Information Frictions in Trade: Online Appendix

Treb Allen[*]
Yale University

November 9, 2011

Abstract

This online appendix introduces several extensions to the model presented in the paper and presents several additional tables and figures.

1 Model extensions

In this section, I outline three extensions of the model. The first subsection contains several derivations of the model extension incorporating intermediaries mentioned in Section 6 of the paper. The second subsection discusses how the introduction of fixed costs of export to the model affects the structural estimation and results. The third subsection demonstrates how the model can be extended to incorporate spatially correlated price shocks without affecting the qualitative results of the model.

1.1 Intermediaries

In this subsection, I first detail the derivations determining the optimal purchase price offered by traders and then provide the details of the procedure used to estimate the trader fixed costs of search.

1.1.1 Determining the price offered by traders to farmers

In this subsection, I show 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. Since the quantity purchased q_t as a function of price offered p_t is

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

[*]Department of Economics, Yale University, New Haven, CT 06511. Email: treb.allen@yale.edu.

that traders with greater s_t purchase a greater quantity of produce follows immediately from the fact that they offer higher buying prices to farmers, i.e. (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 $K'(p) = -\left(1 - F_{\frac{p}{\tau}}(p)\right) < 0$ and hence $K''(p) = F'_{\frac{p}{\tau}}(p) > 0$. 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 the quantity purchased, the higher the per unit revenue. As a result, (c) follows necessarily from (b).

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

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

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

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

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

$$\frac{\partial}{\partial s_t} p_t(s_t) = -\frac{\tilde{\pi}_{12}(p_t, s_t)}{\tilde{\pi}_{11}(p_t, s_t)} \tag{1}$$

Assuming that $\tilde{\pi}_{11}(p_t, s_t) < 0$ (which is necessary to yield an interior solution), equation (1) 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 this is the case requires some onerous algebra and calculus. Note that:

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

Hence:

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

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

$$\frac{\partial}{\partial s} R\left(sq\left(p_t\right)\right) = \tilde{q}\left(p_t\right) R'\left(s\tilde{q}\left(p_t\right)\right)$$
$$\frac{\partial}{\partial s} R'\left(s\tilde{q}\left(p_t\right)\right) s\tilde{q}\left(p_t\right) = \tilde{q}\left(p_t\right) R'\left(s\tilde{q}\left(p_t\right)\right) + \tilde{q}\left(p_t\right)^2 R''\left(s\tilde{q}\left(p_t\right)\right) s.$$

Finally, note that:

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

so that

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

Substituting into equation (2) yields:

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

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

1.1.2 Estimating the trader fixed cost of search

From equation (22) in the paper, the threshold s_t required to participate in the market satisfies the following condition:

$$\pi\left(s_t^*\right) = 0 \Leftrightarrow K^{-1}\left(\frac{f}{q\left(p\left(s_t^*\right); s_t^*\right)}\right) = p\left(s_t^*\right) \Leftrightarrow f = K\left(p\left(s_t^*\right)\right) q\left(p\left(s_t^*\right); s_t^*\right), \tag{3}$$

which is equivalent to equation (24) in the paper, where Q_t^* and p_t^* in the paper correspond to $q\left(p\left(s_t^*\right); s_t^*\right)$ and $p\left(s_t^*\right)$ in equation (3), respectively.

I estimate $p\left(s_t^*\right)$ to be the minimum price I observe a farmer selling to a trader in province i of commodity c in month m in year t, p_{icmt}^*. I can then construct $K_{ict}\left(p_{icmt}^*\right)$ using the definition of $K(\cdot)$ from equation (4) in the paper and the estimated search probabilities and transportation costs estimated in Section 4 of the paper.

To construct $q\left(p\left(s_t^*\right); s_t^*\right)$, I first calculate the fraction of total quantity sold to traders that is sold to traders offering the minimum price, which I refer to as α_{icmt}^*. I then scale α_{icmt}^* by the total quantity sold to traders. To do this, I rely on the fact that the theory implies all active traders will export their purchased commodity. I can then scaleα_{icmt}^* by $\frac{X_{ict}}{\sum_{m=1}^{12} q_{icmt}}$, where X_{ict} is total exports in a given year and q_{icmt} is the total quantity I observe sold to traders within a given market. I then define the estimated trader fixed cost to be:

$$\hat{f}_{icmt} \equiv \alpha_{icmt}^* \frac{X_{ict}}{\sum_{m=1}^{12} q_{icmt}} K_{ict}\left(p_{icmt}^*\right)$$

Finally, to compare the estimated trader and farmer fixed costs of search with those estimated in the baseline model (with are annual estimates rather than monthly estimates), I aggregate to the province-commodity-year level by constructing an average across monthly markets, weighting months according to the total quantity sold.

1.2 Fixed costs of export

In this subsection, I extend the model to incorporate fixed costs of export. I show how fixed costs of export affect the model, the structural estimation, and the results.

1.2.1 Incorporating fixed costs of export into the model

Assume that for a producer to export, she must pay a fixed cost g_i. This fixed cost is incurred prior 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 which 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:

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

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

$$\varphi_i^E(p_i) = \frac{f_i + g_i}{K_i(p_i)}. \tag{4}$$

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, i.e. $\varphi_i^*(p) = \frac{f_i}{K_i(p)}$. 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:

$$\frac{f_i + g_i}{K_i(p_i)} = \frac{f_i}{K_i(p_i^E)} \Leftrightarrow p_i^E = K_i^{-1}\left(\frac{f_i}{f_i + g_i} K_i(p_i)\right) \geq p_i. \tag{5}$$

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:

$$Q_{ij} > 0 \Leftrightarrow \frac{p_j}{\tau_{ij}} \geq p_i + \alpha_i, \tag{6}$$

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

$$Q_{ij} = \frac{\theta_i}{\theta_i - 1} f_i^{1-\theta_i} A_i M_i s_{ij} \sum_{l=1}^{L} \frac{K_i\left(p_{l-1}^{ij}\right)^{\theta_i - 1} - K_i\left(p_l^{ij}\right)^{\theta_i - 1}}{1 - F_{\frac{p}{\tau}}^{i}\left(p_{l-1}^{ij}\right)}, \tag{7}$$

where $p_0^{ij} \equiv p_i^E$ instead of p_i as in equation (6) in the paper.

It is informative to compare equation (6) to the corresponding equation when there is only a fixed cost of export and information is complete, i.e. when information is complete. From equation (4), 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:

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

or, equivalently:

$$Q_{ij} > 0 \Leftrightarrow \frac{p_j}{\tau_{ij}} = p_i + \alpha_i, \tag{8}$$

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

1.2.2 Incorporating fixed costs of export into the structural estimation of the model

The fixed cost can be identified by structurally estimating the wedge between the home price and destination price net of transportation costs at which exports begin to occur. The key insight that allows for identification is that the price wedge does not depend on the particular destination.[1] Following the basic model, I assume that transportation costs can be written as:[2]

$$\ln \tau_{ijct} = \ln \tau_{ijc} + \varepsilon_{ijct}, \tag{9}$$

where the idiosyncratic component $\varepsilon_{ijct} \sim N(0, \sigma^2)$. Then I can estimate both τ_{ijc} and the price wedge α_{ict} from equation (6) by maximizing the following log likelihood function function:

$$\ell\left(\{\alpha_{ict}\}, \{\tau_{ijc}\}\right) = \sum_{j \neq i}^{N} \sum_{t=1}^{T} \left[\begin{array}{c} \mathbf{1}\{Q_{ijct} = 0\} \ln\left(1 - \Phi\left(\frac{1}{\sigma} \ln\left(\frac{p_{jct}}{\tau_{ict}(p_{ict} + \alpha_{ict})}\right)\right)\right) \\ +\mathbf{1}\{Q_{ijct} > 0\} \ln \Phi\left(\frac{1}{\sigma} \ln\left(\frac{p_{jct}}{\tau(p_{ict} + \alpha_{ict})}\right)\right) \end{array} \right]. \tag{10}$$

It is informative to compare equation (10) to the corresponding equation for the complete information case. Since equation (8) differs from equation (6) only in the equality, the complete information log likelihood function differs from equation (10) only in one way: the

[1] If the wedge did depend on the destination, the wedge would still be identified, but the identification would arise from functional form alone.

[2] While year fixed effects could be theoretically included in the estimation, because of computation time they are practically infeasible. Since the year fixed effects estimated in the basic model were small, this is unlikely to substantially affect the estimation results.

normal c.d.f. function in the second term becomes a normal p.d.f. function, just as in the basic model.

Once α_{ict} and τ_{ijc} have been estimated from the extensive margin, estimation of the intensive margin proceeds as in the basic model (i.e. equation 19 in the paper). where the only difference is that $p_0^{ijct} \equiv p_{ict} + \hat{\alpha}_{ict}$ instead of p_{ict} in the empirical analog of equation (7). Furthermore, given the estimated $\hat{\alpha}_{ict}$ and trade openness, the fixed cost of export and the fixed cost of search can be separately identified. To see this, note that $\Lambda_{ict} = \left(\frac{f_{ict}+g_{ict}}{K_i(p_i)}\right)^{1-\theta_i}$ (from the analog of the trade openness equation in the basic model) and $\frac{f_i+g_i}{K_i(p_i)} = \frac{f_i}{K_i\left(p_i^E\right)}$ from equation (5), so that:

$$f_i = \frac{K_i\left(p_i + \alpha_i\right)}{\Lambda_i^{\frac{1}{\theta_i - 1}}} \text{ and } g_i = \frac{K_i(p_i) - K_i(p_i + \alpha_i)}{\Lambda_i^{\frac{1}{\theta_i - 1}}}.$$

As a result, estimation of $\hat{f}_{ict}$ and $\hat{g}_{ict}$ occurs from a simple modification of equation 20 in the paper:[3]

$$\ln \hat{f}_{ict} = \ln K_{ict}\left(p_{ict} + \hat{\alpha}_{ict}\right) - \frac{1}{\theta_i - 1} \ln \Lambda_{ict}$$

$$\ln \hat{g}_{ict} = \ln\left(K_{ict}\left(p_{ict}\right) - K_{ict}\left(p_{ict} + \hat{\alpha}_{ict}\right)\right) - \frac{1}{\theta_i - 1} \ln \Lambda_{ict}.$$

1.2.3 The results of incorporating fixed costs of export

In this section, I briefly report how the inclusion of fixed costs of export affect results of the structural estimation. Table 4 reports the summary statistics of the structural estimates. The estimated fixed costs of export are variable but moderate, with a mean of 8,562 pesos and a median of zero. The inclusion of fixed costs of export reduces the median fixed cost of search from 4,648 to 3,373 and reduces the median transportation cost from 1.33 to 1.23 – both reductions of approximately 30 percent.

The comparison of transportation costs estimated under complete information and incomplete information remains qualitatively unchanged from the paper. Figure 5 illustrates the distribution of the two estimates. The fixed costs of export shifted both estimated transportation costs downward roughly proportionally (the mean transportation costs are 1.57 and 1.28 for complete and incomplete information, respectively), so the incomplete information estimates remain roughly half as large as the complete information estimates. Figure 6 indicates that both estimates and the difference between them are increasing in shipping distance as in the basic paper. Figure 7 indicates that the incomplete information estimates remain much more realistic than the incomplete information estimates given observed freight costs; indeed, overall the estimated transportation costs are almost exactly in the middle of the expected range.

Incorporating fixed costs of export does not alter the paper's conclusion that the majority of the negative relationship between distance and trade flows is due to declines in the search

[3] In the case that $\hat{\alpha}_{ict} = 0$, it is straightforward from equation (5) that $\hat{g}_{ict} = 0$, so no log transformation is necessary.

probability rather than increases in transportation costs. Table 5 reports the results of the new decomposition. The estimated coefficients are nearly identical to the results in the paper. Finally, Figure 8 compares the welfare effects of a 50 percent reduction in the fixed costs of export to a 50 percent reduction the fixed costs of search. Because the fixed costs of export are small relative to the fixed costs of search, the change in welfare is negligible for all farmers. Hence, I conclude that incorporating fixed costs of export into the structural estimation does not affect the central conclusions of the paper in any substantial way.

1.3 Correlated Productivity Shocks

Productivity shocks due to variation in weather are likely to be spatially correlated. As a result, producers may be able to infer the market conditions in other regions from the market conditions in the regions they search, which may cause them to alter their search search strategy. In this subsection, I show that the basic trade model with information frictions generalizes to allow for the correlation of weather shocks within groups of markets (henceforth, islands), as long as the shocks are independent across different islands. Assume that farmers first search across islands, observing the mean island price, then upon choosing an island, search across regions (henceforth, markets) within the island to find a market 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_{ε}^{g}. For simplicity, assume that the within island distribution of ε is the same for all islands, i.e. $F_{\varepsilon}^{g} = F_{\varepsilon}\ \forall g$.

This two-stage search problem can be solved using backwards 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:

$$V^m\left(p_g, \varepsilon_j\right) = \max\left\{\varphi p_g \varepsilon_j, \int V\left(p_g, \varepsilon_j\right) dF_{\varepsilon} - f^m\right\}. \tag{11}$$

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

$$\varphi p_g \bar{\varepsilon}\left(\varphi\right) = \int V^m\left(p_g, \varepsilon_j\right) dF_{\varepsilon} - f^m. \tag{12}$$

Substituting equation (12) into equation (11) yields:

$$V^m\left(p_g, \varepsilon_j\right) = \varphi p_g \max\left\{\varepsilon_j, \bar{\varepsilon}\left(\varphi\right)\right\}. \tag{13}$$

Substituting equation (13) into equation (12) yields:

$$\begin{aligned} \varphi p_g \bar{\varepsilon}\left(\varphi\right) &= \varphi p_g \left[\bar{\varepsilon}\left(\varphi\right) F_{\varepsilon}\left(\bar{\varepsilon}\left(\varphi\right)\right) + \int_{\bar{\varepsilon}(\varphi)} \varepsilon dF_{\varepsilon}\left(\varepsilon\right)\right] - f^m \Leftrightarrow \\ f^m &= \varphi p_g K_{\varepsilon}\left(\bar{\varepsilon}\left(\varphi\right)\right), \end{aligned} \tag{14}$$

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

From equation (14), 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 comprise 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 (13),

$$V^m(p_g;\varphi) \equiv \int V^m(p_g,\varepsilon_j)\, dF_\varepsilon(\varepsilon_j) = \varphi p_g G(\bar{\varepsilon}(\varphi)), \tag{15}$$

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:

$$V^i(p_g;\varphi) = \max\left\{p_g \varphi G_\varepsilon(\bar{\varepsilon}(\varphi)), \int V^i(p;\varphi)\, dF_{\bar{p}}(p) - f^i\right\}. \tag{16}$$

As above, the solution to equation (16) yields a reservation correlated shock $\bar{p}_g(\varphi)$ such that:

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

so that

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

and

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

Again, the solution to equation (16) yields a reservation correlated shock $\bar{p}_g(\varphi)$ such that:

$$\varphi G_\varepsilon(\bar{\varepsilon}(\varphi))\, \bar{p}_g(\varphi) = \varphi G_\varepsilon(\bar{\varepsilon}(\varphi)) \int \max\{p, \bar{p}_g(\varphi)\}\, dF_{\bar{p}}(p) - f^i. \tag{17}$$

As above, combining equations (16) and (17) yields:

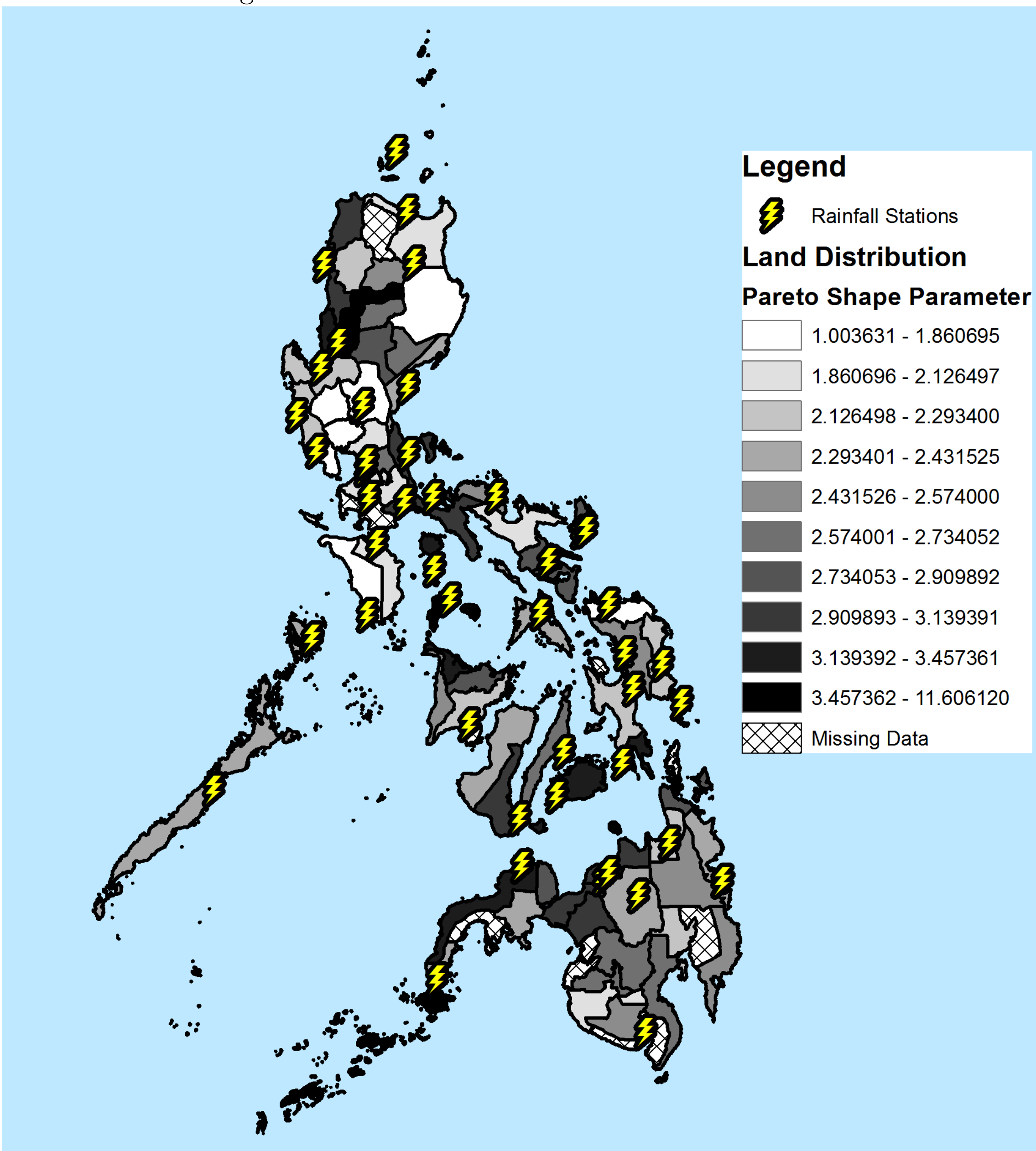
$$f^i = \varphi G_\varepsilon(\bar{\varepsilon}(\varphi))\, K_p(\bar{p}_g(\varphi)), \tag{18}$$

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

It is clear from equation (18) 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.

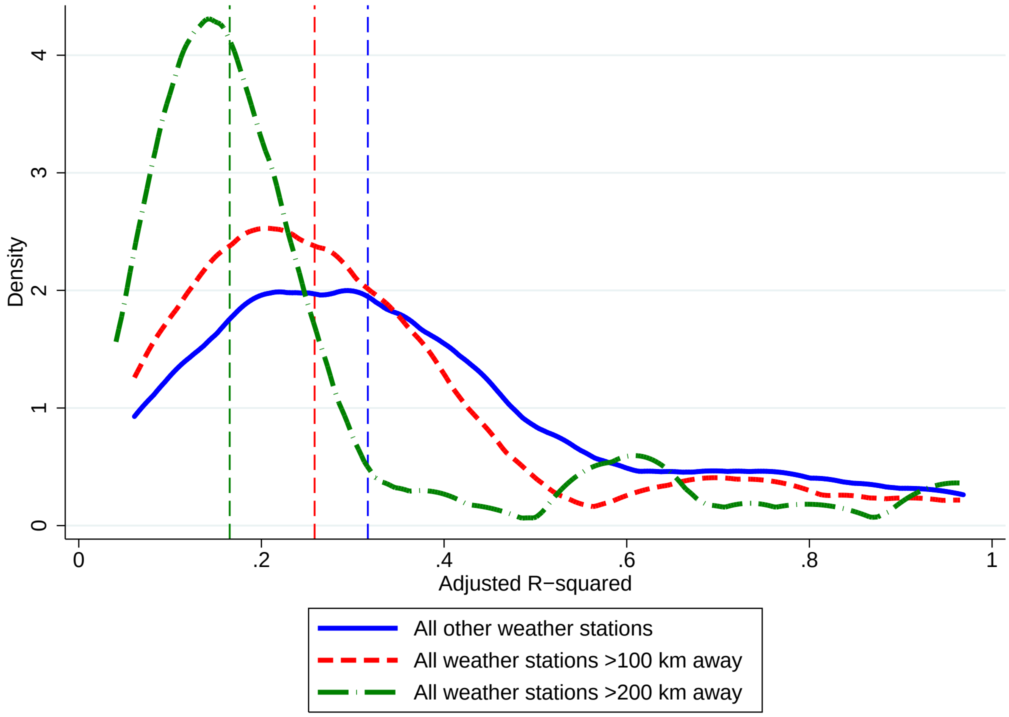
2 Tables and Figures

Figure 1: Rainfall Stations and Land Distribution



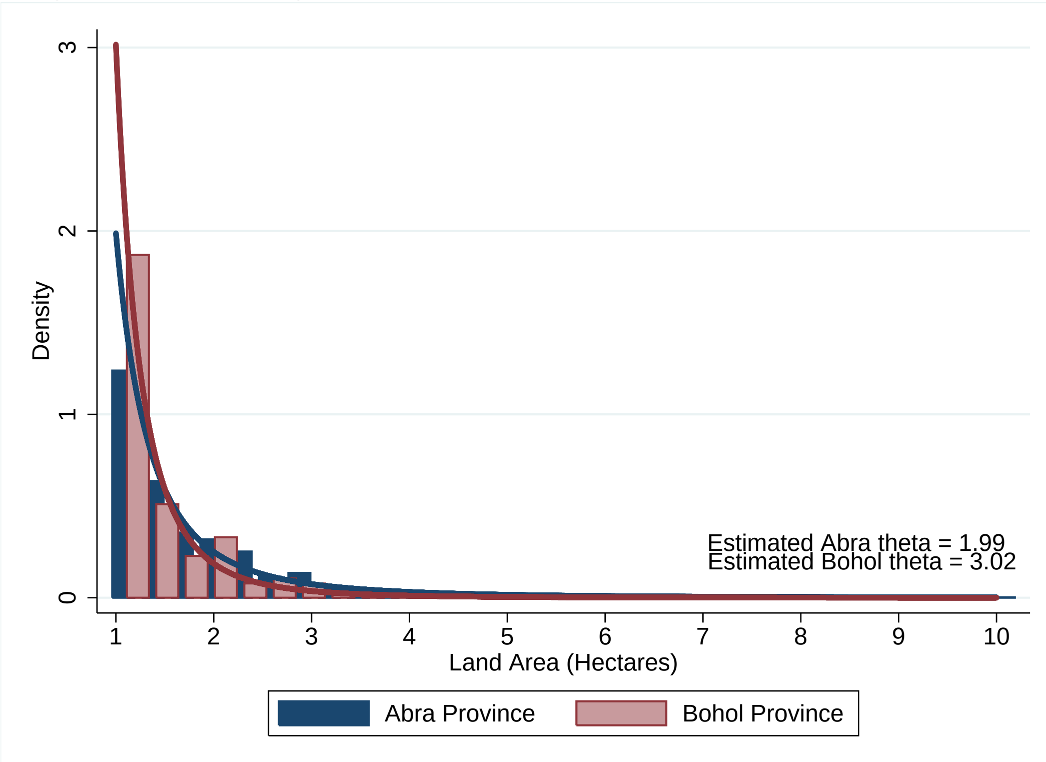
Notes: This figure shows the estimated land distribution shape parameter θ_i for each province. The shading corresponds to each decile and is darker for larger values of θ_i (indicating a greater proportion of small farmers). The figure also indicates the location of the 47 weather stations used to construct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

Figure 2: Spatial variation in rainfall in the Philippines



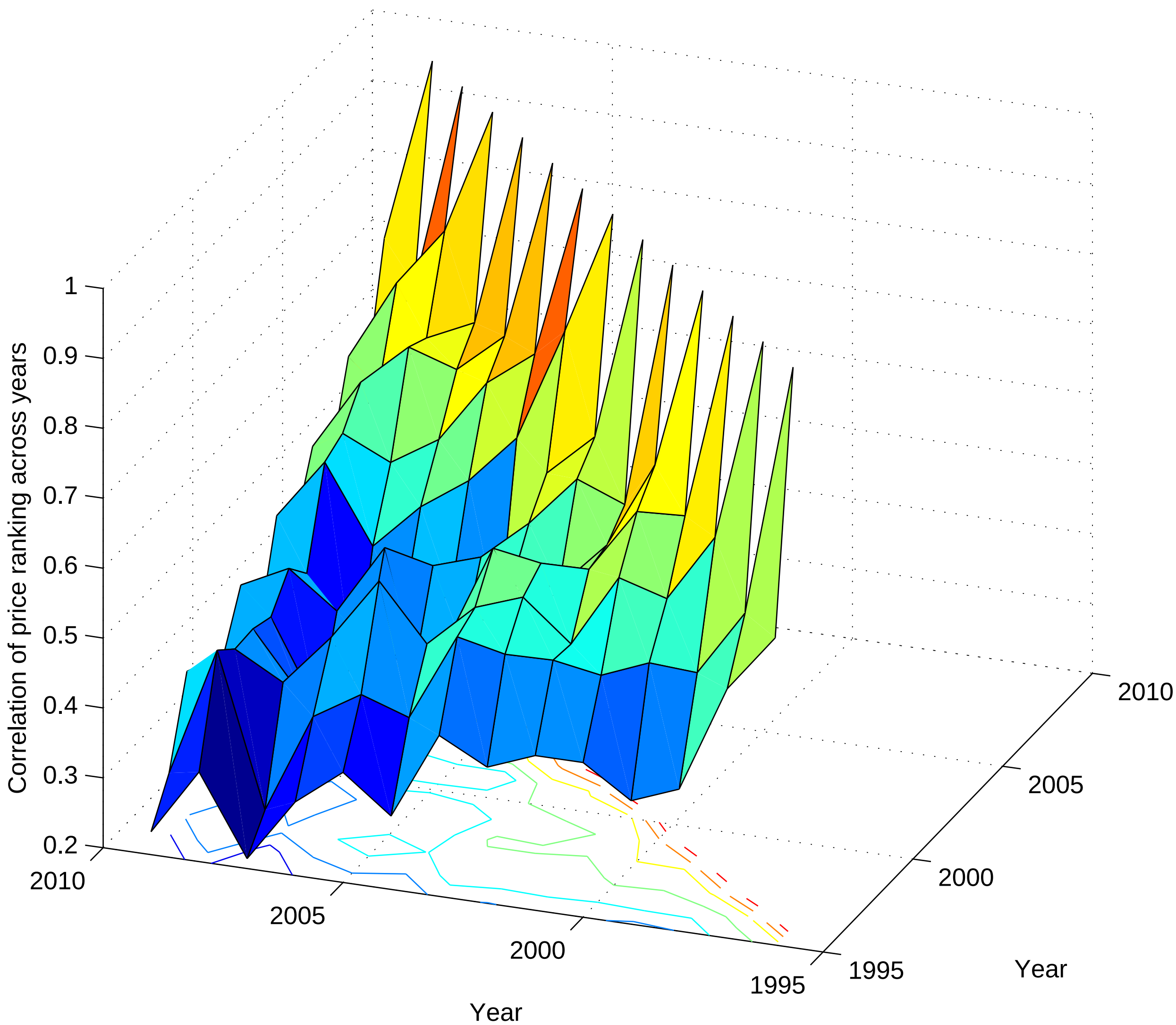
Notes: This figure shows the distribution of the predictive power of the rainfall observed at other weather stations on the observed rainfall of each weather station. Each point in the distribution is the adjusted R^2 of a regression of daily rainfall at one weather station on the daily rainfall at other weather stations for all days of the previous and subsequent week. The solid line includes all other weather stations in the regression; the dashed line includes all weather stations greater than 100 km away in the regression; and the dashed-dotted line includes all weather stations greater than 200 km away in the regression.Day fixed effects are included to control for changes in rainfall over the year.

Figure 3: Estimating the Pareto shape parameter for the land distribution



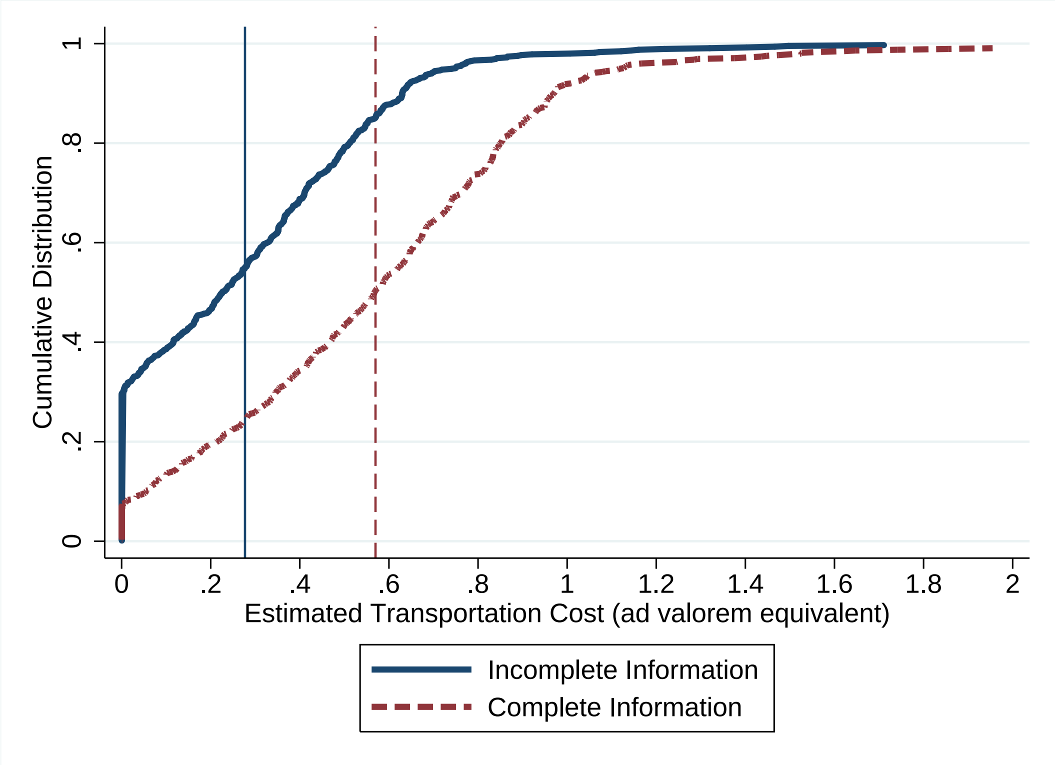
Notes: This bars shows the observed distribution of palay (rice) landholdings from the 1991 Agricultural Census for Abra and Bohol provinces, respectively. The lines show implied Pareto distribution using the maximum likelihood shape parameter.

Figure 4: Correlation of relative prices over time



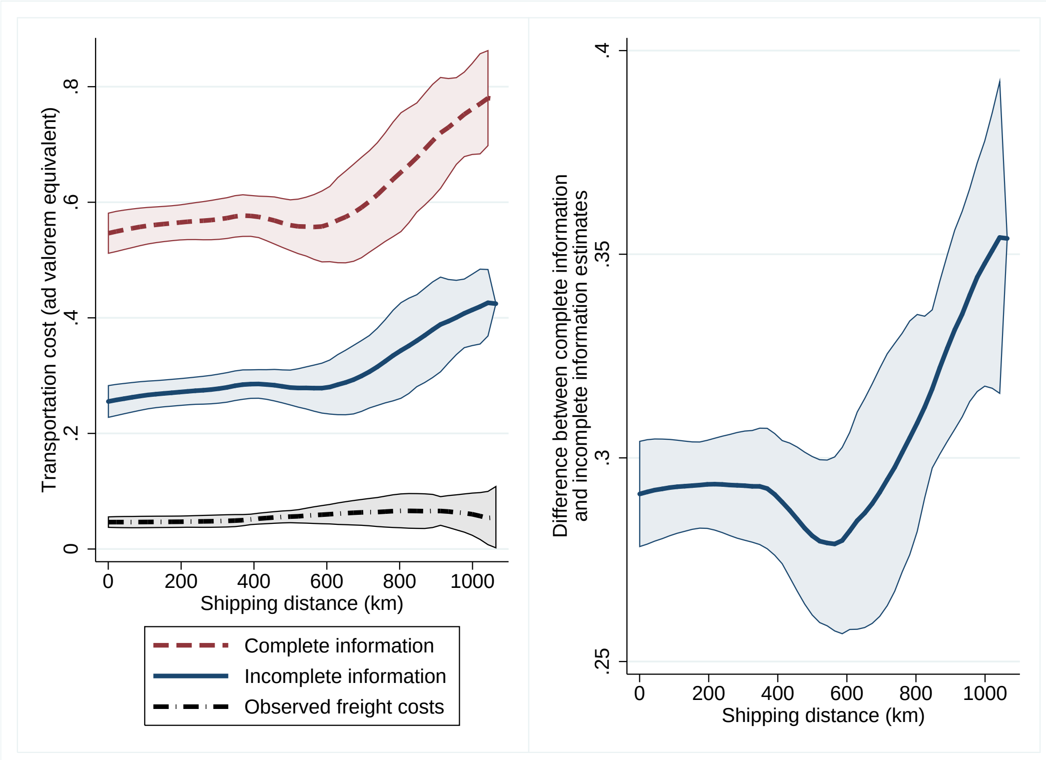
Notes: This figure shows the correlation of the relative ranking of the price within a given province across time. The relative ranking is measured as the empirical cumulative distribution function of the price of a particular commodity in a particular province (i.e. the fraction of prices in other regions below each price). The sample includes all provinces for the 10 commodities in the sample.

Figure 5: Distribution of estimated variable transportation costs (with fixed costs of export)



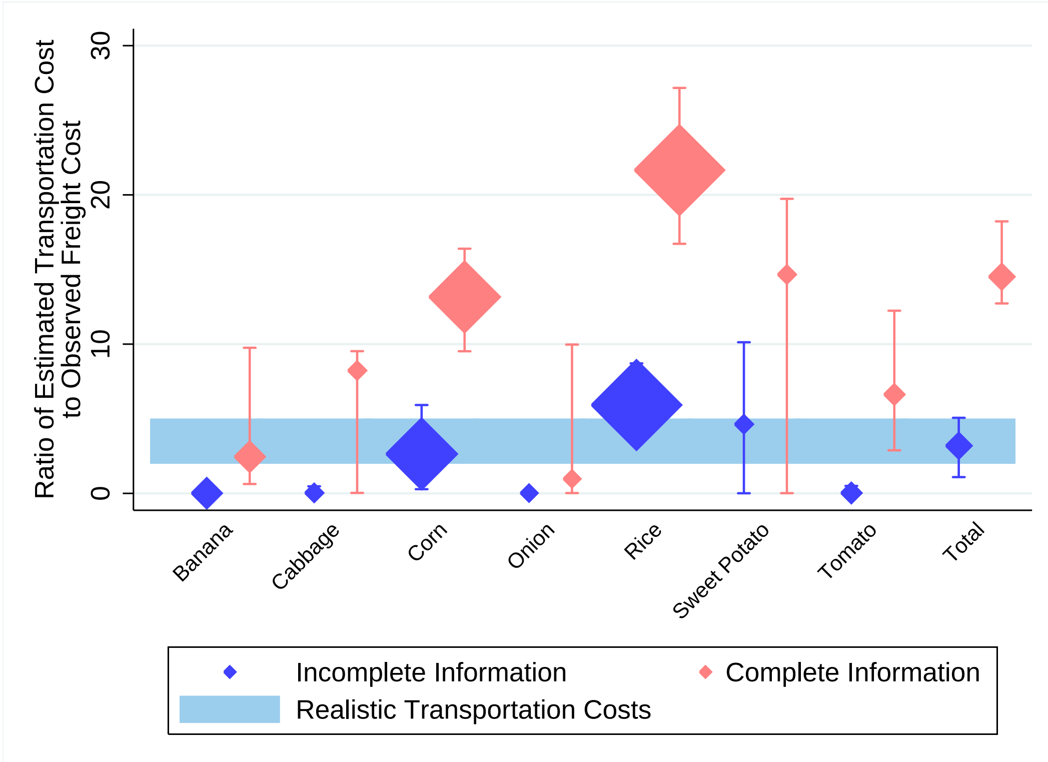
Notes: The figure depicts the cumulative distribution function of estimated variable transportation costs across origin-destination-commodities for complete information and incomplete information when the structural estimation allows for fixed costs of export. The sample includes all origin-destination-commodity triplets with wholesale markets in which trade was observed in some but not all years.

Figure 6: Estimated transportation costs and shipping distance (with fixed costs of export)



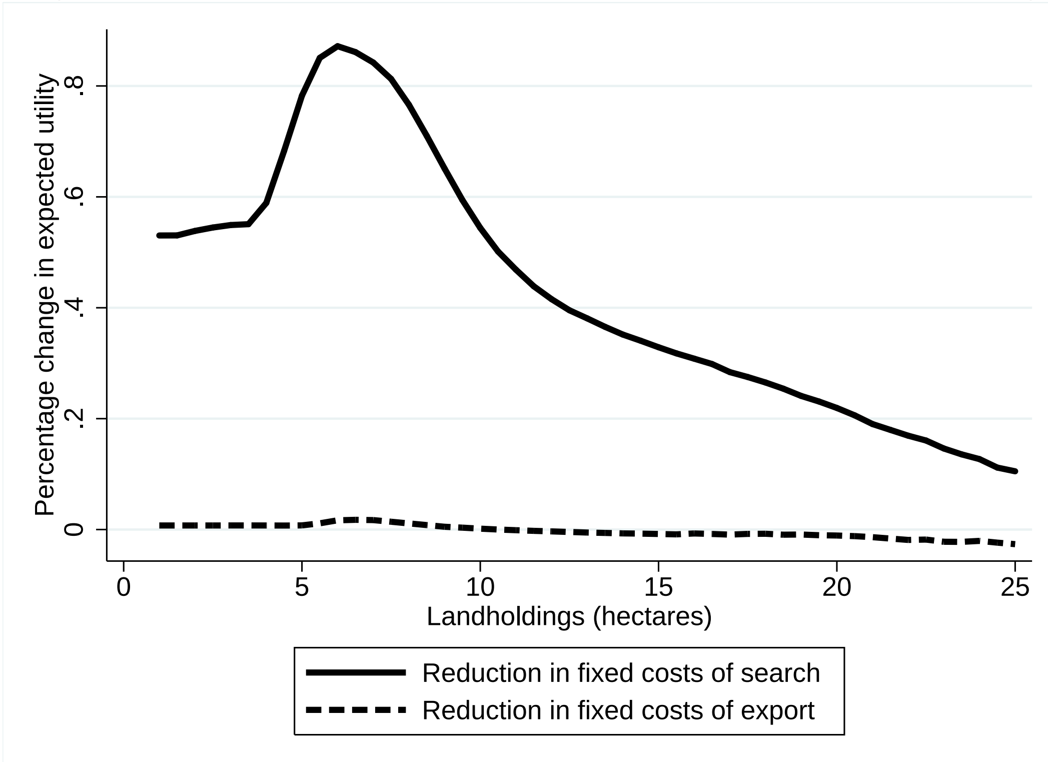
Notes: The left panel depicts the estimated transportation costs across origin-destination-commodities for complete information and incomplete information by shipping distance. The right panel depicts the difference between the complete information estimate and the incomplete information estimate by shipping distance. The structural estimation allows for fixed costs of export. Both panels use a non-parametric regression with an Epanechnikov kernel and 150km bandwidth. The shaded regions indicate the 95% confidence interval. The sample includes all origin-destination-commodity triplets with wholesale markets in which trade was observed in some but not all years. Freight costs are only observed for a subset (59%) of these origin-destination-commodity triplets.

Figure 7: Estimated transportation costs by commodity (with fixed costs of export)



Notes: Diamonds report the median ratio of the estimated transportation cost to the observed freight cost under the assumptions of incomplete and complete information when the structural estimation allows for fixed costs of export. The size of each diamond (except for the total column) is proportional to the number of estimated transportation costs. Error bars report the 95% nonparametric bootstrap confidence interval. Realistic transportation costs are defined as those between two to five times the magnitude of the observed median freight cost. The sample includes all origin-destination-commodity triplets with wholesale markets in which trade was observed in some but not all years and for which freight costs are observed. Commodities with five or fewer origin-destination pairs are not reported in the figure (garlic, mung bean, and pineapple).

Figure 8: Welfare effects of reductions in fixed costs by size of landholding



Notes: The figure reports the change in expected utility of rice farmers from a 50 percent reduction in the fixed cost of search and fixed cost of export, respectively. The welfare effects are calculated as the average across provinces and states of the world, where provinces are weighted according to their farmer population. I use the average fixed cost of search and fixed cost of export across years rather than the median since the median estimated fixed cost of export is zero in all provinces except one.

Table 1: Product differentiation is not causing trade patterns

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	6800	8260

Ordinary least squares. The dependent variable is an indicator if the importing province also exports. In the first column, each observation is an importing province-commodity-year triplet; in the second column, each observation is an importing port-commodity-4th 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. Stars indicate statistical significance: * $p<.10$ ** $p<.05$ *** $p<.01$.

Table 2: Price arbitrage and product differentiation

Dep. var.: change in log destination price ratio	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
Change in log origin	0.620***	0.672***	0.743***	0.749***
price ratio	(0.017)	(0.018)	(0.018)	(0.019)
Change in log origin			-0.541***	-0.718***
price ratio * Homogeneous commodities			(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	1724	1724	1724	1724

Notes: First differences. The dependent variable is the change in the log wholesale price ratio of two commodities in the destination province. Each observation is an 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 whether the estimated coefficient is one (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. Stars indicate statistical significance: * p<.10 ** p<.05 *** p<.01.

Table 3: Changes in Freight Costs over Time

	(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	2686	2686	2686	2686

The dependent variable is the observed freight costs (in iceberg form). Each observation is a 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 ommitted year is 1995. Standard errors are reported in parentheses. Stars indicate statistical significance: * p<.10 ** p<.05 *** p<.01.

Table 4: Summary Statistics of Structural Estimates (with fixed costs of export)

	(1) Trans. cost $\hat{\tau}_{ijc}$	(2) Search prob. $\hat{s}_{ijct}$	(3) Fixed cost of search $\hat{f}_{ict}$	(4) Fixed cost of export $\hat{g}_{ict}$
Mean	1.28	0.075	35698.7	8562.3
Std. Dev.	0.29	0.14	166089.2	71397.7
Median	1.23	0.010	3373	0
Minimum	1	0	1	0
Maximum	3.10	0.99	3242212	1584146
Coeff. of variation across commodities	0.03	0.95	0.55	1.04
Number of estimates	650	4337	992	992
Unit of identification	Origin-Destination-Commodity	Origin-Destination-Commodity-Year	Origin-Commodity-Year	Origin-Commodity-Year

Notes: Transportation costs are reported only for origin-destination-commodity triplets which traded in some but not all years. Search probabilities are identified only for observations in which trade occured. Fixed costs are reported in 2000 Philippines pesos (1 USD is approximately equal to 45 PHP). Coefficients of variation are calculated within origin-destination pairs (within origin provinces for the fixed cost of search).

Table 5: Transportation Costs versus Information Frictions in the Gravity Equation (with fixed costs of export)

Dependent variable:	(1) Log quantity	(2) Info. frictions	(3) Trans. costs	(4) Log freight
Log shipping distance	-0.437***	-0.415***	-0.022***	0.002*
	(0.037)	(0.035)	(0.003)	(0.001)
Origin-Product-Year FE	Yes	Yes	Yes	Yes
Destination-Product-Year FE	Yes	Yes	Yes	Yes
R-squared	0.058	0.058	0.017	0.002
Observations	4337	4337	4337	2686

Notes: Ordinary least squares. The dependent variable is indicated above the columns. Each observation is a origin-destination-commodity-year quadruplet. Freight costs are not reported for all observed trade flows. Standard errors are reported in parentheses. Stars indicate statistical significance: * p<.10 ** p<.05 *** p<.01.