Working paper

International Growth Centre

The Agricultural Productivity Gap in Developing Countries

Douglas Gollin
David Lagakos
Michael E. Waugh

February 2011

DIRECTED BY

FUNDED BY

The Agricultural Productivity Gap in Developing Countries

Douglas Gollin

Williams College

David Lagakos

Arizona State University

Michael E. Waugh

New York University

February 2011

ABSTRACT

According to national accounts data for developing countries, value added per worker is on average four times higher in the non-agriculture sector than in agriculture. Taken at face value this "agricultural productivity gap" suggests that labor is greatly misallocated across sectors in the developing world. In this paper we draw on new micro evidence to ask to what extent the gap is still present when better measures of sector labor inputs and value added are taken into consideration. We find that even after considering sector differences in hours worked and human capital per worker, and alternative measures of sector income constructed from household survey data, a puzzlingly large gap remains.

Email: douglas.gollin@williams.edu, lagakos@asu.edu, mwaugh@stern.nyu.edu. We thank Lisa Starkman and Michael Jaskie for excellent research assistance. For helpful comments we thank Francesco Caselli, Matthias Doepke, Berthold Herrendorf, Larry Huffman, Per Krusell, Kiminori Matsuyama, Maggie McMillan, Ed Prescott, Victor Rios Rull, Richard Rogerson, Mark Rosenzweig, Todd Schoellman, Francis Teal, Aleh Tsyvinski, and Dietz Vollrath, as well as seminar participants at Arizona State, Northwestern, and Yale, and conference participants at the NBER Summer Institute (Economic Growth Group), the SED (Ghent), the Conference on Economic Growth and Cultural Change (Munich), and the World Bank's Annual Bank Conference on Development Economics. This paper was written while Gollin was on leave at the Yale School of Forestry and Environmental Studies. Financial support from the International Growth Centre is gratefully appreciated. All potential errors are our own.

1. Introduction

The agriculture sector accounts for large fractions of employment and value added in developing countries. Almost always, agriculture's share of employment is higher than its share of value added. As a simple matter of arithmetic, this implies that value added per worker is higher in the non-agriculture sector than in agriculture. According to data from national income and product accounts, this "agricultural productivity gap" (APG) is around a factor of four in developing countries, on average.

These large agricultural productivity gaps have several important implications for developing countries. First, with minimal assumptions on production technologies, they imply that labor is misallocated across sectors. Second, they imply that developing countries trail the developed world by a much larger margin in agriculture than in non-agriculture (see, e.g. Caselli (2005), Restuccia, Yang, and Zhu (2008), and Vollrath (2009)). Together, these two implications suggest that the problem of economic development is closely linked to an apparent "misallocation" of workers across sectors, with too many workers in the less-productive agriculture sector.

In this paper, we ask to what extent these gaps are still present when better measures of sector labor inputs and value added are taken into consideration. In other words, we ask how much of the agricultural productivity gaps are due to problems of omitted factors and mismeasurement, as opposed to real differences in output per worker? Several existing studies have argued that these measurement issues may be first-order: Caselli and Coleman (2001), for example, argue that agriculture workers have relatively lower human capital than other workers; Gollin, Parente, and Rogerson (2004) suggest that agriculture output maybe underestimated due to home production; and Herrendorf and Schoellman (2011) claim that measurement error in agricultural value added data are prevalent even across U.S. States. Despite these concerns, the literature does not have a clear answer to how important these measurement issues are in practice in developing countries.

To answer this question, we construct a new database from population censuses and household surveys for a large set of developing countries. We organize our analysis around possible biases that could affect value added per worker in the denominator (employment) and in the numerator (value added). We then use our new database to perform a sequence of adjustments to the data on agriculture's shares of employment and value added. In the first set of adjustments, we use measures of hours worked by sector for 51 developing countries, and measures of human capital by sector for 98 developing countries. We find that taking sector differences in hours and human capital per worker into consideration jointly reduces the size of the average agricultural productivity gap from around four to around two.

We then construct alternative measures of value added by sector using household income surveys from ten developing countries. Our surveys come from the World Bank's Living Standards Measurement Studies (LSMS), which are are designed explicitly to obtain measures of household income and expenditure. They allow us to compute, among other things, the market value of all output—whether ultimately sold or consumed at home—produced by the households. We find that gaps in value added per worker by sector implied by these household income surveys are similar in magnitude to those found in the national accounts. This suggests that mis-measurement of value added in national accounts is unlikely to account for the agricultural productivity gaps implied by national accounts data, at least in these countries.

We then consider a set of other potential explanations for the gaps, including sector differences in labor's share in production, potential discrepancies between income per worker and income per household, and urban-rural differences in the cost of living. We conclude that the agricultural productivity gaps in the developing world are unlikely to be completely explained by any of the measurement issues we address in the paper. What this suggests, we argue, is that a better understanding is needed of why so many workers remain in the agriculture sector, given the large residual productivity gaps that we find in most developing countries. Understanding these gaps will help determine, in particular, whether policy makers in the developing world should pursue polices that encourage movement of the workforce out of agriculture.

We are not the first to point out the existence of large agricultural productivity gaps. Lewis (1955), for example, noted that in developing countries "there is usually a marked difference between incomes per head in agriculture and in industry."[1] These differences in sectoral productivity were viewed as critical by early development economists. Rosenstein-Rodan (1943), Lewis (1955), and Rostow (1960) viewed the development process as fundamentally linked to the reallocation of workers out of agriculture and into "modern" economic activities. More recently, the work of Caselli (2005), Restuccia, Yang, and Zhu (2008), Chanda and Dalgaard (2008), and Vollrath (2009) has shown that the apparent misallocation of workers across agriculture and non-agriculture can account for the bulk of international income and productivity differences. McMillan and Rodrik (2011) argue that reallocations of workers to the most productive sectors would raise income dramatically in many developing countries.

Our contribution is to take a step back and attempt to account for the gaps using richer data on labor and value added at the sector level than in any prior study. In particular, our paper is the first to make use of household survey-based measures of schooling attainment by sector, hours worked by sector, and cost-of-living differences in urban and rural areas. Furthermore, we are the first to compare sector productivity levels computed from "macro" data, based on

[1]The fact that the agriculture productivity gaps are most prevalent in poor countries was first shown by Kuznets (1971), and later documented in richer detail by Gollin, Parente, and Rogerson (2002). Interestingly, Gollin, Parente, and Rogerson (2002) note that the disparities were fairly small in today's rich countries at moments in the historical past when their incomes were substantially lower than at present.

the national accounts, to those implied by "micro" data, based on household surveys of income. Our work is similar in this regard to that of Young (2011), who compares growth rates computed from national accounts data to those computed from household survey data in a set of developing countries.

The paper most closely related to ours is the work of Herrendorf and Schoellman (2011), who ask why agricultural productivity gaps are so large in most U.S. states. A key difference in the conclusions of the two papers is that Herrendorf and Schoellman (2011) argue that systematic under-reporting of agriculture value added is a major factor in accounting for the low relative productivity of agriculture, unlike in our study. The main similarity is that both studies find that sector differences in human capital per worker explain a substantial fraction of the gaps.

Finally, our work relates closely to the recent literature on misallocation and its role in explaining cross-country differences in total factor productivity and output per worker. Seminal examples of this line of research are Restuccia and Rogerson (2008) and Hsieh and Klenow (2009) who focus on the misallocation of capital across firms; or Caselli and Feyrer (2007) who study the misallocation of capital across countries. In contrast, we focus on the potential misallocation of workers *across sectors.* Our focus on the divide between the agriculture and non-agriculture sectors is important because developing countries have the vast majority of their workers in agriculture, suggesting that misallocation between these two sectors may be the most relevant source of sectoral misallocation.

2. Agricultural Productivity Gap — Theory

In this section, we discuss some implications of standard neoclassical theory for data. Consider the standard neoclassical two-sector model featuring constant returns to scale in the production of agriculture and non-agriculture, along with free labor mobility across sectors and competitive labor markets.[2] Free labor mobility implies that the equilibrium wage for labor across the two sectors is the same. The assumption of competitive labor markets implies that firms hire labor up to the point where the marginal value product of labor equals the wage. Since wages are equalized across sectors, this implies that marginal value products are also equalized:

$$p_a \frac{\partial F_a(\mathbf{X})}{\partial L} = \frac{\partial F_n(\mathbf{X})}{\partial L} = w, \tag{1}$$

where subscripts a and n denote agriculture and non-agriculture. Units are chosen here such that the non-agricultural good is the numeraire, p_a is the relative price of the agricultural good, and $\mathbf{X}$ is a vector of inputs (including labor) used in production.

[2] Parametric examples in the literature include Gollin, Parente, and Rogerson (2004), Gollin, Parente, and Rogerson (2007) and Restuccia, Yang, and Zhu (2008).

If the production function displays constant returns to scale, then marginal products are proportional to average products with the degree of proportionality depending on that factors share in production. Defining $1-\alpha_a$ and $1-\alpha_n$ as the shares of labor in production, the constant-returns production functions imply:

$$(1-\alpha_a) \times \frac{p_a Y_a}{L_a} = (1-\alpha_n) \times \frac{Y_n}{L_n}. \tag{2}$$

Noting that $p_a Y_a$ and Y_n equal value added in the agriculture and non-agriculture sector, equation (2) says that value added per worker across the two sectors should be equated (modulo differences in labor shares which we discuss later in Section 6.3). Assuming that labor shares are the same across sectors implies that:

$$\frac{Y_n/L_n}{p_a Y_a/L_a} \equiv \frac{VA_n/L_n}{VA_a/L_a} = 1. \tag{3}$$

If the condition in (3) is not met, then this suggests that workers are misallocated relative to the competitive benchmark. For example, if the ratio of value added per worker between non-agriculture and agriculture is larger than one, we should see workers move from agriculture to non-agriculture, simultaneously pushing up the marginal product of labor in agriculture and pushing down the marginal product of labor in non-agriculture. This process should tend to move the sectoral average products towards equality.

An important point to note in condition (3) is that it does not depend on any assumptions about other factor markets. In particular, labor productivity should be equalized across sectors even in the presence of market imperfections that lead to misallocation of other factors of production. For example, capital markets could be severely distorted, but firm decisions and labor flows should nevertheless drive marginal value products—and hence value added per worker—to be equated. Thus, the model implies that if (3) does not hold in the data, the explanation must lie either in either measurement problems related to labor inputs or in frictions of some kind in the labor market – nothing else.

Writing equation (3) in terms of agriculture's share of employment and output gives:

$$\frac{(1-y_a)/(1-\ell_a)}{y_a/\ell_a} = 1. \tag{4}$$

where $y_a \equiv VA_a/(VA_a + VA_n)$ and $\ell_a \equiv L_a/(L_a + L_n)$. In other words, the ratio of each sector's share in value added to its share in employment should be the same in the two sectors.

The relationship in (4) is the lens through which we look at the data. Under the (minimal) conditions outlined above, we first ask if the condition in (4) holds in cross-country data. One way to think about this exercise is along the lines of Restuccia and Rogerson (2008) and Hsieh and

Klenow (2009) who focus on the the equality of marginal products of capital across firms; or Caselli and Feyrer (2007) who study the equality of marginal products of capital across countries. Here, in contrast, we focus on the value of the marginal product of labor across sectors.

3. The Agricultural Productivity Gap — Measurement and Data

In this section we ask whether, in national accounts data, value added per worker is equated across sectors, as predicted by the theory above. We begin with a detailed—perhaps tedious—description of how the national income and product accounts approach the measurement of agricultural value added and how national labor statistics quantify the labor force in agriculture. We conclude that while there are inevitably some difficulties in the implementation of these measures, there is no reason *ex ante* to believe that the data are flawed.

With these measurement issues clear, we then present the "raw", or unadjusted, agricultural productivity gaps using aggregate value added and employment data. We show that the gap is around a factor of four on average in developing countries, well above the prediction of the theory.

3.1. Conceptual Issues and Measurement: National Accounts Data

The statistical practices discussed below are standard for both rich and poor countries, but there are particular challenges posed in measuring inputs and outputs for the agricultural sector in developing countries. A major concern is that aggregate measures of economic activity and labor allocation in poor countries may be flawed—and may in fact be systematically biased by problems associated with household production, informality, and the large numbers of producers and consumers who operate outside formal market structures. Given these concerns, we focus on the conceptual definitions and measurement approaches used in the construction of national accounts data and aggregate labor measures.

To illustrate the potential problems consider the example of Uganda, a country where household surveys and agricultural census data show that as much as 80 percent of certain important food crops (cassava, beans, and cooking bananas) may be consumed within the farm households where they are grown. Most households are effectively in quasi-subsistence; the government reports that even in the most developed regions of the country, nearly 70 percent of households make their living from subsistence agriculture. In the more remote regions of the country, over 80 percent of households are reported as deriving their livelihoods from subsistence farming (Uganda Bureau of Statistics 2007b, p. 82).

Given these concerns, it is possible that value added measures will by design or construction omit large components of economic activity. As we discuss below this is not the case. Although value added may be measured with error, the conceptual basis for value added measurement is clear and well-defined.

3.2. Measurement of Value Added in Agriculture

Perhaps surprisingly, the small scale and informality of agricultural production in poor countries does not mean that their output goes largely or entirely unmeasured in national income and product accounts. At a conceptual level, home-consumed production of agricultural goods does fall within the production boundary of the UN System of National Accounts, which is the most widely used standard for national income and product accounts. The SNA specifically includes within the production boundary "the production of all agricultural goods for sale or own final use and their subsequent storage" (FAO (1996), p. 21), along with other forms of hunting, gathering, fishing, and certain types of processing. Within the SNA, there are further detailed instructions for the collection and management of data on the agricultural sector.

How is the measurement of these activities accomplished? Accepted practice is to measure the area planted and yield of most crops, which can be surveyed at the national level, and to subtract off the value of purchased intermediate inputs.[3] There are also detailed guidelines for estimating the value of output from animal agriculture and other activities, as well as for the consideration of inventory. Detailed procedures also govern the allocation of output to different time periods.[4] Allowances are made for harvest losses, spoilage, and intermediate uses of the final product (e.g., crop output retained for use as seed). The final quantities estimated in this way are then valued at "basic prices," which are defined to be "the prices realized by [farmers] for that produce at the farm gate excluding any taxes payable on the products and including any subsidies."

Although it is difficult to know how consistently these procedures are followed in different countries, the guidelines for constructing national income and product accounts are clear, and they apply equally to subsistence or quasi-subsistence agriculture as to commercial agriculture. Furthermore, there is no reason to believe that national income and product accounts for poor countries do an intrinsically poor job of estimating agricultural value added (as opposed to the value added in services or manufacturing, where informality is also widespread). Nor is there reason to believe that agricultural value added in poor countries is consistently underestimated,

[3]For some crops, only area is observed; for others, only production is observed. The guidelines provide detailed information on the estimation of output in each of these cases.

[4]The national accounting procedures also provide guidance on the estimation of intermediate input data. In the poorest countries, there are few intermediate inputs used in agriculture. But conceptually, it is clear that purchased inputs of seed, fertilizer, diesel, etc., should be subtracted from the value of output. Data on these inputs can be collected from "cost of cultivation" or "farm management" surveys, where these are available , but the FAO recommends that these data "should be checked against information available from other sources," such as aggregate fertilizer consumption data. Similar procedures pertain for animal products.

rather than overestimated.[5]

3.3. Measurement of Labor in Agriculture

Potential mis-measurement of labor in agriculture is another key concern. Because agriculture in poor countries falls largely into the informal sector, there are not detailed data on employment of the kind that might be found in the formal manufacturing sector. There are unlikely to be payroll records or human resources documentation. Most workers in the agricultural sector are unpaid family members and own-account workers, rather than employees. For example, in Ethiopia in 2005, 97.7 percent of the economically active population in agriculture consisted of "own-account workers" and "contributing family workers," according to national labor force survey made available through the International Labour Organization. A similar data set for Madagascar in 2003 put the same figure at 94.6 percent.

The informality of the agricultural sector may tend to lead to undercounting of agricultural labor. But a bigger concern is over-counting—which would lead to misleadingly low value added per worker in the sector. Over-counting might occur in at least two ways. First, some people might be mistakenly counted as active in agriculture simply because they live in rural areas. In principle, this should not happen; statistical guidelines call for people to be assigned to an industry based on the "main economic activity carried out where work is performed." But in some cases, it is possible that enumerators might count individuals as farmers even though they spend more hours (or generate more income) in other activities. In rural areas in developing countries (as also in rich countries), it is common for farmers to work part-time in other activities, thereby smoothing out seasonal fluctuations in agricultural labor demand. This might include market or non-market activities, such as bicycle repair or home construction.

A second way in which over-counting might occur is if hours worked are systematically different between agriculture and non-agriculture. In this situation, even if individuals are assigned correctly to an industry of employment, the hours worked may differ so much between industries that we end up with a misleadingly high understanding of the proportion of the economy's labor that is allocated to agriculture.[6] We explore this possibility directly in Section 4.1, below.

Note that this type of over-counting would affect sectoral productivity comparisons only if hours worked differ systematically across sectors – so that workers in non-agriculture supply more hours on average than workers in agriculture. At first glance, it might seem obvious that

[5]Nevertheless, many development economists find it difficult to believe that national income accounts data for developing countries can offer an accurate picture of sectoral production. We revisit these concerns later in Section 5, where we construct alternative measures of value added by sector using household survey data from ten developing countries. Although these data have their own limitations, as we discuss later, we find that the large agricultural productivity gaps are present in these household survey data as well.

[6]This is an issue studied in some detail by Vollrath (2010) recently, and dates back to the dual economy theory of Lewis (1955), in which he posited a surplus of labor in agriculture.

this is the case; but much of non-agricultural employment in poor countries is also informal. Many workers in services and even in manufacturing are effectively self-employed, and labor economists often argue that informal non-agricultural activities represent a form of disguised unemployment in poor countries, with low hours worked. To return to the Ethiopian data, in 2005, 88.4 percent of the *non-agricultural* labor force consisted of own-account workers and family labor. Thus, the predominance of self employment and family business holds across sectors. If there are important differences in hours worked across sectors, we cannot simply assume that this results from differences in the structure of employment.

A final way in which over-counting of labor in agriculture might occur is if human capital per worker were higher in non-agriculture than in agriculture. In this were true, we would be overestimating the effective labor input in agriculture compared to non-agriculture. In this case, the underlying real differences in sectoral productivity would be smaller than the measured APGs. We address these possibilities directly in Section 4.3, to follow.

3.4. Raw Agricultural Productivity Gap Calculations

With these measurement issues clear, this section describes the sample of countries, our data sources, and then presents the "raw," or unadjusted, agricultural productivity gaps.

The Sample and Data Sources

Our sample of countries includes all *developing countries* for which data on the shares of employment and value added in agriculture are available. By developing countries, we mean countries for which income per capita, in US Dollars expressed at exchange rates, is below the mean of the world income distribution.[7] We restrict attention to countries with data from 1985 or later, and the majority of countries have data from 1995 or later. We end up with a set of 113 countries which have broad representation from all geographic regions and per-capita income levels within the set of developing countries. In each country we focus our attention on the most recent year for which data are available.

Our main source of data on agriculture's share of employment is the World Bank's World Development Indicators (WDI). We supplement these with employment data by sector compiled by the International Labor Organization (ILO). The underlying source for all these data are nationally representative censuses of population or labor force surveys conducted by the countries' statistical agencies.[8] One advantage of using surveys based on of samples of individuals

[7]This cutoff is arbitrary; however the results of the analysis do not differ meaningfully if we use the classifications of the World Bank or other international organizations.

[8]We exclude a small number of countries in which employment shares in agriculture are based on non-nationally representative surveys, such as urban-only samples, or surveys of hired workers, as opposed to surveys of the entire workforce.

Table 1: Raw Agricultural Productivity Gaps

Measure	Weighted	Unweighted
5th Percentile	1.7	1.1
Median	3.7	3.0
Mean	4.0	3.6
95th Percentile	5.4	8.8
Number of Countries	113	113

Sample is developing countries, defined to be below the mean of the world income distribution. The weighted statistics weight each country by its population.

or households is that they include workers in informal arrangements and the self employed. Surveys of establishments or firms, in contrast, often exclude informal or self-employed producers from their sample.

Workers are defined to be the "economically active population" in each sector. The economically active population refers to all persons who are unemployed or employed and supply any labor in the production of goods within the boundary of the national income accounts (FAO (1996)). There is no minimum threshold for hours worked. This definition includes all workers who are involved in producing final or intermediate goods, including home consumed agricultural goods. In general, employed workers are classified into sectors by their reported main economic activity, and unemployed workers are classified according to their previous main economic activity.

Our data on agriculture's share of value added come from the WDI. The underlying sources for these data are the national income and product accounts from each country. In all cases these data are expressed at current-year local currency units.[9] Industry classifications are made in the majority of cases using the International Standard Industrial Classification System (ISIC).

Raw Agricultural Productivity Gaps

Table 1 reports summary statistics for the raw APGs for our set of developing countries. We refer to these as raw APGs because they are before any adjustments (e.g. for hours worked), unlike the calculations that follow. The first data column describes the APG distribution for the entire sample of 113 countries when weighting by population. Across all countries, the mean

[9] An alternative would be to use a single set of international comparison prices to value the agricultural output of each country. This would be relevant if we were making comparisons of real agricultural output per worker across countries, as in Caselli (2005), Restuccia, Yang, and Zhu (2008), Vollrath (2009) or Lagakos and Waugh (2011). In the current paper, however, we are interested in comparing the value of output produced per worker across sectors within each country.

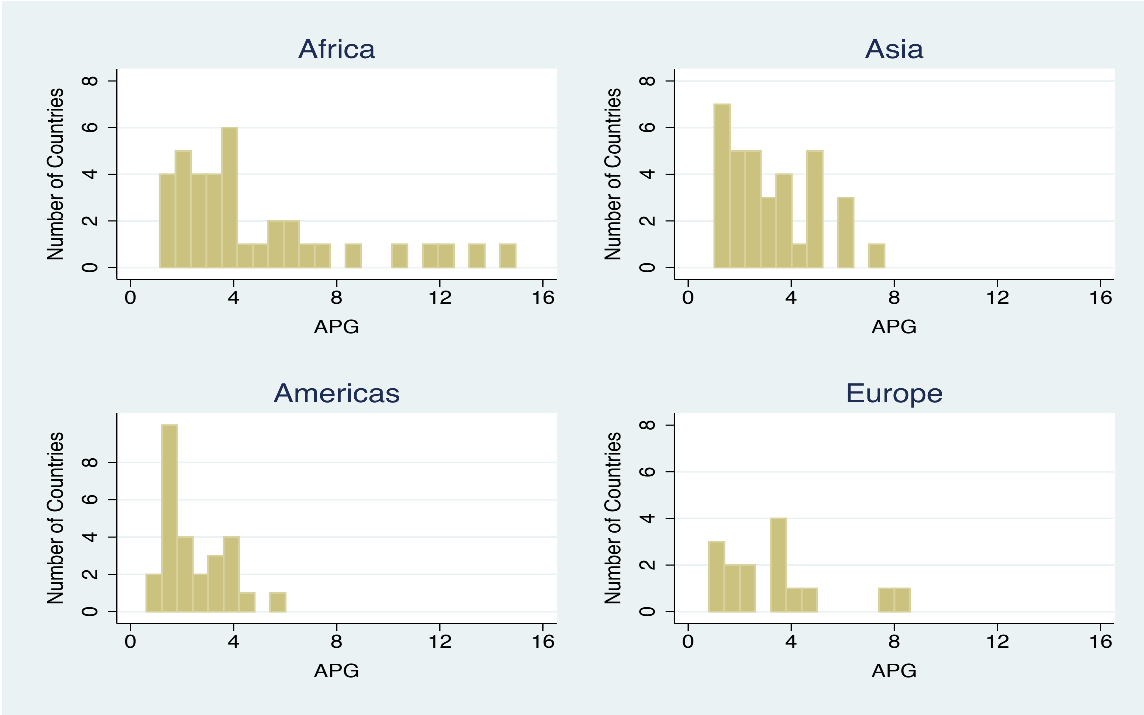


Figure 1: Distribution of APGs by Region

value of the gap is 4.0, implying that value added per worker is approximately four times higher in non-agriculture than in agriculture. The median is slightly lower, at 3.7. Even at the 5th percentile of the distribution, the gap is greater than unity (1.7), implying that in almost all countries for which we have data, the simple prediction of (4) is inconsistent with the data. At the 95th percentile of the distribution, the gap is 5.4.

The second data column of Table 1 presents the same statistics when not weighting. The results are largely similar, with the unweighted mean APG at 3.6 and the median at 3.0. When not weighting, the range of gaps is larger across countries. The 5th percentile is 1.1, and the 95th percentile is now 8.8. Still, the majority of countries have gaps above unity, contrary to the prediction of (4).

Figure 1 shows histograms of the APG by region. Africa has the highest average APG, and all countries with gaps above ten (Burkina Faso, Chad, Guinea, Madagascar and Rwanda) are in Africa. Still, in all regions—Africa, Asia, the Americas and Europe—the average country is well above unity, and each region has a number of countries with gaps above four. These data suggest that the large gaps are not confined to developing countries in one area of the world.

Relative to the discussion in Section 2, it is abundantly clear that the data are not consistent with (4), which would give an APG of one. The raw data suggest very large departures from parity in sectoral productivity levels among these developing countries.

Differences of this magnitude are striking. If we take these numbers iterally, they raise the possibility of very large misallocations between sectors within poor countries. Are such large disparities plausible? Do these numbers reflect underlying gaps in real productivity levels and living standards? Or do they largely reflect flawed measurements of labor inputs and value added? In the following sections, we discuss the new data we bring to bear on the question, and consider a number of ways in which mismeasurement may occur. We will also compare the magnitude of these possible mismeasurements with the observed gaps in productivity.

4. Improved Measures of Labor Inputs by Sector

In this section, we report the results of efforts to adjust the productivity gaps to account for potential differences in the quantity and quality of labor inputs across sectors. We base this analysis on a new database that we constructed, which contains sector-level data on average hours worked and average years of schooling for a large set of developing countries. We construct our data using nationally-representative censuses of population and household surveys, with underlying observations at the individual level.

One part of our data comes from International Integrated Public Use Microdata Series (I-IPUMS), from which we use micro-level census data from 44 developing countries around the world. We also get data on schooling attainment by sector from 51 countries from the Education Policy and Data Center (EPDC), which is a public-private partnership of the U.S. Agency for International Development (USAID) and the Academy for Educational Development. From a number of other countries we get schooling and hours worked from the World Bank's LSMS surveys of households. The remainder of the data comes from individual survey data and published tables from censuses and labor force surveys conducted by national statistical agencies. Table 7 in Appendix A details the sources and data used in each of the 113 developing countries in our data.

4.1. Sector Differences in Hours Worked

We now ask whether the sectoral productivity gaps are explained by differences across sectors in hours worked. We find that in most of the countries for which we have data on hours worked, there are only modest differences in hours worked by sector; on average, workers in non-agriculture supply around 1.2 times more hours than workers in agriculture. Thus, hours worked differences are unlikely to be the main cause of the large APGs we observe.

We measure hours worked for all workers in the labor force, including those unemployed during the survey, for whom we count zero hours worked. The typical survey asks hours worked in the week or two weeks prior to the survey, although some report average hours worked in the previous year. We classify people as workers in either agriculture or non-agriculture, according to their main reported economic activity. For unemployed workers not reporting

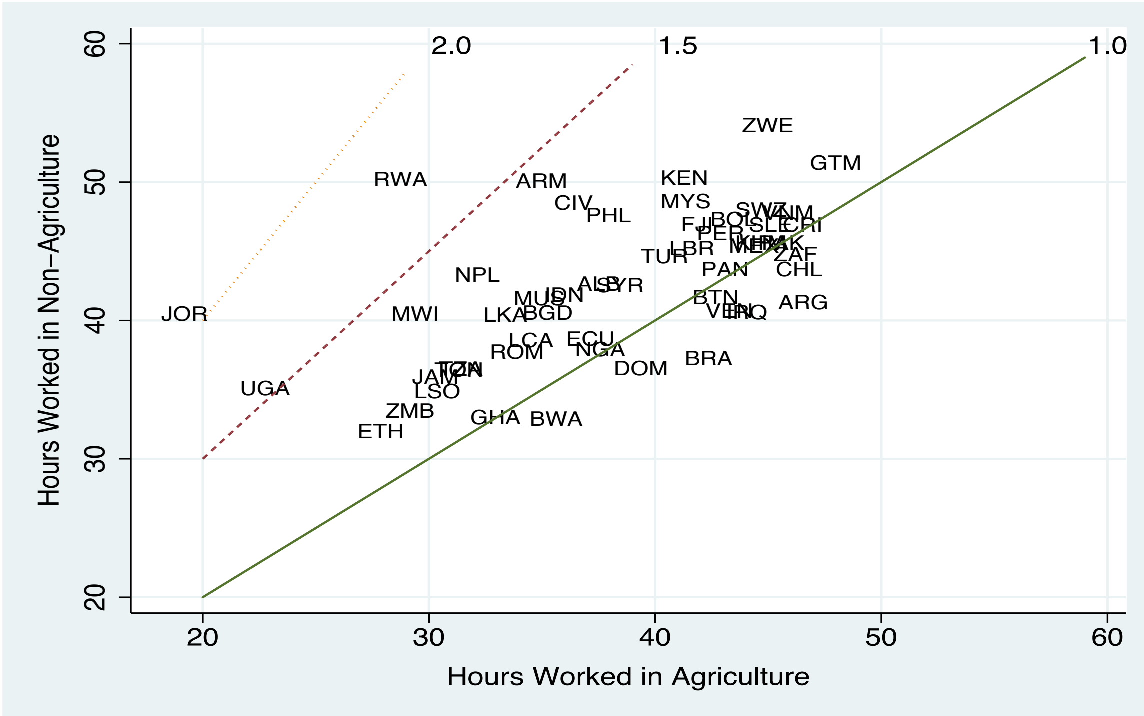


Figure 2: Hours Worked by Sector

a main economic activity, we classify them as agricultural if they live in rural areas, and as non-agricultural if they live in urban areas.

For some countries, we cannot obtain measures of hours by agricultural or non-agricultural employment, but we are able instead to use hours worked by urban-rural status. Table 7 lists the countries for which we use urban-rural status to construct our hours measures. In these countries, as in the others, we count unemployed workers as having worked zero hours.[10] Using urban-rural status in some countries represents a potential limitation of our data, as the non-agricultural (agricultural) workforce and urban (rural) workforce do not correspond exactly to one another. However, in those countries for which we can measure average hours by both urban-rural status and agriculture-non-agriculture status, the two give similar average hours measures.

Figure 2 shows hours worked in non-agriculture, plotted against hours worked in agriculture, for each of the countries with available data. The 45-degree line, marked 1.0, corresponds to a situation where average hours worked are identical in the two sectors. Similarly, the other two lines represent factor of 1.5 and 2.0 differences in hours worked. Most of the observations are clustered closely around the 1.0 line, and all but a few are below the 1.5 line, meaning that hours worked differences across sectors are generally modest. An arithmetic average across countries gives a factor 1.2 difference in hours worked in non-agricultural compared to agriculture.

[10]Our results change very little when using average hours among only employed workers.

This pattern does not vary much across regions, with average ratios of 1.2 for developing countries in Africa, Europe, and Asia, and an average ratio of 1.0 in the Americas. Uganda and Rwanda have the most pronounced differences in hours worked, with roughly 1.7 times as many hours worked in non-agriculture as agriculture in these countries. Notably, these countries also have large APGs.[11] So while hours worked differences overall do not seem to explain much of the large APGs (as that would require an average ratio of around 4.0), in some countries lower hours worked in agriculture seems to be an important part of their large measured gaps.

4.2. Hours Worked: A Further Breakdown

In the calculations above, we classify workers by their primary sector of employment and then attribute all their labor hours to that sector. A potential concern is that individuals classified as agricultural (non-agricultural) work a substantial fraction of their hours in non-agricultural (agricultural) activities. For example, suppose that individuals in agriculture in fact devote a large fraction of their hours to non-agricultural activities. In this case, we would be overcounting their hours worked in agriculture, leading to an underestimate of average labor productivity in agriculture. For this to be quantitatively important, it would need to be the case that a substantial fraction of hours are misallocated in this fashion.

To explore this possibility, we analyze individual-level data from LSMS household surveys for a number of countries with available data. Table 2 shows the results of this analysis. In this table, we show the hours worked in each sector by workers classified as agricultural or non-agricultural. As noted above, the classification of workers is based on their primary sector of employment. However the LSMS data allow us to measure the hours worked by individuals across all their economic activities.

These measures of hours worked show that to an overwhelming degree, those individuals classified as working in agriculture do in fact allocate their time to agricultural activities; similarly, workers classified as non-agricultural allocate almost all of their time to non-agricultural activities. In all of these cases except that of the 1998 Ghana LSMS, we find that agricultural-classified workers devote 95 percent or more of their hours to agriculture; and in every case we find that workers classified as non-agricultural devote at least 94 percent of their hours to non-agricultural activities.

Although we have not carried out these painstaking calculations for all the countries with available micro data, we feel comfortable on the basis of the available evidence that the procedure we are using for calculating hours worked by sector is accurately reflecting the allocation of

[11] Jordan is also an outlier, but does not have a particularly large APG or agricultural employment share.

Table 2: Hours Worked: A Further Breakdown

		Sector of Hours Worked	
Country	Worker Classification	Agriculture	Non-agriculture
Cote d'Ivoire (1988)	Agriculture	35.1	1.0
	Non-agriculture	0.7	49.2
Ghana (1998)	Agriculture	28.8	3.7
	Non-agriculture	2.0	30.6
Guatemala (2000)	Agriculture	47.6	1.3
	Non-agriculture	0.8	49.1
Malawi (2005)	Agriculture	26.4	1.4
	Non-agriculture	2.3	38.2
Tajikistan (2009)	Agriculture	39.5	0.1
	Non-agriculture	0.1	39.3

Note: Workers are classified by sector according to their primary sector of employment. Hours are classified by sector of job for each of the workers' jobs.

hours at the individual level.[12]

4.3. Sector Differences in Human Capital

We next ask to what extent sectoral differences in human capital per worker can explain the observed APGs. We show that while schooling is lower on average among agricultural workers, the differences are not large enough to fully explain the measured gaps.

Our calculations in this section are related to those of Vollrath (2009), who also attempts to measure differences in average human capital between workers in agriculture and non-agriculture. While both sets of calculations have their limitations, ours improve on those of Vollrath (2009) in several dimensions. Most important, our calculations come from nationally representative censuses or surveys with direct information on educational attainment by individual.[13] We also end up with estimates for a much larger set of countries. Finally, we attempt to adjust for

[12]At first glance, these numbers might appear to be inconsistent with the stylized fact that non-farm income represents an important source of earnings for rural households. In fact, our results are entirely consistent with that stylized fact. The reason is simply that "rural" and "agricultural" are different categories. In all of the micro data sets that we have examined, there are substantial fractions of rural households that are classified as non-agricultural. For example, in the 1998 Ghana LSMS data, 29.2 percent of rural workers are classified as non-agricultural, and 44.5 percent of rural income was non-agricultural. In our view, this emphasizes the point that the relevant productivity differences in developing countries are between the agriculture sector and non-agricultural sectors, rather than simply between rural and urban areas.

[13]Those used by Vollrath (2009) are imputed using school enrollment data.

quality differences in schooling across sectors. Our calculations are also similar to those of Herrendorf and Schoellman (2011), who measure human capital differences across sectors in U.S. States.[14]

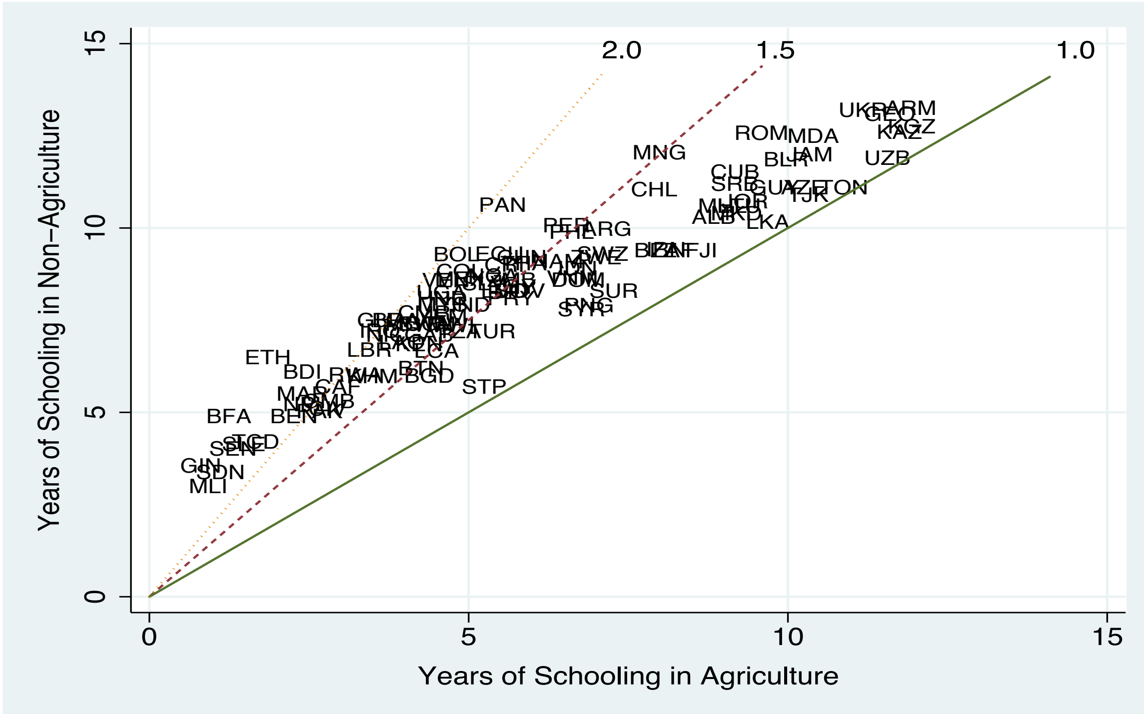
As before, we compute average years of schooling by sector using household survey and census data. As for our hours measures, we use all employed or unemployed people in the agricultural and non-agricultural sectors when possible, and otherwise we use urban-rural status. When direct measures of years of schooling completed are available, we use those. When they are not, we impute years of schooling using educational attainment data. Table 7 details which countries use years of schooling directly and which use educational attainment data.[15] These imputations are likely to yield noisy measures of years of schooling of course, as a category such as "some secondary schooling completed" (for example) could correspond to several values for years of schooling . However, in all countries where we impute schooling, we do so in exactly the same way for non-agricultural and agricultural workers. Thus, the noisiness should in principle not systematically bias our measures of average years of schooling by sector.

Figure 3(a) shows our results for the 98 countries for which we constructed average years of schooling by sector. Again, the 45-degree line, marked 1.0, indicates equality in schooling levels, and the lines 1.5 and 2.0 represent those factor differences in years of schooling. As can be seen in the figure—in literally every country—average schooling is lower in agriculture than non-agriculture. Countries with the highest levels of schooling in agriculture tend to be closest to parity between the sectors. For example, the former Soviet block countries of Armenia, Kazakhstan, Uzbekistan, Georgia, and Ukraine have the highest schooling in agriculture and among the lowest ratios of non-agricultural to agricultural schooling. The ratios are generally higher for countries with less schooling among agriculture workers, with the lowest generally coming in francophone African countries. Mali, Guinea, Senegal, Chad and Burkina Faso have the lowest schooling for agricultural workers and among the highest ratios.

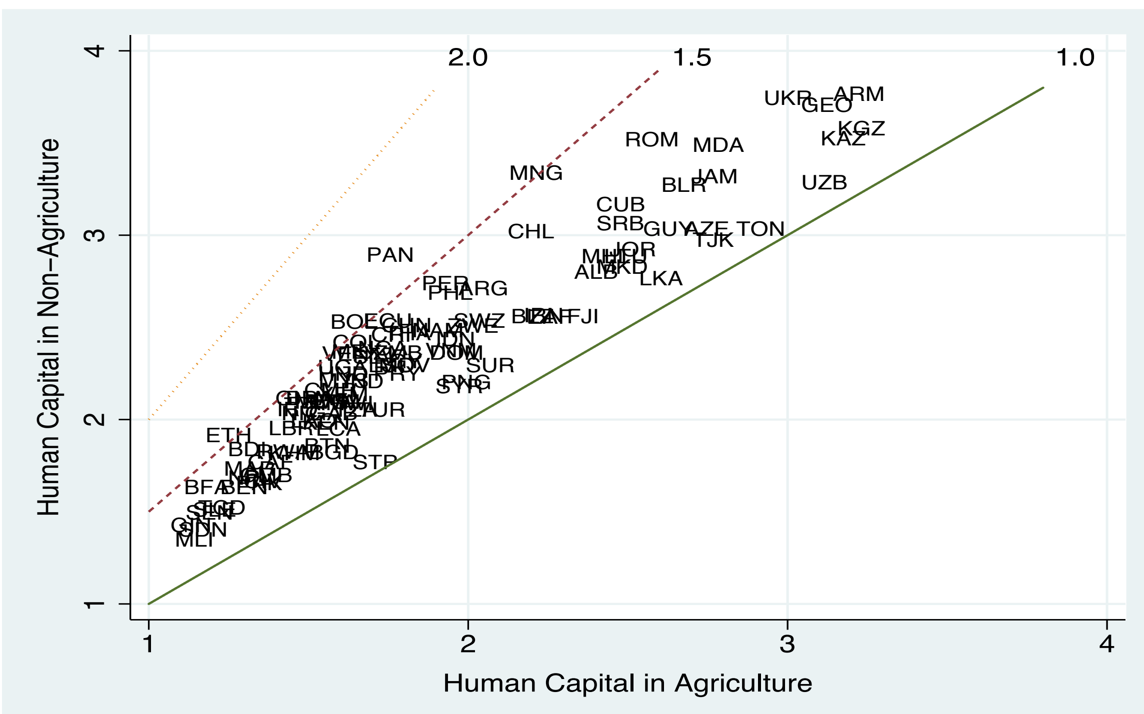
We are interested in the differences in human capital per worker that can be attributed to these differences in schooling. To turn years of schooling into human capital, we consider several different approaches. All of them assume that average human capital in sector j of country i can be expressed as $h_{j,i} = \exp(r_i \cdot s_{j,i})$ where $s_{j,i}$ is average years of schooling in sector j of country

[14] One advantage of the calculations of Herrendorf and Schoellman (2011), relative to those of the current paper, is that they allow for sector differences in human capital arising through sector differences in returns to experience. They find lower returns to experience among agriculture workers than other workers. Measuring returns to experience by sector across the developing world is a task outside the scope of the current paper. Lagakos, Moll, and Qian (2012) use data from a set of countries from all income levels to argue that returns to experience are generally lower in developing countries than in richer countries, and that this increases the importance of human capital in accounting for income differences across countries. However, they do not (yet) measure returns to experience by sector in the developing countries.

[15] The data on educational attainment provide categories such as "some primary schooling completed," rather than specific measures of years of schooling.



(a) Years of Schooling by Sector



(b) Human Capital by Sector

Figure 3: Schooling and Human Capital by Sector

i, and r_i is the return to each year of schooling in country i. Many macro studies simply assign a constant value to r_i across countries—assuming, for example, that each year of schooling increases wages by around 10 percent. A slight variation on this approach is to assume that there is some concavity in years of schooling, so that the first several years of schooling gives a higher return, in terms of human capital accumulation, than subsequent years of schooling.[16]

Figure 3(b) plots the results for human capital by sector using this approach. The resulting estimates of human capital by sector suggest that in virtually all countries, the average non-agricultural worker has between 1.0 and 1.5 times as much human capital as the average agricultural worker. The biggest ratios are still for the countries with the lowest human capital in both sectors, but the differences are less pronounced than those of schooling. This is simply because having (say) twice as many years of schooling implies having considerably less than twice as much human capital (see, e.g., the discussion of Mincer return estimates in Banerjee and Duflo (2005) and Psacharopoulos and Patrinos (2002)). The weighted average across countries is a factor 1.4 difference in human capital of across the two sectors. The average is a little higher in the Americas at 1.5, and lower in Europe at 1.3.[17]

By using the same rate of return to schooling for all countries, we can calculate human capital for a large set of countries. However, one might worry that there are important differences across countries in the rates of return to schooling, and hence in the human capital accumulation of individuals with different years of schooling. To address this concern, we use country-specific estimates of the returns to schooling that have been compiled in three previous studies. Two of these sets of estimates can be traced to Psacharopoulos and Patrinos (2002), who generated a large list of country-specific rates of return, based on Mincer-type regressions. Based on these data, Banerjee and Duflo (2005) offered a modified set of estimates; an updated data set from the World Bank also provides estimates for some additional countries and some modifications to other numbers. Finally, a third set of country-specific estimates of returns to schooling comes from the work of Schoellman (2012). Unlike the other two data sources, Schoellman (2012) bases his estimates on the earnings of migrants to the United States, based on census data. Earnings are observed for migrants with different levels of education, allowing for estimates of country-specific rates of return to schooling.

We calculate sectoral differences in human capital per worker using all three sources of data on country-specific returns to education. Because these three data sets are incomplete in terms of country coverage, we can only calculate the sectoral differences for limited numbers of coun-

[16]This is the approach used, for example, by Hall and Jones (1999) and Caselli (2005).

[17]By comparison, Vollrath (2009) finds that human capital in non-agriculture is higher by a factor of only around 1.2, averaging across countries. In other words, we suggest that more of the agricultural productivity gaps can be explained by human capital differences. The proximate reason for this is that our measures yield higher levels of schooling in both sectors than Vollrath's, but we find a substantially higher level of schooling in non-agriculture than he does, while our measures for the agricultural sector are only slightly higher.

tries. The World Bank data and the Banerjee and Duflo (2005) data give essentially the same results; as a result, we report only the former. The data of Schoellman (2012) show lower returns to schooling in the poorest countries and thus generate different numbers for sectoral human capital levels.

Using the World Bank data, based on Psacharopoulos and Patrinos (2002), we find that sectoral differences in years of schooling translate into a level of human capital per worker that is 1.5 times higher in non-agriculture than in agriculture; in other words, each worker has 50% more human capital in non-agriculture. This compares to a figure of 1.4 when we use a constant 10 percent rate of return to a year of schooling for all countries. The regional differences we find using these data range from 1.5 in Africa to 1.6 in the developing countries of the Americas.

Using the estimates of Schoellman (2012), we find that the sectoral differences in human capital are dampened considerably. Because Schoellman (2012) generally finds low rates of return to schooling in poor countries, and since these are the countries where the sectoral differences in schooling levels are (proportionally) the greatest, the Schoellman (2012) data tend to reduce the importance of schooling differences across sectors. With these estimates, we find that human capital per non-agricultural worker is on average 1.3 times higher than human capital per worker in agriculture. Regional differences are relatively small, with a figure of 1.2 for African countries and 1.3 in Asia.

To summarize our findings in this section, we find that there are substantial differences in human capital per worker across sectors. Because education levels and educational attainment are almost universally lower in agriculture than in non-agriculture, we estimate that workers in the non-agricultural sector have 1.3 to 1.5 times as much human capital than those in agriculture, depending on our source of data. This does appear to be an important source of differences in average labor productivity. However, these differences alone are not able to account fully for the raw gaps observed in the data.

4.4. Adjusting for Education Quality using Literacy Rates

One limitation of the analysis above is that our procedure treats years of schooling among agriculture workers as equally valuable as those among non-agriculture workers. There is evidence, however, that the quality of schooling in rural areas in many developing countries is below that of schooling in urban areas. For example Williams (2005) and Zhang (2006) provide evidence that literacy rates and test scores in mathematics and reading are most often lower in rural schools than urban ones. Thus, our estimates above may tend to overestimate the human capital level of agriculture workers, who in general received their schooling from lower-quality rural schools.

To consider the effect of adjustments for education quality differences, we present a simple

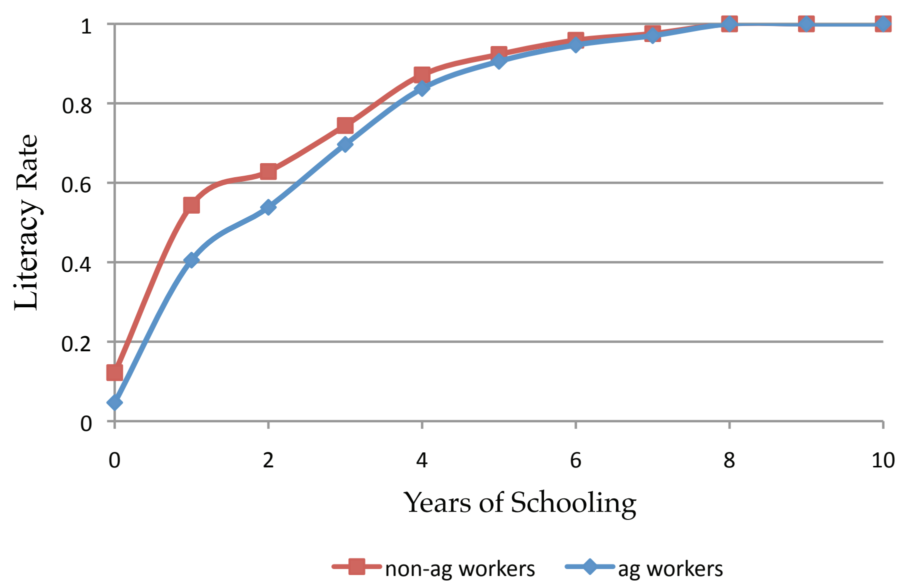


Figure 4: Literacy Rates by Years of Schooling, Uganda

new method of adjusting for quality differences in schooling among agricultural and non-agricultural workers using literacy data. The basic idea is that literacy, particularly in primary schools, is one of the main components of the human capital that students receive through schooling. Thus literacy rates for workers by years of schooling completed in the two sectors are informative about quality differences in schooling received by workers in the two sectors.

What we observe in our micro data are the literacy rates for non-agricultural and agricultural workers in country i conditional on having completed s years of schooling, which we denote $\ell_i^n(s)$ and $\ell_i^a(s)$ for $s = 0, 1, 2....$ If the quality of schooling received were the same for the two groups, then $\ell_i^n(s)$ and $\ell_i^a(s)$ would be the same (at least approximately) for each s. Instead, we find that in almost every country in our sample, $\ell_i^n(s) > \ell_i^a(s)$ for most or all values of s. In other words, literacy rates are higher for non-agricultural workers at most or all schooling levels, and hence an average year of schooling received by the non-agricultural workers must have been more effective than an average year received by the agricultural workers.

Figure 4 illustrates the literacy data by sector for Uganda. The x-axis contains years of schooling completed and the y-axis shows the literacy rates $\ell_i^n(s)$ and $\ell_i^a(s)$ for the two sectors by years of schooling completed. Note that at each year of schooling completed, non-agricultural workers have literacy rates that are at least as high as those of agricultural workers, with the biggest difference coming for the lower years of schooling completed (particularly 1 year.) The differences in literacy are largely absent by about 10 years of schooling completed, with virtually all workers literate by then, hence we cut the graph off then.

To pin down how much more effective a year of urban education is than a rural year in country

i, our method is the following. First we interpolate the literacy outcome data for agricultural workers and create a continuous literacy function of schooling: $\tilde{\ell}_i^a(s)$. This function, which for the case of Uganda is the dotted curve in Figure 4, allows us to evaluate literacy rates for agricultural workers for non-integer years of schooling. We then posit that, in country i, s years of schooling for agricultural workers are as effective as $s\gamma_i$ years of schooling for non-agriculture workers, and set γ_i to the value that solves

$$\min_{\gamma} \sum_{s=1}^{\bar{s}} \left(\tilde{\ell}_i^n(\gamma s) - \tilde{\ell}_i^a(s) \right)^2 . \tag{5}$$

In other words, we pick the value of γ that equates as closely as possible the literacy rates between agricultural workers with s years of schooling and non-agricultural workers with $s\gamma$ years of schooling, over a range of s values up to some value $\bar{s}$. Since primary school ends at 5 years in most countries, and since most workers are literate by then, setting $\bar{s}$=5 seems warranted. In the example of Uganda, we find that $\gamma_{UGA} = 0.82$, meaning that a each year of schooling for agriculture workers is worth 82 percent of a year of schooling for the typical non-agriculture worker in terms of acquiring literacy. We assume therefore that a year of schooling for agriculture workers is worth 82 percent of a year of schooling for non-agriculture workers in terms of acquiring human capital.

Available data allowed us to make similar calculations for 17 other developing countries.[18] The average estimate is 0.87, suggesting real but modest differences in schooling quality across countries. The range of all other estimates runs from a low of 0.62 in Guinea to a high of 0.95 in Bolivia. Mexico, Venezuela and Vietnam are other notably low estimates, all around 0.75. Only Tanzania has an estimate above one; why rural schools appear to fare better than urban ones is a question for which we do not yet have a clear answer.

Figure 5 displays the human capital in each sector for the 17 countries for which we made the quality adjustments, where $h_{a,i}^q = \exp(\hat{\gamma}_i s_{a,i})$ for each country i. Countries above the 45-degree line are those that have higher ratios once the quality adjustments are made. As can seen from the figure, human capital is between 1.2 and 1.6 times higher in non-agriculture, with an average of 1.4. Thus, while adjusting for quality of human capital makes somewhat of a difference relative to the unadjusted calculations with more countries, the differences are modest.

We conclude that these education quality adjustments, while perhaps crude, suggest that quality differences in schooling do not substantively alter the our findings regarding human capital

[18] These countries, and their estimated quality differences (expressed as the number of years of urban schooling equivalent to one year of rural schooling) are, Argentina (0.87), Bolivia (0.95), Brazil (0.95), Chile (0.92), Ghana (0.90), Guinea (0.62), Malaysia (0.93), Mali (0.89), Mexico (0.77), Panama (0.87), Philippines (0.80), Rwanda (0.88), Tanzania (1.25), Thailand (0.90), Uganda (0.82), and Venezuela (0.78).

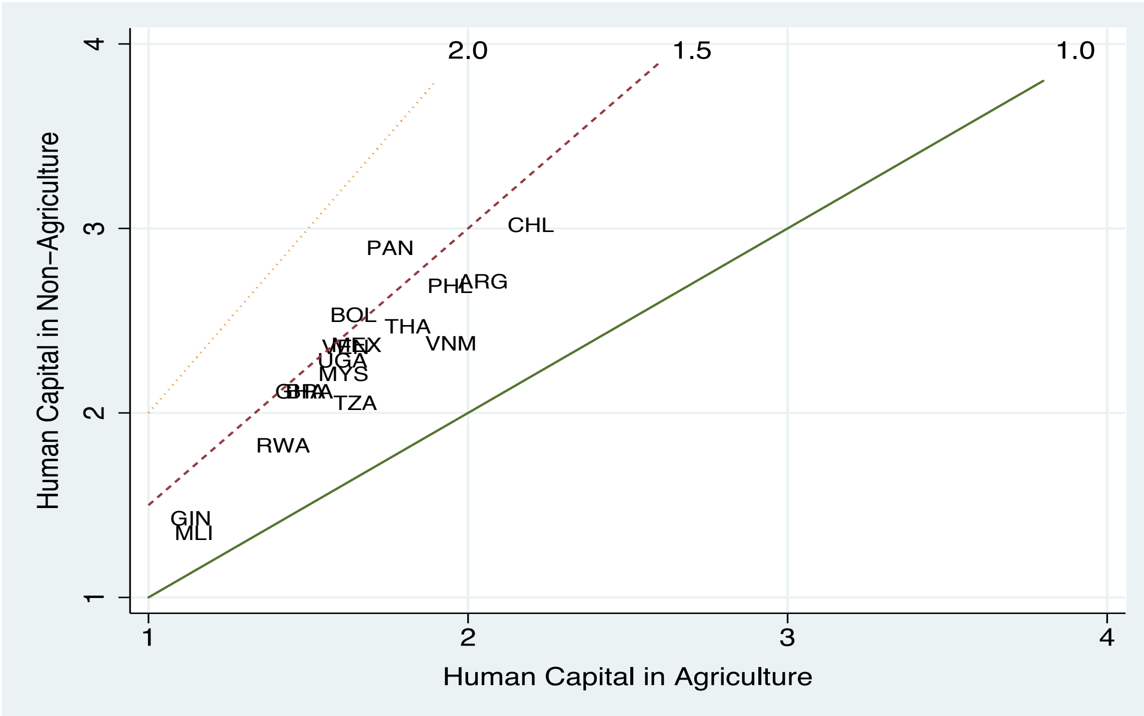


Figure 5: Human Capital by Sector, Adjusted for Quality

differences by sector. In the average developing country, human capital per worker is 1.4 times as high in the non-agriculture sector as the agriculture sector, and this ratio is basically unchanged when we correct for schooling quality using our method.

4.5. Adjusted APGs

We now compute the "adjusted" agricultural productivity gaps, which take into consideration the sector differences in hours worked and human capital. We do not have all these data for all the countries in our sample, and hence we proceed in two ways. First, we compute the adjusted APG for each of the countries for which we have complete data (which here consists of sectoral differences in hours worked and sectoral differences in schooling). Second, we compute the adjusted APG for every country in our sample by imputing any missing data. We do this by assigning any missing value to be the weighted average ratio across all other countries with data.[19] For each country, we construct the adjusted APG by dividing the raw APG by the ratio of hours worked and the ratio of human capital.

Table 3 shows summary statistics of the adjusted APG distributions for countries with complete data and then all countries in our data. The adjusted gaps in the table are based on the assumption of a 10 percent rate of return per year of schooling in all countries. For both complete-data

[19]Most of the imputed values are for ratios of hours worked, since hours measures were available for the fewest countries. Our results do not change substantially when using alternative imputation methods, such as projecting missing data using GDP per capita and regional dummies.

Table 3: Adjusted Agricultural Productivity Gaps

Measure	Complete Data	All Countries
5th Percentile	0.8	0.7
Median	2.2	1.9
Mean	2.1	2.1
95th Percentile	3.9	3.9
Number of Countries	50	113

Sample is developing countries, defined to be below the mean of the world income distribution. "Complete data" means the set of countries with data on hours and human capital. "All countries" means that when data is missing is it imputed as the mean ratio across all countries with data available. All statistics are unweighted.

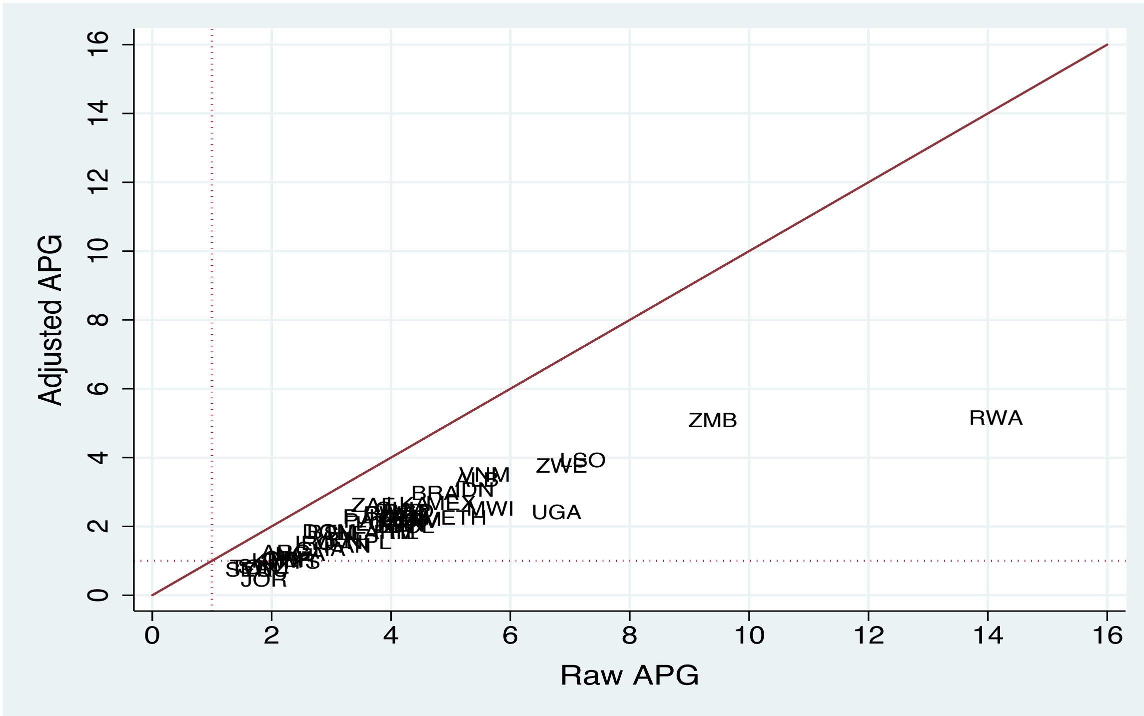
countries and all countries, the mean adjusted gap is 2.1. The median adjusted gaps are 2.2 for the countries with complete data, and 1.9 for all countries. The range runs a 5th percentile of around 0.7 to a 95th percentile of 3.9.

Using country-specific rates of return to schooling, the numbers change slightly. Using the World Bank numbers based on Psacharopoulos and Patrinos (2002), the mean APG for 27 countries with complete data is 1.9; for all countries, the mean is 2.0. Using the returns data of Schoellman (2012), we have 25 countries with complete data. Among these countries, the mean adjusted APG is 2.3, and the median is 2.2. Extending the results to countries without complete data, we get an mean APG of 2.6 and an median of 2.1.

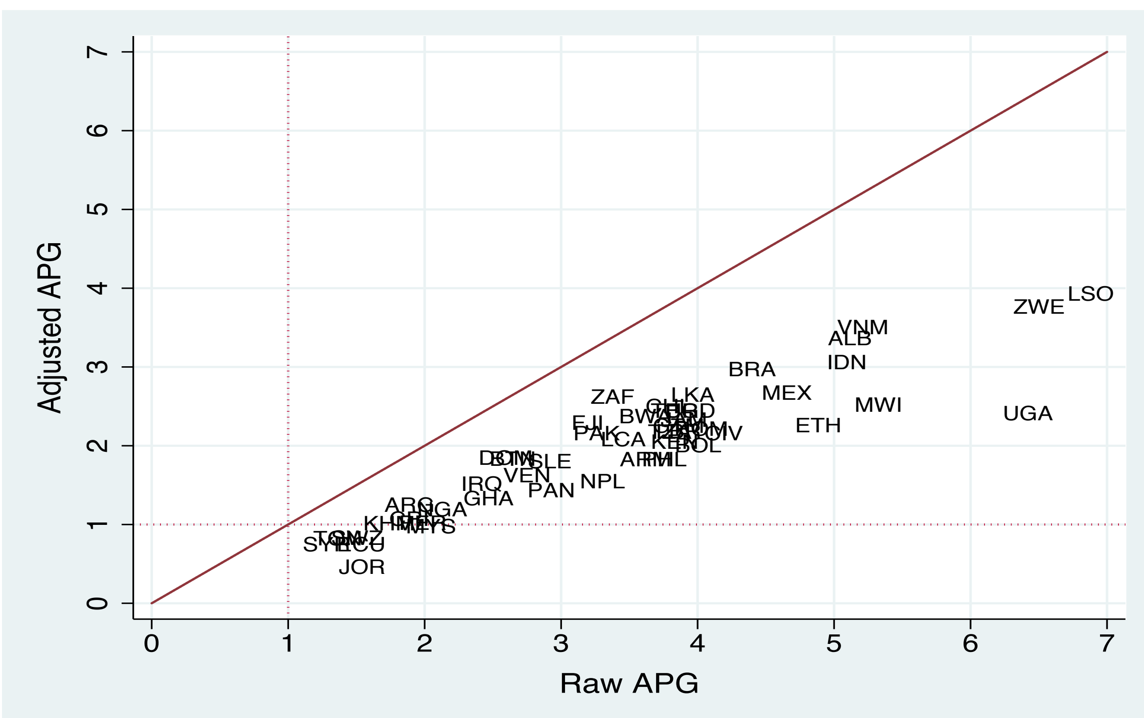
Taken together, these numbers suggest that the typical country has an APG that shrinks dramatically once our adjustments are made. Using any of our three measures for returns to schooling, and either means and medians, the APGs fall from around four on average to around two on average. In virtually all countries, adjusted APGs are substantially lower than their raw counterparts.

Figure 6 provides more detail on how the adjusted and raw APG values differ for the countries for which we have complete data. The top panel of Figure 6(a) shows all countries. Most notably, Rwanda and Zambia have big raw gaps, of 14 and 9.5 respectively, and much smaller gaps after our adjustments, with both countries below 4. Figure 6(b) provides a "close up" of the same countries minus those with raw APG values of over 7. Now one can see that Lesotho and Uganda have initial gaps of around 7, and adjusted gaps of around 2 and 3 respectively. Interestingly, the remainder of the countries tend cluster along a ray of slope one-half from the origin, suggesting that our adjustments explain around one half of their raw gaps.

Explaining roughly one half the raw APG measures represents success for our adjustments.



(a) All Countries with Complete Data



(b) Close up – Countries with Complete data and Raw APG ≤ 7

Figure 6: Raw and Adjusted Agricultural Productivity Gaps

Still, we note that our adjustments thus far only part of the way towards explaining the differences in productivity between sectors. The remaining gap of around two is puzzlingly large. It suggests that the average non-agriculture worker has roughly twice the income as the average non-agricultural worker. The implication is that there should be large income gains from moving workers out of agriculture and into other economic activities.

Interestingly, longitudinal evidence on the effects of migration show large gains in income of roughly the same magnitude that our numbers here suggest. Beegle, De Weerdt, and Dercon (2011) use a unique tracking survey from Tanzania from to study the economic outcomes of a sample of individuals in rural Kagera province from 1994 ten years later in 2004. Unlike in other studies, Beegle, De Weerdt, and Dercon (2011) track workers who migrate to other villages or urban areas anywhere in our outside of Tanzania. What they find is that workers who moved from agriculture to non-agriculture employment increased their income by rough a factor 1.7. By comparison, those that stayed in agriculture saw income increases of around 1.2. These results provide additional evidence that non-agriculture employment is associated with far higher income than agricultural employment.[20]

5. Measures of Value Added by Sector from Micro Data

We now ask to what extent the agricultural productivity gaps implied by national accounts data are an artifact of mismeasurement of agricultural value added in national accounts data in practice. Mismeasurement may manifest itself in several ways. First, while national accounts data in theory should measure home production, agriculture output in practice maybe underestimated due to home production, as argued by Gollin, Parente, and Rogerson (2004). Second, national accounts data may feature other types of bias disproportionately affecting agriculture. For example Herrendorf and Schoellman (2011) argue that the agricultural productivity gaps present in the majority of U.S. states largely arise from mismeasurement due to the treatment of land and proprietors income.[21]

To answer this question, we use *household survey data* for a set of developing countries to construct new alternative measures of value added by sector. These micro-data allow us to com-

[20] Bryan, Chowdhury, and Mobarak (2011) conduct an related controlled experiment in Bangladesh. The authors find that by providing rural farmers with a very small amount of cash, plus a list of potential employers in a nearby city, they are far more likely to migrate during the slow agricultural season. Furthermore, these workers and their families are experience sizable increases in income relative to workers who stayed behind.

[21] Herrendorf and Schoellman (2011) find that in the U.S. national accounts data, payments by farm owners for for rental of land are subtracted from value added in agriculture. Instead, it should be included in agricultural value added, as it is simply a payment to a factor employed in the production of agricultural output. They also find that estimates of proprietors income in non-agricultural businesses are adjusted upwards (substantially) in an attempt to correct for under-reporting, while such a correction is not made for agricultural proprietors. Once these two errors are fixed, and once sector differences in human capital are accounted for, value added per worker is roughly identical across sectors in the majority of U.S. states.

pute income by economic activity for nationally representative samples of households, which we then aggregate to construct value added by agricultural and non-agricultural activity. A key feature of these data is that we observe home production, which may or may not be accounted for properly in the national accounts.

What we find is that the shares of value added computed from the household data are similar to those of the national accounts. As a result, the agricultural productivity gaps we compute using the household data are similar to those implied by the national accounts. While the household survey data are not without their own limitations, as we discuss below, these results suggest that the agricultural productivity gaps in developing countries are real, rather then artifacts of measurement problems with national accounts data.

5.1. Household Income Surveys

The household survey data we use comes from the World Bank's Living Standards Measurement Study (LSMS). The LSMS surveys typically involve the collection of detailed data at the household (and individual) level on income, health, education, and other "outcome" measures; expenditure and consumption; labor allocation; asset ownership; and details on agricultural production, business operation, and other economic activities. The surveys undertaken in different countries do not always follow identical methodologies; nevertheless, substantial efforts have been made to allow for as much international comparability as possible, for example in the treatment of home production. In micro-development economics, data from these household surveys are generally seen as representing a high standard for data quality (Deaton (1997).)

We have ten developing countries for which we can measure value added by sector using household data. These are Armenia (1996), Bulgaria (2003), Cote d'Ivoire (1988), Guatemala (2000), Ghana (1998), Kyrgyz Republic (1998), Pakistan (2001), Panama (2003), South Africa (1993) and Tajikistan (2009). Appendix 1.1 provides more detail about each of the surveys. While small, our set of countries features a variety of geographic locales with four countries from Asia, two from the Americas, one from Europe, and three from Africa. It also features a wide variety of income levels, with three countries below $2,000 PPP income per capita (Cote d'Ivoire, Ghana and Tajikistan), two between $2,000 and $5,000 (Kyrgyz Republic and Pakistan), two between $5,000 and $10,000 (Armenia and Guatemala), and three slightly above $10,000 (Bulgaria, South Africa and Panama.)

5.2. Measuring Value Added from Household Income Surveys

We construct value added in agriculture using the household survey data as follows. Letting i index a household, we define value added in agriculture to be:

$$VA_a = \sum_i y_{a,i}^{SE} + \sum_i y_{a,i}^{L} + \sum_i y_{a,i}^{K}, \tag{6}$$

where $y_{a,i}^{SE}$, $y_{a,i}^{L}$ and $y_{a,i}^{K}$ represent self-employed agricultural income, agricultural labor income, and agricultural capital income of household i. Self-employed agricultural income of household i is defined as:

$$y_{a,i}^{SE} = \sum_{j=1}^{J} p_j \left(x_{i,j}^{home} + x_{i,j}^{market} + x_{i,j}^{invest} \right) - COSTS_{a,i}, \tag{7}$$

where j indexes all agriculture goods in the economy, p_j is the farm-gate price of good j, and the three $x_{i,j}$ terms are the quantities of good j used for home consumption, market sales, and investment. In most cases households with agricultural production report $x_{i,j}^{home}$, $x_{i,j}^{market}$ and $x_{i,j}^{invest}$ for each crop j in kilograms, and report p_j for all crops for which some sales were made. For other crops the surveys report a local or regional average price. $COSTS_{a,i}$ is the cost of intermediate goods purchased, plus hired labor and rented capital (and land) used for production. Conceptually, $y_{a,i}^{SE}$ represents the value of all output produced by i net of any costs.

Agricultural labor income, $y_{a,i}^{L}$, is defined to be all income paid in currency or in kind for labor services rendered by any member of the household in the agriculture sector. Wage income is measured at the individual level and then aggregated to the household level. Agricultural capital income, $y_{a,i}^{K}$, is defined to be all income earned in currency or in kind for rental of housing, land or equipment, plus interest payments. Capital income is measured directly at the household level. Since it is virtually impossible to assign capital income to a particular sector, we assume that all capital income earned by agricultural households is agricultural, and all capital income earned by non-agricultural households is non-agricultural. We in turn classify households as being either agricultural or non-agricultural based on which sector the majority of the household's workers are employed, and in the event of ties, which sector the majority of self-employment plus wage income comes from.

Value added in the non-agricultural sector is defined as:

$$VA_n = \sum_i y_{n,i}^{SE} + \sum_i y_{n,i}^{L} + \sum_i y_{n,i}^{K}, \tag{8}$$

where $y_{n,i}^{SE}$, $y_{n,i}^{L}$ and $y_{n,i}^{K}$ represent self-employed non-agricultural income, non-agriculture employment income, and non-agricultural capital income of household i. Self-employed non-agriculture income is defined as:

$$y_{n,i}^{SE} = REV_{n,i} - COSTS_{n,i} \tag{9}$$

where $REV_{n,i}$ is self-reported revenues in non-agricultural businesses owned by household i, and $COSTS_{n,i}$ is any intermediate or factor cost incurred by these non-agricultural business. Non-agriculture labor income, $y_{n,i}^{L}$, and non-agriculture capital income $y_{n,i}^{K}$ is defined as above, only for the non-agricultural sector. For households with non-agricultural income, revenues and input costs are reported directly in all countries for which we have data.

Conceptually, our value added measures represent the total value of all payments made to factors of production applied to the production of output in each sector. Labor and capital income are unambiguous payment for labor and capital services used for production. The terms $y_{a,i}^{SE}$ and $y_{n,i}^{SE}$ represent payments made to entrepreneurs in the two sectors, and capture a mix of labor and capital income (see Gollin (2002).)

For each country, we compute value added by sector as described above, and then compute agriculture's share of total value added. We then compute agriculture's employment share by classifying each workers by her primary industry of employment. Workers are defined to be all economically active adults aged 15 or older. Using these two shares for each country we can construct the ratio of value added per worker in non-agriculture to agriculture, which is essentially the "micro" analog of our raw APG measures.

Our calculations of value added from these micro data have several advantages relative to national accounts measures. First, they unambiguously include home production of agriculture. Second, they include "informal" income sources, such as small-scale self-employment or informal wage employment, which may not be included completely in national accounts data. Finally, they focus only on domestic households, and exclude (for example) large multinational national resource firms who may contribute a lot to domestic value added in non-agriculture without much effect on the income of domestic residents.

Our value added measures also have several limitations. First, given the relatively small sample sizes (usually several thousand), the surveys are unlikely to capture the income of the very highest income earners in the economy, who may be business owners or simply those with high wage income. Second, as is always true with surveys of income, self-employment income could be under-reported. We worry particularly that non-agricultural self-employment income is under-reported, since non-agricultural business owners are typically asked directly to report their revenues (unlike in agriculture, where farm owners report physical quantities of output, crop by crop.) To the extent that this is the case, our non-agricultural value added measures may be biased downward, and hence so may our APG estimates. Finally, we were only able to make these calculations for ten countries, and each country's estimates are based on income

Table 4: Micro and Macro Data and Agricultural Productivity Gaps

Country	Agriculture Share of Employment	Agriculture Share of Value Added		APG	
	Micro	Macro	Micro	Macro	Micro
Armenia (1996)	34.2	36.8	32.8	0.9	1.1
Bulgaria (2003)	14.1	11.7	18.4	1.2	0.7
Cote d'Ivoire (1988)	74.3	32.0	42.1	4.7	4.0
Guatemala (2000)	40.2	15.1	18.7	3.8	2.9
Ghana (1998)	53.9	36.0	33.3	2.2	2.3
Kyrgyz Republic (1998)	56.9	39.5	39.3	2.0	2.0
Pakistan (2001)	46.9	25.8	22.6	2.5	3.0
Panama (2003)	27.0	7.8	11.8	4.4	2.7
South Africa (1993)	11.0	5.0	7.0	2.3	1.7
Tajikistan (2009)	41.0	24.7	30.1	2.1	1.6

Note: "Micro" means calculated using LSMS household survey data. "Macro" means calculated using national accounts data. APGs are calculated using the shares of value added from micro and macro data, and the shares of employment from micro data.

data for only a few thousand people, which limits the generality of our findings.

5.3. Results: APGs from Household Income Surveys

Table 4 shows the results of the calculations of value added by sector for our set of ten developing countries. The first data column shows the share of workers in agriculture according to the micro survey data.[22] The second and third data columns show agriculture's share in value added according to the macro data (the national accounts) and the micro data.

There are several points to take away from Table 4. First, the shares of value added in agriculture are fairly similar in both the macro and micro data with no apparent bias in either direction. For example, for some countries such as Cote d'Ivoire, Tajikistan, and Bulgaria the micro value added shares are larger than the macro shares. Whereas countries such as Armenia and Ghana have micro value added shares lower then the macro shares.

Second—and most revealing—the micro employment shares (except for Bulgaria) are all larger

[22] We do not report the "macro" employment shares in agriculture for two reasons. First, there is no conceptual difference between how we compute employment shares by sector and how sector employment shares are computed in aggregate statistics. Second, aggregate statistics on employment shares by are not available in most of the years of our surveys.

then the micro value added shares, resulting in APGs greater than one. The final column summarizes this result by showing the micro APGs. The average micro APG is 2.2, with countries such as Cote d'Ivoire, Pakistan, Guatemala, Panama having the largest APGs. For comparison, the second to the last column reports the macro APGs, which use the macro value added shares and the micro employment shares. The average APG from the macro data is 2.6, and for the most part the same countries having the largest gaps in the micro data are those with the largest gaps in the macro data. Thus, both sets of data suggest large APGs, albeit with somewhat smaller gaps computed from the micro data.

We conclude that, in spite of the differences in data and methodology between our calculations and those of the national accounts, the two measures provide surprisingly similar estimates of the size of the APGs in these developing countries. While counties may differ in the size of the employment and value added shares of agriculture, there are no countries for which micro and macro sources paint a substantially different picture of agriculture's share in aggregate value added.[23] Thus, at least for these ten developing countries, substantial gaps in value added per worker by sector appear prominently in household survey data, and the magnitude of the gaps are as large or only modestly smaller than those found in the national accounts data.

6. Other Explanations

Thus far, we have argued that agricultural productivity gaps are still large even after adjusting for improved measures of labor input, and that measures of the shares of value added by sector are similar in household surveys as in national accounts data. In this section, we discuss several additional explanations for these residual agriculture productivity gaps.[24]

6.1. Household Income and Expenditure by Sector

The theory of Section 2 assumes that workers supply labor to one particular sector, and are indifferent between work in the two sectors. In reality, decisions are often made at the household level, and households often diversify income across different types of economic activity.

[23]One potential explanation for the similarity of the micro and macro numbers is that the underlying data sources are in fact the same or similar. In Tanzania, for example, the value added of agriculture based largely on an extensive survey of rural households called the *Agricultural Sample Census*, combined with a second nationally representative household survey called the *Household Budget Survey.* Unfortunately, we do not know how much national statistical agencies in developing countries more generally base their value added estimates on household surveys.

[24]To be sure, a number of other measurement issues remain. One is that the non-agricultural sector includes a number of industries—such as government services—in which output is valued at the cost of inputs and in which labor markets may not be fully competitive. If these sectors receive inflated wages, it will be misleadingly reflected in the data as high productivity. A second measurement problem is that the costs of living for rural semi-subsistence farmers may be overstated by price indices based on local market prices. Many households in poor countries may in fact face very low prices for a range of home-produced goods, so that their realized utility levels are higher than would be suggested by income and expenditure data.

Thus, it could be that households primarily involved in agriculture earn combined incomes from agricultural and non-agricultural activities which are, on average, equal to the total incomes of non-agricultural households. To address this question, we use our LSMS data to ask whether a gap exists between the average income per worker of agricultural households and non-agricultural households.

As above, we define households as being either agricultural or non-agricultural based on where the majority of their workers are employed (and in a tie, which sector the majority of their self-employment plus wage income comes from.) We define income of a household i in sector s as

$$y_i = \sum_{j=a,n} y^{SE}_{j,i} + \sum_{j=a,n} y^{L}_{j,i} + y^{K}_{s,i}. \tag{10}$$

In other words, total income represents self-employed income from businesses in both sectors, wage income from both sectors, and capital income from the sector in which the household is classified. Note that for many households, at least one of the entries is zero. We define the total number of workers by household as the total number of economically active persons aged 15 or older.

In addition, we also measure the average expenditure per worker by sector. The rationale is that total household expenditure may provide a more accurate measure of income than direct measures of income, again due to under-reporting of self-employed income (Deaton (1997); Ravallion (2003)). Appendix 1.1 provides more detail about how expenditure is measured in each survey. To construct the measures of income per worker and expenditure per worker by sector, we use the same LSMS described above in Section 5.1.

Table 5 presents the results, with the last two columns reporting the ratio of income and expenditure per worker for non-agricultural households relative to agricultural households. For convenience, we also reproduce our measures of the APGs using the micro approach from Table 4. The results show that, for the most part, the gaps in income per worker are similar to the gaps in value added per worker. The average gap in income per worker is 2.1 relative to a 2.2 gap in value added per worker. The relative rankings are very similar as well. The countries with the largest APGs also have the largest gaps in income per worker.

Expenditures per worker data paint a similar picture—most countries exhibit gaps in expenditure per worker across sectors. The relative ranking across countries in expenditure gaps is similar to ranking of gaps in income per worker and micro-APGs. However, the magnitudes of the expenditure gaps are often lower. For example, the average expenditure per worker gap is 1.7 relative to 2.1 in income per worker and 2.2 in micro-APGs.[25]

[25]Part of the reason the expenditure gaps may be smaller than income gaps is the existence of certain types of insurance arrangements, which often involve transfers from richer households to poorer ones. One prominent example are remittances from non-agricultural workers (in the city, say) back to relatives in agricultural areas.

Table 5: Micro APGs and Ratios of Income and Expenditure Per Worker

Country	APG Micro	Income per Worker Ratio	Expenditure per Worker Ratio
Armenia (1996)	1.1	0.7	0.9
Bulgaria (2003)	0.7	1.4	1.2
Cote d'Ivoire (1988)	4.0	3.5	3.2
Guatemala (2000)	2.9	3.2	2.4
Ghana (1998)	2.3	2.0	1.9
Kyrgyz Republic (1998)	2.0	1.3	1.8
Pakistan (2001)	3.0	3.2	1.4
Panama (2003)	2.7	2.8	2.1
South Africa (1993)	1.7	1.7	1.2
Tajikistan (2009)	1.6	1.2	1.1

Note: APGs are calculated as Table 4. Ratios of income per worker are calculated as income per worker in agricultural households divided by income per worker in non-agricultural households. Households are classified agricultural if the majority of their workers report agriculture as their primary sector of employment, and in the event of equal numbers of workers in each sector, whether the majority of the household's income comes from agricultural activities. Ratios of expenditure per worker are calculated in the same way but using total household expenditure.

We conclude that household income per worker appears lower for agriculture households than non-agricultural households, with gaps similar in magnitude to those we observed in value added per worker in Section 5. Thus, at least for these ten countries, it appears unlikely that an explanation of the large residual APGs comes down to the distinction between agricultural workers and agricultural households. Put differently, income gaps are present between agriculture and non-agricultural households, not just between agriculture and non-agricultural workers.[26]

These types of transfers could account for the differences between the value added per worker, income per worker and expenditure per worker metrics. In Pakistan, for example, where the largest discrepancy between the gap in income and expenditure per worker, Ilahia and Jafareyc (1999) argue that remittances to rural agricultural households provide for a substantial amount of the consumption of rural agricultural families.

[26] Another candidate explanation is that that the gaps have arisen recently, and workers have simply not had sufficient time to reallocate across sectors in response. However, we find that for virtually all countries for which historical data is available from the WDI, the average APG in the period 1985 to the present is similar in magnitude to the average APG in the period 1960 to 1984.

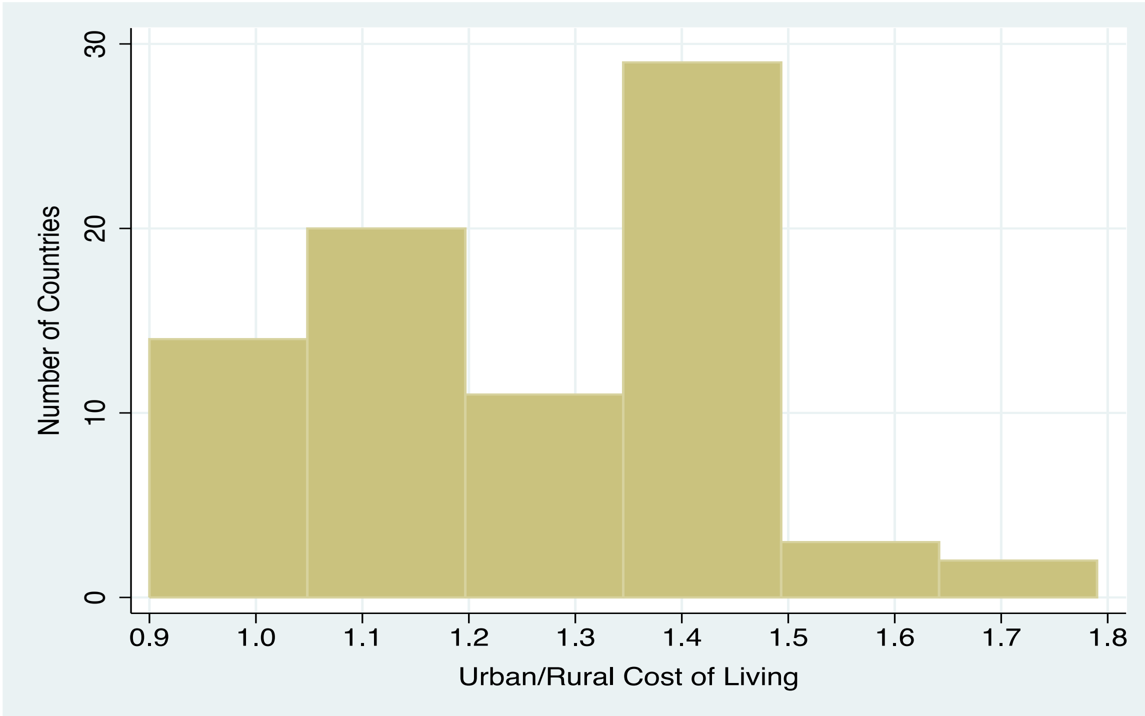


Figure 7: Cost-of-Living Ratio, Urban Areas/Rural Areas

6.2. Cost-of-Living Differences

Another potential explanation for the large residual productivity gaps is that there are differences in the cost of living for agricultural and non-agricultural workers. The prediction of model is that workers are allocated by sector in such a way that their real wage is equated across sectors. Thus far, we have not considered the possibility that the cost of living is lower for agriculture workers, the majority of whom live in rural areas. It could be that agricultural workers have lower value added per worker on average than nonagricultural workers because working in the non-agriculture sector involves a higher cost of the same basket of goods consumed.

To address this issue, we make use of proxies for the cost of living in rural and urban areas that are available for a large number of developing countries. Ravallion, Chen, and Sangraula (2009) use the World Bank's country studies from a set of 78 developing countries to compute the cost of the basket of goods consumed by households living on $1 per day in rural and urban areas. While this basket is not necessarily the same as the basket of the average household in the countries studied, Ravallion, Chen, and Sangraula (2009) argue that most poor households have a basket that is quite similar (i.e. mostly food), and hence a cost of living that is similar. For example, they found very similar urban-rural cost of living differentials when re-computing the cost of a basket consumed by households living on $2 per day.

Figure 7 shows a histogram of the ratio of cost of living in urban areas to rural areas. As can be seen in the figure, the urban cost of living is mostly larger than the rural cost of living. The average developing country has an urban cost of living that is roughly 1.3 times that of rural areas. Thus, while the costs of living do seem to differ between rural and urban areas, it does not appear that these differences are large enough to offset the large residual APGs, which are on average around two in our set of developing countries. Thus, it appears the residual APGs are not explained by differences in the cost of living by sector.

6.3. Sector Differences in Labor's Share in Production

Up to this point we have maintained the assumption that labor shares in production are the same in agriculture and non-agriculture. Could sector differences in labor shares could account for much of the remaining gap? To answer this question, consider a a Cobb-Douglas production function where the importance of labor and other inputs in production differs across sectors:

$$Y_a = A_a L_a^{\theta_a} K_a^{\phi_a} X_a^{1-\theta_a-\phi_a} \quad \text{and} \quad Y_n = A_n L_n^{\theta_n} K_n^{\phi_n} X_n^{1-\theta_n-\phi_n}. \tag{11}$$

One can show that the firms' first order conditions plus free labor mobility imply that sector differences in value added per worker are given by the ratio of the Cobb-Douglas factor shares for labor:

$$\frac{VA_n/L_n}{VA_a/L_a} = \frac{Y_n/L_n}{p_a Y_a/L_a} = \frac{\theta_a}{\theta_n}. \tag{12}$$

Equation (12) suggests that we could explain the remaining sectoral differences in average labor productivity if θ_n is approximately half as large as θ_a. Is labor's share in non-agriculture production half as large as in agriculture? Several pieces of evidence suggest that it is not. The first source is the direct estimates of agricultural labor shares computed using producer-level time series or cross sectional data. Fuglie (2010) provides a recent review of the estimates from around the world. His data imply that the average share of labor relative to land, equipment and structures is 0.58 for China, India, Indonesia, Brazil, Mexico, and sub-Saharan Africa, while the corresponding figures for the U.S. and U.K. are 0.51 and 0.52. In order to explain the residual APGs found in the paper, the non-agricultural labor share would have to be in the range of 0.25 to 0.29, which is highly implausible given that the aggregate labor shares are in the ballpark of two-thirds (see, e.g., Gollin (2002)).

The second piece of evidence against θ_n being substantially lower than θ_a comes from the relationship between aggregate labor shares and income per capita across countries. As Gollin (2002) points out, labor shares – once adjusted for the mixed income of the self-employed – vary relatively little across countries, and the variation is largely uncorrelated with income per capita. If this is the case, and if agriculture's share of GDP varies systematically with income per capita (as is widely understood), then labor shares cannot differ very much between agriculture and non-agriculture; otherwise, we would observe large and systematic variation in aggregate labor shares. We do not, which suggests θ_n and θ_a are similar in size.

The final piece of evidence comes from observations from share tenancy arrangements. In much of the world, large areas of agricultural land are farmed by operators under share tenancy arrangements, in which the operators pay a fraction of gross or net output to land owners in lieu of a cash rent (see, e.g., Otsuka (2007)). These arrangements are informative about cost shares in production since the operator typically provides all the labor, while the land owner provides the land and buildings. In principle, then, the split of gross output between the operator and the land owner, along with the allocation of capital costs and intermediate input costs, will allow for the calculation of the (net) share of labor in production. In practice, it may be difficult to arrive at precise calculations, because relationships between land owners and operators may be quite complicated (see, e.g., Jacoby and Mansuri (2009)). Nevertheless, the gross output share and the cost shares provide a useful – if crude – estimate of the factor shares.

A striking stylized fact in the share tenancy literature is that over time and across countries, most share contracts seem to involve 50-50 splits of both gross output and intermediate inputs. Otsuka (2007) refers to this as the "commonly observed rate," and Otsuka, Chuma, and Hayami (1992) note that "the output sharing rate is almost universally 50 percent under share tenancy in many developing countries."[27] Jacoby and Mansuri (2009) note that in survey data for rural Pakistan, in 1993 and 2001, "nearly three-quarters of share-tenants ... report a 50-50 output sharing rule." The 50-50 split is also common in modern-day agriculture in the United States (see Canjels (1998).) The predominance of the 50-50 split would tend to suggest that θ_a is around one-half. If this is true, than one would require θ_n of around 0.25 to explain the residual APGs, which is again highly implausible given that aggregate estimates are around two-thirds.

6.4. Selection

Yet another potential explanation for the large residual sector productivity gaps is that workers select themselves by sector in such a way that the average worker in the non-agricultural sector earns a higher average wage. Lagakos and Waugh (2011) formalize one version of this story where workers are heterogeneous in ability in each sector and chose where to supply their labor. In a parameterized version of their model, the average wage is higher in the non-agricultural sector in equilibrium, even without barriers to moving out of agriculture. The reason is that the underlying distribution of non-agriculture abilities is more dispersed than the distribution of agriculture ability, and workers with the highest endowments of non-agriculture ability disproportionately to enter the non-agricultural sector.

This type of selection is typically difficult to observe because workers self-select based on latent, unobserved abilities. However, two empirical proxies for the latent abilities modeled in Lagakos and Waugh (2011) are "strength" and "cognitive abilities," where strength is valuable

[27] They further note that the 50-50 split was historically pervasive in many parts of the world, to the extent that the French and Italian words for share tenancy (*metayage* and *mezzadria*, respectively) mean "splitting in half."

in agriculture work, and cognitive abilities are valuable in non-agricultural work (see e.g. Foster and Rosenzweig (1996) and Pitt, Rosenzweig, and Hassan (Forthcoming).) Given these proxies, one piece of evidence in favor of this type of selection story comes from a study by Miguel and Hamory (2009), who draw on a unique data set of cognitive ability scores from rural students in Kenya. The students were given the cognitive ability test in primary school and then later followed. The authors find that the scores are a very strong predictor of who later migrates out of agricultural areas to take non-agricultural employment. Their estimates are that individuals scoring one standard deviation higher on cognitive ability scores are roughly 17 percent more likely to migrate.

Schooling choice and migration patters suggests selection is at work as well. For example, the work of Beegle, De Weerdt, and Dercon (2011) show that students in Tanzania who attend more years of school are more likely to move out of agricultural work and into the urban non-agriculture sector. One interpretation of this finding is that those individuals with greater cognitive abilities were the ones selecting more years of schooling and, subsequently, work in the non-agricultural sector. Moreover, the current paper shows that there are systematic differences in schooling and human capital accumulation across sectors. While this evidence is consistent with certain patterns of selection, understanding better the sources of these differences is an important topic for future research.

7. Conclusion

According to national accounts data from developing countries, value added per worker is on average four times higher in the non-agricultural sector than in agriculture. This agricultural productivity gap, when taken at face value, suggests that labor is greatly misallocated in developing countries. In this paper we ask to what extent the gap is still present when better measures of sector labor inputs and value added are taken into consideration. To answer this question we construct a new data set for a large number of developing countries, with measures of hours worked and human capital per worker by sector, urban-rural cost-of-living differences, and alternative measures of value added per worker constructed from household income surveys.

We find that even after taking all these measurement issues into consideration, a puzzlingly large agricultural productivity gap remains. The value of output per worker in non-agriculture still appears to be nearly twice as high as in agriculture. We conclude that researchers interested in economic development may need a better understanding of which factors, beside average income, determine the allocation of workers by sector in the developing world.

References

BANERJEE, A., AND E. DUFLO (2005): "Growth Theory Through the Lens of Development Economics," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf.

BEEGLE, K., J. DE WEERDT, AND S. DERCON (2011): "Migration and Economic Mobility in Tanzania," *Review of Economics and Statistics*, 93, 1010–1033.

BRYAN, G., S. CHOWDHURY, AND M. MOBARAK (2011): "Seasonal Migration and Risk Aversion," Unpublished Manuscript, Yale University.

CANJELS, E. (1998): "Risk and Incentives in Sharecropping: Evidence from Modern U.S. Agriculture," The New School Center for Economic Policy Analysis Working Paper Number IV.

CASELLI, F. (2005): "Accounting for Cross-Country Income Differences," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf.

CASELLI, F., AND W. J. COLEMAN (2001): "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation," *Journal of Political Economy*, 109(3), 584–616.

CASELLI, F., AND J. FEYRER (2007): "The Marginal Product of Capital," *Quarterly Journal of Economics*, 122(2), 535–568.

CHANDA, A., AND C.-J. DALGAARD (2008): "Dual Economies and International Total Factor Productivity Differences:Channelling the Impact from Institutions, Trade and Geography," *Economica*, 75, 629–661.

DEATON, A. (1997): *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. The World Bank.

FAO (1996): *A System of Economic Accounts for Food and Agriculture*. Food and Agriculture Organization.

FOSTER, A. D., AND M. R. ROSENZWEIG (1996): "Comparative Advantage, Information and the Allocation of Workers to Tasks:Evidence from Agricultural Labor Markets," *Review of Economic Studies*, 63, 347–374.

FUGLIE, K. (2010): *The Shifting Patterns of Agricultural Production and Productivity Worldwide*chap. 4. Iowa State University.

GOLLIN, D. (2002): "Getting Income Shares Right," *Journal of Political Economy*, 110(2), 458–474.

GOLLIN, D., S. L. PARENTE, AND R. ROGERSON (2002): "The Role of Agriculture in Development," *American Economic Review Papers and Proceedings*, 92(2).

——— (2004): "Farm Work, Home Work and International Productivity Differences," *Review of Economic Dynamics*, 7.

——— (2007): "The Food Problem and the Evolution of International Income Levels," *Journal of Monetary Economics*, 54, 1230–1255.

HALL, R. E., AND C. I. JONES (1999): "Why Do Some Countries Produce So Much More Output per Worker than Others?," *Quarterly Journal of Economics*, 114(1), 83–116.

HERRENDORF, B., AND T. SCHOELLMAN (2011): "Why is Agricultural Productivity so Low in the United States?," Unpublished Manuscript, Arizona State University.

HSIEH, C.-T., AND P. J. KLENOW (2009): "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124, 1403–1448.

ILAHIA, N., AND S. JAFAREYC (1999): "Guestworker Migration, Remittances and the Extended Family: Evidence From Pakistan," *Journal of Development Economics*, 58(2), 285–512.

JACOBY, H., AND G. MANSURI (2009): "Incentives, Supervision, and Sharecropper Productivity," *Journal of Development Economics*, 88(2), 232–241.

KUZNETS, S. (1971): *Economic Growth of Nations: Total Output and Production Structure*. Harvard University Press.

LAGAKOS, D., B. MOLL, AND N. QIAN (2012): "Experience Matters: Human Capital and Development Accounting," Unpublished Manuscript, Princeton University.

LAGAKOS, D., AND M. WAUGH (2011): "Selection, Agriculture and Cross-Country Productivity Differences," Unpublished Manuscript, Arizona State University.

LEWIS, W. (1955): *The Theory of Economic Growth*. Irwin.

MCMILLAN, M. S., AND D. RODRIK (2011): "Globalization, Structural Change and Productivity Growth," NBER Working Paper No. 17143.

MIGUEL, E., AND J. HAMORY (2009): "Individual Ability and Selection into Migration in Kenya," Unpublished Manuscript, University of California, Berkeley.

MINNESOTA POPULATION CENTER (2011): "Integrated Public Use Microdata Series, International: Version 6.1," Minneapolis: University of Minnesota.

OTSUKA, K. (2007): "Efficiency and equity effects of land markets," *Handbook of Agricultural Economics*, 3, 2671–2703.

Otsuka, K., H. Chuma, and Y. Hayami (1992): "Land and Labor Contracts in Agrarian Economies: Theories and Facts," *Journal of Economic Literature*, 30(4), 1965–2018.

Pitt, M., M. R. Rosenzweig, and N. Hassan (Forthcoming): "Human Capital Investment and the Gender Division of Labor in a Brawn-Based Economy," *American Economic Review*.

Psacharopoulos, G., and H. A. Patrinos (2002): "Returns to Investment in Education: A Further Update," World Bank Policy Research Working Paper 2881.

Ravallion, M. (2003): "Measuring Aggregate Welfare in Developing Countries: How Well Do National Accounts and Surveys Agree?," *Review of Economics and Statistics*, 85(3), 645–652.

Ravallion, M., S. Chen, and P. Sangraula (2009): "Dollar a Day Revisited," *World Bank Economic Review*, 23(2), 163–184.

Restuccia, D., and R. Rogerson (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments," *Review of Economic Dynamics*, 11(4), 707–720.

Restuccia, D., D. T. Yang, and X. Zhu (2008): "Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis," *Journal of Monetary Economics*, 55, 234–250.

Rosenstein-Rodan, P. (1943): "Problems of Industrialisation of Eastern and South-Eastern Europe," *The Economic Journal*, 53(210), 202–211.

Rostow, W. (1960): *The Stages of Economic Growth: A Non-Communist Manifesto*. Cambridge, Cambridge University Press.

Schoellman, T. (2012): "Education Quality and Development Accounting," *Review of Economic Studies*, 79(1), 388–417.

Vollrath, D. (2009): "How Important are Dual Economy Effects for Aggregate Productivity?," *Journal of Development Economics*, 88(2), 325–334.

——— (2010): "Measuring Aggregate Agricultural Labor Effort in Dual Economies," Unpublished Manuscript, University of Houston.

Williams, J. H. (2005): "Cross-National Variations in Rural Mathematics Achievement: A Descriptive Overview," *Journal of Research in Rural Education*, 20(5).

Young, A. (2011): "The African Growth Miracle," Unpublished Manuscript: London School of Economics.

Zhang, Y. (2006): "Urban-Rural Literacy Gaps in Sub-Saharan Africa: The Roles of Socioeconomic Status and School Quality," *Comparative Education Review*, 50(4).

A. Data Appendix

In this section we provide more detail about the data sources used in the paper. Section 1.1 describes the Living Standards Measurement Surveys and our associated calculations, while 1.2 details the sources for our data on hours worked and years of schooling.

1.1. Living Standards Measurement Surveys

The Living Standards Measurement Surveys (LSMS) have been conducted in a number of developing countries over the last three decades. They are conducted in most cases by the World Bank and the national statistical agencies in the country in question. We make use of ten of these surveys, which are listed in Table 6 below, along with their sample sizes (the number of households surveyed.)

Table 6: LSMS Surveys

Country	Survey	Sample Size
Armenia	Household Budget Survey, 1996	4,920
Bulgaria	Multitopic Household Survey, 2003	3,023
Cote d'Ivoire	Cote d'Ivoire Living Standards Survey, 1988	1,600
Guatemala	National Survey of Living Standards, 2000	7,276
Ghana	Ghana Living Standards Survey, 1998	5,998
Kyrgyz Republic	Kyrgyz Poverty Monitoring Survey, 1998	2,934
Pakistan	Pakistan Integrated Household Survey, 1991	4,800
Panama	Survey of Living Standards of 2003	5,591
South Africa	South Africa Integrated Household Survey, 1993	8,811
Tajikistan	Living Standards Survey, 2009	1,500

Each of the surveys in Table 6 are *nationally representative* surveys of households, and employ a two-stage sampling methodology. In the first stage, a set of "primary sampling units," (PSU) or geographic regions, are chosen at random from an existing list. One can think of PSUs as "villages" in rural areas and "neighborhoods" in urban areas. In the second stage, a set of households are chosen at random from each PSU.[28] Every member in the sampled households are then interviewed. While survey content varies from country to country, all of them cover basic demographic information, educational attainment, employment status in the last week, month and/or year. For households employed in agricultural production of any kind, there is

[28]We use sample weights whenever they are available; several of the surveys are self-weighting, and do not require weights.

an agriculture module which contains detailed questions on agricultural outputs, such as the quantity of each crop grown, and inputs, such as fertilizer use or the cost of hired labor. For households with non-agricultural production, there is an additional module with questions about the type of business, revenues, and input costs.

For each country, we compute the share of employment in agriculture as the fraction of all economically active persons aged 15 or greater whose primary employment is in agriculture, hunting or fishing. We define economically active persons to be those who are either employed or seeking work at the time of the survey. Following the national accounts methodology, we do not use a minimum hours worked threshold. We compute the share of value added by sector using equations (6), (7), (8) and (9) . The prices used to value output are either the locally prevailing price paid to producers, or the households' reported price that it would receive, depending on the crop and survey. We define the input costs to be all expenditures on hired labor, rented equipment or land, repair services, purchased fertilizer or seeds, or any other expenditure listed on the survey questionnaire.

In each survey, some of the aggregation to the household level has been computed by the World Bank and/or the national statistical agency associated with the survey. We make use of these aggregated variables where available. For example, in all surveys the total value of agricultural output by household has already been constructed. In some surveys, labor income at the individual level has already been aggregated to include wage and in-kind payments plus bonuses at all jobs. In others, the input costs of non-agricultural or agricultural businesses have been aggregated to the household level. In all surveys, the World Bank or statistical agencies have imputed missing variables or re-coded outliers (or missing variables) where they deemed appropriate. We drop any additional households with missing income data, although typically this involves a very small number of records. We do not drop any outliers, although we found that our results are robust to dropping the top and bottom 1 percent of households by income or consumption.

Wage income, self-employment income and capital income are all measured before taxes. Wage income is measured at the individual level and then aggregated to the household level. It includes all payments in currency or in kind plus any bonuses received at primary, secondary or other jobs. Capital income is measured directly at the household level, and varies slightly from country to country in what is available. In most countries, capital income includes payments from rental of land, housing, or other property, plus interest payments. Most countries do not record dividends from shares of business owned but not controlled, as income from these sources is likely to be quite small in these countries.

We do *not* include transfer payments, gifts, inheritances or "other income" in any of our income calculations. The rationale is that while these income sources are relevant items in the budget constraints of individual households, they do not correspond to income from a national accounting perspective. Put differently, they do not correspond with production of new goods or services by the households in the survey.

Measures of household expenditure have been aggregated by broad expenditure category by the World Bank or relevant statistical agency in each of the surveys. We compute total household expenditure to be the sum of purchased food and non-food non-durable expenditures, plus service expenditures, plus the value of food or non-food items produced and consumed directly by the households. Some surveys also contain the imputed value of durables owned by the households, which we include if it is available. The yearly flow value of durables is imputed as the price paid for the durable good divided by an estimate of its usable life in years. In all surveys, these computations are made in the same way for agricultural and non-agricultural households.

All of the LSMS surveys are publicly available (often at no charge) from the World Bank website. For more on any individual survey, a basic information document is available for each country, as are the survey questionnaires themselves. Our household-level data (and our computer code used to construct it) is available upon request for each country where our user agreement permits us to share it.

1.2. Hours and Schooling by Sector

Table 7 below details the sources of data on schooling and hours used for each country in our data set. If not otherwise indicated, the samples include only economically active individuals aged 15 or greater, schooling is computed as the number of years of schooling completed, hours worked means the total number of hours worked in the reference week of the sample, and workers are classified by industry using the industry of primary employment if employed, the primary industry of previous employment if unemployed, or rural-urban status if unemployed and previous industry is not available. Whenever urban-rural status was used in place of agricultural non-agricultural status this is marked in the final column of the table. All other deviations are indicated using footnotes.

Our final country-level data set, including the raw and adjusted agricultural productivity gaps, average hours worked by sector, average schooling by sector, and human capital by sector is available upon request.

Table 7: Schooling and Hours Source Information by Country

Country	Variable	Year	Source	Ag/Non-Ag or Urban/Rural
Albania	Schooling[1]	2005	Multiple Indicator Cluster Survey (EPDC)	U
	Hours[5]	1998	Household Living Conditions Survey (ILO)	A
Argentina	Schooling	2001	Census of Population and Housing (IPUMS)	A
	Hours	1995	Encuesta Permanente de Hogares (ILO)	A
Armenia	Schooling[1]	2001	Population and Housing Census (IPUMS)	A
	Hours	2008	Report on Labour Force and Informality	A
Azerbaijan	Schooling[1]	2006	Demographic and Health Survey (EPDC)	U
Bangladesh	Schooling[1]	2007	Demographic and Health Survey (EPDC)	U
	Hours	1989	Labour Force Survey (ILO)	A
Belarus	Schooling	1999	Population Census (IPUMS)	A[3]
Belize	Schooling[1]	2006	Multiple Indicator Cluster Survey (EPDC)	U
Bhutan	Schooling[1]	2005	Population and Housing Census	U
	Hours[4]	2007	Living Standard Survey	U
Bolivia	Schooling[1]	2001	Census of Housing and Population (IPUMS)	A
	Hours	2000	Mecovi Survey	U
Botswana	Schooling[1]	1996	Labour Force Survey	A
	Hours	1996	Labour Force Survey	A
Brazil	Schooling	2000	Demographic Census (IPUMS)	A
	Hours	2000	Demographic Census (IPUMS)	A
Burkina Faso	Schooling[1]	2006	Multiple Indicator Cluster Survey (EPDC)	U
Burundi	Schooling[16]	1998	Enquete Prioritaire	A
Cambodia	Schooling	1998	General Population Census (IPUMS)	A
	Hours[4]	2001	Labour Force Survey	U

Cameroon	Schooling[1]	2006	Multiple Indicator Cluster Survey (EPDC)	U
CAR	Schooling[1]	2006	Multiple Indicator Cluster Survey (EPDC)	U
Chad	Schooling[1]	2004	Demographic and Health Survey (EPDC)	U
Chile	Schooling	2002	Population and Housing Census (IPUMS)	A
	Hours	2002	National Employment Survey (ILO)	A
China	Schooling[1]	1990	National Population Census (IPUMS)	A
Colombia	Schooling	2005	General Census (IPUMS)	A
Costa Rica	Schooling	2000	Population and Housing Census (IPUMS)	A
	Hours[3]	2000	Multi-Purpose Household Survey (ILO)	A
Cote d'Ivoire	Schooling	1988	Living Standards Measurement Survey	A
	Hours	1988	Living Standards Measurement Survey	A
Cuba	Schooling[1]	2002	Population and Dwelling Census (IPUMS)	A
Dominican Republic	Schooling[1]	2007	Demographic and Health Survey (EPDC)	U
	Hours	2007	Encuesta de Fuerza de Trabajo (ILO)	A
Ecuador	Schooling	2001	Census of Population and Dwelling (IPUMS)	A
	Hours	2001	Census of Population and Dwelling (IPUMS)	A
Egypt	Schooling[1]	2000	Demographic and Health Survey (EPDC)	U
El Salvador	Schooling[1]	2006	Encuesta de Hogares de Propositos Multiples	U
Ethiopia	Schooling[1]	2005	Demographic and Health Survey (EPDC)	U
	Hours	2005	Labour Force Survey	U
Fiji	Schooling[17]	1996	Census of Population and Housing	U
	Hours[4]	2005	Employment and Unemployment Survey	U
Gabon	Schooling[1]	2000	Demographic and Health Survey (EPDC)	U
The Gambia	Schooling[1]	2006	Multiple Indicator Cluster Survey (EPDC)	U

Georgia	Schooling[1]	2005	Multiple Indicator Cluster Survey (EPDC)	U
Ghana	Schooling	2000	Population and Housing Census (IPUMS)	A
	Hours	2000	Population and Housing Census (IPUMS)	A
Guatemala	Schooling	2010	National Survey of Employment and Income	A
	Hours	2000	Living Standards Measurement Survey	A
Guinea	Schooling[13]	1996	Census of Population and Housing (IPUMS)	A
Guyana	Schooling[1]	2005	Demographic and Health Survey (EPDC)	U
Honduras	Schooling[1]	2005	Demographic and Health Survey (EPDC)	U
India	Schooling[1]	2004	Socio-Economic Survey (IPUMS)	A
Indonesia	Schooling[1]	2007	Demographic and Health Survey (EPDC)	U
	Hours	2006	National Labour Force Survey (ILO)	A
Iran	Schooling[1]	2006	Census of Population and Housing (IPUMS)	A
Iraq	Schooling[1]	1997	Population Census (IPUMS)	A
	Hours	2007	Household Socio-Economic Survey (LSMS)	A
Jamaica	Schooling	2001	Population and Housing Census (IPUMS)	A
	Hours	2001	Population and Housing Census (IPUMS)	A
Jordan	Schooling[1]	2004	Population and Housing Census (IPUMS)	A
	Hours	2004	Population and Housing Census (IPUMS)	A
Kazakhstan	Schooling[1]	2006	Multiple Indicator Cluster Survey (EPDC)	U
Kenya	Schooling[13]	1999	Population and Housing Census (IPUMS)	A
	Hours	2006	Integrated Budget Household Survey	U
Kyrgyz Republic	Schooling[1]	1999	Population Census (IPUMS)	A

Lao PDR	Schooling[1]	2006	Multiple Indicator Cluster Survey (EPDC)	U
Lesotho	Schooling[1]	2004	Demographic and Health Survey (EPDC)	U
	Hours[4]	2008	Integrated Labour Force Survey	U
Liberia	Schooling[1]	2007	Demographic and Health Survey (EPDC)	U
	Hours[4]	2010	Labour Force Survey	U
Lithuania	Schooling[179]	2000	Population and Housing Census	U
Macedonia	Schooling[1]	2005	Multiple Indicator Cluster Survey (EPDC)	U
Madagascar	Schooling[1]	2008	Demographic and Health Survey (EPDC)	U
Malawi	Schooling	2008	Population and Housing Census (IPUMS)	A
	Hours	2004	Living Standards Measurement Survey	A
Malaysia	Schooling[1]	2000	Population and Housing Census (IPUMS)	A
	Hours	2007	Labour Force Survey (ILO)	A
Maldives	Schooling	2009	Demographic and Health Survey (EPDC)	U
Mali	Schooling	1998	Census of Population and Housing (IPUMS)	A
Marshall Islands	Schooling	1994	Multi-Subject Household Survey	A
Mauritius	Hours[5]	2009	Continuous Multi-Purpose Household Survey	A
Mexico	Schooling	2000	Population and Dwelling Count II (IPUMS)	A
	Hours	2000	Population and Dwelling Count II (IPUMS)	A
Moldova	Schooling[1]	2004	Demographic and Health Survey (EPDC)	U
Mongolia	Schooling[1]	2000	Population and Housing Census (IPUMS)	A
Morocco	Schooling[1]	2004	Demographic and Health Survey (EPDC)	U
Namibia	Schooling[1]	2006	Demographic and Health Survey (EPDC)	U
Nepal	Schooling[1]	2001	National Population Census (IPUMS)	A
	Hours[4]	2008	Labour Force Survey	A
Nicaragua	Schooling[1]	2001	Demographic and Health Survey (EPDC)	U

Nigeria	Schooling[1]	2008	Demographic and Health Survey (EPDC)	U
	Hours[4]	2009	Labour Force Survey	U
Pakistan	Schooling[17]	1998	Housing and Population Census (IPUMS)	U
	Hours[4]	2009	Labour Force Survey	U
Panama	Schooling[1]	2000	Census of Population and Housing (IPUMS)	A
	Hours	2001	Continuous Household Survey (ILO)	A
Papua New Guinea	Schooling[178]	2000	Census National Report	U
Paraguay	Schooling[17]	2002	Censo Nacional de Poblacion y Vivienda	U
Peru	Schooling	2007	Census of Housing and Population (IPUMS)	A
	Hours	2007	Estadisticas del Mercado de Trabajo	U
Philippines	Schooling[5]	1990	Census of Population and Housing (IPUMS)	A
	Hours[5]	1990	Labour Force Survey (ILO)	A
Romania	Schooling[1]	2002	Population and Housing Census (IPUMS)	A
	Hours[3]	2002	Population and Housing Census (IPUMS)	A
Rwanda	Schooling[1]	2002	Census of Population and Housing (IPUMS)	A
	Hours	2006	Integrated Living Conditions Survey	A
Saint Lucia	Schooling[1]	1991	Population and Housing Census (IPUMS)	A
	Hours	1991	Population and Housing Census (IPUMS)	A
Sao Tome and Principe	Schooling[1]	2009	Demographic and Health Survey (EPDC)	U
Senegal	Schooling[3]	2002	Census of Population and Housing (IPUMS)	A
Serbia	Schooling[1]	2006	Multiple Indicator Cluster Survey (EPDC)	U
Sierra Leone	Schooling[1]	2004	Population and Housing Census (IPUMS)	A

	Hours	1989	Labour Force Survey	A
South Africa	Schooling	2007	Community Survey (IPUMS)	A
	Hours	2009	Labour Market Dynamics in South Africa	A
Sri Lanka	Schooling[19]	2001	Census of Population and Housing	U
	Hours[4]	2009	Labour Force Survey	A
Sudan	Schooling[1]	2008	Population and Housing Census (IPUMS)	A
Suriname	Schooling[1]	2006	Multiple Indicator Cluster Survey (EPDC)	U
Swaziland	Schooling[1]	2006	Demographic and Health Survey (EPDC)	U
	Hours[5]	2008	Labour Force Survey	A
Syria	Schooling[1]	2006	Multiple Indicator Cluster Survey (EPDC)	U
	Hours	2010	Labour Force Survey	A
Tajikistan	Schooling[1]	2005	Multiple Indicator Cluster Survey (EPDC)	U
Tanzania	Schooling[1]	2002	Population and Housing Census (IPUMS)	A
	Hours	2009	Integrated Labour Force Survey	A
Thailand	Schooling[5]	2000	Population and Housing Census (IPUMS)	A
Tonga	Schooling[17]	2006	Census of Population and Housing	U
	Hours[4]	2003	Labour Force Survey	A
Turkey	Schooling[1]	2003	Demographic and Health Survey (EPDC)	U
	Hours	2003	Household Labour Force Survey (ILO)	A
Uganda	Schooling	2002	Population and Housing Census (IPUMS)	A
	Hours	2006	National Household Survey	A
Ukraine	Schooling[1]	2007	Demographic and Health Survey (EPDC)	U
Uzbekistan	Schooling[1]	2006	Multiple Indicator Cluster Survey (EPDC)	U
Venezuela	Schooling	2001	Population and Housing Census (IPUMS)	A
	Hours	2001	Population and Housing Census (IPUMS)	A

Vietnam	Schooling	1999	Population Census (IPUMS)	A
	Hours	1999	Labour Force Survey (ILO)	A
Yemen	Schooling[1]	2006	Multiple Indicator Cluster Survey (EPDC)	U
Zambia	Schooling[1]	2007	Demographic and Health Survey (EPDC)	U
	Hours	2005	Labour Force Survey	A
Zimbabwe	Schooling[1]	2006	Demographic and Health Survey (EPDC)	U
	Hours[5]	2009	Labour Force Survey	A

Note: Hours worked and years of schooling data are for all economically active persons aged 15+ unless otherwise noted.
IPUMS is the International Public-Use Microdata Series; EPDC is the Education Policy and Data Center; ILO is the International Labor Organization; LSMS are the Living Standards Measurement Surveys.
[1]Years of schooling imputed from educational attainment.
[2]Hours worked in main occupation.
[3]Agriculture status determined from occupation
[4]Computed from intervalled hours data.
[5]Sample consists of only employed persons.
[6]Sample consists of heads of households only.
[7]Sample includes economically inactive persons.
[8]Sample includes persons aged 5+.
[9]Sample includes persons aged 10+.

The International Growth Centre (IGC) aims to promote sustainable growth in developing countries by providing demand-led policy advice based on frontier research.

Find out more about our work on our website
www.theigc.org

For media or communications enquiries, please contact
mail@theigc.org

Subscribe to our newsletter and topic updates
www.theigc.org/newsletter

Follow us on Twitter
@the_igc

Contact us
International Growth Centre,
London School of Economic and Political Science,
Houghton Street,
London WC2A 2AE

IGC
International Growth Centre

DIRECTED BY

LSE

UNIVERSITY OF OXFORD

FUNDED BY

UKaid from the British people

Designed by soapbox.co.uk