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

Kevin Donovan†
University of Notre Dame

This draft: December 2016
First draft: December 2012

Abstract

Agricultural productivity is crucial for understanding aggregate cross-country income differences. This paper considers the impact of low intermediate input intensity in developing countries. In a dynamic general equilibrium model with idiosyncratic shocks, incomplete markets, and subsistence requirements, farmers in developing countries rationally use fewer intermediate inputs because it limits their exposure to uninsurable shocks. The calibrated model accounts for nearly half of the difference in intermediate shares between the U.S. and India and amplifies differences in per capita GDP by twenty percent relative to an identical model with perfect insurance.

JEL Classification Codes: O11, O41

Keywords: Agriculture, intermediate inputs, misallocation, productivity, risk

*For comments and insights I thank conference and seminar participants at Arizona State, ITAM, Notre Dame, Rochester, the St. Louis Fed, Virginia, The World Bank, the SED Annual Meeting, the CASEE Macro Reunion Conference, and the Econometric Society NASM, especially those from Alex Bick, Berthold Herrendorf, Chris Herrington, Joe Kaboski, David Lagakos, Ed Prescott, B. Ravikumar, Richard Rogerson, and Todd Schoellman. Thanks also to Jainyu Lu, who provided excellent research assistance. The usual disclaimers apply.

†Contact Info: 434 Flanner Hall, Notre Dame, IN 46556. *Email:* kdonova6@nd.edu

1 Introduction

Differences in agricultural labor productivity between the richest and poorest countries are twice as large as differences in aggregate labor productivity. In spite of this, the least developed countries in the world employ over eighty percent of their population in the agricultural sector. Since these countries employ such a large fraction of their population in a particularly unproductive sector, development accounting suggests that understanding agricultural productivity differences are crucial for understanding aggregate differences.[1]

One possible cause of agricultural productivity differences is that farmers in developing countries use fewer intermediate inputs. For example, as a share of harvest value, the value of intermediate inputs used on farms ranges from 4 percent in Uganda to 40 percent in the United States. Moreover, I document in Section 2 that this positive cross-country correlation does not exist in other sectors, suggesting that it may be an important margin for understanding why the agricultural sector exhibits significantly lower labor productivity than the nonagricultural sector in developing countries. The goal of this paper is to provide a theory to understand the cross-country correlation between the agricultural intermediate input share and per capita income, and in turn, quantitatively assess its role for cross-country productivity differences.

I argue that low intermediate input intensity is generated endogenously as a response to low total factor productivity (TFP). Because intermediate decisions are made before the realization of productivity shocks, the absence of insurance markets requires farmers to internalize the impact this choice will have on *ex post* consumption. In particular, purchasing more intermediate inputs leads to lower consumption in the event of a low shock realization. The extent to which this consideration impacts the *ex ante* intermediate choice depends critically on the income level of farmers. Low shock realizations are particularly disastrous for farmers in extremely poor countries, since consumption moves close to subsistence. These farmers are less willing to take on the risk associated with intermediate inputs usage, thus driving down labor productivity in developing countries.

I formalize and quantify this idea with a dynamic general equilibrium model in which both aggregate and sector-specific differences can potentially influence farmer response to

[1] This argument has been made in various forms starting with Restuccia et al. (2008), and also Caselli (2005), Vollrath (2009), and Gollin et al. (2014), among others.

shocks. Farmers produce agricultural output utilizing intermediate inputs, and are subject to incomplete markets and random fluctuations in farm productivity. In this sense, the model is similar to those used to focus on capital misallocation with self-insurance (e.g. Buera et al., 2011; Midrigan and Xu, 2014; Moll, 2014, among others). Here, however, deviations from unconstrained profit maximization are driven not by explicit input market frictions, but instead by the inability to insure *ex post* consumption. The timing of input choices implies that each shock realizations is weighted its risk neutral probability which includes a normalized measure of marginal utility. As TFP decreases in poor countries, income moves closer to subsistence and marginal utility at low shock realizations increases, so that farmers in poor countries put relatively more weight on bad potential outcomes. From the perspective of misallocation (e.g. Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009), this implies a wedge between the expected profit-maximizing marginal value and price of intermediate inputs, and is in fact isomorphic to a tax wedge in a model with no risk.

The result depends critically on the inclusion of a subsistence requirement when the equilibrium agricultural price varies across countries. I show that without subsistence, uninsurable shocks play no role in understanding differences in the agricultural input mix nor aggregate productivity across countries. That is, the agricultural productivity gap between any two economies is identical regardless of insurance against farm-level shocks. The addition of the subsistence requirement then has two important implications. First, I show theoretically that it generates a positive correlation between the intermediate share and aggregate income, consistent with the empirical evidence in Section 2. Second, identical distortions have differential impacts in rich and poor countries when combined with this risk channel. In this sense, the model provides an otherwise absent complementary amplification channel for distortions considered in a complete markets framework, including transportation costs (Adamopoulos, 2011; Gollin and Rogerson, 2014), distortions that vary farm size (Adamopoulos and Restuccia, 2014) or technology choice (Yang and Zhu, 2013), and more general distortions that affect input markets (Gollin et al., 2004; Restuccia et al., 2008).

I then turn to quantify the cross-country impact of this theory. I calibrate the model using a mix of aggregate and micro-level data from India, where the nominal intermediate

input share in agriculture is 11 percent. I jointly match consumption and production volatility to allow for some consumption smoothing in the model, and I also include relevant sector-neutral and agriculture-specific features to isolate the importance of each. I then vary exogenous sector-neutral productivity and the cost of intermediate inputs to U.S. levels. Since these differences exogenously increase labor productivity, I isolate the impact of the theory developed here by asking how much larger productivity differences are relative to an identical model in which shocks are perfectly insured.

The quantitative results imply that the seemingly sub-optimal intermediate input choices in agriculture can be partially explained as a rational response by farmers to risk, and through that distortion, affect aggregate labor productivity across countries. The calibrated model predicts that the Indian economy has an intermediate input share of 0.26, compared to the U.S. intermediate share of 0.40. This is 48 percent of the difference found in the data. This risk-driven distortion then amplifies cross-country productivity differences relative to a model with perfectly insured shocks. Agricultural productivity differences increase by 30 percent from a factor of 34.5 to 45.0, while GDP per capita differences increase from a factor of 6.4 to 7.8 for an amplification of 22 percent. Risk, therefore, plays an important role in understanding both agricultural input mix and productivity across countries.

This paper contributes to a recent macroeconomic literature on the role of agriculture in understanding cross-country income differences, including Gollin et al. (2004), Lagakos and Waugh (2013), Adamopoulos and Restuccia (2014), Herrendorf and Schoellman (2015), Tombe (2015), and Caunedo and Keller (2016). Most closely related is the work of Restuccia et al. (2008), who also focus on the role of intermediate inputs. Building off their work, this paper contributes a micro-founded rationale for distortions in the intermediate input market by focusing on the risk associated with intermediate input choices. Moreover, I show that the distortions emphasized in their work have a larger impact when combined with risk, and can help explain resource misallocation in agriculture. In this sense, the paper more broadly relates to the literature relating establishment-level distortions to aggregate productivity (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009). While recent work has focused on the role of financial development to explain pervasive deviations from undistorted profit maximizing behavior (e.g. Buera et al., 2011;

Midrigan and Xu, 2014; Moll, 2014), uninsurable risk has the potential to similarly distort input choices when combined with a subsistence requirement. The paper is therefore also similar to Angeletos (2007), who shows that the increase in aggregate savings relative to a complete markets benchmark is reversed when capital income is risky in Aiyagari (1994)-style models.

At the same time, a growing micro literature has highlighted the importance of agricultural risk in a variety of countries. Experimental evidence from the introduction of rainfall insurance shows that insurance can incentivize farmers to take more risk across a variety of countries (Mobarak and Rosenzweig, 2012; Cole et al., 2013b; Cai et al., 2014; Karlan et al., 2014), though the proper design of these contracts has been confounded by issues of trust and understanding among participants (Cole et al., 2013a). Other non-insurance work has found similar results, either through measuring downside risk by reconstructing state-specific consumption realizations from panel data (Dercon and Christiaensen, 2011) or through the introduction of seeds that limit downside risk (Emerick et al., 2015). I show that an appeal to risk aversion and incomplete markets is not sufficient for risk to matter in a cross-country context when the equilibrium price adjusts. Moreover, the quantitative results show that scaling such (properly designed) contracts have important amplification effects in equilibrium.

2 Motivating Evidence

Before turning to the model, this section documents facts about the intermediate share across countries. First, I confirm the results of Restuccia et al. (2008) and show that the same results extend to nominally priced intermediate shares in both aggregate and micro-level data. I then show that the same relationship does not hold in other sectors of the economy, suggesting that this relationship is a relevant margin for understanding the agricultural productivity gap.

Define the real intermediate input share as $X^{j*} := p_x^* X^j / p_a^* Y_a^j$, where X^j is intermediate input consumption in agriculture of country j, Y_a^j is agricultural output, and p_x^* and p_a^* are international prices of intermediates and agricultural output. The nominal intermediate share is given by $\widehat{X}^j := p_x^j X^j / p_a^j Y_a^j$, where the only difference is that intermediates

and output are valued at nominal country-specific prices p_x^j and p_a^j. As discussed in the introduction, the influential work of Restuccia et al. (2008) finds that real intermediate shares differ substantially across countries. Using data from Prasada Rao (1993), which is constructed from Food and Agricultural Organization (FAO) statistics and underlies the Restuccia et al. (2008) analysis, Figure 1a reproduces their finding of a strong positive correlation between GDP per capita and the real intermediate share. In addition to the correlation, the level differences are also large compared to other inputs. Merging this data with FAO estimates on capital stock in agriculture I find that the cross-country 90-10 ratio of intermediates per worker is more than double that of capital per worker in agriculture (Table 1). I then compute a simple variance decomposition in the style of Caselli (2005). Relative to capital per worker differences, variation in real intermediates per worker account for three times more of the variance in agricultural gross output per worker.[2]

Table 1: Agricultural Input Differences Across Countries (1985)

	90/10 Ratio	Variance decomposition ($\alpha_k = 0.30$, $\alpha_x = 0.50$)	Variance decomposition ($\alpha_k = 0.30$, $\alpha_x = 0.40$)
Output per worker	59.82	–	–
Capital per worker	82.78	0.11	0.11
Intermediate inputs per worker	196.38	0.38	0.26

Table notes: Data is from FAO and Prasada Rao (1993). The variance decomposition numbers are given by the following: $\frac{var(\log(k^{\alpha_k} x^{\alpha_x}))}{var(\log(y_a))}$ where each row shuts down the other input. For example, row two (capital) is measured by setting $\alpha_x = 0$, or $\frac{var(\log(k^{\alpha_k}))}{var(\log(y_a))}$.

Figure 1b uses the same data and plots the nominal share. Again, there is a strong positive correlation of 0.65. The tenth percentile country, as ranked by GDP per capita, has a nominal intermediate share that is one-fourth of the intermediate share in the United States.[3] The positive correlation with per capita GDP in both real and nominally priced shares implies that the price ratio p_x/p_a does not systematically vary with development,

[2]This is of course not to say that capital differences are irrelevant for agricultural productivity, only that intermediate inputs play an important role. The exercise does not include differences in capital quality, for example. See Caunedo and Keller (2016) for a quantitative evaluation of capital quality differences in agriculture.

[3]If rich countries are producing different crops than developing countries, one might suspect that the result is driven by different production techniques for these different types of output. While I cannot directly test this, I do group countries by latitude to control for the type of agricultural production, and compare within-group variation. The same correlation holds within groups.

though there is substantial variation in the price ratio, as Figure 1c shows.

2.1 Evidence from Micro Data

I next turn to evidence from micro data, using the Living Standard Measurement Studies (LSMS) released by the World Bank in cooperation with local governments. I include all countries with surveys after 2000, staying as close to 2010 as possible. There are 14 countries with sufficient data that, when combined with weights in the data, provide nationally representative samples in each country. I compute the nominal expenditure share of fertilizer and pesticides to corroborate aggregate statistics.[4] I use the median sale price to value harvest quantities in countries where the nominal expenditures are not directly available. Since the data is nationally representative when combined with available weights, aggregating gives the national expenditure shares. Figure 2 combines this data with Penn World Table GDP per capita, and confirms the positive relationship.

2.2 Comparison to Manufacturing and Services

As a last step, I turn back to aggregate data to compare intermediate input shares across sectors using the United Nations *System of National Accounts* (SNA). The U.N. data includes 87 countries in which data is sufficiently complete to construct nominal intermediate shares across the broadly defined sectors of agriculture, manufacturing, and services. These nominal intermediate shares are plotted in Figure 3, along with the nonagricultural sector measured as the total economy net of agriculture. Figure 3a confirms the relationship between the agricultural intermediate input share and per capita GDP. Figures 3c and 3d, however, show the nominal intermediate shares in manufacturing and services exhibit no such relationship. The figures also include the estimated coefficients from the simple linear regression of the sectoral intermediate share on log PPP GDP per capita. Only agriculture has a slope significantly different from zero, implying that the positive relationship between the intermediate share and per capita income is unique to the agricultural sector.

[4] Other intermediate inputs, such as fuel, are only available in some countries so I exclude them here. Also, note that these are not consumption shares. Fertilizer and pesticide consumption is only available in 6 of the countries. Since I focus on inorganic fertilizer and pesticide (i.e. not manure) nominal expenditures and consumption valued at market prices should be similar.

The rest of this paper is devoted to developing and quantifying a model to understand the cause of the correlation in agriculture and assess its impact on cross-country productivity differences.

3 Model

Time is discrete, and a model period is one year. There are two sectors, sector a for agriculture and sector m for manufacturing, which includes all nonagriculture. The manufacturing good is the numeraire, so its output price is normalized to $p_{mt} = 1$ for all t. Within an economy, decisions are made by a measure one of infinitely-lived households.

3.1 Technology

Manufacturing The manufacturing output good can be used as either consumption or intermediate inputs in agricultural production. Production is characterized by a stand-in firm which uses only labor services N_{mt} to produce output according to the constant returns to scale 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. The firm maximizes profits at each date t, so that N_{mt} is the solution to

$$\max_{N_{mt} \geq 0} \; AN_{mt} - w_t N_{mt} \tag{3.1}$$

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

Agriculture Each household is endowed with one farm that requires intermediate inputs x and labor n_a. Production occurs according to the decreasing returns to scale production function $y_{at} = z_t A x_t^{\psi} n_{at}^{\eta}$, where $\psi + \eta < 1$ and A is, again, sector neutral TFP. The shock z_t is a household-specific productivity shock drawn from a time-invariant distribution with cumulative distribution function $Q(z)$ and support on $[\underline{z}, \overline{z}]$.[5,6] The realization of z_t is i.i.d. with respect to both households and time. I assume the law of large numbers

[5] Throughout, it is assumed $\underline{z}$ is high enough to guarantee subsistence can be satisfied for all economies with TFP A in some set $\mathcal{A} \subset \mathbb{R}_+$. The results should be interpreted as holding for economies with TFP in the set $\mathcal{A}$.

[6] Note that increased intermediate intensity does not decrease the variance of shocks. This is supported by micro evidence in both developed and developing countries (Just and Pope, 1979; Traxler et al., 1995).

holds, so that the distribution of shocks across households is certain. Intermediate inputs are purchased from the manufacturing sector, at the price $p_x \geq 1$, which varies across countries. Note that the implicit assumption made is that there exists a technology to turn one unit of manufacturing output into $1/p_x$ units of intermediate input. This is a simple way to capture the fact that intermediate inputs are more expensive in developing countries.

3.2 Household Utility and Decisions

A household values consumption from both sectors a and m, and maximizes expected utility

$$\mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t u(c_{at}, c_{mt}) \right]$$

with discount factor $\beta \in (0, 1)$. The period t utility flow takes the form $u(c_{at}, c_{mt}) = \alpha \log(c_{at} - \bar{a}) + (1 - \alpha) \log(c_{mt})$, where c_{jt} is consumption from sector $j \in \{a, m\}$ and $\bar{a} > 0$ is the subsistence requirement of agricultural consumption. The utility function is consistent with the structural transformation paths when households consume gross output (Herrendorf et al., 2013).

Households do not have access to insurance markets, so that the shock can only be insured against through self-insurance. To this end, they save by storing agricultural output. This storage depreciates at a country-specific rate δ to capture differences in agricultural savings technologies across countries.

3.2.1 Decision Timing

At time $t - 1$, households save b_t units of the agricultural good. A fraction δ depreciates, and the household enters time t with $(1 - \delta)b_t$ units of savings. The period t decision problem is broken down into two stages denoted *ordering* and *production*, which are separated by the realization of the shock z.

In the ordering stage, each household chooses intermediates x_t to use in their farm. After ordering, z_t is realized. All production and consumption occurs in the production stage. First, a household chooses how to allocate labor between the agricultural sector, where they can work on the household farm, and in the manufacturing sector, where they

can work for wage w_t which is taxed at rate $\tau \geq 0$. Note that this allows labor to be used to smooth consumption across shock realizations. A fraction ν of the tax revenue is rebated back to households as a lump-sum transfer $T(b, z)$, while the fraction $1-\nu$ is used to purchase manufacturing output and discarded. After labor is decided, all production takes place. There is a centralized market for buying and selling goods, implying a unique equilibrium price p_a. Profits are made, all factors of production are paid, and consumption and savings choices $(c_{at}, c_{mt}, b_{t+1})$ take place.[7]

3.3 Recursive Problem

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

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

$$v^p(x, b, z) = \max_{c_a, c_m, n_a, b'} \quad \alpha \log(c_a - \bar{a}) + (1-\alpha)\log(c_m) + \beta v^o(b') \tag{3.2}$$

subject to constraint set

$$\begin{aligned} &p_a c_a + c_m + p_a b' = p_a z A x^{\psi} n_a^{\eta} - p_x x + (1-\tau) w (1-n_a) + p_a (1-\delta) b + T(b, z) \\ &b' \geq 0, \qquad c_a \geq \bar{a}, \qquad c_m \geq 0 \end{aligned}$$

where v^o is the value of entering the ordering stage at $t+1$ with b' units of savings in the stationary equilibrium. The production problem in (3.2) defines decision rules as a function of the intermediate choice. Working backwards, the ordering stage value of entering time t with b savings is

$$v^o(b) = \max_{x \geq 0} \quad \int_z v^p(x, b, z) dQ(z). \tag{3.3}$$

This defines the decision rule for intermediate inputs $x(b)$ and therefore the production

[7] I abstract from manufacturing risk here because manufacturing households are significantly richer. As I show in Section 6, the impact of risk is small at sufficiently high levels of savings.

stage decision rules $c_a(b,z)$, $c_m(b,z)$, $n_a(b,z)$, and $b'(b,z)$.

3.4 Stationary Equilibrium

The stationary competitive equilibrium of this economy is defined by an invariant distribution $\mu = \mu^*$, a value function v^o, 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^o solves the households'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) the state contingent transfer balances for all (b,z): $\nu\tau A(1-n_a(b,z)) = T(b,z)$, (4) the law of motion for μ, $\Lambda(\mu)$, implies $\Lambda(\mu^*) = \mu^*$, and μ^* is consistent with $Q(z)$ and decision rules, and (5) markets clear:

(a) Labor market:

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

(b) Agricultural goods market:

$$\int_b \int_z c_a(b,z) dQ(z) d\mu + \delta \int_b b d\mu = \int_b \int_z zAx(b)^{\psi} n_a(b,z)^{\eta} dQ(z) d\mu$$

(c) Manufacturing good market:

$$\int_b \int_z c_m(b,z) dQ(z) d\mu + p_x \int_b x(b) d\mu + (1-\nu)\tau A \int_b (1-n_a(b,z)) dQ(z) d\mu = AN_m$$

4 Characterization and Analytic Results

To clarify the mechanics of the model, this section characterizes the model analytically. Since the goal is to highlight the link between TFP and the intermediate input share, I assume throughout this section that $p_x = 1$ and $\tau = 0$ in all economies.

4.1 Subsistence Requirements and Risk Aversion

The utility function used here is standard in models of structural transformation when agents consume gross output (see Herrendorf et al., 2013), as the non-homotheticity generates the decline of agriculture during structural transformation. The main point in

this section is that this assumption also implies decreasing relative risk aversion, which I exploit to generate variation in intermediate shares across countries. To see this, first define C as the total expenditure on consumption at the production stage, given savings b, intermediate choice x, shock z, and the optimal savings decision rule b'

$$C(x,b,z) := p_a c_a(x,b,z) + c_m(x,b,z) = p_a z A x^{\psi} n_a^{\eta} - p_x x + w(1-n_a) + p_a(1-\delta)b - p_a b'.$$

With total expenditure C, a household purchases enough agricultural consumption 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 α. This implies $c_a(C) = \bar{a} + (\alpha/p_a)(C - p_a\bar{a})$ and $c_m(C) = (1-\alpha)(C - p_a\bar{a})$. Using these decision rules, the utility flow can be rewritten as a function of total consumption expenditures C,

$$\widetilde{u}(C) := u(c_a(C), c_m(C)) = \Omega - \alpha\log(p_a) + \log(C - p_a\bar{a}) \tag{4.1}$$

where $\Omega = \alpha\log(\alpha) + (1-\alpha)\log(1-\alpha)$. Because $\widetilde{u}$ is only a function of C, relative risk aversion, given $\bar{a}$ and price p_a, can be defined as

$$R(C;\bar{a},p_a) = \frac{C}{C - p_a\bar{a}}.$$

If $\bar{a} = 0$, this is the standard result of constant relative risk aversion with a log utility function. However if $\bar{a} > 0$, the utility function instead exhibits decreasing relative risk aversion. With the rewritten period utility flow, the indirect utility function at the ordering stage is

$$v^o(b) = \Omega - \alpha\log(p_a) + \max_{x\geq 0} \int_z \Big[\log\big(C(x,b,z) - p_a\bar{a}\big) + \beta v^o\big(b'(x,b,z)\big)\Big] dQ(z). \tag{4.2}$$

The key tension, as seen from (4.2), is that the optimal intermediate input choice balances the desire for both high income and low exposure to risk, and therefore allows the subsistence requirement to play an important role, which is the focus of Section 4.2.

4.2 Intermediate Shares and Productivity

I show in this section that TFP differences generate differences in intermediate shares by utilizing the interaction of uninsured shocks and subsistence. To make these results as sharp as possible, I consider the static version of the model (identically, $\delta = 1$ for all economies). All proofs are relegated to Appendix E.

I isolate the impact of uninsured shocks by comparing the model developed above (denoted by superscript IM for incomplete markets) with a complete markets version of the economy (denoted by superscript CM for complete markets). The complete markets economy is identical in every respect, except that households can trade a full set of state-contingent assets before the realization of z. How this affects the intermediate input choice can be seen by comparing the first order conditions with respect to x in the IM and CM economies. Farmers maximize expected profit with complete markets, so the ordering stage first order condition is

$$Ap_a^{1/(1-\eta)} F'(x) \int_Z z^{1/(1-\eta)} dQ(z) = 1 \tag{4.3}$$

where $F(x) = x^{\psi/(1-\eta)} \left(\eta^{\eta/(1-\eta)} - \eta^{1/(1-\eta)}\right)$ and $F'(\cdot)$ is the derivative with respect to x. Without the ability to trade these claims, the IM economy first order condition yields

$$Ap_a^{1/(1-\eta)} F'(x) \int_Z z^{1/(1-\eta)} \left(\frac{\widetilde{u}'(C(x,z))}{\mathbb{E}_z[\widetilde{u}'(C(x,z))]} \right) dQ(z) = 1 \tag{4.4}$$

where $\widetilde{u}'$ is the derivative of (4.1) with respect to consumption expenditures C. The two first order conditions are exactly the same except for the addition of marginal utility to the integrand of equation (4.4). While realizations in (4.3) are weighted by the probability of occurrence, realizations in (4.4) are weighted by risk neutral probabilities, which combines the probability of occurrence with deviations from average marginal utility. Households therefore internalize the impact their intermediate choice has on consumption in the absence of insurance markets. Low realizations receive more weight, so that Economy IM tilts the weight assigned by every household toward low shock realizations. Some algebra on (4.4) yields an implicit equation for the aggregate nominal intermediate share in the

economy,

$$\frac{X}{p_a Y_a} = \psi \left[\frac{\int_Z z^{1/(1-\eta)} \left(\frac{\widetilde{u}'(C(x,z))}{\mathbb{E}_z[\widetilde{u}'(C(x,z))]} \right) dQ(z)}{\int_Z z^{1/(1-\eta)} dQ(z)} \right] < \psi \tag{4.5}$$

where the last inequality follows from an application of Jensen's inequality. This gap between the realized and profit maximizing intermediate share is referred to by Morduch (1995) as "income smoothing," as households shift away from risky input choices to smooth consumption across states of the world. Moreover, it is the broad motivation behind the micro literature on rainfall insurance cited in the introduction. However, Proposition 1 shows that the appeal to risk aversion and incomplete markets alone is not sufficient to generate excess income smoothing in poor economies when the equilibrium price of agriculture varies across economies.

Proposition 1. *In the model with uninsurable shocks (Economy IM) and $\bar{a} = 0$, the following results hold in the competitive equilibrium:*

1. *The intermediate share $X/(p_a Y_a)$ is independent of TFP level A*
2. *$n_a(z)$ is independent of A*
3. *For two economies with TFP levels A^1 and A^2, the agricultural output per worker gap is identical to what it would be if both economies had complete markets. That is,*

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

Despite the fact that incomplete markets lower the intermediate share relative to its profit-maximizing level, it remains constant across countries when $\bar{a} = 0$. The result is driven by the equilibrium price of agriculture. As TFP decreases, the equilibrium price increases to incentivize households to take more risk. When $\bar{a} = 0$, these forces exactly offset, and generate risk neutral probabilities that are independent of TFP. As can be seen in equation (4.5), this implies nominal intermediate shares are constant across countries. This prediction has important implications for the division of labor across sectors and predicted labor productivity. The model counterfactually predicts that the employment share in agriculture is constant across countries, but also that the lack of insurance plays no role in understanding cross-country productivity differences. In the absence of the

subsistence requirement, predicted agricultural productivity differences between rich and poor countries are identical regardless of the degree of insurance. Proposition 2 shows that the inclusion of subsistence requirements breaks this result, and that the interaction of uninsured productivity shocks and subsistence requirements can qualitatively replicate the empirical correlation between the nominal intermediate share and TFP from Section 2.

Proposition 2. *In the competitive equilibrium, the intermediate share is increasing in A if and only if $\bar{a} > 0$.*

The same non-homotheticity that generates income effects in structural transformation (and therefore the negative relationship between agricultural labor and TFP in the model) also implies a positive relationship between the nominal intermediate share and aggregate income. Intuitively, the same forces are at work here as with $\bar{a} = 0$, with one important change. While the agricultural price still increases in response to lower TFP, it also now increases the cost of subsistence $p_a\bar{a}$ and thus increases relative risk aversion. This opposing force is not present when $\bar{a} = 0$, and therefore the agricultural price can no longer perfectly offset the drop in TFP. Thus, the model predicts a wedge between the marginal value and price of intermediates that increases as countries get poorer. Indeed, I show in Appendix A that the risk-based wedge is isomorphic to a reduced form tax wedge on intermediate inputs in the corresponding model that abstracts from risky production.

The rest of this paper attempts to quantify the importance of this channel for understanding cross-country productivity differences.

5 Calibration and Testing Model Predictions

Since the model predicts that uninsured shocks both depress intermediate intensity and act as a distortion in developing countries, I turn to quantifying the impact of this risk on cross-country productivity differences. I begin by calibrating the model to India using a combination of micro and aggregate statistics. I then test a number of untargeted moments using Indian micro data to confirm that the channel studied here is operational in the data. The cross-country results are in Section 6, where I compare the Indian model

economy to a U.S.-calibrated economy. This first requires U.S. specific productivity and agricultural distortions, which I discuss in Section 5.2.

I calibrate the model with the ICRISAT Village Level Studies (VLS2) data from India. The ICRISAT VLS2 is a household-level survey in India starting in 2001 and running through 2013, and is a continuation of earlier ICRISAT data from 1975-1984. The VLS2 includes information on household composition, consumption, farm inputs, and harvest values, and includes both a cross-sectional and panel component. Consumption data is not collected in 2007, so I focus on the 2001-2006 panel.

Since rainfall is a key component of risk in the developing world, I first compute the coefficient of variation of annual rainfall to see whether the ICRISAT villages are a reasonable approximation for India. I do so using detailed rainfall data for 13,688 cells in India from 1998–2005 from the Tropical Rainfall Measuring Mission (TRMM), which provides monthly rainfall estimates for the non-arctic world ($[-50, 50] \times [-180, 180]$ latitude-longitude) on 0.25×0.25 degree cells, for a total of 576,000 cells. The map of rainfall coefficient of variation is in Figure 4a, while the density across Indian cells is plotted in Figure 4b. The six ICRISAT villages are all very close to the mean, as 5 of 6 villages fall within 5 percent of the mean cell. The sixth, Shirapur, is somewhat higher at 25 percent above the mean cell. As a further test, Figure 4c confirms that the rainfall pattern is similar across months. It plots the monthly rainfall coefficient of variation averaged across ICRISAT villages and non-ICRISAT village cells. Again, the two sets of locations follow the same pattern.[8]

The model includes ten parameters. Six are common across the two economies, while four determine the distortions facing India that do not exist in the United States. They are discussed in turn.

5.1 Common Parameters

The six common parameters are the production exponents ψ and η, the shock distribution, the rebate on taxes ν, and preference parameters $\bar{a}$, α, and β. I set $\alpha = 0.005$ following Restuccia et al. (2008) and Lagakos and Waugh (2013) and $\beta = 0.96$. I set $\bar{a} = 0.03$, so

[8] In this section, I show only that the ICRISAT villages approximate the average Indian rainfall variation. A broader question is whether or differences in cross-country shock variances can account for differences in intermediate use and productivity. I return to this question in detail in Section 8 and find that they cannot.

that the Indian model economy has 50 percent of the population engaged in agriculture in the stationary equilibrium, consistent with sectoral employment in India (World Bank, 2015).

That leaves the production parameters, shock distribution, and tax rebate. The production parameters cannot be set to match nominal factor shares in the Indian economy, as the realized shares combine both the technological parameters and the distortions I seek to investigate. Instead, I assume the technologies are the same across U.S. and India and choose $\eta = 0.40$ and $\psi = 0.40$ to be consistent with U.S. estimates (Restuccia et al., 2008; Valentinyi and Herrendorf, 2008). Note however that the model predicts that the nominal labor share should equal $\eta = 0.40$ in both the U.S. and India. This is consistent with the ICRISAT data. ICRISAT includes both hired and household labor, valued at gender-specific market wages, which is the counterpart to the model definition. For each household (363 in 2006) I compute the value of labor services as a share of harvest value, and the average is 0.41. The same procedure for the nominal intermediate share implies an average value of 0.11, much lower than the U.S. level, and consistent with substantial distortions in intermediate use in India.

The last two parameters are the variance of the shock distribution and tax rebate share ν. I choose these parameters to match the average standard deviation of growth rates in household-level harvest values and total consumption expenditures. For this, I use the six year panel of households from 2001-2006. The panel consists of data on consumption, harvest, and household characteristics for 236 households across 6 villages. However, the data includes variation in harvest and consumption due to heterogeneity in household size, education, and village-level variation that are not modeled here. To the extent that these are predictable, directly using variance in the data would attribute them to unanticipated shocks. Instead I follow Kaboski and Townsend (2011) and others and purge the data of these factors with the regressions

$$\begin{aligned} \log(Y_{ivt}) &= \alpha^Y + \beta^Y X_{ivt} + \theta^Y_{vt} + \varepsilon^Y_{ivt} \\ \log(C_{ivt}) &= \alpha^C + \beta^C X_{ivt} + \theta^C_{vt} + \varepsilon^C_{ivt} \end{aligned}$$

where θ_{vt} is a village-time fixed effect, and X_{ivt} is a set of controls for household i at

time t that include number of men, women, and children, and age, gender, and education of the household head. Y and C are the household values of harvest and consumption, deflated by the Indian consumer price index. The R^2 on these regressions are 0.33 and 0.74 respectively, showing these features account for a large part of the variation in harvest and consumption. I then compute the sample average for the vector X_{ivt}, denoted $\overline{X}_{ivt}$, and compute the new data as

$$\widehat{\log(Y_{ivt})} = \widehat{\alpha}^Y + \widehat{\beta}^Y \overline{X}_{ivt} + \widehat{\varepsilon}^Y_{ivt} \tag{5.1}$$

$$\widehat{\log(C_{ivt})} = \widehat{\alpha}^C + \widehat{\beta}^C \overline{X}_{ivt} + \widehat{\varepsilon}^C_{ivt}. \tag{5.2}$$

My measure of the growth of harvest and consumption are then the differences $\Delta Y_{ivt} = \widehat{\log(Y_{ivt})} - \widehat{\log(Y_{ivt-1})}$ and $\Delta C_{ivt} = \widehat{\log(C_{ivt})} - \widehat{\log(C_{ivt-1})}$. The estimated harvest growth rates are more volatile than consumption, consistent with the ability of households to partially smooth income shocks. The average standard deviation of harvest growth across households is $\sigma_{\Delta Y} = 1.00$ compared to $\sigma_{\Delta C} = 0.49$ for consumption. Matching these two standard deviations imply a standard deviation of the shock distribution of $\sigma_z = 0.48$ and $\nu = 0.085$. That is, 8.5 percent of tax revenue is rebated to households. While the averages are targeted, Figures 5a and 5b plot the density of harvest and consumption growth rate volatility in both the model and the data to assess whether the distributions match with the data. I compute the implied volatility in the model by simulating 100,000 individuals for six years each, consistent with the ICRISAT panel length. Both implied distributions fit the data well despite not being targeted directly.[9]

5.2 Economy Specific Parameters

There are four dimensions along which India will differ from the U.S. economy: TFP A, the depreciation rate δ, the tax rate τ, and the intermediate input price p_x. I use TFP and p_x numbers from Restuccia et al. (2008). This implies $(A^{India}, p_x^{India}) = (0.22, 2.77)$ and normalized values $(A^{US}, p_x^{US}) = (1, 1)$.

[9] Alternatively, I could calibrate the shock distribution to the United States or some other rich country. However, because poor countries are much closer to subsistence, they respond more strongly to changes in the shock distribution. As I will show, rich countries respond little to changes in the variance of shocks, as they act quite close to profit maximizing households. Therefore, it is more important for the quantitative results to properly match the shock distribution of the poor country, while matching the variance of shocks in rich countries is of little quantitative importance. I therefore assume that they are the same to minimize differences across economies.

The last two parameters, τ and δ, control the relevance of the two smoothing channels in the model. First, τ is the ease by which households can smooth consumption by moving labor across sectors in response to shocks. I set $\tau^{US} = 0$, and calibrate τ^{India} from the ICRISAT data. ICRISAT includes household earnings and days worked in both agricultural and nonagricultural wages. Following a similar strategy to Section 5.1, I run the regressions

$$\log(w^j_{ivt}) = \alpha^j + \beta^j X_{ivt} + \theta^j_{vt} + \varepsilon^j_{ivt} \quad \text{for } j \in \{a, m\}$$

where w^j_{ivt} is the household wage earned in sector $j \in \{a, m\}$. This controls for differences in household composition, education, and age that are not included in the model. The implied wages are therefore

$$\widehat{\log(w^j_{ivt})} = \widehat{\alpha}^j + \widehat{\beta}^j \overline{X}_{ivt} + \widehat{\varepsilon}^j_{ivt}.$$

Taking averages of these implied wages across households gives $\overline{w}^a/\overline{w}^m = 0.45$, for an implied distortion of $\tau^{India} = 1 - 0.45 = 0.55$. The same procedure on the unmodified wages implies an implied distortion of 0.69, much closer to the value of 0.77 used in Restuccia et al. (2008).

The depreciation rate of savings δ determines the availability of smoothing through storing agricultural output. Udry (1995), Fafchamps et al. (1998), and Kazianga and Udry (2006) all point to the importance of crop storage for smoothing consumption in the developing world. There is also evidence that livestock is used to smooth consumption, particularly in India (e.g. Rosenzweig and Wolpin, 1993). Thus, I use the total value of livestock holdings and crop storage as the measure of savings, both of which are available in ICRISAT. In 2006, 54 percent of savings in the average ICRISAT household is livestock, while the other 46 percent comes from stored crops, suggesting both are important.[10] The total market value of livestock and stored crops is 96 percent of total harvest value in the average household, and I set $\delta^{India} = 0.15$ to match this fact. This is consistent with large costs to storage in developing countries. Despite targeting only the average, Figure

[10] ICRISAT also includes the pesticide and fertilizer storage. Consistent with the model used here, there is almost no storage of either. Including these intermediates in savings, only 0.6 percent of savings in the average household is made up of pesticide and fertilizer. Only 8 percent of households store any fertilizer or pesticide.

5c shows that the savings distribution in the model matches the data well. I set $\delta^{US} = 0$, but this has no impact on the results as the U.S. model economy is sufficiently rich such that the savings distribution, at any negative interest rate, is nearly degenerate at zero.

5.3 Savings, Consumption, and Intermediate Use: Model vs. Data

With the calibrated model in hand, I lastly assess model predictions for the relationship between consumption volatility, savings, and intermediate intensity by comparing the Indian model predictions to ICRISAT data. These are out-of-sample tests designed to assess whether the key model predictions are operational in the data. To compute the regressions in the model, I simulate 100,000 households in the Indian stationary equilibrium. I then compute standard errors using a bootstrap procedure to create 1000 samples of 205 households, consistent with the number of households in ICRISAT with the full six year panel and the requisite data.[11]

The results are broken into two sections. The first relates to the relationship between savings and intermediate use. The model predicts that intermediate expenditures and farm yield are positively related to savings, as households can better insure low shock realizations. The second covers the relationship between consumption volatility and intermediate use. Here, the model predicts a positive relationship between intermediate inputs and the coefficient of variation of consumption. I test both sets of predictions, and find that the same relationships hold in the data.

5.3.1 Savings and Intermediate Intensity

I consider the relationship between savings and intermediate use with the regressions

$$x_{2006} = \alpha + \beta b_{2005} + \varepsilon \tag{5.3}$$

$$\left(\frac{p_x x}{p_a y_a}\right)_{2006} = \alpha + \beta b_{2005} + \varepsilon \tag{5.4}$$

$$yield_{2006} = \alpha + \beta b_{2005} + \varepsilon \tag{5.5}$$

where savings is lagged to prevent endogeneity concerns. In all regressions, both dependent and independent variables are normalized by their respective means. Results are

[11] 31 households are missing data on fertilizer use, hence the decrease in sample size from the 236 used in Section 5.1. Redoing the analysis exclusively with this smaller sample does not change the results.

presented in Table 2.

Table 2: Savings and Intermediate Inputs (Model vs. Data)

	Model	Data	Model	Data	Model	Data
Savings	0.368***	0.390***	0.088*	0.008	0.258***	0.106***
	(0.039)	(0.061)	(0.077)	(0.047)	(0.088)	(0.038)
Constant	0.632***	0.610***	0.855***	0.992***	0.741***	0.892***
	(0.045)	(0.112)	(0.095)	(0.088)	(0.094)	(0.071)
Obs	n.a.	205	n.a.	205	n.a.	205
R^2	0.532	0.168	0.020	0.000	0.057	0.020
Dependent variable	Expenditures	Expenditures	Nominal share	Nominal share	Yield	Yield

Table notes: Significance at 0.01, 0.05, 0.1 levels denoted by ***, **, and *. Yield is measured as total harvest value per acre of land. The model standard errors are bootstrapped using 1000 samples of 205 individuals. Dependent and independent variables are normalized by sample mean.

First, regression (5.3) shows that the model and data match well when considering intermediate expenditures, where $\widehat{\beta}^{data} = 0.37$ and $\widehat{\beta}^{model} = 0.39$. Both are significant at one percent. In contrast, the realized intermediate shares are partially driven by unexpected volatility in output. Correspondingly, the R^2 in model regression (5.4) decreases from 0.53 to 0.02 in the model and from 0.17 to 0.00 in the data. Moreover, the estimate of $\widehat{\beta}^{data} = 0.01$ is small and statistically insignificant. The model prediction is somewhat stronger, with $\widehat{\beta}^{model} = 0.09$. This estimate is significant at ten percent, but no stricter, despite the strong relationship between intermediate expenditures and savings.

Lastly, regression (5.5) considers the relationship between realized farm yield, measured as harvest value per acre, and lagged savings. Note that yields – like intermediate shares – are in part driven by random shocks to farm productivity. Indeed, like the intermediate share regression (5.4), both the model and data have low R^2. However, unlike the intermediate share regression, the relationship between yield and savings is significant at one percent in both model ($\widehat{\beta}^{model} = 0.26$) and data ($\widehat{\beta}^{data} = 0.11$). The model can rationalize this result quite simply. First, note that the realized nominal intermediate share and farm yield in the model are

$$\begin{aligned}
\frac{p_x x}{p_a y_a(x,z)} &= p_x z^{\frac{-1}{1-\eta}} x^{\frac{1-\eta-\psi}{1-\eta}} (p_a A)^{\frac{-1}{1-\eta}} \left(\frac{\eta}{w}\right)^{\frac{-\eta}{1-\eta}} \\
p_a y_a(x,z) &= z^{\frac{1}{1-\eta}} x^{\frac{\psi}{1-\eta}} (p_a A)^{\frac{1}{1-\eta}} \left(\frac{\eta}{w}\right)^{\frac{\eta}{1-\eta}} .
\end{aligned}$$

The critical difference is the exponent on x. Since the calibrated model implies $(1-\eta-\psi)/(1-\eta) = 0.33$ and $\psi/(1-\eta) = 0.67$, variation in intermediate expenditures is a larger component of variation in yield relative to the intermediate share. Since regression (5.3) shows that variation in x is tightly related to variation in b, this generates the increased importance of savings for yield relative to the intermediate share.

5.3.2 Consumption Volatility and Intermediate Intensity

I next turn to the relationship between consumption volatility and intermediate use. In both the model and the data, I consider the relationship between the coefficient of variation of consumption and average intermediate use during the six year panel. That is, denoting $\widehat{x}_{it}$ as the realized nominal intermediate share of household i at time t, I run the regression

$$\frac{\sum_{t=2001}^{t=2006} \widehat{x}_{it}}{6} = \alpha + \beta \left(\frac{\sigma^C}{\mu^C}\right)_i + \varepsilon_i, \tag{5.6}$$

The preferred comparison is the consumption estimates from regression (5.2), in which $C_{it} = \exp(\widehat{\log(C_{ivt})})$, for the reasons discussed in Section 5.1. For completeness, I also report results the household consumption directly from ICRISAT (deflated by the Indian CPI). The results are in Table 3. The model is consistent with the positive relationship between the consumption coefficient of variation and the average intermediate share found in the data, in which the model predicts $\widehat{\beta}^{model} = 0.09$, compared to $\widehat{\beta}^{est-data} = 0.15$ using the estimated consumption data.

The key intuition for this result is the positive correlation between the standard deviation and mean of consumption across households, in which the model also matches the data. The model predicts a correlation of 0.88, compared to 0.85 using the consumption data estimated from (5.2) and 0.90 using the direct ICRISAT consumption data. Given the strong positive correlation in the model, the results from regression (5.6) imply that an increase in mean consumption must be met with a less than one-for-one increase in the standard deviation of consumption. That is, the regression

$$\mu_i^C = \alpha + \beta\sigma_i^C + \varepsilon_i \tag{5.7}$$

Table 3: Consumption Volatility and Intermediate Inputs (Model vs. Data)

	Model	Data (estimated)	Data (direct)
c.v. of consumption	0.088*	0.154**	0.231*
	(0.049)	(0.063)	(0.119)
Constant	0.912***	0.758***	0.769***
	(0.051)	(0.117)	(0.135)
Obs	205	205	205
R^2	0.020	0.029	0.018
Corr(σ_C, μ_C)	0.884***	0.846***	0.898***

Table notes: Significance at 0.01, 0.05, 0.1 levels denoted by ***, **, and *. The model standard errors are bootstrapped using 1000 samples of 205 individuals. Column 2 computes consumption as the resulting estimates from regression (5.2). Column 3 computes household consumption directly from ICRISAT data, deflating by the Indian CPI. Dependent and independent variables are normalized by sample mean.

implies an estimate of $\widehat{\beta}^{model} \in (0,1)$ in the model. Table 4 shows the result of this regression in the model and compares it to the data. The model matches the estimated data well with $\widehat{\beta}^{model} = 0.22$ and $\widehat{\beta}^{est-data} = 0.28$, though the unmodified consumption data predicts a higher estimate $\widehat{\beta}^{data} = 0.66$.

Table 4: Mean and Standard Deviation of Consumption (Model vs. Data)

	Model	Data (estimated)	Data (direct)
σ^C	0.223***	0.278***	0.662***
	(0.018)	(0.012)	(0.023)
Constant	0.778***	0.722***	0.338***
	(0.015)	(0.019)	(0.029)
Obs	205	205	205
R^2	0.781	0.715	0.805

Table notes: Significance at 0.01, 0.05, 0.1 levels denoted by ***, **, and *. The model standard errors are bootstrapped using 1000 samples of 205 individuals. Column 2 computes consumption as the resulting estimates from regression (5.2). Column 3 computes household consumption directly from ICRISAT data, deflating by the Indian CPI. Dependent and independent variables are normalized by sample mean.

The model predictions are consistent with the empirical household-level relationship between intermediate input usage, savings, and consumption volatility. In the next section I therefore proceed to assess the quantitative cross-country implications of the model tested here.

6 Quantitative Cross-Country Results

To compare the Indian model economy to the U.S., I vary the calibrated model in the four ways discussed in the previous section: agriculture-specific distortions (p_x, τ, δ) and sector-neutral TFP A. I isolate the effect of risk by comparing the model predictions against an identical model with complete insurance against shocks.

The main cross-country results on productivity and input mix are in Table 5. First, note that the U.S. model economy is sufficiently rich that the move from complete to incomplete markets has little impact. The nominal intermediate share and employment share remain nearly identical.

This is not the case in India. The elimination of complete markets causes the nominal intermediate share in India to drop from 0.40 to 0.26. The model therefore captures $(0.40 - 0.26)/(0.40 - 0.11) = 48$ percent of the nominal intermediate share difference between the U.S. and India.[12] This causes agriculture to become more labor intensive in India, as the employment share in agriculture increases from 34 to 50 percent, an increase of 47 percent.

Table 5: Quantitative Cross-Country Results

	Labor Productivity Gap		$p_x X / p_a Y_a$		N_a *(%)*	
Economy	Agriculture	Aggregate	Rich	Poor	Rich	Poor
Data: U.S./India	77.4	10.6	0.40	0.11	1.6	50.0
Data: 90/10 Ratio	63.7	23.1	0.40	0.09	2.0	82.0
Model with						
Incomplete markets	45.0	7.8	0.40	0.26	1.2	50.0
Complete markets	34.5	6.4	0.40	0.40	1.2	34.1

Table notes: The second row is the differences between the ninetieth and tenth percentile countries, while the first directly compares the U.S. and India. Results are presented for U.S. and Indian model economies with complete and incomplete markets.

The increased labor intensity in Indian agriculture impacts productivity in both agriculture (measured as the U.S.–India ratio of agricultural gross output per worker) and the aggregate (the ratio U.S. priced GDP per capita). The agricultural productivity gap

[12] An alternative exercise is to lower exogenous agricultural productivity until the predicted real intermediate share differences matched the data. Targeting this moment directly implies a productivity too low to satisfy subsistence among the households who receive the lowest shock realization while holding no savings.

between the two countries increases by 30 percent from 34.5 to 45.0, while the aggregate gap increases by 22 percent from 6.4 to 7.8. Thus, the introduction of risk provides a substantial amplification of cross-country productivity differences at both the agricultural and aggregate level.

6.1 Misallocation of Intermediate Inputs

Underlying the aggregate results is the unwillingness of households to take on the consumption risk associated with intermediate inputs. This constraint is most severe among the poor, as they lack the savings to overcome a negative shock. On the other hand, as households get richer, they act more similar to households in rich countries. To see this more clearly, first define the function $\phi^{risk}(b) := \psi/\mathbb{E}_z[\widehat{x}(b)]$. This function is the ratio of the expected profit maximizing intermediate share – equal to ψ for all levels of savings due to the Cobb-Douglas production function – and the expected intermediate share predicted by the model, denoted $\mathbb{E}_z[\widehat{x}(b)]$. As I show in Appendix A, $\phi^{risk}(b)$ measures the extent of intermediate input misallocation in the model.[13]

Figure 6a plots this function in both the U.S. and Indian model economies. In both countries, the tax is decreasing in savings, as richer individuals are better able to self-insure against farm shocks. Thus, the model operates by allowing risk to distort the efficient distribution of intermediate inputs across production units, in the style of Restuccia and Rogerson (2008) or Hsieh and Klenow (2009). While this result holds in both the U.S. and Indian model economies, it is clearly more dramatic in the Indian economy, which drives the amplification effect of risk beyond the implied differences in the complete markets model. This result also plays an important role when considering the impact of changing distortions, and I return to this in Section 7.

6.2 Decomposing the Direct and Indirect Effects of Risk

Next, I turn to decomposing the aggregate effects, which include a direct effect and an indirect general equilibrium effect. This former is the effect of risk on intermediate intensity, whereas the latter is the response of sectoral employment, driven by variation of

[13] More specifically, in Appendix A, I develop a model in which intermediate inputs are chosen after the shock realization. I show that if intermediates are taxed at rate $\phi^{risk}(b)$, this new model generates the same stationary equilibrium as the model developed in the text. Thus, $\phi^{risk}(b)$ measures the extent of misallocation in intermediate inputs generated by risk.

the agricultural price. I decompose these effects in Table 6 with a series of counterfactual Indian economies.

To assess how risk directly dampens intermediate intensity, I ask the following: in a complete markets economy, how expensive would intermediate inputs have to be to match the real intermediate share predicted by the baseline Indian model? This experiment isolates the effect of risk on the economy through intermediate intensity, but maintains the complete markets structure. Therefore, it provides the direct effect of risk through intermediate intensity relative to a complete markets model such as Restuccia et al. (2008).

To compute this, I increase p_x until the predicted real intermediate share is identical to that implied by the baseline incomplete markets Indian economy. I refer to this as the "modified complete markets economy." The experiment generates an economy with a 34 percent higher intermediate price than the baseline complete markets model, as the intermediate price increases from $p_x = 2.77$ to $p_x = 3.70$. That is, the baseline incomplete markets Indian economy has the same real intermediate share as a complete markets economy with a 34 percent higher intermediate price. The aggregate changes implied by this modified complete markets economy are in row one of Table 6. Holding the market structure fixed, lowering the real intermediate share to the level achieved by the baseline incomplete markets model implies a 15 percent increase in agricultural employment and a 14 percent decrease in agricultural productivity.

If the model developed in this paper operated exclusively by decreasing the real intermediate share, its aggregate moments should look similar to those of the modified complete markets model. As one can see by comparing the two rows of Table 6, this is not the case. Instead, it predicts a much larger change in agricultural employment, despite the fact that the two models are constructed to have identical real intermediate shares. The incomplete markets model predicts a 46 percent increase in agricultural employment, compared to only a 15 percent increase in the modified complete markets model. This difference is the indirect general equilibrium effect, as the agricultural price responds more strongly than in the modified complete markets model. Due to decreasing returns to labor, this further decreases agricultural productivity in excess of the direct effect. Agricultural productivity decreases by 24 percent instead of 14 percent.

Taking stock of these changes shows an important role for both the direct and indirect

effects. When one considers to total change in agricultural productivity (a 24 percent drop), these results imply that the direct effect accounts for 58 percent of that change (-0.14/-0.24 = 58 percent). The other 42 percent comes from the fact that generating an identical drop in the real intermediate share requires a higher equilibrium price in the incomplete markets model. This, in turn, implies a larger shift of employment toward agriculture.

Table 6: Decomposition of aggregate effects

%Δ relative to baseline complete markets Indian economy	Exogenously Varied	Equilibrium Response			
	p_x	X/Y_a	N_a	Y_a/N_a	p_a
Modified Complete Markets (equalized real intermediate share)	0.34	−0.14	0.15	−0.14	0.16
Incomplete Markets (baseline calibration)	0.00	−0.14	0.46	−0.24	0.29

Table notes: This table decomposes the direct versus indirect effect of a decrease in p_x in the Indian economy. Both rows are different Indian model economies. All entries are percentage changes in the given moments relative to the baseline complete markets Indian economy. The first row exogenously increases p_x until the real intermediate share (X/Y_a) is equalized in both economies but maintains the complete markets assumption. The second row is the baseline incomplete markets Indian economy.

How reasonable is this general equilibrium effect? The model predicts $(p_x^{India}/p_a^{India})/(p_x^{USA}/p_a^{USA}) = 0.66$. The statistics in Restuccia et al. (2008) derived from Prasada Rao (1993) imply a value of 0.55, so the model actually somewhat under-predicts the relative price gap. Looking back on Figure 1c, which plots the relationship between the price ratio p_x/p_a relative to the U.S., the relatively strict 90 percent confidence interval around the regression line includes the value 0.66 for India. The amplification predicted by the model is therefore well within reason of variation in the data.[14]

7 Heterogenous Impact of Agricultural Distortions

As can be seen in Figure 6a, the model predicts substantial variation in the implied household-level distortion. I next turn to assessing how variation in agricultural distortions change this underlying household-level misallocation. I do so by varying p_x from $p_x = 2.77$ to $p_x = 1$ in both the Indian and U.S. model economies and compute the new

[14]This general equilibrium effect also implies that the aggregate increase in income from a large-scale insurance program is larger than would be expected from a partial equilibrium result. This is in contrast to Buera et al. (2014), who find that the general equilibrium effects of microfinance dampen the partial equilibrium effect on income.

stationary equilibrium.[15] This price plays an important role in a number of recent theories of agricultural productivity, including those that rely on agriculture-specific input market distortions (Gollin et al., 2004; Restuccia et al., 2008), internal trade and transportation costs (Adamopoulos, 2011; Gollin and Rogerson, 2014), and those that link intermediate prices to technology choice (Yang and Zhu, 2013).

The main results are in Table 7.[16] The results show that the impact on productivity depends critically on the level of overall initial productivity in the economy, which is in contrast to theories such as Restuccia and Rogerson (2008) and Hsieh and Klenow (2009). Agricultural productivity increases by 98 percent in India compared to 65 percent in the United States (Table 7). This is driven by differential responses to the underlying household distortion. Figure 6b shows that this distortion decreases by as much as 60 percent among the poorest households in India compared to negligible changes in the U.S..

Table 7: Agricultural productivity as p_x changes

	$A = 0.22$	$A = 1$
$p_x = 2.77$	0.023	0.607
$p_x = 1$	0.044	1.000
% increase	0.913	0.647

Table notes: An entry in this table is agricultural productivity for the economy with the given level of A and p_x. The entry in the lower right ($A = 1$ and $p_x = 1$) is the baseline U.S. economy and is normalized to one.

Table 8 computes the percentage changes in different agricultural moments when the intermediate price decreases from the Indian level $p_x = 2.77$ to the U.S. level $p_x = 1$. First, changes in the real intermediate share play only a small role. In both countries, the increase in p_x causes the real intermediate share to increase, but the change is similar across countries (69 percent increase in the U.S. and 75 percent increase in India). The change in sectoral labor, however, is more stark. While agricultural employment drops by 34 percent in the U.S., it drops by 52 percent in India. Therefore, as in Section 6, these results imply an important role for general equilibrium effects in generating productivity changes.

The results therefore show that changing distortions is not independent of sector-

[15] In the Appendix, I also show the dynamics of the response to a surprise decrease in p_x in the Indian model economy.
[16] I provide identical results for changes in τ and δ in Appendix B.4

Table 8: %Δ in agricultural moments when p_x decreases from 2.77 to 1

	$\frac{p_a X}{p_a Y_a}$	X/Y_a	N_a	p_a
$A = 1$	0.01	0.69	-0.35	-0.48
$A = 0.22$	0.22	0.75	-0.52	-0.39

Table notes: This table provides percentage changes in computed moments for both the rich ($A = 1$) and poor ($A = 0.22$) economies when the intermediate price p_x decreases from the calibrated Indian level $p_x = 2.77$ to the calibrated U.S. level $p_x = 1$.

neutral productivity, once the model includes both risk and subsistence consumption. Moreover, they provide an otherwise absent complimentary channel for other theories considered in the literature.

8 Robustness

8.1 Different Shocks or Different Responses?

The theory developed in this paper assumes that households in different countries face the same shocks but respond differently. Alternatively, one might imagine that poor countries simply face much more uncertainty in their shocks which could potentially generate the results. I investigate this using the Global Agro-Ecological Zones (GAEZ) data produced by the FAO and International Institute for Applied Systems Analysis. While the analysis is detailed completely in Appendix C, I briefly outline the procedure and results here. The GAEZ data is spatial grid data on historical rainfall from 1980 to 2000 at the 5 arc minute resolution, and includes approximately 9 million cells around the world, and also includes the internationally priced value of harvest in each cell for the year 2000. This allows me to compute a harvest-weighted country-level rainfall time series for 147 countries, in which rainfall variation in areas where farming occurs is weighted more heavily.

I find no evidence that cross-country variation in this series is related to GDP per capita or agricultural productivity. A one standard deviation increase in the rainfall coefficient of variation is associated with a one percent decrease in agricultural productivity, not nearly large enough to matter quantitatively. The result echoes recent work by Adamopoulos and Restuccia (2015) who show that natural disadvantages (soil and land quality, for example) are not responsible for low agricultural productivity in poor

countries. Instead, these results suggest that the differential response to risk across countries is key for understanding the relationship between risk and intermediate inputs, not variation in the exogenous shocks themselves.

8.2 Shock Variance and Insurability

In Appendix B, I provide a number of additional robustness checks. To provide some context for the baseline results and robustness to the calibration procedure in Section 5, I vary both the variance of the agricultural shock (Appendix B.1) and the fraction of the shocks that are insurable (Appendix B.2).

Increasing the shock variance actually decreases the agricultural productivity gap, thought the quantitative effect is quite small. The result is due to two roughly offsetting effects. Larger shock variance subjects households to more risk, which pushes productivity down. The opposing effect is due to the low utility weight on agricultural consumption. Since the utility weight on agricultural consumption is low, total demand is inelastic and approximately equal to $\overline{a}$. As the variance increases, it becomes easier for a small set of households to "luck" into high production and thus satisfy economy-wide demand, even in the absence of sufficient input investment. Thus, the equilibrium price remains low, as households need not be incentivized to produce agricultural goods. This low price keeps employment out of agriculture, which in turn increases labor productivity. The two forces work in opposite directions, and generate only a small total change in productivity as the variance changes.

Increasing the share of insured shocks naturally decreases the quantitative effect of risk. However, even when a large portion of the shocks are insurable, there is still a substantial risk-driven distortion. The result highlights the importance of downside risk among the poor. Since the households internalize the effect of risk on consumption, even small shocks are given relatively large weight by the poor, and can have important quantitative effects.

9 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. The model provides a risk-based foundation for misallocation across agricultural production units, but the ability to generate relatively larger distortions in poor economies depends critically on the inclusion of a subsistence requirement. Quantitatively, the model captures half of the difference in intermediate input shares between the richest and poorest countries. This has important quantitative implications for productivity across countries. Relative to an identical model with complete insurance, the distortionary impact of risk amplifies agricultural productivity differences by about thirty percent and aggregate productivity differences by twenty percent.

In addition, the model also predicts that the impact of agricultural distortions depends critically on the overall level of productivity in the economy. Counterfactual experiments show that lowering these distortions can substantially decrease the household-level distortion that risk creates, but has a much larger effect in poor countries. This result is particularly important in light of the fact that the model captures about half of the difference in nominal intermediate shares, leaving room for complimentary explanations. Yang and Zhu (2013), for example, highlight technological choice in agriculture in which farmers can choose to use a technology with no intermediate inputs. A more detailed analysis of the link between such theories will hopefully provide a more complete picture of agricultural input choices and productivity.

References

T. Adamopoulos. Transportation Costs, Agricultural Productivity, and Cross-Country Income Differences. *International Economic Review*, 52(2):489–521, 2011.

T. Adamopoulos and D. Restuccia. The Size Distribution of Farms and International Productivity Differences. *American Economic Review*, 104(6):1667–1697, 2014.

T. Adamopoulos and D. Restuccia. Geography and Agricultural Productivity: Cross-

Country Evidence from Micro Plot-Level Data, July 2015. University of Toronto Working Paper.

S. R. Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics*, 109(3):659–684, 1994.

G. Angeletos. Uninsured idiosyncratic investment risk and aggregate saving. *Review of Economic Dynamics*, 10(1):1–30, 2007.

F. Buera, J. Kaboski, and Y. Shin. The Macroeconomics of Microfinance, March 2014. Working Paper.

F. J. Buera, J. P. Kaboski, and Y. Shin. Finance and Development: A Tale of Two Sectors. *American Economic Review*, 101(5):1964–2002, 2011.

H. Cai, Y. Chen, H. Fang, and L.-A. Zhou. The Effect of Microinsurance on Economic Activities: Evidence from a Randomized Field Experiment. *Review of Economics and Statistics*, 2014. forthcoming.

F. Caselli. Accounting for Cross-Country Income Differences. In P. Aghion and S. Durlauf, editors, *Handbook of Economic Growth*, pages 679–741. 2005.

J. Caunedo and E. Keller. Capital Obsolescence and Agricultural Productivity, July 2016. Working Paper.

S. Cole, X. Giné, J. Tobacman, R. Townsend, P. Topalova, and J. Vickery. Barriers to Household Risk Management: Evidence from India. *American Economic Journal: Applied Economics*, 5(1):104–135, 2013a.

S. Cole, X. Giné, and J. Vickery. How Does Risk Management Influence Production Decisions? Evidence from a Field Experiment, July 2013b. World Bank Working Paper 6546.

S. Dercon and L. Christiaensen. Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia. *Journal of Development Economics*, 96(2):159–173, 2011.

K. Emerick, A. de Janvry, E. Sadolet, and M. H. Dar. Technological innovations, downside risk, and the modernization of agriculture. 2015. Working Paper.

M. Fafchamps, C. Udry, and K. Czukas. Drought and Saving in West Africa: Are Livestock a Buffer Stock? *Journal of Development Economics*, 55(2):273–305, 1998.

Goddard Earth Sciences Data and Information Services Center, NASA. *Tropical Rainfall Measuring Mission (TRMM) 3B43*, 2015.

D. Gollin and R. Rogerson. Productivity, Transport Costs, and Subsistence Agriculture. *Journal of Development Economics*, 107:38–48, 2014.

D. Gollin, S. L. Parente, and R. Rogerson. Farm work, home work and international productivity differences. *Review of Economic Dynamics*, 7(4):827–850, 2004.

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

B. Herrendorf and T. Schoellman. Why is Measured Productivity so Low in Agriculture? *Review of Economic Dynamics*, 2015. forthcoming.

B. Herrendorf, R. Rogerson, and A. Valentinyi. Two Perspectives on Preferences and Structural Transformation. *American Economic Review*, 103(7):2752–2789, 2013.

C. Hsieh and P. J. Klenow. Misallocation and manufacturing tfp in china and india. *Quarterly Journal of Economics*, 124(4):1403–1448, 2009.

R. E. Just and R. D. Pope. Production Function Estimation and Related Risk Considerations. *American Journal of Agricultural Economics*, 61(2):276–284, 1979.

J. P. Kaboski and R. M. Townsend. A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative. *Econometrica*, 79(5):1357–1406, 2011.

D. Karlan, R. D. Osei, I. Osei-Akoto, and C. Udry. Agricultural Decisions After Relaxing Credit and Risk Constraints. *Quarterly Journal of Economics*, 129(2):597–652, 2014.

H. Kazianga and C. Udry. Consumption Smoothing? Livestock, Insurance and Drought in Rural Burkina Faso. *Journal of Development Economics*, 79(2):413–446, 2006.

D. Lagakos and M. E. Waugh. Selection, Agriculture, and Cross-Country Productivity Differences. *American Economic Review*, 103(2):948–980, 2013.

V. Midrigan and D. Xu. Finance and Misallocation: Evidence from Plant-Level Data. *American Economic Review*, 104(2):422–458, 2014.

A. M. Mobarak and M. Rosenzweig. Selling Formal Insurance to the Informally Insured, February 2012. Yale Economics Growth Center Discussion Paper No. 1007.

B. Moll. Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation? *American Economic Review*, 104(10):3186–3221, 2014.

J. Morduch. Income Smoothing and Consumption Smoothing. *Journal of Economic Perspectives*, 9(3):103–114, 1995.

D. Prasada Rao. Intercountry comparisons of agricultural output and productivity, 1993. FAO Economic and Social Development Paper No. 112.

D. Restuccia and R. Rogerson. Policy Distortions and Aggregate Productivity with Heterogeneous Establishments. *Review of Economic Dynamics*, 11(4):707–720, 2008.

D. Restuccia, D. T. Yang, and X. Zhu. Agriculture and aggregate productivity: A quantitative cross-country analysis. *Journal of Monetary Economics*, 55(2):234–250, 2008.

M. R. Rosenzweig and K. I. Wolpin. Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investments in Bullocks in India. *Journal of Political Economy*, 101(2):223–244, 1993.

T. Tombe. The Missing Food Problem: Trade, Agriculture, and International Productivity Differences. *American Economic Journal: Macroeconomics*, 7(3):226–258, 2015.

G. Traxler, J. Falck-Zepeda, J. Ortiz-Monasterio R., and K. Sayre. Production risk and the evolution of varietal technology. *American Journal of Agricultural Economics*, 77 (1):1–7, 1995.

C. Udry. Risk and Saving in Northern Nigeria. *American Economic Review*, 85(5):1287–1300, 1995.

United Nations. *National Account Statistics: Main Aggregates and Detailed Tables*. Table 2.3, 2015. URL http://data.un.org/Data.aspx?q=Table+2.3&d=SNA&f=group_code%3a203.

A. Valentinyi and B. Herrendorf. Measuring Factor Income Shares at the Sectoral Level. *Review of Economic Dynamics*, 11(4):820–835, 2008.

D. Vollrath. How important are dual economy effects for aggregate productivity? *Journal of Development Economics*, 88(2):325–334, 2009.

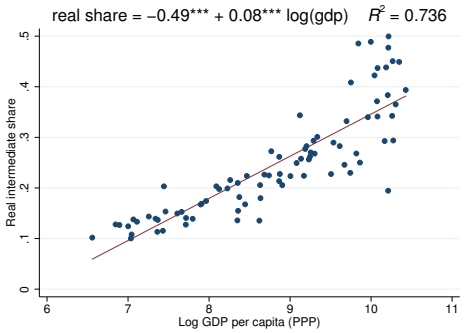
World Bank. *World Development Indicators*. The World Bank, October 2015. URL http://databank.worldbank.org/.

D. Yang and X. Zhu. Modernization of Agriculture and Long-Term Growth. *Journal of Monetary Economics*, 60(3):367–382, 2013.

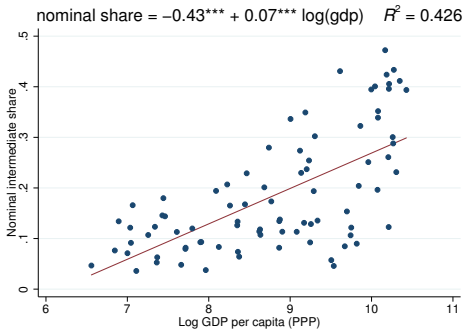
10 Figures

Figure 1: Cross-Country Intermediate Shares (1985)

(a) Real intermediate share



(b) Nominal intermediate share



(c) Nominal price ratio

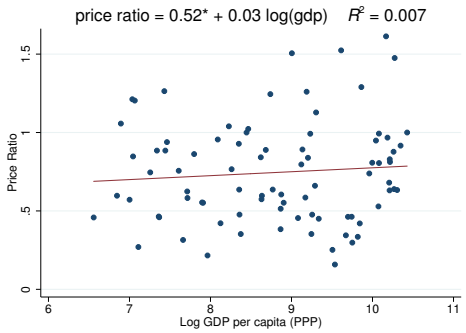
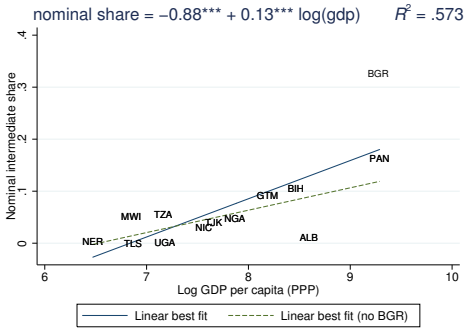


Figure notes: This figure plots the real and nominal intermediate shares, along with the relative price ratio of intermediates to agricultural output. They are derived from Prasada Rao (1993). Significance of coefficient estimates in the linear regression of each moment on log GDP per capita are denoted by ***, **, and *.

Figure 2: Nominal expenditure shares from LSMS micro data

(a) Fertilizer



(b) Fertilizer and Pesticide

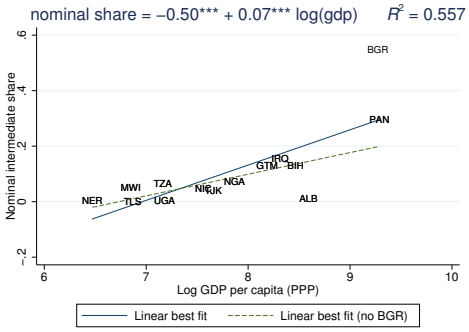


Figure notes: Aggregated intermediate shares using (1) fertilizer and (2) fertilizer and pesticide computed from World Bank LSMS micro data. Also included are the regressions equations derived from regressing these aggregate shares on log GDP per capita. Significance of coefficient estimates at 0.01, 0.05, 0.1 levels denoted by ***, **, and *. Excluding Bulgaria (BGR) still maintains the positive relationship at the one percent level.

Figure 3: Sectoral nominal intermediate shares (2005)

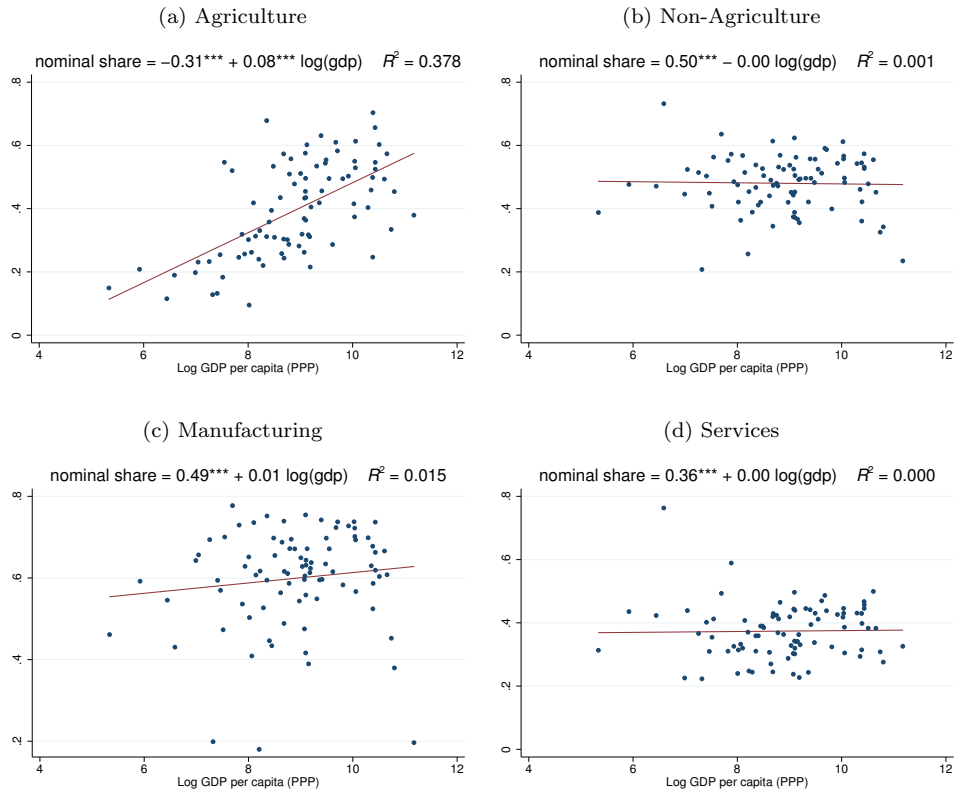
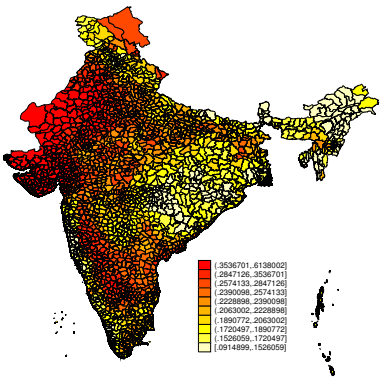


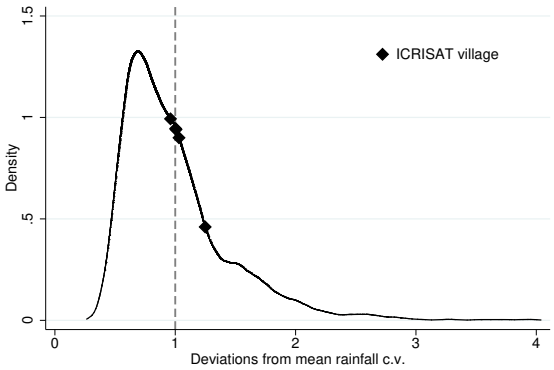
Figure notes: This figure plots nominal intermediate input shares derived from United Nations sectoral data. Non-agriculture is computed as the entire economy net of the agriculture sector. Also included are regression equations for each share regressed against log GDP per capita. Significance of coefficient estimates at 0.01, 0.05, 0.1 levels denoted by ***, **, and *.

Figure 4: Coefficient of Variation of Annual Rainfall in India

(a) Map



(b) Density



(c) Monthly Rainfall Coefficient of Variation

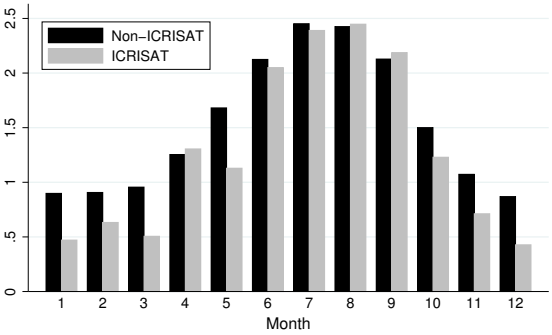


Figure notes: This figure shows the coefficient of variation for approximately 13,688 cells in India, derived from TRMM data. The left panel is a map of India with the implied rainfall c.v. for each cell, with darker red indicating larger variation. The right panel is the density of these estimates. The squares on this graph are estimates for ICRISAT villages used in this analysis.

Figure 5: Fit of Model Distributions

(a) Consumption Growth Volatility (b) Harvest Growth Volatility (c) Savings Rate

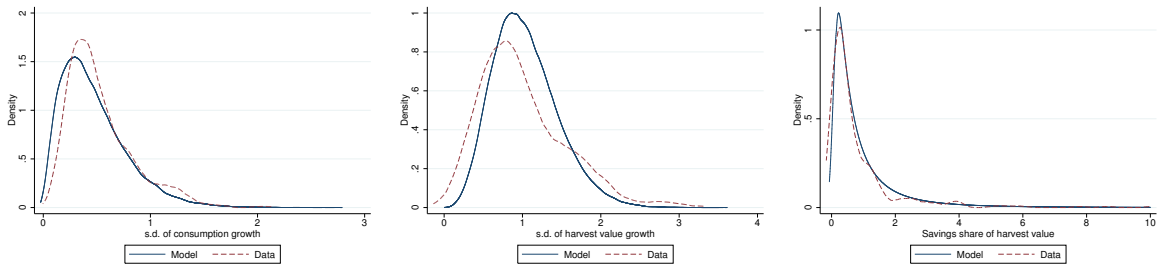


Figure notes: The solid line in all three pictures is the cross-sectional density of volatility in consumption growth at the household level, along with the savings rate as a share of harvest value. All are derived from household-level estimates using ICRISAT panel data. The dashed line are the implied model counterparts, which are computed by simulating 100,000 individuals in the Indian model economy for 6 periods.

Figure 6: Household level distortion $\phi^{risk}(b)$

(a) Baseline distortions (b) %Δ when p_x decreases

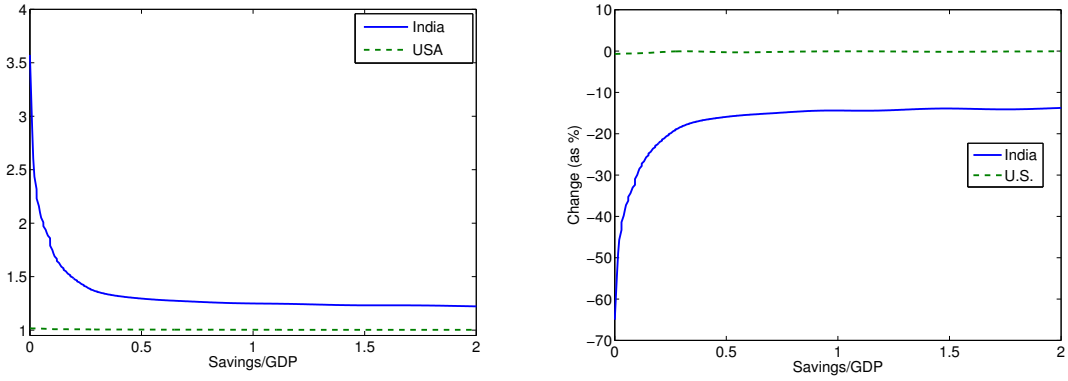


Figure notes: The left panel is the implied household-level distortion generated by risk in both the Indian and U.S. model economies. As discussed in the text, it is the percentage difference between the expected marginal revenue and marginal cost of intermediate use. It is derived formally in Appendix A. The horizontal axis is the value of savings normalized by that economy's GDP. The right panel details how that distortion changes in response to a decrease in the price of intermediates, p_x, from $p_x = 2.77$ to $p_x = 1$. It is computed for both the U.S. and Indian economies.

Appendices (for online publication)

A Uninsured Shocks as a Reduced-Form Distortionary Tax Wedge

In this section, I show that the model developed in the paper is isomorphic to a generic tax wedge. Moreover, this wedge can be decomposed to isolate the impact of risk.

Define the "*ex post* economy" as the economy in which all decisions are made after the realization of shock z. The household problem in the *ex post* economy is

$$\begin{aligned} & v(z,b) = \max_{x,n_a,b'} \quad \alpha \log(c_a - \bar{a}) + (1-\alpha)\log(c_m) + \beta \int_Z v(z',b')dQ(z) \\ s.t. \quad & p_a c_a + c_m + p_a b' = zAx^{\psi} n_a^{\eta} - (1+\phi(z,b))x + w(1-n_a) + p_a(1-\delta)b + \Phi(z,b) \\ & b \geq 0. \end{aligned}$$

The link between the two models is the tax $\phi(z,b)$ on intermediates, which is rebated back to households as $\Phi(z,b)$. The rest of the model is identical to that in the main body of the paper. Proposition 3 shows that the tax $\phi(z,b)$ can be designed to implement the equilibrium of the incomplete markets model developed in this paper, and moreover, is log separable in z and b, so that the risk-driven distortion can be isolated.

Proposition 3. *For an economy with TFP A, there exists a tax function $\phi(z,b)$ such that the equilibrium of the ex post economy is identical to the incomplete markets economy. The tax can be decomposed as $1+\phi(z,b) = \phi^{time}(z) \times \phi^{risk}(b)$ where*

$$\begin{aligned} \phi^{time}(z) &= \frac{z^{1/(1-\eta)}}{\int_Z z^{1/(1-\eta)} dQ(z)} \\ \phi^{risk}(b) &= \frac{\int_Z z^{1/(1-\eta)} dQ(z)}{\int_Z z^{1/(1-\eta)} \left(\frac{\widetilde{u}'(y(x^I(b),z))}{\mathbb{E}_z(\widetilde{u}'(y(x^I(b),z))} \right) dQ(z)} \end{aligned}$$

and $x^I(b)$ is the decision rule the intermediate choice for the baseline incomplete markets model.

Proof. Assume an equilibrium of the baseline model economy (i.e. the model in the main body of the paper) characterized by decision rules $x^I(b)$, $b'^I(b,z)$, invariant distribution $\mu(b)$, and equilibrium price p_a.

Assume that the equilibrium price in the *ex post* economy is equal to p_a. First, this is sufficient to guarantee that the choice of x is the same for all (z, b). The first order condition for x in the *ex post* economy for an individual with savings b and shock realization z, after solving for $n_a(z, b)$, is

$$Ap_A^{1/(1-\eta)} F'(x) z^{1/(1-\eta)} = p_x(1 + \phi(z, b))$$

When

$$1 + \phi(z, b) = \frac{z^{1/(1-\eta)}}{\int_Z z^{1/(1-\eta)} \left(\frac{\widetilde{u}'(y(x^I(b), z))}{\mathbb{E}_z(\widetilde{u}'(y(x^I(b), z))} \right) dQ(z)},$$

it follows that $x^{IM}(b)$ is the only solution to this problem, and is independent of the realization of z. From there, the fact that the transfer is rebated back to the household insures that profit is the same for all individuals with individual state (z, b) in both the *ex post* economy and the baseline economy. The transfer of tax revenue back to households guarantees that income is identical across economies as well. Since income is the same, savings decisions are the same as well, thus implying that the invariant distribution across savings is identical.

Production and income decisions are therefore identical in the two economies. Since markets clear in the incomplete markets economy, they must also in the *ex post* economy. Since p_a is the unique equilibrium price, this implies that the equilibrium in the *ex post* economy is identical to that of the baseline model economy. ■

The portion of the tax $\phi^{time}(b)$ simply accounts for the change in timing between the models, and is irrelevant for the risk-driven misallocation. $\phi^{risk}(b)$ is the share of the distortion generated by increased relative risk aversion among poor households. It implies a positive tax on all households ($\phi^{risk} > 1$), but a higher tax for poor households. Therefore, uninsured shocks work by misallocating resources away from low wealth households in the same way as models of explicit input market distortions, though the distortion instead comes from the inability of households to insure consumption.

B Additional Results

B.1 Changes in the Variance of z

I also investigate the importance of the shock distribution by varying the standard deviation of the underlying normal distribution σ_z, while holding the support $\underline{z}$ and $\overline{z}$ fixed. The results are presented in Table 9.

Table 9: Model Results for Different σ_z

	Labor Productivity Gap		$p_x X / p_a Y_a$		N_a *(%)*	
Economy	Agriculture	Aggregate	Rich	Poor	Rich	Poor
Model with						
$\sigma_z = 0.50$	45.3	7.8	0.40	0.25	1.2	50.0
$\sigma_z = 0.75$	43.2	6.4	0.39	0.25	1.0	34.9
$\sigma_z = 1.00$	41.7	5.7	0.39	0.25	0.7	24.0

Higher standard deviations result in smaller productivity differences. However, all of the difference comes from the amount of labor in agriculture. Intuitively, the result is due to the interaction of the low utility weight on agricultural consumption, α, and subsistence requirements $\bar{a}$. Because α is so low, total agricultural output needs to be roughly $\bar{a}$. When σ_z is low, the price p_a must increase to incentivize people to produce with risky intermediate inputs. As σ_z increases, a larger and larger number of households "luck" into a good shock, and are able to produce $\bar{a}$ and the equilibrium price remains low. This impact is counteracted in part by the fact that households are subject to more risk. Hence, a doubling in the standard deviation of shocks implies only an 8 percent decrease in the agricultural productivity gap.

B.2 Changes in Productivity As Shocks Become Insurable

In the paper, I compare the complete and incomplete markets regimes. Here, I allow for some fraction of the shocks to be insurable. In particular, I re-write the production function as

$$y_a = e^{z_1 + z_2} A x^{\psi} n_a^{\eta}$$

where z_1 and z_2 are both normal random variables with mean zero and $\sigma_1 + \sigma_2 = 0.48$. That is, together, the shocks have the same mean and variance as the shock process in the paper. Here, however, I assume that z_1 is uninsurable while z_2 is insurable. As $\sigma_2 \to 0.48$, the results converge to the complete markets case, while $\sigma_2 \to 0$ implies the baseline results.

Table 10: Changing % of insurable variance

	Labor Productivity Gap		$p_x X / p_a Y_a$		N_a *(%)*	
Economy	Agriculture	Aggregate	Rich	Poor	Rich	Poor
% of variance insurable						
0	45.0	7.8	0.40	0.26	1.2	0.50
20	41.4	7.3	0.40	0.28	1.2	0.45
60	39.0	7.1	0.40	0.31	1.2	0.43
80	38.8	6.9	0.40	0.32	1.2	0.42
100	34.5	6.4	0.40	0.40	1.2	0.34

Even when a large proportion of shocks are insurable, there is still a substantial gap from the results in which all shocks are insurable. This highlights the importance of downside risk. While increasing the proportion of variance insured decreases the likelihood of receiving a low uninsured shock, it does not change the fact that households put more weight on those realizations, which somewhat limits the impact of decreasing uninsurable risk.

B.3 Dynamics of Surprise Response to Decrease in Intermediate Price

To further understand the impact of a change in the distortion p_x, Figure 7 shows the dynamics of the response of the baseline Indian economy to a surprise decrease in p_x from 2.77 to 1.00. The model reaches its new steady state relatively quickly. In doing so, the agricultural aggregates all overshoot their steady state levels. Initially when $p_x = 1$, households are saving too much and thus are willing to take on more risk, as evidenced by the sharp increase in the nominal intermediate share. Savings is a costly activity here, so there is a balance between savings and consumption. As such, savings continually drops

in the economy, as the incentives to self-insure decrease as p_x decreases. As households dissave, the relative price, employment share, and nominal intermediate share all converge to the new steady state.

B.4 Impact of τ and δ across countries

In this section, I provide the counterpart of results presented in Section 7. In the main text, I compute the results for changes in the intermediate input share p_x, while here I provide identical results when I vary the tax on labor τ and the depreciation rate δ. Relative to both p_x and τ, the effects of depreciation are small.

Table 11: Agricultural productivity as τ changes

	$A = 0.22$	$A = 1$
$\tau = 0.55$	0.023	0.663
$\tau = 0$	0.039	1.000
% increase	72.26	50.83

Table 12: Agricultural productivity as δ changes

	$A = 0.22$	$A = 1$
$\delta = 0.15$	0.023	1.000
$\delta = 0.03$	0.026	1.000
% increase	14.4	0.00

C Alternative Explanations

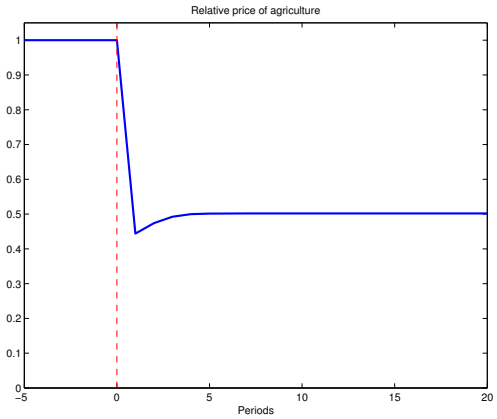
In this section, I detail two other possible explanations.

C.1 Different Shocks or Different Responses?

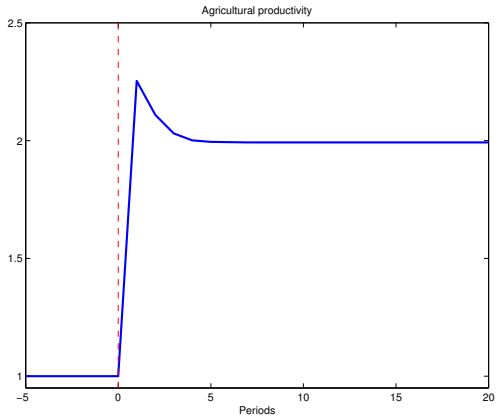
An alternative explanation to the one highlighted in this paper is that poor countries could simply face different exogenous agricultural shocks. I explore this alternative hypothesis using detailed information on historical rainfall fluctuations. Aggregating rainfall data to study cross-country variation present some difficulties however. For one, agriculture

Figure 7: Dynamics of a surprise decrease in p_x in Indian economy at $t = 0$. Initial steady state values are normalized to one.

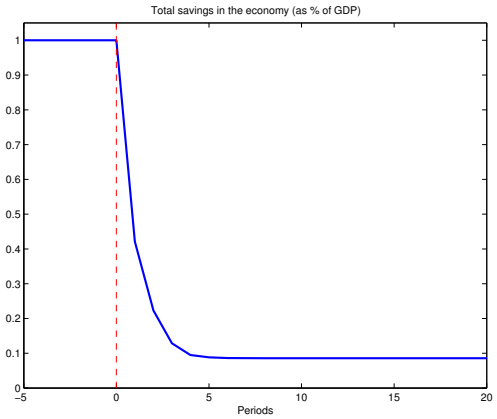
(a) Relative price of agriculture



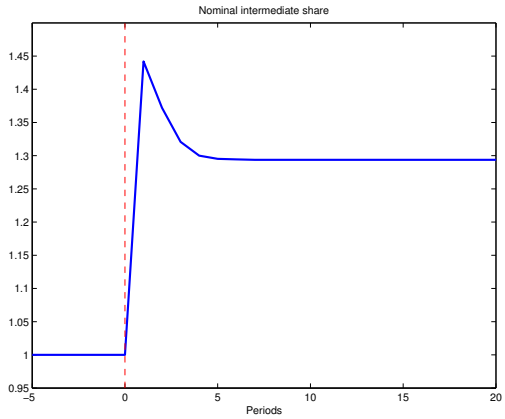
(b) Agricultural productivity



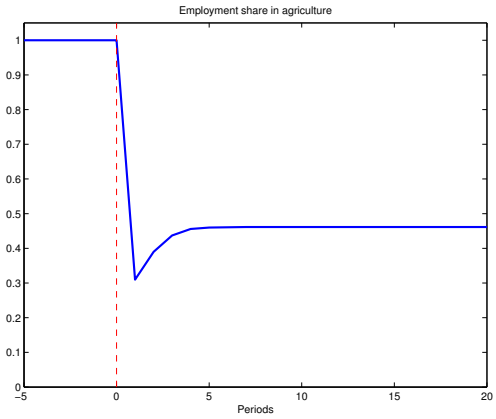
(c) Total Savings/GDP



(d) Nominal intermediate share



(e) Employment share in agriculture



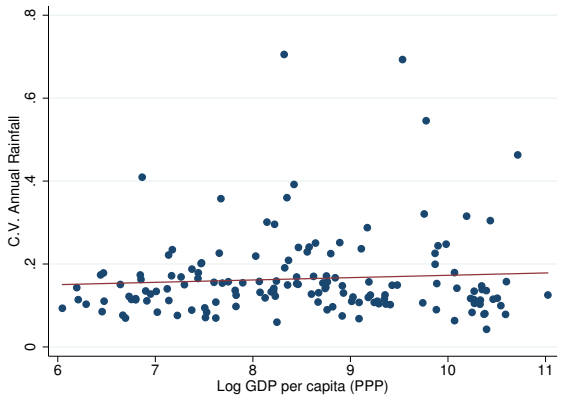
is not uniformly produced across a country's geographic regions. If production occurs in areas that have more stable rainfall, for example, a country-wide average of rainfall variation will overstate the risk faced by farmers. I correct for these issues using the Global Agro-Ecological Zones (GAEZ) data produced by the FAO and International Institute for Applied Systems Analysis. The GAEZ data is spatial grid data on historical rainfall at the 5 arc minute resolution, similar to the TRMM data used in previous sections, and includes approximately 9 million cells. The advantage of the GAEZ is that for the year 2000, it contains the internationally priced value of harvest in each cell. I use this to compute harvest-weighted country-level rainfall for the years 1980-2000, and then compute the country-level variation in rainfall over that period. More specifically, I use arcGIS to assign each cell i to its respective country j, denoting the set of cells in country j as $\mathcal{C}_j$. Then, for each country j, I compute the harvest-weighted annual rainfall as

$$rain_{jt} = \sum_{i \in \mathcal{C}_j} rain_{ijt} \times \left(\frac{Y_{ij}}{\sum_{k \in \mathcal{C}_j} Y_{kj}} \right)$$

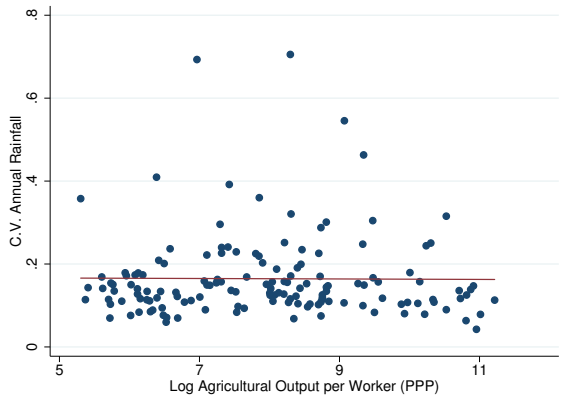
where $rain_{ijt}$ is annual rainfall in cell i in country j in year $t \in \{1980, \ldots, 2000\}$ and Y_{ij} is the internationally priced value of annual harvest in 2000. Figure 8 shows the relationship between the coefficient of variation of the country-level rainfall estimates and both GDP per capita and agricultural output per worker for 147 countries in the year 2000.

Figure 8: Rainfall Variation Across Countries

(a) GDP per capita (PWT)



(b) Agricultural Productivity (FAO)



I find no evidence of a trend in either relationship, and Table 13 confirms this with a simple linear regression of the coefficient of variation of annual rainfall (measured as the z-score) on the two productivity measures. A one standard deviation increase in the rainfall c.v. is associated with a one percent decrease in agricultural productivity, not nearly large enough to matter for agricultural productivity differences across countries. The result echoes recent work by Adamopoulos and Restuccia (2015) who show that natural disadvantages (soil and land quality, for example) are not responsible for low agricultural productivity in poor countries. Instead, these results suggests that the differential response to risk across countries is key for understanding the relationship between risk and intermediate inputs, not variation in the exogenous shocks themselves.

Table 13: Relationship between productivity and rainfall variability

	Log GDP per capita	Log agricultural output per worker
Constant	8.529***	7.956***
	(0.104)	(0.128)
normalized c.v. rainfall	0.090	−0.013
	(0.105)	(0.127)
R^2	0.005	0.000

Table notes: Standard errors are in parentheses. Significance at 0.01, 0.05, 0.1 levels denoted by ***, **, and *. The independent variable is z-score of the coefficient of variation for harvest-weighted annual rainfall 1980-2000.

C.2 Correlation Between Crops and Consumption Volatility

Another alternative is that the harvesting of certain crops are highly correlated with consumption volatility. I therefore compute the average harvest value share of castor, cotton, maize, paddy, pigeon pea, sorghum, and wheat for each household in the ICRISAT panel. I then correlate them with the coefficient of variation and standard deviation of the growth rate of consumption, where consumption is in equation (5.2) in the main text. There is no correlation between crop choice and the two measures of consumption volatility.

Table 14: Partial correlation of harvest share with consumption volatility measures

	σ consumption growth	cv consumption
castor	0.005	0.013
cotton	0.030	0.069
maize	-0.007	0.002
paddy	-0.087	-0.065
pigeon pea	-0.049	-0.055
sorghum	0.020	-0.027
wheat	-0.042	-0.078

Table notes: Significance at 0.01, 0.05, 0.1 levels denoted by ***, **, and *.

D Data Sources and Construction

D.1 Productivity and Intermediate Input Share Statistics

I make use of Prasada Rao (1993), which is the data underlying Restuccia et al. (2008).

Intermediate Shares As in the text, the domestic intermediate share in agriculture of country j is

$$\widehat{X}^j := \frac{p_x^j X^j}{p_a^j Y_a^j} \tag{D.1}$$

This measure is not directly reported in Prasada Rao (1993). He does however, report the real intermediate share in agriculture, defined as

$$\widehat{X}^{j*} := \frac{p_x^* X^j}{p_a^* Y_a^j} \tag{D.2}$$

where p_x^* and p_a^* are international prices of intermediate inputs and agricultural output. Combining equations (D.1) and (D.2), it is possible to write the domestic intermediate share as

$$\widehat{X}^j = \widehat{X}^{j*} \left(\frac{p_x^j / p_x^*}{p_a^j / p_a^*} \right) \tag{D.3}$$

The price ratio in equation (D.3) can be calculated from reported purchasing power parities

$$\begin{aligned} PPP_a^j &= \frac{p_a^j}{p_a^*} \\ PPP_x^j &= \frac{p_x^j}{p_x^*} \end{aligned}$$

where p_a^* and p_x^* are international (unreported) prices and (p_a^j, p_x^j) 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, $PPP_a^{US} = PPP_x^{US} = 1$, implying $\widehat{X}^{US} = \widehat{X}^{US*}$. Therefore, calculating the domestically priced intermediate share of all other countries reduces to

$$\widehat{X}^j = \widehat{X}^{j*} \left(\frac{PPP_x^j}{PPP_a^j} \right) \tag{D.4}$$

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 right hand side of equation (D.4) is the statistic reported in Figure 1b. The horizontal axis, GDP per capita, is real GDP per capita for 1985, variable cgdp from the Penn World Tables version 7.0 (PWT).

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

For the comparison of agriculture to manufacturing and services, I use the publicly available U.N. System of National Accounts. For each sector, I use "Output, at basic prices" as output and "Intermediate consumption, at purchaser's prices" as intermediate inputs. The sectors are defined by aggregating the following industry codes:

1. Agriculture: A, B
2. Manufacturing: C, D, E, F
3. Services: G, H, I, L, M, N, O, P
4. Non-agriculture: Entire Economy - A - B

I use all countries that have a complete set of required data, which is maximized in 2005 (though the results are robust to any other choice of years). The final dataset includes 87 countries. Note that the intermediate share in agriculture derived from the UN statistics and the FAO statistics may differ. This is due to the fact that the UN statistics includes intermediate inputs produced in the agricultural sector, while the FAO statistics only consider nonagricultural intermediate inputs.

D.3 Micro Evidence: LSMS and NASA Weather Data

Weather The weather data is downloaded from the Goddard Earth Sciences Data and Information Center, online here http://goo.gl/nlFomb. The data includes 0.25×0.25 degree monthly rainfall estimates on latitudes [-50, 50] and longitudes [-180,180] every month form January 1998 - current (April 2015) for 576,000 monthly data points around the world.

LSMS The LSMS data is available from the World Bank. See http://iresearch.worldbank.org/lsms/lsmssurveyFinder.htm for a convenient survey locator. Since the goal is to construct variation in weather dating back to only 1998, I restrict attention to surveys completed after 2005. The datasets used must meet four criteria:

1. Have GPS coordinates for households to merge with weather data
2. Fertilizer quantities and either prices or total value to compute prices.
3. One year previous livestock holdings

Malawi, Niger, Tanzania, and Uganda satisfy the three criteria.

E Proofs

E.1 An Additional Lemma for the Proof of Proposition 1

To prove the result, I first characterize the equilibrium of an I 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 I economies characterized by TFP levels A^1 and A^2, both with $\bar{a} = 0$. Denote the equilibrium for economy 1 as $(x^1, n_a^1(z), p_a^1)$. Then the equilibrium for economy 2, $(x^2, n_a^2(z), p_a^2)$ can be characterized as*

$$\begin{aligned} n_a^2(z) &= n_a^1(z) \\ x^2 &= \left(\frac{A^2}{A^1}\right) x^1 \\ p_a^2 &= \left(\frac{A^1}{A^2}\right)^{\psi} p_a^1 \end{aligned}$$

Proof. Two things must be checked for the proposed allocation to be a competitive equilibrium. First, the proposed equilibrium must satisfy the household optimization problem. That is, if $(p_a^1, x^1, n_a^1(z))$ is an equilibrium in economy 1, then $(p_a^2, x^2, n_a^2(z))$ satisfies the farmer's optimization problem in economy 2. Second, markets must clear. These are considered in turn.

Optimization Problem The first thing to check is that the labor choice is identical between the two. Using the decision rules, I can check this using the first order conditions for $n_a^1(z)$ and $n_a^2(z)$.

$$\frac{n_a^1(z)}{n_a^2(z)} = \left(\frac{p_a^1 A^1 (x^1)^{\psi}}{p_a^2 A^2 (x^2)^{\psi}}\right)^{1/(1-\eta)}$$

Plugging in (p_a^2, x^2) implies

$$\frac{n_a^1(z)}{n_a^2(z)} = 1$$

For simplicity, I drop the superscript on $n_a(z)$, with the understanding that they are identical in both economies.

Next up is to check if x^2 satisfy the required first order conditions, given that x^1 satisfies the first order condition in Economy One. Note that when $\bar{a} = 0$, the production utility for a given income y can be written as

$$\begin{aligned} v^p(y) &= \alpha \log(c_a^1) + (1-\alpha)\log(c_m^1) \\ &= \Omega - \alpha \log(p_a^1) + \log(y) \end{aligned} \tag{E.1}$$

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 $j = 1, 2$ as

$$y^j(x, z) = p_a^j A^j z x^{\psi} n_a(z)^{\eta} - x + (1 - n_a(z))A^j$$

Plugging in the proposed equilibrium yields the following relationship

$$y^2(x^2, z) = \left(\frac{A^2}{A^1}\right) y^1(x^1, z) \tag{E.2}$$

Equation (E.1) implies that

$$x^j = \arg\max_x \int_Z \log(y^j(x, z))dQ(z)$$

After plugging in the optimal values for $n_a(z)$, the first order condition for this problem can be written as

$$\int_{\underline{z}}^{\bar{z}} \left(\frac{\psi p_a^j z A^j x^{j\psi-1} n_a(z)^{\eta} - 1}{y^j(x, z)}\right) = 0$$

Plugging in the proposed equilibrium yields a relationship between economies one and two

$$\int_{\underline{z}}^{\bar{z}} \left(\frac{\psi p_a^2 z A^2 x^{2\psi-1} n_a(z)^{\eta} - 1}{y^2(x, z)}\right) = \left(\frac{A^1}{A^2}\right) \int_{\underline{z}}^{\bar{z}} \left(\frac{\psi p_a^1 z A^1 x^{1\psi-1} n_a(z)^{\eta} - 1}{y^1(x^j, z)}\right)$$

Since an equilibrium is assumed in economy one, it follows then that

$$\int_{\underline{z}}^{\bar{z}} \left(\frac{\psi p_a^2 z A^2 x^{2\psi} n_a(z)^{\eta} - 1}{y^2(x, z)}\right) = 0$$

Therefore, the proposed economy two equilibrium satisfies a household's optimization problem.

Market Clearing Aggregate sector a output for economy $j = 1, 2$ is

$$Y_a^j = A x^{j\psi} \mathbb{E}_z(z n_a(z)^{\eta})$$

Thus,

$$\frac{Y_a^1}{Y_a^2} = \left(\frac{A^1}{A^2}\right) \left(\frac{x^1}{x^2}\right)^{\psi} \tag{E.3}$$

Therefore, at the proposed equilibrium,

$$\frac{Y_a^1}{Y_a^2} = \left(\frac{A^1}{A^2}\right)^{1+\psi} \tag{E.4}$$

For any $\bar{a} \geq 0$, the total demand for sector a consumption is given by

$$D_a^j = (1-\alpha)\bar{a} + \frac{\alpha}{p_a^j}\mathbb{E}_z[y^j(X^j, z)] \tag{E.5}$$

Using equation (E.2),

$$\frac{\mathbb{E}_z[y^1(x^1, z)]}{\mathbb{E}_z[y^2(x^2, z)]} = \frac{A^1}{A^2} \tag{E.6}$$

Since $\bar{a} = 0$, equations (E.5) and (E.6) and the prices p_a^1 and p_a^2 imply that

$$\frac{D_a^1}{D_a^2} = \left(\frac{A^1}{A^2}\right)^{1+\psi} \tag{E.7}$$

Since the proof assumes an equilibrium in economy 1, equations (E.4) and (E.7) imply $Y_a^2 = D_a^2$ so that the agricultural output market clears in economy two. Since the labor market in sector m clears trivially, Walras' Law implies that the sector m output market also clears. ■

E.2 Proof of Proposition 1

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

E.2.1 $n_a(z)$ is independent of A

This follows directly from Lemma 1.

E.2.2 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

$$\hat{X}^j = \frac{x^j}{p_a^j Y_a^j} \tag{E.8}$$

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

$$Y_a^j = A^j (x^j)^\psi \mathbb{E}_z (z n_a^j(z)^\eta) \tag{E.9}$$

Using the fact that $n_a^1(z) = n_a^2(z)$ and plugging (E.9) into (E.8) gives

$$\frac{\hat{X}^1}{\hat{X}^2} = \left(\frac{x^1}{x^2}\right)^{1-\psi} \left(\frac{p_a^2}{p_a^1}\right) \left(\frac{A^2}{A^1}\right)$$

Plugging in the equilibrium found in Lemma 1, this gives

$$\begin{aligned} \frac{\hat{X}^1}{\hat{X}^2} &= \left(\frac{A^1}{A^2}\right)^{1-\psi} \left(\frac{A^1}{A^2}\right)^{\psi} \left(\frac{A^2}{A^1}\right) \\ &= 1 \end{aligned}$$

Since A^1 and A^2 are arbitrary, this completes the proof.

E.2.3 No increase in productivity relative to C economy

For any two economies characterized by TFP A^1 and A^2 and complete markets (the C economy), it is easy to show that in equilibrium,

$$\begin{aligned} n_a^1 &= n_a^2 \\ x^2 &= \left(\frac{A^2}{A^1}\right) x^1 \end{aligned}$$

Since this is the same as in the incomplete markets model (the I economy), relative agricultural labor productivity between the two economies is equal in both.

■

E.3 Proof of Proposition 2

Proof. Consider the equilibrium for economy 1 with TFP equal to A^1. Denote this equilibrium $(p_a^1, x^1, n_a^1(z))$. Suppose that the intermediate good share is $\hat{X}^1 < \psi$. Define x^{1C} to be the optimal choice of the farmer who faces p_a^1 but with complete markets. We know

that the intermediate good share is $\hat{X}^{1C} = \psi$. Therefore, the ratio is

$$\frac{\hat{X}^1}{\hat{X}^{1C}} = \frac{\hat{X}^1}{\psi} = \left(\frac{x^1}{x^{1C}}\right)^{(1-\eta-\psi)/(1-\eta)}$$

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

$$\hat{X}^1 = \psi \left(\frac{x^1}{x^{1C}}\right)^{(1-\eta-\psi)/(1-\eta)}$$

Similarly, it follows that in Economy 2,

$$\hat{X}^2 = \psi \left(\frac{x^2}{x^{2C}}\right)^{(1-\eta-\psi)/(1-\eta)}$$

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

$$\frac{x^1}{x^{1C}} > \frac{x^2}{x^{2C}}$$

This follows from the fact that, when $\bar{a} > 0$, relative income net of subsistence,

$$\frac{y^1(z) - p_a^1 \bar{a}}{y^2(z) - p_a^2 \bar{a}}$$

is decreasing in z. ■