative information.

5.1 Basic descriptives on the labour income share

First, national factor income shares are extracted from SAMs, thereby compiling a small pool of data on the labour share. To ensure comparability, I restrict the data to SAMs from IFPRI and UN DESA that both apply the cross-entropy approach. Together, they provide 51 SAMs for 45 developing countries. Unfortunately, the size of the data pool is too small to conduct large-scale data analyses across time and space. Nevertheless, the data provide important basic descriptives about the size and distribution of labour shares in developing countries that can serve as benchmark. Empirical insights into the statics and the dynamics of the labour share can also later be used to perform a robustness check of national account data. Summary statistics of the 51 country-specific SAMs show that the labour share ranges between 0.24 and 0.71, with a mean and a median of 0.46. It is almost normally distributed as shown in Figure 1 which provides its probability density function, obtained from Epachenikov kernel density estimates.

Figure 1: Probability density function of the labour share (from SAMs)

Source: Author’s illustration, based on IFPRI and UN DESA data.
By taking the example of South Africa, the data pool also allows an analysis of the dynamics of the labour income share in the course of a country’s economic development. South Africa is selected as a case study mainly because it is the only country for which at least three SAMs are available, covering a time span of more than ten years, but also because it is a good example of a fast-growing emerging country. Figure 2 shows how South Africa’s real GDP as well as labour’s share in GDP developed from 1993 until 2005. While real GDP increased by almost 50 per cent, the labour income share decreased by 7 percentage points (from 56.3% to 49.8 per cent). This change in the labour share is quite remarkable and is evidence of non-constant factor shares. The downward trend in the labour share of South Africa might just be an exception, but one cannot rule out the possibility that there is a general link between the economic development of a country and its factor income shares.

5.2 Three case studies: Indonesia, Zambia, and Bolivia

Second, I take a closer look at individual SAMs that include an analysis of the self-employed and their factor income. This is to better understand the essence of self-employment in the developing world and to integrate lessons learned into the process of computing the labour share. Hence, the information content of micro- and macro-level data can be exploited in a meaningful manner. Studies on Zambia, Indonesia, and Bolivia, which represent the three major developing regions Africa, Asia, and Latin America, are used for this.

Thiele and Piazolo (2002) construct a SAM for Bolivia for 1997 which decomposes self-employed income into mixed income and employers’ profits, and identifies the sectors in which self-employment prevails. Thurlow et al. (2004), who present a 2001 SAM for Zambia, provide insights into the factor intensities, skill contents, and productivities of each sector. Their study especially sheds light on the agricultural sector, as it differentiates between small- and large-scale farming and itemizes land capital as a production factor. Finally, Yusuf (2006) constructs a SAM for Indonesia for 2003, which calculates the labour and non-labour share of self-employment. Self-employed labour income is imputed on the basis of the average hourly wage in the corresponding employment sector for a similar type of labour (according to skills, activity, and urban–rural
location), multiplied by the number of working hours recorded by labour force surveys. Only self-employed income that exceeds the imputed labour income is counted as non-labour income.

Studying these SAMs reveals that self-employment in developing countries is more prevalent in rural than in urban areas. Self-employed people mostly work in the agricultural sector and are smallholder farmers. Outside agriculture, self-employment usually takes the form of own-account enterprises that primarily engage in wholesale, retail, or hospitality. A smaller share conducts light manufacturing (food manufacturing), especially in rural areas, where processing of agricultural products is common. Employers are mostly active in modern agriculture and the service sector, mainly transport. On the other hand, self-employment hardly appears in the sectors of mining, manufacturing, and textiles.

It is furthermore observed that the sectors in which self-employment prevails are the labour-intensive sectors. For example, in the case of Indonesia, the labour share in the hospitality sector is 0.8, in the agricultural sector 0.65 and in the retail sector 0.58 (Yusuf 2006). At the same time, mining – with 0.2 – has the lowest labour share, and the one in the manufacturing and textiles industry is also only about 0.4. The SAM of Zambia takes a closer look at the agricultural sector and shows that smallholder agriculture is highly labour-intensive: It displays a labour income share of 0.73, followed by a capital share of 0.2 and a land share of 0.07 (Thurlow et al. 2004). Exceptions are urban self-employed and employers who have a relatively high capital share (Thurlow et al. 2004).

It also becomes apparent that, in a given sector, self-employed people pursue a more labour-intensive strategy than larger firms. This is also true for the agricultural sector. Thurlow et al. (2004) distinguish between small-, medium-, and large-scale farms and show that the labour share of agricultural production decreases and the capital share increases with farm size (while the land share does not show a specific pattern). The reason is often seen as the availability of family labour which gives smallholders easy access to manpower (World Bank 2013).

Furthermore, there is some evidence that the self-employed have much lower skills and less education compared to the rest of the workforce, suggesting that they are less productive per unit of labour than employed staff. This seems to be true for self-employed people both inside and outside agriculture. Most labour in smallholder farming is uneducated and the skill content of labour is higher for larger farms. Similarly, urban own-account enterprises are less productive than their larger counterparts. This corresponds to their low educational attainment but is also due to their limited access to (financial) capital (World Bank 2013).

Finally, the SAMs give also some indication of how the labour income of the self-employed compares to employees’ wages. It is found that smallholders and urban self-employed are worse off, suggesting a relatively low labour income (besides a low income from capital). This also corresponds to the finding from Indonesia that the agricultural sector absorbs a large share of the workforce but contributes comparatively little to national GDP. An exception is urban employers, who have quite a high labour as well as capital income; they are small in number though. It can be concluded that self-employed labour income is at most equal but, more likely, lower than that of employees.

Of course, these are just individual cases but findings are consistent with the general literature on self-employment in developing countries (for example, Fields 2014; Fox and Sohnesen 2012; Mead and Liedholm 1998; World Bank 2013).

---

6 This strategy is, among others, also applied by Ivanic (2004) for the Global Trade Analysis Project (GTAP) database.
6 Measuring the labour income share

6.1 Steps of construction

This section describes how the labour income share dataset is constructed. As illustrated above, there are three main alternatives for adjusting the naïve labour share for self-employment. The adjustment is a delicate exercise since it requires making assumptions about self-employed income as well as about its labour income share. To substantiate my assumptions, I rely on the qualitative and quantitative information about self-employment gained from SAMs.

Gollin’s (2002) first adjustment treats all self-employed income as labour income. Although SAMs have shown that typical self-employed activities are associated with high labour share, this approach is likely to overestimate the labour share, given that the capital share in agriculture, retail, or hospitality usually accounts for one-quarter to one-third.

Gollin’s (2002) second adjustment assumes self-employed income to contain the same mix of capital and labour income as the rest of the economy. This, by contrast, rather understates the labour share of developing countries since self-employed people have been shown to systematically differ from employees in their factor shares and rely relatively more on labour.

Unfortunately, data on self-employed income is only reported by less than one-third of developing countries on a regular and comprehensive basis (mainly countries in Latin America, the Caribbean, Eastern Europe and Central Asia). Applying Gollin’s (2002) first or second adjustment thus makes the sample size shrink. Yet, the remaining observations can provide some guidance on the level of adjusted labour shares: they range between 0.26 and 0.87 with a mean of 0.57 (first adjustment) and 0.21 and 0.73 with a mean of 0.46 (second adjustment).

Gollin’s (2002) third alternative imputes the average wage sum of employees to the unincorporated sector relying on the share of self-employed in the workforce. This method might overestimate the labour share. As shown in SAMs, self-employed people, although their work is more labour-intensive, are usually less productive and mostly worse off than employees. By imputing not the total but just a share of employees’ wage bill, it can, however, be a meaningful starting point. As data on the self-employment share is insufficient as well, a proxy that can serve as indirect measure must be chosen beforehand.

I select the share of employment in agriculture as a proxy variable, assuming that most of the self-employed in poor countries are smallholders and most of the farm labour force is self-employed. The case studies of SAMs revealed that there is a clear overlap of self-employment and agricultural employment in developing countries. Of course, this proxy is more appropriate in some regions than in others, depending on the sectoral structure of an economy, and also disregards self-employed activities other than in agriculture. But the correlation of 0.80 between self-employment and the agricultural employment share shows that it serves as a good proxy. Aside from the high congruence, another strength of this proxy variable is the high availability and good quality of data. Data on the agricultural employment share is provided for almost all developing countries by either World Bank WDI or the UN’s Food and Agriculture Organization (FAOSat). Figure 3 illustrates the development of the agriculture employment share by region over time. Similar to what we know about self-employment in the developing world, it shows how the importance of agriculture varies across regions and how it declines in the course of time and with the economic development of a country.
After having chosen a proxy for self-employment, the labour income of the self-employed can be imputed in a next step. The full imputation seems to be suitable for many countries: in Eastern Europe, Central and East Asia, the Middle East, North Africa, Latin America and the Caribbean, this adjustment yields labour shares that range between 0.17 and 0.82 and average at 0.5 or below, which is very close to the labour shares obtained from SAMs. Furthermore, where reported, it ranges between Gollin’s (2002) first and second adjustment. I therefore hold on to this adjustment in these countries. At the same time, however, it yields implausibly high values for other countries (for example, 2.08 in China, 2.4 in Bhutan, or 3.18 in the case of Burkina Faso). There may be three reasons for this: First, countries might already correct for labour income of the self-employed in their reported ‘compensation of employees’ such that any further amendment would mean a double adjustment. Second, the database may simply contain errors or agricultural employment may be inappropriate as proxy, both with the consequence that the adjusted labour share data becomes unreliable as well. Third, and most relevant, the assumption behind this adjustment might not hold for all countries. Indeed, the assumption of same wages is always poor when there are systematic differences between the two sectors of employment.

Bhutan is certainly a case where correction has already been made for the labour income of the self-employed in light of the fact that even the naïve labour share amounts to 0.91. This seems also to be the case for a few other countries. So no further modification is added in countries where the naïve labour share is already reasonably high (greater than 0.21) and an imputation of wages would overshoot (greater than 0.91).7

7 The marking values stem from the most extreme labour shares observed in SAMs, the naïve share and after Gollin’s (2002) first and second adjustment, see Table 1.
In other cases, the underlying data seem to be flawed as there are implausible jumps in ‘employee compensation’ to GDP ratio. A special case is the post-Soviet states, which all show a considerable labour share plunge in the early 1990s. Behind this fall is not only the intense economic transformation but also stagnant statistics: suddenly, a previously non-existent shadow economy sprang up in the former Soviet republics that the national statistical offices were not able to capture (Johnson et al. 1997; Kaufman and Kaliberda 1996). Many formerly official workers began to work as self-employed in the informal economy and no longer appeared in official statistics. To correct for the increasing shadow economy and the related drop in the labour share, I leave the naïve labour share in the years before the plunge so that incorrect upward adjustments are avoided.

The data of most countries with implausibly high labour shares, however, gives reason to conclude that the full imputation of average wages is not appropriate: very high adjusted labour shares (above 0.91 or even above 1) go hand in hand with very low naïve labour shares (below 0.21), suggesting that the actual labour share lies somewhere in between. The most extreme example is Burkina Faso, which has a mean naïve labour share of 0.21 that wage imputation lifts up to 3.18. An adjustment assuming relatively lower average wages for the self-employed seems to be more suitable in these countries. The concerns centered are the poorest of the poor, basically countries in Sub-Saharan Africa and South Asia, where the economy and formal labour markets are least developed and most people engage in low-productivity subsistence farming (FAO 2012; World Bank 2013). For the low- and lower middle-income countries from these two regions, I therefore only impute a share of employees’ wages. Following Bentolila and Saint-Paul (2003), the self-employed are assumed to earn on average two-thirds of employee income. Of course, this assumption is of an arbitrary kind, but the resulting adjusted labour shares appear reliable. They range between 0.11 and 0.77 in Sub-Saharan Africa, and 0.42 and 0.75 in South Asia, and, where available, move between the values reached after the first and second adjustment or below.

After completing these steps, Gollin’s (2002) first adjustment functions as an upper and his second adjustment as a lower bound in countries which report ‘mixed income’ and when the so far adjusted labour share exceeds either of these limits.

The final labour share dataset is thus based on Gollin’s third assumption, using the agricultural employment share as proxy and imputing the full average wage of employees to the self-employed.
in the emerging regions, and a fraction of two-thirds in the less developed regions of South Asia and Sub-Saharan Africa, correcting for obvious measurement errors, leaving the naïve share where self-employed income is already accounted for, and framing the whole with Gollin’s first and second adjustment whenever available.

6.2 Validation of data

Table 1 summarizes the constructed labour share and its components. The final dataset covers 100 low- and middle-income countries from 1990 until 2011. It is an unbalanced panel and has in total 1,396 observations. The labour share ranges from 0.06 to 0.91 with a mean of 0.47, which is much lower than the ‘two-thirds’ standard in economic literature. Due to the systematic differences between self-employed people and employees in developing countries, Gollin’s (2002) third correction runs the risk of overestimating the labour share of developing countries. The summary statistics suggest that this is not the case though. The results are also very close to the findings from SAMs, giving further support to this way of proceeding.

Figure 4: Labour share of South Africa (in %), 1990-2011

Source: Author’s illustration, based on UN SNA, FAOStat, World Bank WDI, IFPRI and UN DESA data.

When looking again at the case study of South Africa (see Figure 4), we can see that Gollin’s (2002) third adjustment, based on the agricultural employment share, is an appropriate way of proceeding. My constructed labour share is similar to that obtained from the country’s SAMs, not only in terms of level but also trend.8 Leaving the naïve share as it is, on the other hand, would understate the labour share.

We can also learn from looking at the temporal development of the constructed labour share in comparison to its components (see Figure 5). As the labour share dataset is mainly based on Gollin’s third adjustment, both show a similar pattern. The composed labour share is also just about the same level as Gollin’s second adjustment, indicating that the imputation of labour income from employee compensation results in an average capital–labour mix of self-employment as in the rest of the economy. The composed labour share is centred between the naïve labour share (no self-employed income) and Gollin’s first adjustment (all self-employed income), which

---

8 Unfortunately, data on mixed income is not available for South Africa.
further suggests that the labour share of self-employment is on average close to one-half. What also immediately becomes visible is the non-constant, negative performance of the labour share over time, which is investigated more closely in the next section.

Figure 5: Different labour share adjustments (in %), 1990-2011

7 Properties of the dataset

7.1 Descriptives on the labour share

After having prepared the labour share data, it can be used to investigate global trends in the development of the labour share over time, as well as differences across developing regions. A first important finding is that labour’s relative income in developing countries has declined over time (see Figure 5), suggesting that labour income has been lagging behind overall productivity increases. My results thus provide strong evidence against the hypothesis of constant labour shares in the long run and are consistent with the findings of previous studies.

The labour share is found to be stable in the early 1990s but starts declining with the end of the Cold War. In the post-1993 era, it falls on average by about 10 percentage points (from 53.7 to 43.6 per cent). The labour share recovers slightly in the late 2000s in the course of the global financial crisis of 2007-8 but continues falling afterwards. This temporarily reversed trend mainly goes back to the countercyclical movement of the labour share, meaning that capital owners usually lose more than wage earners during crises (ILO 2011).

It should be emphasized that the fall in the labour share is independent of the form of measurement. It is noticeable that even the naïve share, which only captures wage employment, is decreasing significantly over time. It is a well-known fact that the labour share has fallen in high-income economies over the last two decades. This is mainly explained by capital-augmenting technological progress and the specialization into capital-intensive commodities in the course of globalization – an argument based on the factor-proportion models of Heckscher, Ohlin, Stolper, and Samuelson (Heckscher 1919; Ohlin 1933; Samuelson 1948; Stolper 1941). To the extent that labour is abundant in developing countries, one would rather expect the labour share in developing
countries to rise with economic growth and international integration. This should be especially the case for the naïve share, considering that development is associated with the expansion of the corporate sector (which is assumed to be labour-intensive). The reasons for the decreasing labour share might be different in various contexts. The literature mainly discusses the effects of increasing automation, worldwide competition, and unemployment, as well as the decreasing bargaining power of labour.

As shown in Figure 6, this downward trend of the labour share has been present in most developing regions. East Asia and the Pacific is the region which experienced the fastest decrease (on average 14 percentage points since 1990), closely followed by Eastern Europe and Central Asia (11 percentage points), and Latin America and the Caribbean (10 percentage points). It was only in the second half of the 2000s that the downturn of labour shares came to a halt. A considerable decline also occurred in Sub-Saharan Africa, where labour shares fell by 6 percentage points between 1990 and 2011. Exceptions to the downward trend are only to be found in South Asia (Sri Lanka, Bhutan, and India), the Middle East, and North Africa, whose labour shares fluctuated but more or less remained on a constant level.

Figure 6: Labour income shares by region, 1990-2011

![Bar chart showing labour income shares by region](image)

Source: Author’s illustration, based on UN SNA, FAOStat and World Bank WDI data.

In terms of levels, South Asia exhibits the highest labour shares (0.7 on average), which is, however, mainly due to Bhutan data. With labour shares ranging between 0.4 and 0.6, East Asia, Eastern Europe and Central Asia, Latin America and the Caribbean are mid-table. Sub-Saharan Africa has slightly lower labour shares moving around 0.4, while labour shares in the Middle East and North Africa remain below 0.4, which is not surprising in view of the fact that most of the oil-producing countries are located in this region.

Figure 7 displays labour shares for different income groups according to the country classification by the World Bank (2011). The level of a country’s labour share seems to be independent of a country’s stage of development and also the negative trend occurs in all income groups. However, it is more pronounced in low-income countries, followed by lower middle-income and finally upper middle-income countries.
7.2 Unit roots

Descriptives on labour share data already provide some evidence against the long-prevailing hypothesis of constant factor shares. Nevertheless, many theoretical models are still based on the Cobb-Douglas production technology or similar constructs that treat the labour share as a persistent variable. For future applications, it is therefore important to be aware of the possible presence of unit roots in the labour share data. If the labour share is used as variable in regression analyses — for example, as dependent variable to explore the reasons behind its negative time trend — and it follows a unit root process together with another variable in the model, the problem of spurious regression may arise (Granger and Newbold 1974). This renders the inferences of conventional methods invalid and therefore would require special estimation techniques, such as a cointegration method or estimation in first differences (Enders 2010; Greene 2003). I therefore test for the presence of unit roots, using a Fisher test statistic as presented in Maddala and Wu (1999):

\[
\lambda = -2 \sum_{i=1}^{N} \ln(\pi_i)
\]

where \(\pi_i\) is the p-value of any unit root test for each cross section \(i\).\(^9\) Within this framework, the augmented Dickey-Fuller (1979) test is performed, which can be considered as the most common method, testing the null hypothesis of unit root against the alternative of stationarity. The Fisher test statistic is preferred over other test statistics, such as Im-Pesaran-Shin, as it is not an asymptotic but an exact test, which does not require a balanced panel. The results, however, should still be treated with caution, given the low power of unit root tests in finite samples such as those we have here (Blander and Dhaene 2012).

Table 2 shows the test results. The test comes in three versions (without constant and trend, with trend and with constant) and is run on the first, second, and third lag, which are appropriate lag

---

\(^9\) The test is \(\chi^2\)-distributed with \(2N\) degrees of freedom.
lengths for annual data. The null of unit root is rejected for all versions of the test at the 1 per cent level of statistical significance, thus providing strong evidence against the persistence of the level of labour share in developing countries.\textsuperscript{10}

Table 2: Augmented Dickey-Fuller tests

<table>
<thead>
<tr>
<th>Lags</th>
<th>No trend</th>
<th>With trend</th>
<th>With constant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\chi^2(174)=341.4^{***}$</td>
<td>$\chi^2(174)=409.1^{***}$</td>
<td>$\chi^2(164)=478.6^{***}$</td>
</tr>
<tr>
<td>2</td>
<td>$\chi^2(170)=253.7^{***}$</td>
<td>$\chi^2(170)=282.4^{***}$</td>
<td>$\chi^2(152)=398.5^{***}$</td>
</tr>
<tr>
<td>3</td>
<td>$\chi^2(162)=319.8^{***}$</td>
<td>$\chi^2(162)=253.4^{***}$</td>
<td>$\chi^2(142)=363.3^{***}$</td>
</tr>
</tbody>
</table>

Note: Degrees of freedom in parantheses, *** $p<0.01$
Source: Author’s compilation.

8 Conclusion

This paper reveals that measuring the labour share of developing countries is neither direct nor straightforward. There clearly is a quality–coverage trade-off regarding its computation, meaning that the more global the coverage, the greater the prevalence of poor-quality data. This may induce researchers to measure the labour share only at the national or, at most, regional level. However, giving up on cross-country measurement of labour income share must not be the consequence. After all, research on the labour share is simply too important to be hampered by a poor data situation. Although different developing regions can hardly be measured by the same yardstick, global datasets are required to analyse broad trends in labour shares. To date, the development of the labour share in low- and lower middle-income countries remains unexplored, which is a clear gap in economic literature. Labour share data can, for example, be used to better research the distributional effects of labour market policies, the trend of factor and personal income inequalities in the process of development, or the distribution of potential gains from globalization. Particularly in developing countries, knowledge about the labour income share can be used to develop poverty reduction strategies. For example, the finding that productivity gains in developing countries do not translate into broad wage increases suggests that the opportunity to increase the living standard of the poor is missed. In fact, the labour share not only influences income inequality within a country but also has significant implications for aggregate demand and thus growth.\textsuperscript{11} In this function, it can serve as a decision-making parameter in favour or against a certain development path, such as export-led growth.

This paper contributes to the literature by providing the necessary data to address these globally relevant issues also for the developing world. SAMs, a micro-funded representation of the economy, provide input and feedback for the construction of labour share data. Yet future research on the labour share depends crucially on more stalwart and robust data. Counter-checking national accounts against microeconomic data can only be a second-best option. It is hence recommended that national statistics offices increase their efforts to gather data on the (informal) self-employment sector. For this, as Jerven (2012) and others demand, more funding and qualified personnel directed towards reliable and regular data collection will be necessary. Until high-quality data is available, it is inevitable that robustness checks should be made on the national accounts data, with SAMs being just one possibility here.

\textsuperscript{10}The degrees of freedom equal the number of countries in the data times two. This number, however, varies across the different tests, because data gaps prevent the test being performed on all countries.

\textsuperscript{11}Labour income is assumed to have a higher consumption share than capital income.
Notwithstanding these constraints, the dataset can provide the first reliable and valuable insights into the capital–labour ratio in the developing world over recent decades. It is found that the average level of the labour share – at 0.47 – is much less than the established ‘two-thirds’ propounded in economic literature. Similarly, in contrast to the long-lasting belief of constant factor shares, it becomes evident that there is a significant downward trend of the labour share since the early 1990s. In this way, my findings confirm Piketty’s (2014) prominent hypothesis of wealth accumulation proceeding faster than economic growth for most of the developing regions as well: income generation more and more shifts from labour to capital, implying nothing other than a shrinking labour share. An increase in human capital seems not to be able to reverse this trend. Future research should focus on the potential reasons behind this development. The downturn has been taking place during times of increased globalization and evolving financialization, suggesting a link between these phenomena. It is therefore proposed to investigate these relationships and explore whether there is not only a correlation but also causation in place.

**Appendix**

**Countries included**

Algeria, Argentina, Armenia, Azerbaijan, Bahrain, Belarus, Benin, Bhutan, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, Bulgaria, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Chile, China, Colombia, Costa Rica, Cote d’Ivoire, Croatia, Cuba, Czech Republic, Democratic Republic of Congo, Djibouti, Dominican Republic, Ecuador, Egypt, Estonia, Fiji, Gabon, Georgia, Greece, Guatemala, Guinea, Honduras, Hungary, India, Iran, Iraq, Jamaica, Jordan, Kazakhstan, Kenya, Kyrgyz Republic, Lao People’s Democratic Republic, Latvia, Lebanon, Lesotho, Lithuania, Macedonia, Malta, Mauritania, Mauritius, Mexico, Moldova, Mongolia, Morocco, Mozambique, Namibia, Netherlands Antilles, Nicaragua, Niger, Nigeria, Oman, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Republic of Korea, Romania, Russia, Saudi Arabia, Senegal, Serbia, Seychelles, Sierra Leone, Slovak Republic, Slovenia, South Africa, South Sudan, Sri Lanka, Sudan, Suriname, Tajikistan, Tanzania, Thailand, Trinidad and Tobago, Tunisia, Turkey, Ukraine, Uruguay, Venezuela, Zimbabwe.

**References**


