Agricultural Risk, Intermediate Inputs, and Cross-Country Productivity Differences*

Kevin Donovan†
Arizona State University
February 2012

Abstract

Agricultural labor productivity differences between the richest and poorest countries are factor of three larger than aggregate productivity differences. In addition, domestically priced intermediate input shares in agriculture differ by a factor of four. I propose a theory in which the under-utilization of intermediate inputs amplifies sector neutral productivity (TFP) differences in the agricultural sector of low income countries. In the face of incomplete markets, idiosyncratic productivity shocks, and subsistence requirements, risk averse farmers in poor countries to put relatively more weight on bad potential shock realizations when making input decisions. Relative to a baseline model with no risk, the model predicts that both agricultural and aggregate productivity decrease by 50% in the poorest countries. This result is driven by the fact that the model accounts for 70% percent of the difference in agricultural intermediate input shares between the richest and poorest countries.

*Thanks to Chris Herrington, Dan Lawver, Adam Nowak, Ed Prescott, Paul Schreck and participants at the ASU Macro Workshop, WSU, the CEA meeting in Ottawa, and the SED meeting in Ghent for comments and insights. A special thanks to Berthold Herrendorf, David Lagakos and Todd Schoellman for countless discussions on this topic. The usual disclaimers apply.

†Contact Info: Department of Economics, Arizona State University, P.O. Box 879801, Tempe AZ 85287-9801. Email: kevin.donovan@asu.edu
1 Introduction

Differences in agricultural output per worker between the richest and poorest countries are significantly larger than differences in aggregate output per worker (Caselli, 2005). In spite of this, over 80% of people are employed in agriculture in the poorest countries. Since poor countries employ a large fraction of the population in a particularly unproductive sector, basic accounting suggests that understanding agricultural productivity differences is key in understanding aggregate differences.

One possible explanation for understanding agricultural productivity differences is that farmers in developing countries are using fewer intermediate inputs. Section 2 documents that the domestically priced ratio of intermediate inputs to agricultural output exhibits a strong positive correlation with per capita GDP, ranging from a low of 0.04 in Uganda to 0.40 in the United States. The goal of this paper is to provide a theory to understand the correlation between domestically priced intermediate input shares in agriculture and income levels across countries, and in turn, assess its role for productivity differences across countries.

To do so, I develop a dynamic general equilibrium model in which farm input decisions are made jointly with consumption choices. This implies that farmers do not take profit as given when choosing consumption, consistent with a large empirical literature reviewed by Morduch (1995). I further assume that farmers face both uninsurable idiosyncratic productivity shocks and subsistence requirements. The inability to insure ex-post consumption creates a deviation from the standard profit maximization problem. Instead of weighting each possible outcome by the probability of occurrence, each outcome is weighted by the product of the probability and the farmer’s marginal utility at that realization. As TFP decreases, the farmer’s income net of subsistence requirements moves closer to zero. This increases marginal utility at low shock realizations relative to farmers in rich countries. Put more simply, farmers in poor countries become more concerned with the disastrous consequences of bad, but unlikely, shocks. In the model, this generates a “wedge” between
profit maximizing marginal value and price of intermediates, even though the Cobb-Douglas farm production technologies differ by only a factor-neutral TFP parameter across countries. When aggregated, this implies that poor countries have both a lower domestically priced intermediate input share and lower labor productivity in agriculture.

To assess the cross-country implications of the mechanics described above, I calibrate the model using a mixture of aggregate and individual level data to replicate key features of agricultural production in developing countries. These features include risky production, for which I use plot level data from six Indian villages from the International Crops Research Institute for the Semi-Arid Tropics Village Level Surveys (ICRISAT VLS) to discipline the distribution of productivity shocks. Relying on previous literature, I further assume that savings is limited to agricultural storage (see, for example Udry, 1995). Storage turns out to be an incredibly costly way to save, due to insect infestations, molds, and any other number of factors. Using a new set of statistics for agricultural storage depreciation, I find that depreciation rates can be as high as 30%, as it is for maize in Zimbabwe. This limits the ability of agricultural storage to be used as a buffer stock to self-insure against risk. Lastly, I include agricultural-specific price distortions that have been documented extensively in the literature (see Restuccia, Yang, and Zhu, 2008; Vollrath, 2009). While agnostic on the cause, I use the model to investigate their role in understanding the impact of agricultural risk. I find that the impact of risk depends critically on their inclusion in the model.

The model generates two main results. With the baseline calibration, the model predicts a factor of 31 difference in agricultural productivity and a factor of 8 difference in aggregate productivity. Empirical evidence puts these values at 64 and 23. This result is generated by differences in intermediate shares across countries. The model predicts that the poorest countries have a domestically priced intermediate share of 0.18, compared to the US intermediate share of 0.40. The model therefore captures seventy percent of the difference found in the data, in which the poorest countries have intermediate input shares of about 0.09. Since technologies differ by only a factor-neutral TFP parameter, any difference in intermediate input shares across countries is driven by the interaction of production risk,
incomplete markets, and subsistence requirements.

The second result deals with the interaction of risk with other distortions present in developing countries. While I show in Section 4 that uninsurable risk is the only distortion needed to theoretically generate a correlation between intermediate input shares and income, it is of course not the only distortion present in poor countries. Other documented distortions include high transportation costs to rural agricultural areas (Adamopoulos, 2011; Gollin and Rogerson, 2010), or institutional differences that generate unequal wages between sectors (Vollrath, 2009; Gollin, Lagakos, and Waugh, 2011). Most similar to this paper is Restuccia, Yang, and Zhu (2008), who consider the role of price distortions in accounting for intermediate input usage across countries. Quantitatively, I find that the aggregate importance of agricultural risk depends critically on the magnitude of these price distortions in agriculture. Relative to these studies, this model presents a new margin through which these distortions can impact productivity. Since high prices decrease expected income, farmers limit their exposure to risk by reducing intermediate input usage. This amplification shows up as lower domestically priced intermediate input shares in poor countries, a result that cannot be generated by price distortions alone. Quantitatively, the model increases the prediction of agricultural output per worker differences between the richest and poorest countries by 52% and aggregate output per worker differences by 63% relative to the world with no risk, but the same price distortions. As a comparison, without price distortions the model predicts an increase in agricultural productivity of 5% relative to the no risk version. I further show that even small changes in price distortions can have dramatic effects on agricultural productivity, implying that the welfare and productivity gains from small improvements in, say, roads are being severely underestimated in standard aggregate models.

In addition to the work on price distortions discussed above, other explanations for cross-country agricultural productivity differences include occupational selection (Lagakos and Waugh, 2010), distortions limiting farm size (Adamopoulos and Restuccia, 2011), barriers that limit specialization through trade (Tombe, 2011), and the possibility of mismeasurement (Herrendorf and Schoellman, 2011). This paper also builds on an empirical literature on
agricultural household intermediate input decisions. First, the model is consistent with empirical evidence that households do not make profit maximizing intermediate input choices, as found in Duflo, Kremer, and Robinson (2008) and Zerfu and Larson (2010). Rosenzweig and Binswanger (1993) and Dercon and Christiaensen (2011) use household level evidence from India and Ethiopia to show that risk plays an important role in explaining this result, with the latter investigating the interaction between risk and fixed costs of purchasing intermediate inputs. This paper investigates the aggregate cross-country implications of agricultural risk in a two sector general equilibrium model, while remaining consistent with the aforementioned empirical results.

The rest of the paper proceeds as follows. Section 2 presents some motivating evidence of differences in intermediate input shares. In a cross section of countries, there is a strong positive correlation between income level and domestically priced intermediate input shares in agriculture. This correlation does not exist in the manufacturing or service sectors. The model is described in Section 3. Section 4 provides some theoretical results to show how the interaction of TFP, uninsurable risk, and subsistence requirements can match cross-country facts on intermediate shares. Turning to quantitative results, Section 5 details the calibration and Section 6 presents the results of the model. Section 7 considers the robustness of the results to different calibrations, and finally, Section 8 concludes.

2 Motivating Evidence

Domestically priced intermediate input shares control for price differences driven by, for example, transportation costs or relatively unproductive intermediate production. Using an analysis similar in spirit to Hsieh and Klenow (2007), I show that even at domestic prices, there is still a strong positive correlation between the intermediate good share in agriculture and the GDP per capita across countries.
2.1 Domestic Intermediate Input Shares

The *domestic intermediate share* in agriculture of country $j$ is

$$
\hat{X}^j := \frac{p^j_x X^j}{p^j_a Y^j_a}
$$

(2.1)

where $X$ is the quantity of nonagricultural intermediate inputs, such as oil and fertilizer, and $Y_a$ is the quantity of agricultural output. The domestic prices faced by the farmer are denoted $p^j_x$ and $p^j_a$. The intermediate input price $p^j_x$ takes into account any distortions between the nonagriculture and agriculture sectors, such as transportation costs. Since I am interested in the decisions of farmers, these are the relevant prices. Note that because $\hat{X}^j$ is unitless, it is directly comparable across countries. The *real intermediate share* in agriculture of country $j$ is defined as

$$
\hat{X}^{j*} := \frac{p^{j*}_x X^j}{p^{j*}_a Y^j_a}
$$

(2.2)

where $p^{j*}_x$ and $p^{j*}_a$ are the international prices of intermediates and agricultural output. The difference between (2.1) and (2.2) is only one of prices. While the domestic intermediate share includes any price distortions in the domestic market, the real intermediate share does not. By equations (2.1) and (2.2), it is possible to write the domestic intermediate share as

$$
\hat{X}^j = \hat{X}^{j*} \left( \frac{p^j_x}{p^{j*}_x} \right) \left( \frac{p^j_a}{p^{j*}_a} \right)
$$

(2.3)

The internationally priced intermediate share and the price ratio in parenthesis are available in Prasada Rao (1993) for a cross section of 84 countries in the year 1985, constructed from Food and Agricultural Organization (FAO) statistics. The two ratios are plotted for the set of 84 countries in Figure 1 with logged real GDP per capita on the horizontal axis, from the Penn World Tables version 7.0 (PWT).\(^1\)

While the price ratio in Figure 1a exhibits no trend, there is a strong trend in the real intermediate share of Figure 1b. This point is emphasized by Restuccia, Yang, and Zhu

---

\(^1\)Prasada Rao (1993) actually reports the purchasing power parities (PPP) of agricultural output and intermediate inputs. These can be translated into measures of relative prices. Here the US is chosen as the numeraire, so $p^{j/US}_x = p^*_x$ and $p^{j/US}_a = p^*_a$ and therefore $\hat{X}^{US} = \hat{X}^{US*}$. See Appendix A for further details.
Figure 1: Components of Domestic Intermediate Share. Source is Prasada Rao (1993) and PWT (2008). Combined with equation (2.3), however, these plots suggest that there is also strong correlation between income and domestic intermediate shares. This is confirmed in Figure 2, which plots the domestic intermediate input share in agriculture.

Figure 2: Domestic Intermediate Share in Agriculture. Source is Prasada Rao (1993) and PWT.

While there is certainly more variance in the domestic intermediate share than the real intermediate share, there is still a clear positive correlation between income level and the domestic intermediate share, with a correlation of 0.65. To give some idea of the difference between rich and poor countries, the domestic intermediate share in Uganda is ten times lower than that of the US. The tenth percentile country, as ranked by GDP per capita, has a domestic intermediate share that is four times lower than the US. Furthermore, the
The differences between the richest and poorest countries are actually larger when measured at domestic prices than when measured at international prices. The correlation between the domestic intermediate share and the real intermediate share is 0.71.

2.2 Comparison to Manufacturing and Services

One key feature of agriculture is that production occurs during only the few harvesting seasons per year. Not constrained by these natural limitations, manufacturing and service sector firms are free to produce output throughout the year. This gives these firms two advantages in dealing with risk. First, if a “yearly” shock is realized, nonagricultural firms can adjust their input bundles throughout the year to respond to the shock realization. Because of the limited harvesting opportunities in agriculture, this is not possible. Second, if each production run brings with it an i.i.d. productivity shock, then the law of large numbers implies that simply increasing the number of production runs acts as a kind of insurance. Again, this is not possible in agriculture, because harvesting is limited to the optimal harvesting seasons. These natural production limitations put agricultural firms at a unique disadvantage to cope with production risk.

Evidence of this can be found by examining intermediate input shares in other sectors of the economy. Using domestically priced output and intermediate consumption statistics from the United Nations System of National Accounts (SNA), I first confirm the results presented in the last section- agricultural intermediate input shares are positively correlated with GDP per capita in a cross section of countries. However, this is not the case in manufacturing or services. Figure 3 plots domestically priced intermediate input shares for three sectors: (a) “Agriculture, hunting, forestry; fishing,” (b) “Manufacturing,” and (c) “Education; health and social work; other community, social and personal services,” which is probably the closest measure of services in the SNA. The plot includes 49 countries with data are available for all sectors, and are listed in Appendix A.

The difference between agriculture and the other two sectors is stark. The intermediate share in agriculture is positively correlated with income, as was also the case using statistics
Figure 3: Domestic Intermediate Shares in Three Sectors for 1985. Source is UN SNA for intermediate input shares and PWT 7.0 for GDP per capita.

derived from the FAO data. The intermediate shares in manufacturing and services, however, do not exhibit this correlation. To summarize, Table 1 presents the results of a simple linear regression of the sectoral domestic intermediate share on log PPP GDP per capita.

While measuring at domestic purchaser prices eliminates the correlation between the intermediate input share and income for manufacturing and services, it remains strongly positive in agriculture. In the linear regression, only agriculture has a slope significantly different from zero. Both results confirm that differences in intermediate input intensity are not driven exclusively by price distortions, since they exist even when these distortions are taken into account. By design, the Cobb-Douglas production function cannot capture this, regardless of the amount of input market frictions built in.\(^2\) The rest of this paper develops a model to understand the cause of this correlation in agriculture and assess its impact on

\(^2\)A constant elasticity of substitution production function with complementarity between capital and intermediate inputs could at least qualitatively deliver this result. If this is the case, then the result can instead be interpreted an amplification mechanism. However, Duflo, Kremer, and Robinson (2008) show that an increase in fertilizer investment would be profitable even without corresponding changes in complimentary agricultural practices. This is the feature highlighted in this paper.
Table 1: Relationship between Domestically Priced Intermediate Input Share and Log GDP per Capita (PPP), by Sector

<table>
<thead>
<tr>
<th></th>
<th>Agriculture</th>
<th>Manufacturing</th>
<th>Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>−0.43***</td>
<td>0.59***</td>
<td>0.21*</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.06)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>Log GDP per capita (PPP)</td>
<td>0.10***</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.52</td>
<td>0.03</td>
<td>0.03</td>
</tr>
</tbody>
</table>

**Table notes:** Standard errors are in parentheses. Significance at 0.01, 0.05, 0.1 levels denoted by ***, **, and *

cross-country productivity differences.

3 Model

The model period is a year, and time is discrete and runs $t = 0, 1, 2, \ldots$. There are two sectors, sector $a$ for agriculture and sector $m$ for manufacturing (the manufacturing sector includes all nonagriculture). Throughout, I normalize the output price of the sector $m$ good to $p_{mt} = 1$ for all $t$. Within an economy, there is a continuum of villages with measure one and each village contains a measure one of infinitely lived members. As discussed in Townsend (1994) and Ogaki and Zhang (2001), individuals are relatively well insured against purely idiosyncratic risk. Covariate risk, such as weather, is more difficult to insure against. Therefore, I assume all decisions are made at the village level.

Looking ahead, part of this paper will investigate the complementarity of production risk and price distortions. Therefore, I lay out the full model here with two distortions based on recent literature. While quantitatively important, I show in Section 4 that these features are not necessary to theoretically generate cross-country differences in intermediate input shares. Following Restuccia, Yang, and Zhu (2008) and Vollrath (2009), I first consider differences in intermediate prices, denoted $p_x$ ($\geq 1$). This captures any inefficiencies in the production or transportation of intermediate goods from manufacturing to agriculture. Second, wages are not equated across sectors, due to costs associated with traveling to the city, insufficient
skills for manufacturing work, uncertain job prospects, urban-rural cost of living differences, or any other number of factors. These are essentially implicit taxes on manufacturing wages. To capture this, I model an explicit proportional tax on manufacturing wage income. Tax revenue is rebated back to the village according to the function $T(b, z)$, which is known to the village. This allows me to focus specifically on the marginal distortion caused by wage differences.\(^3\)

### 3.1 Technology

**Manufacturing** The manufacturing sector is characterized by a stand-in firm which uses only labor services $N_{mt}$ to produce output according to the production function

$$Y_{mt} = AN_{mt}$$

where $A$ is a sector neutral TFP parameter. The parameter $A$ is country-specific, and is a measure of the overall productivity of the economy. For each $t$, the firm chooses $N_{mt}$ to maximize profit

$$\max_{N_{mt} \geq 0} AN_{mt} - w_t N_{mt} \quad (3.1)$$

where $w_t$ is the wage paid per unit of $N_{mt}$. In a competitive equilibrium $w_t = A$ for all $t$.

**Agriculture** Each village is endowed with one farm that requires sector $m$ intermediate inputs $x$ and labor $n_a$. Production occurs according to the decreasing returns to scale production function

$$y_{at} = z_t A x_t^\psi n_a^n, \psi + n < 1$$

where $A$ is, again, sector neutral TFP. The shock $z_t$ is a village-specific productivity shock drawn from a time-invariant distribution with c.d.f. $Q(z)$, with support on $[\underline{z}, \overline{z}]$. The realization of $z_t$ is i.i.d. with respect to both villages and time, and $\mathbb{E}(z_t) = 1$.

\(^3\)If tax revenue is instead thrown into the ocean, the intermediate good share would be changing both because of the lower manufacturing wage and the lower total income in the village. I want to focus only on the former. The tax decreases the opportunity cost of keeping the marginal worker on the farm.
3.2 Village

A village values consumption from both sectors $a$ and $m$, and maximizes total expected village utility given by

$$
E_0 \left[ \sum_{t=0}^{+\infty} \beta^t u(c_{at} - \bar{a}, c_{mt}) \right]
$$

with discount factor $\beta < 1$. The period $t$ utility flow takes the form

$$
u(c_{at} - \bar{a}, c_{mt}) = \alpha \log(c_{at} - \bar{a}) + (1 - \alpha) \log(c_{mt})
$$

where $c_{jt}$ is consumption from sector $j \in \{a, m\}$ and $\bar{a} > 0$ is subsistence requirement of the agricultural good. This assumption plays an important role in this analysis, and is discussed further in Section 3.3.1 after detailing the decision problem.

3.2.1 Decision Timing

At time $t - 1$, the village chooses to save $b_t$ units of the agricultural good. A fraction $\delta$ depreciates, and the village enters time $t$ with $(1 - \delta)b_t$ units of savings. This $\delta$ is allowed to vary across economies to capture differences in savings technologies across countries. The period $t$ decision problem of a village is broken down into two stages denoted planting and harvesting, which are separated by the realization of the idiosyncratic shock $z$.

In the planting stage, each village chooses intermediates $x_t$ to use in their farm. A unit of $x$ can be purchased for a price $p_x \geq 1$. This price is allowed to differ across countries but not time. Note that the assumption here is that there exists a technology that turns one unit of manufacturing output into $1/p_x$ units of intermediate input. Thus, $1/p_x$ defines the productivity of this technology relative to manufacturing output production.

Then, $z_t$ is realized. Recall that this shock is i.i.d. across villages and time. This is assumed for two reasons. First, rainfall is the primary source of income fluctuations in agrarian life. Since the model period is chosen to be a year, it is reasonable to assume the shock is i.i.d.. Second, in Bewley models such as this, the ability to self insure decreases as the persistence of the shock increases. In this sense, I am giving the village the best possible
chance to self-insure by assuming $z_t$ is i.i.d..

Turning to the harvesting stage, the village decides how to allocate labor between the agricultural sector, where they can work on a farm, and the manufacturing sector, where they can work for wage $w_t$. Profits are made, and consumption and savings choices $(c_{at}, c_{mt}, b_{t+1})$ take place. This timing is assumed to capture the fact that off-farm labor is an important form of insurance for farmers (Kochar, 1999; Ito and Kurosaki, 2009).

### 3.3 Recursive Problem

The timing described above implies that the village state variable is savings $b_t$, and the aggregate state is the distribution of savings across all villages, denoted $\mu_t(b)$. Since I will be studying the stationary equilibrium, I suppress the dependence of the decision problem on the aggregate state $\mu_t(b)$.

At the harvesting stage, once the choice of $x$ is made and $z$ realized, the value of entering time $t$ with $(1 - \delta) b$ savings is

$$V^H(x, b, z) = \max_{c_a, c_m, n_a, b'} \alpha \log(c_a - \bar{a}) + (1 - \alpha) \log(c_m) + \beta V^P(b')$$

subject to constraint set

$$p_a c_a + c_m + p_a b' = p_a z Ax^\psi n_a^\psi - p_a x + (1 - \tau) w(1 - n_a) + p_a (1 - \delta) b + T(b, z)$$

$$b' \geq 0$$

where $V^P$ is the value of entering the planting stage at $t + 1$ with $b'$ units of savings in the stationary equilibrium. The first constraint is the village budget constraint, and the second captures market incompleteness. Villages cannot borrow or trade claims to state-contingent asset with other villages, even though there would be gains to doing so. The harvesting problem in (3.2) defines decision rules $c_a(x, b, z), c_m(x, b, z), n_a(x, b, z)$ and $b'(x, b, z)$. Working

---

4This timing does create a minor issue however. The farmer chooses $x$ in the planting period before sector $m$ production occurs. Technically, the farmer commits to using $x$ intermediates, and then plants $x$ regardless of the realization of $z$. However looking ahead, the model period will be calibrated to a year. Since inputs can be purchased just days in advance of cropping, this abstraction seems reasonable.
backwards, the planting stage value of entering time $t$ with $b$ savings is

\[ V^P(b) = \max_{x \geq 0} \int_z V^H(x, b, z) dQ(z) \quad (3.3) \]

For future use, aggregate variables will be denoted by capital letters

\[ N_{at} = \int_b \left[ \int_z n_a(b, z) dQ(z) \right] d\mu_t \quad (3.4) \]
\[ X_t = \int_b x(b) d\mu_t \quad (3.5) \]
\[ Y_{at} = \int_b \left[ \int_z z A x(b)^\psi n_a(b, z)^n dQ(z) \right] d\mu_t \quad (3.6) \]

so that the domestic intermediate input share in agriculture can be written as

\[ \hat{X}_t = \frac{p_x X_t}{p_{at} Y_{at}} \quad (3.7) \]

3.3.1 Discussion

This section provides some discussion about two important features of the model: the form of the utility function and the savings technology.

The Role of Subsistence Requirements The period utility function assumed here is a simplified version of that proposed in Kongsamut, Rebelo, and Xie (2001). Qualitatively, it has two important features. First, it accounts for Engel’s law, so that the fraction of total income spent on agricultural output is decreasing in TFP $A$. Second, it provides an explanation for what Schultz (1953) calls the “food problem.” Countries must produce a certain amount of food to live. Since poor countries are less productive, they must employ a larger fraction of the population in agriculture to produce this critical amount of food. This provides a qualitative answer to why poor countries employ such a large fraction of the population in such an unproductive sector. Quantitatively, Herrendorf, Rogerson, and Valentinyi (2009) have shown that a general form of this utility function can replicate well the structural transformation process in the US. In estimating the utility function with the best fit to the data,
they find that $\bar{a} > 0$ is necessary to generate a good fit.

Given the empirically consistent predictions of the model, variations on this utility function have become commonplace in modeling the agricultural sector. This paper, however, exploits a feature of this utility function that has yet to be explored in a cross-country framework. Namely, subsistence requirements changes the relative risk aversion of a standard constant relative risk aversion (CRRA) utility function. To see this, first define $y$ as the optimal total income at the harvesting stage, given savings $b$, intermediate choice $X$, and shock $z$

$$y(x, b, z) = p_a z A x^\psi n_a^n - p_x x + (1 - \tau) w(1 - n_a) + p_a (1 - \delta) b + T(b, z) - p_a b'$$

Given this $y$, a village purchases enough to satisfy subsistence $\bar{a}$, then splits the rest of their income between the two sectors based on the relative weights assigned by the price $p_a$ and utility parameter $\alpha$.

$$c_a(y) = \bar{a} + \frac{\alpha}{p_a} (y - p_a \bar{a})$$
$$c_m(y) = (1 - \alpha)(y - p_a \bar{a})$$

Using these decision rules, the utility flow can be written as

$$u(c_a - \bar{a}, c_m) = \Omega - \alpha \log(p_a) + \log(y - p_a \bar{a})$$

where $\Omega = \alpha \log(\alpha) + (1 - \alpha) \log(1 - \alpha)$. The relative risk aversion with respect to total income $y$, given $\bar{a}$ and price $p_a$, is then

$$R(y|\bar{a}, p_a) = \frac{y}{y - p_a \bar{a}}$$

If $\bar{a} = 0$, this is a standard log CRRA utility function. However if $\bar{a} > 0$, the utility function instead exhibits decreasing relative risk aversion (DRRA), consistent with the household evidence of Ogaki and Zhang (2001) from both India and Pakistan.
With this form of the period utility function, harvesting utility can be written

$$V^H(x, b, z) = \Omega - \alpha \log(p_a) + \log(y - p_a\bar{a}) + \beta V^P(b'(x, b, z)) \tag{3.8}$$

The choice of $X$ is then the solution to

$$V^P(b) = \Omega - \alpha \log(p_a) + \max_{x \geq 0} \int_z \left[ \log(y(x, b, z) - p_a\bar{a}) + \beta V^P(b'(x, b, z)) \right] dQ(z) \tag{3.9}$$

Equations (3.8) and (3.9) illustrate the key tension between expected income and expected utility in the face of subsistence requirements. While profits drive harvest stage utility, the planting stage choice of $x$ maximizes expected utility, of which income is only one component. Another important component is the risk associated with the choice of $x$. While farm profit increases utility, higher $x$ implies large exposure to risk. To limit this exposure, and thus decrease the variation in harvest utility, the village must decrease the choice of $x$. Thus, the optimal choice of $x$ balances the need for both high income and low exposure to risk. Since $\bar{a} > 0$ implies DRRA, the inclusion of subsistence requirements can alter the way farmers undertake risky investments under different levels of TFP. After defining equilibrium, I show that this is indeed the case. The inclusion of subsistence requirements interacts with TFP differences and uninsurable risk to generate differences in the domestic intermediate share.

**Savings** Since subsistence requirements imply DRRA, it is intuitive then that the savings technology can potentially play an important role. Here, I assume that the only savings technology available is costly storage of the agricultural good, and insurance is not available. The lack of properly functioning insurance markets is certainly not controversial in developing countries. However, there are many ways to save around risk, and savings has been shown to be effective in limiting the impact of risk in Bewley models. This section discusses why this model assumes this primitive savings technology.

First, savings banks are generally not utilized. In addition to paying no interest, Dupas
and Robinson (2011) find that rural savings banks in Kenya actually charge both a start-up fee and a variable fee for every transaction. In 12 of 13 developing countries considered, Banerjee and Duflo (2007) find that less than 14% of all people living on under $1 a day have savings accounts.

Most liquid assets are instead accounted for by livestock and grain storage (Udry, 1995; Swinton, 1988). However, Fafchamps, Udry, and Czukas (1998) show that livestock sales do not seem to be used as a buffer stock in West Africa. This could be due in part to the fact that local markets are poorly integrated, so that local general equilibrium price adjustments make capital goods unable to be used as insurance. Even when they are traded in a way that resembles consumption smoothing, as Rosenzweig and Wolpin (1993) find in India, there is still severe underinvestment in bullocks. During the same West African drought period considered by Fafchamps, Udry, and Czukas (1998), Reardon, Matlon, and Delgado (1988) point out that cereal stocks were almost completely depleted. This suggests that agricultural output storage is the main form of buffer savings in the poorest countries. As one might suspect, storage technologies are heterogeneous between poor and rich countries. In Zimbabwe, for example, almost 30% of maize produced is lost due to storage. This is further detailed in the calibration of Section 5.

3.4 Equilibrium

Turning now to the equilibrium, I study the stationary competitive equilibrium of this economy. This is defined by an invariant distribution $\mu = \mu^*$, a value function $V_P$, decision rules $x, n_a, b', c_a, c_m$, labor choice $N_m$, prices $p_a$ and $w$, and a transfer function $T(b, z)$ such that

1. The value function $V_P$ solves the village's problem given by (3.2) and (3.3) with the associated decision rules

2. $N_m$ solves the sector $m$ firm problem (3.1)

3. Markets clear
(a) Manufacturing labor market:

\[ N_m = 1 - \int_b \int_z n_a(b, z) d\mu dQ(z) \]

(b) Agricultural consumption market:

\[ \int_b \int_z c_a(b, z) dQ(z) d\mu(b) = \int_b \int_z z A x(b)^\psi n_a(b, z)^\eta dQ(z) d\mu \]

(c) Manufacturing consumption market:

\[ \int_b \int_z c_m(b, z) dQ(z) d\mu + px \int_b x(b) d\mu = AN_m \]

4. The state contingent transfer balances for all \((b, z)\)

\[ T(b, z) = \tau w(1 - n_a(b, z)) \]

5. The law of motion for \(\mu\), denoted \(\mu'(\mu)\), is such that \(\mu'(\mu^*) = \mu^*\), and \(\mu^*\) is consistent with \(Q(z)\) and decision rules

4 Analytic Results

This section provides some analytic results to help clarify the mechanics of the interaction between TFP, uninsurable risk, and subsistence requirements. To make these results as sharp as possible, I consider a static version of the model (identically, \(\delta = 1\) for all economies). Within this simplified economy, the model can qualitatively replicate the correlation between intermediate input shares and income detailed in Section 2 if and only if \(\bar{a} > 0\). As mentioned previously, the distortions \(p_x\) and \(\tau\) are not necessary to generate the result. To show this, I also assume that \(\tau = 0\) and \(p_x = 1\) in all economies, so that the only difference between any two model economies is TFP. All proofs are relegated to Appendix B.

To assess the role of risk and incomplete markets, I compare the model to a modified
version with complete markets. In the complete markets model (CM economy), villages are allowed to trade a full set of state contingent assets before the realization of $z$. When these markets are not available (IM economy), production risk translates into consumption risk. The importance of this fact can be seen by comparing the first order conditions in the IM and CM economies with respect to $x$. Because consumption is fully insured against risk with complete markets, farmers maximize expected profit. This implies that the first order condition of the planting problem with respect to $x$ would be

$$Ap_a^{1/(1-\eta)} F'(x) \int_z z^{1/(1-\eta)} dQ(z) = 1 \quad (4.1)$$

where

$$F(x) = x^{\psi/(1-\eta)} \left( \eta^{\eta/(1-\eta)} - \eta^{1/(1-\eta)} \right)$$

However, without the ability to trade these claims (IM economy), the first order condition with respect to $x$ of the village’s planting problem yields

$$Ap_a^{1/(1-\eta)} F'(x) \int_z z^{1/(1-\eta)} \left( \frac{u'(y(x, z) - p_a \tilde{a})}{E_z[u'(y(x, z) - p_a \tilde{a})]} \right) dQ(z) = 1 \quad (4.2)$$

Equation (4.1) shows that the profit maximizing farm considers only the arithmetic mean of $z^{1/(1-\eta)}$. This changes with the addition of incomplete markets. Equation (4.2) shows that the village facing risky consumption considers a weighted average of $z^{1/(1-\eta)}$, where the weight is given by the marginal utility at the realization of $z$ relative to the mean (the “utility weight” at $z$). Those realizations of $z$ that imply a higher than average marginal utility are weighted relatively more heavily by a village that faces uninsurable risk. Similarly, those realizations of $z$ that imply a lower than average marginal utility are weighted less heavily. Thus, the inclusion of uninsurable risk tilts the weight assigned by every village toward “bad” outcomes. This leads naturally to Proposition 1.

**Proposition 1.** For a given $A$, the domestic intermediate share is lower in the IM economy
than the CM economy. That is,

$$\frac{X^{IM}}{p_a^{IM} Y_a^{IM}} < \frac{X^{CM}}{p_a^{CM} Y_a^{CM}} = \psi$$

Graphically, this result can be seen in Figure 4. In the CM economy, the utility weight is irrelevant. Put somewhat more formally, it is equal to one at every realization of $z$, and is given by the dotted line at 1. Once consumption risk is tied to production risk, however, this changes. The utility weight at low $z$ realizations increases, causing a decrease in the domestic intermediate share for all TFP levels $A$, which can be seen in the solid line. The more interesting issue, however, is how the intermediate input share reacts to changes in $A$, since the empirical evidence of Section 2 suggests they should be correlated. First, with $\bar{a} = 0$, uninsurable risk is irrelevant in accounting for the fraction of the labor force in agriculture, the domestic intermediate share, and agricultural productivity differences.

Proposition 2. In the model with uninsurable risk (IM economy) and $\bar{a} = 0$, the following results hold:

1. $n_a(z)$ is independent of $A$

2. The domestic intermediate share $X/(p_a Y_a)$ is independent of $A$

3. For two economies with TFP levels $A^1$ and $A^2$, productivity differences in the IM economy do not increase relative to the CM economy. That is,

$$\frac{(Y_a^{1CM}/N_a^{1CM})}{(Y_a^{2CM}/N_a^{2CM})} = \frac{(Y_a^{1IM}/N_a^{1IM})}{(Y_a^{2IM}/N_a^{2IM})}$$

While Proposition 1 shows that the equilibrium intermediate input share is lower with risk, Proposition 2 shows that when $\bar{a} = 0$, it does not differ across economies. When $\bar{a} = 0$, the period utility function exhibits CRRA, implying the utility weight in equation (4.2) is independent of TFP. This can be seen in the solid line of Figure 4. The utility weight for any realization of $z$ is identical for all levels of TFP. The inclusion of subsistence requirements breaks this result. When $\bar{a} > 0$, the period utility function exhibits DRRA, causing the
utility weight to depend on the level of TFP. Proposition 3 shows that subsistence requirements and uninsurable risk can qualitatively replicate the empirical correlation between the intermediate share and TFP from Section 2.

**Proposition 3.** If $\bar{a} > 0$, the domestic intermediate share is increasing in $A$.

TFP differences, uninsurable risk, and subsistence requirements combine to generate a result that is qualitatively consistent with the evidence provided in Figure 2. Leaving out any one of these features implies a constant intermediate share. The intuition for the result is similar to that of Proposition 1, but in a cross-economy context. Poor farmers have relatively less income than their rich counterparts for all realizations of $z$. With subsistence requirements, this difference increases as $z$ decreases. Since farmers weigh each realization of $z$ by their marginal utility at that realization, farmers in poor economies put relatively more weight on low $z$ than their rich counterparts, as can be seen in Figure 4. This causes the intermediate good share to decrease in economies with low $A$.

Figure 4: Utility Weight for Different Subsistence Levels

Given the theoretical relevance of the interaction between TFP, risk, and subsistence requirements, I now move back to the full dynamic model to investigate the quantitative magnitudes of the results shown here.
5 Quantitative Exercise and Calibration

The first quantitative experiment is to compare the predictions of the US model economy to a “poor” country. To do so, I first normalize $A = 1$ and calibrate a frictionless version of the model ($p_x = 1$ and $\tau = 0$) so that the stationary equilibrium matches a number of features of the US economy in 1985. The baseline calibration of the poor economy differs along four dimensions: TFP $A$, the depreciation rate $\delta$, the intermediate input price $p_x$ and tax rate $\tau$. To compare this to the data, I follow Caselli (2005) and compare the model to statistics of the 90th percentile country relative to the 10th percentile country. I construct an “average” 90th and 10th percentile country by averaging over the top 10 percent of countries, as ranked by GDP per capita in 1985. The 10th percentile country is the average of the bottom 15 to 5 percent of countries. This is detailed in Appendix A.

The second exercise is to vary $p_x$ and $\tau$ in the poor model economy while holding all other parameters fixed. This helps to understand the complementarity between price distortions and risk.\(^5\) This isolates the direct impact of risk for different levels of price distortions.

Section 5.1 presents the parameters that are the same across economies. Section 5.2 details the differences between the two economies in the baseline calibration, which are TFP $A$, storage depreciation $\delta$, intermediate input price $p_x$ and labor wedge $\tau$. Table 3 lists all the parameters chosen.

5.1 Common Parameters

In this section, I detail the parameters that are identical to both economies. They include the production technology (except for TFP), the shock distribution, and utility parameters.

Farm Production Parameters The farm production parameters are the share of intermediates, $\psi$ and labor, $\eta$. These are chosen to match the aggregate intermediate input share and labor share in agriculture in the United States in 1985. The exponent on intermediates is

\(^5\) The goal of this paper is not to explain these distortions but, given that they exist, to understand their interaction with risk. See Adamopoulos (2011) for the role of the transportation sector in accounting for high intermediate input prices in poor countries and Gollin, Lagakos, and Waugh (2011) for a quantitative exploration of the cause of the labor wedge.
set to $\psi = 0.40$ which implies that the intermediate share of slightly less than 0.40 in the baseline economy. This is consistent with Valentinyi and Herrendorf (2008) and Restuccia, Yang, and Zhu (2008), who find that this share is about 0.38. Since labor is chosen after the realization of all uncertainty, the parameter $\eta$ is exactly equal to the payments to labor as a share of gross agricultural output. I choose $\eta = 0.40$, which is consistent with the labor share in Restuccia, Yang, and Zhu (2008). Estimates of this parameter, however, vary widely and Section 7 considers the sensitivity of the results to this parameter.

**Farm Productivity Shock Distribution**  There are two possible choices for choosing the shock distribution. The first is to choose two separate shock distributions— one for the USA and a second for the “poor” economy. The second possibility is to assume that the distribution of shocks is the same between the two economies. It turns out that this decision is quantitatively irrelevant. Because the US economy is so far from subsistence, the distribution of shocks is of little quantitative importance. That is, these villages act similar to profit maximizers. When $A$ decreases however, villages become much more sensitive to this distribution because they are (ex-ante) closer to subsistence. Therefore the distribution is chosen to match the poor economy, and I make the innocuous assumption that the distribution is the same in the rich economy, since the mean is always normalized to one.

To estimate this distribution, I turn to the International Crops Research Institute for the Semi-Arid Tropics Village Level Surveys (ICRISAT VLS). These surveys contain plot level inputs and outputs from ten different Indian villages from the years 1975-1976 to 1983-1984. I use six of the villages that have data starting in 1975-1976. This data set has a few benefits. First, since I calibrate to the year 1985, it is almost perfectly aligned in terms of year coverage. Second, the villages were chosen to give an overview of the different agro-climatic zones in India. Therefore, I am not estimating risk for a village where, for example, it rains every year.

I give an overview of the procedure here, while a more detailed explanation is given

---

6The Rabi season runs through the winter and into the following year. I include this season in the year it starts. Thus, each village contains 9 years of data.
in Appendix C. First, I choose a fixed set of village specific prices. Since my model does not contain any aggregate risk, I do not want to include price fluctuations in my measure of risk. I choose to use the prices from 1975, and denote all inputs and outputs in terms of this 1975 price. From there I construct aggregated village level inputs, which include output, agricultural intermediates, nonagricultural intermediates, labor, capital, land, and agricultural output. One issue is that the data measures labor hours, while the model is calibrated to match the fraction of population in agriculture. While seemingly not the main driver of productivity differences across countries, Gollin, Lagakos, and Waugh (2011) shows that distinction can be important. With that caveat, I treat them as identical here.\footnote{In the model, hours and people can be used interchangeably. However, the model’s calibration and predictions are matched to data on number of people, not hours.} I subtract agricultural intermediate goods from output. Lastly, I combine capital and the value of land together as the fixed factor in the production function, which is normalized to one in the model.

I now have data equivalents of \( Y_a \) (gross output), \( I \) (agricultural intermediate inputs), \( X \) (nonagricultural intermediate inputs), \( N_a \) (agricultural labor), and the fixed factor. Though this is normalized to one in the model, I denote it \( K \) here for expositional purposes. This is the combined value of capital and land. From there, I calculate the Solow residual in village \( v \) at year \( t \) as

\[
\hat{z}_{vt} = \frac{Y^\text{data}_{vt} - I^\text{data}_{vt}}{(X^\text{data}_{vt})^{\psi}(N^\text{data}_{a,vt})^{\eta}(K^\text{data}_{vt})^{1-\psi-\eta}}
\]

Because TFP in this model remains fixed, I use a Hodrick-Prescott filter with smoothing parameter \( \lambda = 6.25 \), consistent with Ravn and Uhlig (2002), to detrend \( z^*_j \) and get values for \( z_{vt} \). Note that \( \mathbb{E}_v(z_{vt}) = 1 \) for all \( v \) by construction. Therefore, I am not biasing the results by assuming average TFP is lower in agriculture than in manufacturing. Since the shock is assumed to be i.i.d. across time and villages, so I can safely assume that each shock realization is drawn from the same distribution.

This procedure generates the discrete empirical probability density function displayed in Figure 5.
Admittedly, there are a relatively small number of observations from which to draw any absolute conclusions about the nature of agricultural risk. However, even within this small sample, notice that while there are some particularly good years, there are also some particularly bad years. Over five percent of the total probability is below $z = .25$, which seems to be roughly consistent with empirical evidence. Dercon (2002), for example, finds that 78% of households surveyed in the Ethiopian Rural Panel Data Survey had weather related harvest failure in the preceding twenty years. Haresaw, a village in the Tigray region of Ethiopia, has rainfall levels less than 40% of the median approximately every ten years (Dercon and Christiaensen, 2011), painting a somewhat more dire picture than here. The key in this model is that even though $Pr(z < .25) \approx .05$, farmers in developing countries weight this outcome more heavily than farmers in developed countries. Given this data, I use a simple distribution in which $z$ takes values in $[0.25, 0.70, 1.05, 1.25, 1.75]$ with respective probabilities $q(z) = [.05, .25, .45, .20, .05]$. This maintains the basic properties derived from the data without putting an exorbitant amount of weight on any particular realization from it.\footnote{Although the lowest value recovered is $z = 0.07$, I abstract from shocks that amount to an almost total loss of output. Including them forces the model to also consider malnutrition, since some villages fall below the subsistence requirements. This implies taking a stand on the complicated nature of bequests and family dynamics of rural life in developing countries. This is beyond the scope of this paper.}
Utility Parameters Since the model period is a year, I set $\beta = 0.96$. The remaining parameters are the weight on agricultural consumption, $\alpha$, and subsistence $\bar{a}$. The calibration strategy is as follows. First, choose $\alpha$, then choose $\bar{a}$ to match a fraction of the population in agriculture of 2.84% in the US model economy, consistent with empirical evidence. The magnitude of $\bar{a}$ plays a key role in this model since it controls the level of relative risk aversion. The higher the value of $\alpha$, the lower $\bar{a}$ can be chosen to match the employment moment.

Herrendorf, Rogerson, and Valentinyi (2009) estimate $\alpha = 0.02$ to match sectoral time series data in the US. This is slightly higher than choices common in the literature. Lagakos and Waugh (2010) choose $\alpha = 0.01$ and Restuccia, Yang, and Zhu (2008) chooses $\alpha = 0.0046$. Given the added importance of $\bar{a}$ in this model relative to these models without risk, I choose to be conservative and set $\alpha = 0.02$. To match a labor share in agriculture of 2.84%, $\bar{a} = 0.043$.

5.2 Economy Specific Parameters

The two economies differ along four dimensions: TFP $A$, depreciation of stored goods $\delta$, tax rate $\tau$, and intermediate input price $p_x$. Recall that $p_x$ changes one for one with the relative productiveness of the technology that turns $Y_m$ into $X$.

TFP For the US economy, TFP is normalized to $A = 1$. I discipline the TFP in the poor country by manufacturing labor productivity. Since manufacturing labor productivity is equal to $A$, I set $A = 0.25$ in the poor economy, which is roughly consistent with nonagricultural labor productivity differences between the richest and poorest countries.

Depreciation of Stored Goods To discipline the deprecation rate of agricultural storage, I use estimates of total storage losses in a number of African countries. Before proceeding to estimates of these storage losses, a distinction must be made between weight and quality losses. Since the model contains no notion of quality, the exact empirical counterpart would be depreciation of the value of agricultural output. However, quality losses are notoriously
difficult to measure, since they can depend on consumers’ preferences and cultural customs (Boxall, 2001; de Lucia and Assennato, 1994). Another issue is that quality and weight do not change one for one. For example, Boxall (2001) points out that some insects feed specifically on the germ of grain. While the germ contains less than 5% of the weight, it contains about 50% of the protein content. Thus, while these infestations may not result in huge weight losses, they do make crops less valuable on market. Boxall (2001) posits that including quality losses could increase the loss from insects by 15-30%.

With that caveat in mind, I focus specifically on weight losses. Post-harvest losses in developing countries are mostly generated before crops leave the farm (i.e. drying and storing crops), while losses in developed countries are mostly generated outside the farm gate (i.e. table waste by consumers). In developed countries, the advent of cold chain storage systems prolong storage life, making on-farm storage losses nearly irrelevant. In developing countries, crops are still dried by the sun and stored in the open. To put a number to these losses, I turn to the African Post Harvest Loss Information System (APHLIS). APHLIS is a network of local experts that aggregates statistics on weight loss into comparable measures across African countries and crops. Table 2 presents the estimated weight loss data for a number of crops in a selection of African countries.

Given these figures, I set $\delta = 0.15$ in the poor economy. It is worth emphasizing that this is a conservative estimate, as quality losses are not included. Increasing $\delta$ further would increase the results. I set $\delta = 0.03$ in the rich economy. Since the rich model economy has little need for precautionary savings, changing this value does not influence the results.

**Tax Rate** Since the US model economy is assumed frictionless, $\tau = 0$. For the poor model economy, I choose $\tau = 0.40$. This is roughly consistent with differences in labor wedges found in Restuccia, Yang, and Zhu (2008) and Vollrath (2009), both of whom study the importance of sectoral wage differences across countries.

---

9See Hodges, Buzby, and Bennett (2011).

10See Hodges et al. (2010) for a more complete review of APHLIS. Considering more countries only emphasizes the results. The large weight losses presented in Table 2 are present in almost all countries in the data set.
Table 2: Post Harvest Weight Loss (%) for Selected Countries and Crops for 2007

<table>
<thead>
<tr>
<th>Country</th>
<th>Maize</th>
<th>Wheat</th>
<th>Sorghum</th>
<th>Millet</th>
<th>Rice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eritrea</td>
<td>17.9</td>
<td>12.9</td>
<td>12.2</td>
<td>10.9</td>
<td>–</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>16.4</td>
<td>12.4</td>
<td>12.4</td>
<td>12.1</td>
<td>11.3</td>
</tr>
<tr>
<td>Kenya</td>
<td>21.1</td>
<td>12.9</td>
<td>12.7</td>
<td>11.9</td>
<td>13.2</td>
</tr>
<tr>
<td>Malawi</td>
<td>19.6</td>
<td>13.4</td>
<td>13.0</td>
<td>12.9</td>
<td>11.6</td>
</tr>
<tr>
<td>Mozambique</td>
<td>21.0</td>
<td>–</td>
<td>12.8</td>
<td>12.6</td>
<td>11.4</td>
</tr>
<tr>
<td>Rwanda</td>
<td>17.5</td>
<td>14.5</td>
<td>12.5</td>
<td>–</td>
<td>11.3</td>
</tr>
<tr>
<td>Sudan</td>
<td>18.0</td>
<td>12.9</td>
<td>12.2</td>
<td>10.7</td>
<td>–</td>
</tr>
<tr>
<td>Tanzania</td>
<td>22.0</td>
<td>14.4</td>
<td>12.5</td>
<td>12.3</td>
<td>11.2</td>
</tr>
</tbody>
</table>

**Median**

<table>
<thead>
<tr>
<th></th>
<th>Maize</th>
<th>Wheat</th>
<th>Sorghum</th>
<th>Millet</th>
<th>Rice</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>19.6</td>
<td>12.9</td>
<td>12.5</td>
<td>12.1</td>
<td>11.4</td>
</tr>
</tbody>
</table>

*Table notes: Data from APHLS*

**Intermediate Input Price**  The US intermediate price is set to $p_x = 1$. In the poor model economy, the intermediate price is $p_x = 2.5$. This is consistent with data from Restuccia, Yang, and Zhu (2008), who use FAO data to show that there is a strong correlation between per capita income and intermediate input prices across countries. Table 3 summarizes the parameters.

6  Quantitative Results

The baseline model results are presented in Table 4. The first two columns presents the model results for the two economies in the world with no production risk (i.e. $Pr[z = 1] = 1$), but still assuming the differences in $(p_x, \tau)$ calibrated above. The second set of columns presents the results of the model when production risk takes the form calibrated in Section 5.1. For comparison, the last column contains the statistical counterparts from the data.

The addition of risk amplifies agricultural productivity differences between the two economies by 52% and aggregate productivity differences by 57%, getting significantly closer
Table 3: Parameter Values for Two Economies

<table>
<thead>
<tr>
<th>Parameter</th>
<th>US</th>
<th>Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\psi$</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td>$\eta$</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>$\bar{a}$</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>$z, Q(z)$</td>
<td>(see text)</td>
<td></td>
</tr>
</tbody>
</table>

Specific

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>1</td>
</tr>
<tr>
<td>$\tau$</td>
<td>0</td>
</tr>
<tr>
<td>$p_x$</td>
<td>1</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.03</td>
</tr>
</tbody>
</table>

This amplification is driven by differences in the two agricultural inputs - intermediate goods and labor. The model with no risk predicts no change in the intermediate input share across countries, an artifact of the Cobb-Douglas form of production. Once risk is added, the intermediate share prediction for the poor economy decreases from 0.40 to 0.18, even while maintaining the Cobb-Douglas production function. The model captures 71% of the difference between the rich and poor countries, in which the poor countries have an average intermediate input share of 0.09. Turning to the fraction of the population working in agriculture, the model with risk again performs significantly better than the model without risk. The prediction of the agricultural labor force increases from 44.1 to 65.4, an increase of 47%.

11 Though not listed here, the baseline model also predicts a factor of 29.1 difference in value added per worker in agriculture when measured at US prices. This is 65% of the difference of 45 found by Caselli (2005) when measured in a similar manner. I focus here on output per worker because first, it is the relevant denominator when considering the intermediate input share, and second, it takes into account the role of intermediate inputs on labor productivity.
## Table 4: Baseline Model Results

<table>
<thead>
<tr>
<th>Economy</th>
<th>Model: no risk</th>
<th>Model: risk</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rich</td>
<td>Poor</td>
<td>Rich</td>
</tr>
<tr>
<td><strong>Output per worker</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>20.7</td>
<td>1.0</td>
<td>31.4</td>
</tr>
<tr>
<td>Aggregate</td>
<td>5.3</td>
<td>1.0</td>
<td>8.3</td>
</tr>
<tr>
<td><strong>Inputs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediate share</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td>Agricultural labor (%)</td>
<td>2.8</td>
<td>44.4</td>
<td>2.8</td>
</tr>
</tbody>
</table>

### 6.1 Interaction Between Risk and Price Distortions

Price distortions can play a major quantitative role in accounting for cross-country productivity differences in models with no risk. This is due to a substitution effect; since the cost of intermediate inputs rises relative to labor, villages substitute labor for intermediate inputs. Since labor exhibits decreasing returns, this drives down agricultural productivity. Adding uninsurable production risk generates an additional margin - an income effect - through which price distortions can affect productivity. As the intermediate price increases, expected income decreases. Because of subsistence requirements and uninsurable risk, villages limit their exposure to risk by further reducing intermediate input usage. In the model, this shows up as a decrease in the domestic intermediate input share.

In this section, I set \((p_x, \tau) = (1, 0)\) in all economies and repeat the exercise. The first thing to notice is the importance of the substitution effect. This can be seen by comparing the \textit{Model: no risk} columns in Tables 4 and 5. Agricultural output per worker differences between the two economies doubles with the introduction of price distortions \(p_x\) and \(\tau\), as emphasized in Restuccia, Yang, and Zhu (2008). Second, adding production risk is of little quantitative importance without price distortions. While the baseline model increases the predictions of agricultural productivity differences by 52 percent, the model with no
Table 5: Model Results with No Price Distortions

<table>
<thead>
<tr>
<th>Economy</th>
<th>Model: no risk</th>
<th>Model: risk</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rich</td>
<td>Poor</td>
<td>Rich</td>
</tr>
<tr>
<td><strong>Output per worker</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agriculture</td>
<td>10.2</td>
<td>1.0</td>
<td>10.7</td>
</tr>
<tr>
<td>Aggregate</td>
<td>5.1</td>
<td>1.0</td>
<td>5.1</td>
</tr>
<tr>
<td><strong>Inputs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediate share</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td>Agricultural labor (%)</td>
<td>2.8</td>
<td>22.1</td>
<td>2.8</td>
</tr>
</tbody>
</table>

price distortions implies an amplification of just 5 percent. While it is still true that the intermediate input share decreases in the poor country, it does so by just 10 percent. This drop is less than one fifth the magnitude of the intermediate input share decrease in the in the model with the calibrated price distortions. To understand the role of risk, it is key to include sector-specific price distortions.

7 Robustness [to be completed]

8 Conclusion

This paper quantifies the role of idiosyncratic production risk in accounting for sectoral output per worker differences in a two sector general equilibrium model. In poor countries, farmers use fewer intermediate inputs, driving down agricultural productivity. Even though Cobb-Douglas farm technologies differ by only a factor-neutral TFP parameter, the model is consistent with the fact that poor countries have lower domestically priced intermediate input shares in agriculture. The model captures about 70% of this difference between the richest and poorest countries. Technically, this result is due to the interaction of uninsurable risk with DRRA preferences driven by subsistence requirements. The model also provides a new
channel through which sector specific distortions can impact productivity. Since distortions decrease income, they feed back into even lower choices of intermediates.

The model can be used to study a number of different extensions. One interesting example is estimating welfare gains from different policy interventions. While the most obvious welfare improving policy change is better insurance markets, this has proved somewhat difficult to implement in reality. Is it possible for better self-insurance mechanisms to provide similar welfare gains? In the model, better self-insurance can arise from better savings technologies or lower price distortions. If the welfare gains are somewhat similar, the results may provide policies that can significantly decrease the impact of risk, without being subject to the issues that have plagued lending programs, such as moral hazard and high default rates. Recent work on mobile phone transfers has shown that they are able to overcome physical distance issues and act as a type of insurance network (Blumenstock, Eagle, and Fafchamps, 2011).

Lastly, while the model focuses on the impact of risk on intermediate inputs, it has the potential to account for other important factors determining agricultural productivity. For one, what generates the implicit labor tax used here? Qualitatively, the addition of a risky manufacturing sector can. If the uninsurable risk in manufacturing is larger than that in agriculture, there will exist a premium for moving to that sector. Future research will explore these ideas.
References


Appendices

A Data Sources and Construction

A.1 Productivity and Intermediate Input Share Statistics

I make use of the publicly available data from Restuccia, Yang, and Zhu (2008) for statistics on aggregate productivity, agricultural productivity, labor, and intermediate input prices. This is augmented with purchasing power parities (PPP) for agricultural output and non-agricultural intermediate inputs from Prasada Rao (1993). The resulting dataset contains 84 countries, which are:

Algeria, Angola, Argentina, Australia, Austria, Bangladesh, Belgium, Bolivia, Brazil, Burkina Faso, Burundi, Cameroon, Canada, Chad, Chile, Columbia, Costa Rica, Côte d’Ivoire, Democratic Republic of the Congo, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Ethiopia, Finland, France, Germany, Ghana, Greece, Guatemala, Guinea, Haiti, Honduras, Hungary, India, Indonesia, Iran, Iraq, Ireland, Israel, Italy, Japan, Kenya, Korea, Madagascar, Malawi, Malaysia, Mali, Mexico, Morocco, Mozambique, Nepal, Netherlands, New Zealand, Nicaragua, Niger, Nigeria, Norway, Pakistan, Papau New Guinea, Paraguay, Peru, Philippines, Portugal, Rwanda, Senegal, Somalia, South Africa, Spain, Sri Lanka, Sudan, Sweden, Switzerland, Syria, Thailand, Tunisia, Turkey, U.K., U.S.A., Uganda, Uruguay, Venezuela, and Zimbabwe.

A.1.1 Productivity

I am interested a measure of the ninetieth percentile country relative to the tenth percentile country, similar to that used in Caselli (2005). As a measure of the rich country, I take average of the top ten percent of countries. Listed from largest to smallest income, they are USA, Canada, Switzerland, Australia, Norway, Netherlands, Belgium, and Germany. As a measure of the “tenth” percentile, I take an average of the countries that make up the bottom fifteen to five percent of countries, as ranked by PPP GDP per capita. They are
Somalia, Rwanda, Mozambique, Uganda, Malawi, Chad, Zaire and Niger.

The productivity statistics are taken from Restuccia, Yang, and Zhu (2008). They are derived from PWT and FAO data. These averages imply a factor of 63.66 difference in agricultural output per worker and 23.18 difference in aggregate output per worker. On average, 82% of the population in the poor countries work in agriculture.

A.1.2 Intermediate Input Shares

As in the text, the domestic intermediate share in agriculture of country \( j \) is

\[
\hat{X}^j := \frac{p^j_i X^j}{p^j_a Y^j_a} \tag{A.1}
\]

This measure is not directly reported in Prasada Rao (1993). He does however, report the real intermediate share in agriculture, defined by equation (2.2). Equation (2.3) in the text shows the relationship between the domestically priced and internationally priced intermediate shares in country \( j \). The price ratio in equation (2.3) can be calculated from reported purchasing power parities

\[
(PPP^j_a = \frac{p^j_a}{p^*_a}, PPP^j_x = \frac{p^j_x}{p^*_x})
\]

where \( p^*_a \) and \( p^*_x \) are international (unreported) prices and \((p^j_a, p^j_x)\) are (unreported) domestic prices for country \( j \). The purchasing power parities are normalized to one in a baseline country, which in Prasada Rao (1993) is the USA. Therefore, \( PPP^US_a = PPP^US_x = 1 \), implying \( \hat{X}^{US} = \hat{X}^{US*} \). Therefore, calculating the domestically priced intermediate share of all other countries reduces to

\[
\hat{X}^j = \hat{X}^{j*} \left( \frac{PPP^j_x}{PPP^j_a} \right) \tag{A.2}
\]

As mentioned, the real intermediate share and the ratio of PPPs are both reported, so this is sufficient to define the domestically priced intermediate input share. The poor group group
of countries has, on average, a domestically priced intermediate input share of 0.09 and a real intermediate input share of 0.13. The left hand side of equation (A.2) is the statistic reported in Figure 2. The horizontal axis, GDP per capita, is real GDP per capita for 1985, variable cgdpc from the Penn World Tables version 7.0 (PWT).

A.2 Three Sector Comparison: UN System of National Accounts

For the comparison of agriculture to manufacturing and services, I use a set of 49 countries from the UN SNA. The 49 countries with sufficient data for all three sectors Austria, Benin, Bolivia, Botswana, Burundi, Cameroon, Canada, Cape Verde, Chile, Hong Kong, Colombia, Cyprus, Denmark, Ecuador, El Salvador, Fiji, Finland, France, Gambia, Germany, Ghana, Hungary, Iceland, Italy, Jamaica, Japan, Jordan, Republic of Korea, Luxembourg, Mauritius, Mexico, Netherlands, New Zealand, Nigeria, Norway, Peru, Portugal, Rwanda, Seychelles, Sierra Leone, Spain, Sri Lanka, Swaziland, Sweden, Syrian Arab Republic, United Kingdom, Uruguay, Venezuela, and Zimbabwe.

From the UN SNA, I use “Output, at basic prices” and “Intermediate consumption, at purchaser’s prices” for the year 1985 for each of the three sectors. Dividing them gives the domestically priced intermediate input share by sector. Figure 3 plots this, along with variable cgdpc for 1985 from PWT on the horizontal axis.

B Proofs

B.1 Proof of Proposition 1

Proof. First, it’s easy to show that

\[ \frac{X_{CM}}{p_{a}^{CM} Y_{a}^{CM}} = \psi \]  \hspace{1cm} (B.1)
Define \( X^* \) be the optimal choice for a farmer facing \( p_a^R \), but facing no risk. Then the first order condition implies,

\[
\frac{X^*}{P_a^{IM} Y_a(X^*)} = \psi
\]

(B.2)

Comparing (B.1) and (B.2), proving the proposition is equivalent to proving that at the price \( p_a^{IM} \), \( X^R < X^* \). This can be seen from the first order conditions. The first order condition in the planting problem is

\[
A p_a^{1/(1-\eta)} F'(X) \int_Z \left( \frac{u'(y(X, z) - p_a \bar{a})z^{1/(1-\eta)}}{E_z[u'(y(X, z) - p_a \bar{a})]} \right) dQ(z) = 1
\]

(B.3)

where \( y(X, z) \) is total optimal income and \( F(X) \) is defined as in the text. Note that \( F(\cdot) \) is concave because \( \psi + \eta < 1 \). Now, consider the profit maximizing problem. The first order condition is

\[
A p_a^{1/(1-\eta)} F'(X^*) \int_Z z^{1/(1-\eta)} dQ(z) = 1
\]

(B.4)

Since \( u(\cdot) \) is concave,

\[
\int_Z \left( \frac{u'(\tilde{y}(X^*, z))z^{1/(1-\eta)}}{E_z[u'(\tilde{y}(X^*, z))]} \right) dQ(z) < \int_Z z^{1/(1-\eta)} dQ(z)
\]

(B.5)

Since \( F(X) \) is concave, it follows that \( X^{IM} < X^* \).

B.2 An Additional Lemma for the Proof of Proposition 2

To prove the result, I first characterize the the equilibrium of an IM economy with TFP \( A^2 \) and \( \bar{a} = 0 \) in terms of an economy with TFP \( A^1 \) and \( \bar{a} = 0 \). This is done in Lemma 1 below.

**Lemma 1.** Consider two IM economies characterized by TFP levels \( A^1 \) and \( A^2 \), both with \( \bar{a} = 0 \). Denote the equilibrium for economy 1 as \((X^1, N^1_a(z), p^1_a)\). Then the equilibrium for
economy 2, \((X^2, N^2_a(z), p^2_a)\) can be characterized as

\[
N^2_a(z) = N^1_a(z) \\
X^2 = \left(\frac{A^2}{A^1}\right) X^1 \\
p^2_a = \left(\frac{A^1}{A^2}\right) p^1_a
\]

**Proof.** Two things must be checked for the proposed allocation to be a competitive equilibrium. First, the proposed equilibrium must satisfy the village optimization problem. That is, if \((p^1_a, X^1, N^1_a(z))\) is an equilibrium in economy 1, then \((p^2_a, X^2, N^2_a(z))\) satisfies the farmer’s optimization problem in economy 2. Second, markets must clear. These are considered in turn.

**Optimization Problem** First I show that the labor choice problem is satisfied. Using the optimal decision rule \(N_a(X, z)\),

\[
\frac{N^1_a(z)}{N^2_a(z)} = \left(\frac{p^1_a A^1(X^1)^\psi}{p^2_a A^2(X^2)^\psi}\right)^{1/(1-\eta)}
\]

Plugging in \((p^2_a, X^2)\) implies

\[
\frac{N^1_a(z)}{N^2_a(z)} = 1
\]

When \(\bar{a} = 0\), the harvesting utility for a given income \(y\) can be written as

\[
V^H(y) = \alpha \log(c^1_d) + (1 - \alpha) \log(c^1_m) \\
= \Omega - \alpha \log(p^1_a) + \log(y)
\]

where \(\Omega = \alpha \log(\alpha) + (1 - \alpha) \log(1 - \alpha)\). Denote the income of a farmer who chooses intermediates \(X\) and gets hit with shock \(z\) in economy 1 as

\[
y^1(X, z) = p^1_a A^1 z X^\psi N^1_a(X, z)^\eta - X + (1 - N^1_a(X, z))A^1
\]
Then the harvesting utility of a farmer in economy one can be written as

\[ V^H(X, z) = \Omega - \alpha \log(p^1_a) + \log(y^1(X, z)) \]

This implies that

\[ X^1 = \arg \max_X \int_Z \log(y^1(X, z))dQ(z) \]

For ease of notation, denote

\[ F(X, N(X)) = X^\psi N_a(X, z)^n \]

The first order condition for the farmer in economy one is

\[ \int_Z \left( \frac{1}{y^1(X^1, z)} \right) \left[ p^1_a A^1 z \left( F^1_X + F^1_N N^1_a(X^1, z) \right) - 1 - A^1 N^1_a(X^1, z) \right] = 0 \] (B.6)

where \( F^1_j \) is the partial derivative of \( F \) with respect to input \( j \) and \( N^1_a \) is the derivative of the labor choice with respect to \( X \) for the economy 1 farmer. Plugging in the proposed equilibrium for economy two yields the following relations

\[ F^2_X = \left( \frac{A^2}{A^1} \right)^{\psi^{-1}} F^1_X \]

\[ F^2_N = \left( \frac{A^2}{A^1} \right)^{\psi} F^1_N \]

\[ N^2_a(z, X^2) = \left( \frac{A^2}{A^1} \right) N^1_a(X^1, z) \]

\[ y^2(X^2, z) = \left( \frac{A^2}{A^1} \right) y^1(X^1, z) \]

The first order condition for a farmer in economy two is

\[ \int_Z \left( \frac{1}{y^2(X^2, z)} \right) \left[ p^2_a A^2 z \left( F^2_X + F^2_N N^2_a(X^2, z) \right) - 1 - A^2 N^2_a(X^2, z) \right] = 0 \] (B.7)

Plugging in the relations found above, and recalling that \( p^2_a = (A^1/A^2)^\psi p^1_a \), implies that
equation (B.7) can be simplified to
\[
\left( \frac{A_1}{A_2} \right) \int_Z \left( \frac{1}{y^1(X^1, z)} \right) \left[ p^1_a A^1 z \left( F^1_X + F^1_N N^1_a(X^1, z) \right) - 1 - A^1 N^1_a(X^1, z) \right] = 0
\] (B.8)

Comparing equations (B.6) and (B.8) implies that since \((X^1, N^1_a(X^1, z))\) solves the farmer’s optimization problem in economy 1, \((X^2, N^2_a(X^2, z))\) solves the farmer’s optimization problem in economy 2.

**Market Clearing**  Aggregate sector \(a\) output for economy \(j = 1, 2\) is
\[
Y^j_a = AX^\psi E_z(z N^j_a(X^j, z)^\eta)
\]
Thus,
\[
\frac{Y^1_a}{Y^2_a} = \left( \frac{A_1}{A_2} \right) \left( \frac{X^1}{X^2} \right)^\psi
\] (B.9)
Therefore, at the proposed equilibrium,
\[
\frac{Y^1_a}{Y^2_a} = \left( \frac{A_1}{A_2} \right)^{1+\psi}
\] (B.10)
For any \(\bar{a} \geq 0\), the total demand for sector \(a\) consumption is given by
\[
D^j_a = (1 - \alpha)\bar{a} + \frac{\alpha}{p^j_a} E_z[y^j(X^j, z)]
\] (B.11)
As pointed out above, \(y^2(X^2, z) = (A^2/A^1) y^1(X^1, z)\). Therefore,
\[
\frac{E_z[y^1(X^1, z)]}{E_z[y^2(X^2, z)]} = \frac{A^1}{A^2}
\] (B.12)
Since \(\bar{a} = 0\), equations (B.11) and (B.12) and the prices \(p^1_a\) and \(p^2_a\) imply that
\[
\frac{D^1_a}{D^2_a} = \left( \frac{A_1}{A_2} \right)^{1+\psi}
\] (B.13)
Since the proof assumes an equilibrium in economy 1, equations (B.10) and (B.13) imply \( Y_2^a = D_2^a \). Since the labor market in sector \( m \) clears trivially, Walras’ Law implies that the sector \( m \) output market also clears.

\[ \square \]

### B.3 Proof of Proposition 2

**Proof.** With Lemma 1 in hand, the three claims follow quickly.

#### B.3.1 \( N_a(z) \) is independent of \( A \)

This follows directly from Lemma 1.

#### B.3.2 The intermediate input share is independent of \( A \)

Denote \( \hat{X}^j \) as the intermediate good share in economy \( j = 1, 2 \), so that \( \hat{X}^j \) is defined as

\[
\hat{X}^j = \frac{X^j}{p_a^j Y_a^j}
\]  

(B.14)

First, note that total agricultural output in economy \( j \) is given as

\[
Y_a^j = A(X^j)^\psi E_z(z N_a^j(z)^\eta)
\]  

(B.15)

Using the fact that \( N_a^1(z) = N_a^2(z) \) and plugging (B.15) into (B.14) gives

\[
\frac{\hat{X}^1}{\hat{X}^2} = \left( \frac{X^1}{X^2} \right)^{1-\psi} \left( \frac{p_2^a}{p_1^a} \right) \left( \frac{A^2}{A^1} \right)
\]  

Plugging in the equilibrium found in Lemma 1, this gives

\[
\frac{\hat{X}^1}{\hat{X}^2} = \left( \frac{A^1}{A^2} \right)^{1-\psi} \left( \frac{A^1}{A^2} \right)^\psi \left( \frac{A^2}{A^1} \right)
\]

\[ = 1 \]

Since \( A^1 \) and \( A^2 \) are arbitrary, this completes the proof.
B.3.3 No increase in productivity relative to CM model

For any two economies characterized by TFP $A^1$ and $A^2$ and no risk, it is easy to show that in equilibrium,

\[
N_a^1 = N_a^2 \\
X^2 = \left( \frac{A^2}{A^1} \right) X^1
\]

Since this is the same as in the risk model, relative agricultural productivity between the two economies is equal in both the risk and the no risk worlds.

\[\blacksquare\]

B.4 Proof of Proposition 3

*Proof.* Consider the equilibrium for economy 1 with TFP equal to $A^1$. Denote this equilibrium $(p^1_a, X^1, N_a^1(z))$. Suppose that the intermediate good share is $\hat{X}^1 < \psi$, where the inequality follows from Proposition 1. Define $X^{1CM}$ to be the optimal choice of the farmer who faces $p^1_a$ and no risk. We know that the intermediate good share is $\hat{X}^{1CM} = \psi$. Therefore, the ratio is

\[
\frac{\hat{X}^1}{X^{1CM}} = \frac{\hat{X}^1}{\psi} = \left( \frac{X^1}{X^{1CM}} \right)^{(1-\eta-\psi)/(1-\eta)}
\]

Thus, we can write $\hat{X}^1$ as

\[
\hat{X}^1 = \psi \left( \frac{X^1}{X^{1CM}} \right)^{(1-\eta-\psi)/(1-\eta)}
\]

Similarly, it follows that in economy 2,

\[
\hat{X}^2 = \psi \left( \frac{X^2}{X^{2CM}} \right)^{(1-\eta-\psi)/(1-\eta)}
\]

These equations show that the intermediate good share is directly related to how “far” the optimal choice of $X$ is from the choice $X^{CM}$. What’s left to show is that when $\bar{a} > 0$ and $A^1 > A^2$,

\[
\frac{X^1}{X^{1CM}} > \frac{X^2}{X^{2CM}}
\]
This follows from the fact that, when \( \bar{a} > 0 \), relative income net of subsistence,

\[
\frac{y^1(z) - p^1 \bar{a}}{y^2(z) - p^2 \bar{a}}
\]

is decreasing in \( z \).

\[\blacksquare\]

C Calibration of Shocks

The data used was collected by ICRISAT. I use the version that was released by Stefan Dercon, via the Oxford University website. It is publicly available at http://www.economics.ox.ac.uk/members/stefan.dercon/icrisat/ICRISAT/oldvls.html.

The ICRISAT VLS is an unbalanced panel set covering 10 villages in India. The data covers the time period 1975 - 1984.

The goal is to calculate the value of the following inputs at the village level: capital \( K \), agricultural intermediates \( I \), nonagricultural intermediates \( X \), human labor hours \( N_a \), and land \( L \). Allowing for some abuse of notation, let these letters also denote the set of all inputs of that type, so \( K \) is the set of all capital goods in the economy, for example.

C.1 Prices

For concreteness, I explain the prices in terms of the set of nonagricultural intermediate good set \( X \).

I use a constant set of prices so that I do not include price fluctuations in my value of uncertainty. Price construction proceeds as follows. First, all prices are considered at the village level. Therefore, fix a village \( v \). I use 1975 as my base year. Prices are available only by imputing them from the quantity and value of inputs used. Therefore, if input \( x \in X \) is used at \( t = 1975 \), the price I use is

\[
p_{v,x} = \frac{TV_{v,x}}{Q_{v,x}}
\]

where \( TV \) and \( Q \) are total value and quantity of input \( x \) at the plot level. Notice that this implies that if input \( x \) is not used every period, I cannot calculate a \( p_{v,x,t} \) at every period.
Thus, if \( x \) is not used at \( t = 1975 \), I need to calculate some price \( p_{v,x,1975} \). I use the following procedure. First, get all prices \( p_{v,x,t} \) for all \( x \in X \) and all \( t \) in \( v \). Denote the set of inputs that have prices available for all years as \( X' \subseteq X \). Construct average change in price over all inputs \( x' \in X' \). This gives me an average price change for 1975 – 1976 (denoted \( g_{75}^{X} \)), another for 1976 – 1977 (\( g_{76}^{X} \)), etc. Note that this average price change is specific to input set \( X \). That is, I do not require \( g_{75}^{X} = g_{75}^{K} \), for example.

Now consider an input \( x \in X \) that is first used in \( v \) at \( t^* \neq 1975 \). Then the first available price for \( x \) is \( p_{v,x,t^*} \). To construct the 1975 price \( p_{v,x,1975} \), I deflate the value by the relevant growth rates

\[
p_{v,x,1975} = \frac{p_{v,x,t^*}}{\prod_{t=1975}^{t^*-1} (1 + g_t^{X})}
\]

Since the prices remain constant over time, I refer to the price of input \( x \) in village \( v \) as \( p_{v,x} \), dropping the time subscript.

C.2 Construction of Inputs

The data includes 5 inputs: Capital, land, human labor, non-agricultural intermediates, and agricultural intermediates.

C.2.1 Capital

I use class code \( E \), farm equipment and implements and class code \( M \), major farm machinery, and class code \( R \), production capital assets. Class code \( E \) includes basic farm equipment such as plows and hoes. Class code \( M \) includes major machinery such as tractors and electric pumps.

A key capital component in agriculture is productive animals. Therefore, I also include bullock labor hours at the plot level, both owned and rented bullocks. The value of an hour of an owned bullock is imputed from the rental rates of hired bullock hours, so they are valued equally. Thus, the value of bullock hours on plot \( p \) owned by family \( f \) in village \( v \) at time \( t \) is given as

\[
B_{f,p,v,t} = r_v^b (B_{f,p,v,t}^o + B_{f,p,v,t}^r)
\]
where the rental rate $r_v^b$ is computed using the technique described above.

Combining these two values gives the total capital input for production on plot $p$ owned by family $f$ in village $v$ at time $t$, denoted $K_{f,p,v,t}$.

$$K_{p,f,v,t} = \left( \sum_{k \in K} p_{k,v} Q_{k,f,p,v,t} \right) + B_{f,p,v,t} \quad \forall (p, f, v, t)$$

### C.2.2 Human Labor

The $Y$ files give hours of male, female, and child labor in the data. Since I calibrate to match the fraction of the population over 15 years of age, I include only male and female labor. Child labor is a small component with the lowest price (i.e. not as productive as an adult laborer). Including it makes no discernible difference. Similar to bullock labor hours, the $Y$ files include disaggregated data on both family and hired workers. Once again though, the value of family labor is imputed from market value, so they are valued equally.

Letting $H^f$ and $H^h$ denote family and hired hours, the total value of labor on plot $(p, f, v, t)$ is given by

$$N_{p,f,v,t} = \sum_{j = M, F} w_{v,j} \left( H^f_{p,f,v,t} + H^h_{p,f,v,t} \right)$$

where $w_{v,j}$ is the value per hour of $j \in \{M, F\}$.

### C.2.3 Non-agricultural Intermediates

Nonagricultural intermediates include pesticides, which are input codes $1A−9A$, and fertilizer (input codes $A−Z$). Therefore, the total non-agricultural intermediate goods on a given plot is

$$X_{p,f,v,t} = \sum_{x \in X} p_{v,x} Q_{x,p,f,v,t}$$

### C.2.4 Agricultural Intermediates

Agricultural intermediates can be included by using the $Y$ files. Organic manure in the data are inputs $1−7$. Seed is denoted as inputs $CA−ZK$. The quantity and values are in the
Y files. As with the other inputs, the total value of agricultural intermediate inputs on plot $p$ owned by farmer $f$ in village $v$ in year $t$ is

$$I_{p,f,v,t} = \sum_{i \in I} p_{v,i} Q_{i,p,f,v,t}$$

**C.2.5 Land**

A key issue with modeling idiosyncratic risk is things that may look like risk for the economic modeler may actually be unmeasured idiosyncratic differences. Since I do not want to include idiosyncratic differences as risk, I include land quality in my construction of land input. To do so, I use plot value in the $PS$ file. This is the value of the land, given as Value of Land (Y-11) in the manual and is 'Rs. 100' per acre. I allow this value to vary over time. Therefore, I do not assume that each plot has the same value in every period. My measure of land $L$ then is

$$L_{p,f,v,t} = TV_{p,f,v,t}^L$$

where $TV_{p,f,v,t}^L$ is the total value of plot $p$ owned by farmer $f$ in village $v$ at time $t$.

**C.3 Output**

Total value of output is given by summing over output values in the $Y$ files by plot level. I include both actual production and by-products produced by farming. Since sometimes more than one crop is planted on a plot, I sum over all outputs on the plot. Letting $Y$ denote the possible set of outputs, total output on a given plot is

$$Y_{p,f,v,t}^a = \sum_{y \in Y} p_{y,v} Q_{y,p,f,v,t}$$

**C.4 Decomposition of Residuals**

As it currently stands, the data are at the $(p, f, v, t)$ level. The next step is to sum over $(p, f)$ to get village $v$ values of inputs and output at year $t$. This now gives me the input
vector \((K, L, N_a, I, X)_{v,t}\). Now, I can calculate the residual

\[
\tilde{z}_{v,t}^* = \frac{Y_{v,t} - I_{v,t}}{N_{\psi}^\eta (K_{vt} + L_{vt})^{1-\psi-\eta}}
\]

where \(\eta\) and \(\psi\) are taken from the calibration in the main text. The rest is explained in the main text.