March 15, 2013

The Efficiency of Human Capital Allocations in Developing Countries

Dietrich Vollrath[†]

University of Houston

ABSTRACT

For a set of 14 developing countries I evaluate whether differences in the marginal product of human capital between sectors - estimated from individual-level wage data - have meaningful effects on aggregate productivity. Under the most generous assumptions regarding the homogeneity of human capital, my analysis shows that equalizing the marginal product of human capital between sectors leads to gains in output of less than 5% for most countries. These estimated gains of reallocation represent an upper bound as some of the observed differences in marginal products between sectors are due to unmeasured human capital. Under reasonable assumptions on the amount of unmeasured human capital the gains from reallocation fall well below 3%. Compared to similar estimates made using data from the U.S., developing countries would gain more from a reallocation of human capital, but the differences are too small to account for a meaningful portion of the gap in income per capita with the United States.

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

1 Introduction

A prominent feature of developing economies is the wide variation in labor productivity between different sectors. Gollin, Lagakos, and Waugh (2012) document that agricultural labor productivity is typically about one-half the level of non-agriculture, echoing the work of Arthur Lewis (1954) on dual economies. Moreover, this productivity gap appear appears to be most pronounced in the poorest countries, a fact noted by Kuznets (1971), and explored further by Gollin, Parente, and Rogerson (2002), Caselli (2005), and Restuccia, Yang, and Zhu (2008). Beyond just agriculture and non-agriculture, MacMillan and Rodrik (2011) document that large differences in labor productivity exist across ten broad sectors within each country they study.[1]

It is an open question whether these disparities in labor productivity reflect real differences in the marginal productivity of labor between sectors. If such differences exist, then aggregate productivity is lower than its potential, offering a partial explanation for low measured total factor productivity in developing countries. Several papers have attempted to estimate the loss in aggregate productivity from such misallocations. Focusing on only the distinction between agriculture and non-agriculture, Chanda and Dalgaard (2008), Vollrath (2009a), and Cordoba and Ripoll (2009) all suggest that there are substantial losses due to misallocations, while Caselli (2005) and Graham and Temple (2006) find much smaller effects.[2]

The existing literature infers differences in the marginal product of labor and/or human capital between sectors based on aggregate level information on output and labor inputs. Gollin, Lagakos, and Waugh (2012) are the most sophisticated in addressing the measurement issues arising from the use of this data. They employ census and survey data to adjust for differences in education and hours worked between sectors, as well as providing evidence that national accounts data provide a reasonable measure of agricultural value-added.[3] After their adjustments, they continue to find large gaps in human capital productivity between agriculture and non-agriculture. Even if these gaps represent real differences in the marginal product of human capital between sectors, it is not immediately clear whether the gaps imply any significant aggregate losses to productivity.

In this paper I analyze the existence and aggregate implications of marginal product gaps

---

[1]MacMillan and Rodrik's calculations also show that shifts of labor between sectors are an important part of economic growth, although the contribution is not necessarily positive. In Latin American and Sub-Saharan Africa sectoral shifts were actually a drag on growth in recent decades.

[2]These papers are part of a wider literature looking at variation in factor returns and the allocation of factors across different uses. Banarjee and Duflo (2005) discuss the wide variation in factor prices found within developing countries. Looking specifically at firms, Hsieh and Klenow (2009) find that misallocation of physical capital and labor across firms in India and China lowers output by around 30% relative to the United States. There is a growing literature on productivity differences across firms and their relationship with aggregate productivity. See Foster, Haltiwanger and Krizan (2001) and Syverson (2011) for overviews of this literature. Similarly, Restuccia and Rogerson (2008) find large effects of dispersion in firm's revenue productivity on aggregate productivity. Jones (2011) discusses how misallocations of factors of production can lead to lower productivity because of the input/output relationships between industries.

[3]Herrendorf and Schoellman (2011) find that measures of agricultural value-added in the United States are actually inaccurate given the accounting treatment of the income of farm proprietors.

between sectors by looking directly at wage data from a set of 14 developing countries.[4] The data are derived from the Living Standard Measurement Surveys organized by the World Bank and are collated by the Rural Income Generating Activity Database produced by Davis et. al. (2010). Despite the name the database contains information on urban workers as well as rural. By using wage data I avoid the disconnect between output and labor force data, and I get a more accurate estimate of the actual wage paid to human capital within a sector.

In the raw data there exists substantial variation between sectors in the wage per day within almost every country I study. Figure 1 shows the average wage in each sector relative to the average wage for each of the 15 surveys I use (there are two surveys for Nicaragua). The nine sectors shown for each country are a standard ISIC (revision 2) breakdown, with the Miscellaneous sector omitted. Agriculture (represented by the dark circles) tends to have wages well below the country average, dipping to only 50% of average wages in Ecuador and Nigeria. Sectors that have consistently high wages are Finance and business services (the open squares) and Utilities (the dark squares), with wages 1.5–2.5 times higher than average in each country.

Of course, much of the variation in wages between sectors in the countries in figure 1 reflects human capital differences, rather than differences in the wage paid per unit of human capital. I will explain the nature of the data and the precise estimation below, but figure 2 plots the residual wage per day in each sector after I have removed the influence of human capital using a simple Mincerian regression for each country that includes education, age, and occupation. As can be seen, there is much less variation in figure 2 within each country. Wages are less dispersed once I control for human capital differences between sectors.

Regardless, there is still some noticeable variation. In Ghana, for example, the wage paid to a unit of human capital in mining is 2.5 times the average across all workers in that country. In Nigeria and Ecuador the wage rate in agriculture is still roughly 50% of the average wage. Tajikistan exhibits several sectors with wages nearly 2 times the national average.

Do the wage differences in figures 1 and 2 imply a significant aggregate loss within developing economies? An answer to that question requires the calculation of the hypothetical wage level under an efficient allocation of human capital across sectors. For that calculation I start by making several assumptions designed to maximize that hypothetical wage level. First, I assume that all units of human capital within a country are perfect substitutes. This implies there is no loss of productivity for a unit of human capital when it shifts sectors. Secondly, I assume that there is no unmeasured human capital. This implies that the wage gaps seen in the figures reflect real differences in wages, rather than differences in unmeasured skills between sectors.

Despite these very strong assumptions, my calculations show that wages under the hypothetical efficient allocation would be less than 5% higher than observed wages for 11 of the 14 countries.

---

[4]The countries are Albania, Bangladesh, Bulgaria, Ecuador, Ghana, Guatemala, Indonesia, Malawi, Nepal, Nicaragua, Nigeria, Panama, Tajikistan, and Vietnam. The surveys from Ecuador, Nicaragua, and Vietnam are from the late 1990's, and the rest are from the 2000's.

For 13 of the 14, the hypothetical gain in wages is less than 11%, and the gain only reaches 15% in Tajikistan. For comparison purposes, the implied gain from reallocation across sectors in the United States, calculated using data from the Current Population Survey in 2000, is approximately 1.8%. While the U.S. appears to have a more efficient allocation, the difference between it and the developing countries is not terribly large. Moreover, the misallocation of human capital does not explain much of the gap in income per capita between these countries and the United States. As an example, income per capita in the U.S. in 2004 was about 22 times that in Nigeria. If human capital were allocated efficiently across sectors in Nigeria, that ratio would only fall to about 19.5. For other countries the explanatory power of misallocation across sectors is even smaller.

Furthermore, given the strong assumptions built into these calculations, the static gain from reallocating human capital across sectors is likely to be much lower than my estimates suggest. I expand on my original estimates by incorporating the possibility that the variation within countries in figure 2 partly reflects unmeasured human capital rather than differences in wages between sectors. If half of the observed wage differences are due to unmeasured human capital, then the hypothetical gains from reallocation fall to below 5% for every country, and the explanatory power of misallocation for cross-country income differences falls as well.

The RIGA data contain information only on wage-earners, and exclude the self-employed - such as farmers - who make up a large portion of the labor force in many developing countries. In an extension I show that under a variety of assumptions about the level of human capital and earnings of self-employed workers, the implied gains from reallocation remain quite low.

Compared to estimates made using aggregate-level data, I find very small potential gains from reallocating labor across sectors, and what gains I do find are still likely to be upwardly biased due to my assumption regarding the perfect substitution of human capital between sectors. From the individual level data it appears that the noticeable variation in wages between sectors does not have important effects on aggregate productivity.

The results suggest that labor markets are operating close to maximum efficiency in allocating human capital between sectors. Recent research, such as papers by Lagakos and Waugh (2010) or Vollrath (2009b), provides explicit theories for large gaps in labor productivity between sectors at the aggregate level despite the assumption of efficient labor markets, consistent with the evidence presented here.

To continue I first outline the method for calculating the hypothetical gain from reallocating human capital between sectors and how to use micro-level data to make that calculation. Following that I go over the data used in the estimations and report the variation in sector-level returns to human capital. With the estimates in hand I am able to calculate the potential gains from reallocation of labor across sectors and discuss the robustness of these results with respect to unmeasured human capital, self-employed workers, and parameter values.

2   The Gains from Reallocation

The earnings gaps observed in figures 1 and 2 suggest the possibility that human capital is misallocated and aggregate productivity is lower than it might otherwise be. In what follows I will describe my method for estimating the hypothetical gain to aggregate productivity from reallocating human capital efficiently. To make this estimate I will be making several aggressive assumptions that favor finding a large gain from reallocation. Hence my estimates should be seen as an upper bound for the gains, and as will be seen later even these upper bounds are empirically small.

To evaluate the static gains from reallocation, let M^* be the maximum level of aggregate earnings obtained under the optimal allocation of human capital across sectors. Let m_i be the observed earnings of individual i. I use R to denote the gains from reallocation and define it as follows:

$$R = \frac{M^*}{\sum_i m_i}. \tag{1}$$

If there is an inefficient allocation of human capital across sectors, then observed earnings will be lower than the hypothetical value M^*, and R will be greater than one. The larger is R, the greater the gains from reallocation. Note that R cannot be less than one, as M^* is by definition the maximum possible level of earnings.

M^* is aggregate earnings under a hypothetical alternative allocation of human capital. In order to calculate M^*, I need to put some structure on how earnings respond to the allocation of human capital across sectors. First, let earnings for individual i working in sector j be described by

$$m_{ij} = w_j h_{ij}, \tag{2}$$

where w_j is the wage paid per unit of human capital, and h_{ij} is the human capital of individual i employed in sector j. I assume that units of human capital within the sector are perfectly substitutable, and so earn an identical rate of return, w_j. I am ignoring issues with the allocation of human capital *within* sectors and only focus on the allocations *between* sectors.

The wage paid per unit of human capital is determined by

$$w_j = Z_j H_j^{-\alpha} \tag{3}$$

where Z_j is a sector-specific productivity term that is invariant to the stock of human capital, H_j, employed in sector j. α measures the elasticity of the wage with respect to the stock of human capital.[5] I will assume that each sector has the same value for α. Further, I will assume that the

---

[5]If one prefers to think of a production function, let $Y_j = (Z_j/(1 - \alpha))H_j^{1-\alpha}$, and then the above equation is simply the marginal product of human capital in sector j.

human capital of workers in sector j is perfectly substitutable, so that I can write

$$H_j = \sum_{i \in j} h_{ij}, \tag{4}$$

the sum over all individuals working in sector j. Finally, I assume that the human capital used in each sector is perfectly substitutable with the human capital in *other* sectors, so that aggregate human capital is

$$H = \sum_{j \in J} H_j, \tag{5}$$

where J is the set of sectors in the economy.

Under these assumptions, the income maximizing allocation of human capital requires that $w_j = w_k$ for any $j \neq k$. Solving across all J sectors, this leads to the following optimal allocation,

$$\frac{H_j^*}{H} = \frac{Z_j^{1/\alpha}}{\sum_{k \in J} Z_k^{1/\alpha}}, \tag{6}$$

where H_j^* is the optimal stock of human capital in sector j and H is the total stock of human capital available in the economy. Simply put, the economy should put more human capital to work in sectors that have high productivity levels, Z_j, relative to the aggregate productivity term.

Total earnings under the optimal allocation will be $M^* = w^* H$, where w^* is the wage rate at the optimal allocation. These earnings can be written as

$$M^* = \left(\sum_{j \in J} Z_j^{1/\alpha} \right)^\alpha H^{1-\alpha}. \tag{7}$$

Given this, the gains of reallocation can be calculated from equation (1).

The assumptions made regarding the perfect substitution of human capital within and between sectors ensure R will be as large as possible given some initial discrepancies in wages between sectors. The assumptions suggest that we can remove agricultural workers, for example, and productively put them to work in finance. This will lower the marginal product of human capital (and the wage) in finance, but raise the marginal product (and wage) in agriculture. Overall there will be an increase in earnings, though, as the marginal product was higher in finance to begin with. If we could not substitute agricultural labor freely into the finance sector, then the potential gain from reallocation would be lower. In the extreme, if agricultural human capital was useless in the financial sector, then the gain would from reallocation would be zero, M^* would be identical to observed earnings, and $R = 1$.

Even with perfect substitution of human capital, differences in w_j between sectors may not be indicative of misallocations between sectors if the wage paid to a unit of human capital is not equal

5

to its value marginal product. For example, the wage paid to human capital in finance may be quite high compared to agriculture, but if that wage reflects rents earned by finance workers, it is still possible that shifting a unit of labor from agricultural to finance will have no impact on aggregate output. Again in the spirit of making the greatest possible case for the role of misallocation, I am assuming that wages in sectors are in fact equal to value marginal products, so that observed differences in w_j reflect real inefficiencies in the allocation of human capital.

A final element influencing the size of R will be α, the elasticity of wages with respect to human capital. This dictates how fast the marginal product falls in the high-wage sector as human capital is increased, and how fast the marginal product rises in the low-wage sector as human capital decreases. In one extreme $\alpha = 0$ and the optimal allocation would be to move every unit of human capital into the highest-wage sector. In the example of Ghana, this would imply that all units of human capital go to work in the mining sector. This seems to be unrealistic, as even economies that we consider to have very efficient labor markets (the United States, for example) are not completely specialized. In the empirical calculations that follow I will consider several values for α in an attempt to see how relevant it is to the implied gains from reallocation.[6]

To proceed to a calculation of M^* and R, I require information on the individual Z_j values for sectors, as well as the aggregate human capital stock H. To acquire these I will need information on individual-level human capital, h_{ij}, as well as the wage paid in each sector, w_j. All these values will be estimated from the individual data available for each country in the RIGA dataset.

Let the log of human capital for individual i in sector j be written as

$$\ln h_{ij} = X_i'\beta + u_j + \epsilon_{ij} \tag{8}$$

where X_i' is a vector of individual characteristics and β is a vector translating those characteristics into units of human capital. u_j is the average unmeasured human capital that a worker in sector j has, and ϵ_{ij} is random variation in human capital unrelated to the other terms.

Given this specification for human capital, I can write individual earnings as

$$\ln m_{ij} = [\ln w_j + u_j] + X_i'\beta + \epsilon_{ij} \tag{9}$$

where I have collected the sector-specific terms in brackets. This equation is amenable to estimation, taking the form of a typical Mincerian wage regression. The actual specification I take to the micro-level data is

$$\ln m_{ij} = \delta_j + X_i'\beta + \epsilon_{ij}. \tag{10}$$

A constant is included in the vector X_i. Finally, I will exclude the J-th sector, so that implicitly

---

[6]The parameter α is intended to capture how the value marginal product of human capital changes with the stock of H_j in a sector. The value marginal product incorporates both the change in physical marginal product as well as the change in the relative price of output in sector j induced by a change in the stock of H_j.

the value of $\delta_J = 0$.

The issue that arises with this regression is that the estimated values of δ_j are picking up $\ln w_j + u_j$, not simply w_j. Practically, earnings in sector j could be high either because w_j is higher than in other sectors (which implies some possible misallocation) or because unmeasured human capital is higher than in other sectors (in which case there is not necessarily a misallocation).

I will consider the importance of this breakdown in a subsequent section. For now, I proceed with the same motive as before, to maximize the possibility of misallocation, and hence maximize the value of R. To do so I assume that $u_j = 0$ for every sector. In other words, any observed difference in earnings between sectors for identical individuals is assumed to be entirely due to differences in the wage rate paid to units of human capital.

The estimated log wage in sector j is therefore

$$\hat{w}_j = exp(\hat{\delta}_j). \tag{11}$$

Given the assumption regarding unmeasured human capital, the estimated value of human capital for an individual i in sector j is

$$\hat{h}_{ij} = X_i'\hat{\beta} + \hat{\epsilon}_{ij}. \tag{12}$$

With these individual human capital values, and the assumption that human capital is perfectly substitutable within sectors, the estimated stock of human capital in sector j is

$$\hat{H}_j = \sum_{i \in j} \hat{h}_{ij}. \tag{13}$$

Similarly, the total human capital stock in the economy is

$$\hat{H} = \sum_i^N \hat{h}_i \tag{14}$$

where N is the total sample of individuals.

Re-arranging equation (3), I can get an estimate of the level of Z_j for a sector as

$$\hat{Z}_j = \hat{w}_j \hat{H}_j^\alpha. \tag{15}$$

Finally, following (1) and (7) the value of R is calculated as

$$R = \frac{\left(\sum_j^J \hat{Z}_j^{1/\alpha}\right)^\alpha \hat{H}^{1-\alpha}}{\sum_i^N m_i} \tag{16}$$

where the estimated values in the numerator are as just described. The estimated gain from

reallocation is built off of underlying Mincerian regressions from a sample of individuals, as those will allow me to determine the stock of human capital and the implied productivity levels of different sectors.

As noted, this calculation is geared towards finding the largest value for R possible. By allowing labor to be fully substitutable I am maximizing the potential gain from reallocation. By assuming that there is no unmeasured human capital within sectors, I am attributing all of the observed differences in sector-level earnings to differences in the wage paid to units of human capital, and hence to differences in marginal products. Both assumptions act to inflate R.

Despite these assumptions, when I actually estimate the value of R for each of the countries available from the RIGA database, I will find relatively small values. Tightening these assumptions, as I will discuss following the initial results, only serves to deflate the values of R even further, implying that the aggregate gain from reallocations of human capital tend to be very small.

3   Estimating Gains from Reallocation

As discussed in the introduction, I will be using individual-level data from the RIGA database of Davis et al (2010). This database combines separate Living Standards Measurement Surveys (LSMS) done under the guidance of the World Bank, providing comparable measures of earnings and hours worked across different countries. For the current purpose, the most important feature of the RIGA database is the availability of earnings information and human capital characteristics that are comparable across individuals.

Here, I am specifically concerned with earnings in the form of wages, and the RIGA database separates labor income from other sources of income (rents, transfers, etc..). In addition, the database reports the wage, by job, for each individual if they are employed in more than one position. This ensures that the wages earned by an individual in different sectors can be distinguished from each other.[7] The wage data are for paid employment, and are reported net of taxes, so that I have a measure of the net return to labor for an individual. RIGA does not, however, provide an equivalent measure of individual self-employment income. I will therefore only be able to test for significant sector-level differences in wage work, and will not be able to offer any direct evidence on the efficiency of labor allocations between wage work and self-employment. Following the main results, though, I consider various assumptions regarding the earnings and human capital of self-employed workers and calculate how the value of R would change if they were included.

For each person engaged in wage work, the RIGA database provides a measure of daily wages. For many individuals, this is what they report directly, while for others this daily wage is imputed from data on monthly or weekly wage income as well as data on days worked in those same time

---

[7]While a very useful feature of the database, it turns out that there are very few individuals engaged in more than one wage-earning job.

periods. Daily wages are not adjusted for hours worked in a given day, as information on hours worked is not available for a large majority of the individuals. Hence there will remain some uncertainty in my results. In terms of the Mincerian regressions I am trying to estimate, this will create a bias in the sector-specific terms δ_j if hours worked per day differ systematically by sector. I may overestimate the variation in δ_j across sectors if there is significant variation in hours worked. As will be discussed below, however, correcting this bias would only reinforce the ultimate findings that sector level return differences have limited effects on aggregate productivity.

Most importantly, the RIGA database harmonizes the various industry classifications used across the LSMS surveys into standard ISIC codes. This provides 10 broad sectors: (1) Agriculture, Forestry, and Fishing, (2) Mining, (3) Manufacturing, (4) Utilities, (5) Construction, (6) Commerce, (7) Transport, Communications, and Storage, (8) Finance and Real Estate, (9) Services, and (10) Miscellaneous.

In terms of measuring the characteristics influencing human capital, X_i, the RIGA database provides a wide range of useful information. I have information on years of education and age, as well as gender. These provide a minimal level of control for human capital. Beyond those, though, RIGA provides a classification of occupation for each job held. Hence I can distinguish individuals by their type of work, independent of sector. This provides a level of control for human capital that has not been available before in macro-level studies that have, at best, been able to account for differences in education.

The different occupations reported are based on the International Labor Organization's (ILO) International Standard Classification of Occupations (ISCO). There are ten major categories, in addition to an Other/Unknown category: (1) Legislators, Senior Officials, and Managers, (2) Professionals, (3) Technicians and Associate Professionals, (4) Clerks, (5) Service Workers and Shop and Market Sales Workers, (6) Skilled Agricultural and Fishery Workers, (7) Craft and Related Trade Workers, (8) Plant and Machine Operators and Assemblers, (9) Elementary Occupations, and (10) Armed Forces Occupations.

In table 1 I've provided summary information on the 15 different surveys (from 14 different countries) used, as well as aggregate-level information about the countries. GDP per capita and the percent rural population are both from the World Development Indicators, with GDP measured at PPP in 2005 international dollars. Education is from Barro and Lee (2010), and represents years completed by those over 25. The surveys represent several ex-Communist countries (Albania, Bulgaria, and Tajikistan), as well as Asian developing nations (Bangladesh, Indonesia, Nepal, and Vietnam), Latin American (Ecuador, Guatemala, Nicaragua, and Panama), and Sub-Saharan Africa (Ghana, Malawi, and Nigeria). These countries range from middle income countries with per-capita GDP above $5,000 a year to below $1,000 as in Malawi and Nepal. Education and the percent rural correspond to these development levels in a manner that is consistent with broad development trends.

The last four columns show summary statistics from the individual country surveys. While there are as few as 1,631 observations (Ghana), most countries have around 5,000 observations to work with. The wage data is reported in local currency units, and so is not directly comparable across countries. However, I have calculated the coefficient of variation (CV) for daily wages in each sample. One can see relatively high variation in wages in Nigeria, Panama, and Albania. Compared to these countries, the Asian countries have relatively low variation, while Latin American variation is relatively large. One thing to note is that there doesn't appear to be a clear relationship between wage variation within the samples and overall development levels.

From the individual surveys, I've computed average years of education for workers, as well as the percent that report living in rural areas. Compared to the country level data, the RIGA surveys show relatively high education levels, and relatively low rural percentages. This is perhaps not surprising, as the surveys are restricted to wage workers, while the country level data covers the entire population. Regardless, there does not appear to be anything alarming about the comparison, and the individual level data from the RIGA database gives us a more accurate picture of those participating in the labor market.[8]

Table 2 reports the average wage of each sector in a country relative to the overall average wage, giving an initial indication of the variation in sector-level wages. As can be seen, aside from Albania and Bulgaria, agricultural wages tend to be around 2/3 of the mean wage. This ratio is similar across Malawi, Nigeria, and Ghana, despite their very different fraction of workers employed in that sector. Compared to agriculture, nearly every other sector has an average wage higher than the overall mean. Thus the main variation in wages tends to be between agriculture on one hand, and the rest of the economy on the other. This is perhaps not surprising given the existing literature on dual economy effects related to the rural/urban or agriculture/non-agriculture difference.

3.1 Estimation of Sector Returns to Labor

To proceed, recall that I will be using the following specification

$$\ln m_{ij} = \delta_j + X_i'\beta + u_{ij} \tag{17}$$

to obtain the estimates of $\hat{\delta}_j$ and $\hat{h}_{ij}$ that are used in the calculation of the gain from reallocation, R. As noted above, for now I am assuming that there is no difference in the level of unmeasured human capital across sectors, and so the estimates $\hat{\delta}_j$ represent the average wage of a unit of human capital in sector j relative to the excluded sector, which in all of my specifications will be the agricultural sector.

An important note is that the above specification does not include a country identifier. I am not estimating this equation for all 15 surveys at once, but rather for each country individually.

---

[8]The distribution of jobs across sectors in each country can be found in the appendix.

This allows the returns paid by a sector to be unique to each country, and not reflect some common sector-level effect on wages. There may be differences in sector-level wages that are common to all countries, perhaps reflecting the unmeasured component of human capital required for work in those sectors. By ignoring those, I will be overstating the variation in δ_j within countries, and as mentioned this will inflate the ultimate calculation of R.

For each country I estimate equation (17) four times, varying the contents of the X_i vector of individual characteristics. The four different specifications are as follows:

1. X_i is empty. This provides a raw estimate of sector-level wage per worker. It is this specification that provides the information in figure 1.

2. X_i contains education (in years), age, age-squared, and gender. These are the primary controls for human capital.

3. X_i contains occupation dummies, as well as the human capital controls in specification 2. Occupation will capture differences in human capital that may exist between workers who have identical education, age, or gender. It is a crude measure of skill outside of years of education.

4. X_i contains occupation dummies, and occupation-specific returns to education, as well as the human capital controls in specification 2. Here, I am allowing for the possibility that some occupations (e.g. Tradesmen) may see a lower return to formal schooling than others (e.g. Professionals and Technicians). This specification is what underlies figure 2 in the introduction.

Each specification is estimated using ordinary least squares, with standard errors calculated using the sampling weight provided by the original LSMS survey. This is simply an explicit way of accounting for heteroskedasticity. Results without weights, but with robust standard errors, provide nearly identical results. Each specification also excludes sectors that have fewer than 10 observations, as there are too few to provide useful estimates of sector dummies.

In each specification I obtain estimates $\hat{\delta}_1$ through $\hat{\delta}_{J-1}$, where J is the total number of sectors in a given country. I do not report the values of $\hat{\delta}_j$, as the specific values are not of primary interest. Recall that wages in a sector are estimated by $\hat{w}_j = exp(\hat{\delta}_j)$, so that variation in the values of $\hat{\delta}_j$ implies variation in the wage paid to a unit of human capital across sectors, and under the assumptions I've made this would suggest a misallocation of human capital.

As a first step consider the test with the null hypothesis $H_0 : \delta_1 = \delta_2 = ... = \delta_{J-1} = 0$. If the δ_j values are all identical this would imply an efficient allocation of human capital across sectors under the assumptions I've made regarding the substitution of human capital. If I reject the null hypothesis, then the sector-level wages (or at least some of them) are statistically different from one another. A rejection thus suggests that there is potentially some scope for finding effects of human

capital misallocation. Note that a rejection implies misallocation only under the assumptions made: perfect substitutability of human capital and the absence of unmeasured human capital. If those assumptions are not true, then rejecting H_0 doesn't necessarily imply anything about labor market efficiency.

The test statistic is distributed as an $F(J-1, N-J-1)$ distribution, where N is the number of observations and J is the number of sectors. For all countries, and under all specifications of the Mincerian regression, the null hypothesis is rejected. Practically, the values of δ_j vary enough between sectors to reject that their differences could be due simply to sampling variation in the RIGA data. The exact values of the F-statistic are reported in the appendix, as they are not themselves of any particular interest, and the p-values are all well below 0.001.

It's possible to add some additional insight into how sector returns to labor vary within countries, and to identify any patterns across countries in which sectors have particularly high or low returns. I convert the estimated dummies, $\hat{\delta}_j$, into comparable measures of sector returns following the work of Krueger and Summers (1989). $\hat{\delta}_j$ captures the wage of sector j relative to the excluded sector, Agriculture. However, it would be useful to look at the sector wage relative to the overall average wage.

Specifically, I calculate this alternative sector-level return, μ_j, as

$$\mu_j = \hat{\delta}_j - \sum_{i=1}^{J} \hat{\delta}_j s_j \tag{18}$$

where $\hat{\delta}_J = 0$ for the reference sector of Agriculture. The value s_j is the share of all workers in sector j. Given the μ_j terms, I then convert this into percentage terms for comparison purposes. To get the percentage difference in sector wages, I have to exponentiate μ_j, and given the sampling error that goes into calculating μ_j, the percentage difference in sector j from the average wage is

$$p_j = \exp\left(\mu_j - \frac{1}{2}\hat{\sigma}_j\right) \tag{19}$$

where $\hat{\sigma}_j$ is the standard error of the estimate of μ_j.[9]

Figure 1 from the introduction is a plot of the values of p_j for each country, using the estimates of δ_j from specification (1), which incorporates no individual controls. It thus represents the widest possible variation in sector-level wages. One can see that Bulgaria and Vietnam have particularly small variation across sectors in wages, while Tajikistan and Malawi appear to have the most. There are sectors with wages roughly 2.5 to 3 times the average wage, while most cluster in the range from 1–2.

If we incorporate the full individual controls of specification (4), then this reduces the variation

---

[9]The standard error of μ_j can be found using the Delta Method, as it is a linear combination of the original δ_j estimates.

in sector wages. This can be seen in figure 2 from the introduction. Compared to figure 1, the values of p_j are clearly more compressed, clustering around the value 1, indicating sector wages equivalent to the average wage in the economy. There are a few outliers, such as the Mining sector in Ghana, but for the most part the outlying sectors in figure 1 have collapsed towards the average wage. This reflects the fact that much of the "raw" difference in sector-level wages is driven by differences in human capital. Once we control for these, the implied sector-level variation declines.

As an alternative way of looking at the data, consider figure 3, which instead shows the different country values of p_j, grouped by industry. The figure is a box and whiskers plot, allowing one to see more clearly the amount of variation in sector-wages across the different country samples. In figure 3, the estimates of p_j are based on specification (1), with no individual characteristics controlled for. One can see that agricultural wages tend to be below average in each country. No other sector has a great tendency to lie below one, meaning they have above-average wages. All the true outliers (denoted by the dots in the figure) lie well above one, indicating relatively large sector wages.

Compare this to figure 4, in which the various sector p_j values are obtained from specification (4), which includes the full human capital and occupation controls. As can be seen, there is a great deal of compression, so that there is less variation across countries in the sector wages. Agriculture remains below one, in general, but not as severely. The commerce sector also has some tendency to be below one, but most other sectors retain wages that are above the average wage.

3.2 Calculating R

Given the estimates of $\hat{\delta}_j$ from the Mincerian regressions in a country, I follow the procedure outlined in section 2 to find $\hat{w}_j$, $\hat{H}_j$, and $\hat{Z}_j$. Using these values I calculate the implied gain from reallocation, R.

I make the initial calculations using an assumed value of $\alpha = 0.3$ for the elasticity of wages with respect to human capital. This would be the value that would hold if production in each sector were Cobb-Douglas and the share of human capital in output was 0.7. This value is consistent with the typical approach used in the existing literature based on macro-level data. In a subsequent section I will show how R changes when the assumed value of α is smaller.

The results are contained in table 3, in columns (1) through (4). Column (1) shows the calculated gain from reallocation, R, under the first specification of human capital, X_i. In this case, there are no human capital controls included in the regressions, which maximizes the variation of the $\hat{\delta}_j$ sector dummies, and therefore maximizes the possible gain from reallocation. One can see that with this limited specification Albanian earnings would rise by 6.4% if human capital were efficiently allocated. Tajikistan has a gain of roughly 34% available. Moving down the table, gains are as high as 22% in Ecuador and 26% in Malawi, although they are under ten percent for many of the countries.

Moving across columns, the more refined specifications for human capital are shown, and as

noted before these limit the variation in $\hat{\delta}_j$, and narrow the gains possible from reallocation. Within each country, the gains from reallocation shrink. By the time I have incorporated the full specification in column (4), the gains of reallocation have become negligible for nearly all the countries. The gain in Tajikistan has fallen to 15%, Ecuador's to 9.6%, and Malawi's to 2.6%.

For Bangladesh, Indonesia, and Panama, the gains from reallocation are under 1% after controlling for human capital fully. Gains are under 5% for most of the countries shown, save Tajikistan, Ecuador, Ghana, and Malawi. Even for these four, though, the gains are around a 10% gain to aggregate wages. More importantly, as noted in section 2, these gains are almost certainly upper bounds on reallocation gains. I am making the severe assumption that human capital is perfectly substitutable between sectors and that unmeasured human capital does not vary systematically across sectors. These assumptions allow for the greatest possible gain from reallocation across sectors. However, even with these assumptions in place the gains are generally only on the order of 5%, rising only to 15% for Tajikistan. Practically speaking, these gains are over-stated because units of human capital are not perfect substitutes and the underlying estimates of δ_j almost certainly capture unmeasured human capital as opposed to differences in wages between sectors.

These relatively small gains of reallocation are not due to assumptions limiting the movement of labor between sectors. As noted, I am allowing human capital to perfectly substitute across sectors, and this leads to rather large movements of human capital between sectors after reallocation. In Tajikistan, nearly 50% of the total human capital of the workforce must be removed from agriculture and reallocated to construction, commerce, and services. Across the Latin American countries, between 5–26 percent of total human capital must be shifted out of agriculture, generally to move into services, construction or utilities. For Albania, Bulgaria, Vietnam, and Ghana, the service sector would lose up to 19% of total human capital. In Ghana, every sector loses human capital save mining, where fully 35% of the human capital is implied to move in a reallocation. For most countries, about 20% of the entire stock of human capital must move sectors. Despite these large shifts, the implied earnings gain, as seen above, is not necessarily that large.[10]

3.3   Cross-Country Comparisons

Leaving aside the absolute size of the gains from reallocation, for inefficiency of human capital allocations to provide a useful explanation for cross-country differences in income per capita, it must be that poorer countries tend to have more inefficient markets. If one goes through table 3, it becomes apparent that there is no distinct relationship between GDP per capita and the estimated size of R. The poorest countries I study, Malawi and Nepal, have gains from reallocation of 2.6% and 5.1%, respectively. The Latin American countries, generally richer, have gains around 2.5% for most, and as large as 9.6% for Ecuador.

---

[10]The appendix contains a table showing the exact shifts in human capital between sectors necessary to reach an efficient allocation in the case where the full set of human capital controls are included.

To make a more concrete comparison, I have made a similar calculation of R for the United States. I use earnings data from the March Current Population Survey from the year 2000. This dataset contains over 161,000 observations on individual-level earnings, and includes not only those working for wages but self-employed workers as well, which I will return to in the following section. The CPS data has similar human capital controls available at the individual level, and I have recoded the reported industry and occupation classifications to match those used in RIGA. See the appendix for full details.

The CPS reports weekly earnings as well as hours worked per week. As I do not have hours worked in the RIGA database, nor do I have days worked per month (or week) in the CPS, I cannot create a stricly comparable measure of wages between the U.S. and the other countries. I report results in table 3 in the final two rows using both weekly earnings and hourly earnings for the U.S.

As can be seen, for weekly earnings the gain from reallocation across sectors is on the order of 5% when no human capital controls are included, but fall to 1.8% with the full controls in column (4). This gain from reallocation is smaller than that found in most countries, but is notable for still being larger than that found in Bulgaria, Bangladesh, Indonesia, and Panama. The implied gains in Albania, Guatemala, Nicaragua, and Malawi are not much larger than 1.8%, suggesting that their human capital is allocated as efficiently as in the United States.

When using hourly earnings, the implied gains from reallocation in the U.S. fall by about half, so that in column (4) with full human capital controls the gain is roughly 1%. This makes the gain smaller than in nearly every country save Indonesia, suggesting that the in general the U.S. has the most efficient allocation of human capital among the countries being compared.

Note that the adjustment to hourly earnings lowered the gains from reallocation substantially in the U.S., implying that weekly earnings differences between sectors may reflect differences in hours worked as opposed to differences in the wage rate paid to human capital. For the developing countries using the RIGA data, recall that I have only daily wages, and not hourly wages. The U.S. results suggest that there could be a meaningful reduction in R if I could examine hourly earnings, as differences in hours worked per day may vary across sectors.

Regardless, even if one simply compares these developing countries with the U.S., the implied reallocation gains pale in comparison to the actual income gaps. U.S. has GDP per capita on the order of 20–50 times higher than the countries in this paper. Gains from reallocation of even 15% are simply not capable of explaining a significant fraction of this gap.

3.4 Unmeasured Human Capital and Reallocation

Recall that the sector-specific intercepts, δ_j, estimated using the Mincerian regressions are actually capturing

$$\delta_j = \ln w_j + u_j. \tag{20}$$

15

In the prior section I simply assumed that there was no unmeasured human capital ($u_j = 0$) in each sector. This assumption is almost certainly false. Abowd, Kramarz, and Margolis (1999) find, using French data, that most of the observed inter-industry wage differentials they find are due to unmeasured individual heterogeneity, and Goux and Martin (1999) find a similar result. Using matched worker-firm data from the state of Washington, Abowd, Finer, and Kramarz (1999) attribute half of the inter-industry wage differentials they find to individual heterogeneity.

In terms of the specification used here, what these papers imply is that variation in δ_j between sectors is driven in part by variation in u_j, unmeasured human capital, as opposed to differences in the wage paid to units of human capital. While by definition I do not have the means to measure u_j, I can provide a crude means of understanding how ignoring it inflates my estimates of the gains from reallocaiton, R.

Assume that some fraction $\gamma \in (0,1)$ of the estimated value $\hat{\delta}_j$ is actually capturing the average unmeasured human capital in sector j. That is, let $\hat{u}_j = \gamma \hat{\delta}_j$. By construction, this means that $\ln \hat{w}_j = (1-\gamma)\hat{\delta}_j$. As I do not have any better information, I assume that this fraction is the same in each sector. Setting $\gamma = 0$ is implicitly what I have already estimated. As γ goes towards one, less of the estimated $\hat{\delta}_j$ represents differences in the wage paid to human capital, and more represents differences in unmeasured human capital. So as γ goes to one, the potential gains of reallocation shrink, as the observed differences in $\hat{\delta}_j$ simply represent our inability to accurately measure the human capital employed in each sector.

I recalculate values of R using varying values for γ to provide an idea of how important unmeasured human capital might be for the gains for reallocation. Specifically, for each individual I now estimate human capital as:

$$\ln \hat{h}_{ij} = \gamma \hat{\delta}_j + X_i' \hat{\beta} + \hat{u}_{ij} \tag{21}$$

which adds in the fraction γ of the estimated sector dummy δ_j to each individual in sector j. The estimate of Z_j is modified to be

$$\ln \hat{Z}_j = (1 - \gamma)\hat{\delta}_j + \alpha \ln \hat{H}_j, \tag{22}$$

where H_j is again the sum of individual human capital employed in sector j, only now that individual human capital is estimated using (21). Given these estimates, I can again calculated values for R, the gain from reallocation.

In figure 5 I've plotted the values of R for the four countries that had the largest gains from reallocation in my original estimation: Ecuador, Ghana, Nigeria, and Tajikistan. As can be seen, as γ rises, the value of R drops rather quickly for each. If γ is equal to 0.25, indicating that one-quarter of the sector-specific return $\hat{\delta}_j$ is actually unmeasured human capital, then the gain for Tajikistan is down to 10%, and for the others it is only around 6-7%. By the time $\gamma = 0.5$ all four countries have gains from reallocation of less than 5%. The work of Abowd, Finer, and Kramarz

(1999) attributed half of inter-industry wage differentials to unmeasured individual human capital, suggesting that $\gamma = 0.5$ is not an unreasonable possibility, and work using French workers and firms suggests γ may be even higher. Regardless, figure 5 shows that even a mild degree of unmeasured human capital makes the gain from reallocation even smaller than originally estimated.

Recall that the other countries all began with gains smaller than the four shown in figure 5, and these shrink demonstrably towards zero as γ rises. The evidence suggests that unmeasured human capital is likely an important source of the apparent sector-level differences in wages, and this means reallocation would have little effect on wages. It seems that wage labor is very close to efficiently allocated across sectors.

3.5 Accounting for Self-Employed Labor

One of the most significant shortcomings of the RIGA data I am using is that it only includes wage workers, and excludes the self-employed. In many developing countries, a large majority are working for their own account, mainly farmers. One concern is that while the gains from reallocation among those employed for wages are small, there could well be large gains to reallocation between the self-employed and those earning wages. In particular, a gap in earnings between self-employed farmers and the rest of the workforce may be indicative of a significant misallocation of human capital. Given the large gaps observed between agricultural and non-agricultural labor productivity in the aggregate data, this distinction is worth exploring further.

I attempt here to gauge the change in R from incorporating the self-employed into my calculations. To do this I consider the self-employed as an additional separate sector of the economy. Adding this self-employed sector into the formula for R,

$$R_{SE} = \frac{\left(\hat{Z}_{SE}^{1/\alpha} + \sum_{j \in J} \hat{Z}_j^{1/\alpha}\right)^{\alpha} \left(\hat{H}_{SE} + \hat{H}\right)^{1-\alpha}}{\hat{M}_{SE} + \sum_i^N m_i} \tag{23}$$

where $\hat{Z}_{SE}$ is the estimated productivity of the self-employed sector, $\hat{H}_{SE}$ is the human capital of the self-employed, and $\hat{M}_{SE}$ are their earnings.

For the human capital of the self-employed, I assume that each self-employed individual has h_{SE} equal the level of h for the average agricultural worker in the country. Aggregate human capital for the self-employed is then $\hat{H}_{SE} = N_{SE} h_{SE}$, where N_{SE} is the number of self-employed workers. To obtain this count I compare the fraction of agricultural workers in my sample (which only includes those working for wages) to the fraction of workers in agricultural reported by the FAO (which includes both wage workers and the self-employed). Letting N be the total number of workers in my existing sample, and N_{Ag} be the number of agricultural workers in my sample, I set N_{SE} such that the ratio $(N_{SE} + N_{Ag})/(N + N_{SE})$ is equal to the FAO's reported ratio. In several cases, the FAO reports a lower ratio of agricultural workers than I have in my sample, which would imply

that N_{SE} is negative. In these cases I presume that the RIGA sample has accurately represented the distribution of workers within the economy, and set $N_{SE} = 0$.

To calculate $\hat{Z}_{SE}$ and $\hat{M}_{SE}$ I require information on the wage paid to self-employed workers. In general, I set this wage to be proportional to the observed agricultural wage, $w_{SE} = \phi w_{Ag}$. In the estimates to follow, I will vary the value of ϕ between 0.75 and 1.25, indicating that self-employed workers earn as little as 75% or as much as 125% of the wage of equivalent agricultural workers. Given the wage, the earnings of self-employed workers are $\hat{M}_{SE} = w_{SE}\hat{H}_{SE}$, and the productivity of the self-employed sector is $\hat{Z}_{SE} = w_{SE}\hat{H}_{SE}^{\alpha}$, follwing the methods described in section 2.

With those pieces of information, I can calculate R_{SE} to see how the gains from reallocation are impacted by allowing for a potentially large group of self-employed workers. Table 3 shows the results in columns (5) through (7), which vary in their assumptions about ϕ, the wage of self-employed workers relative to agricultural workers.

As can be seen, while the values of R_{SE} are generally larger than the baseline results, the increase is not dramatic. Even if wages of self-employed workers are only 75% of that in agriculture, the gains only rise to around 2% for countries like Bangladesh and Indonesia, from a baseline of less than 1%. The most dramatic increase occurs for Nigeria, where with $\phi = 0.75$ the implied gain in wages from reallocation goes to about 21%, whereas it is only 11% when I ignore the self-employed. As the relative wage of self-employed workers rises, the implied gains in Nigeria fall. Recalling that these gains are certainly upper bounds given unmeasured human capital, the allowance for self-employed workers does not appear to materially change the conclusion of this paper that human capital reallocations are not capable of generating significant gains.

One thing to note is that the estimated gain from reallocation is not necessarily declining in ϕ; see the results for Albania, Bulgaria, Bangladesh, and Vietnam. In these countries, human capital is very close to efficiently allocated across sectors. Allowing for self-employed workers who earn more than the average agricultural worker introduces a larger gain from reallocation as we'd expect human capital to flow into self-employment. In the remaining countries, even when I set $\phi = 1.25$ the self-employed are still underpaid relative to the rest of the economy, and so labor would flow out of self-employment if reallocation occurred.

3.6 The Elasticity of Wages

As noted previously, the elasticity of wages with respect to human capital by sector is assumed to be $\alpha = 0.3$ in all the prior calculations. The smaller the elasticity, though, the greater will be the implied gains from reallocation. The reason is that as we add more and more human capital to a high-wage sector, the implied wage does not fall as much, and more human capital can be loaded into the high-wage sector before wages converge across sectors.

Hamermesh (1993) discusses estimates of the own-price elasticity of labor demand, finding a range of [0.15,0.75] plausiable and naming 0.30 as the "best" estimate of this elasticity. With an

elasticity of substitution between capital and labor of one (i.e. Cobb-Douglas) α will be equal to this own-price elasticity. Hence the chosen value of 0.30 matches the existing literature, but values as low as 0.15 may be plausible.[11]

While noting that small values are likely to be unrealistic, I recalculate R for both $\alpha = 0.10$ and $\alpha = 0.20$. In both cases, I do so using the estimates produced using the full set of human capital controls (specification 4) and ignoring unmeasured human capital ($\gamma = 0$). Table 3 reports the results of these alternative calculations in columns (8) and (9).

As can be seen, in column (8), with $\alpha = 0.1$, the value of R is distinctly larger for every country, as expected. However, for ten of the countries the gains are still 10% or less. Only for Tajikistan (44.1%), Nepal (16.6%), Ecuador (22.1%), Ghana (75.6%) and Nigeria (21.4%) are the gains higher than 10%. As these results do not account for unmeasured human capital, the actual gains are certainly smaller than these upper bounds.

For the most part, similar conclusions follow from before. Gains of less than 10% could well be practically important, as certainly wage workers in those countries would appreciate an increase in their wages. However, these gains are again very small relative to the income gaps with the developed world, and they therefore would have trouble explaining much of the existing income gaps. More importantly, there still remains no meaningful relationship between the size of gains and the existing income level. In other words, the gains are larger, but are still not systematically related to development.

For those countries that do have large gains, it is useful to look back at figure 2. Consider Ghana, the country with the largest potential gain. As can be seen, in Ghana the Mining sector has a sector-level wage that lies well outside the group of other sectors, with wages per unit of human capital about 2.5 times the country average. With a very low value for α the reallocation is loading nearly everyone into the Mining sector. Thus the average wage is drawn upwards significantly. Whether this is a realistic reallocation, though, is open to question. In the case where $\alpha = 0.1$ the implied efficient allocation is to have 97.3 percent of all the human capital in Ghana work in the Mining sector. It seems unlikely that it is possible to employ that many people (and their human capital) in mining. So the gain of 75.6% in wages after reallocation seems unreasonable as well.

The spread of wages in Tajikistan, Ecuador, Nepal, and Nigeria also show similar wide gaps in figure 2. These generate the large values for R, but the question remains whether the elasticity of wages is actually so low to generate the gains in columns (8) and (9) of table 3. Most importantly, the reallocation is loading much of the human capital stock into the Finance and Commerce sectors for these countries. The question is whether these sectors could absorb large amounts of labor without decreasing wages much, or if they could even absorb that amount of labor at all before

---

[11]If labor and capital are complements in production, which in the static setting under consideration here may be an appropriate assumption, then α will be larger than the own-price elasticity, and values above 0.30 would be prudent. Only if one considers labor and capital to be highly substitutable, with an elasticity of substitution over four or five, would it make sense to consider values of α below 0.10.

complementarities with existing capital become a problem.

This comes back to a point made earlier regarding the flexibility of human capital. In these calculations I am assuming that the human capital stock is perfectly substitutable between different sectors and all units of human capital are perfectly substitutable with one another. Values of α close to 0.10 generate larger gains by exploiting this flexibility fully. However, the perfect substitutability of different units of human capital is certainly not true and so the gains are surely smaller than those indicated by table 3 for $\alpha = 0.10$. Unmeasured human capital would also play a role here, decreasing the implied gains as shown in figure 5. In the end, the values in columns (8) and (9) are at the extreme end of any plausible range for the gains from labor reallocation.

3.7   Sampling Variation in R

The calculated values of R depend on empirical estimates of individual human capital and sector wages. This means that there is some sampling error involved in calculating R. In addition, the sum of observed wages is calculated from surveys, meaning it is subject to sampling error as well. An under-statement of R could occur if the surveys underlying the RIGA data happened to sample a particularly high-wage set of agricultural workers (for example), and a particularly low-wage set of finance workers. In this case the the wage gap I estimate is under-stated and hence R is under-stated as well. It is also possible that sampling errors in the opposite direction mean that the R value I calculated is over-stated.

Because of how R is calculated, a straightforward way of constructing a standard error for this value is not apparent. To remedy this, I turn to bootstrapping. I repeatedly sample, with replacement, from a given country survey, re-estimate equation (17) each time, and use the results to calculate a new version of R. This gives me an approximation of the distribution of R, which I use to report the range of gains from reallocation.

In table 4 I show the baseline value of R in column (1), and then the lower and upper bounds of a 95% confidence interval in columns (2) and (3). Looking at the upper bounds, one can see that even these represent relatively small gains to wages from reallocation. Most countries have an upper bound of less than 10%, and only Nigeria has an upper bound over 30%.

A more formal test of the gains can be made by utilizing the full distribution of estimated R values. Let the null hypothesis be that $H_0 : R \leq \overline{R}$, where $\overline{R}$ is some arbitrary cut-off value, and the alternative is $H_1 : R > \overline{R}$. I calculate the p-value of this test by appealing directly at the distribution of bootstrapped values of R, asking what proportion of them are less or equal to $\overline{R}$.

In columns (4)–(6) of table 4 I show the results of these tests for each country at several cut-off values of $\overline{R}$. Column (4) shows the test against $\overline{R} = 1.01$, so the null hypothesis is that the gains of reallocation are actually less than a 1% increase in wages. As can be seen, for most countries I can reject this null as the p-values are zero. Practically speaking, none of the five-hundred bootstrapped iterations produced a value of R less than or equal to 1.01. However, for Bulgaria I can only reject

this null at the 7% level, and for Bangladesh, Indonesia, and Panama I cannot reject the null at any typical level of significance. Statistically, the gains of reallocation in these three countries are almost certainly equivalent to a less than 1% increase in wages.

Moving to column (5), the cut-off value is now a gain of 5%, so that $\overline{R} = 1.05$. Here, far more countries fail to reject the null. Only Tajikistan, Ecuador, Ghana, and Nigeria are able to reject that gains are less than this level. For the rest, the null hypothesis cannot be rejected, as in almost every case the p-value is above 0.90. In column (6) the cut-off is increased to $\overline{R} = 1.10$ and there remains only one country, Tajikistan, that can reject the null. For the rest of the countries I cannot reject that gains from reallocation are equivalent to a less than 10% increase in wages. Again, given the likelihood of unmeasured human capital playing an important role, the gains from reallocation are almost certainly lower than what I've estimated, and hence gains are almost certainlyl less than 10%.

4   Conclusion

Large differences in output per worker across sectors suggest that developing countries may be operating well below their maximum efficiency. One possible explanation for this is that human capital is misallocated across different sectors within these countries. This paper calculates the hypothetical aggregate impact of this misallocation. In constrast to the existing macro-level literature that uses sector-level aggregate data, I use micro-level wage data from 14 developing countries to measure the marginal product of human capital within sectors and use that to estimate the aggregate effect of misallocation.

I make several assumptions meant to inflate the hypothetical gain to wages from an optimal allocation of human capital across sectors, but in only a few cases does the potential gain in aggregate output even reach 10%. For most countries the gains are well below 5%. In addition there is no tendency for the gains to become larger as the country becomes poorer, implying that the inefficiency cannot explain cross-country income differences. The implied gains to reallocation in the United States are smaller than in most of the countries I examine, but the difference is miniscule compared to the overall difference in development levels.

Accounting parametrically for unmeasured human capital reduces the implied gains even further, to below 3% for nearly all countries. For almost every country I cannot reject that the gains from reallocation are less than a 5% increase in wages. Allowing for the presence of self-employed workers does not demonstrably change the conclusions.

Despite the appearance of large earnings gaps between sectors in each developing country, the implied gain from reallocating human capital efficiently between sectors is very small. Labor market inefficiencies in these developing countries may exist, but they do not appear to have significant effects on aggregate productivity.

References

[1] Abowd, John, Francis Kramarz, and D. Margolis. 1999. "High Wage Workers and High Wage Firms," *Econometrica*, 67(2):251-334.

[2] Abowd, John, H. Finer and Francis Kramarz. 1999. "Individual and Firm Heterogeneity in Compensation: An Analysis of Matched Longitudinal Employer-Employee Data for the State of Washington," in J. Haltiwanger, J. Lane, and P. Theeuwes (Eds.), *The Creation and Analysis of Employer-Employee Matched Data*. Amsterdam: North Holland.

[3] Banarjee, Abhijit and Esther Duflo. 2005. "Growth Theory through the Lens of Development Economics," in P. Aghion and S. Durlauf (Eds.) *Handbook of Economic Growth*. New York: North Holland.

[4] Caselli, Francesco. 2005. "Accounting for Cross-Country Income Differences," in P. Aghion and S. Durlauf (Eds.) *Handbook of Economic Growth*. New York: North Holland.

[5] 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.

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

[7] Davis, Benjamin, Paul Winters, Gero Carletto, Katia Covarrubias, Estaban Quinones, Alberto Zezza, Kostas Stamoulis, Carlo Azzarri, and Stefania Digiuseppe. 2010. "A Cross-Country Comparison of Rural Income Generating Activities," *World Development*, 38(1):48.63.

[8] Foster, Lucia, John Haltiwanger and C. J. Krizan. 2001 "Aggregate Productivity Growth: Lessons from Microeconomic Evidence," in C. Hulten, E. Dean, and M. Harper (Eds.) *New Developments in Productivity Analysis*. Chicago and London: University of Chicago Press.

[9] Gollin, Douglas, David Lagakos and Michael Waugh. 2011. "The Agricultural Productivity Gap in Developing Countries," working paper.

[10] Gollin, Douglas, Stephen Parente and Richard Rogerson. 2002. "The Role of Agriculture in Development," *American Economic Review Papers and Proceedings*, 92(2):160-164.

[11] Goux, D. and E. Martin. 1999. "Persistence of Inter-Industry Wage Differentials: A reexamination on Matched Worker-Firm Panel Data," *Journal of Labor Economics*, 17:492-533.

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

[13] Hamermesh, Daniel. 1993. *Labor Demand*. Princeton University Press: Princeton, NJ.

[14] Herrendorf, Berthold and Todd Schoellman. 2011. "Why is Measured Productivity so Low in Agriculture?" working paper.

[15] Hsieh, Chang-Tai and Peter J. Klenow. 2009. "Mis-allocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124(4):1403-48.

[16] Jones, Charles I. 2011. "Misallocation, Economic Growth, and Input-Output Economics," NBER working paper 16742.

[17] Krueger, Alan and Lawrence Summers. 1987. "Reflections on the Inter-Industry Wage Structure." in K. Lang and J. Leonard (Eds.), *Unemployment and the Structure of Labor Markets*. Oxford: Basil Blackwell.

[18] Krueger, Alan and Lawrence Summers. 1988. "Efficiency Wages and the Inter-Industry Wage Structure," *Econometrica*, 56(2):259-93.

[19] Kuznetz, Simon. 1971. *Economic Growth of Nations: Total Output and Production Structure*. Harvard University Press.

[20] Lagakos, David and Michael Waugh. 2011. "Selection, Agriculture, and Cross-Country Productivity Differences," working paper.

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

[22] McMillan, Margaret and Dani Rodrik. 2011. "Globalization, Structural Change, and Productivity Growth," NBER working paper w17143.

[23] Restuccia, Diego and Richard Rogerson. 2008. "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments," *Review of Economic Dynamics*, 11(4):707-20.

[24] Restuccia, Diego, Dennis Yang, and Xiaodong Zhu. 2008. "Agriculture and Aggregate Productivity: A Quantititative Cross-Country Analysis," *Journal of Monetary Economics*, 55:234-250.

[25] Syverson, Chad. 2011. "What Determines Productivity?" *Journal of Economic Literature*, 49(2):326-365.

[26] Vollrath, Dietrich. 2009a. "How Important are Dual Economy Effects for Aggregate Productivity?" *Journal of Development Economics*, 11(3):187-228.

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

Appendices

A. Mincerian Regressions

The estimates of the gains from reallocation, R, are based on Mincerian regressions made within each country. These are used both to determine the industry-specific wage premium and the amount of individual human capital. The underlying regressions are not reported in the main body of the paper.

Tables A.1 and A.2 show the estimated industry dummies for each country with no other controls included. Agriculture is the excluded industry. Standard errors are reported in parentheses. The outcome of the joint test that all industry dummies are equal to zero is reported at the foot of the table, along with the p-value of the test. A p-value of 0.000 indicates that the p-value was less than 0.001. These estimates are the basis for the values of R found in column (1) of table 3.

Tables A.3 and A.4 show similar regressions, but using the full set of controls for human capital: years of education, age, age-squared, gender, occupation dummies, and occupation/education interactions. These estimates are the basis for the values of R found in column (4) of table 3.

Results for the regressions in the intermediate cases (columns 2 and 3 of table 3) are not reported for brevity's sake, but are available upon request.

B. United States Data

The estimated gains from reallocation in the United States are based on the Current Population Survey, March extract, 2000. Only those reporting themselves as "Employed - at work" are included in the estimation, consistent with the method used with the RIGA data. Observations were dropped if they did not have age, wage, or gender data.

Several recodes of the CPS data were necessary to have industry and occupation definitions harmonized with the RIGA datasets and ensure an equal comparison.

- Education (variable *grade92* in the CPS)

  · Less than 1st grade (coded 31 in CPS) recoded to 1 year of education

  · Up to fourth grade (32) recoded to 4 years

  · Up to sixth grade (33) recoded to 6 years

  · Up to eigth grade (34) recoded to 8 years

  · 9th grade (35) recoded to 9 years

  · 10th grade (36) recoded to 10 years

  · 11th grade (37) recoded to 11 years

  · 12th with no diploma (38) and high school graduate or GED (39) recoded to 12 years

  · Some college, no degree (40) recoded to 13 years

  · Associates degree (41 and 42) recoded to 14 years

  · Bachelors degree (43) recoded to 16 years

  · Masters degree (44) recoded to 18 years

  · Professional degree (45) recoded to 19 years

· Ph.D. (46) recoded to 21 years

- Occupation (variable *docc80* in CPS)

  · Codes 1,2,3 recoded to 1 (Senior officials and managers)
  · Codes 4–12 recoded to 2 (Professionals)
  · Codes 13–15 recoded to 3 (Technicians)
  · Codes 21–26 recoded to 4 (Clerks)
  · Codes 16–20 and 27–32 recoded to 5 (Sales and service workers)
  · Code 43 recoded to 6 (Skilled agriculture workers)
  · Codes 33,34 recoded to 7 (Craft and related trade workers)
  · Codes 36–39 recoded to 8 (Plant and machine operators)
  · Codes 40,41,42,44,45 recoded to 9 (Elementary laborers)
  · Code 46 recoded to 10 (Armed forces)

- Industry (variable *dind* in CPS)

  · Codes 1,2,46 recoded to 1 (Ag, Forestry, Fisheries)
  · Code 3 recoded to 2 (Mining)
  · Codes 3–28 recoded to 3 (Manufacturing)
  · Code 31 recoded to 4 (Utilities)
  · Code 4 recoded to 5 (Construction)
  · Codes 32,33 recoded to 6 (Wholesale and retail trade)
  · Codes 29,30 recoded to 7 (Transport and communications)
  · Codes 34,35 recoded to 8 (Finance and real estate)
  · Codes 36–45 recoded to 9 (Services)

C. Additional Sector-level Information

Table A.5 reports the percent allocation of jobs reported for a country in the RIGA database in each of the sectors. Table A.6 reports the shift in human capital into (positive entries) or out of (negative entries) sectors within a country to achieve the most efficient allocation. The calculation is made based on the sector wage dummies estimated in the Mincerian regressions using the full set of controls: sector dummies, years of education, age, age-squared, gender, occupation dummies, and occupation/education interactions. The entries in the table show the percentage of total human capital in the entire economy that would have to be moved. For example, in Tajikistan, human capital equivalent to 47.7% of the entire stock of human capital in the country would have to leave agriculture. The fraction of actual agricultural human capital that would have to move is thus much higher.
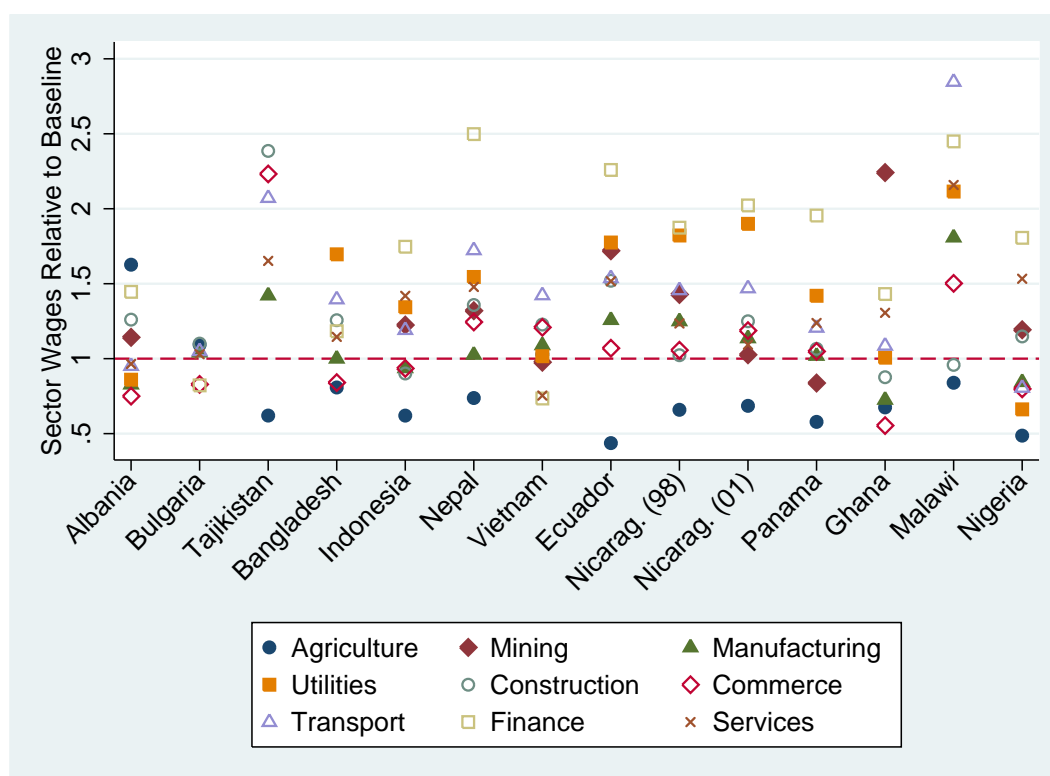
Figure 1: Average earnings per worker per day, by sector, by country

Note: The figure shows, for each country, the average wage per day in a sector relative to the average wage for the entire country. Authors calculations using the RIGA database of Davis et al (2010).
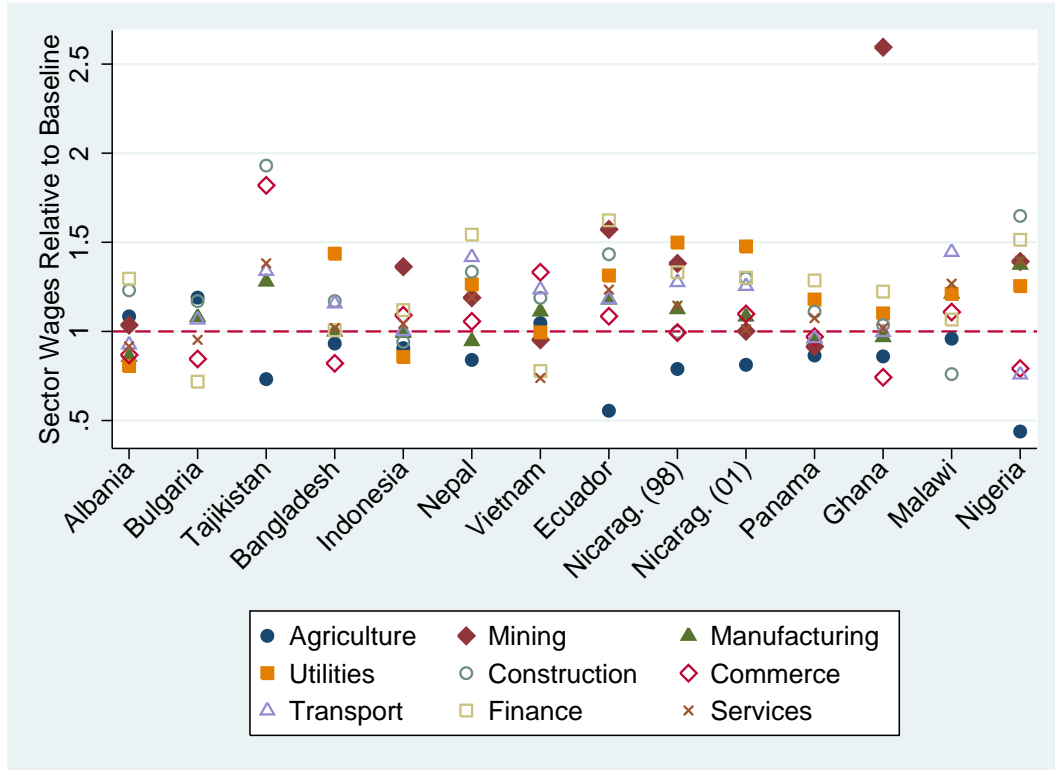
Figure 2: Average earnings per worker per day controlling for human capital, by sector, by country

Note: The figure shows, for each country, the average residual wage per day in a sector relative to the average residual wage for the entire country, where the residual wage is the wage remaining after controlling for human capital characteristics (education, age, and occupation). Authors calculations using the RIGA database of Davis et al (2010).
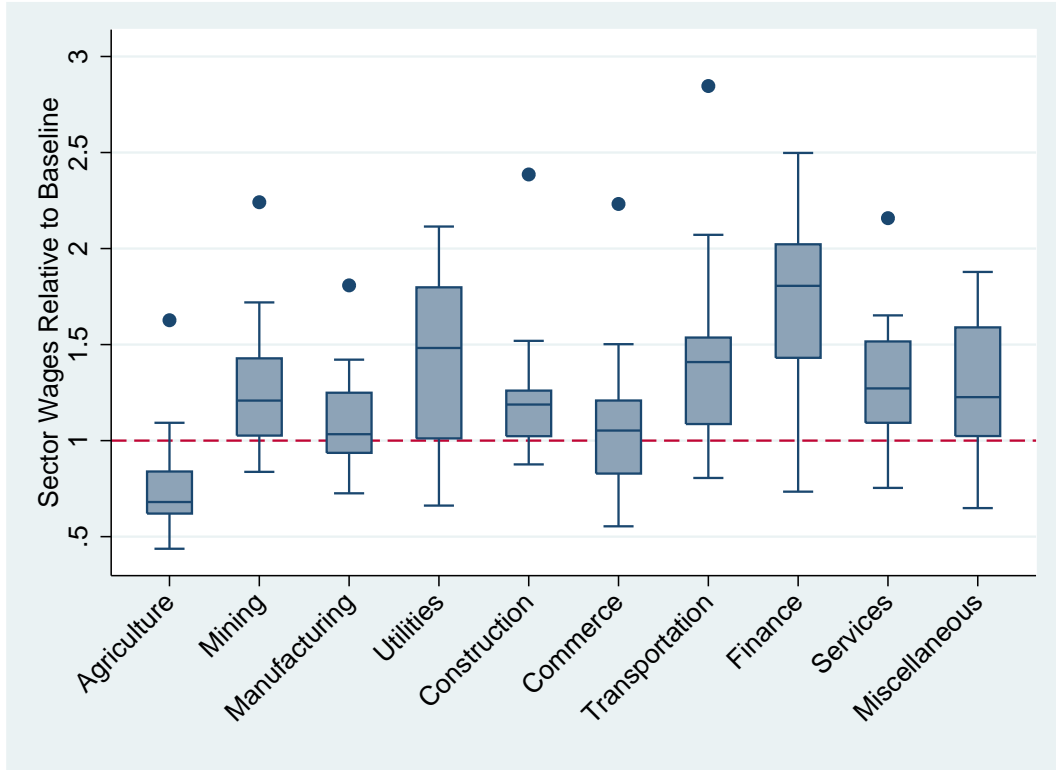
Figure 3: Sector-level Wages, No Human Capital Controls

Note: The figure shows the values of ϕ_j, which is the sector-level wage relative to the average country wage, calculated as in equation (18) in the text. The sector-level wages are calculated using the sector dummies estimated using specification (1), which has no individual characteristics included in the wage regressions.
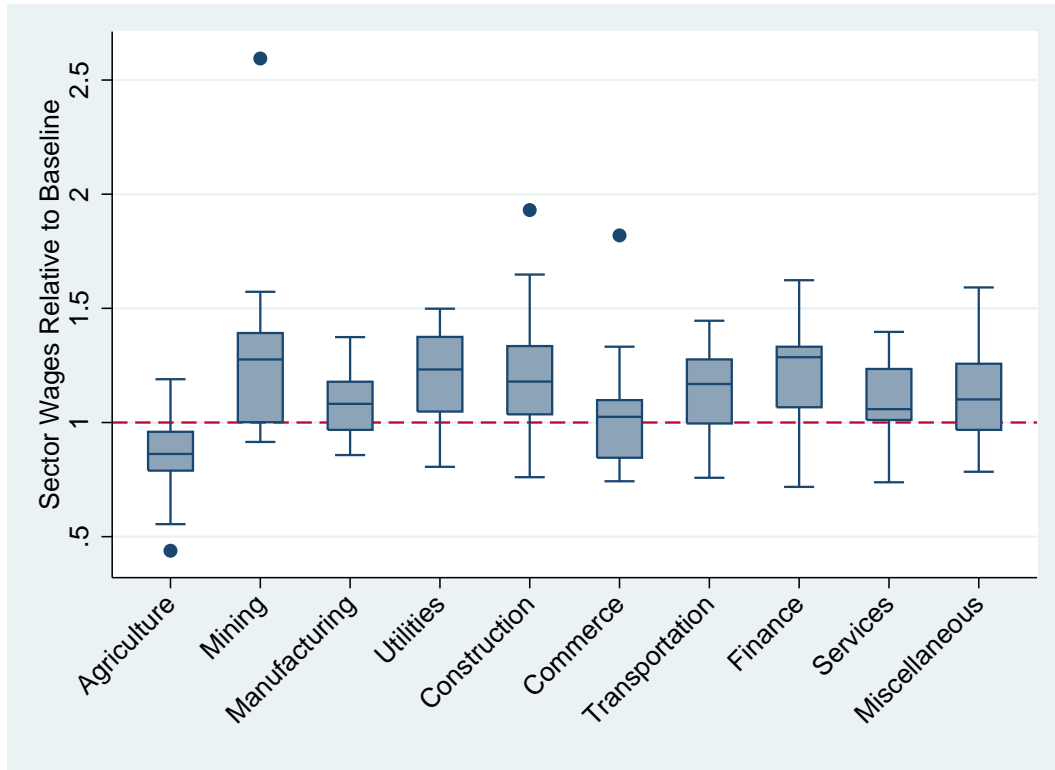
Figure 4: Sector-level Wages, Full Human Capital Controls

Note: The figure shows the values of ϕ_j, which is the sector-level wage relative to the average country wage, calculated as in equation (18) in the text. The sector-level wages are calculated using the sector dummies estimated using specification (1), which has no individual characteristics included in the wage regressions.
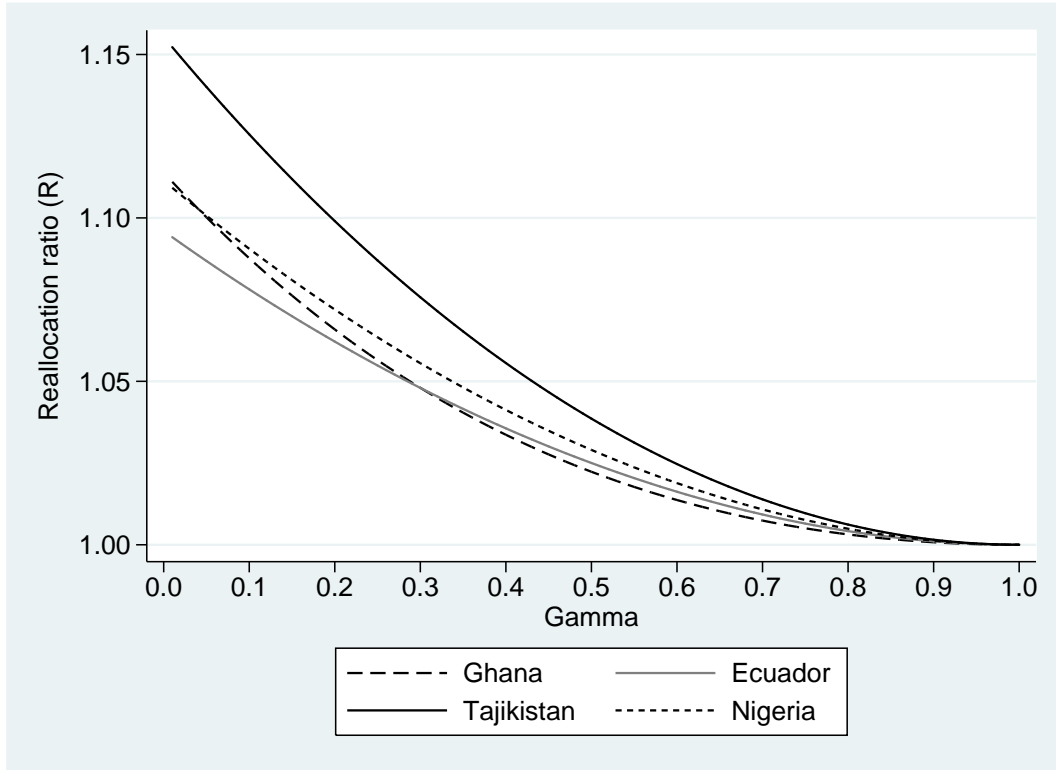
Figure 5: Reallocation Ratio and Unmeasured Human Capital

Note: The figure shows the gains from reallocation (R) plotted against the fraction (γ) of the sector-level return (δⱼ) that is attributed to unmeasured human capital. The four countries in the figure are those with the highest estimated gains from reallocation based on the original analysis with $\gamma = 0$.

Table 1: Summary Statistics, by Country

| Country | Year | Country Data: | | | From RIGA: | | | |
		GDP p.c.	Education	Perc. Rural	Obs.	CV(wage)	Education	Perc. Rural
Ex-Communist								
Albania	2005	$6,107	10.2	55.0	2,396	3.05	11.6	26.9
Bulgaria	2001	$7,664	9.4	31.0	3,397	1.39	11.4	24.3
Tajikistan	2003	$1,250	9.9	74.0	4,593	1.24	11.1	72.4
Asia								
Bangladesh	2000	$1,003	3.7	76.0	6,898	0.80	3.3	64.6
Indonesia	2000	$2,623	4.8	58.0	8,824	1.71	10.7	42.6
Nepal	2003	$919	2.7	85.0	4,397	0.90	3.8	72.5
Vietnam	1998	$1,469	4.5	77.0	6,212	0.58	7.4	63.2
Latin America								
Ecuador	1995	$5,664	6.9	42.0	7,833	1.09	8.6	39.4
Guatemala	2000	$3,960	3.8	55.0	10,173	1.07	5.5	49.5
Nicaragua	1998	$1,982	4.6	46.0	5,069	1.47	5.7	40.6
Nicaragua	2001	$2,169	4.6	45.0	5,412	0.91	6.2	38.4
Panama	2003	$8,240	9.0	31.0	7,745	3.44	9.9	40.4
Sub-Saharan Africa								
Ghana	1998	$1,033	6.3	5.0	1,631	1.21	7.8	48.3
Malawi	2004	$646	3.4	83.0	13,030	1.57	4.8	88.1
Nigeria	2004	$1,702	-	55.0	3,756	11.80	5.8	54.4

Notes: For country-level data, GDP per capita and percent urban are from the World Development Indicators. GDP per capita is PPP in 2005 international dollars. Education data is years attained by population 25 or older, from Barro and Lee (2010). The RIGA data is described in the text.

Table 2: Wages Relative to Overall Mean Wage, by Sector and Country

Country (Year)	Sector:								
	Agriculture	Mining	Manufacturing	Utilities	Construction	Commerce	Transportation	Finance	Services
Albania (2005)	1.62	1.15	0.83	0.87	1.27	0.75	0.96	1.46	0.97
Bulgaria (2001)	1.09	−	1.04	−	1.11	0.83	1.04	0.84	1.03
Tajikistan (2003)	0.62	−	1.43	−	2.39	2.22	2.08	−	1.65
Bangladesh (2000)	0.81	−	0.99	1.71	1.25	0.84	1.39	1.18	1.14
Indonesia (2000)	0.62	1.23	0.93	1.35	0.90	0.93	1.19	1.75	1.40
Nepal (2003)	0.73	1.33	1.01	1.57	1.36	1.24	1.71	2.53	1.46
Vietnam (1998)	1.02	0.98	1.08	1.02	1.23	1.21	1.41	0.74	0.75
Ecuador (1995)	0.43	1.73	1.25	1.80	1.52	1.07	1.54	2.27	1.52
Guatemala (2000)	0.68	0.94	1.01	1.26	1.19	1.12	1.44	2.04	1.23
Nicaragua (1998)	0.64	1.46	1.24	1.83	1.03	1.05	1.45	1.89	1.22
Nicaragua (2001)	0.68	1.04	1.13	1.91	1.25	1.18	1.47	2.03	1.08
Panama (2003)	0.55	0.88	1.01	1.43	1.07	1.05	1.21	1.96	1.23
Ghana (1998)	0.67	2.25	0.71	1.03	0.88	0.56	1.08	1.45	1.31
Malawi (2004)	0.84	−	1.78	2.15	0.96	1.50	2.84	2.47	2.13
Nigeria (2004)	0.48	1.18	0.83	0.68	1.17	0.80	0.81	1.82	1.52

Notes: The table reports the average of the individual wages relative to the overall mean wage, by country and sector. The data are from the RIGA database, Davis et al (2010). See text for more details. Blank cells indicate that no individuals reported themselves as working in that sector.

Table 3: Estimated Efficiency Gain from Re-allocation, by Specification and Country

| Country (Year) | Baseline: | | | | Including self-employed: | | | Wage elasticity: | |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Albania (2005)	1.064	1.071	1.027	1.027	1.036	1.015	1.031	1.100	1.047
Bulgaria (2001)	1.008	1.010	1.015	1.014	1.014	1.014	1.015	1.049	1.024
Tajikistan (2003)	1.342	1.249	1.153	1.155	1.155	1.155	1.155	1.441	1.251
Bangladesh (2000)	1.045	1.027	1.010	1.009	1.023	1.009	1.011	1.044	1.016
Indonesia (2000)	1.102	1.019	1.005	1.006	1.024	1.007	1.007	1.027	1.010
Nepal (2003)	1.129	1.060	1.050	1.051	1.103	1.050	1.030	1.166	1.086
Vietnam (1998)	1.036	1.036	1.037	1.035	1.043	1.024	1.037	1.100	1.056
Ecuador (1995)	1.219	1.141	1.100	1.096	1.113	1.106	1.100	1.221	1.139
Guatemala (2000)	1.097	1.030	1.023	1.025	1.062	1.032	1.019	1.078	1.041
Nicaragua (1998)	1.075	1.034	1.028	1.028	1.028	1.028	1.028	1.081	1.045
Nicaragua (2001)	1.067	1.031	1.025	1.025	1.025	1.025	1.025	1.095	1.043
Panama (2003)	1.065	1.024	1.009	1.009	1.012	1.010	1.009	1.044	1.017
Ghana (1998)	1.140	1.089	1.108	1.114	1.145	1.103	1.082	1.756	1.282
Malawi (2004)	1.261	1.059	1.028	1.026	1.036	1.024	1.024	1.100	1.045
Nigeria (2004)	1.119	1.086	1.100	1.111	1.205	1.183	1.163	1.214	1.154
United States (2000, weekly)	1.051	1.021	1.018	1.018	-	-	-	1.065	1.030
United States (2000, hourly)	1.023	1.009	1.009	1.009	-	-	-	1.034	1.015
Controls included in specification:									
Education and Demographic	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation-specific returns	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Self-empl. wage/agric. wage (ϕ)	-	-	-	-	0.75	1.00	1.25	-	-
Wage elasticity (α)	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.1	0.2

Notes: The value of R is the ratio of wages under the optimal allocation of human capital to the observed level of wages, see text for details. The estimates of R depend upon estimated Mincer equations, using the human capital controls indicated, again see text for details regarding the exact specification. The self-employed columns are calculated including hypothetical self-employed workers earning a fraction, ϕ of the agricultural wage. The number of self-employed is set to match FAO data on the size of the agricultural labor force. Self-employed results for the U.S. are not included as those workers are already in the main sample. The wage elasticity columns refer to the value of α, which dictates the response of wages to the amount of human capital in a sector. The different rows for the U.S. refer to whether weekly earnings or hourly earnings are used in the calculations.

33

Table 4: Sampling Variation of R within Countries

Country (Year)	Baseline R (1)	95% C.I. Lower (2)	95% C.I. Upper (3)	P-values for test $H_0: R \leq \overline{R}$ $\overline{R} = 1.01$ (4)	$\overline{R} = 1.05$ (5)	$\overline{R} = 1.10$ (6)
Albania (2005)	1.027	1.015	1.090	0.000	0.814	0.982
Bulgaria (2001)	1.014	1.008	1.032	0.070	> 0.998	> 0.998
Tajikistan (2003)	1.155	1.129	1.203	0.000	0.000	0.000
Bangladesh (2000)	1.009	1.006	1.015	0.520	> 0.998	> 0.998
Indonesia (2000)	1.006	1.003	1.012	0.870	> 0.998	> 0.998
Nepal (2003)	1.051	1.041	1.074	0.000	0.294	> 0.998
Vietnam (1998)	1.035	1.028	1.045	0.000	0.998	> 0.998
Ecuador (1995)	1.096	1.072	1.142	0.000	0.000	0.546
Guatemala (2000)	1.025	1.017	1.037	0.000	> 0.998	> 0.998
Nicaragua (1998)	1.028	1.023	1.056	0.000	0.936	> 0.998
Nicaragua (2001)	1.025	1.018	1.037	0.000	> 0.998	> 0.998
Panama (2003)	1.009	1.006	1.017	0.530	> 0.998	> 0.998
Ghana (1998)	1.114	1.058	1.194	0.000	0.006	0.298
Malawi (2004)	1.026	1.017	1.042	0.000	> 0.998	> 0.998
Nigeria (2004)	1.111	1.037	1.346	0.000	0.042	0.316

Notes: The first column is the gain from reallocation, R, calculated using the original sample. The 95% confidence interval for R is found using bootstrapping, with 500 replications. The remaining columns show the results for tests that the gains of reallocation, R, are less than or equal to a cutoff level $\overline{R}$. The p-values are obtained numerically from the distribution of R values obtained in the bootstrapping procedure described in the text. > 0.998 values mean that the numerical results indicated that all of the 500 bootstrapped calculations of R were below the cutoff. The values of R are based on regressions using the full set of controls, a value of $\alpha = 0.3$, and excluding any adjustment for self-employed workers.

Appendix Table A.1: Mincerian Regressions, No Human Capital Controls

	(1) Albania	(2) Bulgaria	(3) Tajikistan	(4) Bangladesh	(5) Indonesia	(6) Nepal	(7) Vietnam
Mining	-0.345				0.688	0.587	-0.048
	(0.135)				(0.128)	(0.101)	(0.081)
Manufacturing	-0.667	-0.046	0.836	0.214	0.413	0.331	0.058
	(0.135)	(0.073)	(0.122)	(0.027)	(0.035)	(0.040)	(0.022)
Utilities	-0.627			0.748	0.778	0.754	-0.006
	(0.144)			(0.113)	(0.106)	(0.172)	(0.108)
Construction	-0.246	0.012	1.350	0.442	0.374	0.611	0.177
	(0.134)	(0.097)	(0.079)	(0.028)	(0.040)	(0.022)	(0.022)
Commerce	-0.765	-0.274	1.283	0.043	0.411	0.525	0.162
	(0.139)	(0.074)	(0.068)	(0.077)	(0.042)	(0.066)	(0.037)
Transportation	-0.526	-0.037	1.212	0.547	0.654	0.853	0.326
	(0.148)	(0.084)	(0.111)	(0.023)	(0.050)	(0.098)	(0.043)
Finance	-0.106	-0.266		0.382	1.038	1.233	-0.328
	(0.156)	(0.193)		(0.041)	(0.077)	(0.163)	(0.130)
Services	-0.514	-0.057	0.980	0.350	0.828	0.696	-0.310
	(0.128)	(0.074)	(0.035)	(0.037)	(0.034)	(0.037)	(0.027)
Miscellanesous				0.250	0.698	0.086	
				(0.077)	(0.178)	(0.184)	
Observations	2381	3089	4571	6893	8818	4387	6211
Joint F-stat	14.19	7.54	227.60	98.39	79.72	122.40	42.13
Joint p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: Log daily wage is the dependent variable. Mincerian regressions for each country using only industry dummies as explanatory variables. Standard errors in parentheses. Missing cells imply that fewer than 10 individuals in sample reported as working in that sector. The F-test has a null that all sector dummies are jointly equal to zero. Agriculture is the excluded category.

Appendix Table A.2: Mincerian Regressions, No Human Capital Controls

	(1) Ecuador	(2) Guatemala	(3) Nicaragua	(4) Panama	(5) Ghana	(6) Malawi	(7) Nigeria
Mining	1.378	0.311	0.414	0.421	1.205		0.913
	(0.123)	(0.131)	(0.147)	(0.319)	(0.097)		(0.180)
Manufacturing	1.061	0.403	0.506	0.567	0.078	0.768	0.578
	(0.086)	(0.040)	(0.033)	(0.048)	(0.107)	(0.042)	(0.234)
Utilities	1.414	0.601	1.024	0.904	0.425	0.940	0.356
	(0.158)	(0.195)	(0.093)	(0.110)	(0.227)	(0.179)	(0.311)
Construction	1.250	0.548	0.602	0.612	0.269	0.134	0.877
	(0.084)	(0.038)	(0.047)	(0.050)	(0.127)	(0.053)	(0.190)
Commerce	0.899	0.487	0.550	0.595	-0.192	0.585	0.500
	(0.087)	(0.038)	(0.042)	(0.036)	(0.112)	(0.074)	(0.101)
Transportation	1.262	0.735	0.766	0.736	0.482	1.226	0.511
	(0.091)	(0.052)	(0.088)	(0.056)	(0.113)	(0.100)	(0.118)
Finance	1.648	1.085	1.087	1.220	0.763	1.079	1.320
	(0.101)	(0.058)	(0.101)	(0.061)	(0.154)	(0.127)	(0.132)
Services	1.248	0.591	0.467	0.761	0.664	0.945	1.152
	(0.085)	(0.030)	(0.030)	(0.034)	(0.087)	(0.037)	(0.089)
Miscellanesous	0.399	0.789	0.851	1.188		0.792	0.761
	(0.088)	(0.150)	(0.139)	(0.144)		(0.162)	(0.185)
Observations	7790	10102	5404	7741	1618	13001	3685
Joint F-stat	82.84	103.04	62.23	77.78	36.23	152.19	27.08
Joint p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: Log daily wage is the dependent variable. Mincerian regressions for each country using only industry dummies as explanatory variables. Standard errors in parentheses. Missing cells imply that fewer than 10 individuals in sample reported as working in that sector. The F-test has a null that all sector dummies are jointly equal to zero. Agriculture is the excluded category.

Appendix Table A.3: Mincerian Regressions, All Human Capital Controls

	(1) Albania	(2) Bulgaria	(3) Tajikistan	(4) Bangladesh	(5) Indonesia	(6) Nepal	(7) Vietnam
Mining	-0.006				0.416	0.354	-0.091
	(0.285)				(0.122)	(0.112)	(0.078)
Manufacturing	-0.195	-0.092	0.565	0.070	0.091	0.119	0.061
	(0.284)	(0.095)	(0.118)	(0.039)	(0.040)	(0.046)	(0.029)
Utilities	-0.256			0.437	-0.050	0.422	-0.045
	(0.287)			(0.099)	(0.114)	(0.166)	(0.108)
Construction	0.166	-0.010	0.974	0.228	0.040	0.464	0.127
	(0.283)	(0.116)	(0.090)	(0.042)	(0.057)	(0.030)	(0.033)
Commerce	-0.183	-0.337	0.914	-0.124	0.186	0.233	0.242
	(0.283)	(0.097)	(0.088)	(0.073)	(0.043)	(0.083)	(0.053)
Transportation	-0.117	-0.104	0.611	0.217	0.102	0.527	0.167
	(0.285)	(0.108)	(0.121)	(0.039)	(0.060)	(0.086)	(0.050)
Finance	0.220	-0.485		0.078	0.215	0.622	-0.287
	(0.291)	(0.202)		(0.046)	(0.088)	(0.164)	(0.135)
Services	-0.131	-0.217	0.637	0.091	0.141	0.352	-0.349
	(0.277)	(0.099)	(0.050)	(0.044)	(0.040)	(0.052)	(0.035)
Miscellanesous				0.125	0.402	-0.056	
				(0.060)	(0.171)	(0.158)	
Observations	2381	3089	4571	6893	8816	4387	6211
Joint F-stat	6.80	6.88	52.92	10.92	4.42	33.84	29.72
Joint p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: Log daily wage is the dependent variable. Mincerian regressions for each country using industry dummies as well as age, age-squared, gender, years of education, dummies for occupation, and the interaction of occupation with years of education as explanatory variables. Standard errors in parentheses. Missing cells imply that fewer than 10 individuals in sample reported as working in that sector. The F-test has a null that all sector dummies are jointly equal to zero. Agriculture is the excluded category.

Appendix Table A.4: Mincerian Regressions, All Human Capital Controls

	(1) Ecuador	(2) Guatemala	(3) Nicaragua	(4) Panama	(5) Ghana	(6) Malawi	(7) Nigeria
Mining	1.054	0.191	0.221	0.082	1.112		1.239
	(0.156)	(0.131)	(0.157)	(0.228)	(0.121)		(0.407)
Manufacturing	0.759	0.241	0.288	0.109	0.124	0.234	1.239
	(0.100)	(0.044)	(0.038)	(0.075)	(0.103)	(0.042)	(0.438)
Utilities	0.876	0.102	0.601	0.320	0.265	0.249	1.160
	(0.165)	(0.193)	(0.090)	(0.140)	(0.188)	(0.180)	(0.465)
Construction	0.954	0.445	0.470	0.253	0.194	-0.231	1.427
	(0.101)	(0.049)	(0.044)	(0.080)	(0.125)	(0.049)	(0.453)
Commerce	0.676	0.264	0.302	0.117	-0.140	0.147	0.645
	(0.104)	(0.041)	(0.043)	(0.075)	(0.111)	(0.071)	(0.329)
Transportation	0.761	0.335	0.438	0.110	0.153	0.415	0.608
	(0.107)	(0.062)	(0.086)	(0.082)	(0.108)	(0.096)	(0.349)
Finance	1.081	0.446	0.475	0.401	0.362	0.114	1.301
	(0.119)	(0.057)	(0.088)	(0.090)	(0.134)	(0.123)	(0.351)
Services	0.805	0.308	0.219	0.220	0.173	0.280	1.218
	(0.098)	(0.038)	(0.034)	(0.073)	(0.090)	(0.040)	(0.344)
Miscellanesous	0.562	0.322	0.321	0.617		0.286	0.694
	(0.107)	(0.148)	(0.150)	(0.116)		(0.174)	(0.334)
Observations	7790	10102	5404	7741	1618	13001	3685
Joint F-stat	15.39	14.48	18.73	7.61	16.71	15.39	3.87
Joint p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: Log daily wage is the dependent variable. Mincerian regressions for each country using industry dummies as well as age, age-squared, gender, years of education, dummies for occupation, and the interaction of occupation with years of education as explanatory variables. Standard errors in parentheses. Missing cells imply that fewer than 10 individuals in sample reported as working in that sector. The F-test has a null that all sector dummies are jointly equal to zero. Agriculture is the excluded category.

Appendix Table A.5: Sectoral Allocation of Jobs, in Percents, by Country

Country (Year)	Agriculture	Mining	Manufacturing	Utilities	Construction	Commerce	Transportation	Finance	Services
					Sector:				
Ex-Communist									
Albania (2005)	4.6	2.3	12.3	3.6	18.4	12.3	4.9	1.3	40.5
Bulgaria (2001)	6.8	-	27.0	-	4.8	18.7	10.3	1.7	30.7
Tajikistan (2003)	54.6	-	1.4	-	3.9	4.2	3.3	-	32.5
Asia									
Bangladesh (2000)	36.5	-	21.4	0.6	5.3	2.2	10.3	7.2	14.2
Indonesia (2000)	21.6	0.8	18.9	0.5	8.6	11.1	5.0	1.4	31.8
Nepal (2003)	47.2	0.7	13.0	0.8	16.1	2.7	3.2	0.8	14.5
Vietnam (1998)	31.2	1.3	20.1	0.9	11.5	5.0	4.0	0.7	25.2
Latin America									
Ecuador (1995)	23.9	0.9	13.7	0.7	8.7	16.7	4.4	1.6	23.9
Guatemala (2000)	34.2	0.2	12.4	0.6	7.9	13.1	3.3	3.4	24.7
Nicaragua (1998)	31.3	0.5	10.6	1.1	6.4	13.3	3.9	0.7	31.4
Nicaragua (2001)	29.4	0.8	12.0	1.1	7.8	12.4	3.7	0.8	31.6
Panama (2003)	21.3	0.1	6.1	0.8	7.4	21.7	4.7	1.5	36.2
Sub-saharan Africa									
Ghana (1998)	11.0	3.3	15.2	0.8	4.9	10.8	9.1	4.1	40.7
Malawi (2004)	76.9	-	5.2	0.4	3.6	2.2	1.0	0.4	9.9
Nigeria (2004)	17.4	0.5	3.4	0.7	1.9	22.9	6.9	2.6	40.9

Notes: The table reports the percentage of wage-earning jobs that are in each sector. The data are from the RIGA database, Davis et al (2010). See text for more details. Blank cells indicate that no individuals reported themselves as working in that sector.

Appendix Table A.6: Re-allocation of Human Capital, as a Percent of Total Human Capital, by Country

Country (Year)	Agriculture	Mining	Manufacturing	Utilities	Construction	Commerce	Transportation	Finance	Services
Ex-Communist									
Albania (2005)	0.7	0.1	−6.1	−1.6	21.2	−4.8	−1.5	1.0	−9.0
Bulgaria (2001)	5.2	-	5.8	.	3.0	−7.3	1.7	−1.9	−6.4
Tajikistan (2003)	−47.7	-	0.4	-	16.3	11.3	2.1	-	17.6
Asia									
Bangladesh (2000)	−8.0	-	−1.1	1.4	2.9	−1.1	5.7	−0.3	0.2
Indonesia (2000)	−5.3	1.4	−1.2	−0.3	−1.6	2.5	−0.2	0.7	3.9
Nepal (2003)	−25.3	0.2	−4.9	0.9	15.8	−0.2	6.4	2.1	5.9
Vietnam (1998)	0.5	−0.4	4.7	−0.1	6.4	4.9	3.3	−0.4	−18.9
Latin America									
Ecuador (1995)	−26.5	2.0	2.6	0.7	9.6	−1.3	1.0	3.2	9.9
Guatemala (2000)	−16.7	−0.0	−0.6	−0.3	7.1	0.4	1.1	4.0	4.9
Nicaragua (1998)	−14.8	0.7	2.3	2.0	−1.4	−3.4	5.8	0.8	7.9
Nicaragua (2001)	−13.0	−0.1	1.0	2.7	7.2	2.0	3.9	1.0	−4.8
Panama (2003)	−5.2	−0.0	−1.4	0.5	2.1	−4.3	−1.2	2.4	6.4
Sub-saharan Africa									
Ghana (1998)	−5.1	35.2	−4.9	−0.0	−0.9	−7.0	−3.2	1.3	−15.5
Malawi (2004)	−19.3	-	3.0	0.5	−3.1	0.4	3.6	0.0	14.5
Nigeria (2004)	−13.2	0.1	1.3	0.1	2.0	−15.0	−3.8	1.5	28.6

Notes: The table reports the implied change in human capital in each sector, as a percent of total human capital, from the optimization explained in the text. The reallocations are based on the estimates of human capital with full controls, an elasticity of wages $\epsilon = 0.30$ and the unmeasured human capital share of $\gamma = 0$. Within countries, the values do not sum to zero because the miscellaneous sector is excluded from the table.