Volatility and the Gains from Trade: Online Appendix

Treb Allen and David Atkin

This Online Appendix contains detailed derivations, proofs of propositions and extensions for the model introduced in Section 4 of the main text. Additional information regarding the empirical context of the data in Section 2, robustness related to stylized facts in Section 3, the derivation and estimation of the quantitative model in Section 5 and additional tables and figures mentioned in the main text are contained in the Supplemental Materials posted on both the authors' websites and in the replication files on the journal website.

A.1 Model derivations

In this subsection, we present the derivations of several results in the main paper.

Approximation of real returns (Equation (15)) First, to calculate the income of farmers in a village, we substitute the arbitrage equation (10) into farmers' income as a sum of revenue across all crops to yield:

$$Y_i(s) = \left(\sum_{g \in \mathcal{G}} \alpha_{ig} \left(\frac{\bar{p}_g(s) Q_{ig}(s)}{\alpha_{ig}} \right)^{\frac{\varepsilon_i}{1+\varepsilon_i}} \right)^{\frac{1+\varepsilon_i}{\varepsilon_i}}. \tag{24}$$

Similarly, combining equation (5) with (10) yields the following expression for the period welfare of a farmer in village i:

$$Z_i^f(s) = \frac{1}{L_i} \times \left(\sum_{g \in \mathcal{G}} \alpha_{ig} \left(\frac{\bar{p}_g(s) Q_{ig}(s)}{\alpha_{ig}} \right)^{\frac{\varepsilon_i}{1+\varepsilon_i}} \right) \times \prod_{g \in \mathcal{G}} \left(Q_{ig}(s) \left(\frac{\alpha_{ig}}{\bar{p}_g(s) Q_{ig}(s)} \right)^{\frac{\varepsilon_i}{1+\varepsilon_i}} \right)^{\alpha_{ig}} \tag{25}$$

In the autarky (i.e. $\varepsilon_i = 0$), equation (25) simplifies to $Z_i^{f,aut}(s) \equiv \frac{1}{L_i} \prod_{g \in \mathcal{G}} (Q_{ig}(s))^{\alpha_{ig}}$, as farmers consume what they produce. In free trade (i.e. $\varepsilon_i \to \infty$), equation (25) simplifies to $Z_i^{f,free}(s) \equiv \frac{1}{L_i} \times \left(\sum_{g \in \mathcal{G}} \bar{p}_g(s) Q_{ig}(s) \right) \times \prod_{g \in \mathcal{G}} \left(\frac{\alpha_{ig}}{\bar{p}_g(s)} \right)^{\alpha_{ig}}$,as farmers sell what they produce and purchase what they consume at the central market prices.

We now note that with a large number of villages and idiosyncratic shocks that $\bar{p}_g(s) = \bar{p}_g$, i.e. the central market prices is state invariant. Taking logs of equation (25) then yields:

$$\begin{aligned} \ln Z_i^f(s) =& \ln \left(\sum_{g \in \mathcal{G}} \theta_{ig}^f \times \frac{\alpha_{ig}}{\theta_{ig}} \left(\frac{\bar{p}_g(s) \theta_{ig}}{\alpha_{ig}} A_{ig}(s) \right)^{\frac{\varepsilon_i}{1+\varepsilon_i}} \right) \\ &+ \left(\frac{1}{1+\varepsilon_i} \right) \sum_{g \in \mathcal{G}} \alpha_{ig} \ln A_{ig}(s) + \sum_{g \in \mathcal{G}} \alpha_{ig} \left(\ln \left(\alpha_{ig} \left(\frac{\bar{p}_g \theta_{ig}}{\alpha_{ig}} \right)^{\frac{1}{1+\varepsilon_i}} \right) - \ln \bar{p}_g \right) \end{aligned} \tag{26}$$

We then apply the following second-order approximation implying that the sum of log normal variables is itself approximately log normal (see, e.g. Campbell and Viceira (2002)). Suppose that $\ln \boldsymbol{x}_i(s) \sim N(\boldsymbol{\mu}_i^x, \boldsymbol{\Sigma}_i)$

and $X_i(s) \equiv \ln\left(\sum_{g\in\mathcal{G}} w_{i,g} x_{i,g}(s)\right)$ for some weights $\sum_{g\in\mathcal{G}} w_{i,g} = 1$. Then a second order approximation around the mean log returns is:

$$X_i(s) \approx \ln\left(\sum_{g\in\mathcal{G}} w_{i,g} \exp\left(\mu^x_{i,g}\right)\right) + \sum w_{i,g}(\ln x_{i,g}(s) - \mu^x_{i,g}) - \frac{1}{2}\sum_{h\in\mathcal{G}}\sum_{g\in\mathcal{G}} w_{i,g} w_{i,h} \sigma^x_{i,gh} + \frac{1}{2}\sum_{g\in\mathcal{G}} w_{i,g}\sigma^x_{i,gg}. \tag{27}$$

In our case, we have:

$$\ln x_{ig}(s) \equiv \ln\left(\frac{\alpha_{ig}}{\theta_{ig}}\right) + \frac{\varepsilon_i}{1+\varepsilon_i}\ln\left(\frac{\bar{p}_g \theta_{ig}}{\alpha_{ig}}\right) + \frac{\varepsilon_i}{1+\varepsilon_i}\ln(A_{ig}(s))$$

and $w_{i,g} \equiv \theta^f_{i,g}$ which implies that $\mu^x_{i,g} = \ln\left(\frac{\alpha_{ig}}{\theta_{ig}}\right) + \frac{\varepsilon_i}{1+\varepsilon_i}\ln\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}\right) + \frac{\varepsilon_i}{1+\varepsilon_i}\mu^{A,i}_g$ and $\sigma^x_{i,gh} = \left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)^2 \sigma^A_{i,gh}$. Applying the approximation (27) to the real returns (26) results in:

$$\ln Z^f_i(s) \approx \mu^Z_i + \sum_{g\in\mathcal{G}}\left(\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\theta^f_{i,g} + \left(\frac{1}{1+\varepsilon_i}\right)\alpha_{ig}\right)\left(\ln A_{ig}(s) - \mu^{A,i}_g\right),$$

where

$$\begin{aligned}\mu^Z_i \equiv & \sum_{g\in\mathcal{G}}\left(\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\theta^f_{i,g} + \left(\frac{1}{1+\varepsilon_i}\right)\alpha_{ig}\right)\mu^{A,i}_g + \ln\left(\sum_{g\in\mathcal{G}}\theta^f_{ig}\times\frac{\alpha_{ig}}{\theta_{ig}}\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}\exp\left(\mu^{A,i}_g\right)\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}\right)\\ & -\frac{\varepsilon_i}{1+\varepsilon_i}\sum_{g\in\mathcal{G}}\theta^f_{i,g}\mu^{A,i}_g + \sum_{g\in\mathcal{G}}\alpha_{ig}\left(\ln\left(\alpha_{ig}\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}\right)^{\frac{1}{1+\varepsilon_i}}\right) - \ln\bar{p}_g\right)\\ & + \frac{1}{2}\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)^2\left(\sum_{g\in\mathcal{G}}\theta^f_{i,g}\Sigma^{A,i}_{gg} - \sum_{h\in\mathcal{G}}\sum_{g\in\mathcal{G}}\theta^f_{i,g}\theta^f_{i,h}\Sigma^{A,i}_{gh}\right),\end{aligned} \tag{28}$$

as required.

It immediately follows that farmer utility is (approximately) log normally distributed across states of the world:

$$\ln Z^f_i \sim N\left(\mu^Z_i, \sigma^{2,Z}_i\right),$$

where

$$\sigma^{2,Z}_i \equiv \sum_{g\in\mathcal{G}}\sum_{h\in\mathcal{G}}\left(\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\theta^f_{i,g} + \left(\frac{1}{1+\varepsilon_i}\right)\alpha_{ig}\right)\left(\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\theta^f_{i,h} + \left(\frac{1}{1+\varepsilon_i}\right)\alpha_{ih}\right)\Sigma^{A,i}_{gh}. \tag{29}$$

Optimal crop choice first order conditions (equation (17)) Beginning with the maximization problem:

$$\max_{\{\theta^f_{ig}\}} \mu^Z_i + \frac{1}{2}(1-\rho_i)\sigma^{2,Z}_i \text{ s.t. } \Sigma_{g\in\mathcal{G}}\theta^f_{ig} = 1$$

and substituting in the expressions for μ_i^Z and $\sigma_i^{2,Z}$ from equations (28) and (29) results in:

$$\max_{\{\theta_{ig}^f\}} \ln\left(\sum_{g\in\mathcal{G}}\theta_{ig}^f\frac{\alpha_{ig}}{\theta_{ig}}\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}\exp\left(\mu_g^{A,i}\right)\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}\right)+\left(\frac{1}{1+\varepsilon_i}\right)\sum_{g\in\mathcal{G}}\alpha_{ig}\mu_g^{A,i}+\sum_{g\in\mathcal{G}}\alpha_{ig}\left(\ln\left(\alpha_{ig}\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}\exp\left(\mu_g^{A,i}\right)\right)^{\frac{1}{1+\varepsilon_i}}\right)-\ln\bar{p}_g\right)$$
$$+\frac{1}{2}\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)^2\left(\sum_{g\in\mathcal{G}}\theta_{i,g}^f\Sigma_{gg}^{A,i}-\sum_{h\in\mathcal{G}}\sum_{g\in\mathcal{G}}\theta_{i,g}^f\theta_{i,h}^f\Sigma_{gh}^{A,i}\right)$$
$$+\frac{1}{2}(1-\rho_i)\sum_{g\in\mathcal{G}}\sum_{h\in\mathcal{G}}\left(\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\theta_{i,g}^f+\left(\frac{1}{1+\varepsilon_i}\right)\alpha_{ig}\right)\left(\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\theta_{i,h}^f+\left(\frac{1}{1+\varepsilon_i}\right)\alpha_{ih}\right)\Sigma_{gh}^{A,i}$$

subject to:

$$\Sigma_{g\in\mathcal{G}}\theta_{ig}^f=1.$$

Taking the first order conditions with respect to θ_{ig}^f (note that each farmer makes her crop choice taking the crop choice of other farmers as given) results in the following first order conditions:

$$\frac{\frac{\alpha_{ig}}{\theta_{ig}}\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}\exp\left(\mu_g^{A,i}\right)\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_{g\in\mathcal{G}}\theta_{ig}^f\times\frac{\alpha_{ig}}{\theta_{ig}}\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}\exp\left(\mu_g^{A,i}\right)\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}+\frac{1}{2}\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)^2\Sigma_{gg}^{A,i}+\frac{\varepsilon_i}{(1+\varepsilon_i)^2}\sum_{h\in\mathcal{G}}\alpha_{ih}\Sigma_{gh}^{A,i}$$
$$-\rho_i\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\sum_{h\in\mathcal{G}}\left(\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\theta_{i,h}^f+\left(\frac{1}{1+\varepsilon_i}\right)\alpha_{ih}\right)\Sigma_{gh}^{A,i}=\lambda_i$$

or equivalently:

$$\mu_{ig}^Z-\rho_i\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\sum_{h\in\mathcal{G}}\left(\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\theta_{i,h}^f+\left(\frac{1}{1+\varepsilon_i}\right)\alpha_{ih}\right)\Sigma_{gh}^{A,i}=\lambda_i,$$

where $\mu_{ig}^Z\equiv\frac{1}{\theta_{ig}}\frac{\alpha_{ig}\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}\exp\left(\mu_g^{A,i}\right)\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_{g\in\mathcal{G}}\alpha_{ig}\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}\exp\left(\mu_g^{A,i}\right)\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}+\frac{1}{2}\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)^2\Sigma_{gg}^{A,i}+\frac{\varepsilon_i}{(1+\varepsilon_i)^2}\sum_{h\in\mathcal{G}}\alpha_{ih}\Sigma_{gh}^{A,i}$, as required.

Equilibrium crop choice (equation (18)) We re-write the first order conditions as:

$$\frac{\frac{\alpha_{ig}}{\theta_{ig}}\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}\exp\left(\mu_g^{A,i}\right)\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_{g\in\mathcal{G}}\theta_{ig}^f\times\frac{\alpha_{ig}}{\theta_{ig}}\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}\exp\left(\mu_g^{A,i}\right)\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}=\lambda_i-\left(\frac{1}{2}\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)^2\Sigma_{gg}^{A,i}+\frac{\varepsilon_i}{(1+\varepsilon_i)^2}\sum_{h\in\mathcal{G}}\alpha_{ih}\Sigma_{gh}^{A,i}\right.$$
$$\left.-\rho_i\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\sum_{h\in\mathcal{G}}\left(\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\theta_{i,h}+\left(\frac{1}{1+\varepsilon_i}\right)\alpha_{ih}\right)\Sigma_{gh}^{A,i}\right)\Longleftrightarrow$$
$$\theta_{ig}\propto\alpha_{ig}(\bar{p}_gB_{ig})^{\varepsilon_i}\Longrightarrow$$
$$\theta_{ig}=\frac{\alpha_{ig}(\bar{p}_gB_{ig})^{\varepsilon_i}}{\sum_{h\in\mathcal{G}}\alpha_{ih}(\bar{p}_hB_{ih})^{\varepsilon_i}},$$

where $B_{ig}\equiv\frac{\exp\mu_g^{A,i}}{\left(\lambda_i-\left(\frac{1}{2}\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)^2\Sigma_{gg}^{A,i}+\frac{\varepsilon_i}{(1+\varepsilon_i)^2}\sum_{h\in\mathcal{G}}\alpha_{ih}\Sigma_{gh}^{A,i}-\rho_i\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\sum_{h\in\mathcal{G}}\left(\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)\theta_{i,h}+\left(\frac{1}{1+\varepsilon_i}\right)\alpha_{ih}\right)\Sigma_{gh}^{A,i}\right)\right)^{\frac{1+\varepsilon_i}{\varepsilon_i}}}$, as required.

A.2 Proofs

This subsection contains the proofs of Propositions 1 and 2.

A.2.1 Proof of Proposition 1

We first restate the proposition:

Proposition. *Given any set of preferences* $\{\alpha_{ig}\}_{g\in\mathcal{G}}$*, trade costs* $\{\varepsilon_i\}_{i\in\mathcal{N}}$*, and any state of the world* $s\in S$ *such that quantity produced is* $\{Q_{ig}(s)\}_{i\in\mathcal{N}}^{g\in\mathcal{G}}$*:*

(a) There exists a state equilibrium.

(b) If the trade costs $\{\varepsilon_i\}_{i\in\mathcal{N}}$ are sufficiently close to 1, then that equilibrium is unique.

Proof of part (a) (existence)

Proof. In what follows, we omit dependence of prices $p_{ig}(s)$ and quantities $Q_{ig}(s)$ on state s for clarity. To prove existence, we first show that it is sufficient to focus on the excess demand function of the central market. We then show that the central market excess demand function satisfies all conditions necessary to guarantee existence from Proposition 17.C.1 of Mas-Colell et al. (1995).

We first note that given quantities $\{Q_{ig}\}_{i\in\mathcal{N}}^{g\in\mathcal{G}}$ and the equilibrium central market prices $\{\bar{p}_g\}_{g\in\mathcal{G}}$, village level incomes $\{Y_i\}_{i\in\mathcal{N}}$ are given immediately from equation prices (24); in turn, given village incomes $\{Y_i\}_{i\in\mathcal{N}}$, village level prices $\{p_{ig}\}_{i\in\mathcal{N}}^{g\in\mathcal{G}}$ are then given immediately from equation (10); and finally, given village level prices $\{p_{ig}\}_{i\in\mathcal{N}}^{g\in\mathcal{G}}$, village level consumption $\{C_{ig}\}_{i\in\mathcal{N}}^{g\in\mathcal{G}}$ are given immediately from equation (9). That is, given quantities $\{Q_{ig}\}_{i\in\mathcal{N}}^{g\in\mathcal{G}}$ and the equilibrium central market prices $\{\bar{p}_g\}_{g\in\mathcal{G}}$, it is straightforward to find a set of village prices $\{p_{ig}(s)\}_{i\in\mathcal{N}}^{g\in\mathcal{G}}$ and village consumption $\{C_{ig}(s)\}_{i\in\mathcal{N}}^{g\in\mathcal{G}}$ such that markets clear within each village (and condition 1 of the state equilibrium is satisfied). Hence, all that remains to determine the full state equilibrium is the set of equilibrium central market prices $\{\bar{p}_g\}_{g\in\mathcal{G}}$ such that the central market clears.

To find the equilibrium central market prices, we consider the following central market excess demand function $Z\equiv\{Z_g\}_{g\in\mathcal{G}}:\mathbb{R}^G\to\mathbb{R}^G$:

$$Z_g\left(\{\bar{p}_g\}_{g\in\mathcal{G}}\right):\frac{\bar{\alpha}_g\sum_h\sum_i\bar{p}_h\left(1-\left(\frac{\bar{p}_h}{p_{ih}}\right)^{-1}\right)\left(1-\left(\frac{\bar{p}_h}{p_{ih}}\right)^{-\varepsilon_i}\right)Q_{ih}}{\bar{p}_g}-\sum_i\left(1-\left(\frac{\bar{p}_g}{p_{ig}}\right)^{-\varepsilon_i}\right)Q_{ig}\iff$$

$$Z_g\left(\{\bar{p}_g\}_{g\in\mathcal{G}}\right):\frac{\bar{\alpha}_g\sum_h\sum_i\bar{p}_h\left(1-\left(\alpha_{ih}\left(\frac{\alpha_{ih}^{\frac{1}{1+\varepsilon_i}}Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}}\bar{p}_h^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_l\alpha_{il}^{\frac{1}{1+\varepsilon_i}}Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}}\bar{p}_l^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)^{\frac{1}{\varepsilon_i}}\right)\left(1-\alpha_{ih}\left(\frac{\alpha_{ih}^{\frac{1}{1+\varepsilon_i}}Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}}\bar{p}_h^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_l\alpha_{il}^{\frac{1}{1+\varepsilon_i}}Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}}\bar{p}_l^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)Q_{ih}}{\bar{p}_g}$$
$$-\sum_i\left(1-\left(\alpha_{ig}\left(\frac{\alpha_{ig}^{\frac{1}{1+\varepsilon_i}}Q_{ig}^{\frac{\varepsilon_i}{1+\varepsilon_i}}\bar{p}_g^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_h\alpha_{ih}^{\frac{1}{1+\varepsilon_i}}Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}}\bar{p}_h^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)\right)Q_{ig},\tag{30}$$

where the first term of Z_g is the quantity of good g demanded by the central market at price vector $\{\bar{p}_g\}_{g\in\mathcal{G}}$ (see equation (12)) and the second term is the quantity of good g supplied to the central market at price vector $\{\bar{p}_g\}_{g\in\mathcal{G}}$ (see equation (11)) and the second line uses equations (10) and (24) to substitute out for village level prices.

We now verify that the excess demand function defined by (30) satisfies conditions (i) to (v) of Proposition 17.B.2 of Mas-Colell et al. (1995), which from Proposition 17.C.1 of Mas-Colell et al. (1995) guarantees the existence of a set of central market prices $\{\bar{p}_g(s)\}_{g\in\mathcal{G}}$ and central market consumption $\{\bar{C}_g(s)\}_{g\in\mathcal{G}}$ that clear the central market (i.e. satisfy condition 2 of the state equilibrium).

Condition (i): Continuity. This is self evident from equation (30).

Condition (ii): Homogeneity of degree zero in prices. For any $C>0$, we have:

$$
\begin{aligned}
Z_g(\{C\bar{p}_g\}) &= \frac{\bar{\alpha}_g \sum_h \sum_i \bar{p}_h \left(1-\left(\alpha_{ih}\left(\frac{\alpha_{ih}^{\frac{1}{1+\varepsilon_i}} Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} (C\bar{p}_h)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_l \alpha_{il}^{\frac{1}{1+\varepsilon_i}} Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}} (C\bar{p}_l)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)^{\frac{1}{\varepsilon_i}}\right)\left(1-\alpha_{ih}\left(\frac{\alpha_{ih}^{\frac{1}{1+\varepsilon_i}} Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} (C\bar{p}_h)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_l \alpha_{il}^{\frac{1}{1+\varepsilon_i}} Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}} (C\bar{p}_l)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right) Q_{ih}}{\bar{p}_g} \\
&\quad -\sum_i \left(1-\left(\alpha_{ig}\left(\frac{\alpha_{ig}^{\frac{1}{1+\varepsilon_i}} Q_{ig}^{\frac{\varepsilon_i}{1+\varepsilon_i}} (C\bar{p}_g)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_h \alpha_{ih}^{\frac{1}{1+\varepsilon_i}} Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} (C\bar{p}_h)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)\right) Q_{ig} \\
&= \frac{\bar{\alpha}_g \sum_h \sum_i \bar{p}_h \left(1-\left(\alpha_{ih}\left(\frac{\alpha_{ih}^{\frac{1}{1+\varepsilon_i}} Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_h^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_l \alpha_{il}^{\frac{1}{1+\varepsilon_i}} Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_l^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)^{\frac{1}{\varepsilon_i}}\right)\left(1-\alpha_{ih}\left(\frac{\alpha_{ih}^{\frac{1}{1+\varepsilon_i}} Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_h^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_l \alpha_{il}^{\frac{1}{1+\varepsilon_i}} Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_l^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right) Q_{ih}}{\bar{p}_g} \\
&\quad -\sum_i \left(1-\left(\alpha_{ig}\left(\frac{\alpha_{ig}^{\frac{1}{1+\varepsilon_i}} Q_{ig}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_g^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_h \alpha_{ih}^{\frac{1}{1+\varepsilon_i}} Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_h^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)\right) Q_{ig} \\
&= Z_g(\{\bar{p}_g\}),
\end{aligned}
$$

as required.

Condition (iii): Walras' law. We have:

$$
\begin{aligned}
\sum_g \bar{p}_g Z_g &= \sum_g \bar{\alpha}_g \sum_h \sum_i \bar{p}_h \left(1-\left(\alpha_{ih}\left(\frac{\alpha_{ih}^{\frac{1}{1+\varepsilon_i}} Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_h^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_l \alpha_{il}^{\frac{1}{1+\varepsilon_i}} Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_l^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)^{\frac{1}{\varepsilon_i}}\right)\left(1-\alpha_{ih}\left(\frac{\alpha_{ih}^{\frac{1}{1+\varepsilon_i}} Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_h^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_l \alpha_{il}^{\frac{1}{1+\varepsilon_i}} Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_l^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right) Q_{ih} \\
&\quad -\sum_g \bar{p}_g \sum_i \left(1-\left(\alpha_{ig}\left(\frac{\alpha_{ig}^{\frac{1}{1+\varepsilon_i}} Q_{ig}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_g^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_h \alpha_{ih}^{\frac{1}{1+\varepsilon_i}} Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_h^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)\right) Q_{ig} \\
&= -\sum_h \sum_i \left(\frac{\left(\sum_l \alpha_{il}^{\frac{1}{1+\varepsilon_i}} Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_l^{\frac{\varepsilon_i}{1+\varepsilon_i}}\right)^{\frac{1}{\varepsilon_i}}}{\alpha_{ih}^{-\frac{1}{1+\varepsilon_i}} Q_{ih}^{-\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_h^{-\frac{\varepsilon_i}{1+\varepsilon_i}}}\right) + \sum_h \sum_i \left(\frac{\left(\sum_l \alpha_{il}^{\frac{1}{1+\varepsilon_i}} Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_h^{\frac{\varepsilon_i}{1+\varepsilon}}\right)^{\frac{1+\varepsilon_i}{\varepsilon_i}}}{\alpha_{ih}^{-1}}\right) \\
&= \left(\sum_l \alpha_{il}^{\frac{1}{1+\varepsilon_i}} Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_l^{\frac{\varepsilon_i}{1+\varepsilon_i}}\right)^{\frac{1}{\varepsilon_i}} \left[-\sum_h \sum_i \alpha_{ih}^{\frac{1}{1+\varepsilon_i}} Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_h^{\frac{\varepsilon_i}{1+\varepsilon_i}} + \sum_h \sum_i \alpha_{ih} \sum_l \alpha_{il}^{\frac{1}{1+\varepsilon_i}} Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_l^{\frac{\varepsilon_i}{1+\varepsilon_i}}\right] \\
&= \left(\sum_l \alpha_{il}^{\frac{1}{1+\varepsilon_i}} Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_l^{\frac{\varepsilon_i}{1+\varepsilon_i}}\right)^{\frac{1}{\varepsilon_i}} \left[-\sum_h \sum_i \alpha_{ih}^{\frac{1}{1+\varepsilon_i}} Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_h^{\frac{\varepsilon_i}{1+\varepsilon_i}} + \sum_i \sum_l \alpha_{il}^{\frac{1}{1+\varepsilon_i}} Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \bar{p}_l^{\frac{\varepsilon_i}{1+\varepsilon_i}}\right] \\
&= 0,
\end{aligned}
$$

as required.

Condition (iv): Bounded below. In particular, we need that there is an $s>0$ such that $Z_g(p)>-s$ for all p and all goods g. This is straightforward as the first sum must be nonnegative. To see this, note that in each term we have something of the form$\bar{p}_h(1-x)\left(1-x^{\frac{1}{\varepsilon_i}}\right)$ with $x>0$. If $x>1$, both $1-x$ and $1-x^{\frac{1}{\varepsilon_i}}$ are negative and the term is positive. Similarly, if $x<1$, both terms are positive. If $x=0$, It is zero. For the second sum, we have something of the form $1-x$ for each term with $x>0$. Therefore,

$$Z_g(\{\bar{p}_g\})=\frac{\bar{\alpha}_g\sum_h\sum_i\bar{p}_h\left(1-\left(\alpha_{ih}\left(\frac{\alpha_{ih}^{\frac{1}{1+\varepsilon_i}}Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}}\bar{p}_h^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_l\alpha_{il}^{\frac{1}{1+\varepsilon_i}}Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}}\bar{p}_l^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)^{\frac{1}{\varepsilon_i}}\right)\left(1-\alpha_{ih}\left(\frac{\alpha_{ih}^{\frac{1}{1+\varepsilon_i}}Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}}\bar{p}_h^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_l\alpha_{il}^{\frac{1}{1+\varepsilon_i}}Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}}\bar{p}_l^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)Q_{ih}}{\bar{p}_g}$$
$$-\sum_i\left(1-\left(\alpha_{ig}\left(\frac{\alpha_{ig}^{\frac{1}{1+\varepsilon_i}}Q_{ig}^{\frac{\varepsilon_i}{1+\varepsilon_i}}\bar{p}_g^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_h\alpha_{ih}^{\frac{1}{1+\varepsilon_i}}Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}}\bar{p}_h^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)\right)Q_{ig}$$
$$\geq -\sum_i Q_{ig}$$

Then we can take $s=\max_g\sum_i Q_{ig}$, and $Z_g(\{\bar{p}_g\})\geq -s$ for all g and $\{\bar{p}_g\}$.

Condition (v): Limiting behavior as prices go to zero. Condition (v) requires that if $p^n\to p$, where $p\neq 0$ and $p_g=0$ for some g, then $\max_g\lim_{n\to\infty}Z_g(p^n)\to\infty$. To see this, choose g such that $\lim_n\frac{p_g^n}{p_h^n}<\infty$ for all h; intuitively, p_g^n goes to 0 as fast or faster than any other price p_h^n. Since $p\neq 0$, there must be an h' such that $\lim_n\frac{p_g^n}{p_{h'}^n}=0$. We have that

$$Z_g(p^n)=\frac{\bar{\alpha}_g\sum_h\sum_i p_h^n\left(1-\left(\alpha_{ih}\left(\frac{\alpha_{ih}^{\frac{1}{1+\varepsilon_i}}Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}}(p_h^n)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_l\alpha_{il}^{\frac{1}{1+\varepsilon_i}}Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}}(p_l^n)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)^{\frac{1}{\varepsilon_i}}\right)\left(1-\alpha_{ih}\left(\frac{\alpha_{ih}^{\frac{1}{1+\varepsilon_i}}Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}}(p_h^n)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_l\alpha_{il}^{\frac{1}{1+\varepsilon_i}}Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}}(p_l^n)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)Q_{ih}}{p_g^n}$$
$$-\sum_i\left(1-\left(\alpha_{ig}\left(\frac{\alpha_{ig}^{\frac{1}{1+\varepsilon_i}}Q_{ig}^{\frac{\varepsilon_i}{1+\varepsilon_i}}(p_g^n)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_h\alpha_{ih}^{\frac{1}{1+\varepsilon_i}}Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}}(p_h^n)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)\right)Q_{ig}$$
$$=\bar{\alpha}_g\sum_h\sum_i\frac{p_h^n}{p_g^n}\left(1-\left(\alpha_{ih}\left(\frac{\alpha_{ih}^{\frac{1}{1+\varepsilon_i}}Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_l\alpha_{il}^{\frac{1}{1+\varepsilon_i}}Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}}\left(\frac{p_l^n}{p_h^n}\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)^{\frac{1}{\varepsilon_i}}\right)\left(1-\alpha_{ih}\left(\frac{\alpha_{ih}^{\frac{1}{1+\varepsilon_i}}Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_l\alpha_{il}^{\frac{1}{1+\varepsilon_i}}Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}}\left(\frac{p_l^n}{p_h^n}\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)Q_{ih}$$
$$-\sum_i\left(1-\left(\alpha_{ig}\left(\frac{\alpha_{ig}^{\frac{1}{1+\varepsilon_i}}Q_{ig}^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_h\alpha_{ih}^{\frac{1}{1+\varepsilon_i}}Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}}\left(\frac{p_h^n}{p_g^n}\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{-1}\right)\right)Q_{ig}$$

This goes to ∞ as $n\to\infty$. To see this, consider the h such that $\lim_n\frac{p_h^n}{p_g^n}=\infty$. Then to guarantee $Z_g(p^n)\to\infty$, we simply need that

$$\alpha_{ih}\frac{\sum_l\alpha_{il}^{\frac{1}{1+\varepsilon_i}}Q_{il}^{\frac{\varepsilon_i}{1+\varepsilon_i}}\left(\frac{p_l}{p_h}\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\alpha_{ih}^{\frac{1}{1+\varepsilon_i}}Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}}}$$

does not equal 1 for one of those h and i. If there is any l and h such that $\lim_n\frac{p_l^n}{p_h^n}=0$ and $\lim_n\frac{p_h^n}{p_g^n}=\infty$, then clearly this must be the case as $\frac{p_l}{p_h}=\infty$. The alternative is that there are is some subset $(p_{h_1},\ldots,p_{h_n})$ such that $0<\frac{p_{h_i}}{p_{h_j}}<\infty$ and $\frac{p_g}{p_{h_i}}=0$ for all of the other goods. For Z_g to not explode, these must all equal 0. That gives n equations for a given i

$$\alpha_{ih_j}^{\frac{1}{1+\varepsilon_i}} Q_{ih_j}^{\frac{\varepsilon_i}{1+\varepsilon_i}} p_{h_j}^{\frac{\varepsilon_i}{1+\varepsilon_i}} = \alpha_{ih_j} \sum_k \alpha_{ih_k}^{\frac{1}{1+\varepsilon_i}} Q_{ih_k}^{\frac{\varepsilon_i}{1+\varepsilon_i}} p_{h_k}^{\frac{\varepsilon_i}{1+\varepsilon_i}}, \forall j$$

The only solution to this linear system is $\alpha_{ih_j}^{\frac{1}{1+\varepsilon_i}} Q_{ih_j}^{\frac{\varepsilon_i}{1+\varepsilon_i}} p_{h_j}^{\frac{\varepsilon_i}{1+\varepsilon_i}} = 0$. This contradicts the fact that $p \neq 0$. Therefore, we must have that one of these does not equal to 1, meaning that $Z_g(p^n) \to \infty$.

Since the excess demand function $Z_g\left(\{\bar{p}_g\}_{g\in\mathcal{G}}\right)$ satisfies conditions (i)-(v), recall from above that Proposition 17.C.1 of Mas-Colell et al. (1995) guarantees the existence of a set of central market prices $\{\bar{p}_g(s)\}_{g\in\mathcal{G}}$ and central market consumption $\left\{\bar{C}_g(s)\right\}_{g\in\mathcal{G}}$ that clear the central market (i.e. satisfy condition 2 of the state equilibrium). As condition 1 is then trivially satisfied (see above), this establishes the existence of a state equilibrium. □

Proof of part (b) (uniqueness)

Proof. To establish sufficient conditions for uniqueness, we show that the excess demand function $Z_g\left(\{\bar{p}_g\}_{g\in\mathcal{G}}\right)$ defined in equation (30) satisfies the gross substitutes property $\partial Z_g\left(\{\bar{p}_g\}_{g\in\mathcal{G}}\right)/\partial \bar{p}_h > 0$ for all $h' \neq g$ as long as $\{\varepsilon_i\}$ is sufficiently close to one for all $i \in \mathcal{N}$. Then from Proposition 17.F.3 of Mas-Colell et al. (1995), there exists at most one equilibrium, which, when combined with part (a) (existence) of this proposition, implies that the equilibrium is unique.

We have:

$$\begin{aligned}
p_g \frac{\partial Z_g(p)}{\partial p_{h'}} &= \sum_i \bar{\alpha}_g Q_{ih'} - \bar{\alpha}_g \frac{Q_{ih'}}{1+\varepsilon_i} \alpha_{ih'}^{\frac{\varepsilon_i}{1+\varepsilon_i}} Q_{ih'}^{-\frac{\varepsilon_i}{1+\varepsilon_i}} p_{h'}^{-\frac{\varepsilon_i}{1+\varepsilon_i}} \left(\sum_h \alpha_{ih}^{\frac{1}{1+\varepsilon_i}} p_h^{\frac{\varepsilon_i}{1+\varepsilon_i}} Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \right) \\
&\quad - \bar{\alpha}_g \frac{\varepsilon_i}{1+\varepsilon_i} \left(\sum_h \alpha_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} p_h^{\frac{1}{1+\varepsilon_i}} Q_{ih}^{\frac{1}{1+\varepsilon_i}} \right) \alpha_{ih'}^{\frac{1}{1+\varepsilon_i}} Q_{ih'}^{\frac{\varepsilon_i}{1+\varepsilon_i}} p_{h'}^{-\frac{1}{1+\varepsilon_i}} + \frac{\varepsilon_i}{1+\varepsilon_i} \alpha_{ig}^{\frac{\varepsilon_i}{1+\varepsilon_i}} Q_{ig}^{\frac{1}{1+\varepsilon_i}} p_g^{\frac{1}{1+\varepsilon_i}} \alpha_{ih'}^{\frac{1}{1+\varepsilon_i}} p_{h'}^{-\frac{1}{1+\varepsilon_i}} Q_{ih'}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \\
&= \sum_i \bar{\alpha}_g Q_{ih'} - \bar{\alpha}_g \frac{Q_{ih'}}{1+\varepsilon_i} \alpha_{ih'}^{\frac{\varepsilon_i}{1+\varepsilon_i}} Q_{ih'}^{-\frac{\varepsilon_i}{1+\varepsilon_i}} p_{h'}^{-\frac{\varepsilon_i}{1+\varepsilon_i}} \left(\sum_h \alpha_{ih}^{\frac{1}{1+\varepsilon_i}} p_h^{\frac{\varepsilon_i}{1+\varepsilon_i}} Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \right) \\
&\quad - \bar{\alpha}_g \frac{\varepsilon_i}{1+\varepsilon_i} Q_{ih'} \left(\sum_h \alpha_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} p_h^{\frac{1}{1+\varepsilon_i}} Q_{ih}^{\frac{1}{1+\varepsilon_i}} \right) \alpha_{ih'}^{\frac{1}{1+\varepsilon_i}} Q_{ih'}^{-\frac{1}{1+\varepsilon_i}} p_{h'}^{-\frac{1}{1+\varepsilon_i}} + \frac{\varepsilon_i}{1+\varepsilon_i} \alpha_{ig}^{\frac{\varepsilon_i}{1+\varepsilon_i}} Q_{ig}^{\frac{1}{1+\varepsilon_i}} p_g^{\frac{1}{1+\varepsilon_i}} \alpha_{ih'}^{\frac{1}{1+\varepsilon_i}} p_{h'}^{-\frac{1}{1+\varepsilon_i}} Q_{ih'}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \\
&\geq \sum_i \bar{\alpha}_g Q_{ih'} \frac{\varepsilon_i}{1+\varepsilon_i} \left[\alpha_{ih'}^{\frac{\varepsilon_i}{1+\varepsilon_i}} Q_{ih'}^{-\frac{\varepsilon_i}{1+\varepsilon_i}} p_{h'}^{-\frac{\varepsilon_i}{1+\varepsilon_i}} \left(\sum_h \alpha_{ih}^{\frac{1}{1+\varepsilon_i}} p_h^{\frac{\varepsilon_i}{1+\varepsilon_i}} Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \right) - \alpha_{ih'}^{\frac{1}{1+\varepsilon_i}} Q_{ih'}^{-\frac{1}{1+\varepsilon_i}} p_{h'}^{-\frac{1}{1+\varepsilon_i}} \left(\sum_h \alpha_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} p_h^{\frac{1}{1+\varepsilon_i}} Q_{ih}^{\frac{1}{1+\varepsilon_i}} \right) \right] \\
&\quad + \frac{\varepsilon_i}{1+\varepsilon_i} \alpha_{ig}^{\frac{\varepsilon_i}{1+\varepsilon_i}} Q_{ig}^{\frac{1}{1+\varepsilon_i}} p_g^{\frac{1}{1+\varepsilon_i}} \alpha_{ih'}^{\frac{1}{1+\varepsilon_i}} p_{h'}^{-\frac{1}{1+\varepsilon_i}} Q_{ih'}^{\frac{\varepsilon_i}{1+\varepsilon_i}}
\end{aligned}$$

When $\varepsilon_i = 1$ for all $i \in \mathcal{N}$ we then have:

$$\begin{aligned}
&p_g \frac{\partial Z_g(p)}{\partial p_{h'}} \geq \frac{\varepsilon_i}{1+\varepsilon_i} \alpha_{ig}^{\frac{\varepsilon_i}{1+\varepsilon_i}} Q_{ig}^{\frac{1}{1+\varepsilon_i}} p_g^{\frac{1}{1+\varepsilon_i}} \alpha_{ih'}^{\frac{1}{1+\varepsilon_i}} p_{h'}^{-\frac{1}{1+\varepsilon_i}} Q_{ih'}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \iff \\
&p_g \frac{\partial Z_g(p)}{\partial p_{h'}} > 0,
\end{aligned}$$

since $\frac{\varepsilon_i}{1+\varepsilon_i} \alpha_{ig}^{\frac{\varepsilon_i}{1+\varepsilon_i}} Q_{ig}^{\frac{1}{1+\varepsilon_i}} p_g^{\frac{1}{1+\varepsilon_i}} \alpha_{ih'}^{\frac{1}{1+\varepsilon_i}} p_{h'}^{-\frac{1}{1+\varepsilon_i}} Q_{ih'}^{\frac{\varepsilon_i}{1+\varepsilon_i}} > 0$. Moreover, by continuity, the exists a $\delta > 0$ where, for all

ε_i such that $|\varepsilon_i - 1| < \delta$: we have

$$\frac{\varepsilon_i}{1+\varepsilon_i}\alpha_{ig}^{\frac{\varepsilon_i}{1+\varepsilon_i}} Q_{ig}^{\frac{1}{1+\varepsilon_i}} p_g^{\frac{1}{1+\varepsilon_i}} \alpha_{ih'}^{\frac{1}{1+\varepsilon_i}} p_{h'}^{-\frac{1}{1+\varepsilon_i}} Q_{ih'}^{\frac{\varepsilon_i}{1+\varepsilon_i}} \geq \left| \sum_i \bar{\alpha}_g Q_{ih'} \frac{\varepsilon_i}{1+\varepsilon_i} \left[\begin{array}{c} \alpha_{ih'}^{\frac{\varepsilon_i}{1+\varepsilon_i}} Q_{ih'}^{-\frac{\varepsilon_i}{1+\varepsilon_i}} p_{h'}^{-\frac{\varepsilon_i}{1+\varepsilon_i}} \left(\sum_h \alpha_{ih}^{\frac{1}{1+\varepsilon_i}} p_h^{\frac{\varepsilon_i}{1+\varepsilon_i}} Q_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}}\right) \\ -\alpha_{ih'}^{\frac{1}{1+\varepsilon_i}} Q_{ih'}^{-\frac{1}{1+\varepsilon_i}} p_{h'}^{-\frac{1}{1+\varepsilon_i}} \left(\sum_h \alpha_{ih}^{\frac{\varepsilon_i}{1+\varepsilon_i}} p_h^{\frac{1}{1+\varepsilon_i}} Q_{ih}^{\frac{1}{1+\varepsilon_i}}\right) \end{array} \right] \right|$$

so that $p_g \frac{\partial Z_g(p)}{\partial p_{h'}} > 0$ for all ε_i such that $|\varepsilon_i - 1| < \delta$, as claimed. □

A.2.2 Proof of Proposition 2

We first restate the proposition:

Proposition. *Consider a village i which increases its openness to trade, i.e. ε_i increases by a small amount. Then:*

(1) [Stylized Fact 1] Any increase in openness: (1a) decreases the responsiveness of local prices to local yield shocks; and (1b) increases the responsiveness of local prices to the central market price:

$$\frac{d}{d\varepsilon_i}\left(-\frac{\partial \ln p_{ig}(s)}{\partial \ln A_{ig}(s)}\right) < 0 \text{ and } \frac{d}{d\varepsilon_i}\left(\frac{\partial \ln p_{ig}(s)}{\partial \ln \bar{p}_g}\right) > 0.$$

(2) [Stylized Fact 2] Starting from autarky, an increase in openness: (2a) causes farmers to reallocate production toward crops with higher mean and less volatile yields (as long as $\rho_i > 1$, i.e. farmers are sufficiently risk averse); and (2b) the reallocation toward less volatile crops is attenuated the greater the access to insurance (i.e. the lower ρ_i). Formally, for any two crops $g \neq h$:

$$\frac{d}{d\varepsilon_i}\left(\frac{\partial(\ln\theta_{ig} - \ln\theta_{ih})}{\partial\left(\mu_g^{A,i} - \mu_h^{A,i}\right)}\right)|_{\varepsilon_i=0} > 0,\ \frac{d}{d\varepsilon_i}\left(\frac{\partial \ln\theta_{ig} - \partial\ln\theta_{ih}}{\partial\left(\sum_{h'\in\mathcal{G}}\alpha_{h'}\Sigma_{g,h'}^{A,i} - \sum_{h'\in\mathcal{G}}\alpha_{h'}\Sigma_{h,h'}^{A,i}\right)}\right)|_{\varepsilon_i=0} < 0,$$

$$\text{and } -\frac{d^2}{d\varepsilon_i d\rho_i}\left(\frac{\partial \ln\theta_{ig} - \partial\ln\theta_{ih}}{\partial\left(\sum_{h'\in\mathcal{G}}\alpha_{h'}\Sigma_{g,h'}^{A,i} - \sum_{h'\in\mathcal{G}}\alpha_{h'}\Sigma_{h,h'}^{A,i}\right)}\right)|_{\varepsilon_i=0} > 0.$$

(3) [Stylized Fact 3] Consider a decomposition of the variance of real returns as follows:

$$\sigma_i^{2,Z} = \sigma_i^{2,Y} + \sigma_i^{2,P} - 2cov_i^{Y,P},$$

where

$$\sigma_i^{2,Y} \equiv var(\ln Y_i(s) - c_i(s))$$

is the farmers' nominal income volatility,

$$\sigma_i^{2,P} \equiv var\left(\sum_{g\in\mathcal{G}} \alpha_{ig} \ln p_{ig}(s) + c_i(s)\right)$$

is the farmers' nominal price volatility,

$$cov_i^{Y,P} \equiv cov\left(\ln Y_i^f(s) - c_i(s), \sum_{g\in\mathcal{G}} \alpha_{ig} \ln p_{ig}(s) + c_i(s)\right)$$

is the co-variance between the two and $c_i(s)$ is a nuisance term capturing the aggregate scale of both nominal prices and incomes, which does not affect the aggregate real returns nor the volatility of the real returns. Any increase in openness increases the farmers' nominal income volatility (3a); decreases the farmers' nominal price volatility (3b); and has an ambiguous effect on farmers' real income volatility (3c). Formally, we have:

$$\frac{\partial \sigma_i^{2,Y}}{\partial \varepsilon_i} > 0, \frac{\partial \sigma_i^{2,P}}{\partial \varepsilon_i} < 0, \text{ and } \frac{\partial \sigma_i^{2,Z}}{\partial \varepsilon_i} \lesseqgtr 0.$$

As sufficient condition for farmers' real income volatility to increase with openness, i.e. $\frac{\partial \sigma_i^{2,Z}}{\partial \varepsilon_i} \geq 0$, *is* $\sum_{g\in\mathcal{G}} \theta_{i,g} \left(\sum_{h\in\mathcal{G}} \Sigma_{gh}^{A,i} \alpha_{ih} \right) \geq \sum_{g\in\mathcal{G}} \alpha_{ig} \left(\sum_{h\in\mathcal{G}} \Sigma_{gh}^{A,i} \alpha_{ih} \right)$, *which (loosely speaking) occurs when a farmers' crop allocation is more risky than her expenditure allocation.*

Proof. **Stylized Fact 1.** From equation (10) we have:

$$\begin{aligned} \ln p_{ig}(s) &= -\left(\frac{1}{1+\varepsilon_i}\right) \ln Q_{ig}(s) + \frac{\varepsilon_i}{1+\varepsilon_i} \ln \bar{p}_g(s) + \frac{1}{1+\varepsilon_i} \ln(\alpha_{ig} Y_i(s)) \iff \\ \ln p_{ig}(s) &= -\left(\frac{1}{1+\varepsilon_i}\right) \ln A_{ig}(s) - \left(\frac{1}{1+\varepsilon_i}\right) \ln \theta_{ig} - \left(\frac{1}{1+\varepsilon_i}\right) \ln L_i + \frac{\varepsilon_i}{1+\varepsilon_i} \ln \bar{p}_g + \frac{1}{1+\varepsilon_i} \ln \alpha_{ig} + \frac{1}{1+\varepsilon_i} \ln Y_i(s) \end{aligned}$$

so that:

$$\frac{\partial \ln p_{ig}(s)}{\partial \ln A_{ig}(s)} = -\frac{1}{1+\varepsilon_i}$$

and hence:

$$\frac{d}{d\varepsilon_i}\left(-\frac{\partial \ln p_{ig}(s)}{\partial \ln A_{ig}(s)}\right) = \frac{d}{d\varepsilon_i}\left(\frac{1}{1+\varepsilon_i}\right) = -\frac{1}{(1+\varepsilon_i)^2} < 0.$$

Similarly:

$$\frac{\partial \ln p_{ig}(s)}{\partial \ln \bar{p}_g} = \frac{\varepsilon_i}{1+\varepsilon_i}$$

and hence:

$$\frac{d}{d\varepsilon_i}\left(\frac{\partial \ln p_{ig}(s)}{\partial \ln \bar{p}_g}\right) = \frac{d}{d\varepsilon_i}\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right) = \frac{1}{(1+\varepsilon_i)^2} > 0,$$

as claimed.

Stylized Fact 2a. From equation (18) we have:

$$\theta_{ig} = \frac{\alpha_{ig} (\bar{p}_g B_{ig})^{\varepsilon_i}}{\sum_{h\in\mathcal{G}} \alpha_{ih} (\bar{p}_h B_{ih})^{\varepsilon_i}},$$

where $B_{ig} \equiv \frac{\exp \mu_g^{A,i}}{\left(\lambda_i - \left(\frac{1}{2}\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)^2 \Sigma_{gg}^{A,i} + \frac{\varepsilon_i}{(1+\varepsilon_i)^2} \sum_{h\in\mathcal{G}} \alpha_{ih} \Sigma_{gh}^{A,i} - \rho_i \left(\frac{\varepsilon_i}{1+\varepsilon_i}\right) \sum_{h\in\mathcal{G}} \left(\left(\frac{\varepsilon_i}{1+\varepsilon_i}\right) \theta_{i,h} + \left(\frac{1}{1+\varepsilon_i}\right) \alpha_{ih}\right) \Sigma_{gh}^{A,i}\right)\right)^{\frac{1+\varepsilon_i}{\varepsilon_i}}}$ so that:

$$\ln \theta_{ig} - \ln \theta_{ih} = \ln(\alpha_{ig}) - \ln(\alpha_{ih}) + \varepsilon_i (\ln \bar{p}_g - \ln \bar{p}_h) + \varepsilon_i (\ln B_{ig} - \ln B_{ih})$$

Differentiating this expression with respect to ε_i and evaluating at $\varepsilon_i = 0$ yields:

$$\frac{d}{d\varepsilon_i}\left(\frac{\partial(\ln \theta_{ig} - \ln \theta_{ih})}{\partial\left(\mu_g^{A,i} - \mu_h^{A,i}\right)}\right)|_{\varepsilon_i = 0} = 1 > 0,$$

as claimed.

Stylized Fact 2b. We proceed similarly. Differentiating respect to ε_i and evaluating at $\varepsilon_i = 0$ yields:

$$\frac{d}{d\varepsilon_i}(\partial \ln\theta_{ig} - \partial \ln\theta_{ih})|_{\varepsilon_i=0} = \frac{1}{\lambda_i}(1-\rho_i)\left(\sum_{h'\in\mathcal{G}} \alpha_{h'}\left(\partial\Sigma^{A,i}_{gh'} - \partial\Sigma^{A,i}_{hh'}\right)\right)$$

so that:

$$\frac{d}{d\varepsilon_i}\left(\frac{\partial \ln\theta_{ig} - \partial \ln\theta_{ih}}{\partial\left(\sum_{h'\in\mathcal{G}}\alpha_{h'}\Sigma^{A,i}_{g,h'} - \sum_{h'\in\mathcal{G}}\alpha_{h'}\Sigma^{A,i}_{h,h'}\right)}\right)|_{\varepsilon_i=0} = \frac{1}{\lambda_i}(1-\rho_i),$$

$$\frac{d}{d\varepsilon_i}\left(\frac{\partial \ln\theta_{ig} - \partial \ln\theta_{ih}}{\partial\Sigma^{A,i}_{gg}}\right)|_{\varepsilon_i=0} = \frac{1}{\lambda_i}(1-\rho_i)\alpha_{ig},$$

which is negative as long as $\rho_i > 1$, as claimed.

Stylized Fact 2c. From the previous expression, we immediately have:

$$\frac{d^2}{d\varepsilon_i d\rho}\left(\frac{\partial \ln\theta_{ig} - \partial \ln\theta_{ih}}{\partial\left(\sum_{h'\in\mathcal{G}}\alpha_{h'}\Sigma^{A,i}_{g,h'} - \sum_{h'\in\mathcal{G}}\alpha_{h'}\Sigma^{A,i}_{h,h'}\right)}\right)|_{\varepsilon_i=0} = -\frac{\alpha_{h'}}{\lambda_i}.$$

Stylized Fact 3. Let us first decompose the distribution of real returns into a price term, and income term, and a covariance term. We have:

$$\begin{aligned} Z_i^f(s) &= \prod_{g\in\mathcal{G}} (c_{ig}(s))^{\alpha_{ig}} \\ &= \prod_{g\in\mathcal{G}} \left(\frac{\alpha_{ig} Y_i^f(s)}{p_{ig}(s)}\right)^{\alpha_{ig}} \\ &= Y_i^f(s) \times \prod_{g\in\mathcal{G}} (\alpha_{ig})^{\alpha_{ig}} \times \prod_{g\in\mathcal{G}} (p_{ig}(s))^{-\alpha_{ig}} \end{aligned}$$

so that:

$$\ln Z_i^f(s) = \ln Y_i^f(s) + \sum_{g\in\mathcal{G}} \alpha_{ig}\ln\alpha_{ig} - \sum_{g\in\mathcal{G}} \alpha_{ig}\ln p_{ig}(s).$$

Hence, we can decompose the variance of the real returns as follows:

$$\sigma_i^{2,Z} = \sigma_i^{2,Y} + \sigma_i^{2,P} + 2cov_i^{Y,P}, \tag{31}$$

where:

$$\sigma_i^{2,Y} \equiv var\left(\ln Y_i^f(s) - c_i(s)\right)$$

$$\sigma_i^{2,P} \equiv var\left(-\sum_{g\in\mathcal{G}} \alpha_{ig}\ln p_{ig}(s) + c_i(s)\right)$$

$$cov_i^{Y,P} \equiv cov\left(\ln Y_i^f(s) - c_i(s), -\sum_{g\in\mathcal{G}} \alpha_{ig}\ln p_{ig}(s) + c_i(s)\right)$$

and $c_i(s) \equiv \ln\left(\frac{Y_i(s)}{L_i}\right)^{\frac{1}{1+\varepsilon_i}}$ term captures the aggregate scale of both prices and incomes, which because it affects both terms with opposite signs, does not affect the aggregate returns nor the volatility of the real returns. Let us examine each term in turn.

Focusing first on the income term we have:

$$\ln Y_i^f(s) - \ln\left(\frac{Y_i(s)}{L_i}\right)^{\frac{1}{1+\varepsilon_i}} = \ln\left(\sum_{g\in\mathcal{G}} \theta_{ig}^f A_{ig}(s) p_{ig}(s)\right) - \ln\left(\frac{Y_i(s)}{L_i}\right)^{\frac{1}{1+\varepsilon_i}} \iff$$

$$\ln Y_i^f(s) - \ln\left(\frac{Y_i(s)}{L_i}\right)^{\frac{1}{1+\varepsilon_i}} = \ln\left(\sum_{g\in\mathcal{G}} \left(\frac{\theta_{ig}^f}{\theta_{ig}}\right) \times \alpha_{ig}\left(\frac{\bar{p}_g A_{ig}(s)\theta_{ig}}{\alpha_{ig}}\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}\right)$$

Applying the same second order approximation as in the main text we have:

$$\ln Y_i^f(s) - \ln\left(\frac{Y_i(s)}{L_i}\right)^{\frac{1}{1+\varepsilon_i}} \approx \ln\left(\sum_{g\in\mathcal{G}} \alpha_{ig}\left(\frac{\bar{p}_g \exp(\mu_g^{A,i})\theta_{ig}}{\alpha_{ig}}\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}\right) - \sum_{g\in\mathcal{G}} \theta_{ig} \ln\left(\theta_{ig}^{-1}\alpha_{ig}\left(\frac{\bar{p}_g \exp(\mu_g^{A,i})\theta_{ig}}{\alpha_{ig}}\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}\right)$$
$$+ \sum_{g\in\mathcal{G}} \theta_{ig} \ln\left(\theta_{ig}^{-1}\alpha_{ig}\left(\frac{\bar{p}_g A_{ig}(s)\theta_{ig}}{\alpha_{ig}}\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}\right) - -\frac{1}{2}\sum_{h\in\mathcal{G}}\sum_{g\in\mathcal{G}} \theta_{ig}\theta_{ih}\Sigma_{gh}^{A,i} + \frac{1}{2}\sum_{g\in\mathcal{G}} \theta_{ig}\Sigma_{gh}^{A,i}$$

so that:

$$\sigma_i^{2,Y} = \left(\frac{\varepsilon_i}{1+\varepsilon_i}\right)^2 \sum_{g\in\mathcal{G}}\sum_{h\in\mathcal{G}} \theta_{ig}\theta_{ih}\Sigma_{gh}^{A,i} \tag{32}$$

Now focusing on the price term we have:

$$-\sum_{g\in\mathcal{G}} \alpha_{ig} \ln p_{ig}(s) + \ln\left(\frac{Y_i(s)}{L_i}\right)^{\frac{1}{1+\varepsilon_i}} = \sum_{g\in\mathcal{G}} \alpha_{ig} \ln\left(\left(\frac{\bar{p}_g A_{ig}(s)\theta_{ig}}{\alpha_{ig}}\right)^{\frac{1}{1+\varepsilon_i}} (\bar{p}_g)^{-1}\right) \iff$$
$$= \left(\frac{1}{1+\varepsilon_i}\right) \sum_{g\in\mathcal{G}} \alpha_{ig} \ln A_{ig}(s) + \sum_{g\in\mathcal{G}} \alpha_{ig} \ln\left(\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}\right)^{\frac{1}{1+\varepsilon_i}} (\bar{p}_g)^{-1}\right)$$

so that the variance of the prices can be written as:

$$\sigma_i^{2,P} = \left(\frac{1}{1+\varepsilon_i}\right)^2 \sum_{g\in\mathcal{G}}\sum_{h\in\mathcal{G}} \Sigma_{gh}^{A,i} \alpha_{ig}\alpha_{ih} \tag{33}$$

$$p_{ig} = (A_{ig}\theta_{ig}L_i)^{-\frac{1}{1+\varepsilon_i}} (\bar{p}_g)^{\frac{\varepsilon_i}{1+\varepsilon_i}} (\alpha_{ig}Y_i)^{\frac{1}{1+\varepsilon_i}}$$

Finally, the covariance between prices and incomes can be written as:

$$cov_i^{Y,P} = \frac{\varepsilon_i}{(1+\varepsilon_i)^2} \sum_{g\in\mathcal{G}}\sum_{h\in\mathcal{G}} \theta_{ig}\alpha_{ih}\Sigma_{gh}^{A,i}. \tag{34}$$

It is straightforward to verify that applying the decomposition (31) to expressions (32), (33), and (34) immediately yields expression (29) for the variance of the total real returns.

Now consider a small increase in the openness of a location. How does it affect the variance of farmers' incomes, prices, and the co-variance between the two? We immediately have:

$$\frac{\partial \sigma_i^{2,Y}}{\partial \varepsilon_i} = 2\frac{\varepsilon_i^2}{(1+\varepsilon_i)^3} \sum_{g\in\mathcal{G}}\sum_{h\in\mathcal{G}} \theta_{ig}\theta_{ih}\Sigma_{gh}^{A,i} > 0$$

$$\frac{\partial \sigma_i^{2,P}}{\partial \varepsilon_i} = -2 \frac{1}{(1+\varepsilon_i)^3} \sum_{g \in \mathcal{G}} \sum_{h \in \mathcal{G}} \Sigma_{gh}^{A,i} \alpha_{ig} \alpha_{ih} < 0,$$

as required.

Let us turn now to the variance of the total real returns. Recall from equation (29) that the variance of real returns is:

$$\begin{aligned}
\sigma_i^{2,Z} &\equiv \sum_{g \in \mathcal{G}} \sum_{h \in \mathcal{G}} \left(\left(\frac{\varepsilon_i}{1+\varepsilon_i} \right) \theta_{i,g}^f + \left(\frac{1}{1+\varepsilon_i} \right) \alpha_{ig} \right) \left(\left(\frac{\varepsilon_i}{1+\varepsilon_i} \right) \theta_{i,h}^f + \left(\frac{1}{1+\varepsilon_i} \right) \alpha_{ih} \right) \Sigma_{gh}^{A,i} \iff \\
&= \sum_{g \in \mathcal{G}} \sum_{h \in \mathcal{G}} \left(\omega_i \theta_{i,g}^f + (1-\omega_i) \alpha_{ig} \right) \left(\omega_i \theta_{i,h}^f + (1-\omega_i) \alpha_{ih} \right) \Sigma_{gh}^{A,i},
\end{aligned}$$

where $\omega_i \equiv \frac{\varepsilon_i}{1+\varepsilon_i}$. Note that $\frac{\partial \omega_i}{\partial \varepsilon_i} = \frac{1}{1+\varepsilon_i} - \frac{\varepsilon_i}{1+\varepsilon_i} \frac{1}{1+\varepsilon_i} = \frac{1}{1+\varepsilon_i} \left(1 - \frac{\varepsilon_i}{1+\varepsilon_i} \right) = \frac{1}{(1+\varepsilon_i)^2}$, so that $\frac{\partial \sigma_i^{2,Z}}{\partial \varepsilon_i} = \frac{1}{(1+\varepsilon_i)^2} \frac{\partial \sigma_i^{2,Z}}{\partial \omega_i}$. We then have:

$$\begin{aligned}
\frac{\partial \sigma_i^{2,Z}}{\partial \varepsilon_i} &= \frac{1}{(1+\varepsilon_i)^2} \frac{\partial}{\partial \omega_i} \left(\sum_{g \in \mathcal{G}} \sum_{h \in \mathcal{G}} \left(\omega_i \theta_{i,g}^f + (1-\omega_i) \alpha_{ig} \right) \left(\omega_i \theta_{i,h}^f + (1-\omega_i) \alpha_{ih} \right) \Sigma_{gh}^{A,i} \right) \iff \\
&= \frac{2}{(1+\varepsilon_i)^2} \left(\omega_i \left(\sum_{g \in \mathcal{G}} \sum_{h \in \mathcal{G}} \left(\theta_{i,g}^f - \alpha_{ih} \right) \left(\theta_{i,h}^f - \alpha_{ih} \right) \Sigma_{gh}^{A,i} \right) + \sum_{g \in \mathcal{G}} \sum_{h \in \mathcal{G}} \left(\theta_{i,g}^f - \alpha_{ig} \right) \alpha_{ih} \Sigma_{gh}^{A,i} \right)
\end{aligned}$$

Because $\Sigma_{gh}^{A,i}$ is positive definite, we know that $\sum_{g \in \mathcal{G}} \sum_{h \in \mathcal{G}} \left(\theta_{i,g}^f - \alpha_{ih} \right) \left(\theta_{i,h}^f - \alpha_{ih} \right) \Sigma_{gh}^{A,i} \geq 0$ for any crop allocation $\left\{ \theta_{i,g}^f \right\}$ and expenditure shares $\{\alpha_{ig}\}$. Hence, $\frac{\partial \sigma_i^{2,Z}}{\partial \varepsilon_i} \geq 0$ if:

$$\begin{aligned}
&\sum_{g \in \mathcal{G}} \sum_{h \in \mathcal{G}} \left(\theta_{i,g}^f - \alpha_{ig} \right) \alpha_{ih} \Sigma_{gh}^{A,i} \geq 0 \iff \\
&\sum_{g \in \mathcal{G}} \theta_{i,g}^f \left(\sum_{h \in \mathcal{G}} \Sigma_{gh}^{A,i} \alpha_{ih} \right) \geq \sum_{g \in \mathcal{G}} \alpha_{ig} \left(\sum_{h \in \mathcal{G}} \Sigma_{gh}^{A,i} \alpha_{ih} \right),
\end{aligned}$$

as required. □

A.2.3 Proof of Proposition #3

We first restate the proposition:

Proposition. *1) In the presence of volatility, moving from autarky to costly trade improves farmer welfare, i.e. the gains from trade are positive; 2) moving from a world with no volatility to one with volatility amplifies farmers' gains from trade; but 3) increasing the volatility in an already volatile world may attenuate farmers' gains from trade*

Proof. **Part 1**. From equation (25), the real income of farmer f in village $i \in \mathcal{N}$ in state $s \in S$ with crop allocation $\left\{ \theta_{ig}^f \right\}_{g \in \mathcal{G}}$ can be written as:

$$Z_i^f \left(s; \left\{ \theta_{ig}^f \right\}_{g \in \mathcal{G}} \right) = \frac{\left(\sum_{g \in \mathcal{G}} \theta_{ig}^f \times \frac{\alpha_{ig}}{\theta_{ig}} \left(\frac{\bar{p}_g \theta_{ig}}{\alpha_{ig}} A_{ig}(s) \right)^{\frac{\varepsilon_i}{1+\varepsilon_i}} \right) \prod_{g \in \mathcal{G}} \left(\alpha_{ig} \left(\frac{\bar{p}_g \theta_{ig}}{\alpha_{ig}} A_{ig}(s) \right)^{\frac{1}{1+\varepsilon_i}} \right)^{\alpha_{ig}}}{\prod_{g \in \mathcal{G}} (\bar{p}_g)^{\alpha_{ig}}}. \quad (35)$$

Consider first the case of autarky, where $\varepsilon_i = 0$. From equation (18), a farmers' optimal autarkic crop allocation is simply equal to her expenditure share, i.e. $\theta_{ig}^f = \alpha_{ig}$,so that from equation (35) her autarkic welfare is:

$$Z_i^{f,aut}(s) = \prod_{g\in\mathcal{G}} (\alpha_{ig} \times A_{ig}(s))^{\alpha_{ig}}.$$

Now consider the case of (costly) trade, where $\varepsilon_i > 0$ but farmer f chooses her autarkic crop allocation. Then from equation (35), her real income is:

$$Z_i^f\left(s;\{\alpha_{ig}\}_{g\in\mathcal{G}}\right) = \frac{\left(\sum_{g\in\mathcal{G}} \alpha_{ig} \times \frac{\alpha_{ig}}{\theta_{ig}}\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}A_{ig}(s)\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}\right)\prod_{g\in\mathcal{G}}\left(\alpha_{ig}\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}A_{ig}(s)\right)^{\frac{1}{1+\varepsilon_i}}\right)^{\alpha_{ig}}}{\prod_{g\in\mathcal{G}}(\bar{p}_g)^{\alpha_{ig}}}. \tag{36}$$

Note that from the generalized mean inequality we have:

$$\sum_{g\in\mathcal{G}} \alpha_{ig} \times \frac{\alpha_{ig}}{\theta_{ig}}\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}A_{ig}(s)\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}} \geq \prod_{g\in\mathcal{G}}\left(\frac{\alpha_{ig}}{\theta_{ig}}\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}A_{ig}(s)\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}\right)^{\alpha_{ig}},$$

with equality only in the case where $\frac{\alpha_{ig}}{\theta_{ig}}\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}A_{ig}(s)\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}} = c_i$ for all $g\in\mathcal{G}$. Substituting this inequality into equation (36) immediately implies

$$Z_i^f\left(s;\{\alpha_{ig}\}_{g\in\mathcal{G}}\right) \geq Z_i^{f,aut}(s),$$

again with equality only in the case where $\frac{\alpha_{ig}}{\theta_{ig}}\left(\frac{\bar{p}_g\theta_{ig}}{\alpha_{ig}}A_{ig}(s)\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}} = c_i$ for all $g \in \mathcal{G}$. Intuitively, as long as the equilibrium price vector is not exactly equal to the slope of the production possibility frontier, farmers can gain by selling goods for which they are relatively more productive and buying goods for which they are relatively less productive. As the productivity realizations are log-normally distributed across states of the world, this equality only occurs with measure zero. Hence, for almost all $s\in S$, we have $Z_i^f\left(s;\{\alpha_{ig}\}_{g\in\mathcal{G}}\right) > Z_i^{f,aut}(s)$, which in turn implies that the expected utility of a farmer choosing her autarkic allocation with costly trade is strictly greater than the expected utility of a farmer in autarky choosing her autarkic allocation, i.e. $\mathbb{E}\left[U_i^f\left(\{\alpha_{ig}\}_{g\in\mathcal{G}}\right)\right] > \mathbb{E}\left[U_i^{f,aut}\right]$. Finally, as farmers make their crop choice to maximize their expected utility, their actual expected welfare with costly trade is at least as great as their expected utility holding their crop choice at the autarkic allocation, so that $\mathbb{E}\left[\max_{\{\theta_{ig}\}_{g\in\mathcal{G}}} U_i^f\left(\{\theta_{ig}\}_{g\in\mathcal{G}}\right)\right] > \mathbb{E}\left[U_i^{f,aut}\right]$, i.e. the gains from trade are strictly positive, as claimed.

Part 2. In the absence of volatility, farmers' utility is invariant to ε_i, i.e. there are zero gains from trade. From Part 1, in the presence of volatility, there are strictly positive gains from trade. Taken together, this implies that the presence of volatility amplifies the gains from trade, as claimed.

Part 3. We prove the statement by example, illustrated in Supplemental Materials Table B.7. Consider a world where there are two types of villages (1 and 2) and two crops (A and B). Suppose both villages have equal expenditure shares on both crops in equal proportions and the means of both crops in both villages is identical. Suppose first that crop A in village 1 and crop B in village 2 are "risky" (i.e. have equally volatile yields), whereas crop B in village 1 and crop A in village 2 are "safe" (i.e. have zero yield volatility). In autarky, both village types grow equal amounts of both crops, but with trade, the two types of villages can specialize in the "safe" crops, achieving positive gains from trade (Case 1 in Supplemental Materials Table B.7). Suppose now that we increase the volatility of the safe crop in both village types so that it receives the same yield shock as the risky crop (i.e. the two crops have perfectly correlated yields within each village, although independent yield realizations across villages). As the relative yields between the two crops are always equal in both types of villages, there are no gains from trade (Case 2 in Supplemental Materials Table B.7), illustrating that increasing the volatility in an already volatile world can reduce the gains from trade, as required. ☐

A.3 Model isomorphisms, extensions, and additional results

In this subsection, we present isomorphisms, extensions, and additional results for the model presented in the main paper.

A.3.1 Endogenous capacity constraints

In this subsection, we show how the framework presented in the paper is isomorphic to one in which better traders exchange greater amounts of goods, i.e. have greater capacity for arbitrage. To do so, we suppose that traders with lower trade costs (i.e. lower τ's) are able to offer greater capacity, with the following constant elasticity function:

$$Q(\tau)=c_i\tau^{-\lambda}$$

When $\lambda=0$, capacity is fixed, but for $\lambda>0$ we have the intuitive result that better traders (with lower τ) are able to engage in greater amounts of trade. The constant elasticity form – while analytically convenient – can be viewed as a first-order log-linear approximation to any function where better traders have greater capacity. The scalar c_i is determined to ensure that a single trader handles each unit of production (if traders are buying goods in the village to sell to the market) or consumption (if traders are buying goods in the market to sell to the village). We consider each case in turn.

Suppose first that $\bar{p}\geq p_i$ so that traders buy goods produced in the village and sell them in the market. In this case, it must be that each unit produced in the village is handled by a trader, i.e.:

$$Q_i=\int Q(\tau)dF(\tau).$$

Maintaining the assumption in the main text that the distribution of traders is Pareto distributed with shape parameter ε_i, we have:

$$Q_i=c_i\varepsilon_i\int_1^\infty \tau^{-\lambda-\varepsilon_i-1}d\tau \iff$$
$$Q_i\left(\frac{\lambda+\varepsilon_i}{\lambda}\right)=c_i$$

It is straightforward to calculate the quantity of units the traders purchase in the village and sell to the market:

$$Q_{im}=\int_1^{\frac{\bar{p}}{p_i}} Q(\tau)dF(\tau)\iff$$
$$Q_{im}=\left(1-\left(\frac{\bar{p}}{p_i}\right)^{-(\lambda+\varepsilon_i)}\right)Q_i$$

And the remainder of the production is sold to consumers locally so that:

$$C_i=\left(\frac{\bar{p}}{p_i}\right)^{-(\lambda+\varepsilon_i)}Q_i. \tag{37}$$

Suppose now that $\bar{p} < p_i$ so that traders buy goods in the market and sell them to farmers in the village. In this case, it must be that each unit consumed in the village is handled by a trader, i.e.:

$$C_i = \int Q(\tau) dF(\tau),$$

which yields through an identical derivation as above:

$$C_i \left(\frac{\lambda + \varepsilon_i}{\lambda} \right) = c_i.$$

It is then straightforward to calculate the quantity of units the traders purchase in the market and sell to the village:

$$Q_{mi} = \int_1^{\frac{p_i}{\bar{p}}} Q(\tau) dF(\tau) \iff$$
$$Q_{im} = \left(1 - \left(\frac{p_i}{\bar{p}} \right)^{-(\lambda + \varepsilon_i)} \right) C_i.$$

The remainder of the consumption in the village comes from local production, i.e.:

$$Q_i = \left(\frac{p_i}{\bar{p}} \right)^{-(\lambda + \varepsilon_i)} C_i. \tag{38}$$

Equations (37) and (38) are identical and isomorphic to equation (9) in the main text. This demonstrates that the shape parameter of the Pareto distribution ε_i (where traders are assumed to be infinitely capacity constrained) can be equivalently thought of as a combination of the exogenous heterogeneity of the trade costs across traders and an endogenous component related to the fact that better traders are able to engage in greater amounts of arbitrage.

A.3.2 Expressions for trader and driver incomes

In this subsection, we derive the trader and driver income separately. Let $Y_{ig}^{trader}(s)$ and $Y_{ig}^{driver}(s)$ be the income earned by the trader (from price-arbitrage) and the driver (from the iceberg trade costs), respectively, for the trade of good g between village i and the central market.

It is convenient to first calculate the sum of the trader and driver incomes. Suppose first that the central market price $\bar{p}_g(s)$ exceeds the village price $p_{ig}(s)$, so that trade will flow from the village to the central market. In this case, the sum of trader and driver income can be expressed as:

$$Y_{ig}^{trader}(s) + Y_{ig}^{driver}(s) = (\bar{p}_g(s) - p_{ig}(s))(Q_{ig}(s) - C_{ig}(s))$$

Suppose now that the central market price $\bar{p}_g$ is below the village price $p_{ig}(s)$, so that trade will flow from the central market to the village. In this case, the sum of trader and driver income can be expressed as:

$$Y_{ig}^{trader}(s) + Y_{ig}^{driver}(s) = (p_{ig}(s) - \bar{p}_g(s))(C_{ig}(s) - Q_{ig}(s)).$$

In both cases, when combined with equation (9), the following expression for combined income of traders and drivers is obtained:

$$Y_{ig}^{trader}(s)+Y_{ig}^{driver}(s)=(\bar{p}_g(s)-p_{ig}(s))\left(1-\left(\frac{p_{ig}(s)}{\bar{p}_g(s)}\right)^{\varepsilon_i}\right)Q_{ig}(s). \tag{39}$$

Total trader and driver income can then be calculated by summing across all villages and all goods, as in equation (12).

Now consider the income of traders alone. Suppose first that the central market price exceeds the village price, i.e. $\bar{p}_g(s)\geq p_{ig}(s)$. Then the trader income earned from arbitrage can be calculated by integrating the arbitrage profits across the distribution of trade costs incurred by traders:

$$Y_{ig}^{trader}(s)=\underbrace{Q_{ig}(s)}_{\#\text{ of matches}}\int_1^\infty\underbrace{(\bar{p}_g(s)-\tau p_{ig}(s))}_{\text{arbitrage profits}}\underbrace{\mathbf{1}\{\bar{p}_g(s)\geq\tau p_{ig}(s)\}}_{\text{only trade if profitable}}\underbrace{dF(\tau)}_{\text{trader dist.}}.$$

Given the assumed Pareto distribution of trade costs from equation (6) and equation (9), this expression simplifies to:

$$Y_{ig}^{trader}(s)=\frac{1}{\varepsilon_i}\bar{p}_g(s)(Q_{ig}(s)-C_{ig}(s))+\frac{1}{\varepsilon_i-1}(\bar{p}_g(s)C_{ig}(s)-p_{ig}(s)Q_{ig}(s)). \tag{40}$$

Suppose now that the central market price is below the village price, i.e. $\bar{p}_g(s)\leq p_{ig}(s)$. Then the trader income earned from arbitrage can again be calculated by integrating the arbitrage profits across the distribution of trade costs incurred by traders:

$$Y_{ig}^{trader}(s)=\underbrace{C_{ig}(s)}_{\#\text{ of matches}}\int_1^\infty\underbrace{(p_{ig}(s)-\tau\bar{p}_g(s))}_{\text{arbitrage profits}}\underbrace{\mathbf{1}\{p_{ig}(s)\geq\tau\bar{p}_g(s)\}}_{\text{only trade if profitable}}\underbrace{dF(\tau)}_{\text{trader dist.}}.$$

Again, given the assumed Pareto distribution of trade costs from equation (6) and equation (9), this expression simplifies to:

$$Y_{ig}^{trader}(s)=\frac{1}{\varepsilon_i}p_{ig}(s)(C_{ig}(s)-Q_{ig}(s))+\frac{1}{\varepsilon_i-1}(p_{ig}(s)Q_{ig}(s)-\bar{p}_g(s)C_{ig}(s)). \tag{41}$$

Together, equations (40) and (41) characterize the portion of trade income earned by the trader; the difference between the expressions and the total income to both traders and drivers given in equation (39) is then the income earned by the driver.

A.3.3 A microfoundation for insurance

In the baseline model presented in Section 4, the farmer's utility function is given by equation (5):

$$U_i^f(s)\equiv\frac{1}{1-\rho_i}\left(\left(Z_i^f(s)\right)^{1-\rho_i}-1\right)$$

where ρ_i is the "effective" risk aversion parameter and we show that $\ln Z_i^f(s)\sim N\left(\mu_i^Z,\sigma_i^{2,Z}\right)$, which then implies that farmers' expected utility can be written as in equation (16):

$$\mathbb{E}\left[U_i^f\right]=\left(\frac{1}{1-\rho_i}\right)\left(\exp\left((1-\rho_i)\left(\mu_i^Z+\frac{1}{2}(1-\rho_i)\sigma_i^{2,Z}\right)\right)-1\right). \tag{42}$$

In what follows, we will show that there exists a micro-foundation for the "effective" risk aversion parameter ρ_i whereby farmers purchase insurance from perfectly competitive lenders ("banks"). In this micro-foundation, the "effective" risk aversion parameter ρ_i can then be written as a function of the (fundamental) risk aversion of farmers and a (technological) parameter governing the efficiency of the insurance market. As a result, we can interpret changes to the "effective" risk aversion parameter as technological changes in the access to banks, allowing us to perform normative counterfactual analysis.

Suppose that all farmers have identical and fundamental risk aversion parameters ρ_0 and have access to banks that offer insurance at perfectly competitive rates. To save on notation, in what follows, we will omit the location of the farmer and denote states of the world with subscripts, the probability of state of the world s with π_s. Suppose that the insurance allows pays out one unit of the consumption bundle in state of the world s for price p_s.[27] Hence, consumption in state of the world s will be the sum of the realized consumption in that state and the insurance payout less the cost of insurance: $C_s = Z_s + q_s - \sum_t p_t q_t$, where q_s is the quantity of insurance for state s purchased by the farmer. A farmer's expected utility function ex-post insurance is then:

$$\mathbb{E}\left[U^{f,ins}\right] = \sum_s \pi_s \frac{1}{1-\rho_0}\left((C_s)^{1-\rho_0} - 1\right).$$

Farmers purchase their insurance from a large number of "money-lenders" (or, equivalently, banks). Money-lenders have the same income realizations and preference-structure as farmers and face the same prices, but are distinct from farmers in that they are less risk averse. Let money-lenders' risk aversion parameter be denoted by $\lambda \leq \rho_0$, where we view λ as a technological parameter governing the quality/access farmers have to credit: the better farmers' access to credit, the lower the risk aversion of money-lenders.

Because lenders are also risk averse, farmers will not be able to perfectly insure themselves. Money lenders compete with each other to lend money, and hence the price of purchasing insurance in a particular state of the world is determined by the marginal cost of lending money. We first calculate the price of a unit of insurance in state of the world s. Since the price of insurance is determined in perfect competition, it must be the case that each money lender is just indifferent between offering insurance and not:

$$\sum_{t \neq s} \pi_t \frac{1}{1-\lambda}(Z_t + \varepsilon p_s)^{1-\lambda} + \pi_s \frac{1}{1-\lambda}(Z_t + \varepsilon p_s - \varepsilon)^{1-\lambda} = \sum_t \pi_t \frac{1}{1-\lambda} Z_t^{1-\lambda},$$

where the left hand side is the expected utility of a money-lender offering an small amount ε of insurance (which pays εp_s with certainty but costs ε in state of the world s) and the right hand side is expected utility of not offering the insurance. Taking the limit as ε approaches zero yields that the price ensures that the marginal utility benefit of receiving $p_s \varepsilon$ in all other states of the world is equal to the marginal utility cost of paying $\varepsilon(1-p_s)$ in state of the world s.

$$p_s \varepsilon \sum_{t \neq s} \pi_t Z_t^{-\lambda} = \varepsilon(1-p_s)\pi_s Z_s^{-\lambda} \iff$$

$$p_s = \frac{\pi_s Z_s^{-\lambda}}{\sum_t \pi_t Z_t^{-\lambda}}. \tag{43}$$

Equation (43) is intuitive: it says that the price of insuring states of the world with low aggregate income is high.

Now consider the farmer's choice of the optimal level of insurance. Farmers will choose the quantity of

[27] For simplicity – and without loss of generality as the state of the world defines the price index – we measure both the insurance payout and the prices in real (i.e. price index adjusted) units.

insurance to purchase in each period in order to maximize their expected utility:

$$\max_{\{q_s\}} \sum_s \pi_s \frac{1}{1-\rho_0}\left(\left(Z_s+q_s-\sum_t p_t q_t\right)^{1-\rho_0}-1\right)$$

which yields the following FOC with respect to q_s:

$$\pi_s\left(Z_s+q_s-\sum_t p_t q_t\right)^{-\rho_0} = p_s \sum_t \pi_t\left(Z_t+q_t-\sum_t p_t q_t\right)^{-\rho_0} \iff$$

$$\frac{\pi_s C_s^{-\rho_0}}{\sum_t \pi_t C_t^{-\rho_0}} = p_s. \tag{44}$$

Substituting the equilibrium price from equation (43) into equation (44) and noting that $\mathbb{E}[C^{-\rho_0}] = \sum_t \pi_t C_t^{-\rho_0}$ and $\mathbb{E}\left[I^{-\lambda}\right] = \sum_t \pi_t I_t^{-\lambda}$ yields:

$$\frac{C_s^{-\rho_0}}{\mathbb{E}[C^{-\rho_0}]} = \frac{Z_s^{-\lambda}}{\mathbb{E}[Z^{-\lambda}]}. \tag{45}$$

Because the first order conditions (44) are homogeneous of degree zero in consumption, they do not pin down the scale of ex-post real income, so to ensure that access to insurance only affects welfare through the second moment of returns, we assume that access to insurance does not affect the log mean real returns of farmers, i.e. $\mathbb{E}[\ln C_s] = \mu^Z$. As a result, we can write:

$$C_s = Z_s^{\frac{\lambda}{\rho_0}}\left(\exp\left(\mu^Z\right)\right)^{1-\frac{\lambda}{\rho_0}}, \tag{46}$$

i.e. access to insurance means that the ex-post realized real returns after insurance payouts are a Cobb-Douglas combination of the ex-ante realized returns prior to insurance payouts and the (log) mean real returns. This is intuitive: when money lenders have the same level of risk aversion as the farmers (i.e. $\lambda = \rho_0$), farmers' ex-post returns are equal to their ex-ante returns, i.e. there is no scope for insurance. Conversely, when money lenders are risk-neutral (i.e. $\lambda = 0$), farmers' ex-post returns are simply equal to their mean real returns, i.e. they are perfectly insured. When money-lenders are still risk averse but less so than farmers, there is scope for imperfect insurance, where the scope depends on the degree of risk aversion of the money-lenders. Indeed, equation (46) can be viewed as a first-order log-linear approximation of any insurance technology that reduces the variance of ex-post realized returns around its mean.

Given that the ex-ante realized returns are log-normally distributed $\ln Z_s \sim N\left(\mu^Z, \sigma^{2,Z}\right)$, the ex-post realized returns are also log-normally distributed with:

$$\ln C_s \sim N\left(\mu^Z, \left(\frac{\lambda}{\rho_0}\right)^2 \sigma^{2,Z}\right)$$

so that farmers' expected utility ex post insurance can be written as:

$$\mathbb{E}\left[U^{f,ins}\right] = \frac{1}{1-\rho_0}\left(\exp\left((1-\rho_0)\left(\mu^Z + \frac{1}{2}(1-\tilde{\rho})\sigma^{2,Z}\right)\right)-1\right), \tag{47}$$

where

$$\tilde{\rho} = 1 + (\rho_0 - 1)\left(\frac{\lambda}{\rho_0}\right)^2$$

is the effective level of risk aversion. As a result, we have now shown that the effective level of risk aversion can be written as a function of the innate risk aversion of farmers (ρ_0) and the technological parameter governing their access to insurance markets (as captured by λ), as claimed.

Finally, consider the evaluation of the welfare impact of some counterfactual that changes potentially both the access to insurance markets and the distribution of real returns from $\left\{\lambda_A,\mu_A^Z,\sigma_A^{2,Z}\right\}$ to $\left\{\lambda_B,\mu_B^Z,\sigma_B^{2,Z}\right\}$. The change in expected utility is:

$$(1-\rho_0)\left(\mathbb{E}\left[U_B^{f,ins}\right]-\mathbb{E}\left[U_A^{f,ins}\right]\right)=\exp\left((1-\rho_0)\left(\mu_B^Z+\frac{1}{2}(1-\rho_0)\left(\frac{\lambda_B}{\rho_0}\right)^2\sigma_B^{2,Z}\right)\right)$$
$$-\exp\left((1-\rho_0)\left(\mu_A^Z+\frac{1}{2}(1-\rho_0)\sigma_A^{2,Z}\left(\frac{\lambda_A}{\rho_0}\right)^2\right)\right).$$

We now define what we call the certainty equivalent variation (CEV), which is the hypothetical percentage increase in income an individual would need to receive with certainty that would yield an equivalent change in expected welfare as the counterfactual, holding constant all prices and parameters constant at the baseline. It is straightforward to show that the CEV can be written as:

$$CEV=\left(\mu_B^Z+\frac{1}{2}(1-\rho_0)\left(\frac{\lambda_B}{\rho_0}\right)^2\sigma_B^{2,Z}\right)-\left(\mu_A^Z+\frac{1}{2}(1-\rho_0)\left(\frac{\lambda_A}{\rho_0}\right)^2\sigma_A^{2,Z}\right), \tag{48}$$

or, equivalently, we can write the CEV in terms of the effective risk aversion:

$$CEV=\left(\mu_B^Z+\frac{1}{2}(1-\tilde{\rho}_B)\sigma_B^{2,Z}\right)-\left(\mu_A^Z+\frac{1}{2}(1-\tilde{\rho}_A)\sigma_A^{2,Z}\right),$$

where $\tilde{\rho}_A\equiv 1+(\rho_0-1)\left(\frac{\lambda_A}{\rho_0}\right)^2$ and $\tilde{\rho}_B\equiv 1+(\rho_0-1)\left(\frac{\lambda_B}{\rho_0}\right)^2$ are the effective risk aversion parameters we estimate in Section 5.2. This is the welfare metric we report in Section 5.

A.3.4 Convex transportation costs

In equation (9), we show that under the appropriate set of assumptions, heterogeneous traders and a market clearing condition imply the following no-arbitrage condition:

$$\frac{C_{ig}(s)}{Q_{ig}(s)}=\left(\frac{p_{ig}(s)}{\bar{p}_g(s)}\right)^{\varepsilon_i}$$

i.e. goods flow toward locations with higher relative prices. In this subsection, we provide an alternative setup that generates the same no-arbitrage condition assuming that transportation costs are increasing and convex in the quantity traded.[28] For notational simplicity, we omit the good g and state s notation in what follows.

As in the model in the paper, suppose there is a (small) village i engaging in trade with a (large) market subject to trade costs. Unlike the model in the paper where the trade costs are heterogeneous across traders, suppose now that they increase convexly with the quantity shipped between the village and the market. In particular, let $\bar{M}_i$ denote the quantity of goods imported by village i from the market and $\bar{X}_i$ denote the quantity of goods exported by village i to the market. Suppose that the iceberg trade cost τ_i between the

[28]We are grateful to Rodrigo Adao for pointing out this alternative setup.

village i and its market can be written as:

$$\ln\tau_i = \frac{1}{\varepsilon_i}\ln\left(1+\frac{\bar{M}_i}{Q_i}+\frac{\bar{X}_i}{C_i}\right), \tag{49}$$

where Q_i and C_i are the quantity produced and consumed in village i, respectively. Intuitively, equation (49) says that the greater the flows of goods between the village and the market – relative to the quantity produced in i for imports and relative to the quantity consumed in i for exports – the greater the iceberg trade costs incurred.

Now consider what equation (49) implies when combined with a no-arbitrage condition. Suppose first that the market price exceeds the village price, i.e. $\bar{p} \geq p_i$. In this case, the village will only export the good to the market, i.e. $\bar{M}_i = 0$ and $\bar{X}_i \geq 0$ and the following no-arbitrage condition will hold:

$$\begin{aligned}
\ln\bar{p} - \ln p_i &= \ln\tau_i \iff \\
\ln\bar{p} - \ln p_i &= \frac{1}{\varepsilon_i}\ln\left(1+\frac{\bar{X}_i}{C_i}\right) \iff \\
1+\frac{\bar{X}_i}{C_i} &= \left(\frac{\bar{p}}{p_i}\right)^{\varepsilon_i}
\end{aligned} \tag{50}$$

Now consider the case where the village price exceeds the market price, i.e. $p_i \geq \bar{p}$. In this case, the village will only import the good from the market, i.e. $\bar{M}_i \geq 0$ and $\bar{X}_i = 0$ and the following no-arbitrage condition will hold:

$$\begin{aligned}
\ln p_i - \ln\bar{p} &= \ln\tau_i \iff \\
\ln p_i - \ln\bar{p} &= \frac{1}{\varepsilon_i}\ln\left(1+\frac{\bar{M}_i}{Q_i}\right) \iff \\
1+\frac{\bar{M}_i}{Q_i} &= \left(\frac{p_i}{\bar{p}}\right)^{\varepsilon_i}
\end{aligned} \tag{51}$$

Finally, we impose market clearing in village i, which requires that the total quantity consumed in village i is equal to the total quantity it produces less the net quantity it exports to the market:

$$C_i = Q_i + \bar{M}_i - \bar{X}_i.$$

Combined with either equation (50) or (51), the market clearing condition immediately yields the same equation:

$$\frac{C_i}{Q_i} = \left(\frac{p_i}{\bar{p}}\right)^{\varepsilon_i},$$

which is identical to equation (9) in the main text, as claimed.

A.3.5 Farmer cooperative

In the baseline model, we assume that each farmer makes her crop choice taking the prices as given. Here we explore what would occur if a farmer takes into account the effect of her crop choice on prices, e.g. if all the farmers worked together to form a cooperative. In this case, the cooperative will maximize:

$$\max_{\theta_g}(Y_i(\{\theta_{ig}\}))\prod_g\left(\frac{\alpha_{ig}}{p_{ig}(\{\theta_{ig}\})}\right)^{\alpha_{ig}}$$

subject to:

$$\sum_g \theta_{ig} = 1.$$

Recall:

$$p_{ig} = (A_{ig}\theta_{ig}L_i)^{-\frac{1}{1+\varepsilon_i}}(\bar{p}_g)^{\frac{\varepsilon_i}{1+\varepsilon_i}}(\alpha_{ig}Y_i)^{\frac{1}{1+\varepsilon_i}}$$

$$Y_i(s) = \left(\sum_{g\in\mathcal{G}}\alpha_{ig}\left(\frac{\bar{p}_g(s)Q_{ig}(s)}{\alpha_{ig}}\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}\right)^{\frac{1+\varepsilon_i}{\varepsilon_i}}$$

so that we have

$$\begin{aligned}
Z_i &= (Y_i)\prod_g\left(\frac{\alpha_{ig}}{(A_{ig}\theta_{ig}L_i)^{-\frac{1}{1+\varepsilon_i}}(\bar{p}_g)^{\frac{\varepsilon_i}{1+\varepsilon_i}}(\alpha_{ig}Y_i)^{\frac{1}{1+\varepsilon_i}}}\right)^{\alpha_{ig}} \iff \\
Z_i &= Y_i^{\frac{\varepsilon_i}{1+\varepsilon_i}}\prod_g\left(\frac{\alpha_{ig}(A_{ig}\theta_{ig}L_i)^{\frac{1}{1+\varepsilon_i}}}{(\bar{p}_g)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{\alpha_{ig}} \iff \\
Z_i &= \left(\sum_{g\in\mathcal{G}}\alpha_{ig}\left(\frac{\bar{p}_gA_{ig}L_i\theta_{ig}}{\alpha_{ig}}\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}\right)\prod_g\left(\frac{\alpha_{ig}(A_{ig}\theta_{ig}L_i)^{\frac{1}{1+\varepsilon_i}}}{(\bar{p}_g)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{\alpha_{ig}} \iff \\
Z_i &= \left(\sum_{g\in\mathcal{G}}\alpha_{ig}\left(\frac{\bar{p}_gA_{ig}L_i\theta_{ig}}{\alpha_{ig}}\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}\right)\prod_g\left(\frac{\alpha_{ig}(A_{ig}\theta_{ig}L_i)^{\frac{1}{1+\varepsilon_i}}}{(\bar{p}_g)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right)^{\alpha_{ig}}
\end{aligned}$$

Relative to the case where prices are taken as given, the first order conditions of the farmer cooperative are a little more involved. We have:

$$\begin{aligned}
&\frac{\partial Z_i}{\partial\theta_{ig}} = r_i \iff \\
&\theta_{ig} \propto \varepsilon_i\left(\frac{\alpha_{ig}\left(\frac{\bar{p}_gA_{ig}L_i\theta_{ig}}{\alpha_{ig}}\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_{g\in\mathcal{G}}\alpha_{ig}\left(\frac{\bar{p}_gA_{ig}L_i\theta_{ig}}{\alpha_{ig}}\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right) + \alpha_{ig} \implies \\
&\theta_{ig} = \frac{\varepsilon_i\left(\frac{\alpha_{ig}\left(\frac{\bar{p}_gA_{ig}L_i\theta_{ig}}{\alpha_{ig}}\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_{g\in\mathcal{G}}\alpha_{ig}\left(\frac{\bar{p}_gA_{ig}L_i\theta_{ig}}{\alpha_{ig}}\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right) + \alpha_{ig}}{\sum_g\left(\varepsilon_i\left(\frac{\alpha_{ig}\left(\frac{\bar{p}_gA_{ig}L_i\theta_{ig}}{\alpha_{ig}}\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}{\sum_{g\in\mathcal{G}}\alpha_{ig}\left(\frac{\bar{p}_gA_{ig}L_i\theta_{ig}}{\alpha_{ig}}\right)^{\frac{\varepsilon_i}{1+\varepsilon_i}}}\right) + \alpha_{ig}\right)}.
\end{aligned}$$

Recall that when farmers take prices as given, their equilibrium crop choice is given my equation (14):

$$\theta_{ig} = \frac{(A_{ig}\bar{p}_g)^{\varepsilon_i}\alpha_{ig}}{\sum_{h\in\mathcal{G}}(A_{ih}\bar{p}_h)^{\varepsilon_i}\alpha_{ih}},$$

so this demonstrates that the farmer cooperative chooses a different optimal crop allocation. In particular, the elasticity of the relative crop choice to the central market price $\bar{p}_g$ is smaller for the cooperative (where it is bounded above by $\frac{\varepsilon_i}{1+\varepsilon_i}$) than for the price taking farmers (where it is equal to ε_i). Intuitively, the cooperative purposefully restricts the quantity produced of its high value (high $\bar{p}_g$) crops to ensure greater local prices.

A.4 Comparing the model to a traditional arbitrage model

In this subsection, we describe the methodology used to construct panel (c) of Figure 3 that compares the price arbitrage of our model to a traditional arbitrage model where iceberg trade costs are homogeneous. In both cases, consider a "village" (a district, in the data) whose autarkic relationship between prices and yields follows from CES preferences and the market clearing:

$$\log p_{ig}^{aut} = -\frac{1}{\sigma}\log A_{ig} + \frac{1}{\sigma}\log \beta_i \alpha_{ig} + \frac{1}{\sigma}\log \frac{\sum_{h=1}^{G} p_{ih}^{aut} L_i \theta_{ih} A_{ih}}{L_i \theta_{ig} \sum_{h=1}^{G} \alpha_{ih} (p_{ih}^{aut})^{1-\sigma}}, \tag{52}$$

where we omit the state of the world for readability. Suppose that the village is small in size relative to a market (a state, in the data) which has a price $\bar{p}_g$. Note that given estimates of β, α and σ from Section 5.2 and observed yields $\{A_{ig}\}$, allocations $\{\theta_{ig}\}$, and land areas $\{L_i\}$, there exists a unique (to-scale) set of autarkic prices p_{ig}^{aut} that satisfy equation (52).

A standard "kinked" model First consider a standard trade model, where the village is separated from the regional market by an iceberg trade costs $\tau_i > 1$. Then a standard no-arbitrage condition delivers the following relationship between the equilibrium local prices p_{ig}, the given market price, $\bar{p}_g$ and the autarkic local price p_{ig}^{aut}:

$$\log p_{ig} - \log \bar{p}_g = \begin{cases} \log \tau_i & \text{for } \log p_{ig}^{aut} - \log \bar{p}_g > \log \tau_i \\ \log p_{ig}^{aut} - \log \bar{p}_g & \text{for } \log p_{ig}^{aut} - \log \bar{p}_g \in [-\log \tau_i, \log \tau_i] \\ -\log \tau_i & \text{for } \log p_{ig}^{aut} - \log \bar{p}_g < -\log \tau_i \end{cases} \tag{53}$$

The difference between the equilibrium local prices and the regional market prices then are a "kinked" function of the trade costs between the two (when trade occurs and the no-arbitrage equation holds) and the autarkic price p_{ig}^{aut} (when the trade costs are sufficiently high such that no trade occurs).

Our "smooth" model Now consider our framework, where from equation (60) equilibrium prices are:

$$\log p_{ig} = -\frac{1}{\sigma+\varepsilon_i}\log A_{ig} + \frac{\varepsilon_i}{\sigma+\varepsilon_i}\log \bar{p}_{m(i)g} + \frac{1}{\sigma+\varepsilon_i}\log \beta_i \alpha_{ig} + \frac{1}{\sigma+\varepsilon_i}\log \frac{\sum_{h=1}^{G} p_{ih} L_i \theta_{ih} A_{ih}}{L_i \theta_{ig} \sum_{h=1}^{G} \alpha_{ih} (p_{ih})^{1-\sigma}} \tag{54}$$

Combining equations (54) and (52) we can then write the difference between the equilibrium local price and the central market price

$$\log p_{ig} - \log \bar{p}_g = \frac{\sigma}{\sigma+\varepsilon_i}(\log p_{ig}^{aut} - \log \bar{p}_g) \tag{55}$$

Hence, unlike equation (53), equation (55) states that the local price relative to the market price should smoothly vary with the difference with the local autarkic price relative to the market price. Note that equations (53) and (55) coincide with each other under autarky ($\varepsilon_i = 0$, $\tau_i = \infty$) or free trade ($\varepsilon_i = \infty$, $\tau_i = 1$).

Empirical Strategy The basic idea is to compare the model fit of equations (53) and (55). In order to do so, we have to first solve a few implementation issues. First, as autarkic prices are only identified up to scale, we add a location specific constant c_i to both models, so that the standard "kinked" model becomes:

$$\log p_{ig} - \log \bar{p}_g = \begin{cases} \log\tau_i + c_i & \text{for } \log p_{ig}^{aut} - \log \bar{p}_g > \log\tau_i + c_i \\ \log p_{ig}^{aut} - \log \bar{p}_g & \text{for } \log p_{ig}^{aut} - \log \bar{p}_g \in [-\log\tau_i + c_i, \log\tau_i + c_i], \\ -\log\tau_i + c_i & \text{for } \log p_{ig}^{aut} - \log \bar{p}_g < -\log\tau_i + c_i \end{cases} \tag{56}$$

and our "smooth" model becomes:

$$\log p_{ig} - \log \bar{p}_g = \frac{\sigma}{\sigma + \varepsilon_i} (\log p_{ig}^{aut} - \log \bar{p}_g) + c_i \tag{57}$$

The advantage of the additional constant is that both models are now ensured to have an R-squared statistic between 0 and 1, which will be our statistic for goodness of fit. The second issue is how to measure prices. Because of the potential endogeneity of yields to prices, as in Section 5.2, we use rainfall-predicted yields to construct a rainfall-predicted measure of autarkic prices from equation (52). Also as in Section 5.2, we measure the market price as the quantity weighted average price in all districts within a state except the one being examined to avoid mechanical correlations between market and local prices.

Estimation and Results As in Section 5.2, we allow the trade costs to vary by district-decade. To do so, we conduct the estimation of both the standard "kinked" model and our "smooth" model separately for each district-decade combination. The estimation for our smooth model is simply a linear regression of the log district price (relative to the state leave-one-out price) on the log rainfall-predicted autarkic price (again relative to the state leave-one-out price). The kinked model is similar, but uses a non-linear least squares routine to capture the kinks present in equation (56), where we constrain $\log \tau_i \geq 0$. Note that in both cases, we are estimating just two parameters using the same left hand side and right hand side variables: the constant c_i and a measure of trade costs ($\log\tau_i$ for the standard model and $\frac{\sigma}{\sigma+\varepsilon_i}$ for our model).

We compare the residual sum of squares from both models and normalize this by the variance of the dependent variable to create a comparable version of the R^2 for comparison. Panel (c) of Figure 3 plots the cumulative density of the fits for the two models. The smooth model has a better fit than the kinked model in nearly 71% of all district-decade combinations. The mean R^2 of the smooth and kinked model runs are 0.11 and 0.15 respectively.

A.5 References

Campbell, John Y and Luis M Viceira, *Strategic asset allocation: portfolio choice for long-term investors*, Oxford University Press, 2002.

Mas-Colell, A., M.D. Whinston, and J.R. Green, *Microeconomic theory*, New York: Oxford University Press, 1995.