Econometrica Supplementary Material

SUPPLEMENT TO "INFORMATION FRICTIONS IN TRADE"
(*Econometrica*, Vol. 82, No. 6, November 2014, 2041–2083)

BY TREB ALLEN

This Supplemental Material comprises eight appendixes: Appendix A proves the existence and uniqueness of equilibrium prices; Appendix B presents a detailed description of the data used in the paper; Appendix C considers four alternative search frameworks; Appendix D considers three alternative complete information frameworks that may be consistent with the observe empirical patterns; Appendix E presents five extensions of the model mentioned in the main text; Appendix F presents additional derivations of the model mentioned in the text; Appendix G presents additional results related to the structural estimation; Appendix H presents additional tables mentioned in the main text.

APPENDIX A: EXISTENCE AND UNIQUENESS OF EQUILIBRIUM

IN THIS APPENDIX, I prove the existence and uniqueness of equilibrium prices. I first prove existence and then prove uniqueness.

A.1. *Existence*

The proof of existence comprises four steps. In the first step, I characterize the total supply to a region as a function of prices in all regions. In the second step, I define a function whose fixed point is the set of equilibrium prices. In the third step, I present a lemma that guarantees equilibrium prices occur in a compact set bounded by autarkic prices. In the fourth step, I apply the Schauder fixed point theorem to prove existence.

Step 1. In this step, I characterize the total supply of rice to region j as a function of its price and all other prices. Total domestic supply of rice Q_{jj} is

$$Q_{jj} = \max\left\{\theta_j A_j M_j \int_1^{\varphi^*(p_j)} \varphi^{-\theta_j}\, d\varphi, 0\right\} = \max\left\{\frac{\theta_j}{\theta_j - 1} A_j M_j \left(1 - \left(\frac{K_j(p_j)}{f_j}\right)^{\theta_j - 1}\right), 0\right\}. \tag{19}$$

Since $K_i(\cdot)$ is strictly decreasing and $\theta_j > 1$, Q_{jj} is continuous and strictly increasing in p_j. This is intuitive: the greater is the home price, the lower is the value of search, causing more farmers to sell domestically rather than export. Furthermore, Q_{jj} is continuous and weakly decreasing in p_i for all $i \neq j$, as increases in p_i will increase the value of search if $\frac{p_i}{\tau_{ji}} \geq p_j$, causing domestic producers to export more to i (if $\frac{p_i}{\tau_{ji}} < p_j$, then no exports to i occur so that changes in p_i have no effect on domestic supply).

 DOI: 10.3982/ECTA10984

Total imports I_j require summing over all exporters to j using equations (7) and (9):

$$I_j = \sum_{i \neq j} \mathbf{1}\left\{\frac{p_j}{\tau_{ij}} > p_i\right\} A_i M_i \theta_i f_i^{1-\theta_i} s_{ij} \int_{p_i}^{p_j/\tau_{ij}} K_i(p)^{\theta_i - 2}\, dp \geq 0. \tag{20}$$

Note that I_j is strictly increasing (and continuous) in p_j. This is also intuitive: a greater price in region j will cause more farmers searching j to sell there. In addition, I_j is weakly decreasing (and continuous) in p_i for all $i \neq j$ since an increase in p_i will reduce the quantity of imports arriving from i as long as $p_i < \frac{p_j}{\tau_{ij}}$ (if $p_i \geq \frac{p_j}{\tau_{ij}}$, then an increase in p_i has no direct effect on I_j since i is not exporting to j).[42]

Recall that there are a "large" number of regions, that is, a continuum or a countable infinity of regions. Denote the set of regions by S. (In the uniqueness proof below, I restrict attention to compact S.) Let $\mathbf{p}: S \to \mathbb{R}_+$ denote the function (or infinite vector) of prices for all regions, that is, $\mathbf{p}(i) \equiv p_i$ for all $i \in S$. Let $\mathbf{S}$ denote the set of all functions $\mathbf{g}: S \to \mathbb{R}_+$, that is, $\mathbf{S}$ is a function space. Define the operator $\mathbf{R}: \mathbf{S} \to \mathbf{S}$ as

$$\mathbf{Rp}(j) \equiv Q_{jj}(\mathbf{p}) + I_j(\mathbf{p}) \quad \forall j \in S,$$

that is, for any price function, $\mathbf{R}$ yields a function giving the total supply of rice to each region. From above, $\mathbf{Rp}(j)$ is continuous in all elements of $\mathbf{p}$, strictly increasing in $\mathbf{p}(j)$, and weakly decreasing in all other elements $\mathbf{p}(i)$ for $i \neq j$ (strictly decreasing as long as i exports to j or j exports to i). Since $\mathbf{R}$ is constructed using equations (7) and (9), it satisfies the first property of equilibrium in Section 3.5.

Step 2. In this step, I characterize a function whose fixed point is the set of equilibrium prices. Define the operator $\mathbf{G}: \mathbf{S} \to \mathbf{S}$ as

$$\mathbf{Gp}(j) \equiv D_j \circ \mathbf{Rp}(j) = D_j\big(Q_{jj}(\mathbf{p}) + I_j(\mathbf{p})\big) \quad \forall j \in S,$$

that is, $\mathbf{G}$ yields the set of prices everywhere that would result from the inverse demand function in each region when each region is supplied with $\mathbf{Rp}(j)$. Since $D_j(\cdot)$ is continuous for all r and $\mathbf{Rp}(j)$ is continuous for all elements of $\mathbf{p}$, $\mathbf{G}$ is continuous in all elements of $\mathbf{p}$ as well. Since $\mathbf{Rp}$ satisfies the first property of equilibrium, a set of prices $\mathbf{p}^*$ is a set of equilibrium prices if and only if $\mathbf{Gp}^* = \mathbf{p}^*$, that is, the prices resulting from the inverse demand function given

[42]Continuity is not affected by the presence of the indicator function, as when $\frac{p_j}{\tau_{ij}} = p_i$, $\sum_{l=1}^{L} \frac{K_i(p_{l-1}^{ij})^{\theta_i - 1} - K_i(p_l^{ij})^{\theta_i - 1}}{1 - F_{p/\tau}^i(p_{l-1}^{ij})} = 0$. An increase in p_i also indirectly affects I_j by increasing the value of search of other regions $k \neq i$ that export to both i and j; with a large number of regions in the world, however, these indirect effects can be safely ignored.

supply $\mathbf{Rp}^*$ are the same prices that yield the supply. It remains to show that such a fixed point exists.

Step 3. In this step, I show that no equilibrium price can be greater than the maximum autarkic price, which I use to create a compact set to apply the Schauder fixed point theorem. I do so in the following lemma.

LEMMA 1: *Define* $\tilde{p}^{\max} \equiv \sup_{j\in S} D_j(\frac{\theta_j}{\theta_j-1}A_jM_j)$ *to be the maximum autarky price across all regions and let* $b > \tilde{p}^{\max}$ *be any scalar greater than the maximum autarkic price. Then for all* $j \in S$ *and* $\mathbf{p} \in \{\mathbf{S} | p(j) = b$ *and* $p(i) \in [0, b]\ \forall i \in S\}$, $\mathbf{Gp}(j) < b$.

PROOF: I prove the lemma by contradiction. Suppose not, that is, suppose there exists a $j \in S$ and $\mathbf{p} \in \{\mathbf{S} | p(j) = b$ and $p(i) \in [0, b]\ \forall i \in S\}$ such that $\mathbf{Gp}(j) \geq b$. By the definition of $\tilde{p}^{\max}$, we have

$$\begin{aligned}(21)\qquad & \mathbf{Gp}(j) \geq b > \tilde{p}^{\max} \geq D_j\left(\frac{\theta_j}{\theta_j-1}A_jM_j\right) \\ & \Rightarrow \quad \mathbf{Gp}(j) > D_j\left(\frac{\theta_j}{\theta_j-1}A_jM_j\right).\end{aligned}$$

Since $D_j(\cdot)$ is a strictly decreasing function, equation (21) implies

$$(22)\qquad \mathbf{Rp}(j) < \frac{\theta_j}{\theta_j-1}A_jM_j.$$

Substituting $\mathbf{Rp}(j) = \frac{\theta_j}{\theta_j-1}A_jM_j(1-(\frac{K_j(p_j)}{f_j})^{\theta_j-1}) + I_j(\mathbf{p})$ into equation (22) yields

$$(23)\qquad I_j(\mathbf{p}) < \frac{\theta_j}{\theta_j-1}A_jM_j^{\theta_j-1}\left(\frac{K_j(p_j)}{f_j}\right)^{\theta_j-1}.$$

Equation (23) is intuitive: for j to have a price at least as great as b, it must be that its total domestic supply is less than its domestic supply in autarky (since the autarkic price is lower). For this to be the case, it must have exported more than it imported. Since $p_j = b$ and $p_i \in [0, b]\ \forall i \in S$, $p_j \geq \frac{p_i}{\tau_{ji}}\ \forall i \in S$, that is, region j must have a higher domestic price than the price net of transportation costs that its farmers could receive in any other region. From equation (6), this implies that $K_j(p_j) = 0$, that is, there is zero value of search for farmers in region j since the price in region j is at least as great as the price anywhere else. Hence no farmer in region j exports, so there must be negative imports, that is, $I_j(\mathbf{p}) < 0$. This, of course, is impossible since, from equation (20), imports are nonnegative, that is, $I_j(\mathbf{p}) \geq 0$. As a result, there is contradiction. *Q.E.D.*

Step 4. In this step, I show that the Schauder fixed point theorem guarantees the existence of an equilibrium set of prices. From Step 2, $\mathbf{G}$ is a continuous mapping from the function space $\mathbf{S}$ to itself, and a price function $\mathbf{p}^*$ is an equilibrium price function if and only if $\mathbf{G}\mathbf{p}^* = \mathbf{p}^*$. Let $b > \tilde{p}^{\max}$ be a scalar greater than the maximum autarkic price and define $\tilde{\mathbf{S}} \equiv \{\mathbf{S} | p(i) \in [0, b]\ \forall i \in S\}$ to be the (compact) space of functions bounded between 0 and b. Define the operator $\tilde{\mathbf{G}} : \tilde{\mathbf{S}} \to \tilde{\mathbf{S}} \equiv \tilde{\mathbf{G}}\mathbf{p}(j) = \min\{b, \mathbf{G}\mathbf{p}(j)\}$. Since $\tilde{\mathbf{G}}$ is a continuous mapping from a compact and convex set to itself, by the Schauder fixed point theorem, there exists a $\mathbf{p}^* \in \tilde{\mathbf{S}}$ such that $\tilde{\mathbf{G}}\mathbf{p}^* = \mathbf{p}^*$.

It remains to show that $\mathbf{G}\mathbf{p}^* = \mathbf{p}^*$. Since $\tilde{\mathbf{G}}\mathbf{p}^* = \mathbf{p}^*$, from the definition of $\tilde{\mathbf{G}}$, we have $\mathbf{p}^*(j) = \min\{b, \mathbf{G}\mathbf{p}(j)\}$ for all $j \in S$. For all $j \in S$, it cannot be the case that $\mathbf{p}^*(j) = b$ since Lemma 1 guarantees $\mathbf{G}\mathbf{p}(j) < b$. Hence it must be the case that $\mathbf{p}^*(j) = \mathbf{G}\mathbf{p}^*(j)$ for all $j \in S$ or, equivalently, $\mathbf{G}\mathbf{p}^* = \mathbf{p}^*$, as required.

A.2. *Uniqueness*

An equilibrium function of prices $\mathbf{p}_1^*$ is said to be unique to scale if, for all $\mathbf{p} \in \mathbf{S}$ such that $\mathbf{G}\mathbf{p} = \mathbf{p}$, there exists a constant $\alpha \in \mathbb{R}_{++}$ such that $\mathbf{p}_1^* = \alpha\mathbf{p}$. In what follows, I prove that there exists a constant $t > 0$ such that if, for all $i \in S$ and $j \in S$, $\tau_{ij} < 1 + t$, then the equilibrium function of prices $\mathbf{p}_1^*$ is unique to scale. The proof itself follows three steps. In the first step, I show that normalizing the prices has no effect on the equilibrium. In the second step, I present a lemma showing that when trade costs are sufficiently small, an increase in the price elsewhere, all else equal, decreases the amount consumed in a region. In the third step, I prove uniqueness using a simple contradiction argument motivated by the proof presented in Mas-Colell, Whinston, and Green (1995, p. 613).

Step 1. Prior to proving uniqueness, I first note that normalizing the prices by a scalar does not affect the equilibrium. Formally, if $\mathbf{G}(\mathbf{p}) = \mathbf{p}$ and $\alpha \in \mathbb{R}_{++}$, then $\mathbf{G}(\alpha\mathbf{p}) = 0$, that is, $\mathbf{G}$ is homogeneous of degree 0. To see this, note that from equation (6), $K(\cdot)$ is homogeneous of degree 0. Since scaling prices by α also entails scaling the fixed costs of search by α, equations (19) and (20) immediately imply that $\mathbf{r}(\mathbf{p})$ is homogeneous of degree 0. Since the demand function is also homogeneous of degree 0 (since multiplying all prices by a constant along with wealth does not change the utility maximization problem), $\mathbf{G}$ is homogeneous of degree 0.

Step 2. I now prove the following lemma that shows that the total quantity consumed in region j will be strictly lower if the relative price in at least one other region is greater and the relative price in no other region declines for sufficiently small trade costs.

LEMMA 2: *Consider price functions* $\mathbf{p}_1 : S \to \mathbb{R}_+$ *and* $\mathbf{p}_2 : S \to \mathbb{R}_+$ *such that* (i) *there exists an* $j \in S$ *such that* $\mathbf{p}_1(j) = \mathbf{p}_2(j)$ *and* (ii) *for all* $i \in S$, $\mathbf{p}_1(i) \leq \mathbf{p}_2(i)$, *with the equality strict for at least one $i \in S$. Then there exists some $t > 0$ such that if, for all $i \in S$ and $j \in S$, $\tau_{ij} < 1 + t$:*

$$Q_{jj}(\mathbf{p}_1) + I_j(\mathbf{p}_1) > Q_{jj}(\mathbf{p}_2) + I_j(\mathbf{p}_2).$$

PROOF: Recall that

$$\begin{aligned} &Q_{jj}(\mathbf{p}) + I_j(\mathbf{p}) \\ &\quad = \max\left\{\frac{\theta_j}{\theta_j - 1} A_j M_j \left(1 - \left(\frac{K_j(p_j)}{f_j}\right)^{\theta_j - 1}\right), 0\right\} \\ &\qquad + \sum_{i \neq j} \mathbf{1}\left\{\frac{p_j}{\tau_{ij}} > p_i\right\} A_i M_i \theta_i f_i^{1-\theta_i} s_{ij} \int_{p_i}^{p_j/\tau_{ij}} K_i(p)^{\theta_i - 2}\, dp. \end{aligned}$$

Let $i^* \in S$ denote one (of possibly many) locations such that $\mathbf{p}_1(i^*) < \mathbf{p}_2(i^*)$. Define $\varepsilon \equiv \mathbf{p}_2(i^*) - \mathbf{p}_1(i^*)$ to be the difference between the two price vectors. If $\frac{\mathbf{p}_2(i^*)}{\tau_{ji^*}} > \mathbf{p}_2(j)$ (so that j exports to i^* at the higher price in i^*) or if $\frac{\mathbf{p}_2(j)}{\tau_{i^*j}} > \mathbf{p}_1(i^*)$ (so that j imports from i^* at the lower price in i^*), then it is immediately evident from the above expression that $Q_{jj}(\mathbf{p}_1) + I_j(\mathbf{p}_1) > Q_{jj}(\mathbf{p}_2) + I_j(\mathbf{p}_2)$, that is, increasing the price in i^* reduces the quantity consumed in j as long as i^* and j trade with each other. Define $t \equiv \max\{\frac{1-\mathbf{p}_1(i^*)}{\mathbf{p}_1(i^*)}, \mathbf{p}_1(i^*) - 1 + \varepsilon\} > 0$. Then it is straightforward to show that if, for all $i \in S$ and $j \in S$, $\tau_{ij} < 1 + t$, then i^* and j trade with each other: Suppose $\frac{1-\mathbf{p}_1(i^*)}{\mathbf{p}_1(i^*)} \geq \mathbf{p}_1(i^*) - 1 + \varepsilon$. Then

$$\frac{1}{\mathbf{p}_1(i^*)} = 1 + t > \tau_{i^*j} \quad \Longleftrightarrow \quad \frac{\mathbf{p}_2(j)}{\tau_{i^*j}} > \mathbf{p}_1(i^*),$$

that is, i^* will export to j. Conversely, suppose that $\frac{1-\mathbf{p}_1(i^*)}{\mathbf{p}_1(i^*)} \leq \mathbf{p}_1(i^*) - 1 + \varepsilon$, so that

$$\mathbf{p}_2(i^*) = 1 + t > \tau_{ji^*} \quad \Longleftrightarrow \quad \frac{\mathbf{p}_2(i^*)}{\tau_{ji^*}} > \mathbf{p}_2(j),$$

that is, j will export to i^*. *Q.E.D.*

Step 3. Finally, I use a simple contradiction argument to prove that the equilibrium function of prices $\mathbf{p}_1^*$ is unique to scale for sufficiently small trade costs. Suppose not. Then there exist two functions $\mathbf{p}_1^*$ and $\mathbf{p}_2^*$ such that $\mathbf{G}\mathbf{p}_1^* = \mathbf{p}_1^*$ and $\mathbf{G}\mathbf{p}_2^* = \mathbf{p}_2^*$ such that there does not exist a constant $\alpha \in \mathbb{R}_{++}$ such that $\mathbf{p}_1^* = \alpha \mathbf{p}_2^*$. Without loss of generality, establish the following ordering $\succeq$ on the set of locations S: for any $i \in S$ and $j \in S$, $i \succeq j$ if and only if $\frac{\mathbf{p}_1^*(i)}{\mathbf{p}_2^*(i)} \geq \frac{\mathbf{p}_1^*(j)}{\mathbf{p}_2^*(j)}$. Assume S is compact, so that there exists an $i^* \in S$ such that for all $j \in S$, $i^* \succeq j$. Because

$\mathbf{G}$ is homogeneous of degree 0, without loss of generality I can normalize $\mathbf{p}_1^*$ by $\mathbf{p}_1^*(i^*)$ and $\mathbf{p}_2^*$ by $\mathbf{p}_2^*(i^*)$. As a result, $\mathbf{p}_1^*(i^*) = \mathbf{p}_2^*(i^*) = 1$ and, for all $j \in S$, $\mathbf{p}_2^*(j) \geq \mathbf{p}_1^*(j)$, with the inequality strict for at least one $j \in S$ since there does not exist a constant $\alpha \in \mathbb{R}$ such that $\mathbf{p}_1^* = \alpha \mathbf{p}_2^*$.

Consider region i^*. Since $\mathbf{G}\mathbf{p}_1^* = \mathbf{p}_1^*$ and $\mathbf{G}\mathbf{p}_2^* = \mathbf{p}_2^*$, and $\mathbf{p}_1^*(i^*) = \mathbf{p}_2^*(i^*) = 1$, from the definition of $\mathbf{G}$ we have

$$D_{i^*}\left(Q_{i^*i^*}\left(\mathbf{p}_1^*\right) + I_{i^*}\left(\mathbf{p}_1^*\right)\right) = D_{i^*}\left(Q_{i^*i^*}\left(\mathbf{p}_2^*\right) + I_{i^*}\left(\mathbf{p}_2^*\right)\right),$$

which, from the monotonicity of $D_{i^*}(\cdot)$, implies

$$Q_{i^*i^*}\left(\mathbf{p}_1^*\right) + I_{i^*}\left(\mathbf{p}_1^*\right) = Q_{i^*i^*}\left(\mathbf{p}_2^*\right) + I_{i^*}\left(\mathbf{p}_2^*\right).$$

However, the lemma above shows that for sufficiently small trade costs,

$$Q_{i^*i^*}\left(\mathbf{p}_1^*\right) + I_{i^*}\left(\mathbf{p}_1^*\right) > Q_{i^*i^*}\left(\mathbf{p}_2^*\right) + I_{i^*}\left(\mathbf{p}_2^*\right),$$

which is a contradiction, thereby proving the uniqueness to scale of the equilibrium function of prices $\mathbf{p}_1^*$.

APPENDIX B: DATA DESCRIPTION

This appendix describes in detail the data used.

B.1. *Trade Flows*

Beginning in 1995, the National Statistics Office of the Philippines (NSO) has collected data on the domestic trade in the Philippines using the Domestic Trade Statistics System (DOMSTAT). DOMSTAT covers the flow of commodities over water, air, and rail, of which more than 99% of both the value and the quantity of trade consistently occurs over water. Statistics on trade flows over water are derived from the cargo manifests collected by the Philippines Port Authority (PPA) and contain information on the port of origin, the port of destination, the description of the commodity, the quantity shipped, the value shipped, and (in most years) the total freight costs.

With financial support from the Yale University Economic Growth Center and logistic support from the Business and Services Statistics Division of the NSO, I was able to acquire the annual aggregates of the bilateral port-to-port domestic trade data.[43] For every commodity (classified at the SITC five-digit level), these data included the quantity, value, and freight cost of all shipments from each port of origin to each destination port. I aggregated the port data to the province level to create a data set of province-to-province bilateral trade

[43] In an apparent error, the data also included information on trade flows in the fourth quarter alone, which I used to construct Figure 3.

flows. To create a measure of the bilateral freight costs in the standard iceberg cost form, I calculated the mean (nonmissing) freight cost as a fraction of the total value of shipments of a commodity in a origin–destination province pair in a particular year.[44]

I then identified 51 agricultural commodities using the SITC classification codes, which constitute the sample of analysis for Figure 3. Of these commodities, 10 could be matched with wholesale price and production data. Statistics about these 10 commodities are presented in Table I. As is evident, these commodities constitute a large majority of the Philippines agricultural sector, comprising 65% of the total value of agricultural output and 60% of the total agricultural area.[45] These commodities constitute the primary sample of analysis.

In total, I observe 4,332 nonzero province-origin–province-destination–year–commodity trade flows (of 32,922 potential trade pairs) spanning 1995 through 2009 where the wholesale price is observed in both the origin and destination province, and where production data are available for the origin province.[46] These observations are roughly evenly split between years, ranging from 152 observations in 2009 to 424 observations in 1996. The sample includes 40 origin provinces and 50 destination provinces.

B.2. *Prices*

Wholesale prices of agricultural commodities are collected in 66 markets in 55 provinces throughout the Philippines by the Integrated Agricultural Marketing Information System (AGMARIS). For each commodity in each market in each quarter, respondents are stratified according to the type of trader (e.g., large distributor, provincial assembler, etc.) and assigned into two or three similar groups. In each group, five respondents are interviewed each collection day. The statistics are then aggregated to the commodity–province–

[44]Unfortunately, the freight data are missing for a large (38%) fraction of observations. Furthermore, smaller shipments are substantially less likely to report freight costs. As a result, I refrain from using freight costs directly in regressions of bilateral trade flows. The observed freight costs are helpful, however, as indicators of the magnitude of overall transportation costs.

[45]The major agricultural commodities not included in the data set are coconut and banana, which are produced primarily for export so that domestic trade is limited (BAS (2011)).

[46]In particular, exports from Manila are excluded from the data set because Manila does not produce any agricultural commodity. Its observed exports likely come from provinces in northern Luzon that ship their agricultural commodities overland to Manila. Since the provinces in northern Luzon do not export over water, the sample of origin provinces does not include these provinces. These provinces and Manila, however, remain in the data set as potential destinations for commodities produced elsewhere. Since the analysis relies on observing export shares and destination prices, the estimation of trade costs elsewhere is unaffected by excluding these provinces.

month level and made publicly available on the CountrySTAT Philippines website (http://countrystat.bas.gov.ph). Wholesale prices (rather than farm gate or retail prices) are chosen as the relevant prices for empirical analysis as they are the prices that exporters receive when selling produce to other provinces.

B.3. *Production*

Data on agricultural commodity production come from two surveys administered by the Philippines Bureau of Agricultural Statistics (BAS): the Palay and Corn Production Survey (RCPS) and the Crops (Other than Rice and Corn) Production Survey (CrPS). Both the RCPS and CrPS surveys are administered quarterly in each province in the Philippines. For the RCPS, households are sampled from the largest producing barangay in each municipality and half the remaining barangays. For the CrPS, between three and five farmer/producers are interviewed in the major producing municipalities in each province in each quarter. In both surveys, respondents are asked the volume of production and the area harvested/planted for each commodity. The data are then aggregated to the commodity–province–quarter level (for rice and corn) and commodity–province–year level (for all other crops), and made publicly available on the CountrySTAT Philippines website (http://countrystat.bas.gov.ph). Because the production figure include production by subsistence farmers who do not sell their production in the market, I multiply the observed production by the fraction of land cultivated by nonsubsistence farmers (defined as farmers with more than 1 ha of land) using the 1991 agricultural census, where the fraction is calculated at the commodity–province level.

B.4. *Rainfall*

The Philippine Atmospheric, Geophysical and Astronomical Services Administration (PAGASA) collects daily rainfall data from 47 weather stations located throughout the Philippines (the location of the stations is presented in Figure 8). Using these data, I constructed provincial level daily rainfall measures using an distance weighting technique suggested by Dirks, Hay, Stow, and Harris (1998), that is, $r_{it} = \sum_s w_{is} r_{st}$, where r_{st} is the rainfall measured at station s at time t, r_{it} is the estimated rainfall in province i at time t, and $w_{is} \equiv \frac{dist_{is}^{-\alpha}}{\sum_s dist_{is}^{-\alpha}}$ is a weighting factor depending on the distance between weather station s and province i. I chose the parameter α to maximize the R^2 of a regression of rainfall at each station on the predicted rainfall at the station; it turns out this is maximized at $\alpha = 1$, which is the simple inverse distance weighting method commonly used.

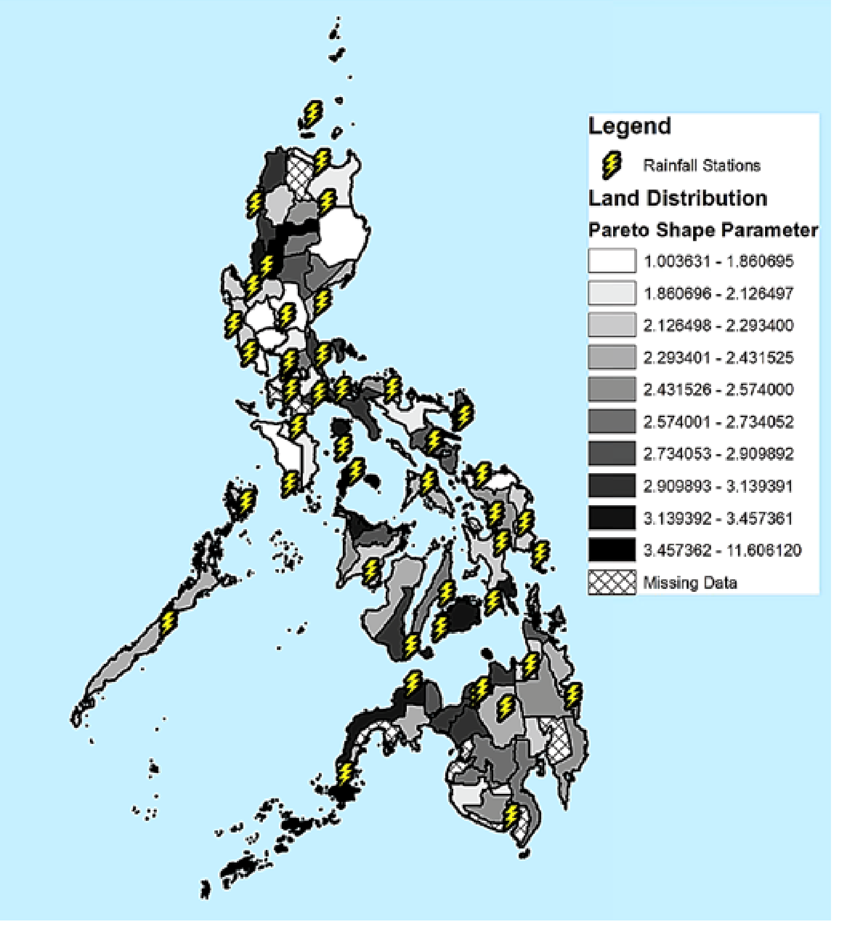

FIGURE 8.—Rainfall stations and land distribution. This figure shows the estimated land distribution shape parameter θ_i for each province. The shading corresponds to each decile and is darker for larger values of θ_i (indicating a greater proportion of small farmers). The figure also indicates the location of the 47 weather stations used to construct the idiosyncratic weather shocks.

B.5. *Cell Phones*

Every cell phone tower in the Philippines must be registered with the National Telecommunications Commission of the Philippines (NTC). Through the substantial efforts of researchers at the Asia Pacific Policy Center, the registration records of the universe of cell phone towers were digitized. As a result, for every cell phone tower in the Philippines built prior to 2010, I observe the province and municipality in which it was built, the day it went into operation, and the technology it used. Using these data, I construct a measure of the number of civilian cell phone towers in operation in every province in every month between 1990 and 2009.[47]

B.6. *Land Distribution*

In February 1992, the Philippines National Statistics Office (NSO) conducted a census of agriculture. This census comprised all plots (greater than 0.1 ha in size) in a randomly chosen 50% of all barangays (outside the National Capital Region) in the nation. The census recorded, among other characteristics, the size of each plot, the crop produced on the plot, and the owner of each plot. With financial support from the Yale University Economic Growth Center and logistic support from the NSO, I have acquired the raw data from the census. For each farmer f in province i producing commodity c recorded in the census, I calculated the observed total land area under his/her cultivation used to cultivate that commodity L_{fic}. To estimate the Pareto distribution shape parameter θ_{ic}, I first restrict the sample to only those farmers who cultivated at least 1 ha of land so as to exclude subsistence farmers who do not sell their produce to the market.[48] I then define θ_{ic} to be a parameter that maximizes the likelihood of observing $\{L_{fic}\}_{f=1}^{F}$ under a Pareto distribution:

$$\theta_{ic} = \arg\max_{\theta>1} \sum_{f=1}^{F} \ln\bigl(\theta L_{fic}^{-(\theta+1)}\bigr) \quad \Rightarrow \quad \theta_{ic} = \left(\frac{1}{F}\sum_{f=1}^{F} \ln L_{fic}\right)^{-1}.$$

Hence, θ_{ic} is simply the inverse of the mean of the log value of land cultivated for commodity c by all farmers in province i.

The estimated shape parameter varies substantially across crops and provinces, with a mean of 3.16 and a standard deviation of 1.92 (see Figure 8 for a map of the distribution of shape parameters). Figure 9 depicts the relationship between the observed landholding distribution and the Pareto distribution with estimated θ_{ic} for two provinces.[49] Two points are evident from the figure. First, the Pareto distribution of landholdings appears to be a good approximation of the true distribution of landholdings. Second, a larger value of θ_{ic} (in this case, for Bohol) is associated with a greater concentration of land among smaller farmers.

[47] I focus only on 900 MHz, 1,800 MHz, 2 GHz, and 3 GHz towers, as these are the frequencies used by for standard global system for mobile (GSM) phones. Towers broadcasting at other frequencies are used primarily for noncivilian purposes (e.g., military, ship communications, etc.).

[48] Including all farmers systematically reduces the magnitude of the estimated shape parameter; however, the parameters estimated for the restricted and unrestricted sample are highly correlated.

[49] Other provinces generate similar figures.

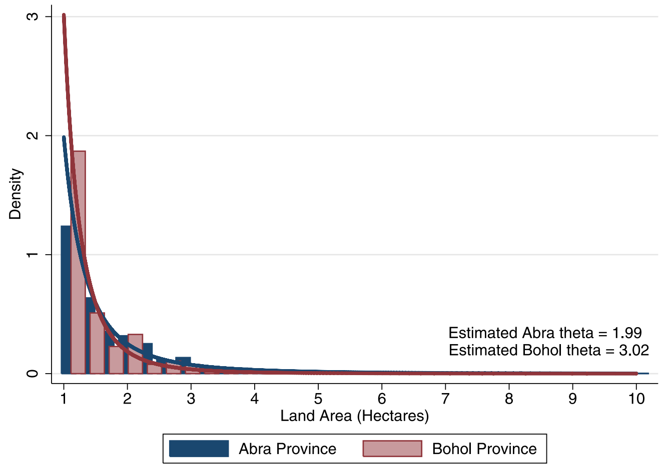

FIGURE 9.—Estimating the Pareto shape parameter for the land distribution. The bars show the observed distribution of palay (rice) landholdings from the 1991 Agricultural Census for Abra and Bohol provinces, respectively. The lines show the implied Pareto distribution using the maximum likelihood shape parameter.

B.7. *Farmer Sales*

Data on individual farmer–trader transactions come from the Farm Price Survey (FPS) administered by the Philippines Bureau of Agricultural Statistics (BAS). The FPS provides the basis for estimation of the farm gate agricultural commodity prices published by the BAS and made available online at CountrySTAT Philippines (http://countrystat.bas.gov.ph). The FPS is administered in the last 10 days of every month in every province in the Philippines. For each commodity produced in each province, at least five farmers in each of the top five producing municipalities in the province are interviewed about the quantity of the commodity they sold in the past month, the price they received, and the total amount (if any) of freight/transportation costs they incurred.

With financial support from the Yale University Economic Growth Center and substantial logistic support from the BAS, the individual transaction records from the FPS for all of the Philippines and all commodities were digitally compiled for years 2000–2009.[50] These data include information on roughly 2.3 million unique farmer–trader transactions in roughly 134,000 province–commodity–year–month markets.

[50]Because of technological limitations, records of transactions in years prior to 2000 were unavailable.

B.8. *Ethnicity and Religion Data*

Data on the ethnicity and religion composition of provinces come from the Philippines 2000 Census of Population and Housing, which covers the entire population in the Philippines. Each individual is classified into one of 95 different religions and one of 148 different ethnicities. I first tabulate the share of the population in a province that belongs to each religion and ethnic group. I then use these shares to calculate the probability that a randomly selected individual in province i has the same ethnicity as individuals in province j, which I use as a measure of the similarity of ethnicities between the two provinces. The variable is the sum over all ethnicities of the product of the share of individuals in province i and j who are a particular ethnicity, that is, $\sum_e s_{ie}s_{je}$, where s_{ie} is the fraction of individuals in province i who are ethnicity e. An identical methodology is used to construct a measure of the similarity of religions between two provinces.

B.9. *Marketing Cost Surveys*

The Philippines Bureau of Agricultural Statistics has published detailed marketing cost structure reports for a number of major crops (BAS (2002a, 2002b, 2003, 2007a, 2007b)). In each report, the total costs of bringing a crop from the farmer to the consumer are calculated for a number of provinces. The average fraction of total marketing costs due to transportation costs across all sampled provinces is 28% for garlic, 29% for onion, 19% for tomato, 27% for potato, and 48% for corn. (Other costs include labor, materials, depreciation, and other operating expenses.) To the extent that observed freight costs fail to capture the transportation costs to and from the port, they underestimate total transportation costs. To the extent that observed freight costs include labor and operating expenses during transport, however, they overestimate total transportation costs. I assume that these two biases cancel out, so that observed freight costs in the trade data should account for roughly 20–50% of total costs associated with shipping a good.

APPENDIX C: ALTERNATIVE SEARCH FRAMEWORKS

In this appendix, I present four alternative models of information frictions in trade. I compare the theoretical predictions of these alternative frameworks to those of the model in the paper so as to assess how strongly the predictions of the model rely on the particular modeling assumptions.

C.1. *Directed Search*

This framework is based on the "directed" search literature. Producers choose in which destination to attempt to sell their produce but face the possibility of being unable to successfully match with a consumer. This match probability depends in part on the number of other producers also searching the island. I show that this framework does not necessarily yield the prediction that larger producers on average sell to destinations with higher prices; neither does the framework easily incorporate trade costs, as it cannot be the case that producers from different origins are indifferent across the same destinations.

Setup

There are a large number of islands $i \in S$. Each island is inhabited by a representative consumer and a heterogeneous measure of profit-maximizing producers. Producers produce a quantity φ of a homogeneous commodity (rice), where φ (landholdings) is assumed to vary across producers according to cumulative distribution function $F_i(\varphi)$.

Producers engage in "directed" search to sell their produce.[51] In particular, the representative consumer on each island posts a price at which he is willing to purchase the produce, and producers choose which island to search. Let p_{ij} denote the price that the representative consumer in island j offers to producers from island i. There are search frictions in the sense that there exists a probability that a producer searching a particular island may not successfully transact with the consumer. As is common in the literature, I assume that the probability a producer successfully transacts depends on the number of other producers who search the island. To make the problem as simple as possible, suppose that the probability a producer from i is able to successfully complete a transaction after searching island j is $\pi_\varphi(n_{ij})$, where n_{ij} is the density of producers from i who search island j. In what follows, I assume that $\frac{\partial \pi_\varphi}{\partial n} < 0$, that is, the greater is the number of producers searching a destination, the lower is the probability that any individual producer is successfully matched.[52]

As is common in the literature, I assume the payoff to the representative consumer is linear (note that this assumption is not required in the main text). In particular, the representative consumer in island j chooses the set of offer prices so as to maximize

$$\sum_i Q_{ij}(1 - p_{ij}),$$

[51]This model is based on the directed search model described by Peters (2000), which is itself based on the model from Moen (1997).

[52]It is common in the literature to also assume that π is convex. For example, Eeckhout and Kircher (2010) write the probability of a successful match as proportional to the inverse of the number of searchers, that is, $\tilde{\pi}(\frac{1}{n})$, and assume that $\tilde{\pi}$ is strictly increasing and strictly concave, which is implied by $\pi(n)$ to be strictly increasing and strictly convex.

where Q_{ij} is the quantity of produce from i that is successfully transacted in island j. Let $n_{ij}(\varphi)$ denote the density of producers from i searching island j of size φ. We then have

$$n_{ij} = \int n_{ij}(\varphi)\, d\varphi$$

and

$$Q_{ij} = \int \varphi \pi_{\varphi}(n_{ij}) n_{ij}(\varphi)\, d\varphi.$$

Equilibrium

The representative consumer in each island chooses the set of prices $\{p_{ij}\}_i$ so as to maximize her payoff, taking into account how the choice of prices will affect the number of producers searching her island and, therefore, the extent of matching frictions. Because search is directed, each representative consumer has to offer a price that is at least as great as the expected profits a producer will receive elsewhere or else the producer will not search there. Because there are a large number of islands, it is assumed that each representative consumer takes the expected profits of a producer as given. Because no consumer will choose prices to offer more than the expected profits of a producer, we have, for all $j \in S$,

$$p_{ij}\pi_{\varphi}(n_{ij}) = u_i(\varphi), \tag{24}$$

where u_i is the expected per-unit profits of a producer from island i of size φ.

Because π is monotonically decreasing, its inverse is well defined so that we can write, for all φ,

$$n_{ij} = \pi_{\varphi}^{-1}\left(\frac{u_i(\varphi)}{p_{ij}}\right),$$

so that $\frac{\partial n_{ij}}{\partial p_{ij}} > 0$, that is, offering higher prices attracts a greater density of producers. Intuitively, for a producer to be indifferent between searching two locations with different prices, it must be that the probability of being successfully matched in the location with the higher price is lower, which in turn requires that there are more producers searching that location.

Substituting equation (24) into the representative consumer's utility function yields

$$\sum_i Q_{ij}(1 - p_{ij}) \quad \iff \quad \sum_i \left(\frac{1}{p_{ij}} - 1\right) \int \varphi u_i(\varphi) n_{ij}(\varphi)\, d\varphi.$$

The representative consumer chooses the price to maximize his utility, taking into account how the choice of price will affect the number of producers searching his island. The first order conditions of the representative consumer with respect to the price p_{ij} yield

$$(25)\qquad \left(p_{ij} - p_{ij}^2\right) = \frac{\displaystyle\int \varphi u_i(\varphi) n_{ij}(\varphi)\, d\varphi}{\displaystyle\int \varphi u_i(\varphi) \frac{\partial}{\partial p_{ij}} n_{ij}(\varphi)\, d\varphi}.$$

Finally, to determine the expected per-unit profit function $u_i(\varphi)$, it must be the case that the total number of producers searching equals the density of producers who exist, that is, for all $i \in S$ and φ,

$$(26)\qquad \sum_j n_{ij}(\varphi) = A_i\, dF_i(\varphi),$$

where A_i is the total measure of producers in island i.

Comparison With the Model in the Paper

The model in the paper, which is based on an undirected search process, yields the prediction that larger producers will search more intensively and, as a result, on average sell to destinations with higher prices. These predictions are borne out empirically at both the macrolevel (see Pattern 4 in Section 2.3) and the microlevel (see Pattern 5 in Section 2.3).

To what extent does this model of directed search yield the same predictions about the relationship between producer size and the search process? First, note that larger producers achieve higher expected per-unit profits if and only if they are more likely to be successfully matched given a certain total number of producers searching. To see this, differentiate (24) with respect to φ,

$$\frac{\partial u_i(\varphi)}{\partial \varphi} = p_{ij} \frac{\partial \pi_\varphi(n_{ij})}{\partial \varphi},$$

so that $\frac{\partial u_i(\varphi)}{\partial \varphi} > 0$ if and only if $\frac{\partial \pi_\varphi(n_{ij})}{\partial \varphi} > 0$.

While the directed search model also can predict that larger producers have greater expected profit (given the assumption $\frac{\partial \pi_\varphi(n_{ij})}{\partial \varphi} > 0$), it does not necessarily follow that larger producers will on average sell to destinations with higher prices. This is because producers' expected profits in a directed search model depend both on the price she receives and the probability of a successful match. Indeed, it is relatively simple to construct an equilibrium where producers of all sizes from a given origin sell equal shares to all destinations. To do so, assume that $n_{ij}(\varphi) = n_{ij}\, dF_i(\varphi)$, so that $\frac{\partial n_{ij}}{\partial \varphi\, \partial p_{ij}} = 0$, that is, the responsiveness of producers to changes in the destination price is invariant to the producer size. The producer indifference condition then be written as

$$n_{ij}(\varphi) = \pi_{\varphi}^{-1}\left(\frac{u_i(\varphi)}{p_{ij}}\right) dF_i(\varphi),$$

which allows us to write the representative consumer utility function

$$\sum_i \left(\frac{1}{p_{ij}} - 1\right) \int \varphi u_i(\varphi) \pi_{\varphi}^{-1}\left(\frac{u_i(\varphi)}{p_{ij}}\right) dF_i(\varphi),$$

so that the first order conditions with respect to p_{ij} become

$$\left(\frac{\int \varphi u_i(\varphi) \pi_{\varphi}^{-1}\left(\frac{u_i(\varphi)}{p_{ij}}\right) dF_i(\varphi)}{\int \varphi u_i(\varphi)^2 (\pi_{\varphi}^{-1})'\left(\frac{u_i(\varphi)}{p_{ij}}\right) dF_i(\varphi)}\right) = -\left(\frac{1}{p_{ij}} - 1\right).$$

Finally, equation (26) can be written as

$$\sum_j \pi_{\varphi}^{-1}\left(\frac{u_i(\varphi)}{p_{ij}}\right) = A_i.$$

The last two equations can be jointly solved to determine the equilibrium prices and expected utility level $u_i(\varphi)$.

Hence, in a directed search model, there is no need for larger producers to be matched with destinations with higher prices than for smaller producers. This is because in equilibrium the wait time adjusts so that all producers are indifferent between selling across locations. Even if larger producers have better match technologies, this only increases their expected per-unit profits; because of the indifference condition, it need not effect where they search.

Extending the Framework to Incorporate Interactions Among Producers From Different Islands

Thus far, it has been assumed that there is no interaction between producers from different islands who search the same destination; in particular, there is no restriction on the relationship between the prices they receive; neither does the number of producers searching from elsewhere affect the probability of being matched. Since producers produce the same homogeneous good regardless of their location and sell to the same destinations, such an assumption seems unrealistic. What would happen if producers from different islands sold their produce in the same markets? If the representative consumer did not differentiate between producers from different markets, then the match probability of a producer of size φ searching island j would be $\pi_\varphi(n_j)$ regardless of the origin of the producer, where n_j is the total number of producers searching island j. Additionally, all producers would be offered the same price by the representative consumer. In what follows, I make the standard assumption in trade models that the price received by a producer from island i selling to island j is the market price in island j net of the transport costs of shipping to island j, that is, $p_{ij} = \frac{p_j}{\tau_{ij}}$, where $\tau_{ij} \geq 1$ is the "iceberg" transportation cost.

With these two assumptions, the producer indifference condition (equation (24)) now becomes

$$\frac{p_j}{\tau_{ij}} \pi_\varphi(n_j) = u_i(\varphi).$$

Consider two producers from islands $i \in S$ and $l \in S$, respectively, both of size φ. Since they are indifferent between selling to any two destinations $j, k \in S$, this implies

$$\frac{\frac{p_j}{\tau_{ij}} \pi_\varphi(n_j)}{\frac{p_j}{\tau_{lj}} \pi_\varphi(n_j)} = \frac{\frac{p_k}{\tau_{ik}} \pi_\varphi(n_k)}{\frac{p_k}{\tau_{lk}} \pi_\varphi(n_k)} \quad \Longleftrightarrow \quad \frac{\tau_{lj}}{\tau_{lk}} = \frac{\tau_{ij}}{\tau_{ik}}. \tag{27}$$

For producers from different islands to be indifferent between selling to all destinations, it must be the case that the relative trade costs to all destination pairs is the same for all locations. This requirement is immediately violated if we assume that for all $i, j \in S$, $\tau_{ij} = 1 \iff i = j$, that is, there exist positive trade costs to every destination except the local destination. Letting $i = j$ and $l = k$ in equation (27) immediately yields the contradiction $\tau_{il} = \frac{1}{\tau_{il}}$. Hence, when trade is costly and the representative consumer does not discriminate between producers arriving from different islands, it cannot be the case that producers are indifferent across the different destinations.

C.2. *Costly Directed Search*

Like the previous framework, producers engage in directed search by choosing the probability of transacting with each destination, knowing the prices everywhere. Unlike the previous framework, there is no risk of not meeting a buyer: instead, I assume that it becomes increasingly costly to target particular destinations. I show that this setup yields very similar qualitative predictions as the model in the paper, but is less tractable and dependent on the chosen functional form of the cost function.

Setup

There are a large number of islands $i \in S$, each inhabited by consumers and producers. Producers are price takers and all produce a homogeneous commodity (rice). Producers produce an amount equal to their productivity φ (landholdings), which is distributed according to the cumulative density function $F_i(\varphi)$. Each island has a perfectly competitive market for the homogeneous good: the price in the market is determined by the consumer's inverse demand function $p = D_i(q)$, where $\frac{\partial}{\partial q} D_i(q) < 0$. The mass of producers producing on an island is A_i, where A_i is a stochastic variable (weather shock).

A producer from i can either sell at home for the market price or pay a cost f_i to export. If the producer decides to export, she then engages in a costly search process to determine where she sells her produce. Let the probability that a producer from island i sells in destination j be π_{ij}. Let $c_i(\{\pi_{ij}\}_j)$ be the total costs incurred by a producer with a chosen set of match probabilities $\{\pi_{ij}\}_j$. I assume that c_i is increasing and convex in each of its arguments (i.e., it becomes increasingly costly to target particular destinations).

Equilibrium

A producer from island i of size φ chooses her match probabilities $\{\pi_{ij}\}_j$ so as to maximize her expected profits:

$$\varphi \sum_j p_{ij}\pi_{ij} - c_i\big(\{\pi_{ij}\}_j\big) \quad \text{s.t.} \quad \sum_j \pi_{ij} = 1.$$

Taking the first order conditions of the corresponding Lagrangian with respect to π_{ij} yields

$$\varphi p_{ij} - \frac{\partial}{\partial \pi_{ij}} c_i\big(\{\pi_{ij}\}_j\big) = \lambda. \tag{28}$$

The second order conditions with respect to π_{ij} require

$$\frac{\partial^2}{\partial \pi_{ij}^2} c_i\big(\{\pi_{ij}\}_j\big) > 0,$$

which is satisfied by assumption. Applying the inverse function theorem to equation (28) immediately yields that $\frac{\partial \pi_{ij}}{\partial p_{ij}} > 0$, that is, producers are more likely to sell to destinations with higher prices. Furthermore, it can be shown that

$$\frac{\partial^2 \pi_{ij}}{\partial p_{ij}\,\partial \varphi} > 0 \quad \iff \quad \frac{\varphi p_{ij}\left(\frac{\partial^3}{\partial \pi_{ij}^3} c_i(\{\pi_{ij}\}_j)\right)}{\left(\frac{\partial^2}{\partial \pi_{ij}^2} c_i(\{\pi_{ij}\}_j)\right)^2} < 1, \tag{29}$$

that is, larger producers respond more to an increase in price if the convexity of the cost function is not increasing too much with π_{ij}.

Let the equilibrium match probabilities satisfying equation (28) be denoted by $\pi_{ij}(\varphi)$. A producer will decide to search if the expected net benefit of doing so exceeds the fixed costs of searching, that is, if

$$\varphi\Bigl(\sum_j p_{ij}\pi_{ij} - p_{ii}\Bigr) - c_i\bigl(\{\pi_{ij}\}_j\bigr) - f \geq 0.$$

Differentiating the left hand size with respect to φ yields

$$\begin{aligned}
&\frac{\partial}{\partial\varphi}\Bigl(\varphi\Bigl(\sum_j p_{ij}\pi_{ij}(\varphi) - p_{ii}\Bigr) - c_i\bigl(\{\pi_{ij}\}_j\bigr) - f\Bigr) \\
&\quad = \Bigl(\sum_j p_{ij}\pi_{ij}(\varphi) - p_{ii}\Bigr) \\
&\qquad + \Bigl(\sum_j \pi'_{ij}(\varphi)\Bigl(\varphi p_{ij} - \frac{\partial}{\partial \pi_{ij}} c_i\bigl(\{\pi_{ij}\}_j\bigr)\Bigr)\Bigr) \quad \Longleftrightarrow \\
&\quad = \Bigl(\sum_j p_{ij}\pi_{ij}(\varphi) - p_{ii}\Bigr) + \Bigl(\lambda \sum_j \pi'_{ij}(\varphi)\Bigr) \quad \Longleftrightarrow \\
&\quad = \sum_j p_{ij}\pi_{ij}(\varphi) - p_{ii},
\end{aligned}$$

where the second line uses the first order conditions and the last line uses the constraint that $\sum_j \pi_{ij}(\varphi) = 1$ for all φ. Hence, the benefit of searching is increasing in producer size as long as the expected price from searching exceeds the home price, that is, $\sum_j p_{ij}\pi_{ij}(\varphi) > p_{ii}$. This implies that there will exist a threshold producer size φ_i^* such that all producers above the threshold search and all the producers below the threshold sell at home, where

$$\varphi_i^* = \frac{f + c_i(\{\pi_{ij}(\varphi_i^*)\}_j)}{\sum_j p_{ij}\pi_{ij}(\varphi_i^*) - p_{ii}}.$$

It is immediately evident that the greater is the fixed cost of search, the larger is the threshold where the producer is indifferent between searching and not searching, that is, $\frac{\partial \varphi_i^*}{\partial f} > 0$.

Finally, total trade from island i to island j can be calculated by integrating across all producers:

$$Q_{ij} = A_i \int_{\varphi_i^*}^{\infty} \varphi \pi_{ij}(\varphi)\, dF_i(\varphi). \tag{30}$$

A Simple Example

Suppose that the cost of search increases quadratically with the search intensity:

$$c_i(\{\pi_{ij}\}_j) = \frac{1}{2}\sum_j \pi_{ij}^2.$$

Then the producer chooses the set of π_{ij} so as to maximize

$$\sum_j \varphi p_{ij}\pi_{ij} - \frac{1}{2}\sum_j \pi_{ij}^2 \quad \text{s.t.} \quad \sum_j \pi_{ij} = 1.$$

For simplicity, I ignore the possibility of corner solutions and suppose that $\pi_{ij} \in (0, 1)$ for all i and $j \in S$. Taking first order conditions with respect to π_{ij} yields

$$\pi_{ij} = \varphi p_{ij} - \lambda.$$

Summing over all j and applying the constraint that the probabilities sum to 1 provides an analytic solution for λ,

$$\lambda = \varphi \bar{p}_{ij} - \frac{1}{S},$$

where $\bar{p}_{ij} \equiv \frac{1}{S}\sum_j p_{ij}$ is the average price producers from i would achieve selling elsewhere. This implies the simple function for the trade shares:

$$\pi_{ij} = \frac{1}{S} + \varphi(p_{ij} - \bar{p}_{ij}). \tag{31}$$

Equation (31) is intuitive: producers direct their search toward destinations with higher prices, with larger producers placing a greater weight on finding destinations with above average prices.

The size of the threshold producer is determined by the quadratic equation

$$\varphi_i^* = \frac{f + \frac{1}{2}\sum_j \left(\frac{1}{S} + \varphi_i^*(p_{ij} - \bar{p}_{ij})\right)^2}{\sum_j p_{ij}\left(\frac{1}{S} + \varphi_i^*(p_{ij} - \bar{p}_{ij})\right) - p_{ii}} \quad \Longleftrightarrow \tag{32}$$

$$0 = (\varphi_i^*)^2\left(\frac{1}{2}\sum_j (p_{ij}^2 - \bar{p}_{ij}^2)\right) + \varphi_i^*(\bar{p}_{ij} - p_{ii}) - \left(f + \frac{1}{2}\sum_j \frac{1}{S^2}\right).$$

Total trade flows are

$$Q_{ij} = A_i \int_{\varphi_i^*}^{\infty} \varphi\left(\frac{1}{S} + \varphi(p_{ij} - \bar{p}_{ij})\right) dF_i(\varphi) \quad \iff$$
$$= \frac{A_i}{S} \int_{\varphi_i^*}^{\infty} \varphi\, dF_i(\varphi) + A_i(p_{ij} - \bar{p}_{ij}) \int_{\varphi_i^*}^{\infty} \varphi^2\, dF_i(\varphi).$$

If producer size is Pareto distributed (i.e., $dF_i(\varphi) = \theta_i \varphi^{-\theta_i - 1}\, d\varphi$), then total trade flows are

$$Q_{ij} = A_i \theta_i \left(\frac{1}{S} \int_{\varphi_i^*}^{\infty} \varphi^{-\theta_i}\, d\varphi + (p_{ij} - \bar{p}_{ij}) \int_{\varphi_i^*}^{\infty} \varphi^{1-\theta_i}\, d\varphi\right) \quad \iff$$
$$= A_i \theta_i \left(\frac{1}{S}\left(\frac{1}{\theta_i - 1}\right)\left(\varphi_i^*\right)^{1-\theta_i} + (p_{ij} - \bar{p}_{ij})\left(\frac{1}{\theta_i - 2}\right)\left(\varphi_i^*\right)^{2-\theta_i}\right),$$

where φ_i^* is determined by equation (32) above.

Comparison With the Model in the Paper

This general search model presented above yields very similar predictions concerning the equilibrium search behavior as the model presented in the paper. As in the paper, the model above predicts that producers will sell more to destinations with higher prices, and if condition (29) is satisfied, larger producers will sell disproportionately more to destinations with higher prices. Furthermore, the presence of a fixed cost of search implies that all producers below a threshold size will sell at home rather than searching; as in the paper, this threshold is increasing in the size of the fixed cost and decreasing in the benefits of exporting. Given this threshold behavior, both the model presented here and the model in the paper allow for total bilateral trade flows to be determined by aggregating across the behavior of all producers above the threshold at which they export. The fact that the theoretical predictions are qualitatively the same between the model in the paper and the general model of search presented here demonstrates that the particular nature of the search process presented in the paper is not necessary to generate the predictions regarding the differences across producers in their search behavior.

While the general search model presented here yields predictions similar to the model in the paper, the model in the paper is more tractable. Even with a very simple cost function (such as the one presented in the example above), bilateral trade flows cannot be written as a closed form function of prices alone. More complicated cost functions (or allowing for the possibility of corner solutions) would not yield closed form solutions of the equilibrium match probabilities.

C.3. *Bilateral Search*

In the paper and in previous frameworks, I have assumed that producers have actively engaged in search while consumers have taken a passive role in transactions. In this framework, I assume that both producers and consumers engage in a search process so as to successfully complete a transaction. I show that when consumers search as well, there will be within-island price dispersion as well as across-island price dispersion, which will lead me to, if anything, underestimate the extent of information frictions.

Setup

There are a large number of islands $i \in S$, each inhabited by a unit mass of consumers and producers. Producers produce a quantity φ of homogeneous commodity (rice), where φ (landholdings) is assumed to vary across producers according to cumulative distribution function $F_i(\varphi)$.

Like the model in the paper, producers sequentially search markets to determine where to sell. Unlike the paper, instead of producers selling their produce in a centralized perfectly competitive market once they have chosen a market, I suppose here that there exist search frictions within the island. I model these within-island search frictions using a "directed" search framework. Suppose there exists a mass of potentially heterogeneous consumers living on the island, and for trade to occur, a producer and a consumer must be successfully matched. The matching process proceeds as follows. Each consumer announces a price at which she is willing to purchase produce from a producer and exerts some costly effort (interpreted as a number of search "attempts") to become matched with a producer. Producers, knowing the price each consumer is offering, choose a particular consumer to whom to attempt to sell their produce. Search frictions arise because there is a positive probability that the match is unsuccessful.

I assume that how successful the search process is depends on how many attempts the consumer makes, e, and the amount of produce targeting the consumer, q. For analytic tractability, I make an assumption common in the literature that the match function is Cobb–Douglas, so that the total quantity s a consumer is able to successfully find is

$$s(n, e) = B_i q^{\alpha} e^{1-\alpha},$$

where $\alpha \in (0, 1)$ and B_i is the island-specific match technology. Given the match function, the probability that a particular unit of the homogeneous good the producer targets to a particular consumer is successfully transacted is $\frac{M(q,e)}{q} = B_i(\frac{e}{q})^{1-\alpha}$.[53]

[53]The assumption that the match function depends on the total quantity of produce being targeted at a particular consumer rather than the total number of producers targeting a particular

I assume that the consumer is choosing between the homogeneous good sold by the producers and an outside numeraire good that provides her with constant marginal utility, and faces linearly increasing marginal costs to searching with slope c so that the utility function of the consumer can be written as

$$U(s, p, e) = u(s) - ps - c\frac{e^2}{2}.$$

In what follows below, I focus on the particularly tractable case where $u(s) = \log s$.

Equilibrium

Let $\bar{p}_i$ be the expected per-unit profit of a searching producer, which each consumer takes as given. A consumer has to make a price offer so that the searching producer has profits of at least $\bar{p}_i$ and has no reason to offer any more. Given the probability of a successful transaction, a consumer will offer a price p such that

$$B_i\left(\frac{e}{q}\right)^{1-\alpha} p = \bar{p}_i.$$

Combining this indifference condition with the matching function yields the total quantity a consumer will successfully purchase as a function of the number of searches she makes, the price she offers, and the expected per-unit profit of a searching producer:

$$s = B_i^{1/(1-\alpha)}\left(\frac{p}{\bar{p}_i}\right)^{\alpha/(1-\alpha)} e. \tag{33}$$

Combining equation (33), which relates the quantity a consumer purchases, with the consumer's first order conditions with respect to the number of searches e and the price she offers p yields the following equation, which determines the optimal offered price as a function of the consumer's search cost c and the expected per-unit profit of a searching producer:

$$p(c) = \alpha u'\left(\frac{1}{c}\left(\frac{1-\alpha}{\alpha}\right)B_i^{2/(1-\alpha)}\left(\frac{p}{\bar{p}_i}\right)^{2\alpha/(1-\alpha)} p\right).$$

consumer is made for analytical simplicity. Note that because producers are profit maximizing, the expected payoff from having each unit being successfully sold with probability p is equal to the expected payoff of the entire transaction being successful with probability p. However, this assumption has the (admittedly unattractive) implication that a particular producer targeting a particular consumer may be successful in transacting some units with that consumer while unsuccessful with transacting other units.

When $u(q) = \log q$, the expression becomes

$$p(c) = \frac{\alpha^{(1-\alpha)/2}}{B_i}\left(\frac{\alpha}{1-\alpha}\right)\bar{p}_i^{\alpha} c^{(1-\alpha)/2}. \tag{34}$$

Equation (34) is intuitive: the greater is the expected per-unit profit of producers, the higher is the price the consumer must offer so as to make producers willing to target the consumer. Consumers with higher costs of searching search less intensively and instead offer higher prices to producers to compensate them for the lower probability of a successful match.

To determine $\bar{p}_i$, we can combine the match function with equation (34) to yield the total quantity producers send to each consumer. It turns out that with the above functional form assumptions, the total amount sent to each consumer is invariant to the search cost c, so that the total amount sent to each consumer is

$$q_i(c) = \frac{\alpha\left(\frac{\alpha}{1-\alpha}\right)^{(1+\alpha)/(1-\alpha)}}{\bar{p}_i}.$$

Let G_i be the cumulative distribution function of the search costs across consumers. By integrating over all consumers, we can determine the total quantity of produce that is sent to island i,

$$Q_i = \int q_i(c)\,dG_i(c) = \frac{\alpha\left(\frac{\alpha}{1-\alpha}\right)^{(1+\alpha)/(1-\alpha)}}{\bar{p}_i},$$

which we can invert to derive the expected per-unit profit of producers as a function of the quantity sent to a particular destination:

$$\bar{p}_i = \frac{Q_i}{\alpha\left(\frac{\alpha}{1-\alpha}\right)^{(1+\alpha)/(1-\alpha)}}. \tag{35}$$

Comparison With the Model in the Paper

As in the model in the paper, producers choose which island to sell to based on their expected per-unit profits in that island. In the model in the paper, this is the competitive market price in that island; in the framework presented here, the expected per-unit profits $\bar{p}_i$ depend also on the probability of successful transactions. Given the same distribution of per-unit profits in both frameworks, the producer's search process across islands remains unchanged.

Furthermore, equation (35) provides a particular functional form for the inverse demand function assumed in the paper. As a result, the equations governing trade flows (as well as the existence and uniqueness proofs) from the paper immediately apply here. In this sense, introducing bilateral search does not affect the theoretical results of the paper.

Introducing bilateral search, however, does impact the mapping between the model and the data. This is because in the presence of bilateral search frictions (that is, an absence of a centralized market in each island), there will be a distribution of transaction prices within each island. In constructing a price for a particular commodity in a particular province at a particular time, the Philippines Bureau of Agricultural Statistics (BAS) calculates (roughly[54]) a weighted average of a number of transaction prices. Let $w(c)$ be the weight the BAS places on a transaction with a consumer of type c. Then the price I observe in the data, p_i, is

$$p_i = \int p(c)w(c)\,dF_i(c) \quad \iff$$
$$p_i = \frac{\alpha^{(1-\alpha)/2}}{B_i}\left(\frac{\alpha}{1-\alpha}\right)\bar{p}_i^{\alpha}\int c^{(1-\alpha)/2}w(c)\,dF_i(c).$$

If all islands have the same search technology and the same cost distribution across consumers (i.e., $B_i = B$ and $dF_i = dF$ for all $i \in S$), then for any $i, j \in S$, we have

$$\frac{\bar{p}_i}{\bar{p}_j} = \left(\frac{p_i}{p_j}\right)^{1/\alpha},$$

so that the difference across islands in relative prices that I observe in the data is actually smaller (since $\alpha \in (0, 1)$) than the difference across islands in the relative expected per-unit profits. That is, the observed price dispersion in the data understates the dispersion in the expected per-unit profits that drive the producer search process. Recall from equation (7) that the fixed cost of search is proportional to the value of search (i.e., $f_i = \frac{K_i(p_i)}{\Lambda_i^{1/(\theta_i-1)}}$). Since the value of search increases with the amount of price dispersion, this implies that the information frictions estimated in the paper assuming competitive markets in each island *underestimate* the extent of information frictions if trade within islands is subject to bilateral search frictions.

While the exact expression of the relationship between p_i and $\bar{p}_i$ depends on the functional form assumptions for the consumer utility function and the

[54]When the BAS gave me the survey data that underlie their calculated market prices, I was cautioned not to attempt to construct market prices myself from the data, as the raw data undergo a substantial amount of "cleaning" during the averaging process.

matching technology, the intuition suggests that this result is more general. The elasticity of the price offered by consumers (p_i) to the expected per-unit profits ($\bar{p}_i$) is less than 1 because in the presence of search frictions, consumers have two margins with which to counter an increase in $\bar{p}_i$: they can raise their price or increase their search intensity. To the extent that consumers do the latter, observed differences in offered prices across islands fail to capture differences in search intensities, thereby understating the differences in expected per-unit profits.

C.4. *Ex ante Search*

In this framework, I assume that producers decide which destinations to search prior to the realization of productivity shocks. Then, after productivity shocks have been realized, producers simply sell to the destination with the highest price that they have searched rather than engaging in a sequential search process. I show that this framework yields similar qualitative predictions as the model in the paper, but is only tractable under extreme symmetry assumptions.

Setup

There are a large number of islands $i \in S$, each inhabited by consumers and producers. Producers are price takers and all produce a homogenous commodity (rice). Producers produce an amount equal to their productivity φ (landholdings), which is distributed according to the cumulative density function $F_i(\varphi)$. Each island has a perfectly competitive market for the homogeneous good: the price in the market is determined by the consumer's inverse demand function $p = D_i(q)$, where $\frac{\partial}{\partial q} D_i(q) < 0$. The mass of producers producing on an island is A_i, where A_i is a stochastic variable (weather shock).

The timing of the model is as follows. First, producers decide which markets to search. A producer from island i searching island j incurs a fixed cost f_{ij}, where it is assumed $f_{ii} = 0$ for all $i \in \{1, \dots, N\}$. After producers have searched the destinations, the weather shocks $\{A_i\}$ are realized. In equilibrium, the realization of the weather shocks will result in variation across islands in prices. Producers then choose to which destination to sell. Producers can only sell to destinations that they have searched.[55] A producer from i selling to j incurs an iceberg trade cost τ_{ij}, where it is assumed that $\tau_{ii} = 1$ for all i.

[55]There may also be noninformation frictions preventing producers from selling to destinations they have not searched; for example, it could be that producers have to establish relationships with traders in a particular destination so as to sell there. However, information frictions must also be present, as the critical assumption of the framework is that producers do not respond to prices in unsearched markets, regardless of their level. Without information frictions, this assumption is untenable: if a producer was aware of an arbitrarily high price in an unsearched region, she would be willing to incur arbitrarily large costs to sell there.

For simplicity, in what follows I consider that all islands are ex ante identical and trade costs to each island are symmetric, that is, $\tau_{ij} = \tau$ and $f_{ij} = f$ for all i and j.

Optimal Search

Once prices have been realized, each producer will simply sell her produce to the destination with the highest price net of trade costs. As a result, prior to the prices being realized, a producer with productivity φ will choose where to search so as to maximize her expected profits. Because searching the home location is costless, all producers will search there. Let $F_i(p)$ be the cumulative distribution function of the prices in island j. Because all destinations are ex ante identical, $F_i(p) = F(p)$ for all i. As a result, instead of choosing which destinations to search, each producer simply chooses the number of destinations (other than their home island) to search,

$$\max_n \frac{\varphi}{\tau} \int_0^\infty p \, dF_{(n)}(p) - nf,$$

where $F_{(n)}(p)$ is the cumulative distribution function of the maximum price received when searching n destinations. It is straightforward to show that $F_{(n)}(p) = F(p)^n$, so that the producer's search decision becomes

$$\max_n \frac{\varphi}{\tau} \int_0^\infty pnF(p)^{n-1} \, dF(p) - nf.$$

For simplicity, ignore integer constraints so that we can take first order conditions. Then the optimal number of destinations to search, $n(\varphi)$, is implicitly defined by the equation

$$\frac{\varphi}{\tau}\left(\int_0^\infty pF(p)^{n(\varphi)-1} \, dF(p) + n(\varphi) \int_0^\infty pF(p)^{n(\varphi)-1} \log F(p) \, dF(p)\right) = f. \tag{36}$$

It can be shown that $\frac{\partial}{\partial\varphi} n(\varphi) > 0$, that is, more productive producers choose to search more locations, as their revenue increases more from being connected to a location with a higher price.

Aggregate Bilateral Trade Flows

Because producers are indifferent about which destination they search, I assume that search is random. As a result, the density of producers from island i searching any island j of productivity φ is $n(\varphi) \, di$. The probability that a producer from i selling to destination $j \neq i$ searches is equal to the probability that the realized price in j, p_j, is greater than every other destination she searches as well as her home price. As a result, the total trade from i to j (where $i \neq j$) is equal to

$$Q_{ij}(p_j) = A_i\, di \int_0^\infty n(\varphi) F(p_j)^{n(\varphi)-1} F\left(\frac{p_j}{\tau}\right) dG(\varphi).$$

Conditioning on the home price p_i as well yields

$$(37) \qquad Q_{ij}(p_i, p_j) = \begin{cases} 0, & \text{if } \dfrac{p_j}{\tau_{ij}} < p_i, \\ A_i\, di \displaystyle\int_0^\infty n(\varphi) F(p_j)^{n(\varphi)-1}\, dG(\varphi)\, di, & \text{if } \dfrac{p_j}{\tau_{ij}} \geq p_i. \end{cases}$$

The probability that a producer of type φ sells at home is the probability that none of the $n(\varphi)$ places she searches has a price (net of transportation costs) greater than p_j, so that the total quantity of producers selling locally is

$$Q_{jj}(p_j) = A_j \int_0^\infty F(p_j \tau)^{n(\varphi)}\, dG(\varphi).$$

Equilibrium Prices and Price Distribution

The equilibrium price in island j is determined by the fixed point expression

$$\begin{aligned} p_j &= D_j\left(\int_{S\setminus j} Q_{ij}(p_j) + Q_{jj}(p_j)\right) \\ &= D_j\left(\int_0^\infty n(\varphi) F(p_j)^{n(\varphi)-1} F\left(\frac{p_j}{\tau}\right) dG(\varphi) \int_{S\setminus j} A_i\, di \right. \\ &\qquad \left. + A_j \int_0^\infty F(p_j \tau)^{n(\varphi)}\, dG(\varphi)\right), \end{aligned}$$

where $D_j(\cdot)$ is the inverse demand function in island j, and $Q_{ij}(p_j)$ and $Q_{jj}(p_j)$ are as defined above. It is straightforward to show that $\frac{\partial}{\partial p} Q_{ij}(p) > 0$ for all $i \neq j$ and $\frac{\partial}{\partial p} Q_{jj}(p) > 0$, that is, as the price increases, the number of producers selling to destination j increases. Given any inverse demand function D_j such that $\frac{\partial}{\partial q} D_j(q) < 0$, the proof of the existence of a set of equilibrium prices follows the one presented in the paper. Hence, we can write the equilibrium price in destination j solely as a function of its own weather shock: $p_j = f_j(A_j)$. Let $H(A_j)$ be the cumulative distribution function of the home weather shock. Then the cumulative distribution function of prices in island j can be written only as a function of its own weather shock,

$$F_j(p) = \int \mathbf{1}\{f_j(A_j) \leq p\}\, dH(A_j),$$

where $\mathbf{1}\{\cdot\}$ is an indicator function.

Comparison With the Model in the Paper

The departure of this framework from the model in the paper is that here search occurs prior to the realization of prices. Despite this difference, many of the model predictions are qualitatively the same. At the microlevel, large producers search more intensively and are more likely to receive higher prices. At the macrolevel, trade occurs from i to j if and only if $\frac{p_j}{\tau_{ij}} \geq p_i$, and when $\frac{p_j}{\tau_{ij}} \geq p_i$, trade flows increase with the destination price. Indeed, the bilateral trade flow equation (37) bears a strong resemblance to the bilateral trade flow equation in the paper.

The major difference in empirical predictions between this model and the one in the paper is that in this framework, conditional on trade flows occurring, the origin price has no effect on bilateral trade flows. Intuitively, if the destination price (net of transportation costs) exceeds the origin price, then anyone searching the destination will necessarily sell there, regardless of the origin price. In contrast, the model in the paper implies that the bilateral trade flow is greater the more different is the origin and the destination price, as the lower is the origin price (conditional on the destination price), the greater is the incentive producers have to search.[56]

The other difference between this framework and the framework in the paper is that this framework is not nearly as analytically tractable. Even with the symmetry assumptions made above, the optimal number of destinations searched is an implicit function of the equilibrium price distribution (see equation (36)). Relaxing the symmetry assumptions substantially complicates the problem, as different locations will have different price distributions, leading producers to choose which destinations to search rather than how many destinations to search. This optimization problem is especially intractable, as the distribution of the maximum prices across a given set of destinations will depend in part on the correlation in the prices across those destinations, which are themselves equilibrium objects. In contrast, aggregate trade flows in the framework presented in the paper are only a function of the origin price, destination price, and a single sufficient statistic (the "value of search") capturing the effect of prices in other destinations.

APPENDIX D: ALTERNATIVE NONSEARCH FRAMEWORKS

In this appendix, I consider alternative stories other than information frictions that may generate the observed patterns in the data. In particular, I first consider a complete information trade model where producers have to incur

[56]One could adjust the timing of this framework so that the search occurs after the realization of the weather shocks, so that producers are aware of the origin price prior to choosing which destinations to search. If this were the case, however, producers would prefer to search destinations sequentially so as to minimize the fixed costs incurred, as in the model in the paper.

a fixed cost of export. I then consider a similar framework but suppose that producers face idiosyncratic transportation costs. I consider both idiosyncratic bilateral iceberg transportation costs and idiosyncratic fixed export costs.

D.1. *Fixed Costs of Export*

In this subsection, I consider a complete information model where farmers incur fixed costs of export. I show that the elasticity of trade flows to changes in the destination price is decreasing in producer heterogeneity in a framework nearly identical to the information frictions model when there is complete information.

Setup

Suppose there is complete information, and trade from i to j is subject to a variable iceberg transportation cost τ_{ij} and a fixed cost of export f_{ij}. Goods are homogeneous, but sold by heterogeneous price-taking producers with quantity $\phi \in [1, \infty)$, where $F_i(\phi) = 1 - \phi^{-\theta_i}$, that is, the distribution of quantities produced is Pareto with a shape parameter θ_i. Larger values of θ_i reflect greater homogeneity in the distribution of quantities produced.

Trade

Producers maximize profits, which since the economy in an endowment economy, is equivalent to maximizing revenue. Since each producer has a fixed quantity to sell and takes prices throughout the world as given, each solves

$$\max_j \frac{p_j}{\tau_{ij}} \phi - f_{ij}.$$

To avoid having to pay two fixed costs, no firm will ever sell to more than one destination. Furthermore, larger firms will be willing to incur a higher fixed cost so as to reach destinations with high prices net of transportation costs. As a result, we can order the destinations by the size of firms from i that choose to sell to the destination. Without loss of generality, label destination 1 as the destination to which the smallest firms from i sell, label destination 2 as the next destination to which slightly larger firms sell, and so forth. Note that if a marginally larger firm sells to a new destination, it must be that the destination has both a higher price and a higher fixed cost, that is, $\frac{p_k}{\tau_{ik}} > \frac{p_{k-1}}{\tau_{ik-1}}$ and $f_{ik} > f_{ik-1}$. We can then define the threshold size ϕ^*_{ik} at which firms begin to export to the kth market:

$$\frac{p_k}{\tau_{ik}} \phi^*_{ik} - f_{ik} = \frac{p_{k-1}}{\tau_{ik-1}} \phi^*_{ik} - f_{ik-1} \quad \iff \quad \phi^*_{ik} = \frac{f_{ik} - f_{ik-1}}{\frac{p_k}{\tau_{ik}} - \frac{p_{k-1}}{\tau_{ik-1}}}. \tag{38}$$

The share of quantity produced in i that is exported to destination k, λ_{ik}, will then simply be the quantity produced by the interval of producers willing to sell to destination k:

$$\begin{aligned}
\lambda_{ik} &= \frac{\displaystyle\int_{\phi^*_{ik}}^{\phi^*_{ik+1}} \phi\, dF(\phi)}{\displaystyle\int_{1}^{\infty} \phi\, dF(\phi)} \\
&= \left(\phi^*_{ik}\right)^{1-\theta_i} - \left(\phi^*_{ik+1}\right)^{1-\theta_i} \\
&= \left(\frac{\dfrac{p_k}{\tau_{ik}} - \dfrac{p_{k-1}}{\tau_{ik-1}}}{f_{ik} - f_{ik-1}}\right)^{\theta_i - 1} - \left(\frac{\dfrac{p_{k+1}}{\tau_{ik+1}} - \dfrac{p_k}{\tau_{ik}}}{f_{ik+1} - f_{ik}}\right)^{\theta_i - 1}.
\end{aligned}$$

As a result, the elasticity of trade flows to changes in the destination price is

$$\begin{aligned}
&\frac{\partial \ln \lambda_{ik}}{\partial \ln \dfrac{p_k}{\tau_{ik}}} \\
&\quad = \frac{\partial \lambda_{ik}}{\partial \dfrac{p_k}{\tau_{ik}}} \frac{\dfrac{p_k}{\tau_{ik}}}{\lambda_{ik}} \\
&\quad = \frac{\dfrac{p_k}{\tau_{ik}}}{\lambda_{ik}} \frac{\partial}{\partial \dfrac{p_k}{\tau_{ik}}} \left(\left(\frac{\dfrac{p_k}{\tau_{ik}} - \dfrac{p_{k-1}}{\tau_{ik-1}}}{f_{ik} - f_{ik-1}}\right)^{\theta_i - 1} - \left(\frac{\dfrac{p_{k+1}}{\tau_{ik+1}} - \dfrac{p_k}{\tau_{ik}}}{f_{ik+1} - f_{ik}}\right)^{\theta_i - 1}\right) \\
&\quad = \frac{\dfrac{p_k}{\tau_{ik}}}{\lambda_{ik}} (\theta_i - 1) \left((f_{ik} - f_{ik-1})^{1-\theta_i} \left(\frac{p_k}{\tau_{ik}} - \frac{p_{k-1}}{\tau_{ik-1}}\right)^{\theta_i - 2} \right. \\
&\qquad \left. + (f_{ik+1} - f_{ik})^{1-\theta_i} \left(\frac{p_{k+1}}{\tau_{ik+1}} - \frac{p_k}{\tau_{ik}}\right)^{\theta_i - 2} \right) \\
&\quad = (\theta_i - 1) \frac{p_k}{\tau_{ik}} \left((f_{ik} - f_{ik-1})^{1-\theta_i} \left(\frac{p_k}{\tau_{ik}} - \frac{p_{k-1}}{\tau_{ik-1}}\right)^{\theta_i - 2} \right. \\
&\qquad \left. + (f_{ik+1} - f_{ik})^{1-\theta_i} \left(\frac{p_{k+1}}{\tau_{ik+1}} - \frac{p_k}{\tau_{ik}}\right)^{\theta_i - 2} \right) \\
&\qquad \Bigg/ \left(\left(\frac{\dfrac{p_k}{\tau_{ik}} - \dfrac{p_{k-1}}{\tau_{ik-1}}}{f_{ik} - f_{ik-1}}\right)^{\theta_i - 1} - \left(\frac{\dfrac{p_{k+1}}{\tau_{ik+1}} - \dfrac{p_k}{\tau_{ik}}}{f_{ik+1} - f_{ik}}\right)^{\theta_i - 1}\right).
\end{aligned}$$

Note that a first order Taylor approximation yields

$$\begin{aligned}
&\left(\frac{\frac{p_k}{\tau_{ik}}-\frac{p_{k-1}}{\tau_{ik-1}}}{f_{ik}-f_{ik-1}}\right)^{\theta_i-1} \\
&\approx\left(\frac{\frac{p_{k+1}}{\tau_{ik+1}}-\frac{p_k}{\tau_{ik}}}{f_{ik+1}-f_{ik}}\right)^{\theta_i-1} \\
&\quad+(\theta_i-1)\left(\frac{\frac{p_{k+1}}{\tau_{ik+1}}-\frac{p_k}{\tau_{ik}}}{f_{ik+1}-f_{ik}}\right)^{\theta_i-2} \\
&\quad\times\left(\left(\frac{\frac{p_k}{\tau_{ik}}-\frac{p_{k-1}}{\tau_{ik-1}}}{f_{ik}-f_{ik-1}}\right)-\left(\frac{\frac{p_{k+1}}{\tau_{ik+1}}-\frac{p_k}{\tau_{ik}}}{f_{ik+1}-f_{ik}}\right)\right),
\end{aligned}$$

so that

$$\begin{aligned}
&\frac{\partial\ln\lambda_{ik}}{\partial\ln\frac{p_k}{\tau_{ik}}} \\
&\approx\frac{p_k}{\tau_{ik}}\left((f_{ik}-f_{ik-1})^{1-\theta_i}\left(\frac{p_k}{\tau_{ik}}-\frac{p_{k-1}}{\tau_{ik-1}}\right)^{\theta_i-2}\right. \\
&\quad\left.+(f_{ik+1}-f_{ik})^{1-\theta_i}\left(\frac{p_{k+1}}{\tau_{ik+1}}-\frac{p_k}{\tau_{ik}}\right)^{\theta_i-2}\right) \\
&\quad\Bigg/\left(\left(\frac{\frac{p_{k+1}}{\tau_{ik+1}}-\frac{p_k}{\tau_{ik}}}{f_{ik+1}-f_{ik}}\right)^{\theta_i-2}\left(\left(\frac{\frac{p_k}{\tau_{ik}}-\frac{p_{k-1}}{\tau_{ik-1}}}{f_{ik}-f_{ik-1}}\right)-\left(\frac{\frac{p_{k+1}}{\tau_{ik+1}}-\frac{p_k}{\tau_{ik}}}{f_{ik+1}-f_{ik}}\right)\right)\right) \\
&\approx\frac{p_k}{\tau_{ik}}\left((f_{ik}-f_{ik-1})^{-1}\left(\frac{\left(\frac{p_k}{\tau_{ik}}-\frac{p_{k-1}}{\tau_{ik-1}}\right)/(f_{ik}-f_{ik-1})}{\left(\frac{p_{k+1}}{\tau_{ik+1}}-\frac{p_k}{\tau_{ik}}\right)/(f_{ik+1}-f_{ik})}\right)^{\theta_i-2}\right. \\
&\quad\left.+(f_{ik+1}-f_{ik})^{-1}\right) \\
&\quad\Bigg/\left(\left(\frac{\frac{p_k}{\tau_{ik}}-\frac{p_{k-1}}{\tau_{ik-1}}}{f_{ik}-f_{ik-1}}\right)-\left(\frac{\frac{p_{k+1}}{\tau_{ik+1}}-\frac{p_k}{\tau_{ik}}}{f_{ik+1}-f_{ik}}\right)\right).
\end{aligned}$$

Since $\phi^*_{ik+1} > \phi^*_{ik}$, we know that $\frac{1}{\phi^*_{ik}} > \frac{1}{\phi^*_{ik+1}}$, which, from equation (38), is equivalent to

$$\frac{\frac{p_k}{\tau_{ik}} - \frac{p_{k-1}}{\tau_{ik-1}}}{f_{ik} - f_{ik-1}} > \frac{\frac{p_{k+1}}{\tau_{ik+1}} - \frac{p_k}{\tau_{ik}}}{f_{ik+1} - f_{ik}},$$

so that

$$\left(\frac{\left(\frac{p_k}{\tau_{ik}} - \frac{p_{k-1}}{\tau_{ik-1}}\right) / (f_{ik} - f_{ik-1})}{\left(\frac{p_{k+1}}{\tau_{ik+1}} - \frac{p_k}{\tau_{ik}}\right) / (f_{ik+1} - f_{ik})} \right) > 1.$$

Comparison With the Model in the Paper

In the paper, I find that both theoretically and empirically the origins with more homogeneous producers are less responsive to changes in destination prices. In this model, however, to a first order, the elasticity of trade shares to changes in destination prices is increasing in θ_i, that is, origins with more homogeneous producers respond more to changes in destination prices.

D.2. *Idiosyncratic Bilateral Transportation Costs*

I now consider what would happen if producers faced idiosyncratic bilateral iceberg transportation costs.

Setup

There are a large number of islands $i \in S$, each inhabited by a unit mass of consumers and producers. Producers produce a quantity φ of a homogeneous commodity (rice), where φ (landholdings) is assumed to vary across producers according to cumulative distribution function $F_i(\varphi)$.

Producers are aware of prices everywhere and choose the destination to which to sell so as to maximize their revenue. Consider producer ν of size φ from island i. This producer can either sell at home and receive a price p_i or pay a fixed cost f_i and sell to destination $j \in S \setminus i$. If the producer sells to island j, she receives a price (net of transportation costs) of $\frac{p_j}{\tau_{ij}} \varepsilon_{ij\varphi}(\nu)$, where $\varepsilon_{ij\varphi}(\nu)$ is a producer–destination-specific idiosyncratic transportation cost shock. In what follows, I assume that $\varepsilon_{ij\varphi}(\nu)$ is i.i.d. across destinations and producers; additionally, following Eaton and Kortum (2002), I assume that $\varepsilon_{ij\varphi}(\nu)$ is distributed according to a Frechet distribution so that $\Pr\{\varepsilon_{ij\varphi}(\nu) < \varepsilon\} = \exp(\varepsilon^{-\theta})$.

Trade

Because producers are aware of prices everywhere, they choose the destination to which to sell so as to maximize their profits,

$$\max\left\{\max_{j\in S\setminus i}\frac{p_j}{\tau_{ij}}\varepsilon_{ij\varphi}(\nu)\varphi - f, p_i\varphi\right\},$$

where the outer maximization determines whether or not the producer exports and the inner maximization determines where the producer exports conditional on exporting. It can be shown that the probability a producer exports is

$$\Pr\left\{\max_{j\in S\setminus i}\frac{p_j}{\tau_{ij}}\varepsilon_{ij}(k)\varphi - p_i\varphi > f\right\} = 1 - e^{-(f/\varphi+p_i)^{-\theta}\sum_{j\in S\setminus i}(p_j/\tau_{ij})^{\theta}},$$

so that, as in the model in the paper, larger producers are more likely to export. Intuitively, because larger producers export a greater amount, it is more likely that they will receive idiosyncratic transportation cost draws that make it profitable to incur the fixed cost of exporting. Conditional on exporting, the probability that a producer exports to a particular destination is

$$\pi_{ij}(\varphi) = \frac{\left(\dfrac{p_j}{\tau_{ij}}\right)^{\theta_i}}{\displaystyle\sum_k\left(\dfrac{p_k}{\tau_{ik}}\right)^{\theta_i}}.$$

Since these probabilities do not depend on the productivity of a producer, the elasticity of trade flows to destination prices is invariant to the origin distribution of landholdings.

Comparison With the Model in the Paper

Like the model in the paper, larger producers are more likely to incur the fixed costs of export. Unlike the model in the paper, however, conditional on exporting, larger producers are no more likely to sell to destinations with higher prices than smaller producers. As a result, the elasticity of trade flows to destination prices will not depend on the distribution of landholdings in the origin island (i.e., $F_i(\varphi)$), which is inconsistent with the empirical evidence presented in the paper.

D.3. *Idiosyncratic Fixed Costs*

I now consider what would happen if producers faced idiosyncratic destination-specific fixed costs of export, in the spirit of Eaton, Kortum, and Kramarz (2011).

Setup

As in the previous section, there are a large number of islands $i \in S$, each inhabited by a unit mass of consumers and producers. Producers produce a quantity φ of homogeneous commodity (rice), where φ (landholdings) is assumed to vary across producers according to the cumulative distribution function $F_\varphi(\varphi)$.

Producers are aware of prices everywhere and choose the destination to which to sell so as to maximize their revenue. Consider producer ν of size φ from island i. This producer can either sell at home and receive a price p_i or pay a fixed cost $f_i \varepsilon_{ij\varphi}(\nu)$ and sell to destination $j \in S \setminus i$, where $\varepsilon_{ij\varphi}(\nu)$ is a producer–destination-specific fixed cost shock. For tractability, in what follows, I assume that $-\varepsilon_{ij\varphi}(\nu)$ is i.i.d. and Gumbel distributed, and the fixed costs of search are equal to 1, that is, $f_i = 1$.

Trade

Because producers are aware of prices everywhere, they choose the destination to which to sell so as to maximize their profits,

$$\max\left\{\max_{j \in S \setminus i} \frac{p_j}{\tau_{ij}} \varphi - f_i \varepsilon_{ij\varphi}(\nu),\, p_i \varphi\right\},$$

where the outer maximization determines whether or not the producer exports and the inner maximization determines where the producer exports conditional on exporting. The probability that a producer exports is equal to the probability that the fixed cost in at least one destination is low enough to make export profitable. Equivalently, this can be written as 1 minus the probability that the export in all destinations is sufficiently higher than the price gap, that is,

$$\begin{aligned}
&\Pr\left\{\max_{j \in S \setminus i} \frac{p_j}{\tau_{ij}} \varphi - f_i \varepsilon_{ij\varphi}(\nu) > p_i \varphi\right\} \\
&\quad = 1 - \exp\left\{-\sum_{j \in S \setminus i} \exp\left\{-\varphi\left(\frac{p_j}{\tau_{ij}} - p_i\right)\right\}\right\},
\end{aligned}$$

so that, as in the model in the paper, larger producers are more likely to export (since F_ε is monotonically increasing). Intuitively, because larger producers export a greater amount, it is more likely that they will receive idiosyncratic fixed cost draws that make it profitable to incur the fixed cost of exporting.

Following McFadden (1973), conditional on exporting, the probability that a producer exports to a particular destination is

$$\pi_{ij}(\varphi) = \frac{\exp\left\{\frac{p_j}{\tau_{ij}}\varphi\right\}}{\sum_{k\in S\setminus i}\exp\left\{\frac{p_k}{\tau_{ik}}\varphi\right\}}.$$

Unlike the previous subsection, the probability that a producer exports to a particular destination does depend on the size of the producer. In particular, the larger is the producer, the greater is the probability that the producer exports to a destination with a higher price net of transportation cost.

Similarly, let $\Pi_{ij}(\varphi)$ denote the probability that a producer of size φ in location i exports to j *unconditional* on exporting. It can be shown that the unconditional probability is similar to the conditional probability p_i with a correction term that compares how attractive destination prices are to the origin price:

$$\Pi_{ij}(\varphi) = \frac{\exp\left\{\frac{p_j}{\tau_{ij}}\varphi\right\}}{\sum_{k\in S\setminus i}\exp\left\{\frac{p_k}{\tau_{ik}}\varphi\right\}}\left(1-\exp\left\{-\exp\left\{\varphi\left(-p_i+\sum_{k\in S\setminus i}\frac{p_k}{\tau_{ik}}\right)\right\}\right\}\right). \tag{39}$$

Equation (39) implies that the probability a producer of a given size exports to a particular destination increases in the destination price and decreases in the origin price; intuitively, increasing the destination price makes a particular destination relatively more attractive than other destinations, while decreasing the origin price increases the probability a particular producers exports.

Finally, total trade flows can be determined by aggregating across the entire distribution of landholdings. As in the model in the text, suppose that there is a mass M_i of producers in location i, a fraction A_i of which produce. Then total trade flows can be written as

$$\begin{aligned} (40)\qquad & Q_{ij} = \int A_i M_i \varphi \Pi_{ij}(\varphi)\, dF_\varphi(\varphi) \quad \iff \\ & Q_{ij} = A_i M_i \int \frac{\exp\left\{\frac{p_j}{\tau_{ij}}\varphi\right\}}{\sum_{k\in S\setminus i}\exp\left\{\frac{p_k}{\tau_{ik}}\varphi\right\}} \\ & \qquad \times \left(1-\exp\left\{-\exp\left\{\varphi\left(-p_i+\sum_{k\in S\setminus i}\frac{p_k}{\tau_{ik}}\right)\right\}\right\}\right)\varphi\, dF_\varphi(\varphi). \end{aligned}$$

I am unaware of any functional form assumption for the distribution of landholdings F_φ that would allow equation (40) to be written in a more tractable form.

Comparison With the Model in the Paper

This model shares many features with the one presented in the paper. Like the model in the paper, larger producers are more likely to incur the fixed costs of export and, conditional on exporting, more likely to sell to destinations with greater prices. In addition, larger producers are more likely to export to destinations with higher prices, which implies that they are more sensitive to changes in destination prices than are small producers.

The major drawback of this framework is in the lack of tractability. It is critical to assume that the fixed costs of search f_i are equal to 1 so as to yield closed form solutions for the probability of export and the probability of trading with any particular destination. The structural estimates of the model in the paper, however, find that there is substantial variation across space in the estimated fixed costs. Furthermore, equation (40), which governs the trade flows, cannot be explicitly solved for a particular distribution, making it difficult to calculate how overall trade flows would respond to changes in the distribution of landholdings.

APPENDIX E: MODEL EXTENSIONS

In this appendix, I outline four extensions of the basic model. In the first extension, I introduce intermediary traders into the model by assuming that farmers first search for traders and traders subsequently search for a destination to which to sell. I show that the central prediction of the model—that larger farmers tend to sell to markets with higher prices—remains unchanged. In the second extension, I introduce a (noninformation related) fixed cost of export and show that the theoretical predictions of the basic model remain qualitatively unchanged. In the third extension, I show how the assumption of a perfectly competitive market within an island can be relaxed to allow for within-island search. The introduction of within-island information frictions is shown to exacerbate the differences in searching intensity between large and small producers. In the fourth extension, I allow for there to be (noninformation related) economies of scale in the price received by producers. I show that the model remains qualitatively unchanged and discuss how the structural estimates in the paper would be affected by economies of scale.

E.1. *Intermediary Traders*

Suppose that farmers, instead of searching across regions to sell directly to consumers, search locally for intermediaries (“traders”) to whom to sell their produce. After purchasing from producers, traders then conduct their own search across regions to determine where to sell to consumers.[57] In what follows, I show that the basic model can be extended to incorporate this additional stage of searching without substantially affecting the central predictions.

The timing of the model is as follows. First, a mass M of farmers[58] (referred to in the feminine) produce a homogeneous good according to the same production technology in the basic model. Second, each farmer chooses either to sell to local consumers directly for the local price p or to search for a trader (referred to in the masculine) to whom to sell her produce. If she chooses to search, she pays a fixed cost g to be paired with a trader, where the probability of being paired with a trader of type t is s_t (the "conspicuousness" of the trader). Upon being paired with a trader of type t, she observes his buying price p_t and chooses either to sell to him or to search again for another trader, continuing until she finds a trader to whom she is willing to sell. Third, after all farmers have completed their search process, all traders take their purchased produce of quantity q_t and conduct their own search process across regions, which is identical to the search process in the basic model. In particular, a trader can either choose to sell in his home region for price p or to search elsewhere.[59] If he chooses to search, he pays a fixed cost f and is matched with a region according to the search probability of that region and can either sell the produce there or continue to search, continuing until he finds a region in which he is willing to sell.

The model is solved by backward induction. In the third stage (trader search), given the quantity purchased, the trader searches across markets. Since the setup is identical to the basic model, the only difference in aggregate trade flows is that the relevant distribution is over traders rather than farmers. In the second stage (farmer search), traders choose the purchasing price they offer so as to maximize expected profits. The greater is the price the trader offers, the more produce he attracts from farmers, but the smaller is his profit margin. In particular, a trader of type s_t chooses p_t to maximize his expected profits π_t,

$$\pi_t(s_t) = \max_{\tilde{p}\in[p,\infty)} \big(R\big(q(\tilde{p}; s_t)\big) - \tilde{p}\big)q(\tilde{p}; s_t), \tag{41}$$

where $R(q)$ is the expected per-unit revenue the trader receives as a function of quantity, and $q(\tilde{p}; s_t)$ is the quantity that the trader purchases as a function of the price he offers and his conspicuousness in the market.

[57] Whether intermediaries promote or inhibit efficiency is a topic of much recent debate in the trade literature; see, for example, Antràs and Costinot (2011) and Bardhan, Mookherjee, and Tsumagari (2013). In my model, because intermediaries have a greater quantity to sell than farmers (since they are purchasing produce from multiple farmers), they are more willing to pay the fixed cost to search other markets. As a result, intermediaries improve the efficiency of trade.

[58] For readability, I omit the subscripts of the region when possible.

[59] In equilibrium, all traders participating in the market will search elsewhere, since selling locally will necessarily entail a loss.

I claim that (a) traders with greater s_t offer higher buying prices to farmers, (b) purchase a greater quantity of produce, and (c) receive a greater expected per-unit revenue from trade. To see this, note that the quantity purchased, q_t, by a trader as a function of the price offered p_t is

$$q(p_t) = s_t \theta f^{1-\theta} AM \int_p^{p_t} L(p)^{\theta-2}\, dp,$$

so that traders with greater s_t purchase a greater quantity of produce, which follows immediately from the fact that they offer higher buying prices to farmers, that is, (b) follows necessarily from (a). Similarly, the per-unit revenue from trade R as a function of quantity purchased, q, is

$$R(q) = K^{-1}\left(\frac{f}{q}\right).$$

Recall that $K'(p) = -(1 - F_{p/\tau}(p)) < 0$ and, hence, $K''(p) = F'_{p/\tau}(p) > 0$. By the inverse function theorem, $(K^{-1})(\cdot)$ is also decreasing and convex. Hence $R'(q) = -(K^{-1})'(\cdot)\frac{f}{q^2} > 0$, that is, the greater is the quantity purchased, the higher is the per-unit revenue. As a result, (c) follows necessarily from (b) and, hence, from (a).

It remains to show that traders with greater s_t offer higher buying prices to farmers. Recall that expected profits are chosen to maximize profits $\pi(p_t, s_t)$, where

$$\pi(p_t, s_t) = \bigl(R\bigl(q(p_t)\bigr) - p_t\bigr)q(p_t).$$

Note that by defining $\tilde{q}(p_t) \equiv \theta b^\theta f^{1-\theta} AM \int_p^{p_t} L(p)^{\theta-2}\, dp = \frac{q(p_t)}{s_t}$, I can write expected profits as

$$\pi(p_t, s_t) = s\bigl(R\bigl(s\tilde{q}(p_t)\bigr) - p_t\bigr)\tilde{q}(p_t).$$

I can also define $\tilde{\pi}(p_t, s_t) \equiv (R(s\tilde{q}(p_t)) - p_t)\tilde{q}(p_t) = \frac{\pi(p_t, s_t)}{s_t}$. Since the optimal price is characterized by the first order condition $\pi_1(p_t, s_t) = 0$, which is equivalent to $\tilde{\pi}_1(p_t, s_t) = 0$, by the implicit function theorem,

$$\frac{\partial}{\partial s_t} p_t(s_t) = -\frac{\tilde{\pi}_{12}(p_t, s_t)}{\tilde{\pi}_{11}(p_t, s_t)}. \tag{42}$$

Assuming that $\tilde{\pi}_{11}(p_t, s_t) < 0$ (which is necessary to yield an interior solution), equation (42) implies that traders with greater s_t will offer higher buying prices to farmers if and only if $\tilde{\pi}_{12}(p_t, s_t) > 0$. To show that this is the case requires some onerous algebra and calculus. Note that

$$\begin{aligned}
&\tilde{\pi}_1(p_t, s_t) = \tilde{q}(p_t)\left(\frac{\partial}{\partial p_t} R'\big(s\tilde{q}(p_t)\big) - 1\right) + \tilde{q}'(p_t)\big(R\big(s\tilde{q}(p_t)\big) - p_t\big) \\
\iff \quad &\tilde{\pi}_1(p_t, s_t) = \tilde{q}(p_t)\frac{\partial}{\partial p_t} R\big(s\tilde{q}(p_t)\big) \\
&\qquad\qquad + \tilde{q}'(p_t)R\big(s\tilde{q}(p_t)\big) - \tilde{q}'(p_t)p_t - \tilde{q}(p_t).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\tilde{\pi}_{12}(p_t, s_t) > 0 \\
\iff \quad &\frac{\partial}{\partial s}\left[\tilde{q}(p_t)\frac{\partial}{\partial p_t} R\big(s\tilde{q}(p_t)\big) + \tilde{q}'(p_t)R\big(s\tilde{q}(p_t)\big)\right] > 0 \\
\iff \quad &\tilde{q}'(p_t)\frac{\partial}{\partial s}\big[\tilde{q}(p_t)R'\big(s\tilde{q}(p_t)\big)s + R\big(s\tilde{q}(p_t)\big)\big] > 0 \\
\iff \quad &\frac{\partial}{\partial s}\big[\tilde{q}(p_t)R'\big(s\tilde{q}(p_t)\big)s + R\big(s\tilde{q}(p_t)\big)\big] > 0 \\
\iff \quad &2R'\big(s\tilde{q}(p_t)\big) + s\tilde{q}(p_t)R''\big(s\tilde{q}(p_t)\big) > 0,
\end{aligned} \tag{43}$$

since $\tilde{q}'(p_t) > 0$ and

$$\begin{aligned}
&\frac{\partial}{\partial s} R\big(sq(p_t)\big) = \tilde{q}(p_t)R'\big(s\tilde{q}(p_t)\big), \\
&\frac{\partial}{\partial s} R'\big(s\tilde{q}(p_t)\big)s\tilde{q}(p_t) = \tilde{q}(p_t)R'\big(s\tilde{q}(p_t)\big) + \tilde{q}(p_t)^2 R''\big(s\tilde{q}(p_t)\big)s.
\end{aligned}$$

Finally, note that

$$\begin{aligned}
R'(q) &= -\big(K^{-1}\big)'\left(\frac{f}{q}\right)\frac{f}{q^2}, \\
R''(q) &= \big(K^{-1}\big)''\left(\frac{f}{q}\right)\frac{f^2}{q^4} + 2\big(K^{-1}\big)'\left(\frac{f}{q}\right)\frac{f}{q^3},
\end{aligned}$$

so that

$$qR''(q) = \big(K^{-1}\big)''\left(\frac{f}{q}\right)\frac{f^2}{q^3} - 2R'(q).$$

Substituting into equation (43) yields

$$\tilde{\pi}_{12}(p_t, s_t) > 0 \quad \iff \quad \big(K^{-1}\big)''\left(\frac{f}{q}\right)\frac{f^2}{q^3} > 0,$$

which holds since $K(\cdot)$ is decreasing and convex, so that $K^{-1}(\cdot)$ is decreasing and convex too.

Farmer search proceeds identically to the basic model, although farmers now search across traders rather than destinations. In particular, each farmer has a reservation price she is willing to accept that is increasing in the size of her landholdings. This has two implications. First, larger farmers are more likely to sell to larger traders since larger farmers search more intensively for better prices and larger traders offer higher prices. The positive correlation between farmer size and trader size means that production from larger farms is ultimately sent to markets with, on average, better prices, just as in the basic model without intermediaries. Second, as in the basic model, there exists a threshold land size such that all farmers with landholdings greater than the threshold will choose to sell to a trader, while farmers with landholdings less than the threshold sell locally.

E.2. *Fixed Cost of Export*

In this subsection, I extend the model to incorporate fixed costs of export. Assume that for a producer to export, she must pay a fixed cost g_i. This fixed cost is incurred prior to searching any particular destination market; once it is incurred, a producer must then pay the fixed information cost f_i to search each subsequent market. The fixed cost g_i could represent the costs associated with procuring a ship to transport the produce (a cost that must be incurred regardless of destination).

The inclusion of the fixed cost g_i will reduce the number of producers willing to export. In particular, consider a producer of size φ in region i deciding between selling locally for p_i and entering the export market. The value to the farmer is

$$V_i(p_i;\varphi) = \max\left\{\varphi p_i, \int V_i(p';\varphi)\, dF^i_{p/\tau}(p') - (f_i + g_i)\right\}.$$

From the basic model, this implies that the threshold landholding above which a producer will choose to export, $\varphi_i^E(p_i)$, is

$$\varphi_i^E(p_i) = \frac{f_i + g_i}{K_i(p_i)}. \tag{44}$$

Since g_i is only incurred once prior to exporting, once a producer has entered the export market, the fixed cost of export no longer affects her value of search as it is a sunk cost. As a result, the threshold landholding above which a producer continues to search after exporting is the same as in the basic model, that is, $\varphi_i^*(p) = \frac{f_i}{K_i(p)}$. Because the fixed cost of export g_i increased the minimum size of the exporting producer, the lowest price that an exporting producer will be willing to accept is $p_i^E > p_i$, where

$$\frac{f_i + g_i}{K_i(p_i)} = \frac{f_i}{K_i(p_i^E)} \iff p_i^E = K_i^{-1}\left(\frac{f_i}{f_i + g_i} K_i(p_i)\right) \geq p_i. \tag{45}$$

Hence, the introduction of a fixed cost of exporting creates a wedge between the domestic price and the minimum export price (net of transportation costs) at which exports occur,

$$Q_{ij} > 0 \iff \frac{p_j}{\tau_{ij}} \geq p_i + \alpha_i, \tag{46}$$

where $\alpha_i \equiv p_i^E - p_i \geq 0$. Similarly, the equation governing the extensive margin of trade becomes

$$Q_{ij} = A_i M_i \theta_i f_i^{1-\theta_i} s_{ij} \int_{p_i+\alpha_i}^{p_j/\tau_{ij}} K_i(p)^{\theta_i - 2}\, dp. \tag{47}$$

It is informative to compare equation (46) to the corresponding equation when there is only a fixed cost of export, that is, when information is complete. From equation (44), only farmers with land holdings greater than $\varphi_i^E(p_i) = \frac{g_i}{K_i(p_i)}$ will export. Since information is complete, all farmers choosing to export will sell to the destination with the greatest price, so that

$$Q_{ij} > 0 \iff \frac{p_j}{\tau_{ij}} = \max_{k\in\{1,\dots,N\}} \frac{p_k}{\tau_{ik}}$$

or, equivalently,

$$Q_{ij} > 0 \iff \frac{p_j}{\tau_{ij}} = p_i + \alpha_i, \tag{48}$$

where $\alpha_i \equiv \max_{k\in\{1,\dots,N\}} \frac{p_k}{\tau_{ik}} - p_i > 0$. Hence, just as in the basic model, incorporating information frictions alters the complete information arbitrage equation by replacing an equality with an inequality.

E.3. *Within-Island Information Frictions*

In the basic model, each island is assumed to have a single perfectly competitive market. In this subsection, I relax this assumption to allow for information frictions within an island. I model these within-island information frictions by assuming that each island has a number of markets that must be searched after an island is chosen.[60] I show that this additional friction exacerbates the difference between large and small producers' search strategies.

Assume that farmers first search across islands, observing the mean island price; then upon choosing an island, they search across markets within the island to find a destination in which to sell. Let the price in a market j in island g be $p_g \varepsilon_j$, where p_g is a scalar common to all regions in island g (the correlated shock) and ε_j is a market-specific idiosyncratic shock. (The basic search process is a special case of this search process where ε_j is equal to 1 for all markets within an island.) Let p_g be i.i.d. across islands with cumulative distribution function $F_{\bar{p}}$ and let ε_j be i.i.d. across regions within island g with cumulative distribution function F^g_ε. For simplicity, assume that the within island distribution of ε is the same for all islands, that is, $F^g_\varepsilon = F_\varepsilon \ \forall g$.

This two-stage search problem can be solved using backward induction. First consider the second-stage decision. In particular, consider a farmer with landholdings φ who has chosen to search island g with correlated shock p_g and is now searching across markets on the island. Let f^i and f^m denote the fixed cost of searching an additional island and market, respectively. The value function of a farmer who has discovered a market with price $p_g \varepsilon_j$ is

$$V^m(p_g, \varepsilon_j) = \max\left\{\varphi p_g \varepsilon_j, \int V(p_g, \varepsilon_j)\, dF_\varepsilon - f^m\right\}. \tag{49}$$

As in the basic model, this problem yields a reservation idiosyncratic shock $\bar{\varepsilon}(\varphi)$ such that a firm is indifferent between selling in that market and searching for another market:

$$\varphi p_g \bar{\varepsilon}(\varphi) = \int V^m(p_g, \varepsilon_j)\, dF_\varepsilon - f^m. \tag{50}$$

Substituting equation (50) into equation (49) yields

$$V^m(p_g, \varepsilon_j) = \varphi p_g \max\{\varepsilon_j, \bar{\varepsilon}(\varphi)\}. \tag{51}$$

Substituting equation (51) into equation (50) yields

$$\begin{aligned} &\varphi p_g \bar{\varepsilon}(\varphi) = \varphi p_g \left[\bar{\varepsilon}(\varphi) F_\varepsilon\big(\bar{\varepsilon}(\varphi)\big) + \int_{\bar{\varepsilon}(\varphi)} \varepsilon\, dF_\varepsilon(\varepsilon)\right] - f^m \quad \Longleftrightarrow \\ &f^m = \varphi p_g K_\varepsilon\big(\bar{\varepsilon}(\varphi)\big), \end{aligned} \tag{52}$$

where $K_\varepsilon(\varepsilon) \equiv \int_\varepsilon (\varepsilon' - \varepsilon)\, dF_\varepsilon(\varepsilon')$.

[60] An alternative interpretation is that producers have to search across consumers who have different reservation prices once they have decided to sell to a particular market.

From equation (52), it is clear that $\frac{\partial}{\partial\varphi}\bar{\varepsilon}(\varphi) > 0$; as in the basic setup, larger farmers have higher reservation prices than small farmers since the fixed cost of search comprises a smaller fraction of total revenue.

Define $V(p_g;\varphi)$ as the expected value of arriving on an island with correlated shock p_g. From equation (51),

$$V^m(p_g;\varphi) \equiv \int V^m(p_g,\varepsilon_j)\,dF_\varepsilon(\varepsilon_j) = \varphi p_g G\big(\bar{\varepsilon}(\varphi)\big), \tag{53}$$

where $G_\varepsilon(\varepsilon) \equiv \varepsilon F_\varepsilon(\varepsilon) + \int_\varepsilon \varepsilon'\,dF_\varepsilon(\varepsilon')$. Since $\bar{\varepsilon}(\varphi)$ is monotonically increasing in φ, so too is $G_\varepsilon(\varphi)$.

The second-stage search hence yields a value of arriving in an island with correlated shock p_g as a function of landholdings φ. In the first stage, farmers will search across island groups, knowing how the observed correlated shock will affect the expected value of the second stage. In particular, consider a farmer who has arrived at an island with correlated shock p_g. Then her value function is

$$V^i(p_g;\varphi) = \max\left\{p_g\varphi G_\varepsilon\big(\bar{\varepsilon}(\varphi)\big), \int V^i(p;\varphi)\,dF_{\bar{p}}(p) - f^i\right\}. \tag{54}$$

As above, the solution to equation (54) yields a reservation correlated shock $\bar{p}_g(\varphi)$ such that

$$\bar{p}_g(\varphi)\varphi G_\varepsilon\big(\bar{\varepsilon}(\varphi)\big) = \int V^i(p;\varphi)\,dF_{\bar{p}}(p) - f^i,$$

so that

$$V^i(p_g;\varphi) = \varphi G_\varepsilon\big(\bar{\varepsilon}(\varphi)\big)\max\big\{p_g, \bar{p}_g(\varphi)\big\}$$

and

$$\int V^i(p;\varphi)\,dF_{\bar{p}}(p) = \varphi G_\varepsilon\big(\bar{\varepsilon}(\varphi)\big)\int \max\big\{p, \bar{p}_g(\varphi)\big\}\,dF_{\bar{p}}(p).$$

Again, the solution to equation (54) yields a reservation correlated shock $\bar{p}_g(\varphi)$ such that

$$\varphi G_\varepsilon\big(\bar{\varepsilon}(\varphi)\big)\bar{p}_g(\varphi) = \varphi G_\varepsilon\big(\bar{\varepsilon}(\varphi)\big)\int \max\big\{p, \bar{p}_g(\varphi)\big\}\,dF_{\bar{p}}(p) - f^i. \tag{55}$$

As above, combining equations (54) and (55) yields

$$f^i = \varphi G_\varepsilon\big(\bar{\varepsilon}(\varphi)\big)K_p\big(\bar{p}_g(\varphi)\big), \tag{56}$$

where $K_p(p) \equiv \int_p (p'-p)F_{\bar{p}}(p')$.

It is clear from equation (56) that $\bar{p}_g(\varphi)$ is strictly increasing in φ. Larger farmers have higher reservation correlated shocks and hence will search across more islands before choosing an island in which to search for a market. This effect is amplified by the existence of $G(\varphi)$, which captures the fact that in the second stage, larger farmers will search more intensively within an island and hence will expect to find, on average, better prices within the island. When searching across islands in the first stage, larger farmers will hence place more value on the search process, giving an additional incentive to find an island with better correlated shocks. Hence, the existence of correlated shocks only serves to further distinguish the search behavior of large and smaller farmers.

E.4. *Economies of Scale*

Suppose that larger farmers received higher prices than smaller farmers because of economies of scale rather than information frictions. To model this, I suppose that a farmer of size $\varphi \in [1, \infty)$ selling to a market with price p_j actually receives a price of $p_j \varphi^{\beta}$, where $\beta > 0$ indicates the existence of economies of scale. Hence, the marginal farmer receives no economy of scale, but larger farmers get higher prices.

Farmer Search

We can then write the search process of an individual farmer as

$$V_i(p; \varphi) = \max\left\{\varphi^{\beta+1} p, \int_{p_i^{\min}}^{p_i^{\max}} V_i(p'; \varphi)\, dF^i_{p/\tau}(p') - f_i\right\}. \tag{57}$$

Hence a farmer will choose to sell if and only if

$$\varphi^{\beta+1} p \geq \int_{p_i^{\min}}^{p_i^{\max}} V_i(p'; \varphi)\, dF^i_{p/\tau}(p') - f_i.$$

Let $\bar{p}_i(\varphi)$ be the reservation price of the farmer who is indifferent between selling and not, that is,

$$\varphi^{\beta+1} \bar{p}_i(\varphi) = \int_{p_i^{\min}}^{p_i^{\max}} V_i(p'; \varphi)\, dF^i_{p/\tau}(p') - f_i, \tag{58}$$

so that equation (57) becomes

$$V_i(p; \varphi) = \max\left\{\varphi^{\beta+1} p, \varphi^{\beta+1} \bar{p}_i(\varphi)\right\}. \tag{59}$$

Then integrating equation (59) over p yields

$$\begin{aligned}
(60) \qquad & \int_{p_i^{\min}}^{p_i^{\max}} V_i(p';\varphi)\, dF^i_{p/\tau}(p') \\
&= \int_{\bar{p}_i(\varphi)}^{p_i^{\max}} \varphi^{\beta+1} p\, dF^i_{p/\tau}(p') + F^i_{p/\tau}(\bar{p}_i(\varphi))\varphi^{\beta+1}\bar{p}_i(\varphi).
\end{aligned}$$

Substituting equation (60) into equation (58) finally yields

$$\begin{aligned}
(61) \qquad & \varphi^{\beta+1}\bar{p}_i(\varphi) = \int_{\bar{p}_i(\varphi)}^{p_i^{\max}} \varphi^{\beta+1} p\, dF^i_{p/\tau}(p') \\
&\qquad + F^i_{p/\tau}(\bar{p}_i(\varphi))\varphi^{\beta+1}\bar{p}_i(\varphi) - f_i \quad \Longleftrightarrow \\
& f_i = \int_{\bar{p}_i(\varphi)}^{p_i^{\max}} \varphi^{\beta+1} p\, dF^i_{p/\tau}(p') - \left(1 - F^i_{p/\tau}(\bar{p}_i(\varphi))\right)\varphi^{\beta+1}\bar{p}_i(\varphi) \quad \Longleftrightarrow \\
& f_i = \varphi^{\beta+1}\int_{\bar{p}_i(\varphi)}^{p_i^{\max}} (p - \bar{p}_i(\varphi))\, dF^i_{p/\tau}(p').
\end{aligned}$$

As in the basic model, it is evident from equation (61) that the reservation price $\bar{p}_i(\varphi)$ is strictly increasing in φ, so that as in the basic model, we can solve for the threshold landholdings such that a farmer is indifferent between selling and continuing to search given a market price p:

$$(62) \qquad \varphi_i^*(p) = \left(\frac{f_i}{K_i(p)}\right)^{1/(1+\beta)},$$

where $K_i(p) \equiv \int_p^{p_i^{\max}} (p' - p)\, dF^i_{p/\tau}(p)$ is the "value of search." Hence, increasing the economies of scale serves to reduce the threshold size of a farmer who is willing to search.

Total Trade

Let $N_i^r(\varphi)$ be the density of farmers of size φ who are still searching after r searches. Then we have that the total trade flows from i to j after r searches is

$$Q_{ij}^r = A_i M_i s_{ij} \int_{\varphi_i^*(p_i)}^{\varphi_i^*(p_j/\tau_{ij})} \varphi N_i^r(\varphi)\, dF^i_\varphi(\varphi).$$

As in the basic model, by imposing the Pareto distribution and the law of large numbers, we have

$$Q_{ij}^r = A_i M_i s_{ij} \theta_i \int_{\varphi_i^*(p_i)}^{\varphi_i^*(p_j/\tau_{ij})} \varphi^{-\theta_i} F_{p/\tau}^i\big(\bar{p}_i(\varphi)\big)^r \, d\varphi \quad \Longleftrightarrow$$

$$Q_{ij}^r = A_i M_i s_{ij} \theta_i \int_{p_i}^{p_j/\tau_{ij}} \big(\varphi_i^*(p)\big)^{-\theta_i} F_{p/\tau}^i(p)^r \frac{\partial}{\partial p} \varphi_i^*(p) \, dp,$$

where the second line used integration by substitution. From equation (62) and the definition of $K_i(p)$, we have

$$\frac{\partial}{\partial p} \varphi_i^*(p) = \frac{1}{1+\beta} \left(\frac{f_i}{K_i(p)}\right)^{-\beta/(1+\beta)} \frac{f_i}{K_i(p)^2} \big(1 - F_{p/\tau}^i(p)\big),$$

so that

$$\begin{aligned} Q_{ij}^r &= A_i M_i s_{ij} \frac{\theta_i}{1+\beta} \int_{p_i}^{p_j/\tau_{ij}} \left(\frac{f_i}{K_i(p)}\right)^{-\theta_i/(1+\beta)} F_{p/\tau}^i(p)^r \\ &\quad \times \left(\frac{f_i}{K_i(p)}\right)^{-\beta/(1+\beta)} \frac{f_i}{K_i(p)^2} \big(1 - F_{p/\tau}^i(p)\big) \, dp \quad \Longleftrightarrow \\ &= A_i M_i s_{ij} \frac{\theta_i}{1+\beta} f_i^{(1-\theta_i)/(1+\beta)} \\ &\quad \times \int_{p_i}^{p_j/\tau_{ij}} K_i(p)^{(\theta_i+\beta)/(1+\beta)-2} F_{p/\tau}^i(p)^r \big(1 - F_{p/\tau}^i(p)\big) \, dp. \end{aligned}$$

Total bilateral trade flows are hence

$$\begin{aligned} Q_{ij} &= \sum_{r=0}^{\infty} Q_{ij}^r \quad \Longleftrightarrow \\ &= \sum_{r=0}^{\infty} A_i M_i s_{ij} \frac{\theta_i}{1+\beta} f_i^{(1-\theta_i)/(1+\beta)} \\ &\quad \times \int_{p_i}^{p_j/\tau_{ij}} K_i(p)^{(\theta_i+\beta)/(1+\beta)-2} F_{p/\tau}^i(p)^r \big(1 - F_{p/\tau}^i(p)\big) \, dp \quad \Longleftrightarrow \\ &= A_i M_i s_{ij} \frac{\theta_i}{1+\beta} f_i^{(1-\theta_i)/(1+\beta)} \\ &\quad \times \int_{p_i}^{p_j/\tau_{ij}} K_i(p)^{(\theta_i+\beta)/(1+\beta)-2} \big(1 - F_{p/\tau}^i(p)\big) \sum_{r=0}^{\infty} F_{p/\tau}^i(p)^r \, dp \end{aligned}$$

$$= A_i M_i s_{ij} \frac{\theta_i}{1+\beta} f_i^{(1-\theta_i)/(1+\beta)} \times \int_{p_i}^{p_j/\tau_{ij}} K_i(p)^{(\theta_i+\beta)/(1+\beta)-2}\, dp.$$

The total openness of a region is

$$\begin{aligned}
\Lambda_i &= \int_{\varphi^*(p_i)}^{\infty} \varphi\, dF_{\varphi}^{i}(\varphi) \\
&= \theta_i \int_{\varphi^*(p_i)}^{\infty} \varphi^{-\theta_i}\, d\varphi \\
&= \frac{\theta_i}{\theta_i - 1} \varphi^*(p_i)^{1-\theta_i} \\
&= \frac{\theta_i}{\theta_i - 1} \left(\frac{K_i(p_i)}{f_i} \right)^{(\theta_i-1)/(1+\beta)},
\end{aligned}$$

so that

$$f_i = \left(\frac{\theta_i}{\theta_i - 1} \right)^{(1+\beta)/(\theta_i-1)} K_i(p_i) \Lambda_i^{-(1+\beta)/(\theta_i-1)}. \tag{63}$$

Since $\Lambda_i \in [0, 1]$, for a given $K_i(p_i)$ and Λ_i, increases in the returns to scale (i.e., a higher β) actually lead to *higher* estimates of the fixed cost f_i. This is because greater returns to scale result in a larger fraction of farmers searching (see equation (62)), so that higher fixed costs are required to match the observed openness.

Note that the elasticity of trade flows to changes in the destination price can be written as

$$\begin{aligned}
\frac{\partial \ln Q_{ij}}{\partial \ln \frac{p_j}{\tau_{ij}}} &= \frac{K_i\left(\frac{p_j}{\tau_{ij}}\right)^{(\theta_i+\beta)/(1+\beta)-2}}{\int_{p_i}^{p_j/\tau_{ij}} K_i(p)^{(\theta_i+\beta)/(1+\beta)-2}\, dp} \\
&= \left(\int_{p_i}^{p_j/\tau_{ij}} \left(\frac{K_i(p)}{K_i\left(\frac{p_j}{\tau_{ij}}\right)} \right)^{(\theta_i+\beta)/(1+\beta)-2} dp \right)^{-1}.
\end{aligned}$$

Since $K_i(p) \geq K_i(\frac{p_j}{\tau_{ij}})$ for all $p \in [p_i, \frac{p_j}{\tau_{ij}}]$, the elasticity $\frac{\partial \ln Q_{ij}}{\partial \ln \frac{p_j}{\tau_{ij}}}$ is decreasing in $\frac{\theta_i+\beta}{1+\beta} - 2$. Since $\frac{\theta_i+\beta}{1+\beta} - 2$ is decreasing in β, this implies that the elasticity is increasing in β, that is, trade becomes more responsive to changes in destination prices the greater are the economies of scale.

What Happens When Observed Prices Are Gross of Economies of Scale?

Suppose we did not observe the market price p_i, but instead observed the average price $\tilde{p}_i$ that all producers received taking into account the economies of scale, that is, $\tilde{p}_i = E[p_i \varphi^\beta]$. How would this affect the estimated fixed costs of search?

It is straightforward to show that the average price net of economies of scale is simply a multiplier of the market price:

$$\tilde{p}_i = \frac{\theta_i}{\theta_i - \beta} p_i.$$

Define $\tilde{K}_i(\tilde{p}_i) \equiv \int_{\tilde{p}_i}^{\tilde{p}_i^{\max}} (\tilde{p}' - \tilde{p}_i)\, dF^i_{p/\tau}(\tilde{p}')$ as the value of search constructed using observed average prices net of economies of scale. It can be shown that $K_i(p)$ is homogeneous of degree 1 in p, which immediately implies that

$$\tilde{K}_i(\tilde{p}_i) = \left(\frac{\theta_i}{\theta_i - \beta}\right) K_i(p_i), \tag{64}$$

that is, the value of search using observed average prices net of economies of scale will simply be the value of search constructed using (unobserved) market prices multiplied by $\frac{\theta_i}{\theta_i - \beta} > 1$. That is, because the observed prices are larger (due to economies of scale) than the actual market prices, the observed value of search will be larger than the value of search based only on market prices. (This is because the value of search function is homogeneous of degree 1 in prices.)

We can now compare the fixed costs of search estimated in the paper ignoring economies of scale (f_i^{est}) to the fixed costs of search given the economies of scale (f_i^{true}). From equation (63), the true fixed cost of search is

$$f_i^{\text{true}} = \left(\frac{\theta_i}{\theta_i - 1}\right)^{(1+\beta)/(\theta_i - 1)} K_i(p_i) \Lambda_i^{-(1+\beta)/(\theta_i - 1)},$$

whereas the fixed cost estimated in the paper, given by equation (64), can be written as

$$f_i^{\text{est}} = \left(\frac{\theta_i}{\theta_i - 1}\right)^{1/(\theta_i - 1)} \left(\frac{\theta_i}{\theta_i - \beta}\right) K_i(p_i) \Lambda_i^{-1/(\theta_i - 1)}.$$

Hence, we will tend to underestimate the true fixed costs of search if and only if

$$f_i^{\text{est}} < f_i^{\text{true}} \iff \Lambda_i < \left(\frac{\theta_i}{\theta_i - 1}\right) \Big/ \left(\frac{\theta_i}{\theta_i - \beta}\right)^{(\theta_i - 1)/\beta}. \tag{65}$$

Generally, there exist parameter constellations of Λ_i, θ_i, and β for which inequality (65) may or may not hold. However, for reasonable values of Λ_i, θ_i, and β (i.e., $\beta \in [0, 0.5]$, $\theta_i \in [1.5, 5]$, and $\Lambda_i \in [0, 0.5]$), inequality (65) will hold, suggesting that, if anything, the estimated fixed costs presented in the paper underestimate the true fixed costs of search if there are economies of scale.

E.5. *Specifying Consumer's Income*

In this model extension, I show how the model can be extended to specify the source of consumer's income, thereby making the model fully general equilibrium. Suppose that consumers in each location i produce a homogeneous, freely traded good with productivity C_i and have Cobb–Douglass preferences over this homogeneous good h and the rice r that is subject to information and transportation cost frictions:

$$U_i = r_i^{\alpha} h_i^{1-\alpha}.$$

Since the homogeneous good is costlessly traded, it has the same price in all locations, which I normalize to 1. Consumers hence have an income of C_i, so that the equilibrium price in region i for rice can be written as

$$p_i = \frac{\alpha C_i}{q_i},$$

so that $D_i(q_i) = \frac{\alpha C_i}{q_i}$.

APPENDIX F: ADDITIONAL MODEL DERIVATIONS

F.1. *Producer Revenue*

In this appendix, I show that a producer's expected revenue is equal to her reservation price. Consider a farmer from region i with production φ and optimal reservation price $\bar{p}_i(\varphi)$. The probability that the producer will sell in a particular region is $1 - F^i_{p/\tau}(\bar{p}_i(\varphi))$. The probability that a farmer sells to the nth searched regions (after the home region) is $(1 - F^i_{p/\tau}(\bar{p}_i(\varphi)))F^i_{p/\tau}(\bar{p}_i(\varphi))^n$. Hence, the expected revenue that a farmer will receive from the search process, $R(\varphi)$, is

$$E\big(R(\varphi)\big) = \sum_{n=0}^{\infty} F^i_{p/\tau}\big(\bar{p}_i(\varphi)\big)^n \big(1 - F^i_{p/\tau}\big(\bar{p}_i(\varphi)\big)\big) \times \left(\frac{\varphi \int_{\bar{p}(\varphi)}^{\infty} p \, dF^i_{p/\tau}(p)}{(1 - F^i_{p/\tau}(\bar{p}_i(\varphi)))} - nf_i \right) \iff$$

$$E\big(R(\varphi)\big) = \frac{1}{(1 - F(\bar{p}_{p/\tau}(\varphi)))} \left(\varphi \int_{\bar{p}(\varphi)}^{\infty} p \, dF(p) - f_i \right).$$

Rearranging yields

$$f_i = \varphi \int_{\bar{p}(\varphi)}^{\infty} \left(p - \frac{E(R(\varphi))}{\varphi} \right) dF(p). \tag{67}$$

Comparing equation (67) to equation (4) yields that the expected revenue a farmer will receive from searching is simply the product of her reservation price and the quantity produced, that is,

$$E\big(R(\varphi)\big) = \varphi \bar{p}(\varphi) = \varphi K^{-1}\left(\frac{f_i}{\varphi}\right),$$

so that the expected per unit revenue is equal to the reservation price, as required.

It is also possible to determine the total revenue by all firms in region i. Since the expected revenue for each firm is its reservation price $\bar{p}(\varphi)$, the total revenue among all firms in a region, R_i, is

$$\begin{aligned} R_i &= A_i M_i \left[p_i \int_{b_i}^{\varphi^*(p_i)} \varphi \, dF_\varphi^i(\varphi) + \int_{\varphi^*(p_i)}^{\infty} \bar{p}(\varphi)\varphi \, dF_\varphi^i(\varphi) \right] \\ &= \theta_i A_i M_i \left[p_i \int_{b_i}^{\varphi^*(p_i)} \varphi^{-\theta_i} \, d\varphi + \int_{\varphi^*(p_i)}^{\infty} \bar{p}(\varphi)\varphi^{-\theta_i} \, d\varphi \right], \end{aligned}$$

where the second equality comes from the assumption that productivities are distributed according to a Pareto distribution. The first term in the brackets is equal to $\frac{p_i}{\theta_i - 1}(1 - \varphi^*(p_i)^{1-\theta_i})$. The second integral can be calculated using integration by parts followed by a change of variables:

$$\begin{aligned} \int_{\varphi^*(p_i)}^{\infty} \bar{p}(\varphi)\varphi^{-\theta_i} \, d\varphi &= \frac{1}{\theta_i - 1} \left[\int_{\varphi^*(p_i)}^{\infty} \varphi^{1-\theta_i} \bar{p}'(\varphi) \, d\varphi - \bar{p}(\varphi)\varphi^{1-\theta_i} \Big|_{\varphi^*(p_i)}^{\infty} \right] \\ &= \frac{1}{\theta_i - 1} \left[\int_{p_i}^{p_i^{\max}} \big(\varphi^*(p)\big)^{1-\theta_i} \, dp + p_i \big(\varphi^*(p_i)\big)^{1-\theta_i} \right]. \end{aligned}$$

Using the fact that $\varphi^*(p) \equiv \frac{f_i}{K(p)}$ yields

$$R_i = \frac{\theta_i}{\theta_i - 1} A_i M_i \left(p_i + f_i^{1-\theta_i} \int_{p_i}^{p_i^{\max}} K(p)^{\theta_i - 1} \, dp \right).$$

By adding and subtracting $\bar{p}$ in the parentheses, this can be written as

$$R_i = \frac{\theta_i}{\theta_i - 1} A_i M_i \left(\bar{p} - \int_{p_i}^{p_i^{\max}} \left(1 - \left(\frac{K(p)}{f_i} \right)^{\theta_i - 1} \right) dp \right).$$

Since the total quantity produced is $\frac{\theta_i}{\theta_i-1}A_iM_i$, the term in the parentheses indicates the per-unit average profits of farmers. In the absence of information frictions, all farmers would sell their produce at the highest price net of transportation cost $p_i^{\max}$, so that $\int_{p_i}^{p_i^{\max}}(1-(\frac{K_i(p)}{f_i})^{\theta_i-1})\,dp$ captures the per-unit loss in profits due to the existence of information frictions. This is intuitive; from equation (11), $(\frac{K_i(p)}{f_i})^{\theta_i-1}$ is the fraction of production that is sold for a price at least as great as p, so that the per-unit loss in profits is at most $p_i^{\max}-p_i$, which would occur if every producer sold domestically. Furthermore, declines in the fixed cost of search result in increases in the reservation price that farmers are willing to accept, increasing $(\frac{K_i(p)}{f_i})^{\theta_i-1}$ and reducing the loss in profits due to information frictions.

F.2. *Elasticity of Trade Is Increasing in Producer Heterogeneity*

In this subsection, I formally derive the elasticity of trade and prove that it is strictly increasing in producer heterogeneity (i.e., decreasing in θ_i). Taking logs of equation (9) yields

$$\ln Q_{ij}=\ln A_i+\ln M_i+\ln\theta_i+(1-\theta_i)\ln f_i+\ln s_{ij}$$
$$+\ln\int_{p_i}^{p_j/\tau_{ij}}K_i(p)^{\theta_i-2}\,dp$$

so that

$$\frac{\partial\ln Q_{ij}}{\partial\ln\frac{p_j}{\tau_{ij}}}=\frac{p_j}{\tau_{ij}}\left(\frac{\partial}{\partial\frac{p_j}{\tau_{ij}}}\int_{p_i}^{p_j/\tau_{ij}}K_i(p)^{\theta_i-2}\,dp\right.$$
$$\left.\Big/\int_{p_i}^{p_j/\tau_{ij}}K_i(p)^{\theta_i-2}\,dp\right)\iff$$

$$\frac{\partial\ln Q_{ij}}{\partial\ln\frac{p_j}{\tau_{ij}}}=\frac{p_j}{\tau_{ij}}\left(\frac{K_i\left(\frac{p_j}{\tau_{ij}}\right)^{\theta_i-2}}{\int_{p_i}^{p_j/\tau_{ij}}K_i(p)^{\theta_i-2}\,dp}\right)\iff$$

$$\frac{\partial\ln Q_{ij}}{\partial\ln\frac{p_j}{\tau_{ij}}}=\frac{p_j}{\tau_{ij}}\left(\int_{p_i}^{p_j/\tau_{ij}}\left(\frac{K_i(p)}{K_i\left(\frac{p_j}{\tau_{ij}}\right)}\right)^{\theta_i-2}dp\right)^{-1},$$

as claimed in the text (and where the second line used the Leibniz rule).

I now show that $\frac{\partial \ln Q_{ij}}{\partial \ln \frac{p_j}{\tau_{ij}}}$ is strictly decreasing in θ_i. This can be done by calculating the cross partial derivative

$$
\begin{aligned}
(68)\qquad & \frac{\partial}{\partial \theta_i}\left(\frac{\partial \ln Q_{ij}}{\partial \ln \frac{p_j}{\tau_{ij}}}\right) \\
&= \frac{\partial}{\partial \theta_i}\frac{p_j}{\tau_{ij}}\left(\int_{p_i}^{p_j/\tau_{ij}}\left(\frac{K_i(p)}{K_i\left(\frac{p_j}{\tau_{ij}}\right)}\right)^{\theta_i-2} dp\right)^{-1} \iff \\
&= -\frac{p_j}{\tau_{ij}}\left(\int_{p_i}^{p_j/\tau_{ij}}\left(\frac{K_i(p)}{K_i\left(\frac{p_j}{\tau_{ij}}\right)}\right)^{\theta_i-2} dp\right)^{-2} \\
&\quad \times \left(\frac{\partial}{\partial \theta_i}\int_{p_i}^{p_j/\tau_{ij}}\left(\frac{K_i(p)}{K_i\left(\frac{p_j}{\tau_{ij}}\right)}\right)^{\theta_i-2} dp\right) \iff \\
&= -\frac{p_j}{\tau_{ij}}\left(\int_{p_i}^{p_j/\tau_{ij}}\left(\frac{K_i(p)}{K_i\left(\frac{p_j}{\tau_{ij}}\right)}\right)^{\theta_i-2} dp\right)^{-2} \\
&\quad \times \left(\int_{p_i}^{p_j/\tau_{ij}}\frac{\partial}{\partial \theta_i}\left(\frac{K_i(p)}{K_i\left(\frac{p_j}{\tau_{ij}}\right)}\right)^{\theta_i-2} dp\right) \iff \\
&= -\frac{p_j}{\tau_{ij}}\left(\int_{p_i}^{p_j/\tau_{ij}}\left(\frac{K_i(p)}{K_i\left(\frac{p_j}{\tau_{ij}}\right)}\right)^{\theta_i-2} dp\right)^{-2} \\
&\quad \times \left(\int_{p_i}^{p_j/\tau_{ij}}\ln\left(\frac{K_i(p)}{K_i\left(\frac{p_j}{\tau_{ij}}\right)}\right)\left(\frac{K_i(p)}{K_i\left(\frac{p_j}{\tau_{ij}}\right)}\right)^{\theta_i-2} dp\right).
\end{aligned}
$$

Recall from equation (6) that $K_i(p)$ is positive and strictly decreasing in p. This implies that $K_i(p) \geq K_i(\frac{p_j}{\tau_{ij}})$ for all $p \in [p_i, \frac{p_j}{\tau_{ij}}]$ (with the equality strict if $p < \frac{p_j}{\tau_{ij}}$). Hence $\ln(\frac{K_i(p)}{K_i(\frac{p_j}{\tau_{ij}})}) \geq 0$ for all $p \in [p_i, \frac{p_j}{\tau_{ij}}]$ and $\ln(\frac{K_i(p)}{K_i(\frac{p_j}{\tau_{ij}})}) > 0$ for all $p \in [p_i, \frac{p_j}{\tau_{ij}})$, which in turn implies that $\int_{p_i}^{p_j/\tau_{ij}} \ln(\frac{K_i(p)}{K_i(\frac{p_j}{\tau_{ij}})})(\frac{K_i(p)}{K_i(\frac{p_j}{\tau_{ij}})})^{\theta_i-2}\, dp > 0$. Since

$K_i(p)$ is positive, equation (68) implies

$$\frac{\partial}{\partial\theta_i}\left(\frac{\partial \ln Q_{ij}}{\partial \ln \dfrac{p_j}{\tau_{ij}}}\right) < 0,$$

that is, $\frac{\partial \ln Q_{ij}}{\partial \ln \frac{p_j}{\tau_{ij}}}$ is strictly decreasing in θ_i.

F.3. *The No-Arbitrage Equation Improves as the Fixed Costs of Search Fall*

In this section, I show that as the fixed costs of search fall in a particular location, the difference in price between that location and the price net of transportation costs in any location to which it exports will fall, that is,

$$-\frac{\partial}{\partial f_i}\left(\frac{p_j}{\tau_{ij}} - p_i\right) < 0,$$

where I suppress the crop subscript for readability. As in the main text, suppose there are a large number of locations; this implies that there is no effect of the fixed cost of search in country i on p_j for any $j \neq i$. Hence, it is sufficient to show that

$$\frac{\partial p_i}{\partial f_i} < 0,$$

that is, lowering the fixed cost of search increases the price in location i. Since $p_i = D_i(r_i)$ and $\frac{\partial}{\partial r_i} D_i < 0$, this is true if and only if $\frac{\partial r_i}{\partial f_i} > 0$, that is, lowering the fixed cost of search reduces the quantity consumed locally. From Appendix A.1, we can write the quantity consumed in location i using the expression

$$\begin{aligned} r_i = {} & \frac{\theta_i}{\theta_i - 1} A_i M_i \left(1 - \left(\frac{K_i(p_i)}{f_i}\right)^{\theta_i - 1}\right) \\ & + \sum_{j \neq i} \mathbf{1}\left\{\frac{p_i}{\tau_{ji}} \geq p_j\right\} A_j M_j \theta_j f_j^{1-\theta_j} s_{ji} \int_{p_j}^{p_i/\tau_{ji}} K_j(p)^{\theta_i - 2}\, dp, \end{aligned}$$

where the first term captures the amount of local production that is consumed locally and the second term captures the amount of production imported. Note that the right hand side is implicitly a function of r_i through the price p_i. Using the implicit function theorem, we then have

$$\frac{\partial r_i}{\partial f_i} = -\frac{\partial G_i(r_i, f_i)}{\partial f_i} \Big/ \frac{\partial G_i(r_i, f_i)}{\partial r_i},$$

where $G_i(r_i, f_i) \equiv r_i - \frac{\theta_i}{\theta_i-1} A_i M_i (1 - (\frac{K_i(p_i)}{f_i})^{\theta_i-1}) + \sum_{j\neq i} \mathbf{1}\{\frac{p_i}{\tau_{ji}} > p_j\} A_j M_j \theta_j f_j^{1-\theta_j} \times s_{ji} \int_{p_j}^{p_i/\tau_{ji}} K_j(p)^{\theta_i-2}\, dp$. Note that

$$\frac{\partial G_i(r_i, f_i)}{\partial f_i} = -\theta_i A_i M_i K_i(p_i)^{\theta_i-1} f_i^{-\theta_i} < 0$$

and

$$\begin{aligned}
&\frac{\partial G_i(r_i, f_i)}{\partial r_i} \\
&\quad = 1 - \left(\frac{\theta_i A_i M_i f_i^{1-\theta_i} K_i(p_i)^{\theta_i-2}}{\displaystyle\sum_{j\neq i} \mathbf{1}\left\{\frac{p_j}{\tau_{ij}} \geq p_i\right\} s_{ij}} \right. \\
&\qquad \left. - \sum_{j\neq i} s_{ji} \mathbf{1}\left\{\frac{p_i}{\tau_{ji}} \geq p_j\right\} A_j M_j \theta_j f_j^{1-\theta_j} \frac{1}{\tau_{ji}} K_j\left(\frac{p_i}{\tau_{ji}}\right)^{\theta_j-2} \right) \frac{\partial D_i(r_i)}{\partial r_i}.
\end{aligned}$$

The second term captures the effect of a change in the local fixed cost of search on the local price through the change in net exports. Note that we can write the equation as

$$\begin{aligned}
&\frac{\partial G_i(r_i, f_i)}{\partial r_i} \\
&\quad = 1 - \left(1 - \frac{\displaystyle\sum_{j\neq i} s_{ji} \mathbf{1}\left\{\frac{p_i}{\tau_{ji}} \geq p_j\right\}}{\displaystyle\sum_{j\neq i} s_{ij} \mathbf{1}\left\{\frac{p_j}{\tau_{ij}} \geq p_i\right\}} \left(\frac{\theta_j A_j M_j f_j^{1-\theta_j} \frac{1}{\tau_{ji}} K_j\left(\frac{p_i}{\tau_{ji}}\right)^{\theta_j-2}}{\theta_i A_i M_i f_i^{1-\theta_i} K_i(p_i)^{\theta_i-2}} \right) \right) \\
&\qquad \times \theta_i A_i M_i f_i^{1-\theta_i} K_i(p_i)^{\theta_i-2} \sum_{j\neq i} s_{ij} \mathbf{1}\left\{\frac{p_j}{\tau_{ij}} \geq p_i\right\} \frac{\partial D_i(r_i)}{\partial r_i}.
\end{aligned}$$

Note that when all countries are identical and trade is costless, the second term in the parentheses is equal to 1, in which case $\frac{\partial G_i(r_i,f_i)}{\partial r_i} = 1$. When trade is costly and countries are identical, because the numerator is multiplied by $\frac{1}{\tau_{ji}}$, the second term in the parentheses will be less than 1, which because $\frac{\partial D_i(r_i)}{\partial r_i} < 0$, implies that $\frac{\partial G_i(r_i,f_i)}{\partial r_i} > 1$. More generally, the second term is a measure of the differential effect of a local price change on imports relative to exports; as long as countries are not too asymmetric, this effect will be second order so that $\frac{\partial G_i(r_i,f_i)}{\partial r_i} > 0$. As a result, we have $\frac{\partial r_i}{\partial f_i} > 0$, as required.

APPENDIX G: ADDITIONAL STRUCTURAL ESTIMATION RESULTS

This appendix provides additional details of the structural estimation: the first subsection proves that the beliefs used in the estimation procedure result in the smallest possible estimated fixed cost of search; the second subsection presents Monte Carlo simulations of the estimation procedure; the third subsection presents estimation results under alternative assumptions regarding the price distribution beliefs.

G.1. *Price Distribution Beliefs*

In this subsection, I prove that any distribution of prices where the expected price in destination j at time t is equal to the observed price is a mean-preserving increase in spread of the distribution of prices used in the structural estimation and will, hence, raise the value of search and the estimated fixed costs of search.

Recall that a searching farmer in location i is randomly assigned a destination to search, where the probability of searching destination j is s_{ij} (and $\sum_{j\neq i} s_{ij} = 1$). Because the search probabilities are independent of the realized prices, the cumulative distribution function (c.d.f.) of the farmer's belief about the distribution of prices from which she draws, $F^i_{p/\tau}$, can be written as the product of the search probabilities and the distribution of prices in each possible location,

$$F^i_{p/\tau}(p) = \sum_j s_{ij} \times F_{jt}(p\tau_{ij}), \tag{69}$$

where $F_{jt}(\cdot)$ is the c.d.f. of the farmer's belief about the distribution of prices in location j at time t. It is then sufficient to characterize the beliefs of the farmer over the set of these distributions $\{F_{jt}\}$ so as to characterize the beliefs of the farmer about $F^i_{p/\tau}$.

Consider the set (of a set) of distributions $\mathcal{F}$ such that for all $\{F_{jt}\} \in \mathcal{F}$, $\int p\, dF_{jt}(p) = p_{jt}$ for all $j \in \{1, \dots, N\}$ and $t \in \{1, \dots, T\}$, that is, $\mathcal{F}$ contains all possible beliefs farmers could have about the distribution of prices across time periods and locations such that their belief about the average price in each location and each time period is correct. I say a set of beliefs $\{F_{jt}\}$ is *accurate* if $\{F_{jt}\} \in \mathcal{F}$.

Define $\{F^*_{jt}\}$ to be the set of distributions such that

$$F^*_{jt}(p) = \mathbf{1}\{p_{jt} \leq p\} \quad \text{for all } j \in \{1, \dots, N\} \text{ and } t \in \{1, \dots, T\},$$

where $\mathbf{1}\{\cdot\}$ is an indicator function. $\{F^*_{jt}\}$ is the set of distributions where the beliefs for each price in each time period are equal to the actual prices in each time period. Note that $\{F^*_{jt}\}$ is (trivially) accurate, since the distribution of beliefs has a unit mass point at the true prices. The set of beliefs $\{F^*_{jt}\}$ is used so as to calculate the value of search in the paper.

PROPOSITION 3: *All sets of accurate beliefs* $\{F'_{jt}(p)\} \in \mathcal{F}$ *are mean-preserving increases in spread of* $\{F^*_{jt}(p)\}$ *and will* (*weakly*) *increase the value of search*.

PROOF: I first prove that for any $\{F'_{jt}(p)\} \in \mathcal{F}$ such that $\{F'_{jt}(p)\} \neq \{F^*_{jt}(p)\}$, the resulting distribution of prices when searching is a mean-preserving increase in spread from $\{F^*_{jt}(p)\}$ (the result is trivial if $\{F'_{jt}(p)\} = \{F^*_{jt}(p)\}$). I then prove that any mean-preserving increase in spread of $F^i_{p/\tau}(p)$ results in a higher value of search $K_i(p)$.

First, I prove that any accurate beliefs different than $\{F^*_{jt}(p)\}$ yield a mean-preserving increase in spread from the beliefs generated by $\{F^*_{jt}(p)\}$. Consider any other set of accurate beliefs $\{F'_{jt}(p)\} \in \mathcal{F}$ such that $\{F'_{jt}(p)\} \neq \{F^*_{jt}(p)\}$. Since both sets of beliefs are accurate, we have for all $j \in \{1, \dots, N\}$ and $t \in \{1, \dots, T\}$,

$$\begin{aligned} (70) \qquad & \int_0^{p^{\max}} p\, dF^*_{jt}(p) = \int_0^{p^{\max}} p\, dF'_{jt}(p) \quad \iff \\ & \int_0^{p^{\max}} \big(F'_{jt}(p) - F^*_{jt}(p)\big)\, dp = 0, \end{aligned}$$

where $p^{\max}$ is the maximum of the support of F^*_{jt} and F'_{jt} (or, if the supports differ, the maximum of the two) and the second line uses integration by parts. Denote by $F^*_{p/\tau}$ and $F'_{p/\tau}$ the believed distribution of prices conditional on searching given beliefs $\{F^*_{jt}(p)\}$ and $\{F^*_{jt}(p)\}$, respectively, for a particular $t \in \{1, \dots, T\}$. Using equations (69) and (70), we then have, for any $t \in \{1, \dots, T\}$,

$$\begin{aligned} (71) \qquad & \int_0^{p^{\max}} \big(F'_{p/\tau}(p) - F^*_{p/\tau}(p)\big)\, dp \\ & = \int_0^{p^{\max}} \sum_j s_{ij}\big(F'_{jt}(p) - F^*_{jt}(p)\big)\, dp \quad \iff \\ & = \sum_j s_{ij} \int_0^{p^{\max}} \big(F'_{jt}(p) - F^*_{jt}(p)\big)\, dp \quad \iff \\ & = 0, \end{aligned}$$

that is, the mean believed distribution of prices conditional on searching is the same for both sets of beliefs. Because the two sets of beliefs differ, there exists at least one $j \in \{1, \dots, N\}$ and $t \in \{1, \dots, T\}$ such that there exists a $\tilde{p} \in [0, p^{\max}]$ such that $F'_{jt}(\tilde{p}) \neq F^*_{jt}(\tilde{p})$. Suppose $\tilde{p} < p_{jt}$. Then because $F'_{jt}(\tilde{p}) \in [0, 1]$ and $F^*_{jt}(\tilde{p}) = 0$, it must be the case that $F'_{jt}(\tilde{p}) > F^*_{jt}(\tilde{p})$. However, because both beliefs are accurate, for equation (70) to hold, there must exist a $\tilde{\tilde{p}} \geq p_{jt}$ such that $F'_{jt}(\tilde{\tilde{p}}) < F^*_{jt}(\tilde{\tilde{p}})$. Because a similar argument holds if we suppose $\tilde{p} \geq p_{jt}$, this implies that for any $F'_{jt} \neq F^*_{jt}$, it must be the case that there exists a $\tilde{p} < p_{jt}$ such that $F'_{jt}(\tilde{p}) > F^*_{jt}(\tilde{p})$ and a $\tilde{\tilde{p}} \geq p_{jt}$ such that $F'_{jt}(\tilde{\tilde{p}}) < F^*_{jt}(\tilde{\tilde{p}})$. Furthermore, because both F'_{jt} and $F^*_{jt}(\tilde{p})$ are monotonically increasing and both equal 1 at $p^{\max}$, this implies that for all $\tilde{p} < p_{jt}$, we have $F'_{jt}(\tilde{p}) \geq F^*_{jt}(\tilde{p})$, and for all $\tilde{p} \leq p_{jt}$, we have $F^*_{jt}(\tilde{p}) \geq F'_{jt}(\tilde{p})$. This, along with equation (70), implies (see, e.g., Ljungqvist and Sargent (2004, p. 142)) that for all $\tilde{p} \in [0, p^{\max}]$,

$$\int_0^{\tilde{p}} \left(F'_{jt}(p') - F^*_{jt}(p')\right) dp' \geq 0.$$

Since this holds for any $F'_{jt} \neq F^*_{jt}$, we can prove a similar property for the total believed distribution of prices conditional on searching:

$$\begin{aligned} \int_0^{\tilde{p}} \left(F'_{p/\tau}(p) - F^*_{p/\tau}(p)\right) dp &= \int_0^{\tilde{p}} \sum_j s_{ij}\left(F'_{jt}(p) - F^*_{jt}(p)\right) dp \quad \Longleftrightarrow \\ &= \sum_j s_{ij} \int_0^{\tilde{p}} \left(F'_{jt}(p) - F^*_{jt}(p)\right) dp \quad \Longleftrightarrow \\ &\geq 0. \end{aligned} \tag{72}$$

Equations (71) and (72) are the two necessary and sufficient properties so that $F'_{p/\tau}$ is a mean-preserving increase in spread over $F^*_{p/\tau}$.

Second, I prove that any mean-preserving increase in spread of $F^i_{p/\tau}(p)$ results in a (weakly) higher value of search $K_i(p)$. To see this, note that integration by parts yields an alternative representation of the value of search:

$$K_i(p) \equiv \int_p^{p_i^{\max}} (p' - p)\, dF^i_{p/\tau}(p') = \left(p_i^{\max} - p\right) - \int_p^{p_i^{\max}} F^i_{p/\tau}(p')\, dp'.$$

Suppose that F_2 is a mean-preserving increase in spread of F_1, let $K_1(p)$ refer to the value of search under F_1, and let $K_2(p)$ refer to the value of search under F_2. From the definition of a mean-preserving spread, we have

$$\int_0^{p^{\max}} \left(F_1(p) - F_2(p)\right) dp = 0$$

and, for all $p^* \in [0, p^{\max}]$,

$$\int_0^{p^*} \left(F_1(p) - F_2(p)\right) dp \leq 0.$$

Setting $p^* = p_i$ and subtracting the latter from the former yields

$$\int_0^{p^{\max}} \big(F_1(p) - F_2(p)\big)\,dp - \int_0^{p_i} \big(F_1(p) - F_2(p)\big)\,dp \geq 0 \quad \iff$$
$$\int_{p_i}^{p^{\max}} F_1(p)\,dp \geq \int_{p_i}^{p^{\max}} F_2(p)\,dp$$

so that

$$\begin{aligned} K_1(p_i) &= \big(p^{\max} - p_i\big) - \int_{p_i}^{p^{\max}} F_1(p)\,dp \\ &\leq \big(p^{\max} - p_i\big) - \int_{p_i}^{p^{\max}} F_2(p)\,dp = K_2(p_i), \end{aligned}$$

that is, a mean-preserving increase in spread increases the value of search. *Q.E.D.*

G.2. *Monte Carlo Simulations*

In this subsection, I perform a large number of Monte Carlo simulations of the estimation strategy to assess the extent of the incidental parameter problem. For each simulation, I generate a random set of (time-invariant) transportation costs, search probabilities, fixed costs of search, distribution of landholdings and (time-varying) quantities produced, and prices for 40 regions. For each month in each year, I then calculate bilateral trade flows between each region using the model. I then redo the estimation in an attempt to recover the transportation costs, search probabilities, and fixed costs of search given the *annual* simulated trade flow data and the observed prices, quantities produced, and distribution of landholdings. Given the 40 regions, this requires estimating 1,560 (39×40) bilateral transportation costs and search probabilities as well as 40 fixed costs of search. I repeat the procedure 100 times and calculate a number of statistics measuring the accuracy and precision of the estimates. I also repeat the procedure for different panel lengths ($T \in 5, 15, 30$), as it is the time-series variation that is used to identify the transportation costs, for which there exists an incidental parameters problem.

The results of the simulations are presented in Table X. Estimated transportation costs exhibit very little bias even with only five years of data, suggesting that the incidental parameter problem in the first stage is small. Having additional years of data, however, does substantially improve the correlation between the estimated parameters and the actual parameters.

The estimated fixed costs of search, however, are biased downward, while the estimated search probabilities are biased (slightly) upward. This appears to be a result of the variance of the distribution of the estimated transportation costs

TABLE X

MONTE CARLO SIMULATION RESULTS[a]

Estimate	Transportation Costs	Search Probabilities	Fixed Cost of Search
		Number of Years: $T = 5$	
Correlation	0.3333	0.4712	0.8438
Bias	−3.75E−04	0.0217	−2.3718
Mean log difference	−0.013	0.5131	−0.898
Mean square error	0.0887	0.0032	16.7199
		Number of Years: $T = 15$	
Correlation	0.5885	0.7256	0.9735
Bias	0.0146	0.024	−1.3922
Mean log difference	−3.84E−04	0.8444	−0.4052
Mean square error	0.0533	0.0014	5.8281
		Number of Years: $T = 30$	
Correlation	0.7182	0.799	0.98
Bias	0.0079	0.023	−1.3594
Mean log difference	1.79E−04	0.8716	−0.3898
Mean square error	0.0263	0.001	5.347
Using Actual Transportation Costs			
		Number of Years: $T = 15$	
Correlation	N/A	0.9206	0.9892
Bias	N/A	0.0014	0.2902
Mean log difference	N/A	0.1361	0.0532
Mean square error	N/A	1.33E−04	0.9839

[a]The table reports results averaged over 100 Monte Carlo simulations and all bilateral pairs. Each simulation includes 40 countries, necessitating the estimation of 1,560 bilateral search probabilities and bilateral transportation costs.

exceeding the variance of the true transportation costs because the transportation costs are measured with error. It turns out that this increased variance in transportation costs biases the value of search downward; intuitively, because iceberg transportation costs are bounded below by 1, the increased variance in estimated transportation costs reduced the average destination price net of transportation costs. Since the value of search was biased downward, a lower estimated fixed costs of search was required to rationalize the observed fraction of production that was exported. To verify that the estimated first stage is resulting in bias in the second stage, the last panel of Table X presents the results of the second-stage estimation using the actual transportation costs rather than those estimated in the first stage for $T = 15$. The estimated fixed cost of search and the search probabilities are almost perfectly correlated with the true values and the mean log difference is small: 0.136 for search probabilities and 0.0532 for the fixed costs of search. This suggests that the bias is a finite-sample bias: if the estimated transportation costs are consistent, then the variance of the error in the estimated transportation cost goes to zero, so the estimated distribution of destination prices net of transportation cost approaches the true distribution.

To assess the extent of this bias empirically, in Appendix G.3, I reestimate the fixed costs of search and search probabilities by supposing that the transportation costs are constant and equal to the estimated average transportation cost.

G.3. *Alternative Second-Stage Structural Estimates*

In this subsection, I present three alternative structural estimates of the search probability and the fixed costs of search.

The first two sets of results deviate from the paper in the assumption regarding producers' beliefs about the distribution of destination prices. Recall from the paper that the c.d.f. of the prices from which producers expect to draw conditional on searching is

$$F^i_{p/\tau}(p) = \sum_{j \neq i} s_{ij} F_{jt}(p\tau_{ij}),$$

where $F_{jt}(p) \equiv \mathbf{1}\{p_{jt} \leq p\}$ and p_{jt} is the realized price in destination j at time t. That is, it is as if producers knew prices everywhere and the sole uncertainty arises from the random search process. Suppose instead that these producers have additional uncertainty over the price in each destination. In particular, suppose that the (log of) p_{jt} can be written as a known function g of observables X_{jt} and an error term $\varepsilon_{jt} \sim F_\varepsilon$:

$$\ln p_{jt} = \ln g(X_{jt}) + \ln \varepsilon_{jt}.$$

Suppose that the producer knows g, X_{jt}, and F_ε, but does not observe ε_{jt}. Then we can write the believed distribution of prices in destination j at time t as

$$\tilde{F}_{jt}(p) \equiv \Pr\{p_{jt} \leq p\} = F_\varepsilon\big(\ln p - \ln g(X_{jt})\big).$$

Note that $\tilde{F}_{jt}(p)$ is a mean-preserving increase in the spread of $F_{jt}(p)$; from Proposition 3, we know that this will (weakly) increase the value of search. In what follows, I assume that $\ln \varepsilon_{jt} \sim N(0, \sigma)$ and consider two alternative specifications for $g(X_{jt})$. In the first alternative, $g(X_{jt}) = \delta_t + \delta_j$, that is, the producers know only the average (crop-specific) price in each time period (month) t and location (province) j. In the second alternative, $g(X_{jt}) = \beta p_{jt-1} + \sum_{k=1}^{3} \gamma_k p^k_{jt} + \delta_t + \delta_j$, where p^k_{jt} is the price in the market that is the kth closest to market j, that is, in addition to know time-period and location averages, producers know the price in a location last period as well as contemporary prices in nearby locations. This alternative, while not realistic, allows us to assess the importance of spatial and temporal correlation in prices. In both

TABLE XI

ALTERNATIVE STRUCTURAL ESTIMATES[a]

	Search Probability	Fixed Cost of Search (PHP/kg)	Fixed Cost of Search (PHP/ha)
		Original Estimation (in the paper)	
Mean	0.048	7.44	29,946
Std. Dev.	0.104	20.07	78,544
Median	0.013	1.35	4,173
Minimum	0.000	0.06	135
Maximum	0.871	129.52	414,337
		Uncertainty of Destination Prices: Destination/Time Fixed Effects	
Mean	0.048	8.66	31,521
Std. Dev.	0.109	25.41	87,065
Median	0.011	1.63	4,793
Minimum	0.000	0.00	3
Maximum	0.962	209.46	670,088
		Uncertainty of Destination Prices: Fixed Effects, Lagged Prices, and Nearby Prices	
Mean	0.035	9.54	38,533
Std. Dev.	0.095	26.17	101,477
Median	0.002	1.10	4,062
Minimum	0.000	0.00	0
Maximum	1.000	201.45	644,463
		Average Estimated Transportation Cost	
Mean	0.063	12.56	52,168
Std. Dev.	0.134	40.20	146,828
Median	0.013	2.59	7,672
Minimum	0.000	0.09	135
Maximum	0.999	356.43	995,816

[a]This table compares the structural estimates in the paper to the structural estimates under four alternative sets of assumptions for the second stage of the estimation procedure.

cases, I estimate the coefficients using OLS and estimate σ from the observed standard deviation of the estimated residuals.

The second and third panels of Table XI present the resulting structural estimates. Compared to the original structural estimates (reproduced in the first panel for comparison), the estimates are quite similar. The median fixed cost of search, which is 4,173 pesos in the original estimation, is 4,793 pesos with destination and time fixed effects, and is 4,062 pesos when also including lagged and nearby prices. The average search probability is 4.8% in both the original estimates and with just fixed effects, although that falls to 3.5% when including lagged and nearby prices. Hence, alternative belief structures do not yield substantially different estimates.

The third alternative specification is chosen to assess the extent of the bias in estimates of the fixed cost of search and search probabilities found in the Monte Carlo simulations. In the third specification, rather than calculating the destination price using estimated bilateral transportation costs, I set all bilateral transportation costs to the average estimated transportation cost of 1.47. This (mechanically) prevents estimation error in the transportation costs from affecting the distribution of destination prices net of transportation costs, which was causing the fixed costs of search to be biased downward. The fourth panel of Table XI presents the results. As in the Monte Carlo simulations, the estimated fixed cost of search increases: the median rises from 4,173 pesos to 7,672 pesos. Contrary to the Monte Carlo simulations, however, the average size of the search probabilities increases from 4.8% to 6.3%.

APPENDIX H: ADDITIONAL TABLES

TABLE XII

SUMMARY OF INTERVIEWS[a]

Farmers ID	Total Area (hectares)	Parcels	Crops Grown	Trader Relationship	Cell Phone Use	Knowledge of Prices
1	0.5	1	Rice	Same trader for past 20 years	Does not own cell phone	Does not know prices anywhere else
2	1.2	1	Rice, okra, eggplant	Talks with 2–3 traders before choosing who to sell to	Calls traders in nearby town	Does not know prices anywhere else
3	? (large)	10	Rice	Mills his own rice, sells it at his own store	?	Knows price in nearby towns, but not outside municipality
4	1.3	1	Rice	Sells to the same local trader every year	Does not own cell phone	Knows price in nearby towns, but not outside province

(*Continues*)

TABLE XII—*Continued*

Traders ID	Processing	Buying Price	Crops Purchased	Farmer Relationship	Cell Phone Use	Knowledge of Prices	Sales	Storage
1	No mill	Changes price when larger traders change price	Rice and corn	Different farmers each year	?	Knows prices of larger local traders but not elsewhere	Sells directly to consumers in store he/she owns	Does not store
2	Mill on site	Changes price when government changes price floor	Rice	Many of the same farmers each year	Uses cell phone to find out prices in other regions	Knows prices in nearby towns	Sells to traders from other municipalities	May store for 1–2 months if price is going up
3	Mill on site	Charges a premium over the market price	Rice and corn	?	Has a land line	Knows prices in nearby towns, but not in other regions	Sells to other traders in nearby towns	Does not store
4	Mill on site	Market prices	Rice, sugar, flour (annual: purchase 75,000 metric tons of unmilled rice)	Purchase from farmers at buying stations and other traders, will purchase from other regions in the Philippines as well	Calls other regions throughout Philippines to keep track of prices; thinks mobile phones has increased competition between mills and made prices more responsive	Knows prices throughout Philippines	Ships to other provinces or sells locally, depending on prices	Does not store
5	Mill on site	Market price, updated regularly (sometimes daily)	Rice	Purchases from local farmers and other traders, will import from elsewhere if local supply is low	Has cell phone, uses it to talk more often with traders elsewhere	Communicates with traders in other regions on daily basis	60% shipped elsewhere, 40% sold to local retailers	Has storage facilities, may use storage to wait for price to rise

[a] I conducted all interviews in January 2011 in the Camarines Sur province in the Bicol region of the Philippines.

TABLE XIII

PRODUCT DIFFERENTIATION IS NOT CAUSING TRADE PATTERNS[a]

Dep. Var.: Exported Commodity	(1) Province–Province, Annual	(2) Port–Port, 4th Quarter
Commodity homogeneity (Rauch (1999) classification)	0.209**	0.183***
	(0.090)	(0.059)
R-squared	0.031	0.033
Observations	6,800	8,260

[a] Ordinary least squares. The dependent variable is an indicator if the importing province also exports. In column 1, each observation is an importing province–commodity–year triplet; in column 2, each observation is an importing port–commodity–fourth quarter triplet. A larger value of commodity homogeneity indicates a greater degree of homogeneity. Standard errors clustered at the commodity level are reported in parentheses. Asterisks indicate statistical significance: $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

TABLE XIV

PRICE ARBITRAGE AND PRODUCT DIFFERENTIATION[a]

Dep. Var.: Change in Log Destination Price Ratio	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
Change in log origin price ratio	0.620***	0.672***	0.743***	0.749***
	(0.017)	(0.018)	(0.018)	(0.019)
Change in log origin price ratio $*$ Homogeneous commodities			−0.541***	−0.718***
			(0.040)	(0.058)
First differences	Yes	Yes	Yes	Yes
Test coefficient = 1 (p-value)	0.000	0.000	0.000	0.000
R-squared	0.438	0.938	0.492	0.651
Observations	1,724	1,724	1,724	1,724

[a] First differences. The dependent variable is the change in the log wholesale price ratio of two commodities in the destination province. Each observation is a commodity-pair–exporter–importer–year quadruplet. The change in the origin price ratio is instrumented with the mean and standard deviation of monthly rainfall within the year interacted with a commodity-pair fixed effect to allow the effect to differ across commodity pairs. The p-value of the test of whether the estimated coefficient is 1 (as is implied by complete information price arbitrage) is reported above. Homogeneous commodities is the interaction of the homogeneity of the two commodities in the pair using the Rauch (1999) classification. Standard errors are reported in parentheses. Asterisks indicate statistical significance: $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

TABLE XV

CHANGES IN FREIGHT COSTS OVER TIME[a]

	(1)	(2)	(3)	(4)
Annual time trend	0.001***	0.001**		
	(0.000)	(0.000)		
Year 1996			0.002	0.003
			(0.007)	(0.007)
Year 1997			0.010	0.011
			(0.007)	(0.008)
Year 1998			0.006	0.005
			(0.008)	(0.008)
Year 1999			0.013	0.011
			(0.008)	(0.008)
Year 2001			0.024***	0.021**
			(0.008)	(0.009)
Year 2002			0.014*	0.013
			(0.008)	(0.008)
Year 2003			0.010	0.008
			(0.008)	(0.008)
Year 2004			0.011	0.010
			(0.008)	(0.008)
Year 2005			0.019**	0.017*
			(0.008)	(0.009)
Year 2006			0.021**	0.019**
			(0.009)	(0.009)
Year 2007			0.014	0.011
			(0.009)	(0.010)
Year 2008			0.028***	0.024**
			(0.010)	(0.010)
Year 2009			0.010	0.009
			(0.011)	(0.011)
Origin FE	Yes	Yes	Yes	Yes
Destination FE	Yes	Yes	Yes	Yes
Commodity FE	Yes	Yes	Yes	Yes
Origin–destination FE	No	Yes	No	Yes
F-value year FE jointly 0			1.541	1.096
p-value year FE jointly 0			0.095	0.357
R-squared	0.030	0.236	0.034	0.239
Observations	2,686	2,686	2,686	2,686

[a] The dependent variable is the observed freight costs (in iceberg form). Each observation is an origin–destination–commodity–year quadruplet. Only observations reporting freight costs are included; freight is unavailable for the year 2000. In columns 3 and 4, the omitted year is 1995. Standard errors are reported in parentheses. Asterisks indicate statistical significance: $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$.

REFERENCES

ANTRÀS, P., AND A. COSTINOT (2011): "Intermediated Trade," *Quarterly Journal of Economics*, 126 (3), 1319–1374. [38]

BARDHAN, P., D. MOOKHERJEE, AND M. TSUMAGARI (2013): "Middlemen Margins and Globalization," *American Economic Journal: Microeconomics*, 5 (4), 81–119. [38]

BUREAU OF AGRICULTURAL STATISTICS (BAS) (2002a): "Marketing Cost Structure for Corn," Discussion Paper, Philippines Bureau of Agricultural Statistics. [12]

——— (2002b): "Marketing Cost Structure for Tomato," Discussion Paper, Philippines Bureau of Agricultural Statistics. [12]

——— (2003): "Marketing Cost Structure for Potato," Discussion Paper, Philippines Bureau of Agricultural Statistics. [12]

——— (2007a): "Marketing Cost Structure for Garlic," Discussion Paper, Philippines Bureau of Agricultural Statistics. [12]

——— (2007b): "Marketing Cost Structure for Onion," Discussion Paper, Philippines Bureau of Agricultural Statistics. [12]

——— (2011): "Selected Statistics on Agriculture," Discussion Paper, Philippines Bureau of Agricultural Statistics. [7]

DIRKS, K., J. HAY, C. STOW, AND D. HARRIS (1998): "High-Resolution Studies of Rainfall on Norfolk Island: Part II: Interpolation of Rainfall Data," *Journal of Hydrology*, 208 (3–4), 187–193. [8]

EATON, J., AND S. KORTUM (2002): "Technology, Geography, and Trade," *Econometrica*, 70 (5), 1741–1779. [33]

EATON, J., S. KORTUM, AND F. KRAMARZ (2011): "An Anatomy of International Trade: Evidence From French Firms," *Econometrica*, 79 (5), 1453–1498. [34]

EECKHOUT, J., AND P. KIRCHER (2010): "Sorting and Decentralized Price Competition," *Econometrica*, 78 (2), 539–574. [13]

LJUNGQVIST, L., AND T. J. SARGENT (2004): *Recursive Macroeconomic Theory*. Cambridge: MIT Press. [58]

MAS-COLELL, A., M. WHINSTON, AND J. GREEN (1995): *Microeconomic Theory*. New York: Oxford University Press. [4]

MCFADDEN, D. (1973): "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Economics*, ed. by P. Zarembka. New York: Academic Press. [36]

MOEN, E. R. (1997): "Competitive Search Equilibrium," *Journal of Political Economy*, 105 (2), 385–411. [13]

PETERS, M. (2000): "Limits of Exact Equilibria for Capacity Constrained Sellers With Costly Search," *Journal of Economic Theory*, 95 (2), 139–168. [13]

RAUCH, J. (1999): "Networks versus Markets in International Trade," *Journal of International Economics*, 48 (1), 7–35. [65]

Dept. of Economics, Northwestern University, 2003 Sheridan Road, Evanston, IL 60208-2600, U.S.A.; treb.allen@northwestern.edu.

Manuscript received August, 2012; final revision received July, 2014.