Appendix:
"The Equilibrium Impact of Agricultural Risk on Intermediate Inputs and Aggregate Productivity"
by Kevin Donovan
Review of Economic Studies
August 2020

Table of Contents for Online Appendix

A Uninsured Shocks as a Reduced-Form Distortionary Tax Wedge

In this Appendix, I show that the model developed in the paper is isomorphic to a generic tax wedge in two models: (1) a complete markets model and (2) a model where farmers choose x after the shock realization. Furthermore, I show that the only difference in the implied distortion is an adjustment for the timing difference that depends only on the shock realization. This shows that the ability to insure *ex post* consumption manifests in a way that looks identical to explicit input market distortions.

A.1 In a Complete Markets Model

In the complete markets economy, households can trade a full set of state-contingent assets before the realization of the shock z. Thus, every household maximizes expected profit. Profit is

$$\pi(z,x,b) = \max_{n_a \geq 0} \; zx^{\psi}n_a^{\eta} - (1+\phi^{risk}(b))x + w(1-n_a)$$

where $\phi^{risk}(b)$ is a reduced-form tax levied on those with savings b. With complete markets, total income is given by

$$y(b) = \max_{x \geq 0} \mathbb{E}_z[\pi(z,b)] + p_a(1-\delta)b + \Phi^{risk}(b)$$

Notice that the only reason income depends on b is through the tax and transfer scheme. The household problem in this model is

$$\begin{aligned} v(b) = \max_{b' \geq 0} \quad & \Big(\alpha \log(c_a - \bar{a}) + (1-\alpha)\log(c_m) + \beta v(b')\Big) dQ(z) \\ s.t. \quad & p_a c_a + c_m + p_a b' = y(b). \end{aligned}$$

The rest of the model is identical to that in the main body of the paper. Proposition 1 shows that this tax $\phi^{risk}(b)$ and transfer $\Phi^{risk}(b)$ can be constructed to generate the incomplete markets equilibrium in a complete markets model.

Proposition 1. *For an economy with TFP A, there exists a tax function $\phi(b)$ such that the equilibrium of the complete markets economy is identical to the incomplete markets economy.*

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

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 complete markets economy is equal to p_a. Combining the first order conditions for x and n_a in the complete markets economy gives

$$Ap_a^{1/(1-\eta)}F'(x)\mathbb{E}[z^{1/(1-\eta)}] = p_x(1+\phi(b)).$$

When

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

it follows that $x^I(b)$ is the only solution to this problem. From there, the fact that the transfer is rebated back to the household insures 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 complete markets economy. Since p_a is the unique equilibrium price, this implies that the equilibrium in the complete markets economy is identical to that of the baseline model economy. ■

As one can see from the definition of $\phi^{risk}(b)$, this reduced-form tax wedge is 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.

A.2 In a Model with *Ex Post* Input Choices

The result in the previous section easily extends to a model in which inputs are chosen after the shock realization. This is useful to link the distortion generated in this model to models in which households know the shock realization before making input choices, but are subject to some *ex post* distortion. This is standard, for example, in models that generate misallocation through financial frictions (e.g. Buera, Kaboski and Shin, 2011).

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

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

Again, there is tax $\phi(z,b)$ on intermediates, which is rebated back to households as $\Phi(z,b)$. Unlike the previous result, the tax and transfer now depends on the shock realization z. Proposition 2 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 multiplicative in z and b, so that the risk-driven distortion can be isolated.

Proposition 2. *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 $\phi^{risk}(b)$ is the same distortionary function in the complete markets model. That is,*

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

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

Proof. The proof proceeds identically to the complete markets proof in the last section, except to note that the first order condition is now

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

Now, the realization of z enters directly into the first order condition, and thus the tax ϕ must also depend on z. 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. The rest of the proof is identical. ∎

Note that the portion of the tax $\phi^{time}(b)$ simply accounts for the change in timing between the two models. $\phi^{risk}(b)$ contains all the same properties as in the previous section, and controls the risk-induced distortion in the model.

B Additional Results

B.1 Intermediate Shares in U.N. Data

To test the robustness of the results in Section 2, I redo the exericse using data from 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. The benefit relative to the WIOD data used in the main text is larger country coverage, primarily among poor countries. The cost is that it does not allow one to separate the production sector of intermediates consumed.

These nominal intermediate shares are plotted in Figure 1, along with the nonagricultural sector measured as the total economy net of agriculture. Figure 1a confirms the relationship between the agricultural intermediate input share and per capita GDP. Figures 1c and 1d, 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, thus finding similar results to those from the WIOD.

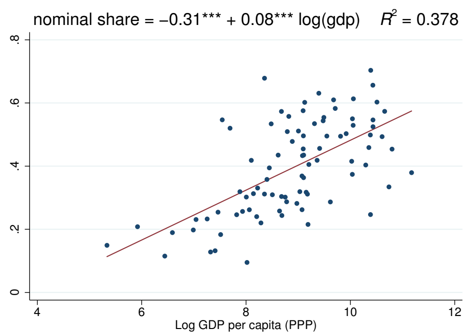
There are two key differences between the UN and the WIOD data. First, while both the UN and WIOD data are derived from national accounts, the WIOD uses a stricter rule for data quality, thus limiting sample size.[1] The UN data provides a larger number of extremely poor countries (the poorest WIOD country is India), though at the expense of some data quality uncertainty. I use the UN data in the main text to provide a broader view of the cross-sectional intermediate use, but the fact that the two datasets provide similar results suggests data quality is not the key issue.

The second issue is that the UN data does not distinguish the sector in which intermediates are produced, reporting only total intermediate consumption by sector. To maintain consistency, the results in Figure 2 do not either. However, the more detailed WIOD data allows me to decompose intermediate consumption in agriculture by sector of production as in the main text. Intermediates produced in the non-agricultural sector drive the correlation with income, as assumed in the model.

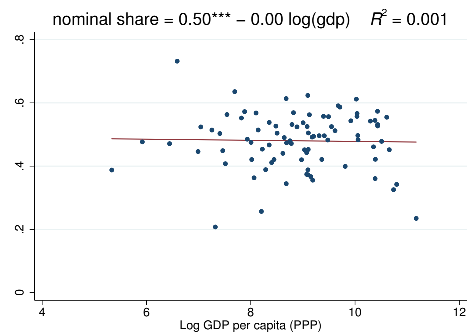
[1] Timmer et al. (2015) provides a detailed description of construction. Note that some of this data quality requirement is to allow the WIOD to construct I-O tables, which is not relevant for the UN data.

Figure 1: Sectoral nominal intermediate shares (2005)

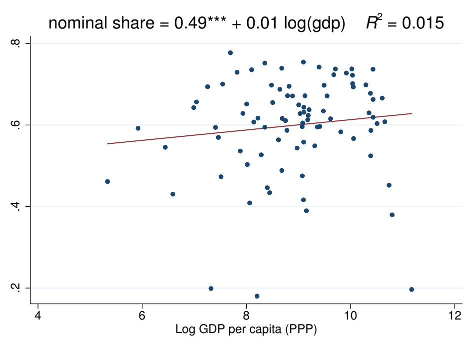
(a) Agriculture



(b) Non-Agriculture



(c) Manufacturing



(d) Services

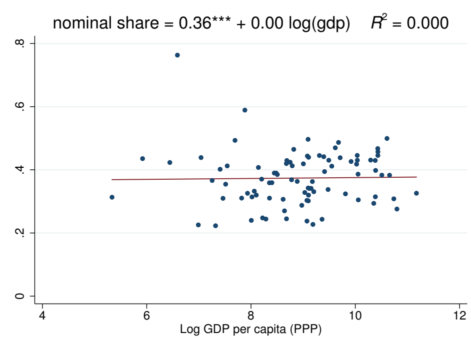


Figure notes: This figure plots nominal intermediate input shares derived from United Nations sectoral data. Non-agriculture is computed as the entire economy net of the agriculture sector. Also included are regression equations for each share regressed against log GDP per capita. Significance of coefficient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

B.2 Further Results from World Input-Output Database

This section further breaks down the non-agricultural sector results of Section 2 using WIOD data. The WIOD predicts a slightly negative relationship in manufacturing (Figure 2a) and slightly positive relationship in services (Figure 2b). Even with the slightly positive result in services, the quantitative magnitude is significantly larger in agriculture. The slope coefficient is more than 3 times higher in agriculture than services, and the R^2 is 4.5 times larger.

Figure 2: Sectoral nominal intermediate shares (2005)

(a) Manufacturing (b) Services

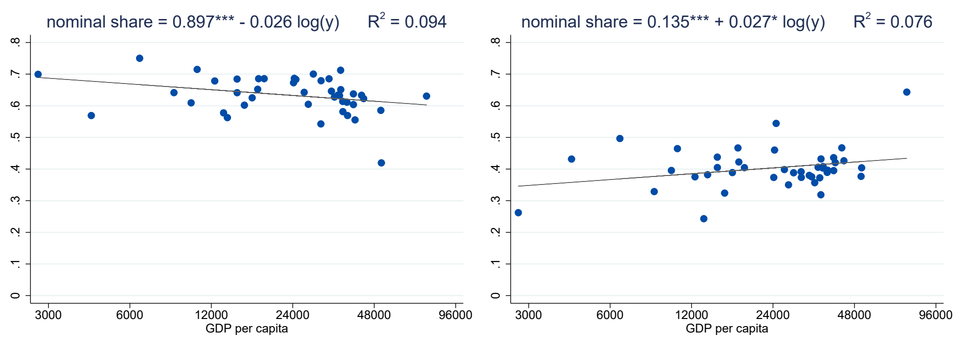


Figure notes: This figure plots nominal intermediate input shares derived from World Input-Output Database. It reproduces the results the main paper, which uses UN data. Non-agriculture is computed as the entire economy net of the agriculture sector. Also included are regression equations for each share regressed against log GDP per capita. Significance of coefficient estimates at 0.01, 0.05, 0.1 levels denoted by ***, **, and *.

B.3 Aggregate Intermediate Shares Constructed from Micro Data

As a final check of the data, I compute the aggregate intermediate shares by aggregating micro data. I use 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.[2] 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 3 combines this data with Penn World Table GDP per capita, and confirms the positive relationship.

Figure 3: Nominal expenditure shares from LSMS micro data

(a) Fertilizer (b) Fertilizer and Pesticide

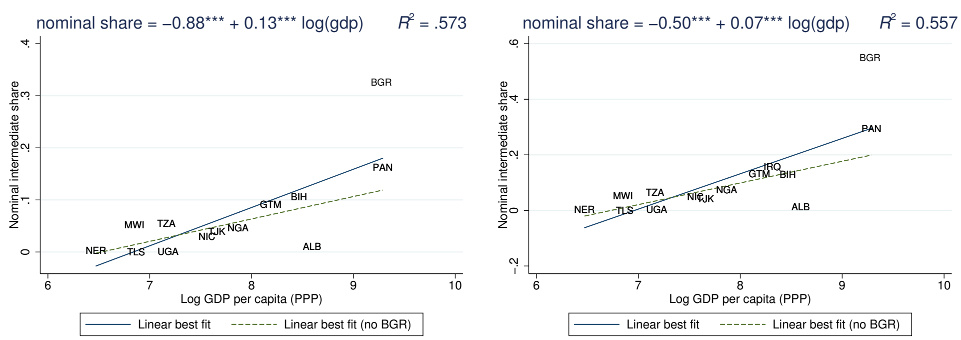


Figure notes: Aggregated intermediate shares using (1) fertilizer and (2) fertilizer and pesticide computed from World Bank LSMS micro data. Also included are the regressions equations derived from regressing these aggregate shares on log GDP per capita. Significance of coefficient estimates at 0.01,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

[2] 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.

B.4 Capital Variation Across Countries

The empirical results in Section 2 show that there are large differences in intermediate intensity across countries. To provide some context for the scale of differences, I compare them here to cross-country differences in agricultural capital.

I use data from the FAO (FAOSTAT, 2018) on agricultural inputs (capital, employment, land, and intermediates) and output to compute simple variance decompositions in the style of Caselli (2005). In particular, for the year 2000, I compute a measure of production as

$$y_{k,l,x} = k^{\theta_k} l^{\theta_l} x^{\theta_x}$$

where lowercase variables stand for those normalized by employment (e.g., $y \equiv Y/N$). I then consider two other production functions setting $\theta_k = 0$ and $\theta_x = 0$,

$$\begin{aligned} y_{l,x} &= l^{\theta_l} x^{\theta_x}, \\ y_{k,l} &= k^{\theta_k} l^{\theta_l}. \end{aligned}$$

Following Caselli (2005) I compute a simple measure of the fraction of true production variance captured by these models:

$$success = \frac{var(\log y)}{var(\log y^{data})}$$

where the numerator uses one of the production functions above ($y_{k,l,x}$, $y_{l,x}$, or $y_{k,l}$). I set $\theta_x = 0.40$, $\theta_k = 0.18$ and $\theta_l = 0.066$. The latter two imply capital and land shares of 0.30 and 0.11 in value added, consistent with Valentinyi and Herrendorf (2008). The results are in Table 1. Variation in intermediates accounts for a larger share of the variance than capital. Not including intermediates decreases the share of variance accounted for from 44 to 10 percent. The same procedure with capital decreases the share of variance accounted for from 44 to 21 percent. Consistent with this, the 90/10 ratio across countries is significantly larger for intermediates than for capital.

Lastly, note that there is still a large role of agricultural TFP here, as the "inputs only" model accounts for 43.6 percent of the total variation in production. As discussed in the text, risk affects both margins here. It lowers the real intermediate share, thus affecting input usage, and moreover, affects the residual agricultural TFP term through its affect on misallocation.

I emphasize that this result should not be taken to suggest that agricultural capital is unimportant in the development process, only that there is substantial variation in the intermediate intensity of agricultural sectors across countries. Chen (2020), for example, shows a qualitatively similar pattern in capital shares across countries that

Table 1: Agricultural Input Differences Across Countries (2000)

	Success	90/10 ratio of missing input
Production function used:		
All inputs ($y_{k,l,x}$)	0.436	–
No intermediates ($y_{k,l}$)	0.095	135.87
No capital ($y_{l,x}$)	0.213	28.72

Table notes: Data is from FAOSTAT (2018). The 90/10 ratio is computed across countries with available data for the missing input in the production function.

I find with intermediate inputs. Moreover, agricultural capital is notoriously difficult to measure and as with most capital stocks, quality differences can play an important role. Caunedo and Keller (forthcoming) study this in detail and find an important role for capital quality in understanding sectoral productivity differences across countries.

B.5 Recalibration with Low β

Because of the exogenous interest rate, the discount factor β here is set exogenously. Here, I show that this choice is not consequential. I re-calibrate the full model assuming $\beta = 0.90$. Primarily, this requires a lower savings depreciation rate, with δ declining from $\delta = 0.10$ to $\delta = 0.037$. The other parameters change only slightly. Table 2 reproduces the gains form complete insurance in India, and shows that the gains are nearly identical to the results in the main text.

Table 2: Aggregate Indian Moments with $\beta = 0.90$

	Labor Productivity			
Economy	Agriculture	GDP per capita	$p_x X / p_a Y_a$	N_a (%)
Incomplete markets	1.00	1.00	0.31	44.3
Complete markets	1.15	1.10	0.40	36.5
%Δ	15.4	10.2	29.8	−17.6

Table notes: Incomplete markets productivity measures are normalized to one.

Similarly, the cross-country differences are nearly identical to those in the text.

Table 3: Quantitative Cross-Country Results with $\beta = 0.90$

	Labor Productivity Gap (U.S./India)		$p_x X / p_a Y_a$		N_a (%)	
Economy	Agriculture	GDP per capita	U.S.	India	U.S.	India
Data	51.6	8.2	0.40	0.10	1.4	44.3
Model with						
Incomplete markets	36.2	8.1	0.40	0.31	1.4	44.3
Complete markets	31.6	7.5	0.40	0.40	1.4	36.5

B.6 Varying Distortionary Channels in the Indian Economy

I provide a number of alternative alternative scenarios here, highlighting how the various cross-country differences interact. I focus on the Indian economy, since these changes only affect India, not the U.S.. In Appendix B.7, I highlight how the same changes highlighted here induce different responses between the U.S. and India.

To study the relative importance of various scenarios, I vary all of the parameters that differ between the U.S. and India. I also allow agricultural and manufacturing productivity to differ individually, and I denote them by A_a and A_m respectively. The results are in Table 4.

The first set of results highlight individual parameter changes. The smallest increase is from changing δ, which would result from better savings or storage technologies in India.[3] This increases agricultural productivity by 9 percent. The most impactful change would be an increase in exogenous agricultural productivity, which would increase agricultural labor productivity by 731 percent. Note, interestingly, that it has a substantially negative effect on the real intermediate share as the agricultural price falls by 91 percent.

The second set of results varies multiple parameters simultaneously. First note that the relative importance of sectoral distortions is small relative to exogenous productivity. Removing all sectoral distortions increases agricultural productivity more than two-fold, while removing exogenous TFP differences (but keeping sectoral distortions) increases agricultural labor productivity 16-fold.

Lastly, note that these various exogenous factors are important for matching the price ratio p_x/p_a, which as discussed in the main text, plays a an important role. Eliminating exogenous productivity differences causes the model to vastly over-predict the relationship, while removing the agricultural-sector distortions causes the model to under-predict this price ratio across countries.

[3]Note, however, that even cash storage can be subject to substantial costs. See, for example, Jakiela and Ozier (2016) and Goldberg (2017) on the cost of social and family pressure.

Table 4: Varying Indian Model Distortions and Implications for Indian Equilibrium

	Parameter Values					Equilibrium Changes				
	τ	p_x	δ	A_m	A_a	p_a	N_a	$(p_x X)/(p_a Y_a)$	X/Y_a	Y_a/N_a
Baseline	0.32	2.77	0.10	0.20	0.20	1.000	1.000	1.000	1.000	1.000
Individual parameter changes:										
Set $\delta = 0$	0.32	2.77	0	0.20	0.20	0.92	0.86	1.14	1.04	1.09
Set $p_x = 1$	0.32	1	0.10	0.20	0.20	0.55	0.53	1.14	1.76	1.80
Set $\tau = 0$	0.32	2.77	0.10	0.20	0.20	1.16	0.77	1.06	1.23	1.26
Set $A_m = 1$	0.57	2.77	0.10	1	0.20	2.07	0.37	1.16	2.40	2.40
Set $A_a = 1$	0.32	2.77	0	0.20	1	0.12	0.13	1.32	0.16	8.31
Multiple parameter changes:										
Only TFP differences	0	1	0	0.20	0.20	0.66	0.43	1.19	2.19	2.19
Only sectoral distortions	0.32	2.77	0.10	1	1	0.28	0.05	1.20	0.34	18.19
Only non-agricultural productivity differences	0	1	0.15	0.2	1	0.09	0.06	1.20	0.30	16.24

Table notes: Baseline equilibrium outcomes are normalized to one, for ease of comprehension.

B.7 Amplifying Other Macroeconomic Distortions

A further implication of the model is that the impact of removing a distortion will have substantially larger effects in poor countries than in rich. This is particularly relevant in light of an important literature that has introduced agricultural-specific distortions in complete markets models. I therefore vary the intermediate price p_x, which subsumes a number of different policies considered by the literature.[4]

I decrease the intermediate price p_x from $p_x^{India} = 2.77$ to $p_x^{US} = 1$ in two economies: one in which TFP is $A^{India} = 0.20$ and one with $A^{US} = 1$. Table 5 shows the differential effect on key aggregate moments.

Table 5: % Change in Aggregate Moments in Response to p_x Decline

	$A = 0.20$	$A = 1$
GDP per capita	23.4	0.3
Agricultural Productivity (Y_a/N_a)	80.2	61.5
Agricultural Employment (N_a)	-47.0	-34.4
Real Intermediate Share (X/Y_a)	75.9	71.7
Nominal Intermediate Share (p_xX/p_aY_a)	14.4	0.9

Table notes: An entry in this table is change (as a percentage) in each aggregate moment moving from $p_x = 2.77$ to $p_x = 1$. GDP per capita is measured at baseline ($p_x = 1$) U.S. model prices.

Agricultural productivity increases by 80 percent, compared to a 62 percent increase in the U.S. economy. A similar result holds at the aggregate level. The intuition here is exactly the misallocation margin highlighted in the paper. In the U.S., farmers are roughly profit maximizers at any level of p_x. Thus, there is no quantitatively relevant variation in the real intermediate share across households. One can see this by noting the extremely small change in the nominal intermediate share. Thus, decreasing p_x affects the U.S. economy in much the same way it would in the complete markets model of Restuccia, Yang and Zhu (2008). In India, on the other hand, the decline in p_x has the additional effect of decreasing misallocation. Again, note the large increase in the nominal intermediate share. This generates a substantially larger drop in employment and a slightly larger increase in the real intermediate share (which is tempered by the agricultural price decline).

Overall, the results therefore show that changing distortions is not independent of sector-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, suggesting they are potentially more damaging to these economies than previously estimated.

[4]These include input market distortions (Gollin, Parente and Rogerson, 2004; Restuccia, Yang and Zhu, 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).

B.8 Alternative DRRA Utility Function

The main body of the paper uses a utility function of the form

$$u(c_a, c_m) = \alpha \log(c_a - \overline{a}) + (1 - \alpha)\log(c_m).$$

This utility function is commonly utilized in models of structural transformation, as the non-homothenticity introduced by subsistence $\overline{a}$ implies an income elasticity of less than one for agricultural consumption. This induces structural change. In addition, in generates decreasing relative risk aversion.

In this section, I consider an alternative utility function that admits decreasing relative risk aversion. In particular, I consider

$$u(c_a, c_m) = 1 - \exp\left(-\frac{(c_a^{\alpha} c_m^{1-\alpha})^{1-\sigma}}{1-\sigma}\right)$$

This is a simple extension of the power risk aversion (PRA) utility function introduced in Xie (2000). Denoting C as total consumption expenditures as in the text, one can show that the corresponding $\widetilde{u}(C)$ can be written as

$$\widetilde{u}(C) = 1 - \exp\left(\Omega \left(\frac{1}{p_a}\right)^{\alpha(1-\sigma)} \left(\frac{C^{1-\sigma}}{1-\sigma}\right)\right)$$

where $\Omega = (\alpha^{\alpha}(1-\alpha)^{1-\alpha})^{1-\sigma}$ is a constant. It follows that relative risk aversion is

$$R(C, p_a) = \sigma + \Omega p_a^{\alpha(\sigma-1)} C^{1-\sigma}.$$

It is instructive to note that the price appears in the relative risk aversion term here, as it does in the original specification, even in the absence of subsistence requirement.[5] Decreasing relative risk aversion in C then requires $\sigma > 1$.

The PRA utility function has the benefit of not relying on subsistence requirements, thus sidestepping any existence-related issues introduced by the subsistence requirement. As written here, it cannot generate structural transformation through sector neutral productivity changes, as agricultural consumption is always a constant fraction of total consumption expenditures. I do this to highlight that the qualitative patterns in this paper extend to a broader class of utility functions that do not rely on the structural change properties of the baseline utility function. One could introduce a broader CES aggregator for sectoral consumption with sector-specific technological progress, as in Ngai and Pissarides (2007), to generate such results.

I set $\sigma = 2$ then keep the rest of the calibration fixed net of one parameter – the

[5] Recall that in the main text, $R(C, p_a) = C/(C - p_a\overline{a})$.

utility weight α on agriculture. This plays an important role in generating the gap in nominal intermediate shares, so I vary α as well.[6] The results are in Table 6.

Table 6: Baseline Aggregate Results with PRA Utility

	Labor Productivity Gap		$p_x X / p_a Y_a$		N_a *(%)*	
Economy	Agriculture	Aggregate	Rich	Poor	Rich	Poor
Share of Total Expenditures in Agriculture						
$\alpha = 0.005$	17.1	5.1	0.40	0.39	0.2	0.2
$\alpha = 0.05$	17.4	5.2	0.39	0.35	2.1	2.2
$\alpha = 0.10$	18.0	5.3	0.39	0.34	4.3	4.4
$\alpha = 0.15$	18.3	5.4	0.38	0.32	6.5	6.6
$\alpha = 0.25$	18.8	5.7	0.37	0.30	11.1	11.5

The results show the importance of the specific form of the utility function in generating the results. While the theoretical results (the relationship between TFP and the nominal intermediate share) can be replicated with other DRRA utility functions, the quantitative need not necessarily be as large. Given the importance of the utility function in generating distortions in agriculture, an interesting route for future work is to better understand and extend the set of utility functions that generate the types of effects highlighted in this paper, similar to the work of Boppart (2014) and Comin, Lashkari and Mestieri (forthcoming) on structural transformation.

B.9 Changes in Productivity As Shocks Become Insurable

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 $\sigma_1 + \sigma_2 = 0.32$. 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 $\sigma_2 \to 0.32$, the results converge to the complete markets case, while $\sigma_2 \to 0$ implies the baseline results. The results are in Table 7.

[6] The utility function used here is clearly counterfactual, as the share of expenditures devoted to agriculture is constant across countries and equal to α. As such, α does not have a clear data moment counterpart. I vary it only to show that the variation in nominal intermediate shares across countries can, in principal, be quite large.

Table 7: Varying the share of insurable shocks in the India economy

	p_a	N_a	$(p_x X)/(p_a Y_a)$	X/Y_a	Y_a/N_a
0% insurable	1.00	1.00	1.00	1.00	1.00
50% insurable	0.99	0.96	1.14	1.06	1.01
75% insurable	0.92	0.87	1.24	1.08	1.08
100% insurable	0.86	0.81	1.28	1.10	1.16

Table notes: Baseline equilibrium outcomes are normalized to one, for ease of comprehension.

Note that the results are not linear – if 75 percent of the shocks are insured, agricultural labor productivity increases by 8 percent, or about 50 percent of difference between the complete and incomplete markets baselines. This follows from the fact that incomplete insurance maintains risk of low consumption. Since shock realizations are weighted by marginal utility, households still put relatively substantial weight on these outcomes.

C Alternative Theories and Robustness

C.1 Credit Constraints

As Appendix A shows, the model developed in this paper generates the same type of cross-sectional distortion generated by models with explicit input market distortions. One important class of models along these lines are those that distort production through collateral constraints (Buera, Kaboski and Shin, 2011; Moll, 2014). I therefore test predictions of a collateral constraints model, and use the ICRISAT data to distinguish it from the baseline model in the text.

The collateral constraints model is identical to the baseline model in the main text, except that (1) all decisions are made after the realization of the shock z and (2) households are subject to a collateral constraint. Formally, the recursive problem of the household is

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

Note that now x is measurable with respect to z. That is, households make all decisions after the realization of the shock. However, they are subject to a collateral constraint in which they can only borrow a multiple λ of savings to purchase intermediates. The parameter $\lambda \in [1,\infty)$ governs the severity of the constraint.[7] I set $\lambda = 1.5$ to match a nominal intermediate share of 0.31, as the model in the main text predicts, though the choice of λ does not change the interpretation of the results presented here. The rest of the calibration is the same. I then compute the new stationary equilibrium, and as in the main text, create a sample of 500,000 individuals to draw on for statistics.

Tables 8 and 9 reproduce the same results from the main text, but now include the new results from the collateral constraints model. For clarity of exposition, I also include the relevant point estimates and standard errors from the model-derived relationships in the main text.

Table 8 first shows that the model can, at least qualitatively, replicate the relationship between intermediates, yields, and savings in the data. The collateral constraints model captures the same general trends as the risk model, as it relates to the importance of savings, though the small bootstrap samples limit the statistical significance of estimates. The qualitative trends, however, are intuitive. Like the risk model, the

[7] This constraint is another micro-foundation of the reduced form distortion $\phi(z,b)$ discussed in Appendix A.

wedge between the marginal revenue and price of intermediates declines with savings, and thus increases intermediate expenditures and farm yields.

Table 8: Savings and Intermediate Inputs (Collateral Constraints Model vs. Data)

	Expenditures	Expenditures	Nominal share	Nominal share	Yield	Yield
Savings	0.426	0.414	0.050	-0.018	0.327	0.382
	(0.086)***	(0.021)***	(0.006)***	(0.020)	(0.080)***	(0.019)***
Savings point estimate from risk model	0.426	–	0.069	–	0.301	–
	(0.016)***		(0.010)***		(0.024)***	
Obs	n.a.	2389	n.a.	2367	n.a.	2389
R^2	0.103	0.335	0.048	0.085	0.048	0.334
Model or Data?	Model	Data	Model	Data	Model	Data

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.

On the other hand, Table 9 shows that the collateral constraints model *cannot* match the positive relationship between the nominal intermediate share and the coefficient of variation of consumption expenditures. The collateral constraints model predicts a strong, negative relationship between the two variables.

Table 9: Consumption Volatility and Intermediate Inputs (Collateral Constraints Model vs. Data)

	Model (risk)	Model (credit)	Data (estimated)	Data (direct)
c.v. of consumption	0.027	-0.045	0.093	0.121
	(0.031)	(0.006)***	(0.048)*	(0.063)*
Obs	n.a.	n.a.	478	478
R^2	0.002	0.166	0.008	0.301

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. Dependent and independent variables are normalized by sample mean.

The intuition for the failure of the collateral constraints model along this dimension is that collateral constraints limit the ability of the poor to respond to shocks. For example, if a low-savings farmer receives an unexpectedly high shock, she cannot respond by adjusting intermediates x. Combined with her low savings, this means that her consumption volatility closely follows shock volatility over the relatively short time horizon considered. Farmers with high savings, on the other hand, can adjust x in response to z, as the collateral constraint is not binding. This provides an extra consumption smoothing device unavailable to the poorest farmers. Combined, these forces generate a negative relationship between the coefficient of variation and the nominal intermediate share. On the other hand, in the baseline risk model, households with high savings *endogenously* expose themselves to more variability in production. The critical distinction is that the risk model links distortions in the consumption market with production.

C.2 Different Shocks or Different Responses?

C.2.1 Construction of Country-Level Weather Shocks

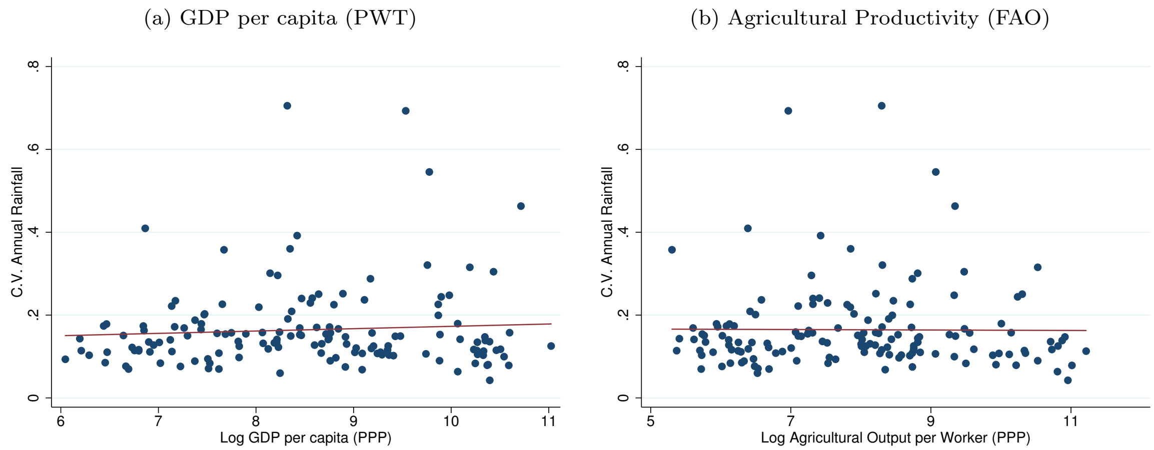
An alternative explanation to the one highlighted in this paper is that poor countries could simply face different exogenous agricultural shocks. I explore this alternative hypothesis using detailed information on historical rainfall fluctuations. Aggregating rainfall data to study cross-country variation present some difficulties however. For one, agriculture 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 4 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.

I find no evidence of a trend in either relationship, and Table 10 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 (2018) 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 rela-

Figure 4: Rainfall Variation Across Countries



tionship between risk and intermediate inputs, not variation in the exogenous shocks themselves.

Table 10: 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.2.2 Changes in Assumed Variance

I also investigate the importance of the shock distribution by varying the standard deviation of the underlying normal distribution σ_z in India, to study whether increasing this variance relative to the U.S. would generate a larger gap in productivity.

Since the mean of the lognormal distribution depends on σ_z, I hold fixed the mean by adjusting each shock value. That is, denoting $\mathbb{E}[z; \sigma_z]$ as the expected value of z

given the implied standard deviation, I consider the modified shock distribution

$$\widetilde{z} = \frac{\mathbb{E}[z; 0.32]}{\mathbb{E}[z; \sigma_z]} \times z, \qquad \forall\ z$$

Below, I provide results for two scenarios in which $\sigma_z^{India} = 0.64$, twice its baseline value. The first holds the remaining parameterization fixed at its baseline level, while the second recalibrates the full model under the assumption of a higher shock variance.

Holding the Remaining Parameterization Fixed Table 11 shows the results when the remaining calibration is held fixed, reproducing the cross-country results from the main text (the U.S. model economy maintains the baseline value $\sigma_z = 0.32$).[8]

Table 11: Quantitative Cross-Country Results with High σ_z in India

	Labor Productivity Gap (U.S./India)		p_xX/p_aY_a		N_a *(%)*	
Economy	Agriculture	GDP p.c.	U.S.	India	U.S.	India
Data	51.6	8.2	0.40	0.10	1.4	44.3
Baseline Model	35.9	8.2	0.40	0.31	1.4	44.3
India has $\sigma_z = 0.64$	27.7	6.8	0.40	0.32	1.4	34.1

Table 11 shows that higher standard deviations result in smaller productivity differences. This result is driven by two opposing forces in the model. First, all else equal, increasing the variance of shocks forces farmers to take more risk, which increases the distortion. One can see this in the "partial equilibrium" (PE) columns of Table 13, which fixes the price at the initial Indian equilibrium. As σ_z increases, the distortion – measured by decline in the nominal intermediate share – increases.

Table 12: Partial (PE) and General Equilibrium (GE) in India for Different σ_z

	Y_a/N_a		p_xX/p_aY_a		N_a *(%)*		p_a	
Economy	GE	PE	GE	PE	GE	PE	GE	PE
Baseline ($\sigma_z = 0.32$)	1.00	1.00	0.31	0.31	44.3	44.3	4.85	4.85
High Shocks ($\sigma_z = 0.64$)	1.30	1.00	0.32	0.26	34.1	108.7	3.74	4.85

Table notes: Partial equilibrium (PE) results assume price p_a is fixed at baseline equilibrium price. Agricultural productivity (Y_a/N_a) is normalized to one in the baseline equilibrium for ease of comprehension.

However, Table 13 also shows that price adjustments have a countervailing effect. This results from the interaction of two model features: the low utility weight on

[8] Since the lognormal mean is increasing in σ_z, not doing this adjustment only amplifies the effects in Table 11.

agricultural consumption, α, and the subsistence requirement $\bar{a}$. Because α is so low, total agricultural output needs to be roughly $\bar{a}$. When σ_z is low, the price p_a must be high enough 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 low price decreases the cost of subsistence and allows Indian farmers to take further risk. This increases the nominal intermediate share in the GE case, relative to the PE case. Overall, these two effects combine to generate the relatively small gap in Table 11.

Updating the Remaining Calibration I next update the remaining calibration, net of the standard deviation σ_z. The key change induced by $\sigma_z = 0.64$ is that the model overshoots the savings rate. With the higher shock variance, the model predicts that the average household holds 1.39 times harvest, compared to the baseline value of 0.77. Thus, a recalibration requires a larger cost of savings to bring the model back in line with its empirical moment. The updated results are in Table 13, and shows that the results are nearly identical to the baseline model. However, ti requires a value of $\delta = 0.39$, implying that the model requires massive costs to savings to jointly rationalize the high shock variance with the (relatively) low savings.

Table 13: Quantitative Cross-Country Results with High σ_z in India (Recalibrated)

	Labor Productivity Gap (U.S./India)		$p_x X/p_a Y_a$		N_a *(%)*	
Economy	Agriculture	GDP p.c.	U.S.	India	U.S.	India
Data	51.6	8.2	0.40	0.10	1.4	44.3
Baseline Model	35.9	8.2	0.40	0.31	1.4	44.3
India has $\sigma_z = 0.64$	36.0	8.2	0.40	0.28	1.3	44.3

C.3 Allowing Off-Farm Specialization

In this model, I allow individuals to choose their sector before the shock is realized. Income in both sectors is subject to i.i.d. shocks z_a and z_m drawn from identical truncated log-normal distributions.

Each period, households can choose which sector to enter before the shock is realized. Manufacturing workers work only in the manufacturing sector. Agricultural households can farm, but also work in the manufacturing sector after learning z_a. However, this comes at a cost τ to their labor productivity, which is a specific interpretation of the reduced form wedge highlighted in the main text.

Define $v(b) = \max\{v_a(b), v_m(b)\}$ as the *ex ante* value of entering the period with b savings. The value functions v_a and v_m are discussed in turn.

Value of Agriculture After choosing to farm, a household's problem similar to that in the main text. That is, the value of entering farming before the shock is realized is

$$v_a(b) = \max_{x \geq 0} \int_z v^p(x, b, z) dQ(z). \tag{C.1}$$

where

$$v^p(x, b, z) = \max_{c_a, c_m, n_a, b'} \alpha \log(c_a - \bar{a}) + (1 - \alpha) \log(c_m) + \beta v(b') \tag{C.2}$$

subject to constraint set

$$\begin{aligned} & p_a c_a + c_m + p_a b' = p_a z A x^{\psi} n_a^{\eta} - p_x x + (1 - \tau) A (1 - n_a) + p_a (1 - \delta) b \\ & b' \geq 0, \qquad c_a \geq \bar{a}, \qquad c_m \geq 0 \end{aligned}$$

Value of Manufacturing The value of working in the manufacturing sector, before learning shock z_m, is

$$v_m(b) = \int_{z_m} \left(\max_{c_a, c_m, b'} \alpha \log(c_a - \bar{a}) + (1 - \alpha) \log(c_m) + \beta v(b') \right) dG_m(z_m) \tag{C.3}$$

subject to constraint set

$$\begin{aligned} & p_a c_a + c_m + p_a b' = A z_m + p_a (1 - \delta) b \\ & b' \geq 0, \qquad c_a \geq \bar{a}, \qquad c_m \geq 0, \end{aligned}$$

which holds for each realizations of z_m.

Note that in both cases, households can re-optimize their sectoral choice each period.

Results I compare the model predictions relative to the model in the main text with no occupational choice, holding the remaining calibration fixed. In the updated selection model, I set $\sigma_m = 1.15 \times \sigma_z$, though this value is not important for the main result that the baseline model generates smaller gains from insurance. Table 14 presents the results for the Indian economies.

Table 14: Gains from Completing the Market in India with *Ex Ante* Labor

	%Δ *Labor Productivity*	
Economy	Agriculture	GDP per capita
Baseline Model	16.3	11.4
With Selection	18.6	11.9

Table notes: The results presented are given as the percentage increase in labor productivity moving from incomplete to complete markets. The first row is the baseline model presented in the main text. The second row presents the results from the new model with *ex ante* selection. Real GDP per capita is measured at the equilibrium price from the associated U.S. model economy.

Table 14 shows that allowing for sectoral choice decreases agricultural productivity relative to the baseline model and very slightly decreases the nominal intermediate share (i.e., increases the distortion). Overall, however, the models are quite similar in their predictions.

The rationale for this result is that because manufacturing is riskier, only the rich enter manufacturing (consistent with empirical evidence).[9] For a given distribution of savings, this selection effect removes the least distorted households from agriculture, and thus increases the average household-level distortion. This naturally lowers average agricultural productivity. This induces a second, general equilibrium effect. Because the equilibrium needs to now make up supply previously produced by these undistorted farmers, the price must rise. This further negatively affects remaining farmers, who become more risk averse. One countervailing force arises as well – farmers want to save to exit agriculture. This increases average savings, and thus tempers the negative impact of selection somewhat.

The qualitative ranking of the two models does not depend on the fact that manufacturing is riskier than agriculture, only that the rich select into the manufacturing sector.

[9] This is because agricultural households have two smoothing tools – savings and the *ex post* labor market. This second force does not exist for manufacturing workers.

C.3.1 *Ex Ante* Labor Demand

Another possibility is that farmers are required to lock in their labor demand before the shock realization. I modify the production function to now be

$$y_a(z) = zAq^{\theta}x^{\psi}n_a(z)^{\eta} \tag{C.4}$$

where q is now *ex ante* labor demand. I assume that both types of wages are paid the same price. Like intermediates, this *ex ante* labor can be paid out of *ex post* profits, so that the model isolates the shift of labor pre-shock realization. The remaining model is identical.

The baseline model is $\theta = 0$, with $\eta = 0.42$. Here, I ask whether shifting some of that weight to *ex ante* labor affects the results. To do so, I hold $\theta + \eta = 0.42$, and then calibrate θ to match the share of the wage bill (including valued family labor) in ICRISAT that is *ex ante* labor. While this dichotomy is naturally difficult to pin down exactly, I take *ex ante* labor to be any labor used in planting, sowing, pesticide or fertilizer application, ploughing, irrigating, weeding, or land preparation. This implies that 42 percent of wages are *ex ante* in 2014 ICRISAT. As I will show in the quantitative results below, this exact value is not critical, as the gains from completing the market are increasing in θ. Thus, the baseline model is a lower bound. The results are in Table 15.

Table 15: Gains from Completing the Market in India with *Ex Ante* Selection

		%Δ *Labor Productivity*	
Economy	Value of θ	Agriculture	GDP per capita
Baseline Model	0	16.3	11.4
Updated Calibration	0.51	20.1	11.5
Updated Calibration	0	10.3	2.5

Table notes: The results presented are given as the percentage increase in labor productivity moving from incomplete to complete markets. The first row is the baseline model presented in the main text. The second row presents the results from the new model with the production function given by (C.4), recalibrated to hit the same moments as the baseline version with $\theta = 0.51$. The third row holds fixed the new calibration, but sets $\theta = 0$. Real GDP per capita is measured at the equilibrium price of the associated U.S. model economy.

This table is most easily read by first comparing rows 2 and 3. These are two economies, both under the same updated calibration net of the θ parameter. The gains from completing the market nearly double when $\theta = 0.51$, compared to the baseline assumption of no *ex ante* labor ($\theta = 0$, row 3). The rationale for this result is that *ex ante* labor limits one possible smoothing technology for households – the ability to vary labor after the shock is realized. Thus, the gains from providing complete insurance are higher when farmers are required to choose labor *ex ante*.

Comparing rows 1 and 2 then compares two economies calibrated to hit the same moments. They differ only in the value of θ – the extent to which labor is required *ex ante*. Here, we again see that the baseline assumption is a conservative one. The difference between the two economies is smaller, however. That is, the gains are smaller in row 3 ($\theta = 0$, updated calibration) than row 1 ($\theta = 0$, baseline calibration). The rationale for this result is that the updated calibration requires a larger shock variance. This follows from farmers' inability to respond to shocks with labor, which mutes the observed variance of harvest realizations. As discussed in Appendix C.2, this actually makes farmers more productive in equilibrium, lowering the gains from insurance.

Overall, these results show that the baseline model takes a conservative stand on the gains from insurance by assuming that all labor is chosen *ex ante*.

C.4 The Role for Permanent Productivity Differences

In the text, the model used induces heterogeneity only through temporary shocks. Recent work in misallocation shows that permanent productivity differences can potentially mute the impact of misallocation. Buera and Shin (2011), Moll (2014), and Midrigan and Xu (2014) study this issue in relation to collateral constraint models. I consider that issue here, and show it to be relatively less important than in previous work focused on the manufacturing sector.

I study the persistence of household-level harvest, denoted $y_{i,v,t}$ for household i in village v at year t. I run a series of regressions using the raw ICRISAT data, which take the form

$$\log(y_{i,v,t}) = \alpha + \beta \log(y_{i,v,t-j}) + \delta_t + \gamma_v + \varepsilon_{i,v,t} \quad \text{for j = 1, 2, 3.} \tag{C.5}$$

The estimates $\widehat{\beta}$ compute the predicted persistence of harvest over time. The idea here follows Midrigan and Xu (2014), and uses the decay in $\widehat{\beta}$ over time to calibrate permanent versus temporary components of firm-level shocks. The results are in Table 16. I compare the empirical regressions with identical regressions drawn from the model.[10]

Table 16: Persistence in Harvests Over Time

	(1) $t-1$	(2) $t-2$	(3) $t-3$
Data	0.749	0.674	0.668
	(0.014)***	(0.017)***	(0.020)***
Model	0.472	0.365	0.293
Obs. (data)	2353	1884	1412
R^2 (data)	0.640	0.558	0.547

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

Intuitively, substantial differences in permanent productivity cause this autocorrelation to decay slowly. For example, Midrigan and Xu (2014) find autocorrelations of 0.90 and 0.85 after one and three years, respectively, among Korean manufacturing firms. They use this to infer quantitatively important differences in permanent firm-level productivity. Here, however, the empirical results suggest a much lower autocorrelation. After one year, the empirical autocorrelation is 0.749 and after three it is 0.668. This suggests a substantially smaller role for permanent productivity

[10] Naturally the model regressions do not include village fixed effects. Note also that the time fixed effects are irrelevant in the model regressions because the model is in a stationary equilibrium.

differences.

However, the baseline model under-predicts persistence. I therefore modify the model to allow for permanent productivity differences to test the robustness of the model. The production function takes the form

$$y_t = \exp^{(\theta + z_t)} A x_t^{\psi} n_{at}^{\eta}$$

where $z_t \sim N(0, \sigma_z)$ is the i.i.d. temporary shock, and $\theta \sim N(0, \sigma_\theta)$ is household-level permanent productivity. The rest of the model remains the same, except the value function is now a function of permanent productivity θ and savings b. Note that the baseline model is the special case in which $\sigma_\theta = 0$. I restrict $\sigma_\theta + \sigma_z = 0.32$ as in the main text, and then choose σ_θ to match the results in Table 16, while leaving the remaining parameters unchanged. In particular, I minimize the sum of squared errors

$$\sum_{j=1}^{3} (\widehat{\beta}_j^{model} - \widehat{\beta}_j^{ICRISAT})^2$$

where $j = 1, 2, 3$ is the j-period lagged coefficient in regression (C.5). Holding the rest of the calibration constant, this procedure implies $\sigma_z = 0.23$ and $\sigma_\theta = 0.32 - \sigma_z = 0.09$, and generates $(\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\beta}_3) = (0.769, 0.689, 0.635)$. That is, matching the autocorrelation decay in the data implies that 72 percent of the standard deviation of productivity is still driven by temporary shocks.

Table 17 presents the cross-country results in which both countries are subject to variation in permanent productivity. I consider two – the first is the calibrated $\sigma_\theta = 0.09$ required to match the ICRISAT data, and the second is $\sigma_\theta = 0.16$, which implies a one-period persistence of 91 percent similar to those found in rich countries (including Midrigan and Xu, 2014). As expected, larger permanent productivity differences reduces the aggregate and agricultural productivity gap. This is the analog of the self-financing results found in Buera and Shin (2011) and Moll (2014). Note, however, that the gap is small. The agricultural productivity gap declines by only 6 percent at the calibrated level.

The relatively small impact is driven by the convex relationship between productivity and the importance of permanent productivity. Figure 5 plots Indian agricultural productivity (Y_a/N_a) as permanent productivity becomes a larger share of total cross-sectional productivity variation. Because the data imply that production decays quickly, and thus a relatively small role for permanent productivity, the economic changes relative to the baseline model are small. Buera and Shin (2011) find a similar convex shape when considering how the impact of collateral constraints changes

Table 17: Changing Permanent Productivity

Economy	*Labor Productivity Gap* Agriculture	GDP p.c.	$p_x X/p_a Y_a$ Rich	Poor	N_a (%) Rich	Poor
% of std from permanent productivity						
Baseline: 0%	35.9	8.2	0.40	0.31	1.4	44.3
Calibrated: 28%	34.1	8.0	0.40	0.34	1.4	43.0
"Rich country" persistence: 50%	33.0	7.7	0.40	0.36	1.4	39.5

with more persistent shocks, though the convexity is even more stark there. Thus, permanent productivity differences play only a small quantitative role.[11]

Figure 5: Indian Agricultural Productivity

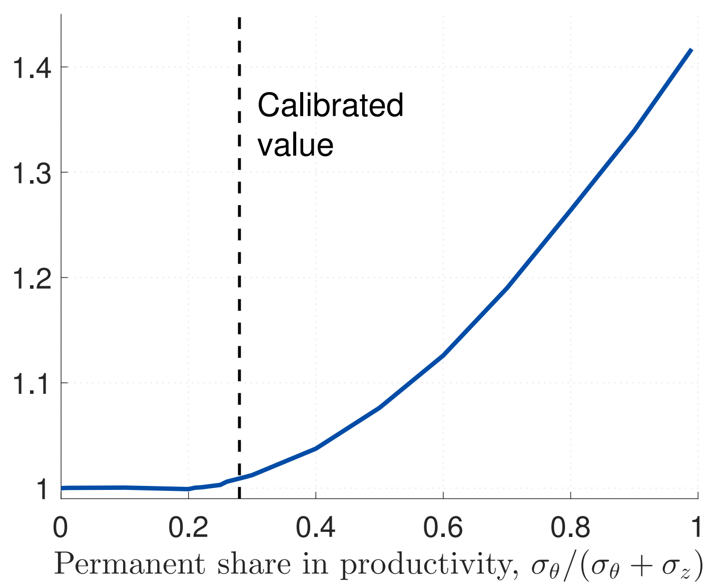


Figure notes: Baseline value $\sigma_\theta = 0$ has agricultural productivity normalized to one.

This difference in income persistence across sectors is interesting in its own right, and could potentially generate differences in sectoral productivity through the channels discussed here. This is left for future work.

[11] It is worth emphasizing that re-calibrating the model to match the same baseline targets with permanent productivity differences only magnifies the small quantitative role for the empirically relevant differences permanent productivity. The rationale is as follows. Larger variance in permanent productivity decrease agricultural employment (see Table 17), which in turn requires a higher $\overline{a}$ to match the Indian employment share. This increases relative risk aversion and thus increases misallocation, tempering the gains highlighted in Table 17.

C.5 Measurement Error

The critical use of harvest values in the model is to calibrate the variance of the shock z (σ_z^2). If harvests include noise, this overstates the calibrated shock variance. I recalibrate the model assuming that 20 percent of observed variation in harvests is measurement error, implying that the true variation is $0.8 \times 0.55 = 0.44$. I recalibrate the model with this updated moment, and find that $\sigma_z = 0.245$, or approximately 77 percent of the baseline value $\sigma_z = 0.32$. The gains from insurance in the Indian economy are in Table 18.

Table 18: Gains from Completing the Market in India with Mismeasurement

	Incomplete Markets Labor Productivity		%Δ *Labor Productivity*	
Economy	Agriculture	GDP per capita	Agriculture	GDP per capita
Baseline Model	1.00	1.00	16.3	11.4
With Mismeasurement	1.00	1.00	14.4	9.7

Table notes: The results presented are given as the percentage increase in labor productivity moving from incomplete to complete markets. The first row is the baseline model presented in the main text. The second row presents the results from the new model with mismeasurement. Real GDP per capita is measured at the equilibrium price from the associated U.S. model economy. Baseline model labor productivity is normalized to one in the incomplete markets economy.

The first two columns provide measures of labor productivity in the incomplete markets economies. The final two columns show the percentage gains from providing insurance. These columns show that the model with mis-measurement slightly lowers the gains from insurance from 16 percent to 14 percent.

Note, however, that the effect is not coming from the fact that the new model has higher labor productivity than the original. In fact, they have nearly identical labor productivity.[12] The similarity between the two results from two competing forces. First, for a given calibration, lower variance actually increases the gains from insurance. This is because of how the equilibrium price interacts with inelastic agricultural demand in this model. However, lower variance simultaneously lowers the incentives to save. Thus, savings falls. In response, the calibration requires a lower value of δ to hit the same moment. The calibrated value of δ declines from $\delta = 0.10$ to $\delta = 0.049$. This pushes productivity back up, as households are less distorted in their savings decision. Cumulatively, the two forces roughly offset.

The decline in the gains from insurance must therefore come from changes to the complete markets version of the economy. Indeed, this is the case. Lowering the shock variance requires the complete markets economy to use more labor to hit the same amount of production. Agricultural labor in the complete markets model increases by 3 percent in the model with mis-measurement. Combined with decreasing returns to

[12] This is not a theoretical result, they differ very slightly.

scale, this lowers labor productivity, albeit slightly.

Taken together, the gap between the incomplete and complete market economies falls with measurement error not due to its impact on the baseline economy, but because of the way it simultaneously affects the complete markets economy.

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, Yang and Zhu (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 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 as the nominally priced intermediate share. 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.

E Proofs

E.1 An Additional Lemma for the Proof of Proposition 2

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 2

Proof. I begin with proving the case of $\bar{a} = 0$. This follows quickly from Lemma 1. I further show that labor is independent of TFP, as an additional result to frame the Proposition.

E.2.1 When $\bar{a} = 0$

$n_a(z)$ is independent of A This follows directly from Lemma 1.

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

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

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.

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.2.2 When $\bar{a} > 0$

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. ■

References

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

Adamopoulos, Tasso, and Diego Restuccia. 2018. "Geography and Agricultural Productivity: Cross-Country Evidence from Micro Plot-Level Data." NBER Working Paper 24532.

Boppart, Timo. 2014. "Structural Change and the Kaldor Facts in a Growth Model with Relative Price Effects and Non-Gorman Preferences." *Econometrica*, 82(6): 2167–2196.

Buera, Francisco J., and Yongseok Shin. 2011. "Self-insurance vs. self-financing: A welfare analysis of the persistence of shocks." *Journal of Economic Theory*, 146(3): 845–862.

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

Caselli, Francesco. 2005. "Accounting for Cross-Country Income Differences." In *Handbook of Economic Growth.* , ed. P. Aghion and S.N. Durlauf, 679–741.

Caunedo, Julieta, and Elisa Keller. forthcoming. "Capital Obsolescence and Agricultural Productivity." *Quarterly Journal of Economics.*

Chen, Chaoran. 2020. "Technology Adoption, Capital Deepening, and International Productivity Differences." *Journal of Development Economics*, 143(1).

Comin, Diego, Danial Lashkari, and Martí Mestieri. forthcoming. "Structural Change with Long-run Income and Price Effects." *Econometrica*.

FAOSTAT. 2018. "Food and Agriculture Organization Statistics." FAO.

Goldberg, Jessica. 2017. "The Effect of Social Pressure on Expenditures in Malawi." *Journal of Economic Behavior and Organization*, 143(1): 173–185.

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

Gollin, Douglas, Stephen L. Parente, and Richard Rogerson. 2004. "Farm Work, Home Work and International Productivity Differences." *Review of Economic Dynamics*, 7(4): 827–850.

Jakiela, Pamela, and Owen Ozier. 2016. "Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economics." *Review of Economic Studies*, 83(1): 231–268.

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

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

Ngai, L. Rachel, and Christopher A. Pissarides. 2007. "Structural Change in a Multisector Model of Growth." *American Economic Review*, 97(1): 429–443.

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

Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu. 2008. "Agriculture and aggregate productivity: A quantitative cross-country analysis." *Journal of Monetary Economics*, 55(2): 234–250.

Timmer, Marcel P., Erik Dietzenbacher, Bart Los, Robert Stehrer, and Gaaitzen J. de Vries. 2015. "An Illustrated User Guide to the World Input-Output Database: the Case of Global Automotive Production." *Review of International Economics*, 23(3): 575–605.

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

Xie, Danyang. 2000. "Power Risk Aversion Utility Functions." *Annals of Economics and Finance*, 265(1): 265–282.

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