Risk, Selection and Productivity Differences

Wenbiao Cai

University of Winnipeg

October 4, 2015

Abstract

Cross-country productivity difference in agriculture is 10 times larger than that in nonagriculture. At the same time, wage in agriculture relative to nonagriculture also increases with income per worker. This paper proposes and quantifies a theory of occupational choice under uncertainty that is able to reconcile these two observations simultaneously. The theory embeds the Roy model of self-selection into a Harris-Todaro model of migration. I show that employment risks present in nonagriculture affect the pattern of specialization and generate wage wedge between agriculture and nonagriculture, in ways that are consistent with data.

JEL Codes: H23; L52; O44;

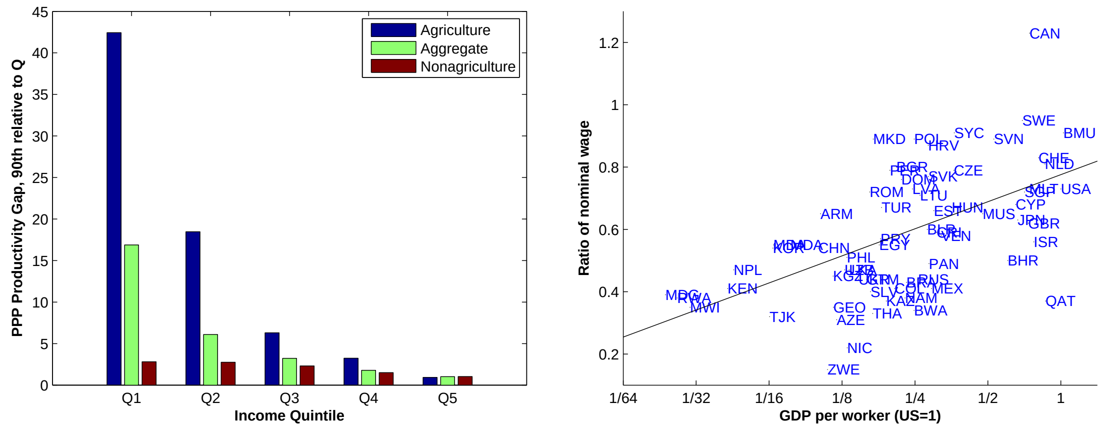
Keywords: Income Differences; Roy model; Harris-Todaro

Preliminary and incomplete. Comments welcome. Email: we.cai@uwinnipeg.ca. I would like to thank Manish Pandey for many helpful discussion.

1 Introduction

Two puzzles about cross-country productivity differences are (1) productivity difference is 10 times larger in agriculture than in nonagriculture. Between the richest 10 percent of countries and the poorest 10 percent of countries, the gap in PPP output per worker is a factor of 45 in agriculture, compared to only a factor of 4.5 in non-agriculture. See panel (a) of Figure 1; (2) within a country, nominal wage is lower in agriculture than in nonagriculture, and more so in poor countries. See panel (b) of Figure 1.

Figure 1: Cross-Country Productivity Differences



(a) Productivity gap by income quintile (b) Nominal wage, agr./non-agr.

Understanding these puzzles is an important step towards answering a fundamental question in economics: why are some countries richer than others. Caselli (2005) shows, by means of an accounting exercise, that low aggregate income in poor countries is primarily driven by high employment in unproductive agriculture. These observations lead to the question I ask and answer in this paper: why is there a large employment and simultaneously low productivity and low wage in agriculture of poor countries? My theory highlights the importance of income risk and its interaction with aggregate TFP, in shaping employment, wage and labor productivity in agriculture and nonagriculture.

I consider a model economy that is populated by a continuum of individuals with two-dimensional skills: one determines their productivity in agriculture and the other one nonagriculture, similar to that in Roy (1951). The level of development for the economy is characterized by its total factor productivity (TFP), which is common to both agriculture and nonagriculture. On the preferences side, the utility function is concave and features a minimum consumption of goods produced in agriculture. Therefore, a poor country with low TFP is subject to what Schultz (1964) describes as the "food problem", and needs a large share of labor in agriculture to produce this minimum level of agricultural output. As more individuals move into agriculture, average agricultural skills reduce while those remain in nonagriculture have a stronger comparative advantage in nonagricultural production. As a result, for a given difference in TFP between a rich economy and a poor one, the difference in agricultural skill amplifies TFP difference into a larger gap in output per worker in agriculture. In nonagriculture, the difference in nonagricultural skill compensates for TFP difference and generates a smaller gap in output per worker.

This aspect of the model has been illustrated in Lagakos and Waugh (2013). They show that such a model is capable of generating larger PPP productivity difference in agriculture than in nonagriculture, and therefore is able to explain the first puzzle. However, I show that the benchmark model in Lagakos and Waugh (2013) could not account for the second puzzle, i.e., nominal wage is lower in agriculture than nonagriculture and the gap decreases with income per worker. The reason is that absent frictions, average wage between agriculture and non-agriculture equalize in a competitive equilibrium. Therefore, the model generates a counterfactual prediction that ratio of wage is roughly constant across levels of development.

I extend the benchmark model in Lagakos and Waugh (2013) by adding uncertainty in earnings from nonagriculture. Specifically, individuals who work in nonagriculture are subject to idiosyncratic shocks that affect their take-home income. In contrast, working in agriculture yields a deterministic income. This set-up is similar in spirit to Harris and Todaro (1970), who argue that employment risks in the urban sector is driving the observed rural-urban wage gap. The difference is that in my model, uncertainty operates through the intensive margin as random shocks to earnings, instead of the extensive margin as the probability of employment in the original Harris-Todaro paper. I restrict the income shocks to have a zero mean, therefore there is aggregate loss in income. I further assume that there is no market for state-contingent contracts, and hence individuals cannot insure against this risk. In this environment, the division of labor takes a different form than that in Lagakos and Waugh (2013). Conditional on skill endowment, individuals choose a sector that yields *higher expected utility*, not higher expected earnings.

I first characterize the effect of uncertainty on equilibrium division of labor and relative wage. Two implications are immediate. First, holing prices constant, uncertainty induces a larger share of labor in agriculture. The reason is that individuals who are indifferent between sectors previously will strictly prefer agriculture now, because they are risk averse and the same expected income in nonagriculture generates strictly lower expected utility. The same logic would imply the second implication, that average wage is lower in agriculture compared to non-agriculture. Since individuals equalize expected utility between the two sectors, nonagriculture commands a premium in earnings relative to agriculture in order to compensate for the bearing of income risks.

Next I show that the effect of uncertainty interacts with the TFP and preferences in a way that enables the model to simultaneously account for both productivity puzzles. The subsistence term in the utility function implies that individuals are more risk averse when TFP (and hence income) is lower. That is because a negative shock to income carries a larger loss in utility when consumption is dangerously close to subsistence. Therefore, working in nonagriculture and being subject to earnings shocks in poor countries requires a bigger premium in wage relative to agriculture. As a result, the model generates wage gaps between agriculture and nonagriculture that decrease with income per worker.

The same mechanism also operates on individuals with different endowments of skills. In particular, individuals with low skills (both agricultural and non-agricultural) will be *more* hesitant to move out of agriculture unless their comparative advantage in nonagricultural production is strong enough. Therefore, within an economy uncertainty acts to increases average skill in nonagriculture, and reduces average skills in agriculture. Now compare an economy with high TFP and another one with low TFP. The effect of income risks is perceivably stronger in the low TFP economy, because again individuals (especially those with low skills) are consuming close to subsistence, while in the high TFP economy even the least skillful individuals are consuming safely above subsistence. In this way, income risks reduce average agricultural skill and increase average nonagricultural skill *more* in the low TFP economy. Consequently, for a given difference in TFP between rich and poor countries, income risks present *further* amplifies difference in agricultural skills and *further* shrinks difference in nonagricultural skill. This feature of the model helps generate a larger difference in output per worker in agriculture and a smaller difference in nonagriculture, through an additional mechanism that is absent in Lagakos and Waugh (2013).

To assess quantitative implications from the model, I first calibrate the model to U.S. data. I assume that skills have a marginal distribution that is Fréchet, and allow for correlation between different skills. I approximate the income shocks by a normal distribution with a zero mean. These distributional parameters and the subsistence parameter in preferences are chosen such that the model matches the share of labor in agriculture, the variance of wage in agriculture, and the variance of the permanent and transitory components of wage in nonagriculture.

I use the model to study productivity differences between rich and poor countries. Countries are identical except for their levels of TFP. In particular, for the time being I restrict the process of income shocks to be the same across countries. To identify the level of TFP for each country, I pick the level of TFP such that the model generates aggregate income per worker relative to the U.S. that is the same as in the data. In other words, the model by construction replicates the world income distribution. I then assess other predictions from the model, and begin with productivity by sector. Between the 90th percentile and the 10th percentile country, the model generates a factor 32 difference in output per worker in agriculture, compared to 45 in the data. For nonagriculture, the model generates a factor 12 difference in output per worker, compared to 4.5 in the data.[1] At the same time, the model predicts that the relative wage of agriculture (to nonagriculture) increases from 0.38 in the poorest 10 percent countries to 0.67 in the richest countries. Put differently, in the data the correlation between income per worker and the relative wage is 0.62.

[1]The respective number reported in Lagakos and Waugh (2013) is 29 and 13.

The model generates a correlation that is 0.68. In sum, the model is able to account for simultaneously the two puzzles about international productivity differences outlined at the beginning of this paper.

The model also generates an array of predictions that are quantitatively consistent with data. For the share of labor in agriculture, the model predicts the it decreases from 80 percent in the poorest countries to less than 2 percent in the richest countries. In countries where a larger share of labor is employed in agriculture, the relative price of agricultural output is also higher - roughly twice higher in the poorest 10 percent than in the richest 10 percent countries. Using producer price of sectoral output from each country, I show that this prediction from the model is quantitatively consistent with data.

This paper relates to two strands of literature that examine cross-country productivity differences from a sectoral perspective. The first one is about understanding the asymmetric productivity differences in agriculture and nonagriculture across countries. Examples along this line are Gollin et al. (2007), Restuccia et al. (2008), Lagakos and Waugh (2013) and Adamopoulos and Restuccia (2014). My paper complements these studies by highlighting the role of income risks in shaping employment and labor productivity in agriculture.

The second strand tries to understand the labor productivity and wage differentials between agriculture and nonagriculture, and how these differentials correlate with levels of development. Gollin et al. (2014) measure difference in value added per worker between agriculture and nonagriculture - the agricultural productivity gap - across countries. They establish that such a gap decreases with income per worker, a stylized fact that closely resemble the second productivity puzzle highlighted in this paper. Herrendorf and Schoellman (2014) and Cai and Pandey (2015) argue that value added per worker is subject to mismeasurement, in both rich and poor countries. Instead, they use census data to measure wage difference between sectors. They find that wage gap is larger in poor countries than in rich ones, but in everywhere human capital cannot account for all of the wage gap.[2] This paper provides a potential explanation for the wage gap, and why it correlates posi-

[2] Hnatkovskay and Lahiri (2013) also reach similar conclusion for India.

tively with income per worker.

Another paper that highlights the role of risks in understanding cross-country productivity difference in Donovan (2013). In that paper, uninsurable risks associated with modern agricultural technologies discourages farmers in poor countries from adopting such technologies, leading to low labor productivity in agriculture of these countries. This paper instead focuses on the risks associated with employment outside agriculture. Gollin et al. (2004) is an earlier attempt to simultaneously reconcile both productivity puzzles. The mechanisms at work are distinctly different between this paper and theirs. They highlight the role of home production while this paper emphasizes the importance of uninsurable income risks outside agriculture. I view my paper as complementary to theirs.

The remainder of the paper is organized as follows. Section 2 lays out the model. Section 3 presents the quantitative results, and Section 4 concludes.

2 Model

The economy is populated with a continuum of measure one individuals, each endowed with skills $X = (x_a, x_m)$. The skill x_a determines one's productivity in agriculture and the skill x_m determines productivity in non-agriculture. These skills are random draws from a known distribution $G(X)$.

Preferences Preferences are given by

$$U(c_a, c_m) = \eta log(c_a - \bar{a}) + (1 - \eta) log(c_m),$$

where c_a and c_m is consumption good produced in agriculture and non-agriculture. The parameter $\bar{a} > 0$ represents the minimum consumption requirement of agricultural goods. Therefore, preferences are non-homothetic.

Technology There are representative firms producing output in agriculture (y_a) and nonagriculture (y_m) using the following technology

$$y_a = AL_a,$$
$$y_m = AL_m,$$

where A is total factor productivity, L_a and L_m is skill-augmented labor employed in each sector defined as follows

$$L_a = \int_{j \in \Omega_a} x_a^j dG,$$
$$L_m = \int_{j \in \Omega_m} x_m^j dG,$$

where Ω_a and Ω_m is the set of individuals working in agriculture and nonagriculture.

There are competitive markets where firms can hire labor at efficiency wage w_a and w_m. Consider an individual with skill (x_a, x_m), her earnings from working in agriculture are deterministic and given by $w_a x_a$. In contrast, her earnings from working in nonagriculture are subject to random *i.i.d* shocks denoted by ξ, which is the fraction of earnings lost. Therefore, her take-home earnings from nonagriculture is $w_m x_m(1-\xi)$. I assume that the income shocks are distributed $H(\xi)$ and $\int \xi dH(\xi) = 0$. Therefore, there is no aggregate loss in income. In addition, there is no market for contingent contracts such that this income risk cannot be insured away.

Optimization Firms choose the amount of labor to maximize profit. Let p denote the price of output in agriculture to nonagriculture, then the linear technology implies that $w_m = A$ and $w_a = pA$. Since income risk is uninsurable, individuals choose a sector that maximizes their expected utility, not expected income. Let $V_a^j(X)$ denote the value function of an individual with ability $X = (x_a^j, x_m^j)$ who chooses to work in agriculture, and $V_m^j(X)$ the value function associated with working in non-agriculture. We have

$$V_a^j(X) = \max_{\{c_a^j, c_m^j\}} U(c_a^j, c_m^j)$$
$$s.t.: pc_a^j + c_m^j = pAx_a^j$$

Similarly, for an individual working in non-agriculture, we have

$$V_m^j(X) = \max_{\{c_a^j, c_m^j\}} \int U(c_a^j, c_m^j) dH(\xi)$$
$$s.t.: pc_a^j + c_m^j = Ax_m^j(1-\xi)$$

Finally, the individual's optimization problem can be written as

$$\max_{\{\Pi^j\}} V_a^j(X)\Pi^j + V_m^j(X)(1-\Pi^j)$$

where Π^j is the indicator function such that $\Pi^j = 1$ if $V_a^j(X) > V_m^j(X)$, and 0 otherwise. The set of individuals working in agriculture and nonagriculture is then given by $\Omega_a = \{j|\Pi^j = 1\}$ and $\Omega_m = \{j|\Pi^j = 0\}$.

2.1 Equilibrium

A competitive equilibrium is a collection of price (p)*, allocation for firms* (L_a, L_m)*, allocation for individuals* (c_a^j, c_m^j, Π^j) *such that given prices, (i) firms maximize profit; (ii) individuals maximize expected utility and (iii) markets clear.*

I now turn to characterizing equilibrium allocation of labor in the economy. It is useful, in fact, to start with an economy with no income shocks (or there is complete market to insure against such shocks). In this case, the model becomes identical to that in Lagakos and Waugh (2013).

No Income Shocks In this economy, the division of labor between sectors take a simple form. That is, individuals choose a sector that provides the highest income. Therefore an individual with skill $X = (x_a, x_m)$ chooses agriculture if and only if

$$pAx_a > Ax_m,$$

and the set of individuals that are indifferent between sectors have skills endowment such that $pAx_a = Ax_m$.

With Income Shocks With income shocks, a sector that yields a higher expected income might not generate higher expected utility because individuals are risk-averse, especially when consumption is closer to subsistence. The value function associated with working in agriculture is

$$V_a(X) = log(pAx_a - p\bar{a}) + \eta log(\eta) + (1-\eta)log(1-\eta) - log(p).$$

The value function associated with working in non-agriculture is

$$V_m(X) = \int log(Ax_m(1-\xi) - p\bar{a})dH(\xi) + \eta log(\eta) + (1-\eta)log(1-\eta) - log(p).$$

Therefore, an individual chooses agriculture if and only if

$$log(pAx_a - p\bar{a}) > \int log(Ax_m(1-\xi) - p\bar{a})dH(\xi).$$

Comparing the characterization of labor allocation between the economy without shock and the one with shocks leads to the following proposition

Proposition 1. *Given price p, individuals that are indifferent between agriculture and non-agriculture in the economy without shocks strictly prefer the agricultural sector in the economy with income shocks.*

Proof: For individuals indifferent between agriculture and non-agriculture in the no-shocks economy, we have $pAx_a = Ax_m$. Then it follows that

$$
\begin{aligned}
log(pAx_a - p\bar{a}) &= log\left(\int \{Ax_m(1-\xi) - p\bar{a}\} dH(\xi)\right) \\
&> \int log(Ax_m(1-\xi) - p\bar{a}) dH(\xi)
\end{aligned}
$$

where equality follows from the assumption $\int \xi dH(\xi) = 0$, and the inequality follows from the Jensen's Inequality.

Define $x_a = I(x_m) = exp\big(\int log(Ax_m(1-\xi) - p\bar{a}) dH(\xi)\big) + \frac{\bar{a}}{A}$. The function $I(x_m)$ yields for a given non-agricultural skill the required agricultural skill such that an individual with these skills is indifferent between agriculture and non-agriculture. It is obvious that the function is strictly increasing in x_m. The following proposition establishes two properties of the function $I(x_m)$.

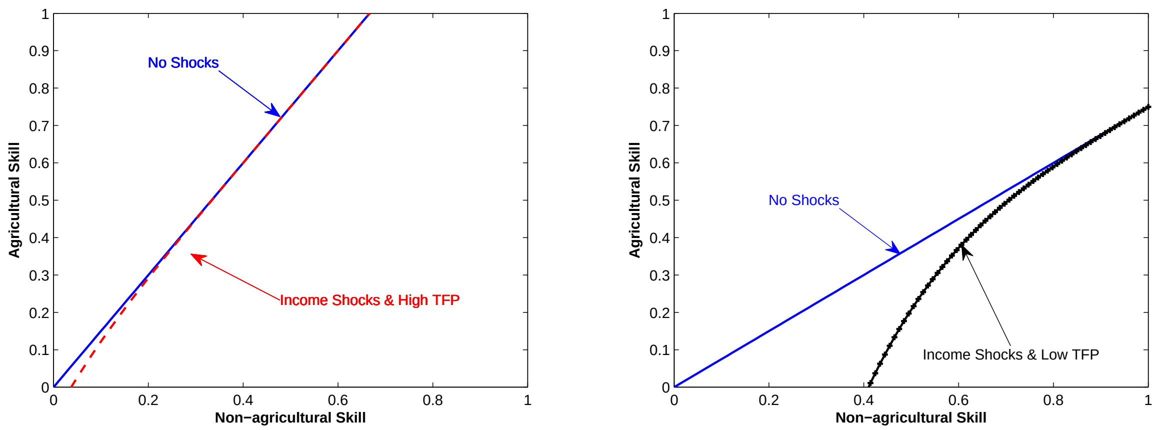
Proposition 2. *Assume that* $Ax_m(1-\xi) - p\bar{a} > 0 \; \forall \; (x_m, \xi)$*, then* $\frac{\partial^2 I}{\partial^2 x_m} < 0$, $\frac{\partial^2 I}{\partial x_m \partial A} < 0$

Proof: see appendix.

The economics of proposition 2 are as follow. It is useful to start with the economy with no shocks. Individuals who are indifferent between agriculture and nonagriculture have skills such that $pAx_a = A_x m$. This is represented by the blue solid line in Figure 2, with non-agricultural skills on the horizonal axis and agricultural skills on the vertical axis. The left panel represents an economy with high TFP, and the right panel represents an economy with low TFP. In both cases, the division of labor in this economy is such that individuals above the blue line choose agriculture, and nonagriculture otherwise. As the picture illustrates, the economy with low TFP will has a larger share of employment in agriculture.

Now introduce income shocks to earnings in nonagriculture. From proposition 1 we know that holding price fixed, those who are previously indifferent between sectors will now strictly prefer agriculture. In other words, there is migration of worker from nonagriculture to agriculture. The question is how does this migration depend on one's skills, or maybe more importantly, comparative advantage? The first part of proposition 2 shows that as individual move down skills levels, one needs stronger comparative advantage in non-agricultural production to remain in nonagriculture. Holding TFP fixed, the function $I(x_m)$ is represented by the red dashed line for the high TFP economy, and by a black crossed line for the low TFP economy. First note that both red and the black lines lie below the blue ones, implying that in both economies there is migration of workers into agriculture. Since the function $I(x_m)$ is concave as Proposition 2 shows, the selection into agriculture is not uniform across skill levels. In particular, those with high skills in both agriculture and nonagriculture are more likely to remain in nonagriculture because, relative to their earnings, the random income shocks are not substantial, especially given that such shocks have a zero mean. In contrast, those with low skills in both agriculture and nonagriculture are more likely to move to agriculture to avoid uncertainty in income. For these individuals, a mean-preserving spread of earnings result in significant loss in expected utility because they are risk averse and their level of consumption is closer to subsistence. The result of this selection is lower average agricultural skills in agriculture and higher average non-agricultural skills in nonagriculture.

Figure 2: Division of Labor



Proposition 2 also implies that the difference between an economy without income shocks and one with depends on the level of TFP. In Figure 2, this is represented by the difference between the blue line and the red (black) line in the high (low) TFP economy. The second part of proposition 2 states that the black curve has a deeper slope. In other words, mass of individuals who switch to agriculture in response to income shocks is higher in the economy with lower TFP. The intuition again is simple - individuals in an economy with lower TFP are consuming closer to subsistence and therefore more risk averse. Therefore, a negative income shock carries a larger weight in terms of utility. As a result, given the same agricultural skill one needs even stronger comparative advantage in non-agricultural production, compared to that required in an economy with higher TFP, to remain in nonagriculture. The result is income risks induce a larger decline in average agricultural skill and a larger increase in nonagricultural in an economy with low TFP, relative to one with high TFP.

To further illustrate this, consider two economies with TFP A_r and A_p, such that $A_r > A_p$. Let y_a and y_m denote output per worker in agriculture and nonagriculture, then we have

$$(y_a^r/y_a^p) = (A_r/A_p)\left(\overline{x_a^r}/\overline{x_a^p}\right),$$
$$(y_m^r/y_m^p) = (A_r/A_p)\left(\overline{x_m^r}/\overline{x_m^p}\right),$$

where $\overline{x_a^i}$ is average agricultural skill and $\overline{x_m^i}$ is average nonagricultural skill, $i = p, r$. From Lagakos and Waugh (2013) we know that the model would generate $\overline{x_a^r} > \overline{x_a^p}$, and $\overline{x_m^r} < \overline{x_m^p}$. The question is, for a given difference in TFP, how does the presence of risks alter skill differences between countries? First of all, the presence of income risks in nonagriculture would force individuals with weak comparative advantage in nonagriculture to move into agriculture. Since this mechanism is stronger among individuals with low skills (in both agriculture and nonagriculture), the end result is that higher average nonagricultural skill and agriculture and lower average agricultural skill in agriculture. That is, $\overline{x_a^r}$ and $\overline{x_a^p}$ both decline while $\overline{x_m^r}$ and $\overline{x_m^p}$ both increase. Although this would occur in both economies, the role of income uncertainty is more important in countries with low aggregate TFP. Therefore, the decline in $\overline{x_a^p}$ and the increase in $\overline{x_m^p}$ are larger in the economy with lower aggregate TFP. Therefore, income risks tend to amplify the difference in output per worker in agriculture, while at the same time reduce the difference in output per worker in nonagriculture.

Income risks also affect average wage between agriculture and nonagriculture. Since individuals are risk averse, working in nonagriculture and being subject to income uncertainty implies that average wage has to be higher in nonagriculture to compensate for risk bearing. Therefore, the presence of risks in nonagriculture generates a wedge in wage between sectors, and the size of the wedge reflect the utility costs associated with income uncertainty from working in nonagriculture. From proposition 2, we know that the utility cost associated with income uncertainty is higher in an economy with lower TFP, therefore a larger premium in wage is required to induce individuals to work in nonagriculture. Therefore, the model would generate a larger wage gap in an economy with lower TFP and a smaller gap in an economy with high TFP. Note that the benchmark model from Lagakos and Waugh (2013) cannot generate this feature.[3]

Of course, so far the discussion has been based on holding relative prices fixed across economies. The differences in the division of labor would in turn effect relative prices in a general equilibrium. Assessing the equilibrium effects is quantitative in nature, and is the objective of the next section.

3 Quantitative Analysis

3.1 Calibration

I calibrate the model to U.S. data. Aggregate TFP is set to 1 by normalization. For preferences, I set $\eta = 0.0036$ to generate a long-run share of expenditure on agricultural good that is commonly assumed in other papers studying cross-country productivity differences, e.g., Restuccia et al. (2008), Lagakos and Waugh (2013) and Adamopoulos and Restuccia (2014).

[3]See section A.2 in appendix for detail.

Functional forms Following Lagakos and Waugh (2013), I assume that the joint distribution of skills takes the following form

$$
\begin{aligned}
& G(x_a, x_m) = C[G^a(x_a), G^m(x_m)], \\
\text{where} \quad & G^a(x_a) = e^{-x_a^{-\lambda_a}} \quad \text{and} \quad G^m(x_m) = e^{-x_m^{-\lambda_m}}, \\
\text{and} \quad & C[u, v] = \frac{-1}{\rho} log \left\{ 1 + \frac{(e^{-\rho u} - 1)(e^{-\rho v} - 1)}{e^{-\rho} - 1} \right\}.
\end{aligned}
$$

The marginal distribution of skill is Fréchet with scale parameter λ_a and λ_m, and a common location parameter that is zero. The parameter ρ governs the correlation between the two skills. The smaller is λ_a and λ_m, the larger the variance of innate skill is. Moreover, a positive (negative) correlation corresponds to a case when an individual who is more productivity in agricultural production is also more (less) likely to productive in nonagricultural production.

I assume that income shock ξ follows a truncated normal distribution with mean 0 and variance σ^2, over the interval $[-\bar{\xi}, \bar{\xi}]$. As a first pass, I set $\bar{\xi} = 1/2$. This ensures that all individuals can afford the minimum consumption of agricultural goods in all economies.

Calibration targets I am left with five parameters whose values remained undermined. Preferences parameter $\bar{a}$, skill distribution parameters $\lambda_a, \lambda_m, \rho$, income shock parameter σ. These parameters are chosen jointly to match five moments in the data: (1) the share of labor in agriculture (0.02); (2) average wage in agriculture relative to nonagriculture (0.71); (3) variance of log wage in agriculture (0.144); (4) variance of log wage in nonagriculture (0.38) and (5) the permanent component of wage variance in nonagriculture (0.224). These statistics about wage in agriculture and nonagriculture are computed using data from CPS. For details, see appendix in Lagakos and Waugh (2013). Figure 7 plots the distribution of normalized log wage in agriculture, along with predictions from the model. Figure 8 does the same for nonagriculture.

3.2 Cross-country Comparison

Now I use the model to think about sectoral productivity differences across countries. I use data from Restuccia et al. (2008) and the sample has 81 countries. Countries are identical except for their levels of TFP. In particular, they have the same unconditional distribution of skills and the same process of income shocks. For each country, I then iterate on TFP until the model generates aggregate GDP per worker (measured at model international price) relative to the U.S. that is the same as in the data. Therefore, the model by construction reproduces the world income distribution.

Sectoral Productivity I first explore whether the model could account for the first productivity puzzle, i.e., larger productivity difference in agriculture than in nonagriculture. Table 1 presents the ratio of output per worker in agriculture and nonagriculture between the 90th percentile and the 10th percentile country. Note again that by construction, the model generates a 22-fold difference in real GDP per worker between the 90th and the 10th percentile countries. Like in the data, the model generates a much bigger gap in real output per worker for agriculture, and a smaller gap for nonagriculture. In the model, the 90th-10th gap in output per worker for agriculture is 32, and 12 for nonagriculture. The last row of Table 1 also reports the corresponding numbers from Lagakos and Waugh (2013). Relative to theirs, predictions from my model with income shocks are closer to the data in terms of cross-country productivity differences.

Table 1: Productivity differences by sector, model and data

	90th-10th Ratio		
	Agriculture	Aggregate	Nonagriculture
Data	45	22	4.5
Model	32	22	12
Lagakos and Waugh (2013)	29	22	13

In additional to TFP, the difference in output per worker at the sector level reflects difference in average sector-specific skills across countries. Table 2 summarizes average skill in agriculture and nonagriculture for the 10th and the 90th percentile countries. For agriculture, average agricultural skill is almost twice higher in the 90th percentile country, relative to that in the 10th percentile country. The opposite is true when it comes to average nonagricultural skill - it is almost twice higher in the 10th percentile county. Therefore, the model generates larger productivity difference in agriculture than in nonagriculture.

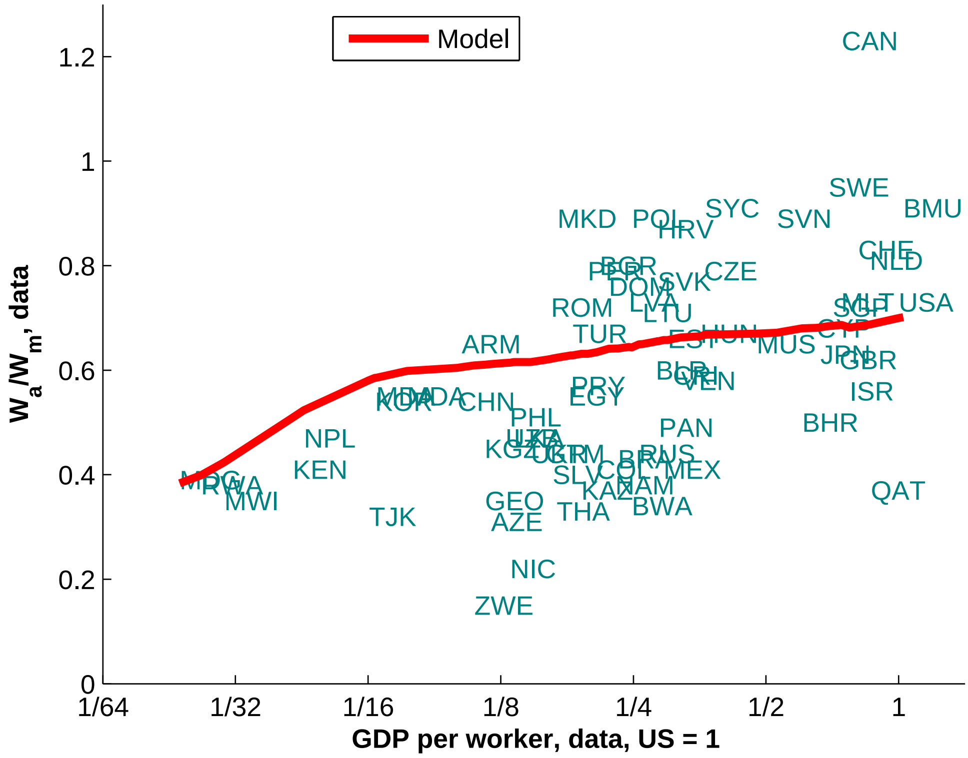
Table 2: Average skill difference by sector, model

	Average Skill	
	Agriculture	Nonagriculture
90th percentile	1.76	1.00
10th percentile	1.02	1.76
90th-10th Ratio	1.72	0.71

Relative wages Next I examine the model's prediction about relative wage between agriculture and nonagriculture. Figure 3 plots the ratio of average wage between agriculture and nonagriculture in the data, along with predictions from the model. To make fair comparison, the wage data is taken directly from Lagakos and Waugh (2013). The model is able to generate a positive correlation between the relative wage and GDP per worker. In the data, the correlation between the two is 0.62. The model generates a correlation that is 0.68. This aspect of the model is novel relative to Lagakos and Waugh (2013). They recognize also the stylized fact that relative wage increases with the level of development, but their model could not generate a significant increase in relative wage when income per worker increases (See Figure 1 in section E of their paper).

In summary, the model is able to generate both (1) larger productivity difference in agriculture than in nonagriculture and (2) increasing relative wage of agriculture with income by relying on only exogenous difference in sector-neutural TFP. In particular, the model's ability to generate wage gap between sectors is not based on explicit frictions such as barriers to inter-sectoral mobility. Instead, the theory highlights the role of employment risks, a theme articulated in Harris and Todaro (1970) and the lack of complete insurance market. Of course, in addition to labor productivity and wage, the difference in economic development encompasses broader implications. It is important that the model is consistent with these other implications as well. In paragraphs that follow, I cross-check the

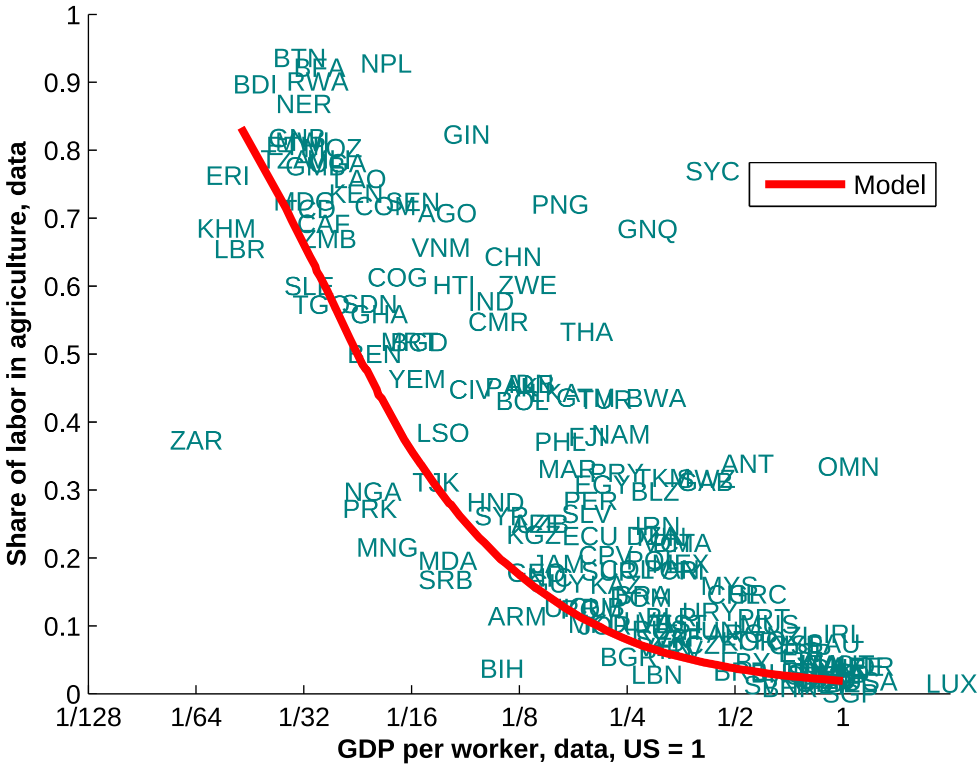
Figure 3: Relative wage, model and data



model's other testable predictions, starting with the share of labor in agriculture.

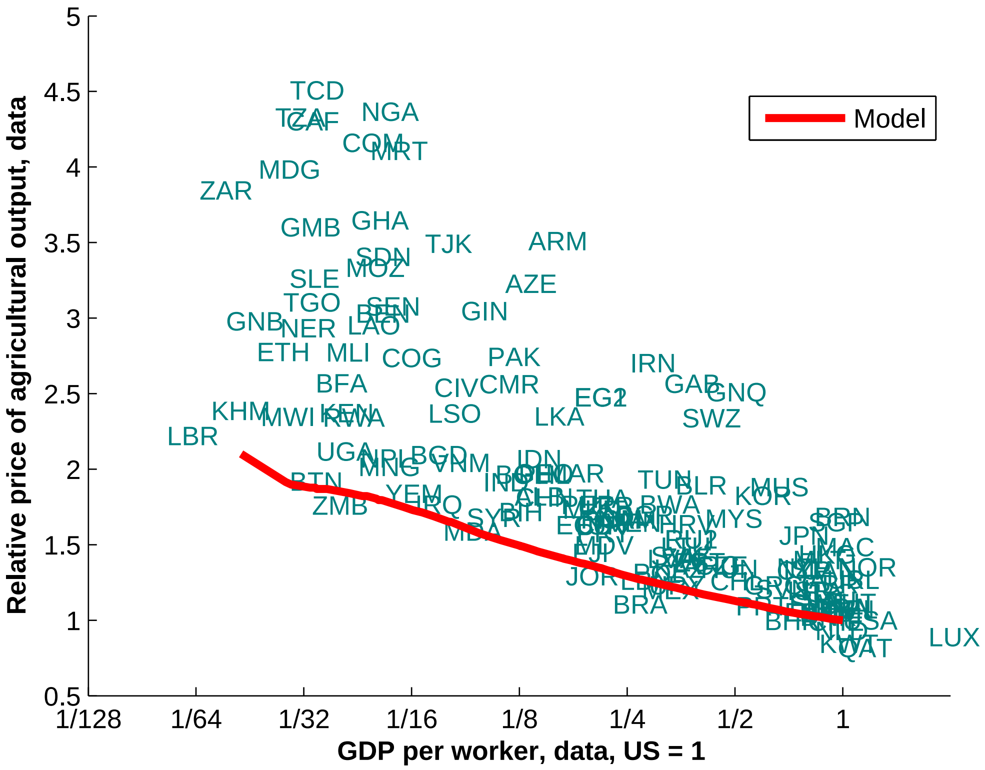
Employment in agriculture Another stylized fact about economic development is that the share of employment in agriculture falls with income. The prediction from the model in this dimension is quantitatively consistent with the data. Figure 4 plots on the horizontal axis GDP per worker relative to the U.S., and on the vertical axis the share of labor in agriculture in the data, along with predictions from the model. As in the data, the model generates a large share of employment (80 percent) in the poorest set of countries. Agricultural employment drops rapidly as the economy develops, a well-known feature of structural transformation. In the richest 10 percent of countries, the share of labor in agriculture is below 5 percent in both data and the model.

Figure 4: Share of employment in agriculture, model and data



Relative prices In the model, the fall in agricultural employment with respect to income is driven by the non-homothetic preferences. Specifically, the minimum consumption requirement in the preferences dictates that as income falls, a larger proportion of expenditure is spent on goods produced in the agricultural sector. Another implication is that relative price of agricultural output is higher in countries with lower aggregate income. Figure 5 plots the relative price of agricultural good relative to nonagricultural good in the data, along with predictions from the model. The relative prices in the data are producers prices aggregated from ICP. Relative price for the United States is normalized to 1 in the plot. The model predicts that this relative price is twice higher in the poorest set of countries, relative to rich ones. While the model is able to capture the decline in relative price with respect to income per worker, it is also noted that in among poor countries, the dispersion in price is considerably larger than those in the model. For example, in some countries the relative price is as much as 5 times higher than that in the U.S.

Figure 5: Price of agriculture relative to nonagriculture



3.3 The Importance of Risk

In the model, income risks interact with TFP to affect the division of labor. How important is risk in explaining cross-country productivity difference? In other words, how would a model without risk perform relative to the benchmark in terms of reconciling the two productivity puzzles? To answer this question, I consider an economy without income risks and compare its implications against those from the benchmark model. To achieve this, I set the income shocks $\xi = 0$. For the time being, I leave the calibration the same and focus on the marginal effects of risk on labor allocation, productivity and wage gap. Table presents the model predictions about productivity differences between then 90th and the 10th percentile country, both from the benchmark model and the alternative model with no risk.

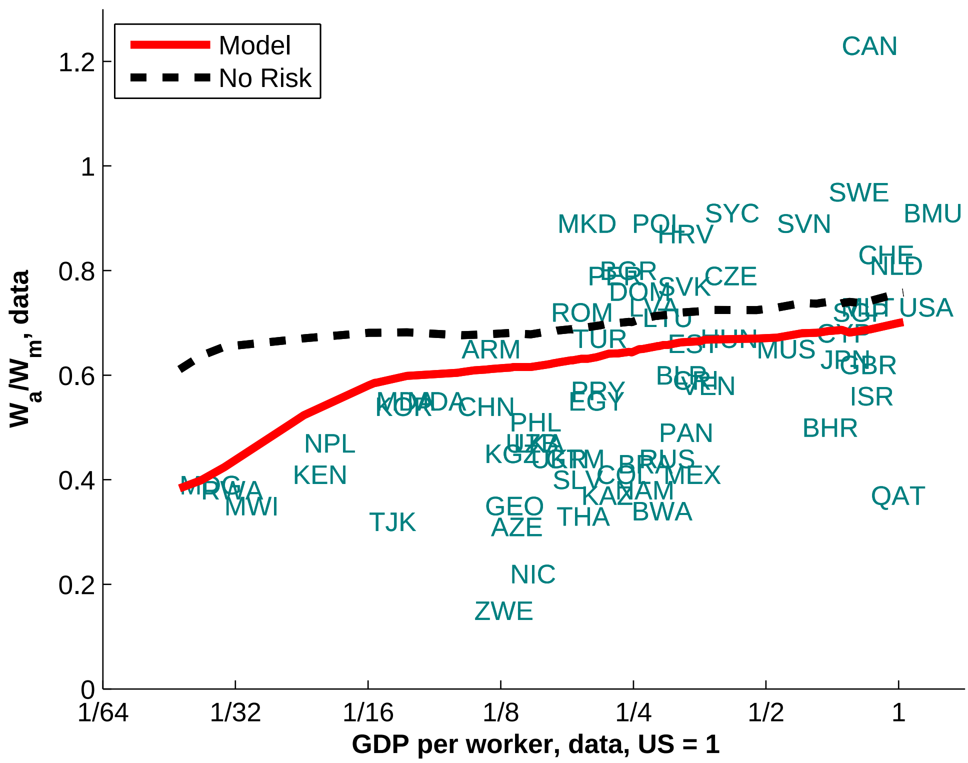
Table 3: Productivity and skill differences

	90th-10th Ratio		
Output per worker			
	Agriculture	Aggregate	Nonagriculture
Model	32	22	12
Model w/o risk	31	22	13
Average skill			
	Agriculture	Aggregate	Nonagriculture
Model	1.79	-	0.68
Model w/o risk	1.72	-	0.71

Relative to the benchmark, the model without risk generates a smaller productivity difference in agriculture and a larger difference in nonagriculture, which is going in the wrong direction in terms of reconciling the first productivity puzzle. In other words, employment risk plays a role in the model in such a way that it generates predictions that are consistent with data. The reason, as explained before, is that employment risks generate a more stark difference in terms of specialization between rich and poor countries. This could be visualized again in Table 3. One of the prediction from the model is that average agricultural skill is higher in rich countries than in poor countries, and the opposite is true for non-agricultural skill. In the benchmark model, the 90th-10th ratio of agricultural skill is 1.79, while that of nonagricultural skill is 0.68. When employment risk is removed, the ratio of agricultural skill reduces to 1.72 while that of nonagricultural skill increases to 0.71. Therefore, the model without employment risk generates a smaller productivity difference in agriculture and a larger difference in nonagriculture.

Another important role played by employment risk is that it endogenously generates a wedge in wage between agriculture and nonagriculture. That is, in equilibrium wage in nonagriculture commands a premium over that in agriculture in order to compensate for the risk, even though such risk on average does not reduce wage in nonagriculture. More importantly, such employment risks are conceivably more important in economies with low aggregate income, as individuals in these economies are living closer to subsistence, and hence, negative shocks to income translate into a higher utility loss. Consequently, the required wage premium for workers in nonagriculture is higher. Thus, the model generates a wage wedge between agriculture and nonagriculture that decreases with the level of development. This prediction is absent in a alternative model without employment risks.

Figure 6: Relative wage, model and data



To verify this, Figure 6 plots relative wage against GDP per worker relative to the U.S., along with predictions from the benchmark model and those from a riskless model. A model without employment risks clearly fails to generate a rise in relative wage with income per worker. This is also noted in Lagakos and Waugh (2013). Therefore, employment risks play a quantitatively important role in explaining this pattern of relative wage across countries.

3.4 Country-specific Income Risks

[To be written.]

3.5 Risk in Agriculture

[To be written.]

4 Conclusion

[To be written.]

References

Adamopoulos, T. and D. Restuccia (2014). The Size Distribution of Farms and International Productivity Differences. *American Economic Review 104*(6), 1667–97.

Cai, W. and M. Pandey (2015). The Agricultural Productivity Gap in Europe. *Economic Inquiry 53*(4), 1807–1817.

Caselli, F. (2005). Accounting For Cross Country Income Differences. In P. Aghion and S. D. Durlauf (Eds.), *Handbook of Economic Growth*, pp. 679–741. ELSEVIER.

Donovan, K. (2013). Agricultural Risk, Intermediate Inputs, and Cross-Country Productivity Differences. Working paper, University of Notre Dame.

Gollin, D., D. Lagakos, and M. E. Waugh (2014). The Agricultural Productivity Gap. *Quarterly Journal of Economics 129*(2), 939–993.

Gollin, D., S. L. Parente, and R. Rogerson (2004). Farm work, Home Work and International Productivity Differences. *Review of Ecconomic Dynamics 7*(4), 827–850.

Gollin, D., S. L. Parente, and R. Rogerson (2007). The Food Problem and the Evolution of International Income Levels. *Journal of Monetary Economics 54*(4), 1230–1255.

Harris, J. R. and M. P. Todaro (1970). Migration, unemployment and development: a two-sector analysis. *The American economic review 60*(1), 126–142.

Herrendorf, B. and T. Schoellman (2014). Why is Measured Productivity so Low in the Agriculture. *Forthcoming. Review of Economic Dynamics*.

Hnatkovskay, V. and A. Lahiri (2013). Structural Transformation and the Rural-Urban Divide. Working paper, University of British Columbia, Vancouver School of Economics.

Lagakos, D. and M. E. Waugh (2013). Selection, Agriculture, and Cross-Country Productivity Differences. *American Economic Review 103*(2), 948–80.

Madrian, B. C. and L. Lefgren (1999). A note on longitudinally matching current population survey (cps) respondents.

Restuccia, D., D. T. Yang, and X. Zhu (2008). Agriculture and Aggregate Productivity: A Quantitative Cross-country Analysis. *Journal of Moneytary Economics 55*(2), 234–250.

Roy, A. D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers 3*(2), 135–146.

Schultz, T. W. (1964). *Transforming Traditional Agriculture*. New Haven: Yale University Press.

A Proofs

A.1 Proposition 2

The assumption guarantees that individuals always consume above subsistence. Differentiating the function I with respect to x_m yields

$$\frac{\partial I}{\partial x_m} = \frac{1}{pA} exp\Big(\int log(Ax_m(1-\xi)-p\bar{a})dH(\xi)\Big)\int\left(\frac{A}{Ax_m(1-\xi)-p\bar{a}}\right)dH(\xi),$$

It is trivial to note that $\frac{\partial I}{\partial x_m} > 0$. Now the second-order derivative is

$$\begin{aligned}\frac{\partial^2 I}{\partial^2 x_m} =& \frac{1}{pA} exp\Big(\int log(Ax_m(1-\xi)-p\bar{a})dH(\xi)\Big)\times \\ & \left[\left(\int\left(\frac{A}{Ax_m(1-\xi)-p\bar{a}}\right)dH(\xi)\right)^2 - \int\left(\frac{A}{Ax_m(1-\xi)-p\bar{a}}\right)^2 dH(\xi)\right] \\ & < 0,\end{aligned}$$

where the last inequality follows from the fact that the function $f(x) = x^2$ is convex and the Jensen's Inequality.

The cross derivative is given by

$$\begin{aligned}\frac{\partial^2 I}{\partial x_m \partial A} =& \frac{1}{p} exp\left(\int log(Ax_m(1-\xi)-p\bar{a})dH(\xi)\right) \times x_m \\ & \left[\left(\int \left(\frac{1}{Ax_m(1-\xi)-p\bar{a}}\right) dH(\xi)\right)^2 - \int \left(\frac{1}{Ax_m(1-\xi)-p\bar{a}}\right)^2 dH(\xi)\right] \\ & < 0.\end{aligned}$$

Again, the last inequality is a direct result of Jensen's Inequality.

A.2 Sectoral wage gap in Lagakos and Waugh (2013)

In a complete market, the share of labor in agriculture is $n_a = Prob(px_a > x_m) = \int_0^\infty \int_0^{px_a} dg(x_m)dg(x_a)$. It is shown in Lagakos and Waugh (2013) that

$$n_a = \frac{1}{p^{-\theta}+1}.$$

Correspondingly, the share of labor in nonagriculture is $n_m = 1 - n_a = \frac{p^{-\theta}}{p^{-\theta}+1}$. Then the ratio of wage in agriculture relative to nonagriculture is

$$\begin{aligned}\frac{py_a/n_a}{y_m/n_m} &= \frac{pE(x_a|px_a > x_m)}{E(x_m|px_a < x_m)} \\ &= \frac{p(1+p^{-\theta})^{\frac{1}{\theta}}\gamma}{(1+p^{\theta})^{\frac{1}{\theta}}\gamma} \\ &= 1.\end{aligned}$$

B Sectoral Wage

The main source of data is March supplements to the Current Population Survey. To increase sample size, I pool together samples between from 2000 to 2006. The samples are restricted to include individuals between age 18 and 65 with positive earnings and hours. I calculate total earnings for an individual as wage and salary income plus 0.6 times business income plus 0.4 times farm income. Hourly wage is then computed as total earnings divided by hours worked last year.

The structure of CPS allows for matching individuals between two consecutive surveys. This is explained in detail in Madrian and Lefgren (1999). The econometric model could be stated as

$$w_{it} = \alpha_i + X\beta + \epsilon_i t,$$

where w_{it} is wage at time t, X is a set of controls, and ϵ_{it} is the random error. In this case, the variance of the permanent component could be estimated as the covariance of wage across two periods, $cov(w_{it}, w_{i,t+1})$.

Figure 7: Distribution of hourly earnings in agriculture, U.S.

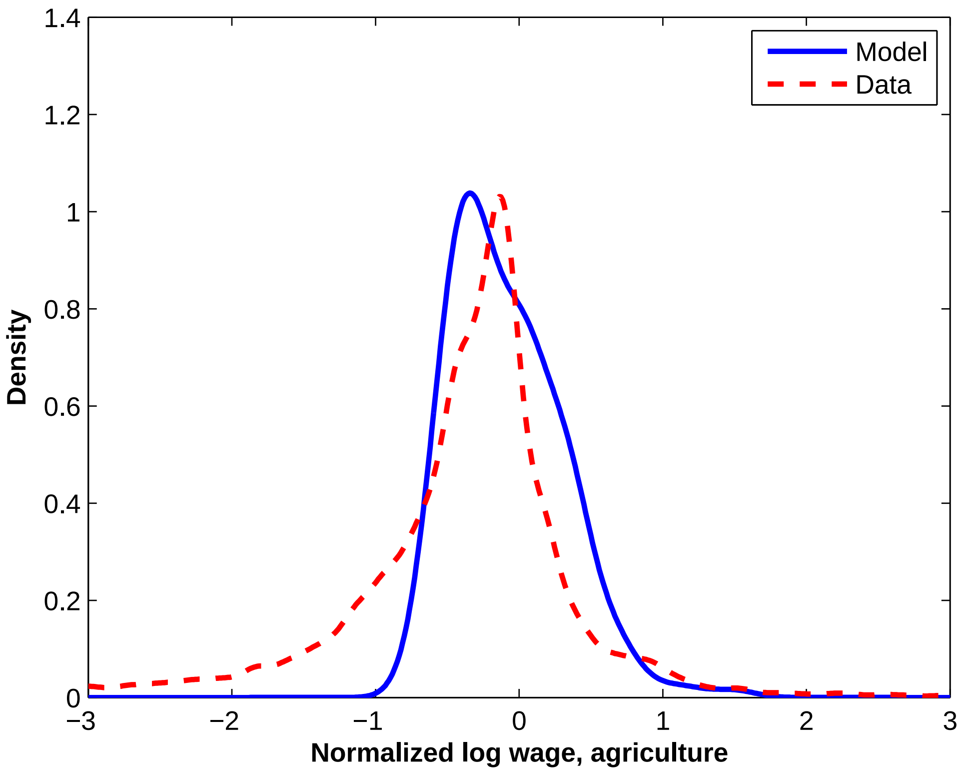


Figure 8: Distribution of hourly earnings in nonagriculture, U.S.