Specialization, Economic Development and Aggregate Productivity Differences

David Lagakos

Federal Reserve Bank of Minneapolis,
and Arizona State University

Michael E. Waugh

Federal Reserve Bank of Minneapolis,
and New York University

June 21, 2009

ABSTRACT

Cross-country labor productivity differences are large in agriculture and much smaller in non-agriculture. We argue that these relative productivity differences arise when subsistence consumption needs prevent workers in poor countries from specializing in the sector in which they are most productive. We formalize our theory by embedding the Roy (1951) model of ability into a two-sector general-equilibrium growth model in which the agents' preferences feature a subsistence food requirement. The model predicts that productivity differences in agriculture will be relatively larger than in non-agriculture, even though countries differ only in a sector-neutral efficiency term. A parameterized version of our model suggests that our theory is quantitatively important in explaining why agriculture productivity differences are so large relative to those in non-agriculture.

Email: lagakos@asu.edu, waugh@minneapolisfed.org. For helpful comments we thank Narayana Kocherlakota, Erzo Luttmer, Ellen McGrattan, Guillermo Ordoñez, Jim Schmitz and participants at the 2009 Midwest Macro Meetings. All errors are our own. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1 Introduction

Cross-country labor productivity differences are large in agriculture and much smaller in non-agriculture relative to aggregate differences (Caselli, 2005; Restuccia, Yang and Zhu, 2008). Development accounting exercises have shown that these sector productivity differences are key in accounting for aggregate productivity differences. If agricultural labor productivity were hypothetically raised to the U.S. level in every country, or if the share of labor in agriculture were hypothetically lowered to the U.S. level, then international variation in aggregate productivity would be virtually eliminated (Caselli, 2005). These results suggest that understanding productivity differences in agriculture and non-agriculture are at the heart of understanding world income inequality.

In this paper we provide a theory of relative labor productivity differences in the agriculture and non-agriculture sectors. We argue that these relative productivity differences arise when sector-neutral efficiency differences combine with subsistence food consumption needs to generate variation in the extent to which workers specialize in the sector where they are most productive.

The basic idea is that countries with low sector-neutral efficiency must deploy a large fraction of their workforce into the agriculture sector to satisfy subsistence food needs. As a result many of those working in agriculture are those whose comparative advantage is *not* in agricultural work, but rather in non-agricultural tasks such as writing newspaper articles, doing economic research, or teaching yoga classes. In countries with high sector-neutral efficiency a smaller fraction of workers are in agriculture, and those remaining in agriculture are those who are relatively most productive at farm work. As a result, physical productivity differences are relatively larger in agriculture than in the aggregate, and relatively smaller in non-agriculture.

We formalize our theory by embedding the Roy (1951) model of ability into a simple two-sector general-equilibrium growth model. Our theory has two main ingredients. First, workers are heterogenous in their ability to produce output in the two sectors. Second, preferences have a subsistence food requirement. Countries differ only in a sector-neutral efficiency term; preferences and the distribution of talents are identical across countries. The main qualitative prediction of the model is that productivity differences in agriculture are relatively larger than aggregate differences, and that non-agriculture productivity differences are smaller than in the aggregate. The novel feature of this result is it follows from optimal behavior only, as opposed to exogenous country-specific sectoral productivity differences, or barriers to agricultural production, as emphasized by other studies (e.g. Restuccia, Yang, Zhu, 2008).

We then parameterize the model to study its quantitative implications. To discipline the degree of worker heterogeneity, our calibration strategy uses observations on the distribution of wages to determine the distributions of ability in our model. For the non-agriculture talent distribution we use cross-sectional wage data from the United States, where virtually every worker is in non-agriculture. For the agriculture talent distribution we use estimates calculated from piece-rate wages for farmers. We discipline preferences by using estimates of subsistence food requirements, and shares of income devoted to food in the developed world.

Our main exercise is to allow sector-neutral efficiency to vary across countries and study the impled differences in agriculture and non-agriculture productivity. When we feed sector-neutral efficiency differences into our model generating a factor of 22 difference in aggregate income, which corresponds to the differences between the 90th and 10th percentile of country income distribution, the model predicts a factor of 28 gap in agriculture labor productivity, and a factor of 10 gap in non-agriculture labor productivity. In the data there is a factor of 45 gap in agriculture and a factor of 5 gap in non-agriculture. Thus, our model explains one-fourth of the productivity differences in agriculture and two-thirds of the productivity differences in non-agriculture between the 90th and 10th percentile of countries. We find that our model explains less of the sector productivity differences between the 90th percentile and countries with intermediate income levels, because differences in shares of labor in agriculture (and hence specialization differences) are smaller between these countries. We conclude that a completely frictionless economy with countries differing only by sector-neutral efficiency can generate meaningfully large differences in agriculture productivity and small differences in non-agriculture productivity between the richest and poorest countries.

We also show that the model performs well quantitatively in matching relevant development facts. Specifically, the model matches the cross-country data on the share of labor in agriculture, and it performs moderately well in matching agriculture's share in gross domestic product and replicating the fact that poor countries have higher relative prices for agriculture goods. Even though our model has no barriers to labor mobility across sectors, it generates gaps in average wages between agriculture and non-agriculture workers that are in line with the data. A failure of our model is its inability to generate a relative agriculture wage that decreases in income, as found in the data. We discuss how a simple modification of the model with schooling costs that decline with development, in the spirit of Caselli and Coleman (2001), can help improve our model's fit in this dimension.

We conclude by providing direct evidence that our mechanism was at work in the development experiences of the United States and Britain. Two dimensions in which sector abilities are (plausibly) observable are sex and age: historians and development economists have argued that women and children have a comparative disadvantage in agricultural work relative to adult men. Our theory thus predicts that, during a structural transformation, women and children leave farm work and enter the industrial sector at a faster rate than men. We cite evidence that this is fact what happened in Britain and the United States in the 18th and 19th centuries.

Our theory has new implications for the way economists think about agricultural productivity in the developing world. Concretely, our model suggests that low aggregate productivity is not *caused* by large fractions of workers working in the relatively unproductive agriculture sector. Instead, low measured productivity in agriculture and large agricultural labor shares are *consequences* of low sector-neutral productivity. The distinction is important because it helps determine the extent to which future research efforts on aggregate productivity differences should focus on the determinants of productivity in agriculture per se, as opposed to more general potential determinants. The policy implications between the two views are different as well. While accounting exercises suggest that fixing agriculture is crucial to raising overall productivity, our theory predicts that improvements in technology, institutions, or social infrastructure (Hall and Jones, 1999; Acemoglu, Robinson and Johnson, 2002) are the key to improving living standards.

2 Agriculture's Role in Accounting for Aggregate Productivity

In this section we highlight the important role of agricultural in understanding aggregate productivity. Specifically, we reproduce the findings of Caselli (2005) to illustrate how differences in labor productivity and shares of workers in agriculture account for much of the variation in aggregate output per worker across countries.

Table 1: Agriculture and Labor-Productivity Accounting

Panel A: Labor Productivity Differences

Sector	Ratio of 90th-10th Percentile
Aggregate	22
Agriculture	45
Non-Agriculture	4

Panel B: Percent of Labor in Agriculture

Country Income Percentile	Percent
90th	2.8
10th	78.3

Source: Caselli (2005)

Panel A of Table 1 shows that labor productivity differences in agriculture are larger than aggregate differences, and that non-agriculture productivity differences are much smaller. The ratio of agricultural output per worker in the 90th and 10th percentiles of the income distribution is 45, compared to just 4 in non-agriculture. As a frame of reference, the ratio for the aggregates is 22. Panel B summarizes the well-known fact that poor countries have a much larger fraction of their work force in agriculture. A country whose per-capita income is in the 90th percentile has just 2.8% of its workers in agriculture, while the 10th percentile country has 78.3% of its workers in agriculture.

These two facts together suggest that labor productivity differences in the aggregate are almost completely accounted for by differences in agricultural productivity and shares of labor in agriculture. Caselli formalizes this argument by computing the hypothetical variance of cross-country aggregate output per worker assuming that agricultural productivity in all countries were equal to the U.S. level. His answer is just a factor of 1.6, down from the actual factor of 22! In other words, international labor productivity differences would be virtually eliminated. A similar experiment computes the hypothetical variance of aggregate output per worker assuming all countries had the U.S. share of workers in agriculture. This experiment yields a factor of 4.2 differences between the 90th and 10th percentile, which again is vastly lower than the 22 seen in the data.

One potential explanation for labor productivity differences in agriculture is physical capital per worker differences across countries. Caselli argues that labor productivity differences almost entirely represent total-factor productivity (TFP) differences. As he puts it, "the factor-only model explains virtually nothing of the observed per-capita income variance in agriculture: it's entirely a story of TFP differences, even more so than for aggregate GDP." (Caselli, 2005, page 49.) In independent work, Chanda and Dalgaard (2008) perform a similar set a counterfactual exercises using capital stock data from agriculture and non-agriculture, and conclude that around 85% of international TFP differences can be accounted for by TFP differences in agriculture relative to non-agriculture.

The results of this section suggest that cross-country differences in agriculture productivity and the share of workers in agriculture are central factors in accounting for aggregate productivity differences. It remains an unanswered question *why* the agriculture sector exhibits so much more variation in productivity across countries than the non-agriculture sector.

3 A Theory of Relative Agriculture Productivity

In this section we develop a model of productivity differences in agriculture relative to non-agriculture. The main result is that sector-neutral efficiency differences across countries lead to relatively larger productivity differences in agriculture and relatively smaller differences in the non-agriculture sector than the efficiency differences themselves. Proofs of all results of are available in Appendix A.

3.1 Households

There are measure one of agents, indexed by i, who differ by ability, as will be explained below. Preferences are given by

$$U^i = \log(c_a^i - \bar{a}) + \theta \log(c_n^i), \tag{1}$$

where c_a^i is food consumption, c_n^i is non-food consumption, $\bar{a}$ is a parameter representing a subsistence food requirement, and θ governs the relative taste for non-food consumption.

Each agent is endowed with one unit of time which she supplies inelastically to the labor market. Each agent is also endowed with a vector of talents $\{z_a^i, z_n^i\}$ which represent the efficiency of one unit of labor in sectors a and n. The population density of talents is drawn from a distribution $G(z_a, z_n)$ with support on the positive reals, positive variance for each talent, and imperfect correlation between the two draws. Agents earn wage income w^i, which is described in more detail below. The budget constraint is

$$p_a c_a^i + c_n^i \leq w^i \tag{2}$$

where p_a is the relative price of food, and the non-agricultural good is taken as the numeraire.

3.2 Production

There is a competitive market in each of the two sectors, and each has its own sector aggregate production function. Both sector technologies are freely available and operated by competitive entrepreneurs. The technologies are given by:

$$Y_a = A\tilde{L}_a \quad \text{and} \quad Y_n = A\tilde{L}_n \tag{3}$$

in agriculture and non-agriculture, where A captures sector-neutral efficiency, and $\tilde{L}_a$ and $\tilde{L}_n$ represent the total number of effective labor units employed in the two sectors.

Let Ω^a and Ω^n denote the sets of agents electing to work in agriculture and non-agriculture. The sector aggregate labor inputs $\tilde{L}_a$ and $\tilde{L}_n$ are defined as

$$\tilde{L}_a \equiv \int_{i \in \Omega^a} z_a^i \, dGi \quad \text{and} \quad \tilde{L}_n \equiv \int_{i \in \Omega^n} z_n^i \, dGi$$

and represent the sum of all talent working in the respective sectors. Notice that our labor input differs from those of standard macro models in that ours sums up worker productivities, rather than workers themselves. The total number of workers in each sector is defined as

$$L_a \equiv \int_{i\in\Omega^a} dGi \quad \text{and} \quad L_n \equiv \int_{i\in\Omega^n} dGi.$$

3.3 Producer and Agent Optimization

Agents take as given prices and a wage schedule which maps talents into sector-specific wage offers. The problem for an agent is first to pick which sector to work in, and then to maximize (1) subject to (2).

Because of competition in production markets, the schedule of wages offered to a worker with talents z_a^i and z_n^i is equal to:

$$w_a^i = p_a A z_a^i \quad \text{and} \quad w_n^i = A z_n^i \tag{4}$$

in the agricultural and non-agricultural sectors. A simple cutoff rule in *relative* talent determines the optimal occupational choice for each agent. Working in non-agriculture is optimal for agent i if and only if

$$\frac{z_n^i}{z_a^i} \geq p_a. \tag{5}$$

Thus, the agents that enter non-agriculture are those whose talent there is sufficiently high relative to their talent in agriculture. Let the resulting wage under the optimal sector choice be defined as $w^i \equiv \max\{w_a^i, w_n^i\}$.

The remainder of the agent's problem is standard, and optimal demands are:

$$c_a^i = \frac{w^i + \bar{a} p_a \theta}{p_a(1+\theta)} \quad \text{and} \quad c_n^i = \frac{\theta(w^i - \bar{a} p_a)}{1+\theta}. \tag{6}$$

Due to the subsistence consumption constraints, agents consume relatively more food when their wage is lower.

An equilibrium of the economy consists of a relative food price, p_a, and allocations for all agents such that labor and output markets clear. Labor productivity in equilibrium is given by Y_a/L_a in agriculture, and Y_n/L_n, and represent the physical quantity of output produced per worker in each sector.

3.4 Productivity Differences Larger in Agriculture than Non-agriculture

In this section we illustrate the main result of the paper: for countries that differ in sector-neutral efficiency levels, agriculture productivity differences are larger than the efficiency differences, and non-agriculture productivity differences are smaller. We first establish that relative food prices decline in the efficiency level, which is important for the other results to follow.

Proposition 1 *Consider two economies, rich and poor, with efficiency terms A^R and A^P such that $A^R > A^P$. Then the relative price of agriculture is higher in the poor economy: $p_a^P > p_a^R$.*

To see the intuition for why p_a^P has to be higher than p_a^R, imagine in contradiction that they were the same. For expositional purposes, assume markets clear in the rich country. Then, by (5), the sector labor supply cutoffs would be the same in both countries, and hence so would the share of workers electing to supply labor in the agriculture sector. But because of the subsistence food requirement, the poorer economy demands a much larger fraction of food. Hence output markets would not clear in the poor economy. In order to induce enough workers to supply labor in agriculture in the poor economy, it must be true that p_a^P is greater than p_a^R.

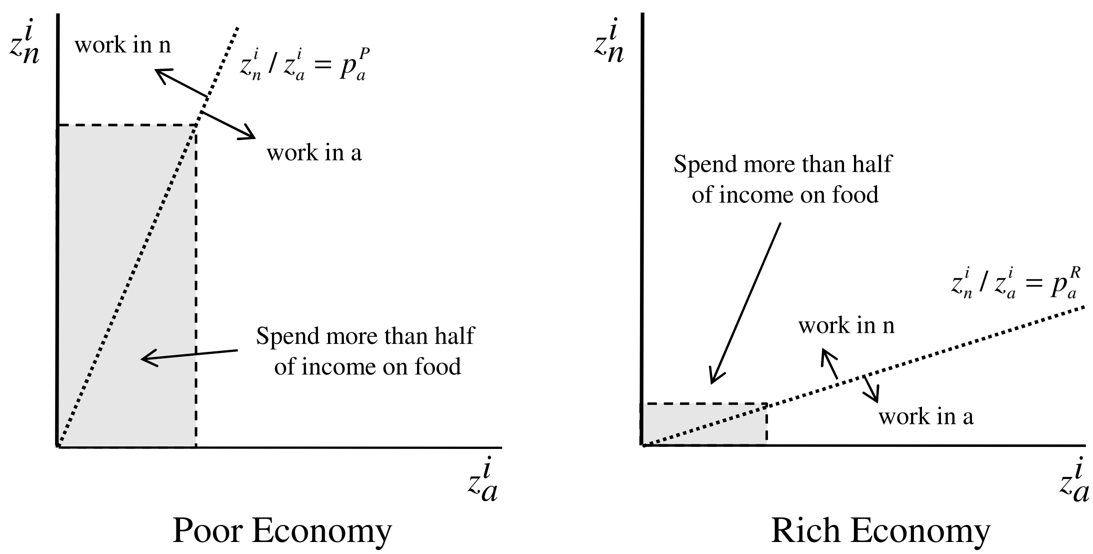

Figure 1: Optimal Sector Choice in Rich and Poor Economies

Figure 1 illustrates optimal sector choice in equilibrium. Each point on the figure represents one conceivable draw of (z_a, z_n), corresponding to a pair of sector-specific talents. The dotted lines stemming from the origin describe the set of talent pairs for which agents are indifferent between the two sectors, i.e. when z_n^i / z_a^i equals p_a^P and p_a^R respectively. Points above the lines represent agents for which working in sector n is optimal, and points below the lines meaning that working in a is optimal. As in Proposition 1, because $p_a^P > p_a^R$, more agents work in non-agriculture in the richer economy. The shaded regions describe the set of agents that spend more than half their income on food.[1] The poor economy has a larger fraction of such agents because of the subsistence food requirement. We can now establish the paper's main qualitative result.

Proposition 2 *Consider two economies, rich and poor, with efficiency terms A^R and A^P such that $A^R > A^P$. Then differences in agricultural labor productivity are larger than the efficiency differences, which are larger than non-agriculture labor productivity differences:*

$$\frac{Y_a^R / L_a^R}{Y_a^P / L_a^P} > \frac{A^R}{A^P} > \frac{Y_n^R / L_n^R}{Y_n^P / L_n^P}.$$

Figure 2 illustrates the intuition behind Proposition 2. The curved grey frontiers represent the production possibilities frontiers (PPF) in the rich economy and poor economy, with the rich economy's PPF is simply shifted out by a factor A^R / A^P from that of the poor economy. The equilibria of the two economies are depicted as points on the frontier, with the poor country choosing a relatively higher ratio of Y_a to Y_n, and the rich economy choosing a point with a lower ratio.

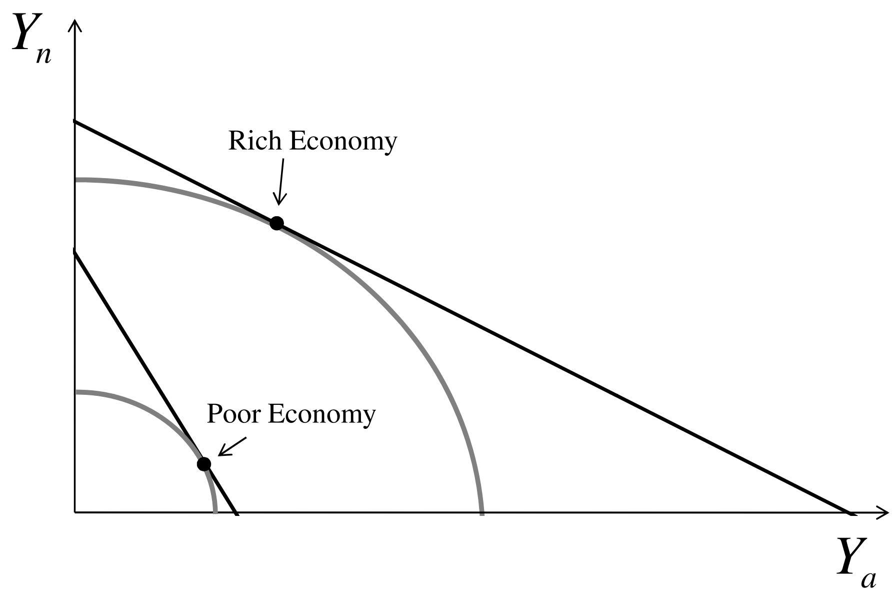

Figure 2: Equilibrium in Rich and Poor Economies

The fact that the PPFs are concave is key to understanding Proposition 2. The concavity of the PPFs arise because of the heterogeneity in worker ability. For points with a high ratio of

[1]The choice of one half income spent on food is arbitrary, and just meant to convey the higher food share in the poor country.

Y_a to Y_n, such as the poor economy's equilibrium, agriculture productivity is low because the high relative price of food assures that many workers with relatively low ability in agriculture choose to work in agriculture. On the other side, the high food price assures that the average worker in non-agriculture has relatively high ability, since by (5) only those with high absolute ability in non-agriculture enter that sector. Hence, non-agriculture productivity is relatively high. In the rich economy, in contrast, the lower relative food price assures that the average ability is higher in agriculture, and lower in non-agriculture, than in the poor country.

Note that for Propositions 1 and 2 to hold, worker heterogeneity and non-homothetic preferences are necessary. If we relax either assumption then neither relative prices nor sector productivity vary with A. With homogenous agents the PPFs would be linear, and only a relative price equal to one would prevail in equilibrium. Hence labor productivity would be unchanged in each sector as one moves along the PPF. With homothetic preferences (but heterogenous agents) the PPFs would be concave, but the optimal production bundles of poor and rich countries would lie along one ray from the origin. Since shifting out the frontiers in sector-neutral way would lead to the same PPF slope along any given ray, it follows that the rich and poor countries would have identical relative prices, and identical sectoral productivity.

4 Quantitative Analysis of the Model

In this section we assess the quantitative importance of the mechanism explained above for understanding agriculture and non-agriculture labor productivity differences across countries. We also assess the model's cross-country predictions for shares of labor in agriculture, shares of GDP in agriculture, relative prices of agricultural goods, and wages of agriculture workers relative to non-agriculture workers.

4.1 Parameterization

To parameterize the model we must select a distribution of talents, $G(z_a, z_n)$, plus values for the two taste parameters $\bar{a}$ and θ; the efficiency term A can be normalized to 1 for the United States.

For the talents we assume that the log of z_a and the log of z_a are each drawn independently from a generalized Pareto distribution.[2] We choose this function form for the flexibility it offers in matching second and third moments of the cross-sectional wage distribution, to which the

[2] We explore the case of non-zero correlation in Appendix A.

talent distribution we will parameterized. The talent distributions have probability density functions:

$$g(z_a) = 1 - (1 + \xi_a z_a / \sigma_a)^{-1/\xi_a} \quad \text{and} \quad g(z_n) = 1 - (1 + \xi_n z_n / \sigma_n)^{-1/\xi_n} \tag{7}$$

for $z_a > 0$, $z_n > 0$, $\sigma_a > 0$, $\sigma_n > 0$, $\xi_a \in \mathbb{R}$, and $\xi_n \in \mathbb{R}$. Roughly speaking, the scale parameters σ_a and σ_n control the variance of talents, while the shape parameters ξ_a and ξ_n control the skewness.

To select values for the non-agriculture talent distribution, we turn to the wage distribution in the United States, using data from the 2007 Current Population Survey (CPS) for non-farm workers. Because virtually all U.S. workers are in the non-agriculture sector, we can identify the non-agriculture talent distribution directly. We select ξ_n to match the skewness of the U.S. log non-farm wage distribution, and σ_n to match one half the standard deviation of log non-farm wages, giving us a standard deviation of non-agricultural abilities of 0.35.

Our assumption is that one half of log wage variation is due to ability differences across workers. Our choice is conservative given Card's (1999) assessment that schooling and work experience leave two-thirds of wage variation unexplained, even when attempts are made to correct for the endogeneity of schooling choice. Our standard deviation of ability is also in line with several estimates of ability heterogeneity estimated using structural models of schooling choice and labor earnings. Hendricks and Schoellman (2009) estimate an ability distribution with standard deviation 0.50, while the model of Belzil and Hansen (2003) implies an ability standard deviation of 0.31.

For the non-agriculture talent distribution, we use estimates of the ability distribution from Foster and Rosenzweig (1996), which are estimated using a structural model of farm activity choice and piece-rate wage data from farm workers in the Philippines. Foster and Rosenzweig (1996) estimate a standard deviation of farm worker ability of 0.085 in 1984. Since the Philippines in 1984 had 45% of its workers in agriculture, we choose σ_a to get an agricultural wage distribution with a standard deviation of wages equal to 0.085 for a country with 45% of workers in agriculture in our model. For lack of evidence on the skewness of the non-agricultural wage distribution, we set ξ_a to give the same skewness as the non-agriculture talent distribution.

Finally, for the preference parameters, we pick θ to give a long-run food expenditure share of 0.8%. This delivers a ratio of U.S. agricultural value added to GDP of 1.36, the 2007 U.S. figure. This choice is in line with other similar models of structural choice: Caselli and Coleman (2001) and Duarte and Restuccia (2008) pick a value of 1%, while Restuccia, Yang and Zhu (2008) pick a value of 0.5%. We set $\bar{a}$ to match a subsistence consumption need of 34% of average income in a model country with 7.5% of the U.S.'s per capita GDP. This is consistent with the independent estimates of subsistence food consumption requirements of Rosenzweig and Wolpin (1993), and

Atkeson and Ogaki (1996), both of which use panel data from a sample of rural housholds from India (which had 7.5% of the U.S. per capita GDP in 1984).[3]

4.2 Quantitative Predictions for Sector Productivity Differences

With the model parameterized, we now ask what it predicts quantitatively for agriculture productivity differences relative to non-agriculture differences in the cross section of countries. Specifically, we solve the model over a range of A values covering the world income distribution, and compute its predictions for relative output per worker in the aggregate and the two sectors.

Table 2: Labor Productivity Differences

Sector	Ratio of 90th-10th Percentile		
	Data	Model	Percent Explained
Agriculture	45	27.7	25
Aggregate	22	22	-
Non-Agriculture	4	10.4	66

Data Source: Caselli (2005)

Table 2 shows the model's predictions for the ratio of the 90th to 10th percentile of countries in the model and data. The differences in aggregate output per worker (expressed as GDP per worker at Gheary-Khamis international prices) is a factor of 22 in the model and data by construction. The model predicts agriculture output per worker differences should be a factor of 27.7, and in non-agriculture it predicts a factor of 10.4 difference. In the data these ratios are (as described in Section 2) a factor of 45 and 4 respectively. The third column of the table shows that this corresponds to the model explaining 25% of agricultural differences and 66% of non-agricultural differences, relative to aggregate differences. The model performs better in non-agriculture because the model's variance of talent is larger in non-agriculture than agriculture. The results in Table 2 show that differences in patterns of specialization are quantitatively important to understanding relative sector productivity differences between the richest and poorest countries.

Table 3 illustrates the model's predictions for developing countries with relatively higher average income. Specifically, it shows the model's prediction for the 90th-50th ratio and 90th-25th

[3]Rosenzweig and Wolpin estimate a subsistence requirement of 1,469 rupees per agent per year. Townsend (1994) reports that average agent size in the sample is 6.7 and that average income per person in the Indian sample is 635 rupees.

ratio. In the latter case aggregate productivity in the model and data differ by a factor of 9.4, again by construction. In the 90th-25th case, the model predicts a factor of 12.9 in agriculture and 7.5 in non-agriculture, compared to 31.1 and 2.7 in the data. The model predicts 16% and 28% of the agricultural and non-agricultural productivity differences, relative to the aggregate, as in the data, which is still large, but substantially lower than for the 90th-10th percentile ratio.

Table 3: Labor Productivity Differences – Intermediate Income Levels

Sector	Ratio of 90th.-25th Percentile		Percent Explained
	Data	Model	
Agriculture	31.1	12.9	16
Aggregate	9.4	9.4	-
Non-Agriculture	2.7	7.5	28
	Ratio of 90th-50th Percentile		
	Data	Model	
Agriculture	11.1	3.5	5
Aggregate	3.1	3.1	-
Non-Agriculture	1.9	3.0	10

Data Source: Caselli (2005)

In the 90th-50th case, the model fares much worse. The aggregate differences are chosen to be a factor of 3.1 as in the data. The model predicts differences in agriculture and non-agriculture of 3.5 and 3.0, compared to 11.1 and 1.9 in the data. This amounts to explaining just 5% and 10% of the sector differences, respectively.

Why does the model fare so much less successfully when explaining differences between rich countries and those at intermediate income levels? The answer has to do with differences in the shares of labor in agriculture. Consider the case of the 90-50 differential. The percent of workers in agriculture in the 50th percentile country is just 9%, compared to 2% in the 90th percentile country. Thus, the average productivity of workers in agriculture is only slightly lower in the 50th percentile country. In contrast, in the 10th percentile country, 74% are in agriculture, and thus the average worker has substantially lower productivity than the average agricultural worker in the 90th percentile country. Hence, the model's explanatory power is larger for differences between the the richest countries and the poorest countries than for the richest and middle income countries.

4.3 Assessing The Model's Other Quantitative Predictions

The model generates large productivity differences in agriculture relative to non-agriculture across countries, at least between the richest and poorest countries. We now ask whether it is successful in matching other relevant features of the cross-country data. In particular, we compute the model's predictions for the share of labor in agriculture, the share of GDP spent in agriculture, the relative price of agriculture, and the wages of agriculture workers relative to non-agriculture workers in the cross-section of countries, and assess whether they are quantitatively consistent with the data.

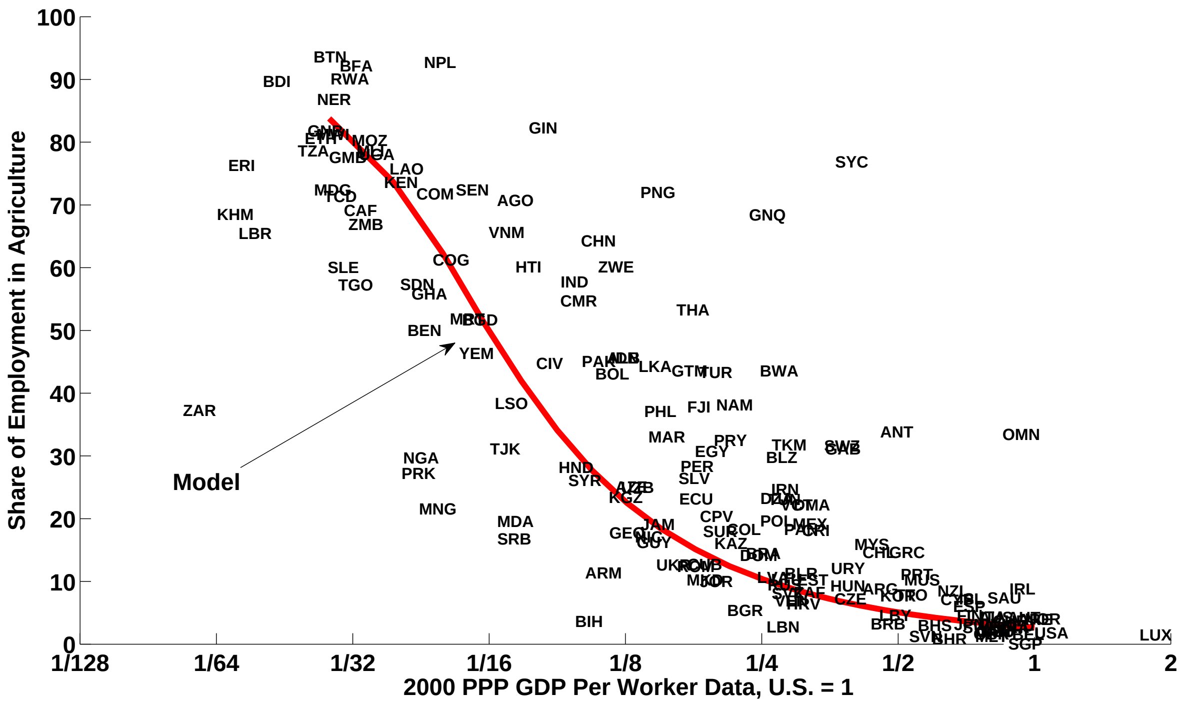

Figure 3: Shares of Labor in Agriculture in Model and Data

Figure 3 shows the model's predictions for the cross-section of countries for the share of labor in agriculture, along with the actual data. The horizontal axis displays purchasing-power parity GDP per worker for 2000 (on a log scale), and the y-axis displays the percent of workers in agriculture. As in the data, our model predicts that the poorest countries should have shares in the range of 70% to 90% of all workers, down to less than 10% for the richest. The model also captures the convex nature of this curve, which is driven in the model by the concavity of preferences along with the subsistence constraint in food. We conclude that this feature of the data is successfully captured by our model.

Next we turn, in Figure 4, to the model's predictions for the share of GDP in agriculture. While similar to the labor shares shown above, note that in the data the GDP shares in agriculture are

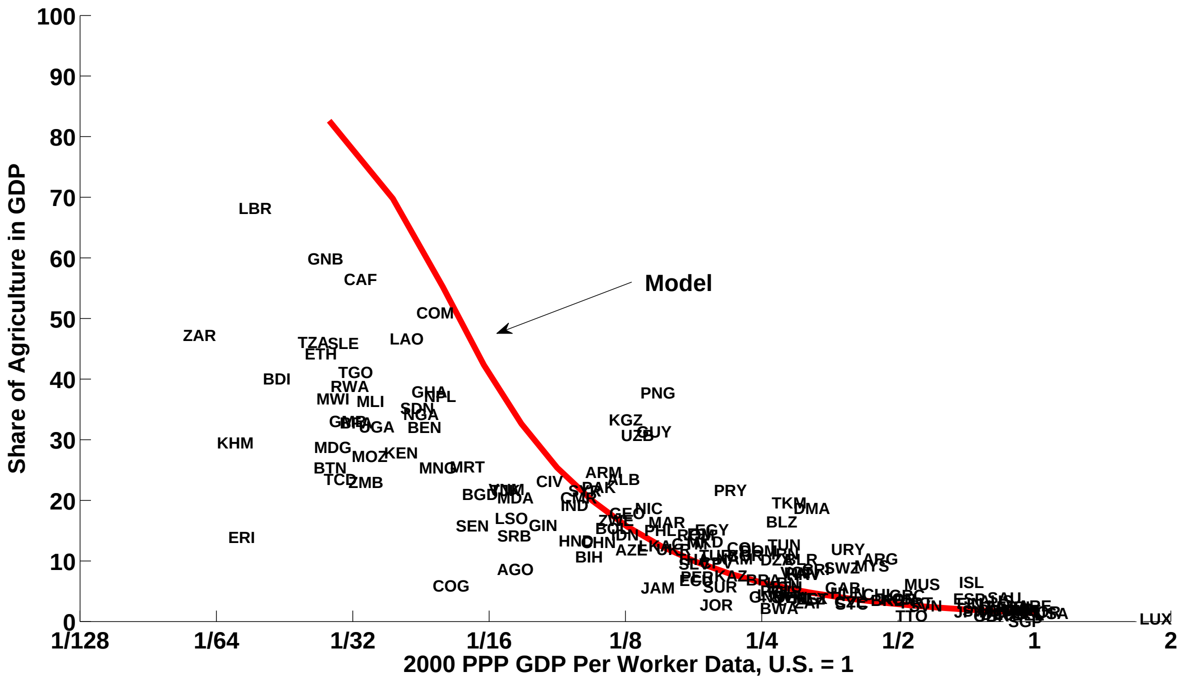

Figure 4: Agriculture Share in GDP in Model and Data

systematically lower than the labor shares in agriculture. In Kenya, for example, agriculture employs 74% of the workers but produces just 28% of GDP. While our model does a reasonable job of capturing this GDP shares for agriculture in the countries with around 1/8 the U.S. income level or higher, it substantially over-predicts the GDP share in agriculture in countries with lower income.

One reason for the model's inconsistencies with the data in this dimension may be that agricultural GDP itself is mis-measured in the poorest countries. If households spend much of their time in home production of agricultural goods, then measured GDP of agriculture will understate true agricultural output (Gollin, Parente, Rogerson, 2004). A second possibility is that the model over-predicts the relative prices of agriculture in the poorest countries, which would make the value of agricultural output higher than in the data. We explore this possibility next.

The relative food prices in the model and data are shown in Figure 5. The vertical axis contains the relative price of agricultural goods (expressed in log base 2) with the U.S. value normalized to one. The prices were constructed using 2005 data available from the Penn World Tables. As in Proposition 1, the model predicts that relative prices of food should be higher in poor countries than rich countries. This feature is in fact captured by the data as well. The 10th percentile country in the data has a relative price of food that is roughly 3.3 times as high as the relative price in the United States, and the trendline shows a clear decreasing relative price in GDP per worker. While the model is generally in line with the data, it over predicts relative

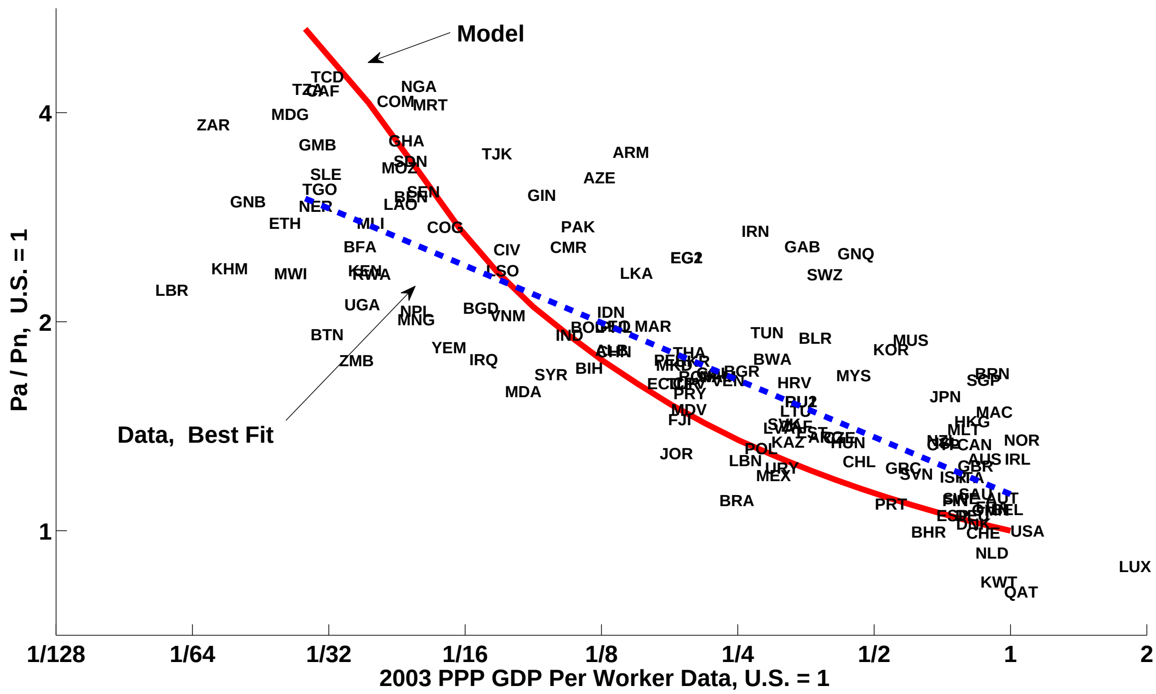

Figure 5: Relative Food Prices in Model and Data

prices in the poorest countries, with a relative price of around 5 in the 10th percentile country.

The final prediction we consider is the average wage in agriculture relative to non-agriculture. Figure 6 shows the model's predicted relative agriculture wage along with data on the ratio of nominal average wages in agriculture compared to nonagriculture in a set of countries, which we collected from the International Labor Organization (ILO).

Two features are of note in the data. First, the average wage of agriculture workers is lower than in non-agriculture, with an average ratio of around 0.6. Our model also produces an average of around 0.6, in line with the data. Our model's success in this dimension comes because the variation in agriculture ability is smaller than the variance in non-agriculture ability. Thus, the average worker in non-agriculture has higher ability relative to the mean of the non-agriculture distribution than the average agriculture worker does relative to the agriculture mean. This is perhaps easiest to see in the case where all workers have identical ability in agriculture (say equal to one). In this case, the average wage in agriculture will be p_a. Those choosing non-agriculture will be, by (5), those whose non-agriculture ability is at least as high as p_a.

The second noteworthy aspect of the figure is that the relative agriculture wages increase in GDP per capita. In principle some of this upward slope could be since the wages are nominal, and there is evidence that the cost of living is lower in rural areas, particularly in less-developed countries, where transportation costs are highest. Our model predicts a downward slope for

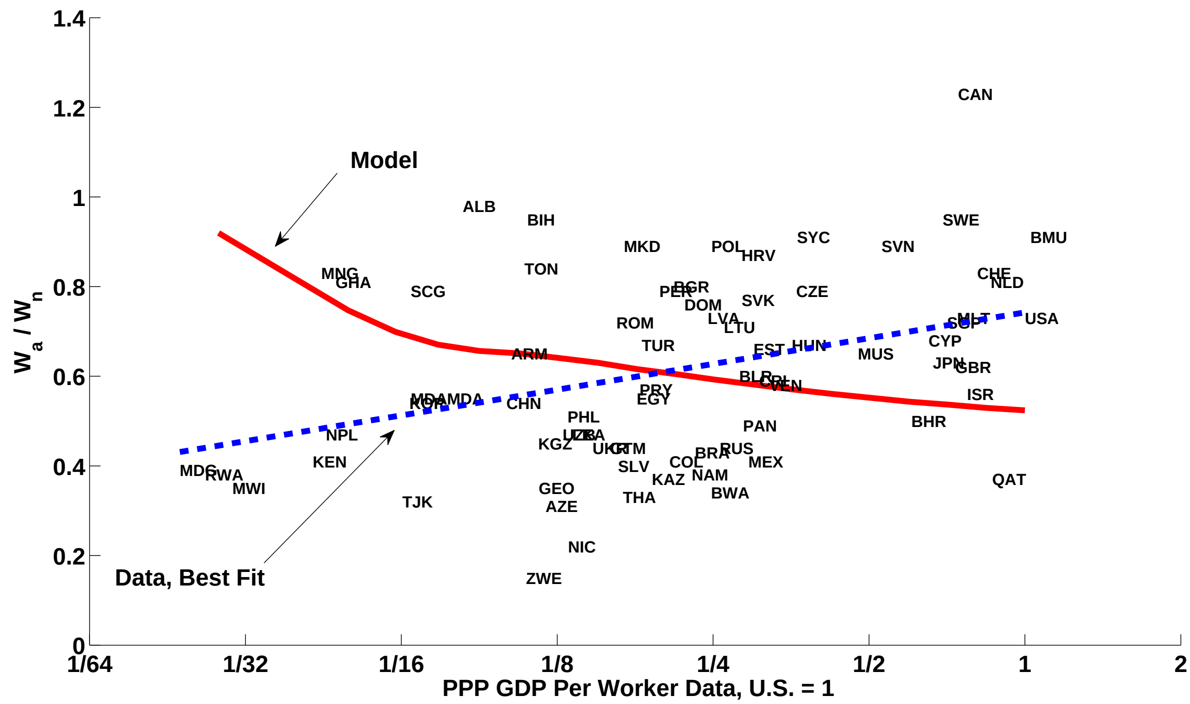

Figure 6: Average Wages in Agriculture Relative to Non-agriculture in Model and Data

relative wages with GDP per capita, in contrast to the data. This is driven mainly by the decreasing relative food prices in our model, which lowers the wage of all agricultural workers. In the next section we explore potential ways to amend the model to be more in line with these data.

5 Fixing the Model's Limitations

In this section we discuss possible ways to improve the model's predictions for relative prices and relative wages. We first discuss how the model behaves once we add the possibility of trade, and second, we explore an extension which allows for the possibility of schooling.

5.1 Adding Trade to the Model

In this section we ask how allow the model's predictions would change once we allow for trade. We draw two conclusions. First, allowing for frictionless trade would introduce strongly counterfactual assumptions about shares of labor in agriculture and relative prices across countries. Thus, to be useful, any extension to allowing trade should have to include trade frictions. Second, adding trade with plausible frictions could help reconcile the model's shortcomings with respect to relative prices and wages. Yet, we conjecture that such a model would have a modest effect on the *quantitative* nature of the model's predictions.

First consider a version of the model where each country has frictionless access to trade in world markets. Then the following is true.

Proposition 3 *Imagine that the rich and poor economies could buy or sell as much as they wanted on world markets at a relative food price p_a^W. Then the following must hold:*

$$\frac{Y_a^P/L_a^P}{Y_a^R/L_a^R} = \frac{Y_n^P/L_n^P}{Y_n^R/L_n^R} \quad \text{and} \quad \frac{L_a^R}{L_a^R + L_n^R} = \frac{L_a^P}{L_a^P + L_n^P}. \tag{8}$$

Proposition 3 says that under frictionless trade, two things are true. First, the extent of specialization would be the same in both countries, and hence labor productivity differences between the rich and poor countries would be the same in agriculture and non-agriculture. Second, the shares of labor in agriculture would be equated across countries. Both are true because, under a common relative price, the sector labor supply cutoff (5) is identical in both countries, and hence the composition of workers in each sector are identical as well.

But the prediction of labor shares being equal across countries is strongly counterfactual. As is well known, a substantially higher fraction of labor in poor countries is in agriculture than rich countries. In Section 2, for example, we cite evidence that the United States has just 2.8% of its labor in agriculture compared to over 78.3% of the labor in a country at the 10th percentile of the world per-capita income distribution. Thus, we conclude that adding frictionless trade will substantially reduces the quality of the model's predictions.

We claim that adding frictional trade to the model could change its qualitative predictions in three important ways. First, it would flatten out the model's predictions for relative food prices across countries, since trade would tend to lessen international price variation. Second, it would lower the model's predictions for the importance of specialization differences in explaining sector productivity differences, i.e. those in Table 2, since less price variation would lead to smaller specialization differences across countries. Third, it would decrease the extent to which the model's relative agriculture wage declines with income, since food prices would now be flatter in income.

Nevertheless, we conjecture that adding frictional trade would have a modest quantitative effect on the model's predictions in these three dimensions. The reason is that any frictions added to the model would have to be in line with relative food prices across countries, yet the baseline model is not too far away from the data in terms of prices. Thus the model's predictions for relative agriculture productivity would be changed little, as would the relative wage of farm workers. In future work we plan to explore this conjecture more fully.

5.2 Adding Schooling to the Model

Another potential way to improve the model's predictions, particularly for relative farm wages, would be to allow for the possibility of schooling. In the current model, as a country's efficiency increases, the sole reason labor moves from agriculture to non-agriculture is that the price of agriculture falls. Caselli and Coleman (2001) offer an additional factor that induces workers to leave agriculture, namely, a falling price of eduction, which increases the prospect of leaving agriculture to become a skilled non-agriculture worker. Caselli and Coleman argue that in the United States, the falling cost of education was a major factor leading to reallocation of workers into non-agricultural work. Furthermore, they show that without a falling education cost, standard models of structural change cannot reconcile the rising relative wages of non-agriculture workers which occurred over the late 19th and 20th centuries.

This finding suggests that allowing for the possibility of schooling in our model could help fix the model's counterfactual prediction that wages in agriculture decrease relative to non-agriculture in GDP per capita. Specifically, we could allow for schooling that increases a worker's effective number of labor units in non-agriculture work at some cost, where the cost of that schooling declines with general efficiency, A. Amending the model in this direction could also help fix the model's relative price predictions. As general efficiency increases, two forces would now induce workers out of agriculture: the falling relative price of food and the decreasing cost of improving one's productivity in non-agriculture. With the latter effect in the mix, the relative price variation across countries would now have to be smaller in order to reconcile differences in the share of workers in agriculture comparable to those in the current model.

6 Historical Evidence: Males versus Females

In this section we provide some direct evidence in support of our theory. While many dimensions of ability heterogeneity are not observable, along two particular observable dimensions, namely age and sex, there is concrete evidence of ability differences in agricultural and non-agricultural tasks. Historians and development economists have argued that women and children generally have a comparative disadvantage at farm work than adult men. As one piece of evidence, Goldin and Sokoloff (1982, 1984) show that wages were much lower for women in farm work in the United States, earning roughly one third to one half as much as men in farm work in the nineteenth century. Foster and Rosenzweig (1996) provide complementary evidence from a sample of farmers from the Philippines, among which they estimate males to have an absolute productivity advantage in several types of agricultural tasks.[4]

[4] They estimate a one-factor model with two tasks: ploughing and weeding. They find that men had an absolute advantage in both and a comparative advantage in ploughing.

Our theory thus predicts that women and children should have been the first to move off the farms and into non-agricultural work. This is in fact what happened. In Britain, according to Allen (1994), the fraction of farmers that were women or children declined substantially during Britain's industrial revolution. Table 4 shows Allen's calculations for the composition of farm workers in England and Wales between 1700 and 1851. In 1700, a full 62.0% of farm workers in England were women and children, with the balance adult men. By 1800 this percent fell to 55.3%, and by 1851 it was down to 36.3%.

Table 4: Composition of English Farm Workers

	1700	1800	1851
Men	38.3	44.7	63.7
Women and Children	62.0	55.3	36.3
Total	100.0	100.0	100.0

Data Source: Allen (1994)

Goldin and Sokoloff (1982, 1984) provide evidence that this pattern held in the United States as well, using evidence from the manufacturing sector, which was a major component of the non-agricultural economy in the nineteenth century. In 1820, in the Northeast United States, roughly 55% of manufacturing workers were women and children. By 1890, this figure was down to 21%. The interpretation given by Goldin and Sokoloff is that as manufacturing work became available, women took manufacturing jobs at a faster rate than men, who stayed in agriculture work relatively longer.

Furthermore, Goldin and Sokoloff argue that the primary reason women moved into manufacturing relatively faster than men is that women had a comparative disadvantage at agriculture work, just as our theory predicts. To support their argument, Goldin and Sokoloff estimate that in 1820, women earned roughly 30% as men in the Middle Atlantic region, and roughly 37% as much as men in New England. By 1850, they estimate relative wages of 51% in the Middle Atlantic and 46% in New England. While numerous factors were at play in this period, the authors argue that their finding of rising female wages "is consistent with the observations of many contemporaries of the early nineteenth century who reported that the relative productivity (and wages) of women and children compared to adult men was low in the agriculture and traditional sectors of the pre-industrial northeastern economy (1982, page 759)."

As additional support for the comparative advantage theory, Goldin and Sokoloff provide evidence that women faced greater comparative disadvantage in farming in the North than in the South, and entered manufacturing to a much greater extent in the North than in the South. The difference in the comparative disadvantage of women stemmed from the types of farm work common in the two regions. In the North, where strength-intensive wheat farming was prevalent, women earned around one third as much as men in the 1820. In the South, where Cotton and Tobacco farming were most common, women earned around one half as much as men, as dexterity played a more important role in farming these crops. Just as the theory predicts, as the U.S. stuctural change progressed in the second half of the 19th century, Northern women entered into factory work to a much larger extent than those in the South.

7 Effects of Policies That Restrict Labor Mobility

Our model implies that government policies that restrict migration within a country are likely to reduce productivity. One component of these policies that has received relatively less attention is the government's role in selecting *which* agents may move and which may not. Our theory says that letting the market allocate workers is more efficient than letting the government allocate them.

China is an important example of a growing country that is underdoing structural change and has restricted the flow of workers out of agricultural areas, typically in central China, into urban areas, typically on the eastern coast. One component of this *hukou* system is that the local and federal government officials decide which workers are permitted to migrate and which must remain in rural areas (See e.g. Au and Henderson (2006) and Lau, Qian, and Roldand (2000).)[5]

In future work we plan to use the model estimate the potential welfare cost of policies that restrict which individuals may leave agriculture and which may not. A rough idea is as follows. Imagine an economy like China in 1985 with 70% of its workforce in agriculture, and imagine it moves to a new steady state in which GDP per worker is (say) doubled. This corresponds roughly to Chinese growth from 1985 to the present. We can compute the welfare costs under two scenarios which restrict migration out of agriculture. The first randomly picks some fraction ω of agricultural agents that may move from agriculture to non-agriculture in such a way that markets clear, and the second lets the market reallocate agent. The experiment is in the same vein as that of Alder (2009), who assesses the efficiency losses from non-assortative matching of managers to projects.

[5] In addition to policies restricting which workers leave agriculture, there is evidence of social norms which restrict worker movements across sectors. Hayashi and Prescott (2008) cite evidence of social customs in the period before World War II which kept first born sons of farmers in agriculture.

8 Conclusion

We argue that cross-country productivity differences in agriculture are larger than in non-agriculture because of differences in the extent to which workers specialize in sectors in which they are talented. In poor countries, virtually everyone works in agriculture, even though many of those workers have a comparative advantage that is *not* in farm work, but rather in non-agricultural tasks such as acting, teaching, or writing newspaper articles. In rich countries, in contrast, those remaining in agriculture are those who are relatively most productive at farm work. As a result, labor productivity differences are relatively larger in agriculture than the aggregate, and smaller in non-agriculture, even though countries differ only in general, sector-neutral, efficiency.

Our theory has new implications for the way economists think about agricultural productivity in the developing world. In contrast to other papers that emphasize barriers to efficient production in farming, we argue that low productivity in agriculture could represent the optimal response to low general efficiency in the face of subsistence food requirements. In this case it is optimal to employ many workers in agriculture who are less talented in farm labor than other tasks. Concretely, our paper suggests that the source of low agriculture productivity might not be entirely found in the agriculture sector itself. It could, for example, be due to weak institutions, poor protection of property rights, or poor social infrastructure, as emphasized by a growing macroeconomics literature (e.g. Hall and Jones, 1999; Acemoglu, Robinson and Johnson, 2002).

A Model Appendix

A.1 Proof of Proposition 1

Let p_a^1, Y_a^1 and Y_n^1 be the equilibrium relative price and quantities in an economy with general efficiency A^1. Let $A^2 > A^1$, and denote by p_a^2, Y_a^2 and Y_n^2 the equilibrium of an economy with efficiency A^2.

Suppose that $p_a^2 = p_a^1$. Then by (5), each agent i chooses to work in the same sector in A^2 as in economy A^1. Thus output in each sector would be scaled up by a factor equal to the ratio of the efficiency terms: $Y_a^2/Y_a^1 = Y_n^2/Y_n^1 = A^2/A^1$. But by (6), we know that agents must demand a higher fraction of non-agriculture goods in economy A^2 than A^1. Thus $Y_n^2/Y_a^2 > Y_n^1/Y_a^1$. But this implies that $Y_n^2/Y_n^1 > Y_a^2/Y_a^1$, which is a contradiction. Thus $p_a^2 \neq p_a^1$.

The only way to be consistent with the agent solutions', (6), is for more agents to supply labor in the non-agriculture sector in economy A^2 than economy A^1. By (5), this occurs if and only if $p_a^2 < p_a^1$. ■

A.2 Proof of Proposition 2

Define $\rho(A)$ to be the relative productivity of agriculture to non-agriculture in an economy with efficiency A, i.e. $\rho(A) \equiv \frac{Y_a}{L_a} / \frac{Y_n}{L_n}$. It suffices to prove that $\rho'(A) > 0$. Recall that by (5), a agent with non-agriculture productivity z_n works in agriculture if and only if $z_a > z_n/p_a$, and that a agent with agriculture productivity z_a enters non-agriculture if and only if $z_n > z_a p_a$. Thus we can write

$$\rho(A) = \left(\frac{\int_0^\infty \int_{z_n/p_a}^\infty z_a \, g(z_a, z_n) \, dz_a \, dz_n}{\int_0^\infty \int_{z_n/p_a}^\infty g(z_a, z_n) \, dz_a \, dz_n} \right) / \left(\frac{\int_0^\infty \int_{z_a p_a}^\infty z_n \, g(z_a, z_n) \, dz_n \, dz_a}{\int_0^\infty \int_{z_a p_a}^\infty g(z_a, z_n) \, dz_n \, dz_a} \right).$$

By Proposition 1, we know that p_a is decreasing in A, and hence the left-side ratio is increasing in A, and the right-side ratio is decreasing in A. Thus $\rho'(A) > 0$. ■

A.3 Model's Predictions under Correlated Productivity Draws

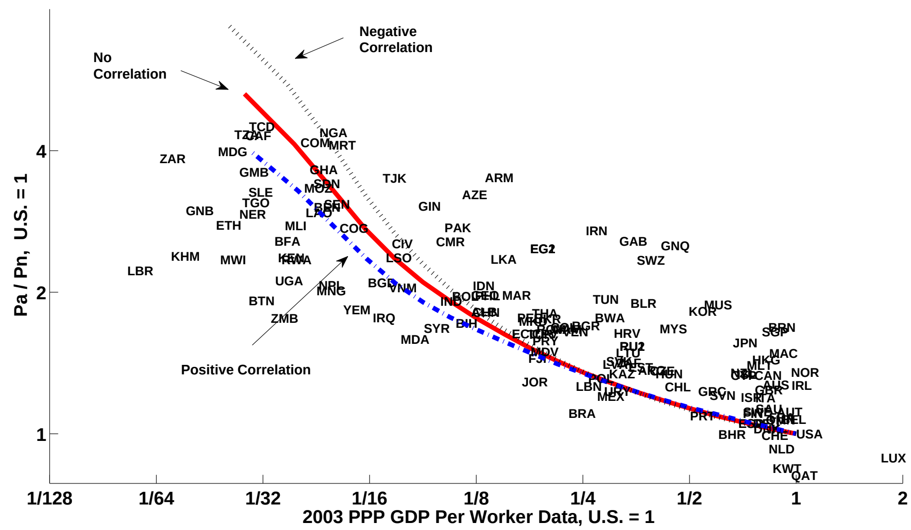

Figure 7: Relative Wage Predictions with Correlated Productivity Draws

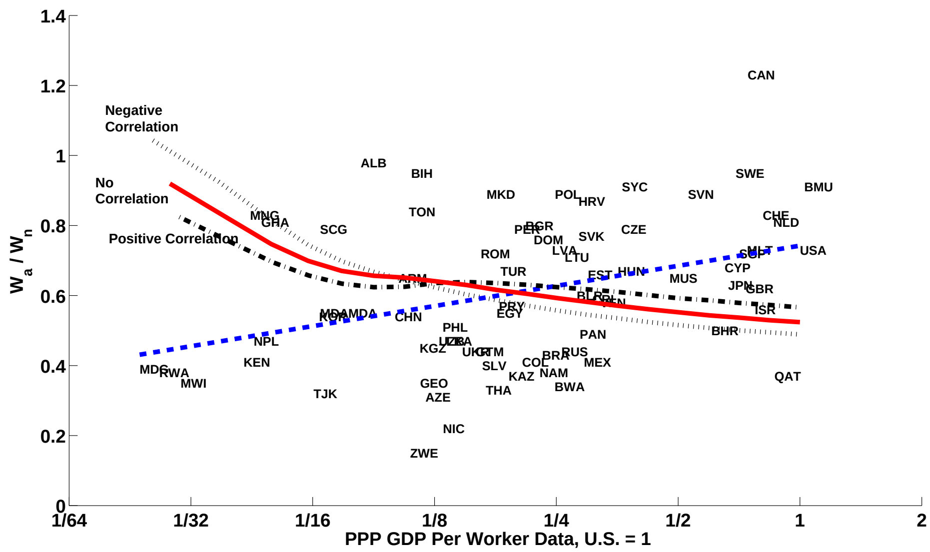

Figure 8: Relative Wage Predictions with Correlated Productivity Draws

B Data Appendix

- **GDP Per Worker** — This data is from the Penn World Table version 6.2. series "rgdpch".
- **Labor Share in Agriculture** — This data comes from Table A.3 in the FAO Statistical Yearbook 2004 online edition.
- **Agriculture Share in GDP** — This data comes from Table G.1 in the FAO Statistical Yearbook online edition.
- **Relative Agriculture Prices** — This data is derived from author's calculations with original data from the World Bank's 2005 International Comparison Program online database. The sector "agriculture" is defined to be food and non-alcoholic beverages, alcoholic beverages and tobacco codes (1101 and 1102). "Non-agriculture" is defined as all individual consumption code (11) minus food, non-alcoholic beverages, alcoholic beverages and tobacco.
- **Relative Wages** — Wage data is from LABORSTA and the series wage data by economic activity is used. Agriculture corresponds directly with ISIC Revision 2 and 3 categories "Agriculture, hunting and forestry" and "Fishing".

References

Acemoglu, Daron, Simon Johnson, and James Robinson. 2002. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics* 117 (4): 1231–94.

Alder, Simeon D. 2009. "In the Wrong Hands: Complementarities, Resource Allocation, and Aggregate TFP."

Allen, Robert. 1994. "Agriculture During the Industrial Revolution." Chapter 5 of *The Economic History of Britain Since 1700*, edited by Roderick Floud and Donald McCloskey, Volume 1, 96–122. Cambridge University Press.

Atkeson, Andrew, and Masao Ogaki. 1996. "Wealth-Vary Intertemporal Elasticities of Substitution: Evidence from Panel and Aggregate Data." *Journal of Monetary Economics* 38:507–534.

Au, Chun-Chung, and J. Vernon Henderson. 2006. "How migration restrictions limit agglomeration and productivity in China." *Journal of Development Economics*.

Belzil, Christian, and Jorgen Hansen. 2002. "Unobserved Ability and the Return to Schooling." *Econometrica* 70 (5): 2075–2091.

Card, David. 1999. "The causal effect of education on earnings." Chapter 30 of *Handbook of Labor Economics*, edited by O. Ashenfelter and D. Card.

Caselli, Francesco. 2005. "Accounting for Cross-Country Income Differences." In *Handbook of Economic Growth*, edited by P. Aghion and S. Durlauf.

Caselli, Francesco, and Wilbur John Coleman II. 2001. "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation." *Journal of Political Economy* 109 (3): 584–616.

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.

Foster, Andrew D., and Mark R. Rosenzweig. 1996. "Comparative Advantage, Information and the Allocation of Workers to Tasks: Evidence from Agricultural Labor Markets." *Review of Economic Studies* 63 (3): 347–374.

Goldin, Claudia, and Kenneth Sokoloff. 1982. "Women, Children, and Industrialization in the Early Republic: Evidence from the Manufacturing Censuses." *Journal of Economic History* 42 (4): 741–774.

———. 1984. "The Relative Productivity Hypothesis of Industrialization: The American Case, 1820 to 1850." *Quarterly Journal of Economics* 99 (3): 461–487.

Gollin, Douglas, Stephen L. Parente, and Richard Rogerson. 2004. "Farm work, home work and international productivity differences." *Review of Economic Dynamics*, vol. 7.

Hall, Robert E., and Charles I. Jones. 1999. "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics* 114 (1): 83–116 (February).

Hayashi, Fumio, and Edward C. Prescott. 2008. "The Depressing Effect of Agricultural Institutions on the Prewar Japanese Economy." *Journal of Political Economy*, vol. 116.

Hendricks, Lutz, and Todd Schoellman. "Student Abilities During the Expansion of U.S. Education, 1950-2000."

Lau, Lawrence J. 2000. "Reform without Losers: An Interpretation of China's Dual-Track Approach to Transition." *Journal of Political Economy*, vol. 108.

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

Rosenweig, Mark R., and Kenneth I. Wolpin. 1993. "Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investments in Bullocks in India." *Journal of Political Economy* 101:223–244.

Roy, A. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3:135–46.

Townsend, Robert M. 1994. "Risk and Insurance in Village India." *Econometrica*, vol. 62 (3).