February 10, 2013

# Measuring Aggregate Agricultural Labor Effort in Dual Economies

Dietrich Vollrath
University of Houston

Abstract

Wide differences in labor productivity are observed between agriculture and industry in most developing countries. Resear ch suggests that these differences – often denoted a "dual economy" effect – can explain a significant portion of low output per capita levels in these countries. A central input to the labor productivity calculation is the aggregate labor effort in the agricultural sector. Using findings from the Rural Income Generating Activity (RIGA) database, I reconsider the measure of labor productivity in agriculture and industry. I use several methods to extract information on labor effort and human capital from the household data in RIGA, and this is used to estimate the aggregate labor effort in the agricultural sector. With these new estimates, dual economy effects are found to be less severe for most of the RIGA countries. Using these estimates to adjust a wider sample of country-level data shows that the share of variation in output per capita explained by dual economy effects is around half of previous estimates.

JEL Codes: O12, O15, O47, O57, J31

Keywords: Dual economy, agricultural labor productivity, aggregate productivity

Contact information: 201C McElhinney Hall, U. of Houston, Houston, TX 77204, devollrath@uh.edu.

# 1 Introduction

A puzzling feature of many developing countries is the co-existence of a low labor productivity sector – often rural and agricultural – alongside a higher productivity "modern" sector. As described originally by Lewis (1954), this dual economy would appear to imply some kind of inefficient allocation of labor, and a variety of theories have tried to explain its appearance both in a static (Harris and Todaro, 1970; Stiglitz, 1974) and dynamic setting (Hayashi and Prescott, 2008; Caselli and Coleman, 2001; Mourmouras and Rangazas, 2007; Proto, 2007; Vollrath, 2009b).[1]

Current empirical work has attempted to gauge the aggregate impact of these inefficiencies (see Chanda and Dalgaard, 2008; Cordoba and Ripoll, 2009; Restuccia, Yang, and Zhu, 2008; Vollrath, 2009a), regardless of their actual origin.[2] The general finding is that the inefficiency of labor allocation between sectors can lead to output per capita 50% below potential in some developing countries. Across countries dual economy effects can explain up to 30% of variation in log output per capita.

To arrive at these numbers data on the current labor allocation between agriculture and industry is necessary. The preferred source for this information is the "Agricultural Population" data series provided through the FAO and based on information from the United Nations and International Labor Organization. This data is preferable to assuming the rural population is equal to the agricultural labor force, but it has potential issues as well.

In particular, the agricultural population data presume that an individual expends 100% of his or her time on agricultural work, and this runs contrary to survey evidence suggesting that individuals in developing countries typically diversify their time across many activities. Individuals who spend 51% of their time working in agriculture may be classified as agricultural workers, while those working only 49% of the time on agriculture may fall into the non-agricultural category.

To capture efficiency we would ideally like a measure of the actual labor *effort* exerted within each sector, regardless of the actual number of bodies engaged. The availability of new survey level data from Davis et al (2010) in the Rural Income Generating Activity (RIGA) database provides a first opportunity to make this kind of adjustment to the agricultural population data. This paper uses that data to make a more accurate estimate of agricultural labor effort within developing countries, and asks whether dual economy effects persist after incorporating these new estimates.

[1]See Temple (2005) for an overview of the issues involved in incorporating dual economy effects into standard models of growth and development.

[2]Work by Graham and Temple (2006) and Caselli (2005) finds smaller effects of labor allocations on aggregate outcomes.

RIGA provides data on the share of total income going to households by category, and divides activities into "agricultural" and "non-agricultural". By applying a very simple behavioral model of time allocation to households, we can back out the share of time spent on each type of activity from the share of income earned in each. Simply put, I assume that households optimize the allocation of their time, and therefore equate the marginal product of that time across different uses. This assumption implies that income shares and time shares are identical. The assumption that households optimize their time allocation still leaves open the possibility that there are inefficiencies *across* households that reflect dual economy effects.

With an estimate of the time allocation of households, I then aggregate up to the total labor effort engaged in agriculture by weighting each household by the amount of human capital it possesses. This amount is calculated directly from the RIGA data itself, with Mincerian regressions run for each country seprately to estimate the return to education (separately by gender). Given the return to education, I use the years of education in each household to calculate their total human capital and use that as the weight in the aggregation.

This adjusted series on agricultural labor effort shows that the implied differences in labor productivity between sectors are reduced. Naive comparisons of agricultural labor productivity based on simple counts of workers by sector show that it is roughly 20-33% of non-agricultural labor productivity. Using my estimates of the aggregate labor effort engaged in agriculture, I find that the agricultural productivity per unit of labor is between 60-70% of non-agricultural productivity for most countries, although it is important to note that for several countries relative agricultural productivity actually falls. Overall, it appears that agriculture is much closer to parity with non-agriculture once I have accounted for time allocations and human capital.

The adjustments I am making here are similar in spirit to those in Gollin, Lagakos, and Waugh (2012), who attempt to provide more accurate measurements of the human capital engaged in agriculture across a sample of countries. They have a wider sample of countries, as they rely on census-level information along with surveys to construct human capital stocks. The primary difference of this paper from theirs is that I employ the household-level data on the split of income between different sources to extract time allocations and human capital allocations between sectors, while they rely on the stark distinction of agricultural workers from non-agricultural workers employed in the census data they have. To the extent that households and individuals mix their time between sectors, my results will be more accurate. If individuals tend to specialize in activities, then our results will not be distinctly different. As I cannot measure individual-level time allocations, and have to infer them from household information, it is quite possible that I am attributing to individuals a split in their time between sectors that does not in fact exist. Regardless, both approaches provide better information regarding the true allocation of human capital between sectors, and hence better information on the size of dual economy effects.

After the initial analysis, the RIGA countries are then used to estimate a general adjustment factor that can be taken to a wider set of countries. Given the adjustment factor, I recalculate the efficiency effects of dual economies using my prior methodology in Vollrath (2000a). Compared to the prior results, dual economy inefficiencies only push output 25-30% below potential in the worst-afflicted developing countries. Across countries, the inefficiencies now account for about 12% of variation in log output per capita. This is still a sizable portion of the variation, but is only about half of the original estimates.

To proceed, I first define more precisely what I mean by agricultural labor effort, and how it differs from the size of the agricultural work-force. I then describe how I determine both the time allocations and human capital of households that are used in the calculation of agricultural labor effort. Following that I discuss the RIGA data in more detail and show the results for relative productivity using my new series for labor effort. The last section of the paper uses the results from the survey data in RIGA to make adjustments to a wider sample of countries and revisits the question of how large the aggregate effects of dual economies might be.

## 2 Agricultural Labor Effort and Duality

The main question surrounding the dual economy is the apparent difference in productivity between sectors. What we appear to see is that

$$\frac{Y_A}{L_A} < \frac{Y_N}{L_N} \tag{1}$$

where $Y_i$ is output, $L_i$ is the labor allocated to each sector, and $A, N$ refer to the agricultural and non-agricultural sectors.

One possibility is that while labor productivity differs, marginal productivity is equal across sectors and there is no efficiency loss. If we make the assumption typical in aggregate analysis that both sectors operate with Cobb-Douglas technologies, then one explanation is that labor's share in agriculture is simply much higher than in non-agriculture. In other words, (1) is an efficient outcome so long as

$$\alpha_A \frac{Y_A}{L_A} = \alpha_N \frac{Y_N}{L_N} \tag{2}$$

and $\alpha_A > \alpha_N$. As I have discussed previously (Vollrath, 2009), while theoretically possible this seems unlikely to explain the extremely wide variation in labor productivity we observe in the data. For example, to explain why agricultural labor productivity is only 12.5% of non-agricultural labor productivity in some Sub-Saharan African countries requires that $\alpha_A$ be 8 times larger than $\alpha_N$. If labor's share in agriculture is even 0.8, this implies labor's share in non-agriculture is only 0.1, which seems unlikely.

Thus there appears to be some inefficiency due to differences in labor productivity. However, this assumes that what we are calling "labor" is correct. On the margin efficiency requires that the return to a unit of labor effort is equal across sectors. What we would expect for efficiency is that

$$\frac{Y_A}{E_A} = \frac{Y_N}{E_N}, \tag{3}$$

where $E_A$ and $E_N$ measure the labor effort put into a sector. If $E_A \neq L_A$ and $E_N \neq L_N$ then it's quite possible for us to observe a difference in output per worker between sectors even though (3) holds.

The goal here is to provide a more precise measurement of $E_A$ and $E_N$ so that we can evaluate the possibility that output per unit of effective labor is equalized across sectors even though output per worker is not. I focus on two dimensions of labor effort: time and human capital.

Specifically, let labor effort in the agricultural sector be equal to

$$E_A = \sum_i t_{Ai} h_i \tag{4}$$

where $t_{Ai}$ is the fraction of their time that person $i$ spends working in the agricultural sector, and $h_i$ is their total human capital, and includes the total time they spend working. The labor effort that is put into agricultural work depends on the time allocations of all individuals as well as their level of human capital. Non-agricultural labor effort is determined symmetrically,

$$E_N = \sum_i (1 - t_{Ai}) h_i, \tag{5}$$

so that each individual is spending the remainder of their time working in non-agriculture.

To determine $E_A$ and $E_N$, then, we need information on time allocations of workers as well as a measure of their human capital. Using $E_A$ and $E_N$, we can then calculate output per effective unit of labor between sectors and see if in fact a dual economy exists in the countries under study.

## 2.1 Individual Time Allocation

To extract a measure of labor effort from the RIGA data requires some assumptions regarding how individuals allocate time. This is because we have information on household income by activity, but not their time allocation to those activities.

Households face two possible uses for their time endowment. Their income from the two uses is described by

$$y_{Ai} = w_{Ai} h_i t_{Ai} \tag{6}$$

$$y_{Ni} = w_{Ni} h_i (1 - t_{Ai}) \tag{7}$$

where the household has one unit of time to allocate between agricultural and non-agricultural work. Assuming that households care about maximizing their income, then the solution for a household depends on the relative size of $w_{Ai}$ and $w_{Ni}$. If $w_{Ai} > w_{Ni}$, then a household will specialize in agricultural production ($t_{Ai} = 1$), while if $w_{Ai} < w_{Ni}$, it will specialize in non-agricultural production ($t_{Ai} = 0$). For households where $w_{Ai} = w_{Ni}$, then they are indifferent between working in the two sectors.

Given data on $y_{Ai}$ and $y_{Ni}$ I can back out the actual time allocation made to each sector. If $y_{Ai} > 0$ and $y_{Ni} = 0$, then the household is clearly setting $t_{Ai} = 1$. On the other hand, if $y_{Ai} = 0$ and $y_{Ni} > 0$, then it must be that $t_{Ai} = 0$. If we observe that both $y_{Ai} > 0$ and $y_{Ni} > 0$, then the household is splitting its time between sectors. Given that we have assumed the only reason for them to split their time is because $w_{Ai} = w_{Ni}$, then it must be the case that

$$\frac{y_{Ai}}{y_{Ni}} = \frac{t_{Ai}}{1 - t_{Ai}}. \tag{8}$$

Hence we can back out the time allocations for any household given information on their earnings in the two different sectors.

This gives me the time allocations, but to get the aggregate measure of labor effort by sector I also require information on the human capital of households. To infer this I will use information from the RIGA database to run simple Mincerian regressions and back out the level of human capital from those. More specifically, for each household I have information on the number of working females aged 15-60 ($f_i$) and the number of working males aged 15-60 ($m_i$). I also know the average years of education ($e_i$) for all working age individuals in the household. Given their total income $y_i$, the Mincerian relationship I will estimate is given by

$$\ln y_i = \beta_0 + \beta_1(f_i e_i) + \beta_2(m_i e_i) + \epsilon_i. \tag{9}$$

The product $f_i e_i$ is the aggregate number of years of education for female workers, and $m_i e_i$ is the aggregate years for males. By separating them into two terms, I am letting the return to education for female and male workers differ.

Given this specification, the human capital of household $i$ is given by

$$h_i = exp(\hat{\beta}_1(f_i e_i) + \hat{\beta}_2(m_i e_i)). \tag{10}$$

This will be an admittedly crude measure of human capital, as I am constrained to only the average years of education in the household and I am assuming that each worker has the average amount. It also presumes that years of education are perfectly substitutable, which is typical of Mincerian regressions.

Given that $\epsilon_i$ is likely to be correlated with the obseved level of education and the gender breakdown of households, my estimates of $\hat{\beta}_1$ and $\hat{\beta}_2$ are most likely biased. However, if I instead appeal to typical measures of the return to years of education in the literature (e.g. $\beta_1 = \beta_2 = 0.10$), then the results I present are essentially unchanged.

Given the time allocation $t_{Ai}$ and the level of $h_i$, I can then aggregate up to find $E_A$ and $E_N$ within a country and use that to evaluate whether output per effective unit of labor is equalized across sectors.

# 3 Estimates of Time and Human Capital

I use data from the Rural Income Generating Activity (RIGA) database, available from the World Bank and described in some detail in Davis et al (2010). The authors pooled data from several of the Living Standards Measurement Studies at the World Bank to extract data on income, by activity, for households in the subject countries. The documentation of the RIGA database describes in detail the conversions and assumptions made to assign activities to different summary categories of income. An important note to make is that despite its name, the RIGA database contains information on both urban and rural households.

I have surveys from 14 different countries, and given multiple surveys from some countries I have a total of twenty country/years to evaluate. Table 1 reports several summary statistics for the country/years I have available. As can be seen the countries cover a broad geographic range, incorporating developing areas in Central and South America, Sub-Saharan Africa, and South Asia. The surveys generally have between

5,000 and 10,000 households, although Nigeria, Pakistan, and Vietnam has distinctly larger sample sizes. Education ranges from a low of 2.63 years per person in Ghana in 1992 to 7.68 years in Panama in 2003.

The average years of education for those of working age is, not surprisingly, higher than the overall average years of education. In all cases, though, none of the surveys report an average over ten years, although Indonesia and Panama are close at 9.9 and 9.2 years, respectively. Average household size is clustered around 4–5 individuals, with an outlier of Pakistan at 7.31 in 2001. Looking at households, the percentage reporting themselves as rural is generally over 50%, with exceptions of Bolivia, Nicaragua, and Panama, although even in these cases the percentage rural is well over 40%. As a comparison, I have also used the household-level data on number of individuals to calculate the total rural population relative to overall population. As can be seen in the final column this is generally very similar to the percent of households reporting themselves as rural. This implies very small differences in household size between rural and urban areas in these surveys.

RIGA breaks household income down into seven primary categories: own crop production, own livestock production, agricultural wage employment, non-farm wage employment, non-farm self employment, transfer income, and other income. To be consistent with my goal of studying labor allocations I study only income associated with labor effort. Practically, that means I exclude transfer income and other income from my calculations. Other income includes earnings from capital and land, while transfer income is as its name implies. Labor income is therefore defined as the sum of own crop and livestock production, agricultural wage employment, non-farm wage employment, and non-farm self employment. Land rents and other capital income are not included in the crop and livestock production values, so these reflect the outcome of labor effort of households. Of that total labor income, $y_{Ai}$ is defined as own crop and livestock production, and agricultural wage work, while $y_{Ni}$ is non-farm wage employment and non-farm self employment.

To get an idea of how income breaks down across households, note that from equation (8) that $t_A = y_{Ai}/(y_{Ai} + y_{Ni})$. Table 2 gives a distribution of households by $t_A$. In row one, for example, one can see that in Bangladesh 37% of households report zero labor income from agriculture, and hence zero time spent in the agricultural sector. 30% report some agricultural income, and for 34% of the households all of their labor income comes from the agricultural sector. The important thing to note in this table is that a very large fraction of households earn some income from agricultural work in every country. For most of the countries, roughly one-third of all households report earning some income from the agricultural sector. Their labor effort cannot be neatly defined as agricultural or non-agricultural, as it can for those with either 0% or 100% agricultural income.

The final column of table 2 reports the share of households that earn more than 75% of their income from agriculture, what Davis et al (2010) refer to as agricultural specialization. As one can see, these values are generally not much larger than the fraction reporting 100% agricultural income. This implies is that, in general, household that report $t_A$ between zero and one tend to have values closer to zero. Households that are not complete agricultural specialists tend to be concentrated in non-agricultural activities.

Table 3 shows summaries of the calculations of labor effort within each country/year I have data for. The first column is from the FAO, for comparison, and shows agricultural workers as a fraction of the total economically active population in a country. The next column is based on my calculations from the RIGA data for a country/year. For each household I backed out the value $t_{Ai}$ as described above. The table shows the mean value of $t_{Ai}$ for each country. As can be seen, the means tend to be below the $L_A/L$ fraction from the FAO, with the exception of the Central American countries Nicaragua and Panama.

On aveage, housholds are spending a smaller fraction of their time on agricultural work than would be indicated by a naive appeal to the FAO data. However, these mean values of $t_{Ai}$ do not give a full look at the distribution of the time allocations. The third column of data shows the median value of $t_{Ai}$ within each country/year. There is no definitive pattern here, but several interesting comparisons. In Bolivia, for example, the median household earns no income from agricultural work. Given that the mean value of $t_{Ai}$ is higher in Bolivia, it implies that for those households that do work in agriculture, their allocation of time is much higher than 0.22. There are several countries in which the median falls below the mean, indicating a concentration of agricultural labor effort in a small number of households.

On the other hand one can see countries like Ghana in 1998, Kenya, Malawi, Nigeria, Tanzania, and Vietnam, where the mean is lower than the median. In these countries the median household tends to be rather specialized in agricultural work, but the lower mean implies that the households that are not working heavily in agriculture are highly specialized in non-agriculture. In Malawi, for instance, the median family gets all of its income from agriculture, but the mean is only 0.73. Those families that are below the median must be far below the median to bring the mean down that far. It indicates that for these African nations (and Vietnam), there tends to be a bifurcation of labor effort across households into the two sectors.

For each country/year, I perform the Mincerian regression that I described above, and extract a fitted value of $h_i$ for each household. Using that and the information on a household's time allocation $t_{Ai}$ I am able to calculate the aggregate values $E_A$ and $E_N$. These represent the aggregate allocations of human capital to the two sectors. Column four of table 3 displays the fraction of total labor effort $(E_A + E_N)$ that is allocated to agriculture.

As can be seen, these fractions are almost invariably below the ratio $L_A/L$ based on workers. That is, in nearly every situation the allocation of labor effort to agriculture is less than implied by the number of people ennumerated as agricultural workers. In some cases the adjustment is dramatic, as in Nepal where $L_A/L = 0.93$ but $E_A/(E_A + E_N) = 0.48$, a factor of one-half. In many cases the implied allocations of labor effort are less than half of those found in the FAO's count of workers.

Several interesting exceptions do occur. In Nicaragua and Nigeria, the fraction $E_A/(E_A + E_N)$ is above that of $L_A/L$. This may indicate several things. First, it could be that $L_A/L$ is understated in these countries. Given its level of development, the fact that the FAO reports only 30% of labor engaged in agriculture in Nigeria seems like a distinct understatement. In Nicaragua as well, the fraction $L_A/L$ appears to be relatively low compared to nations such at Guatemala and Bolivia that have higher levels of GDP per capita.

Figures 1 and 2 plot the shares $L_A/L$ and $E_A/(E_A+E_N)$ against income per capita for each country/year. In both, one can see the distinct inverse relationship between agricultural labor effort and development levels. It is the poorest countries that have the greatest levels of labor effort in agriculture, an unsurprising finding.

If one examines the two figures, however, it becomes apparent that the relationship in figure 1 based on the number of workers is noisier in its relationship to output per capita. That is, there is a far greater spread in the fraction of labor in agriculture at any given level of income than we see in figure 2. The fractions of labor effort I calculated and used in that plot are more tightly related to development levels, and Nicaragua no longer appears to be an outlier. Nigeria switches from being a distinct outlier below the fitted relationship in 1 to being an outlier on the upside in figure 2. Regardless, using the implied share $E_A/(E_A + E_N)$ gives a much stronger fit than using the data on workers alone.

## 4 Dual Economy Effects

Given a better measure of agricultural labor effort, does this change any conclusions regarding the presence of a dual economy? To examine this I calculate the relative productivity of agriculture to non-agriculture under various assumptions regarding the labor effort exerted.

In table 4 one can see information regarding the GDP per capita and percent of value added coming from agriculture for each survey. Following that information are several columns of calculations reflecting the relative productivity of agricultural work. The first is the simple comparison of labor productivity, assuming that agricultural labor effort is equal to the number of agricultural workers reported in the FAOSTAT database. Specifically, I have calculated

$$\text{Relative Productivity } = \frac{Y_A/L_A}{Y_N/L_N}. \tag{11}$$

One can see the typical results in the literature here. Agricultural labor productivity is lower than non-agricultural productivity in nearly every case. It is as low as 4% in Nepal, and there are numerous cases of relative productivity falling below 20%. It is only in Nicaragua and Nigeria that relative labor productivity in agriculture appears to be equivalent to, or even exceed, non-agriculture. As noted previously, however, the numbers regarding agricultural labor force in these countries appear to be skewed and these likely do not reflect the reality of the situation. Ignoring those outliers for the moment, in the rest of the countries agricultural labor productivity is generally less than half that of non-agriculture, suggesting the possibility of large gains to reallocating labor away from agriculture.

In the next column I make a similar calculation, only now using the implied level of labor effort in agriculture based on the household surveys. In particular I calculate

$$\text{Relative Productivity } = \frac{Y_A/E_A}{Y_N/E_N}. \tag{12}$$

Having adjusted for the time that households spend on different activities, and weighting them by their human capital, the productivity differences between sectors now appear to be less severe. If one compares the two calculations, in every case excluding Nigeria and Nicaragua the relative productivity of agriculture is higher. In some cases the adjustments are quite striking. In Bolivia, for example, while labor productivity in agriculture is only 23% of non-agriculture, the relative productivity of labor effort in agriculture is 92% of non-agriculture. In many cases the relative productivity is now twice as much when measured using simple counts of workers.

The source of this difference is that while there are many individuals denoted agricultural workers in the FAOSTAT data, these individuals (a) are not spending their full allotment of time in agricultural work and (b) have lower human capital than non-agricultural workers. Hence the actual amount of labor effort exerted in agriculture is lower than one would estimate based only on counts of workers. The lower agricultural labor effort (and higher non-agricultural labor effort) raises the relative productivity of agriculture by a distinct amount.

Nigeria and Nicaragua were outliers in the calculations using simple labor counts, but note that relative productivity based on $E_A$ and $E_N$ leaves these countries with relative productivity numbers that look normal, so to speak, with respect to the rest of the table. It seems most likely that the official agricultural population statistics are incorrect, and that the RIGA data are giving us a more accurate picture of the labor effort engaged in agriculture in these two countries.

Figure 3 plots the relative productivity based on the labor effort $E_A$ and $E_N$ against GDP per capita. As can be seen, there is no distinct relationship between relative productivity and overall development. In particular, it does not appear to be the case that as countries get richer the productivitiy gap between sectors is closing. If anything, there may be a slight negative relatiionship, suggesting that the productivity gaps are worse in relatively developed countries.

The evidence shows that adjusting, however crudely, the labor effort engaged in agriculture leads to a less severe comparison of relative productivity between agriculture and non-agriculture. In particular, the especially striking differences observed in the very poorest countries (e.g. Nepal and Sub-Saharan Africa) are muted when a better accounting for agricultural labor effort is made.

# 5 Adjusting Aggregate Calculations

Given the evidence of the previous section, it suggests that the implied aggregate effects of dual economies may be less severe than normally thought. As mentioned, papers by Chanda and Dalgaard (2008), Cordoba and Ripoll (2009), and Vollrath (2009a) all found that the labor productivity differences between sectors could be important explanations for low aggregate productivity levels in developing countries. However, they take the labor force information from FAOSTAT as given, and in most cases do only crude adjustments for human capital differences across sectors.

With the current results I can update these calculations to see how much dual economy effects still matter for aggregate productivity. To do so I need some way of translating the adjustments made for the country/years in this paper into adjustments that can be made to countries without individual-level surveys. The issue is that in the countries without surveys I do not have decent information on the stock of human capital, but rather simply a count of the economically active population.

I assume for the purposes of this section that the total count of economically active workers is proportional to the total stock of human capital in an economy. This is a relatively common assumption made in most models of growth and development, which typically assume that total human capital is simply some per-worker amount of human capital times the number of workers. The survey information I employed to this point imply that this assumption is not correct. However, even for the countries that I have been studying, the sum of household-level human capital stocks is almost identical to what one would find by multiplying average household human capital by the number of households. Practically speaking, assuming that total human capital is proportional to the number of workers is not materially important.

From table 3 I know the proportion of labor effort that is exerted in agriculture for the twenty country/year surveys. Multiplying that proportion by the stock of economically active workers in the countries, taken from the FAO, gives me an estimate of the effective number of workers employed in agriculture. To be more clear, in Bangladesh the FAO reports a total economically active population of about 58 million in 2000. Of those, 55% are reported as agricultural workers by the FAO. However, based on my survey calculations, I know that only 37% of total labor effort in Bangladesh is exerted in agriculture. Therefore I estimate the effective number of agricultural workers in Bangladesh as $58 \times 0.37 = 21$ million.

I do a similar calculation for all twenty country/years that I have information for. In figure 4 I plot the adjusted number of agricultural workers against the reported number from the FAO. Not surprisingly, there is a strong positive relationship based merely on the fact that larger countries will have more workers. The fitted relationship in figure 4 has the estimates

$$\text{Adjusted Agric. Pop.} = \underset{(1440.0)}{1436.1} + \underset{(0.084)}{0.608} \times \text{ Agric. Pop.}, R^2 = 0.75. \tag{13}$$

The estimates are nearly identical even if Nigeria is excluded. Estimating the relationship in log-log form does not materially impact the results of the analysis to follow. What this simple regression indicates is that the effective agricultural workforce is only 60% as large as the reported number of agricultural workers. Based on this admitted crude relationship, I can infer an adjusted agricultural population for any country given the FAO reported size of the agricultural population.

Note that if a country has a particularly mis-stated number of agricultural workers, as in Nigeria, then this will deliver a mistaken estimate of the effetive workforce. The simple regression I have run does not provide a better way of measuring the number of agricultural workers. Rather, it is a way of adjusting a given agricultural workforce to reflect differences in time allocations and human capital.

With those caveats in mind, I can use the simple regression to obtain adjusted agricultural workforce data for other countries, and use those adjusted numbers to calculate the aggregate effect of dual economy effects. I use the sample of countries for which I calculated dual economy effects in Vollrath (2009a). This sample of 42 countries was used to show that the mis-allocation of labor between sectors could explain approximately 30% of variation in log output per capita. The methodology is explained in full in that paper, but the essential idea is to calculate the maximized level of output per capita in each country by equating the marginal product of labor between sectors. The variation across the sample in these maximized levels of log output per capita is then compared to the variation in log output per capita in the data. The ratio of these was found to be 0.68, meaning that 68% of the variation in log output per capita remained after eliminating labor mis-allocation. 32% of the variation was explained by the mis-allocation itself.

Previously I took the agricultural labor force (and non-agricultural labor force as a residual) from the FAO, and this was used in the accounting to back out the sector levels of TFP. I now update this to use the adjusted agricultural labor force from equation (13), and then the sector level TFP's are re-calculated. What will be happening is that by lowering the implied labor effort in agriculture, this will necessarily raise the agricultural TFP level backed out of the production function. Non-agricultural TFP will be lower given the higher labor effort in this sector. Given the higher agricultural TFP and lower non-agricultural TFP, the optimal allocation of labor effort will be skewed more towards agriculture than before and hence the gain from optimally allocating labor will be smaller. There is less room to increase income by reallocating labor effort across sectors when the labor productivity gap across sectors is smaller.

The new calculations show that the variation in the maximized log output per capita is now 88% of the variation in observed log output per capita. In other words, mis-allocation of labor effort between sectors only explains 12% of cross-country income variation, not 30% as I found before. This is still a significant piece of variation; consider as a comparison that education usually is thought to account for about 10-15% of cross-country income variation.

A more accurate accounting of the agricultural labor force lowers the implied inefficiency in dual economies. Further refinements to measures of agricultural labor effort could well reduce the impact of dual economy effects further. It's certainly possible that what appears to be an inefficient allocation of labor to agriculture - based on the labor productivity comparison between sectors - is a result of over-counting the labor effort of workers in that sector. A better estimate of agricultural human capital and time allocations reduced the dual economy effects appreciably, and further refinements to these calculations may shrink them further.

# 6 Conclusion

Dual economies appear to exist in many developing countries, and the impact of the inefficient allocation of labor across sectors could have significant aggregate implications. Empirically, the size of dual economy effects depends on the total agricultural labor effort relative to non-agricultural labor effort. To this point, however, much of the empirical work on this subject has used raw counts of workers by sector.

To get a better handle on the actual labor effort involved in agriculture, I have used data from the RIGA project, which contains information on the various income generating activities of households in 14 different developing nations, several across multiple years. With some simple behavioral assumptions the income data in RIGA can be used to back out time allocations to agriculture and non-agriculture. Combined with human capital estimates from Mincerian regressions, I construct a new series of agricultural labor effort.

The new series are used to re-calculate the discrepancy between agricultural and non-agricultural labor productivity. In nearly every case the gap between these two shrinks relative to a baseline that uses FAO data on agricultural population as the measure of labor effort. For these countries the inefficiencies of the dual economy are less pronounced, but do not disappear completely. For several countries, though, the implied inefficiency actually increases with the RIGA estimated labor effort data. In a cross-country setting, adjusting labor force data using RIGA information as a guide leads to a finding of less variation in output per capita being attributable to dual economy effects.

This paper is a preliminary study, and several outstanding issues remain to be addressed. Specifically, better information on the agricultural labor effort can be extracted from the full distribution of household data in the RIGA database. Depending on the distribution of activity across households, this could either raise or lower the implied inefficiency of dual economy effects. In addition, the adjustment factor applied to the cross-country data can be refined, in particular in light of the fact that the RIGA data do not indicate a universal relationship of agricultural labor effort to the size of the agricultural population reported by the FAO. A richer empirical relationship will be necessary to establish the aggregate effects of dual economies. Finally, a richer behavioral model for rural individuals that distinguishes between wage-labor and own-farm work will possibly provide further insight to the origins of dual economy effects as well as the possibility of *within*-household inefficiencies that currently are not measured.

# References


[1] Caselli, Francesco and John W. Coleman. 2001. "The U.S. Structural Transformation and Regional Convergence: A Re-interpretation." *Journal of Political Economy* 109(31):584-616.

[2] Caselli, Francesco. 2005. "Accounting for Cross-Country Income Differences," in *Handbook of Economic Growth*, ed. Phillipe Aghion and Steven Durlauf. North Holland, New York.

[3] Chanda, Areendam and Carl-Johan Dalgaard. 2008. "Dual Economies and International Total Factor Productivity Differences: Channelling the Impact from Institutions, Trade, and Geography." *Economica* 75:629-661.

[4] Cordoba, Juan Carlos and Marla Ripoll. 2009. "Agriculture and Aggregation" *Economics Letters*, 105(1):110-112.

[5] Davis, Benjamin, Paul Winters, Gero Carletto, Katia Covarrubias, Estaban Quinones, Alberto Zezza, Kostas Stamoulis, Carlo Azzarri and Stefania Diguiseppe. 2010. "A Cross-Country Comparison of Rural Income Generating Activities," *World Development*, 38(1):48-63.

[6] Gollin, Douglas, David Lagakos and Michael Waugh. 2012. "The Agricultural Productivity Gap in Developing Countries," working paper.

[7] Graham, Bryan and Jonathan Temple. 2006. "Rich Nations, Poor Nations: How Much can Multiple Equilibrium Explain?" *Journal of Economic Growth* 11(1):5-41.

[8] Harris, J. R. and Michael P. Todaro. 1970. "Migration, Unemployment and Development: A two-sector Analysis." *American Economic Review* 60:126-42.

[9] Lewis, W. Arthur. 1954. "Economic Development with Unlimited Supplies of Labor." *The Manchester School* 22:139-191.

[10] Mourmouras, Alex and Peter Rangazas. 2007. "Reconciling Kuznets and Habbakuk in a Unified Growth Theory." IUPUI Working Paper 2007-04.

[11] Proto, Eugenio. 2007. "Land and the Transition from a Dual to a Modern Economy," *Journal of Development Economics*, 83:88-108.

[12] Restuccia, Diego and Dennis Tao Yang and Xiaodong Zhu. 2008. "Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis." *Journal of Monetary Economics* 55:234-50.

[13] Stiglitz, Joseph E. 1974. "Alternative Theories of Wage Determination and Unemployment in LDCs: The Labour Turnover Model." *Quarterly Journal of Economics* 88:194-227.

[14] Temple, Jonathan. 2005. "Dual Economy Models: A Primer for Growth Economists." *The Manchester School* 73(4):435-478.

[15] Vollrath, Dietrich. 2009a. "How Important are Dual Economy Effects for Aggregate Productivity?" *Journal of Development Economics* 88:325-334.

[16] Vollrath, Dietrich. 2009b. "The Dual Economy in Long-run Development," *Journal of Economic Growth*, 14:287-312.

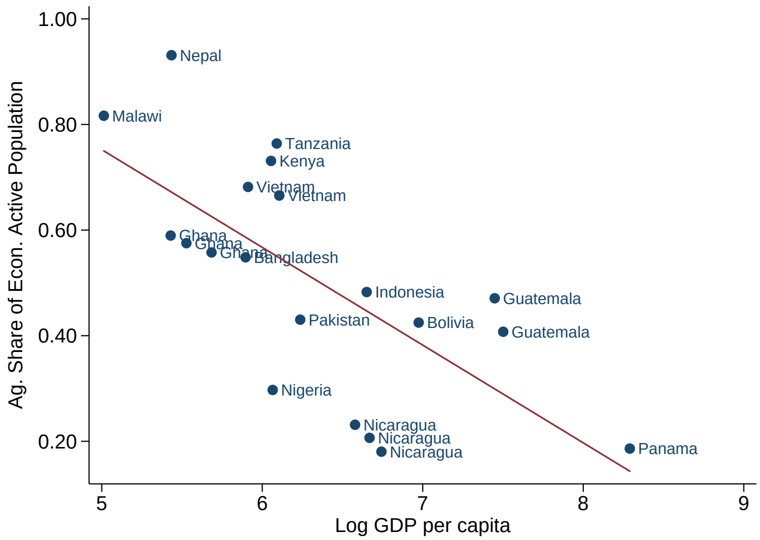


Figure 1: Agricultural Labor and Development Levels

Note: Agricultural labor share is from the FAOSTAT database. GDP per capita is from the World Development Indicators. The multiple observations for countries reflect the different survey years used.

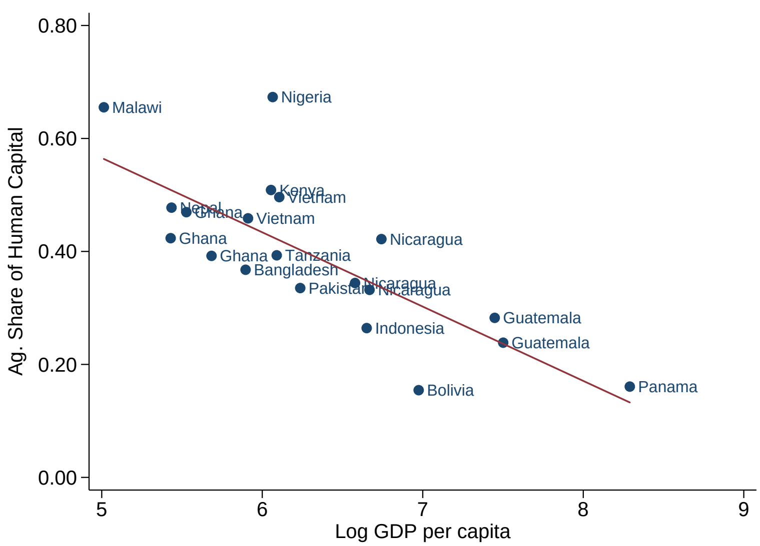


Figure 2: Agricultural Human Capital and Development Levels

Note: The fraction of human capital in agriculture is calculated as in the text. GDP per capita is from the World Development Indicators. The multiple observations for countries reflect the different survey years used.

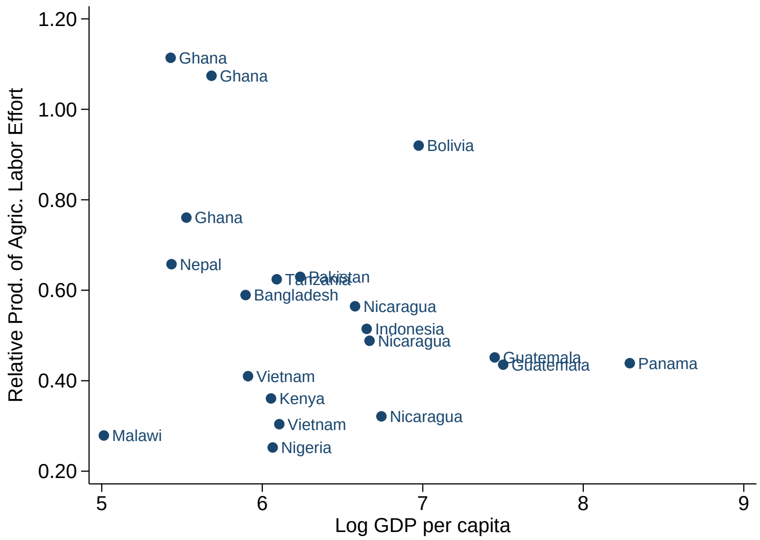


Figure 3: Relative Productivity and Development Levels

Note: The relative productivity of agriculture is from table 4. GDP per capita is from the World Development Indicators. The multiple observations for countries reflect the different survey years used.

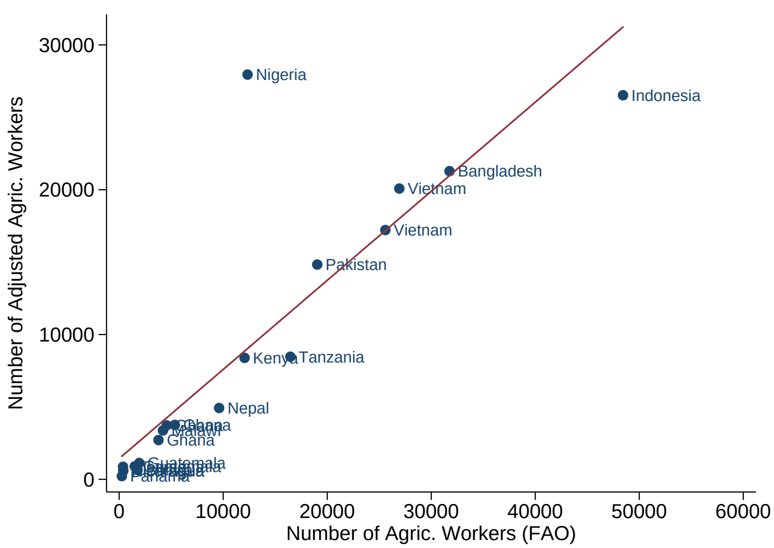


Figure 4: Adjusted Agricultural Work Force Compared to Reported Work Force

Note: The adjusted number of agricultural workers is the economically active population of a country multiplied by the fraction of labor effort exerted in agriculture from table 3. The number of agricultural workers from the FAO is from the FAOSTAT database for the same year as the survey. The line is fitted by OLS.

Table 1: Summary Data for RIGA Surveys

| Country | Year | Obs | Average Years Education | Average Years Educ. 15-60 | Average Hhold Size | Percent Hhold Rural | Percent Pop. Rural |
|---|---|---|---|---|---|---|---|
| Bangladesh | 2000 | 6,925 | 2.99 | 3.54 | 5.22 | 0.68 | 0.68 |
| Bolivia | 2005 | 3,645 | 6.52 | 8.12 | 4.27 | 0.44 | 0.44 |
| Ghana | 1992 | 3,714 | 2.63 | 3.71 | 4.44 | 0.66 | 0.68 |
| Ghana | 1998 | 4,877 | 3.51 | 4.91 | 4.34 | 0.65 | 0.69 |
| Ghana | 2005 | 7,197 | 3.92 | 5.24 | 4.30 | 0.61 | 0.66 |
| Guatemala | 2000 | 6,492 | 3.92 | 4.62 | 5.25 | 0.55 | 0.59 |
| Guatemala | 2006 | 12,659 | 3.67 | 4.33 | 5.14 | 0.58 | 0.62 |
| Indonesia | 2000 | 8,865 | 7.36 | 9.90 | 5.42 | 0.52 | 0.51 |
| Kenya | 2005 | 10,581 | 4.86 | 7.06 | 4.73 | 0.66 | 0.74 |
| Malawi | 2004 | 10,035 | 3.96 | 4.82 | 4.52 | 0.89 | 0.89 |
| Nepal | 2003 | 4,807 | 3.55 | 4.06 | 5.20 | 0.73 | 0.75 |
| Nicaragua | 1998 | 3,601 | 4.03 | 4.99 | 5.51 | 0.49 | 0.51 |
| Nicaragua | 2001 | 3,695 | 4.42 | 5.41 | 5.43 | 0.47 | 0.50 |
| Nicaragua | 2005 | 6,713 | 4.72 | 5.66 | 5.31 | 0.50 | 0.51 |
| Nigeria | 2004 | 16,471 | 3.90 | 4.80 | 4.96 | 0.79 | 0.80 |
| Pakistan | 2001 | 14,384 | 2.98 | 4.01 | 7.31 | 0.63 | 0.63 |
| Panama | 2003 | 5,285 | 7.68 | 9.18 | 4.31 | 0.47 | 0.50 |
| Tanzania | 2009 | 3,039 | 5.21 | 6.02 | 5.18 | 0.65 | 0.67 |
| Vietnam | 1998 | 5,780 | 5.60 | 7.46 | 4.74 | 0.72 | 0.73 |
| Vietnam | 2002 | 27,076 | 5.90 | 7.58 | 4.51 | 0.77 | 0.78 |

*Notes: Author's calculations using data from Davis et al (2010), the RIGA database. Observations refers to the number of households reporting that agricultural activities make up between 0 and 100 percent (inclusive) of their total labor income. Average years of education and average years of education for those 15-60 are means of the household level averages reported in RIGA. Household size is a count of all members of a household regardless of age. Percent of households rural is the fraction of households reporting rural status, regardless of their size. The percent population rural is based on the total number of individuals reported living in households defined as rural.*

Table 2: Agricultural Income Within Households

| Country | Year | Fraction of Households with $s_A$ of: $s_A = 0$ | $0 < s_A < 1$ | $s_A = 1$ | $s_A > 0.75$ |
|---|---|---|---|---|---|
| Bangladesh | 2000 | 0.37 | 0.30 | 0.34 | 0.35 |
| Bolivia | 2005 | 0.56 | 0.40 | 0.04 | 0.10 |
| Ghana | 1992 | 0.32 | 0.30 | 0.38 | 0.44 |
| Ghana | 1998 | 0.26 | 0.26 | 0.48 | 0.51 |
| Ghana | 2005 | 0.34 | 0.28 | 0.38 | 0.43 |
| Guatemala | 2000 | 0.35 | 0.33 | 0.32 | 0.36 |
| Guatemala | 2006 | 0.43 | 0.37 | 0.20 | 0.26 |
| Indonesia | 2000 | 0.53 | 0.23 | 0.24 | 0.27 |
| Kenya | 2005 | 0.27 | 0.24 | 0.49 | 0.52 |
| Malawi | 2004 | 0.08 | 0.35 | 0.57 | 0.66 |
| Nepal | 2003 | 0.14 | 0.46 | 0.40 | 0.45 |
| Nicaragua | 1998 | 0.23 | 0.41 | 0.36 | 0.40 |
| Nicaragua | 2001 | 0.19 | 0.49 | 0.31 | 0.36 |
| Nicaragua | 2005 | 0.20 | 0.52 | 0.28 | 0.37 |
| Nigeria | 2004 | 0.21 | 0.13 | 0.66 | 0.69 |
| Pakistan | 2001 | 0.45 | 0.27 | 0.29 | 0.32 |
| Panama | 2003 | 0.55 | 0.33 | 0.12 | 0.18 |
| Tanzania | 2009 | 0.26 | 0.34 | 0.41 | 0.48 |
| Vietnam | 1998 | 0.19 | 0.49 | 0.32 | 0.42 |
| Vietnam | 2002 | 0.21 | 0.39 | 0.39 | 0.44 |

*Notes: Author's calculations using data from Davis et al (2010), the RIGA database. Labor income is the total of own-farm production (crops and livestock), agricultural wages, non-agricultural wages, and non-agricultural self employment. The numbers reported are the raw fraction of households reporting agricultural income (agricultural wages and own-farm production) as a proportion of total income ($s_A$) in the range indicated in each column.*

Table 3: Agricultural Labor Effort Within Countrires

| | | Agricultural Effort Measured as: | | | |
|---|---|---|---|---|---|
| | | FAO | RIGA | RIGA | RIGA |
| Country | Year | $L_A/L$ | Mean $t_{Ai}$ | Median $t_{Ai}$ | $E_A/(E_A+E_N)$ |
| Bangladesh | 2000 | 0.55 | 0.42 | 0.17 | 0.37 |
| Bolivia | 2005 | 0.42 | 0.22 | 0.00 | 0.15 |
| Ghana | 1992 | 0.59 | 0.49 | 0.37 | 0.42 |
| Ghana | 1998 | 0.58 | 0.57 | 0.84 | 0.47 |
| Ghana | 2005 | 0.56 | 0.48 | 0.34 | 0.39 |
| Guatemala | 2000 | 0.47 | 0.43 | 0.20 | 0.28 |
| Guatemala | 2006 | 0.41 | 0.33 | 0.05 | 0.24 |
| Indonesia | 2000 | 0.48 | 0.33 | 0.00 | 0.26 |
| Kenya | 2005 | 0.73 | 0.56 | 0.92 | 0.51 |
| Malawi | 2004 | 0.82 | 0.73 | 1.00 | 0.66 |
| Nepal | 2003 | 0.93 | 0.58 | 0.63 | 0.48 |
| Nicaragua | 1998 | 0.23 | 0.46 | 0.28 | 0.34 |
| Nicaragua | 2001 | 0.21 | 0.43 | 0.19 | 0.33 |
| Nicaragua | 2005 | 0.18 | 0.43 | 0.22 | 0.42 |
| Nigeria | 2004 | 0.30 | 0.72 | 1.00 | 0.67 |
| Pakistan | 2001 | 0.43 | 0.38 | 0.06 | 0.34 |
| Panama | 2003 | 0.19 | 0.24 | 0.00 | 0.16 |
| Tanzania | 2009 | 0.76 | 0.54 | 0.63 | 0.39 |
| Vietnam | 1998 | 0.68 | 0.53 | 0.56 | 0.46 |
| Vietnam | 2002 | 0.67 | 0.55 | 0.58 | 0.50 |

*Notes: The first column is from the FAOSTAT database, and measures the number of economically active agricultural workers relative to the total economically active population. The remaining columns are author's calculations using data from Davis et al (2010), the RIGA database.* $t_{Ai}$ *is the implied share of time spent on agricultural work for household* $i$. $E_A/(E_A+E_N)$ *is the fraction of total labor effort exerted in the agricultural sector, calculated as described in the text.*

Table 4: Dual Economy Comparisons Under Different Labor Definitions

| Country | Year | GDP per capita ($) | % Agric. Value Added | Relative Productivity of Agric. Calculated Using: $(Y_A/L_A)/(Y_N/L_N)$ | $(Y_A/E_A)/(Y_N/E_N)$ |
|---|---|---|---|---|---|
| Bangladesh | 2000 | 364 | 0.26 | 0.28 | 0.59 |
| Bolivia | 2005 | 1069 | 0.14 | 0.23 | 0.92 |
| Ghana | 1992 | 228 | 0.45 | 0.57 | 1.11 |
| Ghana | 1998 | 251 | 0.40 | 0.50 | 0.76 |
| Ghana | 2005 | 294 | 0.41 | 0.55 | 1.07 |
| Guatemala | 2000 | 1717 | 0.15 | 0.20 | 0.45 |
| Guatemala | 2006 | 1810 | 0.12 | 0.20 | 0.44 |
| Indonesia | 2000 | 773 | 0.16 | 0.20 | 0.51 |
| Kenya | 2005 | 426 | 0.27 | 0.14 | 0.36 |
| Malawi | 2004 | 150 | 0.35 | 0.12 | 0.28 |
| Nepal | 2003 | 229 | 0.38 | 0.04 | 0.66 |
| Nicaragua | 1998 | 719 | 0.23 | 0.98 | 0.56 |
| Nicaragua | 2001 | 787 | 0.20 | 0.93 | 0.49 |
| Nicaragua | 2005 | 848 | 0.19 | 1.07 | 0.32 |
| Nigeria | 2004 | 431 | 0.34 | 1.23 | 0.25 |
| Pakistan | 2001 | 511 | 0.24 | 0.42 | 0.63 |
| Panama | 2003 | 3983 | 0.08 | 0.37 | 0.44 |
| Tanzania | 2009 | 442 | 0.29 | 0.12 | 0.62 |
| Vietnam | 1998 | 369 | 0.26 | 0.16 | 0.41 |
| Vietnam | 2002 | 449 | 0.23 | 0.15 | 0.30 |

*Notes: GDP per capita and agriculture's share of value added are from the World Development Indicators, for the year matching the survey. The three final columns are the author's calculations of the ratio of agricultural productivity to non-agricultural productivity using different measures of the labor effort engaged in agriculture. $L_A/L$ is calculatd using the agricultural labor share from the FAO. $E_A/E$ is the labor effort in agriculture calculated in this paper from the RIGA data. $E^*_A/E^*$ is the alternative labor effort calculation where the human capital of a household is proxied by its total income, rather than years of education.*