# Information Frictions in Trade*

Treb Allen†
Yale University

*Job Market Paper*

December 7, 2011

**Abstract**

It is costly to acquire information about markets in other places, especially in developing countries. In this paper, I examine the effect of such information frictions on trade. I embed a process where heterogeneous producers sequentially search across regions to determine where to sell their produce into a perfect competition Ricardian trade model. Information frictions explain the empirical failure of price arbitrage and provide new insight into how market conditions affect trade flows. Using a data set I assemble on regional agricultural trade in the Philippines, I show that observed trade flows and prices suggest the presence of substantial information frictions. I then structurally estimate the model to disentangle information frictions from transportation costs. I find that (1) estimated transportation costs are half as large as those implied by complete information models and more consistent with observed freight costs; and (2) the vast majority (93 percent) of the "gravity" relationship between trade flows and distance can be attributed to information frictions rather than transportation costs.

**Keywords:** Information, trade costs, agricultural trade, Philippines, search, intermediaries, gravity

**JEL Classification:** D83, F10, F14, O13, Q11, Q13, Q15

*I am extremely grateful to Costas Arkolakis, David Atkin, Penny Goldberg, Mark Rosenzweig, and Chris Udry for their substantial guidance and support throughout this project. Beatriz Armendariz, Arsenio Balisacan, Bruce Blonigen, Gharad Bryan, Lorenzo Caliendo, Thomas Chaney, Marcela Di Blasi, Camilo Dominguez, Dave Donaldson, Doireann Fitzgerald, Doug Gollin, Snaebjorn Gunnsteinsson, Gordon Hanson, Rachel Heath, Tim Kehoe, Sam Kortum, Nimfa Mendoza, Melanie Morten, Giuseppe Moscarini, Adam Osman, Andrés Rodriguez-Clare, Peter Schott, Alexandre Skiba, and Yaniv Stopnitzky provided helpful comments and suggestions, as did seminar participants at the EIIT Conference, the Minneapolis Federal Reserve, the University of the Philippines, Williams College, and Yale University. I am grateful to Margaret Bautista, Sharon Faye A. Piza, Vincent Olaivar, and Ermina Tepora for their help with data collection and the Yale Economic Growth Center for providing financial support for acquiring the data. This material is based upon work supported under a National Science Foundation Graduate Research Fellowship. All errors are my own.

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

# 1 Introduction

Why does the price of a good differ across space? The standard explanation is that trade is costly.[1] Indeed, nearly all trade theories rely on a no-arbitrage condition that equates the price gap between two regions with the trade cost between those regions.[2,3] Recent empirical work by Jensen (2007), Aker (2010), and Goyal (2010), however, suggests that price variation is due not only to the cost of trading goods, but also to the cost of acquiring information about prices elsewhere ("information frictions"). Since producers must first discover the destination price before engaging in arbitrage, information frictions affect prices through the producers' search process rather than directly entering the no-arbitrage condition like traditional trade costs. Despite recognizing that information frictions exist, we currently lack a theoretical understanding of how they affect trade.[4] Understanding the role of information frictions is of particular importance to policy makers, as policies reducing information frictions differ substantially from policies reducing transportation costs.

In this paper I contribute to the understanding of how information frictions affect trade in three ways. First, I develop a trade model that explicitly incorporates the search process producers undergo to acquire information about market conditions elsewhere. The model generates new predictions of how trade flows respond to market conditions and highlights the distinct effects of information frictions and transportation costs. Second, using a data set I assemble on regional agricultural trade flows in the Philippines, I show that observed trade patterns and price dispersion cannot be explained by standard trade models but are consistent with my model incorporating information frictions. Finally, using the same data set, I structurally estimate the model to quantitatively assess the relative importance of information frictions and transportation costs. I conclude that information frictions are at least as important as transportation costs in determining trade flows.

[1]This explanation dates back to at least Heckscher (1916).

[2]By trade costs, I mean all costs associated with transporting a good from one location to another, including freight costs, the time spent in transit, policy barriers (e.g. tariffs), risk of damage, insurance costs, and local distribution costs. In what follows, I refer to these trade costs as transportation costs in order to distinguish them from information frictions.

[3]The price arbitrage condition arises from the assumption that producers sell to the destination with the greatest price net of trade costs, which is present in all seminal models of trade with perfect competition, e.g. Samuelson (1954), Anderson (1979), and Eaton and Kortum (2002). The no-arbitrage condition is also present in trade models with monopolistic competition and CES preferences, since producers price their products at a constant markup over marginal cost, e.g. Krugman (1980) and Melitz (2003). Even in models with imperfect competition where producers find it optimal to vary their mark-up across destinations (e.g. Brander and Krugman (1983), Krugman (1986), Melitz and Ottaviano (2008), Alessandria and Kaboski (2011), and Simonovska (2011)), the no-arbitrage condition requires that the price gap does not exceed the transportation cost (see Feenstra, Markusen, and Rose (2001)).

[4]Related work includes the firm learning models of Albornoz, Pardo, Corcos, and Ornelas (2010), Eaton, Eslava, Krizan, Kugler, and Tybout (2011), and Arkolakis and Papageorgiou (2011) and the network models of Rauch (1999), Rauch and Trindade (2002), Rauch and Trindade (2003), Rauch and Casella (2003), and Chaney (2011b).

To introduce information frictions, I embed a sequential search process based on the seminal job-search models of McCall (1970) and Mortensen (1970) into a many-region trade model with heterogeneous producers. In the model, all producers make the same homogeneous commodity and sell it in perfectly competitive regional markets. After production, producers can either sell locally or search for a better price elsewhere. If a producer decides to search, she pays a fixed cost and observes the price (net of transportation costs) in another market. The producer can then sell in that market or pay the fixed cost and search again. A producer finds it optimal to sell in the first market where she discovers a price greater than her reservation price. Because the fixed cost of search comprises a smaller proportion of total revenue the greater the quantity produced, larger producers have higher reservation prices. The heterogeneity in reservation prices results in a threshold producer size that determines the range of producers willing to sell to any given destination. Bilateral trade flows can be calculated by aggregating across the range of willing producers that search a particular destination.

The introduction of information frictions yields several key implications for trade flows and prices. First, because it is costly to learn about prices elsewhere, price arbitrage opportunities remain in equilibrium. As a result, information frictions can explain several empirical puzzles – including why regions import and export the same commodity and why price shocks are imperfectly transmitted to trading partners – without having to appeal to product differentiation or imperfect competition.

Second, since producers weigh the opportunity cost of searching against selling to a particular destination, market conditions elsewhere have a direct effect on trade. In particular, since larger producers search more intensively, regions inhabited primarily by large producers will export a greater share to destinations with high relative prices than regions inhabited primarily by small producers. As a result, a standard complete information model under-predicts (over-predicts) exports from more-productive (less-productive) regions to these more attractive destinations.

Third, information and transportation frictions affect bilateral trade flows differently, allowing the two to be disentangled empirically. Transportation costs alone affect whether or not a region exports any of a commodity to another region, while both transportation costs and information frictions affect the quantity exported. The intuition is straightforward: bilateral trade occurs only when the price gap between destination and origin exceeds the cost of transporting the good since individual producers only export when it is profitable to do so. In contrast, total trade flows depend on the number of producers who have discovered the arbitrage opportunity as well as the cost of transporting the good.

In the second part of the paper, I evaluate the model using a comprehensive data set on the universe of domestic shipments of agricultural commodities in the Philippines from 1995-2009. The Philippines is an ideal setting to test the model because the empirical context closely matches the model assumptions. I first provide evidence that information frictions exist by showing that the observed trade flows and prices are consistent with my model but inconsistent with a standard trade model with complete information.

I proceed by structurally estimating the model to quantify the relative importance of transportation costs and information frictions. The average estimated ad valorem transportation cost is 36 percent, which is roughly half the size of traditional estimates and much more consistent with observed freight costs and detailed marketing surveys. In addition, the vast majority (93 percent) of the observed decline of bilateral trade flows in shipping distance is a result of decreases in the probability of search rather than increases in transportation frictions. A counterfactual exercise suggests that comparable reductions in transportation costs and information frictions result in similar increases in aggregate welfare, but a reduction in information frictions has the added benefit of reducing inequality.

Finally, I extend the basic model to incorporate intermediary traders. While the predictions regarding regional trade flows remain qualitatively unchanged, this extension allows me to directly examine the search process undertaken by farmers. Using a data set of more than two million individual farmer sales, I provide micro-level evidence of information frictions. In particular, I show that reductions in the cost of search as a result of the introduction of mobile phones disproportionately increased the probability that smaller farmers searched for traders.

The paper is organized as follows. In the next section, I present the model. In Section 3, I describe the empirical context and provide reduced-form evidence that information frictions exist. In Section 4, I structurally estimate the model and quantify the relative importance of information and transportation frictions. In Section 5, I perform counterfactual simulations to assess the welfare implications of information frictions. In Section 6, I extend the model to incorporate intermediaries and test the predictions using micro data on farmer sales. Section 7 concludes.

# 2 Model

In this section I present the model. In the first subsection, I describe the setup; in the second, I describe each producer's optimal search behavior; in the third, I aggregate producer behavior to determine aggregate trade flows; in the fourth, I close the model by proving the existence and uniqueness of prices that equilibrate supply and demand; and in the fifth, I highlight two key empirical implications of incorporating information frictions.

## 2.1 Setup

There are a large number of regions in the world, each inhabited by consumers and a mass $M_i$ of producers.[5] All producers in all regions produce the same homogeneous good and maximize profits. In each region, resident consumers purchase the homogenous good in a perfectly competitive market.[6] Producers may either sell in their local market or search for another market in which to sell their produce. Since the empirical portion of the paper focuses on the agricultural sector, I refer to producers as farmers and the homogenous good as rice. In what follows, I use $i$ to refer to the origin region and $j$ to refer to the destination region.

**Demand:** Consumers in region $j$ are endowed with exogenous wealth and have a continuous utility function that is increasing and concave in rice.[7] As a result, I can write the inverse demand function that yields the price of rice $p_j$ as a function of the total rice consumed in region $r_j$, i.e. $p_j = D_j(r_j)$. It is straightforward to show that $D_j(r_j)$ is non-negative, continuous, and strictly decreasing in $r_j$.

**Production:** Regions are subject to idiosyncratic productivity shocks $A_i \in (0, 1]$, where $A_i$ indicates the fraction of farmers who produce. If a farmer produces in a period, she produces an amount equal to her productivity $\varphi$.[8] In an agricultural framework with constant returns to scale technology, $\varphi$ can be interpreted as the landholdings of the farmer.[9] Productivity $\varphi \in [1, \infty)$ is heterogeneous across farmers and distributed according to the cumulative distribution function $F^i_\varphi$, which is assumed to be a Pareto distribution with shape parameter $\theta_i > 1$, i.e. $F^i_\varphi(\varphi) = 1 - \varphi^{-\theta_i}$.[10]

**Search:** Farmers know their local price and the true distribution of prices (net of transportation costs), but must engage in a sequential search process to learn the realized prices elsewhere. The search process begins after production and works as follows. A farmer can

[5] I follow Lucas and Prescott (1974) in defining "large" as a continuum or a countable infinity of regions. The large number of regions is necessary to ensure that the distribution of prices faced by producers is invariant to the particular realization of productivity shocks.

[6] One way of rationalizing perfectly competitive regional markets is to assume that all consumers are costlessly informed of offer prices of at least two producers selling in their regional market. In this case, Burdett and Judd (1983) show that all producers selling to the region charge the same competitive price.

[7] In Section 5, I explicitly model the income process of consumers when I embed the model in a general equilibrium framework.

[8] In a more complicated framework, producers would choose the quantity to produce given their productivity in order to maximize expected profits. I pursue such a tactic when I incorporate intermediaries in Section 6. Since larger producers find better prices on average, the cost of production must be convex in quantity produced in order for the profit function to be concave in quantity produced.

[9] I assume that the landholdings of farmers are determined exogenously, which is a realistic assumption in a context like the Philippines where most land is acquired by inheritance and land markets are largely missing (e.g. Estudillo, Quisumbing, and Otsuka (2001)). To mitigate any concern of land sales responding to trade flows, in the empirical portion of the paper I use the distribution of landholdings four years prior to the beginning of the trade data.

[10] The Pareto distribution is a reasonable approximation of the the observed land distribution in the Philippines; see Appendix B.7 for details.

either sell at home or pay a fixed amount $f_i$ to search, which reveals the price (net of transportation costs) in a region $j$ and allows the farmer to sell there. The probability that a farmer from region $i$ searches region $j$ is $s_{ij} > 0$, where $\sum_{j \neq i} s_{ij} = 1$.[11] I assume that the probability of search $s_{ij}$ is the same for all farmers in a particular region and constant throughout the search process. After searching, the farmer can either sell to that region or pay the fixed cost $f_i$ to search again. The process continues until the farmer sells her produce. I model transportation costs in the standard iceberg form, where $\tau_{ij} \geq 1$ units of a good must be shipped from region $i$ in order for one unit to arrive in region $j$, where $\tau_{ii} = 1$ for all $i$.[12]

The model has two types of information frictions: the fixed cost of search $f_i$ and the search probability $s_{ij}$. The fixed cost of search $f_i$ captures all costs associated with gathering sufficient information about a market in order to sell there. Since a farmer pays the fixed cost prior to searching a particular destination, the fixed cost is destination invariant.[13] The search probability $s_{ij}$, in contrast, captures the propensity of farmers to search certain destinations more often than other destinations. Farmers may be more likely to search a particular region for a number of reasons, both economic (e.g. higher expected prices or lower shipping costs) and non-economic (e.g. common religion, family ties, etc.).

The assumption that the search process is stochastic deserves some discussion, as it is crucial for the tractability of the model. There are at least two concerns. First, a stochastic search process may appear to be suboptimal: why would a farmer not immediately search the destination with the greatest expected price net of transportation costs? The answer is straightforward: if all other farmers also search this destination first, then any initial arbitrage opportunity would be eroded by their exports. As a result, any individual farmer can profit by deviating and first searching the second-most attractive destination. Indeed, as long as there are a sufficient mass of farmers exporting, the only symmetric equilibrium search strategy is a mixed strategy.[14]

The second concern is that a farmer may discover prices during the search process that offer her additional information about prices in regions not yet searched, creating an incentive to alter her search probabilities mid-search. For example, a farmer may reduce the probability

[11] With a continuum of regions, the search probability becomes a search density $s_i(j)$ such that $\int_J s_i(j)\,dj = 1$, where $J$ is the set of destinations. For simplicity, in what follows I use the notation for a countably infinite number of regions.

[12] In the online appendix, I extend the model to incorporate fixed costs of export in addition to variable transportation costs. The insights of the model and empirical results remain qualitatively unchanged.

[13] If the realized fixed cost of search varied by destination, then a farmer would undertake the search process if and only if the gains from doing so exceeded the expected fixed cost of search, i.e. $f_i = \sum_{j \neq i} s_{ij} f_{ij}$.

[14] In particular, the mass of exporting farmers in a particular region $i$ must be sufficiently large so that the expected price net of transportation costs in the most attractive destination conditional on all exporting farmers from $i$ selling there is less than the expected price net of transportation costs in the second most attractive destination conditional on no farmers selling there. In ongoing work, I allow for the search process to be determined endogenously in order to examine the determinants of information frictions.

of searching regions nearby a region she discovers has a low price. In the online appendix, I extend the model to explicitly incorporate such spatial correlations. The extension builds a two-stage search model motivated by the particular geography of the Philippines, where farmers first search across islands (across which prices are assumed to be uncorrelated), and then search across markets within an island (where prices may be correlated). I show that this extension delivers the same qualitative predictions of the basic search model.

## 2.2 Optimal search behavior

In this subsection, I determine the optimal search behavior of a farmer given the distribution of prices. Consider a farmer with landholdings $\varphi$ from region $i$ who has positive production. Define $F^i_{\frac{p}{\tau}}(p)$ to be the cumulative distribution function of prices net of transportation costs that the farmer believes she will draw from if she chooses to search.[15] Let $p$ be the price (net of transportation costs) the farmer has discovered. Then the value function of the farmer is:

$$V_i(p;\varphi) = \max\left\{ \underbrace{\varphi p}_{\text{sell}}, \underbrace{\int_{p_i^{\min}}^{p_i^{\max}} V_i(p';\varphi)\, dF^i_{\frac{p}{\tau}}(p') - f_i}_{\text{search again}} \right\},$$

where $p_i^{\min} \equiv \min_{j\neq i}\left\{\frac{p_j}{\tau_{ij}}\right\}$ and $p_i^{\max} \equiv \max_{j\neq i}\left\{\frac{p_j}{\tau_{ij}}\right\}$ are the worst and best prices net of transportation costs that a farmer can encounter.

The farmer's optimal strategy yields a reservation price $\bar{p}_i(\varphi)$ such that the farmer will choose to sell if $p \geq \bar{p}_i(\varphi)$.[16] Because of the stochastic nature of the search process, the reservation price is invariant to the number of regions already searched, so no farmer will prefer to sell to a previously searched region. I show in Appendix A.1 that the equilibrium condition governing the reservation price is:

$$f_i = \varphi \int_{\bar{p}_i(\varphi)}^{p_i^{\max}} \left(p' - \bar{p}_i(\varphi)\right) dF^i_{\frac{p}{\tau}}(p'). \tag{1}$$

Equation (1) states that at the optimal reservation price the cost of continuing to search is equal to the marginal benefit of continuing to search. Hence, as the fixed cost of search approaches zero, the reservation price that a farmer is willing to accept increases. In the limit

[15]For what follows, $F^i_{\frac{p}{\tau}}(p)$ need not be everywhere differentiable. For ease of notation in what follows, however, I use $\int g(p)\, dF^i_{\frac{p}{\tau}}(p)$ to refer to the Lebesgue integral, i.e. for any function $g(\cdot)$, $\int g(p)\, dF^i_{\frac{p}{\tau}}(p) = \sum_{j\neq i} s_{ij} g\left(\frac{p_j}{\tau_{ij}}\right)$.

[16]When $p = \bar{p}_i(\varphi)$, the farmer is indifferent between searching and selling, so her optimal decision is indeterminate. I assume that some fraction of farmers choose to search again. Since the mass of indifferent farmers is zero, this does not affect the quantity traded.

where the fixed cost is equal to zero, the reservation price for all farmers regardless of land size is $p_i^{max}$, so that all farmers sell to the destination with the highest price net of transportation costs. As a result, it can be shown (see Appendix A.5) that the model converges to a complete information Ricardian trade model as the fixed cost of search approaches zero.

With positive search costs, equation (1) also implies that the threshold cutoff $\bar{p}$ is strictly increasing in $\varphi$; i.e. larger farmers have greater reservation prices. Intuitively, larger farmers have more to sell so their returns from discovering a better price are greater. As a result, they search more intensively than smaller farmers.[17] Since $\bar{p}$ is strictly increasing in $\varphi$, I can invert equation (1) to yield the landholdings $\varphi^*(p)$ of the farmer that is indifferent between selling and continuing to search at price $p$:

$$\varphi_i^*(p) \equiv \frac{f_i}{K_i(p)}, \tag{2}$$

where

$$K_i(p) \equiv \int_p^{p_i^{\max}} (p' - p)\, dF^i_{\frac{p}{\tau}}(p'). \tag{3}$$

Since $K_i(p)$ measures the per unit benefit of continuing to search as a function of the current price $p$, I refer to it as the "value of search." Note that $K_i(p)$ is strictly decreasing: the greater the offer in hand, the lower the value of continuing to search.[18] As a result, $\varphi_i^*(p)$ is strictly increasing: the greater the offer price $p$, the larger the landholdings of the farmer that is indifferent between continuing to search and selling. In other words, as the price in hand increases, larger farmers become willing to sell at that price.

Unlike in Melitz (2003), the threshold landholding size $\varphi_i^*(p)$ indicates the *maximum* land size such that a farmer would be willing to sell to a particular destination rather than the *minimum* productivity required to enter a market. The difference arises because in this model, farmers have a fixed amount of produce to sell, so the decision to sell to one destination comes at the cost of selling elsewhere, whereas in Melitz (2003) firms have constant marginal costs, allowing them to decide how much to produce in each market independently of all other markets. Hence, the model developed here is more realistic in settings where production cannot easily be scaled to respond to changes in market demand, e.g. agriculture.

[17] It can be shown (see appendix A.2) that the expected per-unit revenue that a farmer receives from searching is her reservation price $\bar{p}(\varphi)$. Since the reservation price is strictly increasing in quantity produced, the expected per-unit revenue is also increasing in quantity produced.

[18] In particular, $\frac{\partial}{\partial p} K_i(p) = -\left(1 - F^i_{\frac{p}{\tau}}(p)\right)$.

## 2.3 Aggregate trade

Given the optimal search behavior of each farmer, it is possible to characterize total bilateral trade flows by aggregating across all farmers within a region.

### 2.3.1 The presence of trade

Consider first whether or not any trade from $i$ to $j$ occurs. Only farmers with reservation price $\bar{p}(\varphi) \geq p_i$ will decide to search rather than sell to their own region. As a result, if $\frac{p_j}{\tau_{ij}} < p_i$, then no farmer searching region $j$ will choose to sell there, since all farmers who choose to search are unsatisfied with prices lower than their local price. Conversely, if $\frac{p_j}{\tau_{ij}} \geq p_i$, then any farmer with reservation price $\bar{p}(\varphi) \in \left[p_i, \frac{p_j}{\tau_{ij}}\right]$ (or, equivalently, landholdings $\varphi \in [\varphi_i^*(p_i), \varphi_i^*\left(\frac{p_j}{\tau_{ij}}\right)]$) who searches region $j$ will choose to sell there. Since the distribution of landholdings is continuous, such a farmer will search region $j$ with a probability of one. As a result:

$$Q_{ij} > 0 \Leftrightarrow \frac{p_j}{\tau_{ij}} \geq p_i, \tag{4}$$

where $Q_{ij}$ denotes the quantity exported from $i$ to $j$. Equation (4) implies that positive exports occur if and only if there exists a price arbitrage opportunity.

### 2.3.2 The quantity of trade

Now consider the quantity of rice exported from $i$ to $j$. Suppose that $\frac{p_j}{\tau_{ij}} \geq p_i$ so that some trade occurs. By summing across the infinite number of possible search paths and integrating over the distribution of landholdings, it can be shown (see section A.3 of the appendix) that:

$$Q_{ij} = \frac{\theta_i}{\theta_i - 1} A_i M_i f_i^{1-\theta_i} s_{ij} \sum_{l=1}^{L} \frac{K_i\left(p_{l-1}^{ij}\right)^{\theta_i - 1} - K_i\left(p_l^{ij}\right)^{\theta_i - 1}}{1 - F_{\frac{p}{\tau}}^i\left(p_{l-1}^{ij}\right)}, \tag{5}$$

where $p_0^{ij} \equiv p_i$, $p_L^{ij} \equiv \frac{p_j}{\tau_{ij}}$, and $p_{l-1}^{ij} \leq p_l^{ij}\ \forall l \in \{1, ..., L\}$ are the set of prices net of transportation costs to all other regions that are between $p_i$ and $\frac{p_j}{\tau_{ij}}$.

Equation (5) shows that bilateral exports are increasing in the total quantity produced ($\frac{\theta_i}{\theta_i - 1} A_i M_i$), decreasing in the fixed cost of search $f_i$, and increasing in the probability of search $s_{ij}$. The summation term captures how prices affect bilateral trade flows (henceforth the "predicted trade share").[19] Since $\frac{\partial}{\partial p} K_i(p) = -\left(1 - F_{\frac{p}{\tau}}^i(p)\right)$, as the difference in ordered prices

[19] It is straightforward to show that the magnitude of the elasticity of bilateral trade flows to transportation costs is decreasing as the difference between the destination price net of transportation costs and the origin price increases. This is consistent with the finding of Novy (2010) that bilateral trade flows respond less to changes in transportation costs the greater the size of trade flows.

$p_l^{ij} - p_{l-1}^{ij}$ converges to zero (i.e. as $F^i_{\frac{p}{\tau}}(p)$ approaches a differentiable function), the predicted trade share converges to a more intuitive expression:

$$\lim_{\max_l\left(p_l^{ij} - p_{l-1}^{ij}\right)\to 0} \sum_{l=1}^{L} \frac{K_i\left(p_{l-1}^{ij}\right)^{\theta_i - 1} - K_i\left(p_l^{ij}\right)^{\theta_i - 1}}{1 - F^i_{\frac{p}{\tau}}\left(p_{l-1}^{ij}\right)} = (\theta_i - 1)\int_{p_i}^{\frac{p_j}{\tau_{ij}}} K_i(p)^{\theta_i - 2}\, dp \tag{6}$$

Hence, trade flows are greater the larger the difference between the origin price and the destination price. The difference in how prices affect trade flows, however, depends on prices in other regions through the value of search. Regions with many small farmers (i.e. high $\theta_i$), will concentrate exports to destinations with low relative prices (i.e. high $K_i\left(\frac{p_j}{\tau_{ij}}\right)$), whereas regions with many large farmers will concentrate exports to destinations with high relative prices. This occurs because smaller producers search less intensively than larger producers.

It will prove helpful later to decompose bilateral trade flows into the total amount traded and the share which is exported to each destination. Define $\lambda_{ij} \equiv \frac{Q_{ij}}{\sum_{j\neq i} Q_{ij}}$ as the share of exports from region $i$ destined for region $j$ and $\Lambda_i \equiv \frac{\sum_{j\neq i} Q_{ij}}{\sum_j Q_{ij}}$ to be the fraction of production that is exported. From equations (2) and (5), it can be shown that:

$$\lambda_{ij} = s_{ij} K(p_i)^{1-\theta_i} \sum_{l=1}^{L} \frac{K_i\left(p_{l-1}^{ij}\right)^{\theta_i - 1} - K_i\left(p_l^{ij}\right)^{\theta_i - 1}}{1 - F^i_{\frac{p}{\tau}}\left(p_{l-1}^{ij}\right)} \tag{7}$$

$$\Lambda_i = \left(\frac{K_i(p_i)}{f_i}\right)^{\theta_i - 1}. \tag{8}$$

The share of exports sent to a destination depends both on the probability that it is searched and its predicted trade share (normalized by the value of search at the home price). In contrast, the fraction of production exported (the "openness" of a region) depends only on the fixed cost of search, the value of search at the local price, and its distribution of landholdings.

## 2.4 Equilibrium

Thus far, I have taken prices in each region as given. In this subsection, I show that there exists a set of equilibrium prices that are consistent with both the optimal search behavior of farmers and consumer demand. Equilibrium prices$\{p_i\}$ are characterized by the following three properties:

1. Given prices and beliefs, supply is governed by the optimal search of farmers, i.e. trade flows are determined by equations (4) and (5).

2. Farmers have rational beliefs concerning the distribution of prices, i.e. $F^i_{\frac{p}{\tau}}(p) = \sum_{j\neq i} s_{ij}\mathbf{1}\left\{\frac{p_j}{\tau_{ij}} \leq p\right\}$.

3. The supply of rice to each region is consistent with the demand for rice at that price, i.e. $p_j = D_j \left( \sum_i Q_{ij} \right)$.

Let a group of regions be *connected* if the graph of the undirected trade network of those regions is connected. The following proposition guarantees that equilibrium prices and trade flows exist.

**Proposition 1** *For any set of consumer preferences that are increasing and concave in rice, there exists a set of prices and trade flows that satisfy equilibrium conditions 1 and 2. Furthermore, for any equilibrium trade network, prices within any connected group of regions are unique up to a scaling factor.*

**Proof.** See Appendix A.4. ■

Intuitively, differences in domestic production and demand result in differences in autarkic prices across regions. Farmers engage in price arbitrage by exporting to destinations with high prices, causing prices to converge. The set of equilibrium prices yields just enough arbitrage opportunities to incite trade flows that ensure there is no excess supply or demand in any region.

## 2.5 The empirical implications of incorporating information frictions

How does my model compare to a standard perfect competition trade model where producers simply sell to the destination with the highest price? There are two key differences that result in empirically testable implications. First, arbitrage opportunities remain in equilibrium. In a complete information model, the no-arbitrage condition requires that if region $i$ exports to region $j$, then the price ratio is equal to the transportation cost, i.e. $\frac{p_j}{p_i} = \tau_{ij}$. With information frictions, however, trade flows imply that the price ratio is at least as great as the transportation cost, i.e. $\frac{p_j}{p_i} \geq \tau_{ij}$.

This difference in price arbitrage has two observable implications. First, in a standard trade model, the same region never both imports and exports the same commodity.[20] With information frictions, however, region $j$ will import from region $i$ and export to region $k$ whenever $p_i \leq \frac{p_j}{\tau_{ij}}$ and $p_j \leq \frac{p_k}{\tau_{ik}}$. Second, in a standard trade model, as long as trade continues, the no-arbitrage condition implies that changes to the price in the origin region will transmit one-to-one to changes in the destination region. With information frictions, however, shocks to the origin price may have negligible effects on the destination price if the quantity exported is small relative to the total amount supplied to the destination, e.g. if the search probability is low.

[20]This is a direct implication of the triangle inequality. The proof is by contradiction: suppose region $j$ imports from region $i$ and exports to region $k$. The triangle inequality requires that $\tau_{ik} < \tau_{ij}\tau_{jk}$. From the price arbitrage equation, $p_i = \frac{p_j}{\tau_{ij}}$ and $p_j = \frac{p_k}{\tau_{jk}}$, which yields $p_i = \frac{p_k}{\tau_{ij}\tau_{jk}}$. By the triangle inequality, this implies that $p_i < \frac{p_k}{\tau_{ik}}$, which violates price arbitrage, as producers selling domestically in region $i$ could profit by instead exporting to region $k$.

The second key difference between the model developed above and a standard perfect competition trade model arises from the relationship between trade flows and prices. In particular, the Donaldson (2008) extension of the Eaton and Kortum (2002) (D-EK) yields an explicit functional form for bilateral trade flows:[21]

$$\ln Q_{ij} = \delta_i + \ln X_j + \alpha \ln \frac{p_j}{\tau_{ij}}, \tag{9}$$

where $\delta_i$ is an origin fixed-effect, $X_j$ is the total expenditure on the commodity in the destination region, and $p_j$ is the CES price aggregator across all varieties of the commodity. Equation (9) implies that the partial elasticity of bilateral trade flows to the destination price net of transportation costs is constant.

With information frictions, however, market conditions in other potential destinations have a direct impact on bilateral trade flows.[22] To see this, consider how the price in region $k$ affects trade from region $i$ to region $j$. A higher price in region $k$ increases the value of search of producers in region $i$, affecting trade flows from region $i$ to region $j$ in two ways: first, it induces smaller farmers to export; second, it increases the reservation price of those farmers already searching, making them more picky about which region they sell to. From equation (5), the first effect dominates in regions with many small farmers ($\theta_i > 2$), increasing exports to region $j$; in contrast, the second effect dominates in regions with many large farmers ($\theta_i < 2$), reducing exports to region $j$.

The presence of information frictions thus suggests that the D-EK equation governing bilateral trade is misspecified. In particular, equation (9) under-predicts bilateral trade flows from regions with many small farmers (and over-predicts bilateral trade flows from regions with many large farmers) to destinations with relatively low prices. In contrast, the biases are reversed for destinations with relatively high prices. As a result, the estimated partial elasticity of bilateral trade flows to prices ($\alpha$) will be biased downwards for regions with more small farmers (higher $\theta_i$) and biased upwards for regions with more large farmers (lower $\theta_i$).[23]

[21] With complete information, trade flows cannot be written explicitly as a function of prices without resorting to some form of product differentiation (see Anderson and van Wincoop (2004)). Accordingly, the D-EK model assumes that each observed commodity (e.g. rice) is in fact an amalgamation of a continuum of different varieties.

[22] Anderson and van Wincoop (2003) make a similar point by showing that a correctly specified gravity equation must control for the origin and destination "multilateral resistance." Because equation (9) includes an origin fixed effect and the destination price (which is a direct measure of the multilateral resistance in the destination), it is correctly specified in their framework. My model implies that the gravity equation is misspecified even if an origin and destination fixed effect are included.

[23] The bias arises because of the negative correlation between the destination price and the value of search $K_i(p)$ for $p \in \left[p_i, \frac{p_j}{\tau_{ij}}\right]$. This negative correlation occurs because $K_i(p)$ is decreasing in relative price of $p$ and identification of $\alpha$ comes from the log deviation of $p_j$ from the average destination price due to the inclusion of the origin fixed effect.

# 3 Evidence of information frictions

In this section, I test the empirical implications of information frictions using a data set I have assembled on regional agricultural trade in the Philippines. I first describe the empirical context and briefly describe the data set. I then provide reduced form evidence that information frictions exist by showing that the observed prices and trade flows are inconsistent with a standard complete information trade model.

## 3.1 Empirical context

The Philippines is an ideal setting to examine information frictions, as the context closely matches three key assumptions of the model. First, production is subject to idiosyncratic productivity shocks. The productivity shocks create idiosyncratic variation across regions in the quantity produced, resulting in idiosyncratic variation in prices. Since farmers are unable to predict prices elsewhere, they must engage in a search process. In the Philippines, weather plays an important role in determining agricultural productivity, so that variation in weather results in productivity shocks.[24] Furthermore, given the island geography of the Philippines, weather shocks vary substantially across regions, so that local weather yields only limited information about weather elsewhere.[25] As a result, the spatial variation in prices differs substantially over time, so that in any given year farmers have little information about prices elsewhere.[26]

The second key assumption of the model is that production decisions occur prior to decisions about where to sell the produce, so that farmers take their quantity produced as given when searching for a market. In agricultural production in the Philippines, the most important decisions impacting the quantity produced occur at or near the time of planting, which happens months or even years before sales occur. As a result, farmers have little ability to adjust their production in response to price shocks that occur after planting.

The final key assumption of the model is that it is costly to learn about market conditions elsewhere. There is strong circumstantial evidence suggesting that this is the case in the Philippines. In qualitative interviews, rice farmers and small traders reported that they had little

[24] This paper is not the first to note the importance of rainfall shocks in the Philippines to agriculture; for example, Yang and Choi (2007) uses rainfall shocks as an instrument to see how remittances respond to income shocks.

[25] In the online appendix, I provide evidence that the observed rainfall at any given weather station is only weakly predicted by rainfall at all other weather stations.

[26] The relative price of a commodity in a particular province varies substantially over time. The correlation between the previous year and the current year price rankings (measured as the empirical cumulative distribution function of the price) varies from 0.64 for 1996-1997 to 0.83 from 2002-2003. The correlation also declines sharply over time: the correlation between the relative price in 1995 and 2010 is 0.2. See the online appendix for a complete correlation of price rankings across time. A fixed effects regression of the price ranking on year demeaned monthly mean and standard deviations of rainfall indicates that 58 percent of the variation of the ranking within a province-commodity over time can be explained by variation in rainfall.

knowledge of prices in nearby cities, let alone in other provinces. Larger traders were better aware of market conditions elsewhere, but exerted a substantial effort to keep their information up to date. Of the traders that were aware of prices elsewhere, all said that they learned the prices by directly contacting sellers in other markets.

## 3.2 Data

Data on the flow of goods within a country is rare, especially in the developing world. Because of its island geography, I have been able to assemble data detailing the universe of commodities shipped by water throughout the Philippines.[27] I summarize the data I have collected here; see Appendix B for a detailed description. The data provide me the quantity and value of every commodity (at the SITC 5-digit level) shipped each year from every port to every other port in the Philippines from 1995 to 2009. The core data set used in the subsequent sections consists of 4,337 observations of annual bilateral trade flows between provinces of 10 major agricultural commodities. The data set is summarized in Table 1. Figure 1 illustrates the network of rice trade flows; as is evident, each province trades with many other provinces.

I combine the trade data with a number of other data sources, giving me information on (1) province-year-commodity wholesale prices; (2) province-year-commodity production and yields; and (3) a census of farms producing each commodity in each province. The price data allow me to observe the market price that producers would receive (excluding transportation costs) should they choose to sell to a particular destination.[28] The production data allows me to determine the fraction of produce exported for each commodity. Finally, the agricultural census allows me to estimate the shape parameter of the land distribution for each commodity in each province (see Appendix B.7 for details). In addition, I collect data on daily rainfall at 47 rainfall stations spread throughout the Philippines which provide a measure of exogenous productivity shocks.

## 3.3 Arbitrage opportunities exist

In Section 2.5, I showed that the complete information no-arbitrage equation yields two empirical predictions: first, the same region should never import and export the same commodity; and second, changes to the price in an origin region should result in equivalent changes to the price in the destination region. In this subsection, I show that neither of these predictions are empirically validated, which suggests that arbitrage opportunities exist.

[27]Similar data for trade via air is unavailable. As trade via air constitutes less than 1% of total trade flows in terms of both quantity and value (NSO, 2001), its exclusion from the following analysis is unlikely to substantially affect the results.

[28]As a result, I do not have to rely on the unit prices reported in the trade data, which are unobserved when trade flows do not occur.

In the Philippines, it is commonplace that a province both imports and exports the same commodity. The top panel of Figure 2 indicates that on average half of importing provinces also export the same commodity in the same year. The results do not appear to be driven by the aggregation of trade data: the bottom panel of Figure 2 indicates that 26 percent of importing ports export the same commodity within a three month period.[29]

The observed variation in prices across provinces is also inconsistent with the complete information no-arbitrage condition. To see this, suppose that transportation costs can be written as:[30]

$$\ln \tau_{ijct} = \ln \tau_{ij} + \delta_t + \delta_c + \varepsilon_{ijct}. \tag{10}$$

If region $i$ exports both commodities $c$ (corn) and $d$ (rice) to region $j$ in periods $t$ and $t-1$, then the complete information no-arbitrage condition implies that changes in the log price ratio of $c$ and $d$ in the origin should co-move one-to-one with changes in the log price ratio in the destination region:

$$\Delta \ln \left( \frac{p_{jct}}{p_{jdt}} \right) = \beta \Delta \ln \left( \frac{p_{ict}}{p_{idt}} \right) + \upsilon_{ijt}, \tag{11}$$

where $\upsilon_{ijt} \equiv \Delta \frac{\varepsilon_{ijct}}{\varepsilon_{ijdt}}$ is the change in the ratio of the log of the idiosyncratic component of the transportation costs of corn and rice, $\Delta$ denotes the difference between $t$ and $t-1$, and the complete information no-arbitrage condition implies $\beta = 1$.[31] To see if equation (11) holds, I examine how the destination price ratio responds to exogenous changes in the origin price ratio due to local weather shocks. In particular, I use a two stage least squares (2SLS) estimation strategy to estimate $\beta$ by instrumenting $\Delta \ln \left( \frac{p_{ict}}{p_{idt}} \right)$ with a vector of changes in province-level rainfall, de-meaned by year to control for Philippines-wide weather shocks.[32]

Table 2 presents the estimation results of equation (11). Columns 1 and 2 present the OLS and 2SLS results, respectively, for the relative price of corn to rice in the Philippines. The estimated elasticities (0.38 for OLS and 0.27 for 2SLS) are precisely estimated and substantially below one, allowing me to strongly reject ($p < 0.001$) that the complete information no-arbitrage

[29] In the online appendix, I argue that the patterns do not appear to be driven by product differentiation either, as the patterns are actually more prevalent amongst more homogeneous agricultural commodities using the Rauch (1999) classification.

[30] This assumption, while restrictive, is much less restrictive than most assumed functional forms of trade costs in the literature. For example, Waugh (2010) assumes that trade costs are a function of distance, a dummy variable for a shared border, and an exporter fixed effect.

[31] Equation (11) resembles the "pass-through" test of how the relative price of a commodity in two countries responds to shocks in the exchange rate between those countries (e.g. Goldberg and Verboven (2001)). In that literature, arbitrage may fail because of firms using market power to vary their mark-ups across destinations, which is unlikely in the context of agricultural commodity prices with many producers. In the online appendix, I extend equation (11) to include many commodities and allow the estimated coefficient to vary by the degree of homogeneity of the commodity pair using the Rauch (1999) classification. I find that the price arbitrage equation actually does worse for more homogeneous goods (i.e. $\beta$ is lower), suggesting that the results are not due to imperfect competition.

[32] This procedure has the added benefit of correcting for any bias due to (classical) measurement error.

equation holds in the context of the Philippines.

How well does this test fare in other contexts where information frictions are likely to be smaller? In columns 3 and 4 of Table 2, I apply the same estimation procedure for the relative price of corn and hay amongst U.S. states (see Appendix B.9 for details). While the estimated elasticities in the U.S. are much greater than in the Philippines (0.56 for OLS and 0.831 for 2SLS), I am still able to reject that the elasticity is equal to one at the $p$ =0.001 level. This suggests that information frictions remain a concern even in the United States, albeit a smaller one than in the Philippines.

## 3.4 Trade flows depend on market conditions elsewhere

Recall from Section 2.5 that the D-EK model implies that bilateral trade flows have a constant partial elasticity to the destination price. With information frictions, in contrast, total trade flows depend not only on the destination price, but also on market conditions elsewhere. As a result, the estimated partial elasticity of the destination price will be biased upward for regions with many small farmers and downward for regions with many large farmers.

To see if this bias exists empirically, I include the interaction of the origin land distribution shape parameter $\theta_{ic}$ with the destination log price in equation (9):

$$\ln Q_{ijct} = \beta_1 \ln p_{jct} + \beta_2 \ln \tau_{ijct} + \beta_3 \ln X_{jt} + \gamma\theta_{ic} \times \ln p_{jct} + \delta_{ict} + \varepsilon_{ijct}. \tag{12}$$

Because regions with a greater proportion of small farmers have a greater $\theta_{ic}$, the existence of information frictions implies that $\gamma < 0$.[33]

The estimation of equation (12) is difficult for two reasons. First, because transportation frictions are not observed, I cannot condition directly on $\tau_{ijct}$. With the assumed functional form of transportation costs from equation (10), however, I can consistently estimate $\gamma$ by including an origin-destination fixed effect $\delta_{ij}$, which restricts identification to variation in prices and trade flows within a bilateral pair. Second, because trade flows and prices are jointly determined, ordinary least squares (OLS) estimation is subject to simultaneity bias. In order to circumvent this problem, I use a two stage least squares (2SLS) estimation technique to isolate changes in the destination price due solely to local weather shocks. From the model, weather shocks in the destination ($A_j$) affect destination prices by changing total production, but only affect bilateral trade flows through the change in prices, as required by the exclusion restriction.[34]

[33] In the D-EK model, the elasticity of trade flows to the destination price is larger the greater the homogeneity of the productivities of firms, as changes in the destination price induce a greater density of firms to begin exporting. A similar effect is also present in the monopolistic competition model of Chaney (2008). In contrast, my model implies the elasticity to the destination price declines as the distribution of farmers becomes more homogeneous, as smaller farmers who search less intensively comprise a greater share of total production.

[34] In the D-EK model, trade shares depend only on prices, transportation frictions, and expenditures. In my

In particular, I instrument for the destination log price with a set of rainfall measures in the destination province, de-meaned by year and province to capture only idiosyncratic variation.[35]

The results are presented in Table 3. To correct for the fact that the land distribution parameter $\theta_{ic}$ is itself an estimate from the agricultural census (see Appendix B.7 for details), I calculate the standard errors using a bootstrap procedure, recalculating the $\theta_{ic}$ for each bootstrap repetition. The first two columns present the estimates of the basic trade equation (9) using OLS and 2SLS, respectively. As predicted by D-EK, bilateral trade flows are increasing in both destination price and destination expenditure, although neither variable is statistically significant. Consistent with simultaneity bias, the estimated coefficient of the destination price from 2SLS is larger than the OLS coefficient.

Columns 3 and 4 of Table 3 present the results of regression (12). As predicted by my model, the interaction term $\gamma$ is negative and statistically significant in both cases. Furthermore, the estimated magnitude of $\gamma$ is substantial: the 2SLS estimate indicates that a one standard deviation (1.92) increase in $\theta_{ic}$ from the mean (3.16) more than halves the elasticity of bilateral trade flows to prices from 1.35 to 0.57. The last two columns include a destination-year fixed effect instead of destination expenditures to circumvent any simultaneity concerns of including the destination expenditures on the right hand side; the estimated coefficients on the destination price and the interaction term remain statistically significant and the magnitudes change only slightly.

Since the response of prices and trade flows to productivity shocks is inconsistent with standard complete information models but consistent with a model incorporating information frictions, the reduced form evidence suggests that information frictions exist. In the next section, I structurally estimate the model presented in Section 2 to assess how important information frictions are.

# 4 Information frictions versus transportation costs

In this section, I separately identify information and transportation frictions using observed market prices and bilateral trade flows. I then use these estimates to quantify the contribution of information frictions to observed price dispersion and the "gravity" relationship between trade flows and shipping distance.

model, trade shares depend only on price, transportation frictions, and the search probability. In both cases, the exclusion restriction is satisfied as long as rainfall shocks do not affect trade frictions. While severe weather shocks may temporarily affect trade (e.g. by damaging ports), it is unlikely that they would have a substantial effect on total annual trade flows.

[35]The rainfall measures include the mean and standard deviation of rainfall within each month in the province. I allow idiosyncratic rainfall shocks to have different effects on the prices of each commodity by interacting the rainfall shock with a commodity fixed effect.

## 4.1 Disentangling information frictions and transportation costs

Recall that there are three frictions in the model: (1) iceberg transportation costs $\tau_{ij}$, (2) search probabilities $s_{ij}$, and (3) the fixed cost of search $f_i$. I consider how each one affects trade flows in turn.

Transportation costs $\tau_{ij}$ are the only friction to affect whether or not trade occurs. From equation (4), there are positive trade flows from $i$ to $j$ if and only if the relative price $\frac{p_j}{p_i}$ exceeds the transportation cost $\tau_{ij}$. Hence, it is possible to identify the transportation cost $\tau_{ij}$ by determining the threshold relative price $\frac{p_j}{p_i}$ at which trade begins to occur. Since this requires observing whether or not trade occurs for multiple observations of relative prices, identification arises from variation over time.

Given transportation costs $\tau_{ij}$, search probabilities $s_{ij}$ can be identified from observed export shares. From equation (7), the share of exports from region $i$ sent to region $j$ depends only on the predicted trade shares and the search probability $s_{ij}$. Hence, given observed prices and transportation costs identified from the extensive margin, equation (7) can be implicitly solved for $s_{ij}$. Intuitively, if two destinations have the same price (net of transportation costs), but a region exports twice as much to the first as to the second, then the first destination must be searched twice as often.

Finally, the fixed cost of search $f_i$ is identified from the openness of a region. Given observed prices, $\tau_{ij}$ and $s_{ij}$, $K_i(p_i)$ can be calculated using equation (3), so that $f_i$ can be identified using the observed trade openness and equation (8). Intuitively, if two regions are otherwise identical but the first exports a greater fraction of its production, then it must have a lower fixed cost of search.

The reason that information frictions and transportation costs can be disentangled by comparing the existence of trade flows and the quantity traded is intuitive. If trade is observed, it must have been profitable, so that the difference in prices must have exceeded the costs of transportation. At the same time, the number of traders that take advantage of a particular arbitrage opportunity depends both on how attractive the opportunity is and the number of traders aware of the opportunity, so the intensive margin yields insight into the extent of information frictions.

## 4.2 Estimation

In this subsection, I describe the estimation procedure.

### 4.2.1 Transportation frictions

The transportation cost from region $i$ to region $j$ can be identified by observing the price ratio in the two regions and whether or not exports occurred. I assume that transportation costs can be written in a generalized form of equation (10):

$$\ln \tau_{ijct} = \ln \tau_{ijc} + \delta_t + \varepsilon_{ijct}, \tag{13}$$

where the idiosyncratic component $\varepsilon_{ijct} \sim N(0, \sigma^2)$.[36] This assumption allows me to use a maximum likelihood (ML) estimation procedure to identify $\ln \tau_{ijc}$ and $\delta_t$ based on variation in observed trade patterns and prices within an origin-destination-commodity triplet over time. From equation (4), the ML estimator of the log transportation costs $\ln \hat{\tau}_{ijc}$ is:[37]

$$\ln \hat{\tau}_{ijc} = \arg \max_{\ln \tau \in [1,\infty)} \sum_{t=1}^{T} \left[ \begin{array}{c} \mathbf{1}\left\{Q_{ijct} = 0\right\} \ln \left(1 - \Phi\left(\frac{1}{\sigma} \ln \left(\frac{p_{jct}}{e^{\delta_t} \tau p_{ict}}\right)\right)\right) \\ +\mathbf{1}\left\{Q_{ijct} > 0\right\} \ln \Phi\left(\frac{1}{\sigma} \ln \left(\frac{p_{jct}}{e^{\delta_t} \tau p_{ict}}\right)\right) \end{array} \right], \tag{14}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

It is informative to compare this estimation strategy to the standard procedure of inferring transportation costs from price dispersion. With information frictions, positive trade flows imply that the ratio of the destination to origin price *exceeds* the transportation cost; in contrast, with complete information, positive trade implies the ratio *equals* the transportation cost. As a result, the equivalent equation for estimating the transportation costs under complete information simply replaces the standard normal cumulative density function in the second term with the standard normal probability density function. Hence, traditional inference overestimates the true transportation costs in the presence of information frictions.

### 4.2.2 Information frictions

Given estimates of transportation costs, search probabilities $\{s_{ijct}\}$ and the fixed costs of search $\{f_{ict}\}$ can be identified from equation (7) governing export shares and equation (8) governing trade openness, respectively. Since I do not observe the idiosyncratic component of transportation costs $\varepsilon_{ijct}$, I cannot calculate the true value of $K_{ict}(\cdot)$ or $F^{ict}_{\frac{p}{\tau}}(\cdot)$ in order to directly estimate the search probabilities and fixed cots of search. Instead, I use a method of simulated moments procedure (see McFadden (1989)) to find the set of (log) information frictions $\ln \hat{s}_{ijct}$ and $\ln \hat{f}_{ict}$ that satisfy the sample mean of the (log) of equations (7) and (8) over a large number $V$ of

[36] The year fixed effect controls for changes in transportation costs across time that affect all bilateral pairs equally. In the online appendix I provide evidence that there is little systematic variation in the observed freight costs over time. Reassuringly, the estimated year fixed effects are small in magnitude and exhibit no pattern.

[37] If region $i$ exports commodity $c$ to region $j$ in every period, then equation (14) will yield $\hat{\tau}_{ijc} = 1$; conversely, if region $i$ never exported commodity $c$ to region $j$ in any time period, then equation (14) will yield $\hat{\tau}_{ijc} = \infty$. Hence, for equation (14) to yield a $\hat{\tau}_{ijc} \in (1, \infty)$ requires that there are both periods in which region $i$ exports commodity $c$ to region $j$ and periods in which it does not.

draws[38] of the unobserved component of transportation costs $\varepsilon_{ijct}$:

$$\ln \lambda_{ijct} = \ln \hat{s}_{ijct} + \frac{1}{V} \sum_{v=1}^{V} \left[ \begin{array}{c} (\theta_i - 1) \ln K_{ict}^{v} (p_{ict}) \\ - \ln \sum_{l=1}^{L} \frac{K_{ict}^{v} \left( p_{l-1}^{ijct,v} \right)^{\theta_i - 1} - K_{ict}^{s} \left( p_{l}^{ijct,v} \right)^{\theta_i - 1}}{1 - F_{\frac{p}{\tau}}^{ict,v} \left( p_{l-1}^{ijct,v} \right)} \end{array} \right] \quad (15)$$

$$\ln \hat{f}_{ict} = \frac{1}{\theta_i - 1} \ln \Lambda_{ict} - \frac{1}{V} \sum_{v=1}^{V} \ln K_{ict}^{v} (p_{ict}), \quad (16)$$

where $\varepsilon_{ijct}^{\nu}$ indicates the $v^{th}$ simulated draw of the unobserved component of $\tau_{ijct}$, $\ln \tau_{ijct}^{\nu} \equiv \ln \hat{\tau}_{ijc} + \hat{\delta}_t + \varepsilon_{ijct}^{\nu}$ is the resulting $v^{th}$ simulated transportation cost, $K_{ict}^{v}(p) \equiv \sum_{j \neq i} s_{ijct} \max \left\{ \frac{p_{jct}}{\tau_{ijct}^{v}} - p, 0 \right\}$ and $F_{\frac{p}{\tau}}^{ict,v}(p) \equiv \sum_{j \neq i} s_{ijct} \mathbf{1} \left\{ \frac{p_{jct}}{\tau_{ijct}^{\nu}} \leq p \right\}$ are the resulting value of search and price distributions from the $v^{th}$ simulated draw, and $p_l^{ijct,v}$ is the $l^{th}$ relative price observed between $p_{ict}$ and $\frac{p_{jct}}{\tau_{ijct}^{v}}$ given the simulated transportation costs $\tau_{ijct}^{v}$ and the observed prices.

There are several things to note about equations (15) and (16). First, since $K_{ict}^{v}(p)$ not only depends on $s_{ijct}$ but on the search probabilities of all other destinations with prices (net of transportation costs) between $p_{ict}$ and $\frac{p_{jct}}{\tau_{ijct}^{v}}$, equation (15) must be solved simultaneously for all destinations of a particular origin-commodity-year. Second, since the identification of $s_{ijct}$ relies on the intensive margin of trade flows, $s_{ijct}$ can only be identified when trade flows occur.[39] Third, the model requires that $i$ exports to $j$ if and only if $p_{ict} \leq \frac{p_{jct}}{\tau_{ijct}}$. As a result, the random draws used in the simulation must ensure that this equation is satisfied. In particular, if trade is (is not) observed, then $\varepsilon_{ijct}^{r}$ is drawn from a normal distribution truncated above (below) by $\ln \frac{p_{jct}/\hat{\tau}_{ijct}}{p_{ict}}$.

### 4.2.3 Results

Table 4 presents the summary statistics of each of the estimated parameters. Estimated transportation costs are substantial, averaging 36 percent (in ad valorem terms). Estimated search probabilities are small, with a mean of 9.2 percent. Since the sum of the search probabilities across all searched destinations is equal to one, this implies that producers search 10.9 destinations on average.[40] Since producers are more likely to have contacts in the major shipping centers of Manila and Cebu, it is reassuring to note that the search probabilities of these

[38] In the results that follow, $V = 100$.

[39] This does not affect the estimation, as $K_{ict}^{\nu}(\cdot)$ does not depend on the search probabilities of destinations where $\frac{p_{jct}}{\tau_{ijct}} < p_{ict}$ and $F_{\frac{p}{\tau}}^{ict,v}(\cdot)$ only depends on these search probabilities inasmuch as they affect the probability of discovering a price below the home market price, which can be calculated using the estimated search probabilities and the constraint that the entire search probabilities sum to 1.

[40] To see this, note that $\sum_{j=1}^{J} s_{ij} = 1 \Leftrightarrow J = \frac{1}{\bar{s}_{ij}}$, where $\bar{s}_{ij}$ is the mean search probability. Since the estimated mean search probability is a consistent estimator of the average $\bar{s}_{ij}$ , by the Slutsky theorem, its inverse is a consistent estimator of the average number of destinations searched.

provinces are especially large, averaging 13.3 percent and 16.7 percent, respectively.

While the estimated fixed cost of search is highly variable across different origin-commodity-years, the median fixed cost of search of 4,647 pesos ($103) is modest. While this amount is certainly more than the cost of a phone call, it seems realistic for the entire cost of determining the market conditions in a potential destination, which includes the time and expense incurred from discussions and negotiations with possibly multiple wholesale purchasers and shipping companies. However, the fixed costs are substantial relative to farmer income (crop income averaged 12,150 pesos in the 2000), preventing a substantial portion of farmers from selling in other markets. I return to the issue of fixed costs in Section 6 when I introduce intermediary traders into the model.

It is also informative to see how much estimates of frictions vary across commodities. I find that the estimates of information frictions are similar across commodities, with an average coefficient of variation of 0.91 within an origin-destination-year for the search probability and 0.53 within an origin-year for the fixed cost of search. This variation is much smaller than the variation in the estimated transportation costs across commodities within origin-destination pair (1.38), suggesting that producers of different commodities differ less in the information frictions they face than in the transportation costs they incur.

## 4.3 Estimated transportation costs

In Section 4.2.2, I showed that transportation costs estimated using the complete information model would be larger than those estimated from the incomplete information model. Figure 3 depicts the distribution of the estimated transportation costs under complete and incomplete information.[41] By relaxing the assumption of complete information, the distribution of estimated transportation costs shifts to the left, causing the average estimated transportation cost to nearly halve from 69 percent to 36 percent (in ad valorem terms). Hence, roughly half the observed price dispersion normally ascribed to transportation costs is actually due to information frictions.

The left panel of Figure 4 shows that both the complete and incomplete information estimates increase only moderately with shipping distance. The right panel of Figure 4 examines how the contribution of information frictions to price dispersion (i.e. the difference between complete information estimates and estimated with information frictions) changes with shipping distance. The positive slope suggests that information frictions increase with shipping distance.

Which estimates are more realistic? To answer this question, I compare the estimated transportation costs to observed freight costs. Detailed analysis of the market structure of several

[41] For comparability, I constrain the estimate of $\hat{\sigma}$ to be the same in both the complete and incomplete information cases. The maximum likelihood estimate of $\sigma$ is 0.39.

commodities in the Philippines (BAS 2002a,b, 2003, 2007a,b) find that direct transportation costs comprise roughly one fifth to one half of the total cost of bringing a good to market, suggesting that realistic transportation costs should be between 2 and 5 times the size of observed freight costs.

Figure 5 reports the median ratio of estimated transportation costs to observed freight costs for each commodity. For all but one commodity, the confidence interval of the ratio is within the range of 2-5 for the incomplete information estimates, suggesting that they are of a realistic magnitude. For many commodities (especially for rice and corn, which comprise the majority of the sample), however, the complete information estimated transportation costs are much higher than five times the observed freight costs. Overall, the median ratio between estimated incomplete information transportation costs and observed freight costs is 5.8 versus 16.7 for estimated complete information transportation costs. This suggests that the incomplete information estimated transportation costs are much more realistic than those estimated assuming complete information.

## 4.4 Gravity

In this subsection, I compare the role of information and transportation frictions in the gravity equation. A large variety of trade models (see Anderson (2011) and Arkolakis, Costinot, and Rodríguez-Clare (forthcoming)) yield the following empirical formulation of the "gravity" equation:

$$\ln Q_{ij} = \beta \ln T_{ij} + \delta_i + \delta_j + \varepsilon_{ij}$$

where $\delta_i$ and $\delta_j$ are origin and destination fixed effects, respectively, and $\beta$ is the partial-elasticity of bilateral trade flows to variable bilateral trade frictions $T_{ij}$.[42] Traditionally, bilateral trade frictions are assumed to be functions of observable bilateral variables, especially shipping distance (Anderson and van Wincoop, 2004). For example, if $T_{ij} = (dist_{ij})^{\gamma}$, where $dist_{ij}$ is the shipping distance from $i$ to $j$, then the gravity equation can be written as:

$$\ln Q_{ij} = \alpha \ln dist_{ij} + \delta_i + \delta_j + \varepsilon_{ij}, \quad (17)$$

where $\alpha \equiv \beta\gamma$. Column 1 of Table 5 reports the result of equation (17) for agricultural trade in the Philippines; an increase in distance of 10 percent is associated with a 4.4 percent decline in trade flows.[43] I refer to this negative relationship between bilateral trade flows and shipping

[42] While trade flows are traditionally measured in value, here I write them in terms of quantities to be consistent with the model. Since trade is measured in free-on-board (fob) prices in the data set, the inclusion of the origin fixed effect ensures that the results are equivalent.

[43] To calculate the shipping distance between province $i$ and province $j$, I use GIS software to calculate the minimum distance over water between any two provinces using a least cost distance algorithm, where travel over water is assumed to have a uniform cost and travel over land is assumed to be infinitely costly.

distance as "gravity."

What portion of gravity is due to information frictions rather than transportation costs? From equation (5), bilateral trade flows can be decomposed into the search probability, predicted trade share, and an origin fixed effect:

$$\ln Q_{ij} = \ln s_{ij} + \ln \sum_{l=1}^{L} \frac{K_i \left(p_{l-1}^{ij}\right)^{\theta_i - 1} - K_i \left(p_l^{ij}\right)^{\theta_i - 1}}{1 - F_{\frac{p}{\tau}}^{i} \left(p_{l-1}^{ij}\right)} + \delta_i \tag{18}$$

Hence, an increase in shipping distance could have one of two effects on bilateral trade flows: it could reduce the probability that a destination is searched or it could increase the transportation cost, reducing the predicted trade share. By regressing the log search probability and the predicted trade share separately on shipping distance, it is possible to determine what proportion of gravity is due to each of the terms.[44]

Columns 2 and 3 of Table 5 report the results of the regression of the log search probability and the predicted trade share, respectively, on the log shipping distance, conditional on origin-commodity-year and destination-commodity-year fixed effects. An increase in shipping distance of 10 percent is associated with a 4.1 percent decline in the search probability and a 0.3 percent decline in the predicted trade share. Hence, the vast majority (93 percent) of the observed gravity relationship is due to a decline in the probability of search rather than an increase in the cost of transport.

While this may seem surprising, column 4 of Table 5 shows that a 10 percent increase in shipping distance increases observed freight costs by only 0.1 percent. If transportation costs increase proportionally with observed freight costs, this implies that the elasticity of trade flows to transportation costs would have to be huge to generate the observed gravity relationship between trade flows and distance. This suggests that information frictions rather than transportation costs are the primary reason for the decline of trade flows with distance.[45]

Since information frictions are responsible for half the observed variation in prices and the vast majority of the "gravity" relationship between trade flows and distance, I conclude that information frictions are at least as important as transportation costs in determining trade flows.

[44] Hillberry and Hummels (2008) use the same methodology to decompose the effect of shipping distance on the number of shipments and the average value of each shipment.

[45] Since the search probability is likely to be positively correlated with the number of contacts producers have in a particular destination, this empirical result is consistent with the theoretical prediction of Chaney (2011a) that the aggregate distribution of exporter's networks of contacts in other regions declines in distance.

# 5 Implications of information frictions: counterfactual experiments

What implications do information frictions have for public policy in the Philippines? In this section, I simulate the model to calculate the welfare effects of several potential government policies. I first compare the effects of reducing information frictions to the effects of reducing transportation costs. I then examine the effect of a land reform program that redistributes land to smaller farmers.

The simulation procedure works as follows. I first draw a vector of productivity shocks for each province from the distribution implied by the observed variation in yields within province over time. Given the geography of the Philippines,[46] I find the set of bilateral trade flows and prices that satisfy the equilibrium conditions presented in Section 2.4 and calculate the equilibrium welfare for each farmers and consumer. I then repeat the procedure for a large number of productivity shocks to find each person's expected utility.[47]

To determine prices and calculate welfare, I embed the trade model in a general equilibrium framework. Because of its economic importance and the fact that it comprises the largest portion of observed trade flows, I assume that all farmers produce rice, whose trade is subject to information frictions and transportation costs. In addition, I assume that each consumer produces a unit of a costlessly traded numeraire good ("services"), from which he derives his income. Both farmers and consumers have CES preferences over rice and services, generating the inverse demand function:

$$p_i = \left( \frac{r_i P_i^{1-\sigma}}{\alpha Y_i} \right)^{-\frac{1}{\sigma}},$$

where $\sigma$ is the elasticity of substitution, $\alpha$ is the budget share of rice, $Y_i$ is the total income of farmers and consumers in region $i$, $r_i$ is the total supply of rice to region $i$, and $P_i \equiv \left((1-\alpha) + \alpha p_i^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$ is the CES price aggregator. I choose $\alpha = 0.2$ to match the reported budget share of rice in the household Family Income and Expenditure Survey (see Appendix B.4 for details) and $\sigma = 4$ to match the observed price variation.

The mass of farmers and consumers in each province are chosen to match population figures of farmers and non-farmers. The transportation frictions $\tau_{ij}$, the search probabilities $s_{ij}$, and the fixed cost of search $f_i$ are those estimated for rice in Section 4.2.[48]

[46] By geography, I mean the set of information frictions, transportation costs, distribution of landholdings and population of farmers and traders in all provinces.

[47] For simplicity, I assume that productivity shocks are independent across provinces.

[48] Since the fixed costs of search and the search probabilities are estimated separately for each year, I use the median estimates across all years. As such, the counterfactuals should be interpreted as relative to a "typical" year in the period 1995-2009.

## 5.1 Reducing information frictions versus transportation costs

I first compare the effects of policies reducing information and transportation frictions. To make the counterfactual policies comparable, I consider a 50 percent reduction in all fixed costs of search and a 50 percent reduction in all transportation costs.

Table 6 summarizes the aggregate impacts of the counterfactual policies. The first column presents the results of the baseline simulations. On average, 13.7 percent of rice production is traded, which results in a coefficient of variation of prices of 0.11 across provinces. The second and third columns present the effects of reducing the fixed costs of search and transportation costs, respectively. Not surprisingly, both policies increase total trade flows and reduce price dispersion. The reduction in trade frictions benefits farmers, while the reduction in price dispersion benefits consumers, resulting in increases in the average social welfare of both groups.[49] As approximately 35 percent of the Philippines population is employed in agriculture (BAS, 2011), the average increase in welfare for the whole population is 0.12 percent from the reduction in search costs and 0.17 percent from the reduction in transportation costs.[50] Hence, the simulation suggests the government should focus on reducing information frictions rather than transportation frictions as long as the reduction in information frictions can be achieved at less than two-thirds the cost of a comparable reduction in transportation costs.

The aggregate effects mask important distributional implications. Figure 6 presents the welfare effects on rice farmers as a function of their total landholdings. Because the fixed cost of search comprises a greater fraction of revenue for small farmers, they benefit disproportionately more from its reduction; indeed, the largest farmers become worse off owing to increased competition. In contrast, since reductions in the transportation cost reduce the variable cost of exporting and large farmers export more, the benefits are concentrated amongst large farmers. As a result, simulations imply that reductions in the fixed cost of search reduce inequality whereas reductions in transportation costs exacerbate it.

## 5.2 Land reform

Another policy that is oftentimes touted as improving efficiency while reducing inequality (see e.g. Binswanger, Deininger, and Feder (1995)) is the redistribution of land from large farmers to small farmers. What do information frictions imply for such a reform? To answer the question, I consider a counterfactual policy where the Pareto shape parameter governing land distribution

[49] To construct the social welfare, I use a utilitarian welfare function where all farmers and consumers are weighted equally. Since large farmers consume more than small farmers, they comprise a greater portion of total farmer utility.

[50] The aggregate welfare impacts are small in magnitude because welfare gains only arise from reducing variation in consumption since aggregate production is held constant. Furthermore, overall uncertainty is low since services are traded costlessly and comprise a large portion of the budget.

is increased by 50 percent while keeping the total land cultivated constant, thereby increasing the concentration of land held by small farmers.[51]

The fourth column of Table 6 presents the effects of such a reform. Since smaller farmers are less willing to search other markets, land reform reduces the efficient allocation of rice across provinces. As a result, trade flows fall and price dispersion increases, both by roughly one-third. The decline in trade reduces aggregate farmer welfare by 0.6 percent, while the increase in price volatility reduces aggregate consumer welfare by 0.5 percent. Hence, while land reform may reduce inequality, it does so at the cost of a loss in aggregate welfare.[52]

# 6 Incorporating intermediaries

In the Philippines, farmers rarely ship their own produce to other provinces; instead, they normally sell their produce to traders who then export to other markets. This may result in some concern about the analysis conducted thus far; for example, if both large and small farmers sell to the same set of traders, it is unclear why the land size of the farmer affects where her produce is shipped. In this section, I extend the basic model to incorporate such intermediaries and show that the basic predictions of the model remain unchanged. I first present the model extension. I then provide micro-level evidence of the search process using data on more than two million farmer sales. Finally, I use the extension of the model to separately estimate the fixed costs of search for farmers and traders.

## 6.1 Model

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.[53] 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[54] (referred to in the

[51]Since the total land cultivated is $M_i \frac{\theta_i}{\theta_i - 1}$, to keep the total land cultivated constant while increasing $\theta_i$ requires increasing the mass of farmers as well; intuitively, the laborers working on the land of large farmers become the owners of the land.

[52]Of course, with a non-utilitarian social welfare function that gives additional weight to the utility of small farmers, the reduction in inequality could be welfare enhancing even with the loss of efficiency.

[53]Whether intermediaries promote or inhibit efficiency is a topic of much recent debate in the trade literature, see e.g. Antràs and Costinot (2011) and Bardhan, Mookherjee, and Tsumagari (2010). 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.

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

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 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 search again for another trader, continuing until she finds a trader to which she is willing to sell.[55] 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 search elsewhere.[56] 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 backwards 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 in order to maximize expected profits. The greater the price the trader offers, the more produce he attracts from farmers, but the smaller his profit margin. In particular, a trader of type $s_t$ chooses $p_t$ to maximize his expected profits $\pi_t$:

$$\pi_t\left(s_t\right) = \max_{\tilde{p}\in[p,\infty)} \left(R\left(q\left(\tilde{p};s_t\right)\right) - \tilde{p}\right) q\left(\tilde{p};s_t\right), \tag{19}$$

where $R\left(q\right)$ is the expected per unit revenue the trader receives as a function of quantity and $q\left(\tilde{p};s_t\right)$ is the quantity that the trader purchases as a function of the price he offers and his conspicuousness in the market. Let $p_t\left(s_t\right)$ denote the buying price that maximizes equation (19). I show in the online appendix that traders with greater $s_t$ offer higher buying prices to farmers, i.e. $\frac{\partial}{\partial s_t} p_t\left(s_t\right) > 0$, since more conspicuous traders attract more farmers for a given increase in the purchase price.

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 landownings. This has two implications. First,

[55] In ongoing research, I extend the model to incorporate bargaining between the farmer and trader so that the price offered by the trader may differ depending on the farmer. Bargaining is a key component of the recent model of middlemen with search frictions by Wong and Wright (2011).

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

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, there exists a threshold land size $\varphi^*(p) \equiv \frac{g}{L(p)}$ (where $L(p) \equiv \int_p (p_t - p)\, dF_{p_t}(p_t)$ is the farmer value of search across traders) such that all farmers with landholdings $\varphi > \varphi^*(p)$ will choose to sell to a trader, while farmers with landholdings less than the threshold forego search and sell locally. I test for the existence of such a threshold in the following section.

## 6.2 Micro-level evidence of search

Incorporating intermediaries allows me to test for the existence of search behavior using micro-data on farmer sales. I have assembled more than two million observations of farmer sales for more than 200 agricultural crops within the period 2000-2009 (see Appendix B.8 for details). For each farmer that sold her produce in a particular month, data is collected on the price the farmer received, the quantity the farmer sold, and importantly, how much (if any) freight costs the farmer incurred. Since a farmer who incurs freight costs is more likely to have brought her goods to a market to sell to a trader, positive freight costs provide a proxy variable of whether a farmer engaged in search for a trader.[57]

One of the central predictions of the search model is that farmers choose to search if and only if the quantity that they sell is greater than a particular threshold. To test if this is the case, I regress whether or not a farmer $f$ in province $i$ selling commodity $c$ in month $m$ in year $t$ chooses to search for a trader on the (log) quantity that she sold:

$$search_{ficmt} = \beta \ln q_{ficmt} + \delta_{icmt} + \varepsilon_{ficmt},$$

where $\delta_{icmt}$ is a province-commodity-month-year fixed effect so that identification arises only from within-market variation. The first column of Table 7 presents the results. The probability of searching for a trader is strongly increasing in the quantity of produce sold; a 10 percent increase in the quantity sold is associated with a 0.29 percentage point (1.6 percent) increase in the probability of searching for a trader.

While consistent with the model, such a relationship may be due to other factors; for

[57] While the model supposes that a farmer who chooses not to search for a trader simply sells to local consumers, qualitative interviews with rice farmers suggest only a small minority sell directly to consumers. Instead, the majority of farmers that choose to forgo the search process sell to neighbors with whom they have a longstanding relationship. These *suki* relationships are common throughout the Philippines and often oblige the farmers to sell their produce only to their *suki* in exchange for credit access or other concessions (see Davis (1973) and Hendriks (1994)). Since selling to her *suki* means that the farmer did not search for a seller and farmers do not pay freight costs to sell to their neighbor, whether or not a farmer incurred freight costs remains a good measure of whether or not a farmer searched for a trader.

example, smaller farmers may be more likely to be obliged to sell to a particular person (and hence forego searching) because they have received loans from that person. Stronger evidence of the existence of search comes from the introduction of mobile phones. I have collected data on the location and the month of construction of all cell phone towers in the Philippines (see Appendix B.6 for details). If mobile phones reduce the fixed cost of searching for a trader, then the threshold land size should decline. As a result, farmers selling less produce should begin to search, increasing the overall number of farmers searching while reducing the positive correlation between quantity sold and the probability of search.

To test this, I regress:

$$search_{ficmt} = \beta_1 \ln q_{ficmt} + \beta_2 phone_{imt} + \gamma \ln q_{ficmt} \times phone_{imt} + \delta_{ict} + \delta_{cmt} + \varepsilon_{ficmt},$$

where $phone_{imt}$ is an indicator function equal to one if there exists a cell phone tower in province $i$ in month $m$ in year $t$, $\delta_{ict}$ is a province-commodity-year fixed effect and $\delta_{cmt}$ is a commodity-month-year fixed effect. The inclusion of $\delta_{ict}$ ensures that the effect of the introduction of mobile phones is identified only by comparing months prior to the construction of the first cell phone tower to the months after the construction of the tower within the same year, limiting the concern of the endogeneity of tower placement.[58] The inclusion of $\delta_{cmt}$ controls for aggregate variation in market conditions of commodity $c$ within a given month, e.g. seasonality of production. From the discussion above, the search model predicts $\beta_1 > 0$, $\beta_2 > 0$, and $\gamma < 0$.

The second column of Table 7 presents the results. In the months after the construction of the first cell phone tower, the fraction of farmers searching for traders increased by 5.7 percentage points relative to the months prior to the construction of the first cell phone tower within a given year. Furthermore, after the construction of a tower, the relationship between trader search and log quantity sold falls by 0.014, a decline of 30 percent. This result suggests that the introduction of cell phones led smaller farmers to search for traders. The third column includes a province-commodity-month-year fixed effect in order to control for any concern about the endogeneity of the month a tower was constructed. While the fixed effect does not allow for the identification of $\beta_2$, the interaction $\gamma$ remains negative and statistically significant. Hence, the farmer sales data provides micro-level evidence that information frictions exist.

## 6.3 Farmer and trader fixed costs of search

By incorporating traders into the model, I am able to separate the fixed cost of search estimated in Section 4.2 into the cost paid by the farmer to search for traders and the cost paid by traders to search across markets.

[58] An alternative specification comparing the 6 months prior to cell phone tower construction to the 6 months after yields very similar results.

The farmer fixed cost of search $g$ can be estimated from the fraction of farmers searching for traders. Since the search process for farmers is identical to the basic model, the fraction of farmers searching for traders, $\Lambda$, is equivalent to equation (8):

$$\Lambda = \left(\frac{L(p)}{g}\right)^{\theta-1}, \tag{20}$$

where $\theta$ is the shape parameter of the Pareto distribution of farmer land size. Since I observe the fraction of farmers searching for traders within a given market along with their land distribution and the distribution of prices, it is straightforward to estimate the farmer fixed cost of search $g$ from equation (20).

The trader fixed cost of search $f$ can be estimated from the minimum purchasing price observed. Since traders will only participate in the market if their expected profits are positive, only traders that are sufficiently conspicuous will choose to purchase produce from farmers. The marginal trader will have zero expected profit, which implies the fixed cost of search is exactly equal to his expected benefit of search:

$$f = Q_t^* K(p_t^*), \tag{21}$$

where $Q_t^*$ is the quantity purchased by the marginal trader and $p_t^*$ is the price he offers. Since I observe the lowest price offered by traders, I can estimate the value of search $K(p_t^*)$ using observed regional prices and the trade frictions estimated from Section 4.2. Furthermore, since I observe the fraction of quantity farmers sell to traders offering price $p_t^*$, I can estimate the total quantity purchased by marginal traders, allowing me to estimate the trader fixed cost of search using equation (21).[59] Since there may be more than one marginal trader, the estimated trader fixed cost of search should be interpreted as an upper bound.

Table 8 summarizes the results. The median estimated fixed cost of search for farmers is 1,105 pesos ($25). In contrast, the median of the trader fixed cost of search is 98,928 pesos ($2,200). While the trader fixed cost is an upper bound, the results suggest that the majority of search costs are associated with search across rather than within markets.

# 7 Conclusion

The goal of this paper has been to show that information frictions have important theoretical and empirical implications for trade. Theoretically, I have developed a trade model incorporating the costly search process producers undergo to learn about market conditions elsewhere. The model shows that the existence of information frictions leads to significant deviations from

[59]See the online appendix for details.

standard trade models, namely a failure of the no-arbitrage equilibrium condition and a departure from the standard equation governing bilateral trade flows. Furthermore, the model yields a method to separately identify transportation costs and information frictions using only observed trade flows and prices.

Empirically, I have shown that in the particular context of agriculture in the Philippines, information frictions are important. They contribute as much to observed price dispersion as transportation frictions, and are responsible for the vast majority of the negative relationship between trade flows and distance. Furthermore, simulations suggest that reducing information frictions would result in similar increases in aggregate welfare as comparable reductions in transportation costs but would also reduce inequality.

This paper provides the first step in examining the role of information frictions in trade. While the focus has been on agricultural trade flows, it is reasonable to expect that information frictions are important in other settings as well. A fruitful direction for future research would be to examine how information frictions affect trade flows in contexts such as manufacturing, where products are differentiated and firms have market power.

# References


Aker, J. (2010): "Information from markets near and far: Mobile phones and agricultural markets in Niger," *American Economic Journal: Applied Economics*, 2(3), 46–59.

Albornoz, F., H. Pardo, G. Corcos, and E. Ornelas (2010): "Sequential exporting," Working paper.

Alessandria, G., and J. Kaboski (2011): "Pricing-to-market and the failure of absolute PPP," *American Economic Journal: Macroeconomics*, 3(1), 91–127.

Anderson, J. (1979): "A theoretical foundation for the gravity equation," *The American Economic Review*, 69(1), 106–116.

——— (2011): "The gravity model," *Annual Review of Economics*, 3(1).

Anderson, J., and E. van Wincoop (2003): "Gravity with gravitas: A solution to the border puzzle," *The American Economic Review*, 93(1), 170–192.

——— (2004): "Trade costs," *Journal of Economic Literature*, 42(3), 691–751.

Antràs, P., and A. Costinot (2011): "Intermediated trade," *The Quarterly Journal of Economics*, 126(3), 1319–1374.

Arkolakis, C., A. Costinot, and A. Rodríguez-Clare (forthcoming): "New trade models, same old gains?," *The American Economic Review*.

Arkolakis, C., and T. Papageorgiou (2011): "Selection, growth, and learning," Working paper.

Bardhan, P., D. Mookherjee, and M. Tsumagari (2010): "Middlemen margins and globalization," Working paper.

BAS (2002a): "Marketing cost structure for corn," Discussion paper, Philippines Bureau of Agricultural Statistics.

——— (2002b): "Marketing cost structure for tomato," Discussion paper, Philippines Bureau of Agricultural Statistics.

——— (2003): "Marketing cost structure for potato," Discussion paper, Philippines Bureau of Agricultural Statistics.

——— (2007a): "Marketing cost structure for garlic," Discussion paper, Philippines Bureau of Agricultural Statistics.

——— (2007b): "Marketing cost structure for onion," Discussion paper, Philippines Bureau of Agricultural Statistics.

——— (2011): "Selected Statistics on Agriculture," Discussion paper, Philippines Bureau of Agricultural Statistics.

Binswanger, H., K. Deininger, and G. Feder (1995): "Power, distortions, revolt and reform in agricultural land relations," *Handbook of development economics*, 3, 2659–2772.

Brander, J., and P. Krugman (1983): "A reciprocal dumping model of international trade," *Journal of International Economics*, 15(3-4), 313–321.

Burdett, K., and K. Judd (1983): "Equilibrium price dispersion," *Econometrica*, pp. 955–969.

Chaney, T. (2008): "Distorted gravity: The intensive and extensive margins of international trade," *American Economic Review*, 98(4), 1707–1721.

——— (2011a): "The gravity equation in international trade: An explanation," Working paper.

——— (2011b): "The network structure of international trade," Working paper.

Davis, W. (1973): *Social relations in a Philippine market: self-interest and subjectivity.* Univ of California Press.

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.

Donaldson, D. (2008): "Railroads of the Raj: Estimating the impact of transportation infrastructure," Working paper.

Eaton, J., M. Eslava, C. Krizan, M. Kugler, and J. Tybout (2011): "A search and learning model of export dynamics," Working paper.

Eaton, J., and S. Kortum (2002): "Technology, geography, and trade," *Econometrica*, 70(5), 1741–1779.

Estudillo, J., A. Quisumbing, and K. Otsuka (2001): "Gender differences in land inheritance, schooling and lifetime income: Evidence from the rural Philippines," *Journal of Development Studies*, 37(4), 23–48.

Feenstra, R., J. Markusen, and A. Rose (2001): "Using the gravity equation to differentiate among alternative theories of trade," *Canadian Journal of Economics*, pp. 430–447.

Goldberg, P., and F. Verboven (2001): "The evolution of price dispersion in the European car market," *The Review of Economic Studies*, 68(4), 811.

Goyal, A. (2010): "Information, direct access to farmers, and rural market performance in central India," *American Economic Journal: Applied Economics*, 2(3), 22–45.

Heckscher, E. (1916): "Växelkursens grundval vid pappersmyntfot," *Ekonomisk Tidskrift*, 18(10), 309–312.

Hendriks, M. (1994): "Trade arrangements and interlinked credit in the Philippines," *Financial landscapes reconstructed: the fine art of mapping development.*, pp. 209–221.

Hillberry, R., and D. Hummels (2008): "Trade responses to geographic frictions: A decomposition using micro-data," *European Economic Review*, 52(3), 527–550.

Jensen, R. (2007): "The digital provide: Information (technology), market performance, and welfare in the south Indian fisheries sector," *The Quarterly Journal of Economics*, 122(3), 879–924.

Krugman, P. (1980): "Scale economies, product differentiation, and the pattern of trade," *The American Economic Review*, 70(5), 950–959.

——— (1986): "Pricing to market when the exchange rate changes," .

Lucas, R., and E. Prescott (1974): "Equilibrium search and unemployment," *Journal of Economic Theory*, 7(2), 188–209.

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

McCall, J. (1970): "Economics of information and job search," *The Quarterly Journal of Economics*, 84(1), 113–126.

McFadden, D. (1989): "A method of simulated moments for estimation of discrete response models without numerical integration," *Econometrica*, pp. 995–1026.

Melitz, M. (2003): "The impact of trade on intra-industry reallocations and aggregate industry productivity," *Econometrica*, 71(6), 1695–1725.

Melitz, M., and G. Ottaviano (2008): "Market size, trade, and productivity," *Review of Economic studies*, 75(1), 295–316.

Mortensen, D. (1970): "Job search, the duration of unemployment, and the Phillips curve," *The American Economic Review*, 60(5), 847–862.

Novy, D. (2010): "International trade and monopolistic competition without CES: Estimating translog gravity," Working paper.

NSO (2001): "Commodity flows in the Philippines," Discussion paper, Philippines National Statistics Office.

Rauch, J. (1999): "Networks versus markets in international trade," *Journal of International Economics*, 48(1), 7–35.

Rauch, J., and A. Casella (2003): "Overcoming informational barriers to international resource allocation: Prices and ties," *The Economic Journal*, 113(484), 21–42.

Rauch, J., and V. Trindade (2002): "Ethnic Chinese networks in international trade," *Review of Economics and Statistics*, 84(1), 116–130.

——— (2003): "Information, international substitutability, and globalization," *The American Economic Review*, 93(3), 775–791.

Samuelson, P. A. (1954): "The transfer problem and transport costs, II: Analysis of effects of trade impediments," *The Economic Journal*, 64(254), pp. 264–289.

Simonovska, I. (2011): "Income differences and prices of tradables," Working paper.

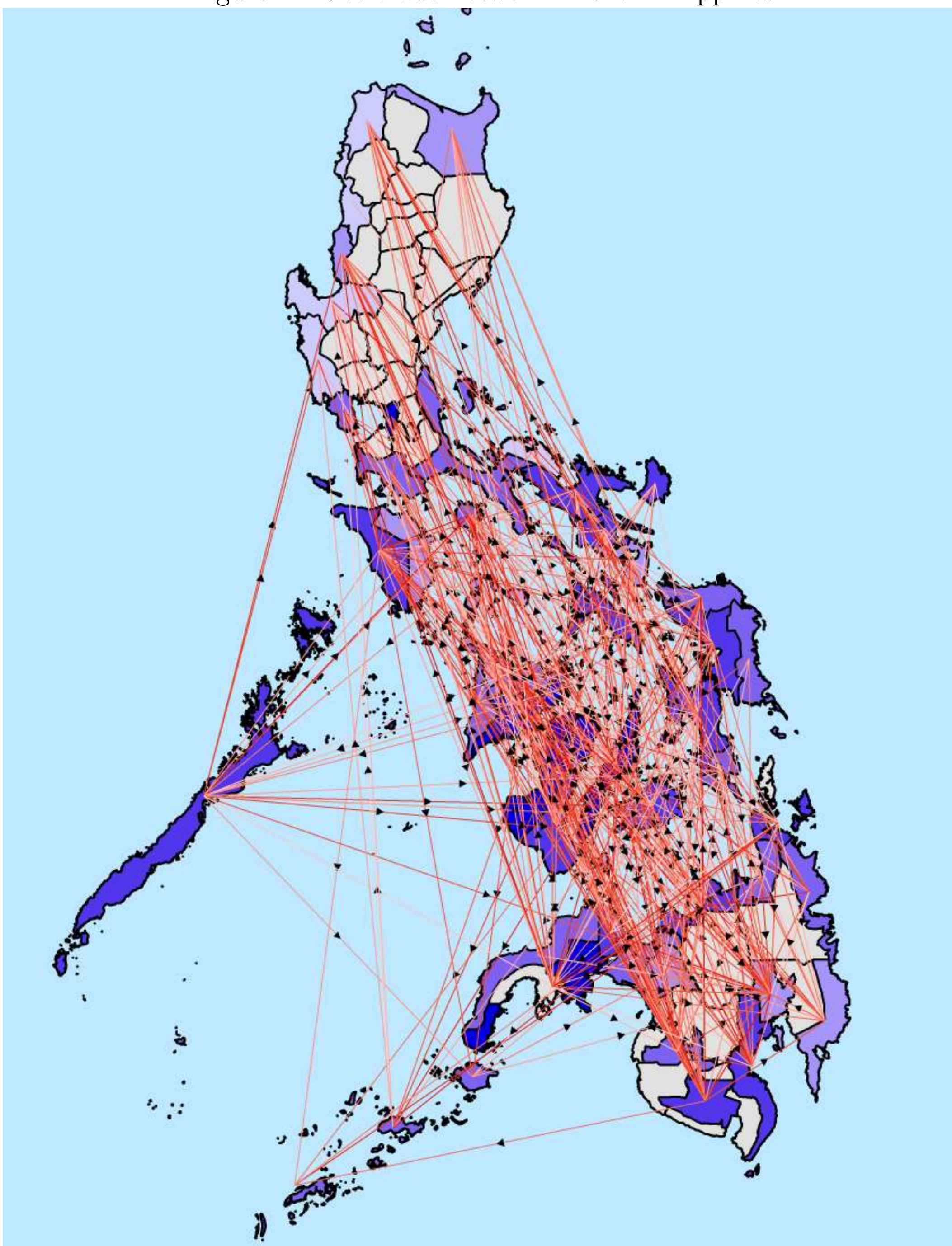
Waugh, M. (2010): "International trade and income differences," *The American Economic Review*, 100(5), 2093–2124.

Wong, Y., and R. Wright (2011): "Buyers, sellers, and middlemen: Variations on search-theoretic themes," Working paper.

Yang, D., and H. Choi (2007): "Are remittances insurance? Evidence from rainfall shocks in the Philippines," *The World Bank Economic Review*, 21(2), 219.

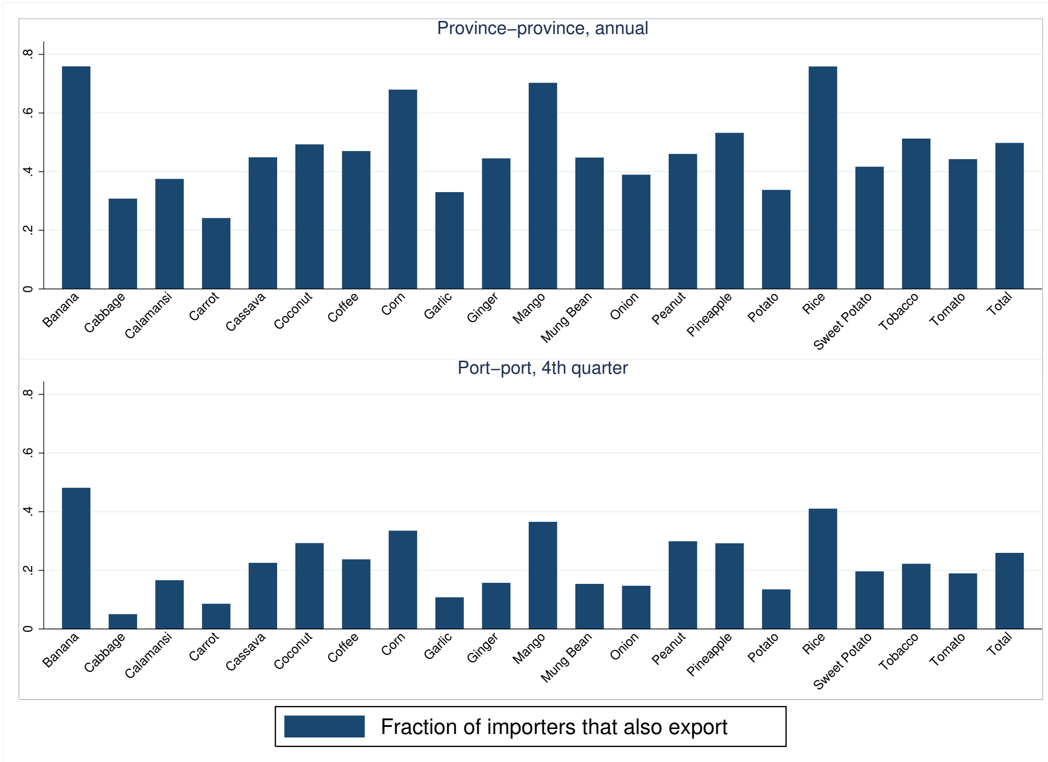
# Tables and figures

Figure 1: Rice trade network in the Philippines



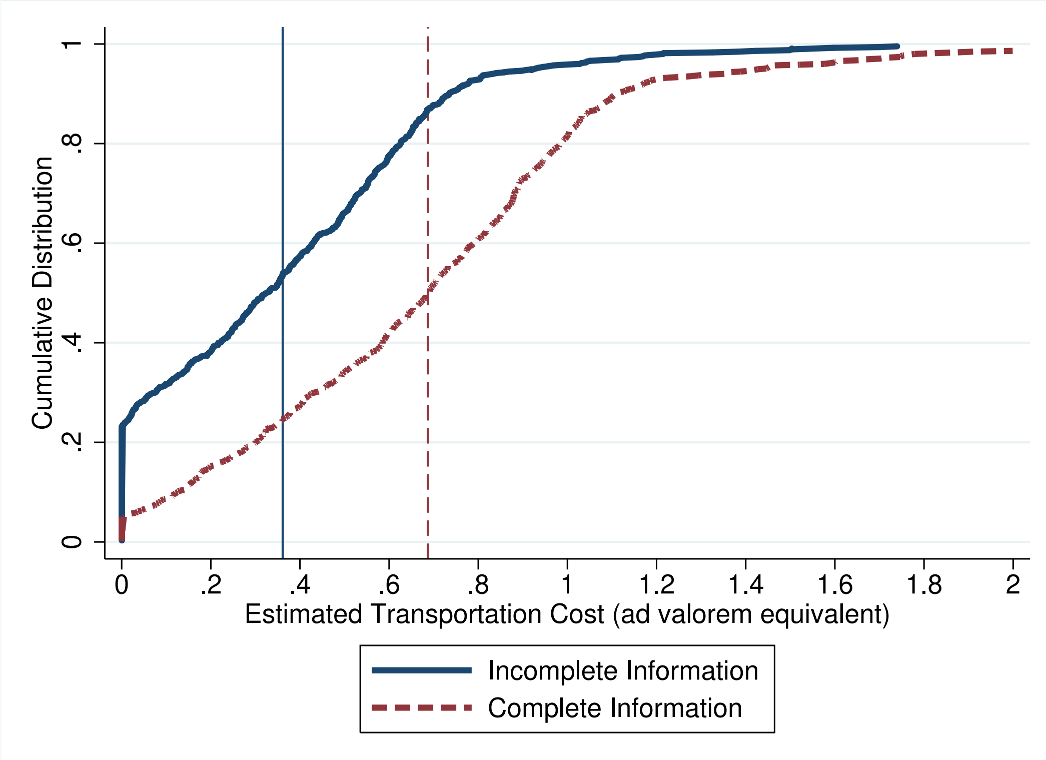
*Notes:* This figure shows the rice trade network in the Philippines. The shading of the provinces indicates the total observed rice trade flows (imports plus exports) from 1995-2009, where a darker shading indicates greater trade flows and gray indicates no observed trade flows. The lines indicate that trade in rice flows occurred between the two provinces in at least one year, with the arrow indicating the direction and the darker red lines indicated a greater amount of trade.

Figure 2: Agricultural trade patterns in the Philippines by commodity



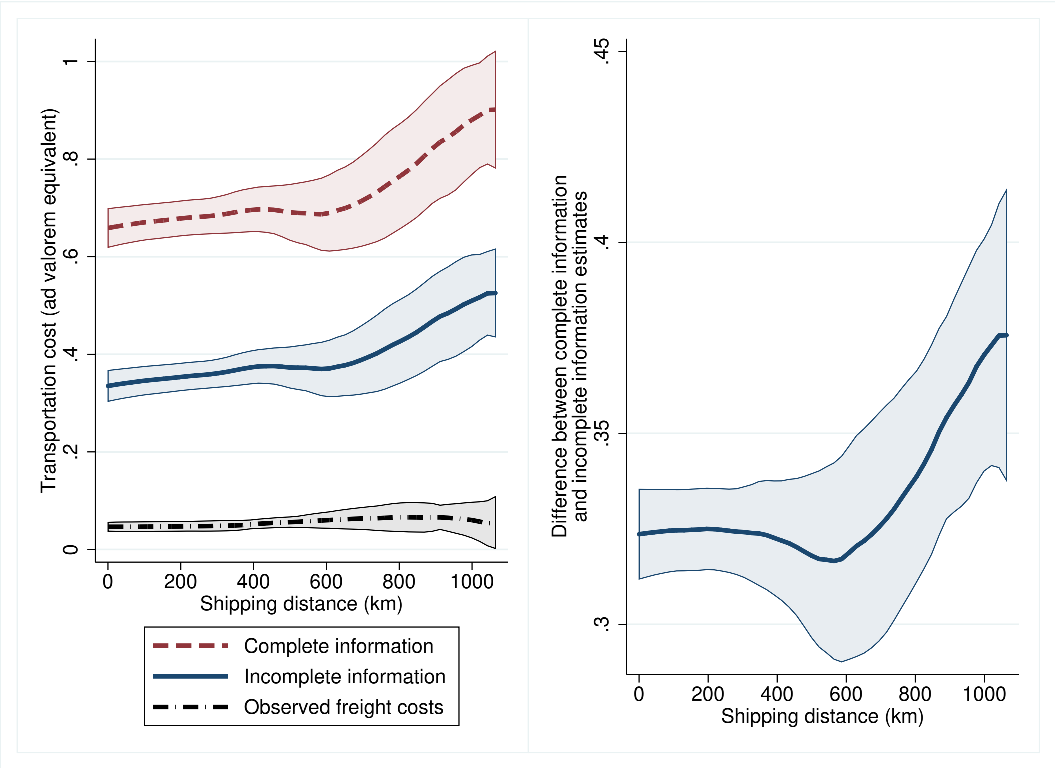

*Notes:* This figure shows the fraction of importers that also export the same commodity. Cebu and Manila are excluded to avoid instances of entrepôt. The sample for the top chart includes all annual bilateral agricultural trade flows between provinces. The sample for the bottom chart includes all bilateral trade flows occurring in the 4th quarter of each year disaggregated to port-to-port level. In both cases, all agricultural commodities are included, rather than just those with observed market prices.

Figure 3: Distribution of estimated transportation costs



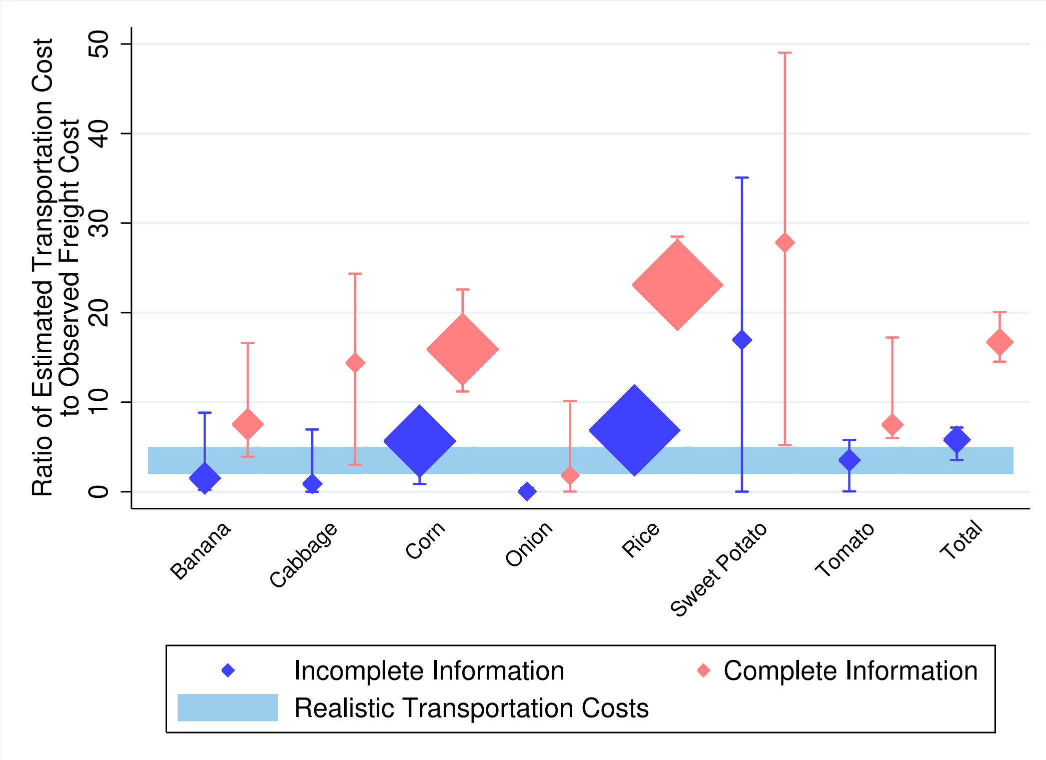

*Notes*: This figure depicts the cumulative distribution function of estimated transportation costs across origin-destination-commodities for complete information and incomplete information. The sample includes all origin-destination-commodity triplets with wholesale markets in which trade was observed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

Figure 4: Estimated transportation costs and shipping distance



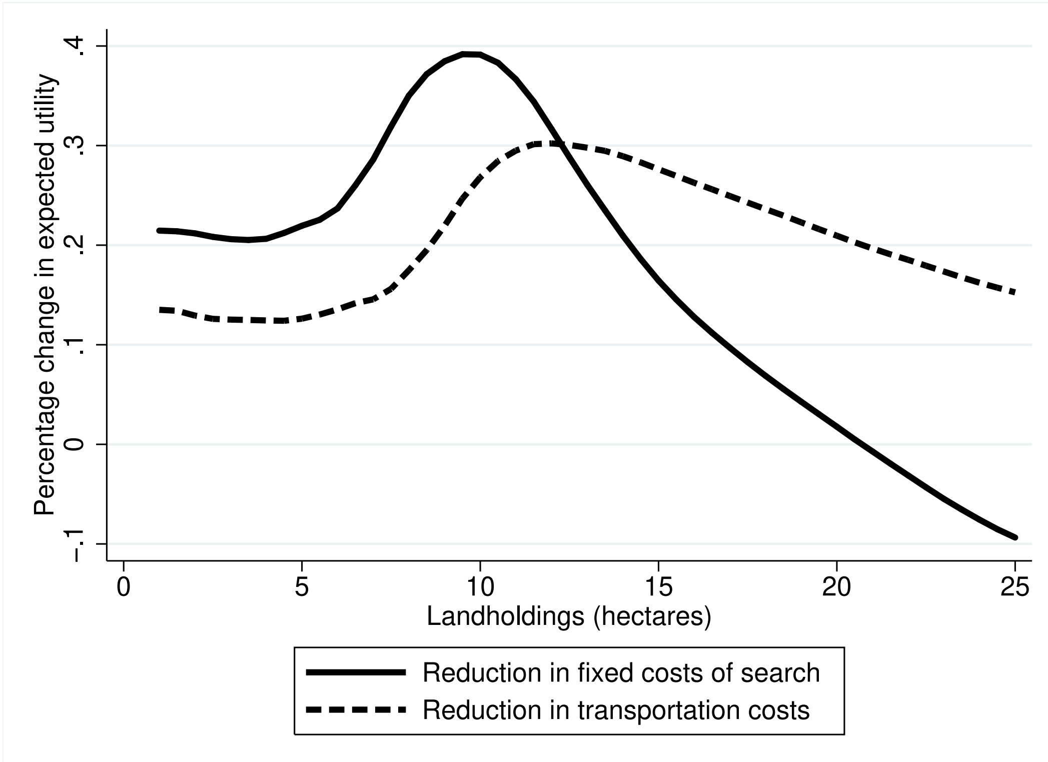

*Notes*: This figure shows how the estimated transportation costs are correlated with shipping distance. The left panel depicts the estimated transportation costs across origin-destination-commodities for complete information and incomplete information by shipping distance. The right panel depicts the difference between the complete information estimate and the incomplete information estimate by shipping distance. 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 5: Estimated transportation costs by commodity




*Notes*: Diamonds report the median ratio of the estimated transportation cost to the observed freight cost under the assumptions of incomplete and complete information. 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 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 6: Welfare effects of reductions in trade costs by size of landholding




*Notes*: This figure reports the change in the expected utility of rice farmers from a 50 percent reduction in the fixed cost of search and transportation costs, 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.

Table 1: Agricultural Commodities in the Trade Data

| Commodity | (1) Percentage of land area | (2) Percentage of output | (3) Provinces producing | (4) Province markets | (5) Percentage of trade value | (6) Number of observations |
|---|---|---|---|---|---|---|
| Banana | 3.4 | 13.8 | 75 | 22 | 0.5 | 170 |
| Cabbage | 0.1 | 0.2 | 58 | 22 | 0.0 | 36 |
| Corn | 20.2 | 11.3 | 75 | 71 | 48.0 | 1,360 |
| Garlic | 0.0 | 0.1 | 27 | 16 | 0.0 | 70 |
| Mung Bean | 0.3 | 0.2 | 73 | 10 | 0.0 | 7 |
| Onion | 0.1 | 0.6 | 39 | 17 | 0.1 | 92 |
| Pineapple | 0.5 | 1.7 | 71 | 17 | 0.0 | 39 |
| Rice | 34.5 | 35.8 | 75 | 77 | 49.8 | 2,417 |
| Sweet Potato | 0.9 | 0.8 | 75 | 19 | 0.0 | 37 |
| Tomato | 0.1 | 0.4 | 74 | 25 | 1.5 | 109 |
| Total | 60.1 | 64.9 | 82 | 82 | 100 | 4,337 |

Columns 1 and 2 are the average share of total land area and total agricultural output value from 2008-2010 taken from BAS (2011). Columns 3 and 4 report the number of provinces reporting any production or market price, respectively, in any year between 1990-2009. Columns 5 and 6 report the percentage of value and number of observations, respectively, of each commodity in the trade dataset. The trade data set includes all bilateral trade flows where: 1) the market price in the origin and destination are observed; and 2) the amount produced in the origin is observed.

Table 2: Evidence of information frictions: Failure of price arbitrage

| | Philippines | | U.S. | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Dep. var.: change in log destination price ratio | OLS | 2SLS | OLS | 2SLS |
| Change in log origin price ratio | 0.341*** | 0.284*** | 0.561*** | 0.831*** |
| | (0.031) | (0.080) | (0.025) | (0.051) |
| First Differences | Yes | Yes | Yes | Yes |
| Test coefficient = 1 (p-value) | 0.000 | 0.000 | 0.000 | 0.001 |
| R-squared | 0.134 | 0.047 | 0.293 | 0.153 |
| Observations | 808 | 808 | 1182 | 1182 |

*Notes:* First differences. The dependent variable is the change in the log wholesale price ratio of corn to rice in the destination province. Each observation is an exporter-importer-year triplet. The change in the origin price ratio of corn to rice is instrumented with the mean and standard deviation of monthly rainfall within the year. The p-value of the test whether the estimated coefficient is one (as is implied by complete information price arbitrage) is reported above. Column 3 and 4 reports the results for the price ratio of corn to hay in U.S. states for 2003-2009, using trade partners from the 2007 Commodity Flow Survey. Standard errors are reported in parentheses. Stars indicate statistical significance: * p<.10 ** p<.05 *** p<.01.

Table 3: Evidence of information frictions: Failure of the gravity equation

| Dep. var.: log quantity exported | (1) OLS | (2) 2SLS | (3) OLS | (4) 2SLS | (5) OLS | (6) 2SLS |
|---|---|---|---|---|---|---|
| Log destination price: $\ln p_{jct}$ | 0.136 | 0.867 | 1.083*** | 2.633*** | 1.042** | 3.045*** |
| | (0.336) | (0.567) | (0.399) | (0.525) | (0.408) | (0.473) |
| Log destination expenditures: $\ln X_{jt}$ | 0.885* | 0.840 | 0.832 | 0.741 | | |
| | (0.510) | (0.512) | (0.520) | (0.524) | | |
| Origin land dist.* Dest. price: $\theta_{ic} \times \ln p_{jct}$ | | | -0.256*** | -0.406*** | -0.251*** | -0.451*** |
| | | | (0.047) | (0.048) | (0.041) | (0.045) |
| Mean dependent variable | 12.191 | 12.191 | 12.191 | 12.191 | 12.191 | 12.191 |
| Origin-Commodity-Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Origin-Destination FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Destination-Year FE | No | No | No | No | Yes | Yes |
| R-squared | 0.061 | 0.094 | 0.141 | 0.155 | 0.000 | 0.095 |
| Observations | 4337 | 4337 | 4337 | 4337 | 4337 | 4337 |

*Notes:* Two-stage least squares. The unit of observation is an origin-destination-commodity-year quadruplet. Destination prices are instrumented with idiosyncratic rainfall shocks in the destination. Bootstrapped standard errors in parentheses. Stars indicate statistical significance: * p<.10 ** p<.05 *** p<.01.

Table 4: Summary Statistics of Structural Estimates

| | (1) Transportation cost $\hat{\tau}_{ijc}$ | (2) Search probability $\hat{s}_{ijct}$ | (3) Fixed cost of search $\hat{f}_{ict}$ |
|---|---|---|---|
| Mean | 1.36 | 0.092 | 54313.1 |
| Std. Dev. | 0.35 | 0.17 | 267784.1 |
| Median | 1.33 | 0.015 | 4647.5 |
| Minimum | 1 | 0 | 1 |
| Maximum | 3.23 | 1 | 4974043 |
| Coeff. of variation across commodities | 1.38 | 0.91 | 0.53 |
| Number of estimates | 650 | 4337 | 992 |
| Unit of identification | Origin-Destination-Commodity | Origin-Destination-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 across commodities within origin-destination pairs, origin-destination-year triplets, and origin-year pairs for transportation costs, search probabilities, and fixed costs of search, respectively.

Table 5: TRANSPORTATION COSTS VERSUS INFORMATION FRICTIONS IN THE GRAVITY EQUATION

| Dependent variable: | (1)<br>Log quantity | (2)<br>Info. frictions | (3)<br>Trans. costs | (4)<br>Log freight |
|---|---|---|---|---|
| Log shipping distance | -0.437*** | -0.411*** | -0.026*** | 0.002* |
| | (0.037) | (0.035) | (0.004) | (0.001) |
| Origin-Product-Year FE | Yes | Yes | Yes | Yes |
| Destination-Product-Year FE | Yes | Yes | Yes | Yes |
| R-squared | 0.058 | 0.057 | 0.021 | 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.

Table 6: COMPARING COUNTERFACTUAL POLICIES

| | (1)<br>Baseline | (2)<br>Reduction in search costs | (3)<br>Reduction in transportation costs | (4)<br>Land redistribution |
|---|---|---|---|---|
| Trade flows | 13.67 | 15.33 | 14.26 | 8.54 |
| (percentage of production) | (0.846) | (0.890) | (0.991) | (0.754) |
| Price dispersion | 0.110 | 0.101 | 0.093 | 0.143 |
| (coefficient of variation) | (0.006) | (0.006) | (0.004) | (0.006) |
| Farmer utility | N/A | 0.07 | 0.42 | -0.58 |
| (percentage change from baseline) | | (0.091) | (0.211) | (0.108) |
| Consumer utility | N/A | 0.14 | 0.04 | -0.52 |
| (percentage change from baseline) | | (0.038) | (0.062) | (0.100) |

*Notes:* The reduction in search costs and transportation costs refer to the effect of reducing all fixed costs of search or transportation costs by 50 percent, respectively. The land redistribution policy increases the Pareto distribution parameter of the observed land redistribution by 50 percent (which concentrates landholdings with smaller farmers) while keeping overall production constant. Changes in farmer and consumer utility are calculated using a utilitarian social welfare function weighting all individuals equally. Standard errors calculated from 100 simulations with stochastic productivity draws are reported in parentheses.

Table 7: FARMER SEARCH AND THE IMPACT OF MOBILE PHONES

| Dep. var.: farmer searched for trader | (1) OLS | (2) OLS | (3) OLS |
|---|---|---|---|
| Log quantity sold | 0.029*** | 0.041*** | 0.047*** |
| | (0.001) | (0.001) | (0.004) |
| Mobile phone access | | 0.057*** | |
| | | (0.005) | |
| Mobile phone access*log quantity sold | | -0.014*** | -0.018*** |
| | | (0.001) | (0.004) |
| Province-Commodity-Year FE | Yes | Yes | Yes |
| Commodity-Year-Month FE | Yes | Yes | Yes |
| Province-Commodity-Year-Month FE | Yes | No | Yes |
| Mean of Dep. Var. | 0.178 | 0.178 | 0.178 |
| R-squared (within) | 0.007 | 0.007 | 0.008 |
| Observations | 2357257 | 2357257 | 2357257 |

*Notes:* The dependent variable is an indicator variable equal to one if the farmer incurred freight costs. The unit of observation is a farmer sale in a province-commodity-year-month. Mobile phone access is an indicator variable equal to one if there existed a cell tower in the province in that particular month. Standard errors clustered at the province-commodity-year-month are the reported in parentheses. Stars indicate statistical significance: * p<.10 ** p<.05 *** p<.01.

Table 8: INTERMEDIARIES AND THE ESTIMATED FIXED COST OF SEARCH

| | Basic Model | Incorporating Intermediaries | |
|---|---|---|---|
| Estimated fixed cost | (1) Farmer | (2) Trader | (3) Farmer |
| Mean | 74404 | 4626212 | 2879 |
| Std.Dev. | 365009 | 13909665 | 6633 |
| Median | 10307 | 98928 | 1105 |
| Minimum | 4 | 0 | 0 |
| Maximum | 4974043 | 107737736 | 85820 |
| Number of estimates | 267 | 267 | 267 |

*Notes:* Fixed costs are reported in 2000 Philippines pesos (1 USD is approximately equal to 45 PHP). Each observation is a province-commodity-year for which fixed costs are estimated in both the basic model and the model extension incorporating intermediaries. The fixed costs of search are estimated for the model extension incorporating intermediaries when individual farmer sales data exist and some but not all farmers sell to traders.

# Appendix

The appendix is composed of two subsections. In the first subsection, I provides details of the derivations and proofs for the model. In the second subsection, I describe the data.

## A Model derivations and proofs

In this subsection, I provide details of the derivations of the model and the proof of the existence of an equilibrium.

### A.1 Search

The value to the farmer is:

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

Since the solution yields a reservation price $\bar{p}_i(\varphi)$ such that the farmer will choose to sell if $p \geq \bar{p}_i(\varphi)$ , the value can be written as:

$$V_i(p;\varphi) = \begin{Bmatrix} \int V_i(p';\varphi)\, dF^i_{\frac{p}{\tau}}(p') - f \text{ if } p < \bar{p}_i(\varphi) \\ \varphi p \text{ if } p \geq \bar{p}_i(\varphi) \end{Bmatrix}. \tag{22}$$

The optimal reservation price satisfies:

$$\varphi \bar{p}_i(\varphi) = \int V_i(p';\varphi)\, dF^i_{\frac{p}{\tau}}(p') - f_i. \tag{23}$$

Substituting equation (22) into (23) indicates that at the reservation price, the fixed cost of search should equal the value of continuing to search:

$$f_i = \varphi \int_{\bar{p}_i(\varphi)} \left(p' - \bar{p}_i(\varphi)\right) dF^i_{\frac{p}{\tau}}(p'). \tag{24}$$

### A.2 Producer profits

Consider a farmer from region $i$ with production $\varphi$ and optimal reservation price $\bar{p}_i(\varphi)$. The probability that the producer will sell in a particular region is $1 - F^i_{\frac{p}{\tau}}(\bar{p}_i(\varphi))$. The probability that a farmer sells to the $n^{th}$ searched regions (after the home region) is $\left(1 - F^i_{\frac{p}{\tau}}(\bar{p}_i(\varphi))\right) F^i_{\frac{p}{\tau}}(\bar{p}_i(\varphi))^n$.

Hence, the expected revenue that a farmer will receive from the search process, $R(\varphi)$ is:

$$E(R(\varphi)) = \sum_{n=0}^{\infty} F^i_{\frac{p}{\tau}}(\bar{p}_i(\varphi))^n \left(1 - F^i_{\frac{p}{\tau}}(\bar{p}_i(\varphi))\right)\left(\frac{\varphi \int_{\bar{p}(\varphi)}^{\infty} p dF^i_{\frac{p}{\tau}}(p)}{\left(1 - F^i_{\frac{p}{\tau}}(\bar{p}_i(\varphi))\right)} - n f_i\right) \Leftrightarrow$$

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

Rearranging yields:

$$f_i = \varphi \int_{\bar{p}(\varphi)}^{\infty} \left(p - \frac{E(R(\varphi))}{\varphi}\right) dF(p). \tag{26}$$

Comparing equation (26) to equation (1) yields that the expected revenue a farmer will receive from searching is simply the product of her reservation price and the quantity produced, i.e.

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

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 price $\bar{p}(\varphi)$, the total revenue amongst all firms in a region, $R_i$ is:

$$R_i = A_i M_i \left[p_i \int_{b_i}^{\varphi^*(p_i)} \varphi dF^i_{\varphi}(\varphi) + \int_{\varphi^*(p_i)}^{\infty} \bar{p}(\varphi)\varphi dF^i_{\varphi}(\varphi)\right] =$$

$$\theta_i A_i M_i \left[p_i \int_{b_i}^{\varphi^*(p_i)} \varphi^{-\theta_i} d\varphi + \int_{\varphi^*(p_i)}^{\infty} \bar{p}(\varphi)\varphi^{-\theta_i} d\varphi\right],$$

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 $\frac{p_i}{\theta_i - 1}\left(1 - \varphi^*(p_i)^{1-\theta_i}\right)$. The second integral can be calculated using integration by parts followed by a change of variables:

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

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 $\frac{\theta_i}{\theta_i-1} A_i M_i$, 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 $\int_{p_i}^{p_i^{\max}} \left( 1 - \left( \frac{K_i(p)}{f_i} \right)^{\theta_i - 1} \right) dp$ captures the per-unit loss in profits due to the existence of information frictions. This is intuitive; from equation (8), $\left( \frac{K_i(p)}{f_i} \right)^{\theta_i - 1}$ 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 $\left( \frac{K_i(p)}{f_i} \right)^{\theta_i - 1}$ and reducing the loss in profits due to information frictions.

## A.3 Intensive margin

To calculate the total trade flows from $i$ to $j$ when $\frac{p_j}{\tau_{ij}} > p_i$ requires summing the trade flows over all possible search paths. Let $Q_{ij}^r$ be the quantity exported by firms that have searched region $j$ after searching $r$ other regions (not including the home region). Note that:

$$Q_{ij}^0 = A_i M_i s_{ij} \int_{\varphi_i^*(p_i)}^{\varphi_i^*\left(\frac{p_j}{\tau_{ij}}\right)} \varphi dF_\varphi^i(\varphi).$$

For $r > 0$, the expression becomes more complicated, as the prices discovered prior to searching region $j$ affect the mass of firms arriving in in region $j$. For $r = 1$, we have:

$$Q_{ij}^1 = A_i M_i s_{ij} \left[ F_{\frac{p}{\tau}}^i(p_i) \int_{\varphi_i^*(p_i)}^{\varphi_i^*\left(\frac{p_j}{\tau_{ij}}\right)} \varphi dF_\varphi^i(\varphi) + \int_{p_i}^{\frac{p_j}{\tau_{ij}}} \int_{\varphi_i^*(\tilde{p})}^{\varphi_i^*\left(\frac{p_j}{\tau_{ij}}\right)} \varphi dF_\varphi^i(\varphi)\, dF_{\frac{p}{\tau}}^i(\tilde{p}) \right].$$

The first term in the brackets corresponds to those firms where the previously searched region had a price less than the home price; as a result, no firms sold there so they all arrived in region $j$. The second term corresponds to those firms that discovered a price (net of transportation costs) $\tilde{p} \in \left[p_i, \frac{p_j}{\tau_{ij}}\right]$. As a result, all firms with $\varphi \in (\varphi_i^*(p_i), \varphi_i^*(\tilde{p})]$ chose to sell in the first searched region and hence would not arrive in region $j$.

Extending this logic to $r$ interim searches, we have:

$$Q_{ij}^r = A_i M_i s_{ij} \left[ \left( F_{\frac{p}{\tau}}^i(p_i) \right)^r \int_{\varphi_i^*(p_i)}^{\varphi_i^*\left(\frac{p_j}{\tau_{ij}}\right)} \varphi dF_\varphi^i(\varphi) + \int_{p_i}^{\frac{p_j}{\tau_{ij}}} \int_{\varphi_i^*(\tilde{p})}^{\varphi_i^*\left(\frac{p_j}{\tau_{ij}}\right)} \varphi dF_\varphi^i(\varphi)\, d\left( F_{\frac{p}{\tau}}^i(\tilde{p}) \right)^r \right], \quad (27)$$

as the probability of receiving a price below $p$ for $r$ consecutive independent draws is simply $\left( F_{\frac{p}{\tau}}^i(p) \right)^r$. Since farmers could arrive in region $j$ after any number of searches, total trade flows can be calculated by summing over all possible $r$ :

$$Q_{ij} = \sum_{r=0}^{\infty} Q_{ij}^{r}.$$

It is straightforward to calculate the infinite sum of the first term of $Q_{ij}^r$. The second term is more involved.[60] Substituting equation (2) and the assumption about Pareto distributions into equation (27) yields:

$$Q_{ij}^{r} = \frac{\theta_i}{\theta_i - 1} f_i^{1-\theta_i} A_i M_i s_{ij} \left[ \begin{array}{c} \left(F_{\frac{p}{\tau}}^{i}\left(p_i\right)\right)^{r} \left(K_i\left(p_i\right)^{\theta_i - 1} - K_i\left(\frac{p_j}{\tau_{ij}}\right)^{\theta_i - 1}\right) + \\ \int_{p_i}^{\frac{p_j}{\tau_{ij}}} \left(K_i\left(\tilde{p}\right)^{\theta_i - 1} - K_i\left(\frac{p_j}{\tau_{ij}}\right)^{\theta_i - 1}\right) d\left(\left[F_{\frac{p}{\tau}}^{i}\left(\tilde{p}\right)\right]^{r}\right) \end{array} \right].$$

Splitting the second integral into two parts and canceling like terms yields:

$$Q_{ij}^{r} = \frac{\theta_i}{\theta_i - 1} f_i^{1-\theta_i} A_i M_i s_{ij} \left[ \begin{array}{c} \left(F_{\frac{p}{\tau}}^{i}\left(p_i\right)\right)^{r} K_i\left(p_i\right)^{\theta_i - 1} - \left(F_{\frac{p}{\tau}}^{i}\left(\frac{p_j}{\tau_{ij}}\right)\right)^{r} K_i\left(\frac{p_j}{\tau_{ij}}\right)^{\theta_i - 1} \\ + \int_{p_i}^{\frac{p_j}{\tau_{ij}}} K_i\left(\tilde{p}\right)^{\theta_i - 1} d\left(\left[F_{\frac{p}{\tau}}^{i}\left(\tilde{p}\right)\right]^{r}\right) \end{array} \right].$$

Summing over $r$, we have:

$$Q_{ij} = \frac{\theta_i}{\theta_i - 1} f_i^{1-\theta_i} A_i M_i s_{ij} \left[ \frac{K_i\left(p_i\right)^{\theta_i - 1}}{1 - F_{\frac{p}{\tau}}^{i}\left(p_i\right)} - \frac{K_i\left(\frac{p_j}{\tau_{ij}}\right)^{\theta_i - 1}}{1 - F_{\frac{p}{\tau}}^{i}\left(\frac{p_j}{\tau_{ij}}\right)} - \sum_{r=0}^{\infty} \int_{p_i}^{\frac{p_j}{\tau_{ij}}} K_i\left(\tilde{p}\right)^{\theta_i - 1} d\left(\left[F_{\frac{p}{\tau}}^{i}\left(\tilde{p}\right)\right]^{r}\right) \right]. \tag{28}$$

Note that $\int_{p_i}^{\frac{p_j}{\tau_{ij}}} K_i\left(\tilde{p}\right)^{\theta_i - 1} d\left(\left[F_{\frac{p}{\tau}}^{i}\left(\tilde{p}\right)\right]^{r}\right) = \sum_{k \neq i} \mathbf{1}\left\{p_i < \frac{p_k}{\tau_{ik}} \leq \frac{p_j}{\tau_{ij}}\right\} \left( \begin{array}{c} \frac{1}{1 - F_{\frac{p}{\tau}}^{i}\left(\frac{p_k}{\tau_{ik}}\right)} - \\ \frac{1}{1 - \left(F_{\frac{p}{\tau}}^{i}\left(\frac{p_k}{\tau_{ik}}\right) - s_{ik}\right)} \end{array} \right) K_i\left(\frac{p_k}{\tau_{ik}}\right)^{\theta_i - 1}$,

so that:

$$\sum_{r=0}^{\infty} \int_{p_i}^{\frac{p_j}{\tau_{ij}}} K_i\left(\tilde{p}\right)^{\theta_i - 1} d\left(\left[F_{\frac{p}{\tau}}^{i}\left(\tilde{p}\right)\right]^{r}\right) = \sum_{k \neq i} \mathbf{1}\left\{p_i < \frac{p_k}{\tau_{ik}} \leq \frac{p_j}{\tau_{ij}}\right\} K_i\left(\frac{p_k}{\tau_{ik}}\right)^{\theta_i - 1} \left( \begin{array}{c} \frac{1}{1 - F_{\frac{p}{\tau}}^{i}\left(\frac{p_k}{\tau_{ik}}\right)} - \\ \frac{1}{1 - \left(F_{\frac{p}{\tau}}^{i}\left(\frac{p_k}{\tau_{ik}}\right) - s_{ik}\right)} \end{array} \right).$$

---

[60] Without an assumption of the distribution of productivities, it is still possible to bound $Q_{ij}$:

$$\frac{A_i M_i s_{ij}}{1 - F_{\frac{p}{\tau}}^{i}\left(p_i\right)} \int_{\varphi_i^*\left(p_i\right)}^{\varphi_i^*\left(\frac{p_j}{\tau_{ij}}\right)} \varphi dF_{\varphi}^{i}\left(\varphi\right) < Q_{ij} < \frac{A_i M_i s_{ij}}{1 - F_{\frac{p}{\tau}}^{i}\left(\frac{p_j}{\tau_{ij}}\right)} \int_{\varphi_i^*\left(p_i\right)}^{\varphi_i^*\left(\frac{p_j}{\tau_{ij}}\right)} \varphi dF_{\varphi}^{i}\left(\varphi\right).$$

To see this, note that the second term in the brackets of equation (27) is bounded above by $\left(\left(F_{\frac{p}{\tau}}^{i}\left(\frac{p_j}{\tau_{ij}}\right)\right)^{r} - \left(F_{\frac{p}{\tau}}^{i}\left(p_i\right)\right)^{r}\right) \int_{p_i}^{\frac{p_j}{\tau_{ij}}} \int_{\varphi_i^*\left(p_i\right)}^{\varphi_i^*\left(\frac{p_j}{\tau_{ij}}\right)} \varphi dF_{\varphi}^{i}\left(\varphi\right)$ and below by 0, so that $q_{ij}^{r} \in \left[A_i M_i s_{ij} \left(F_{\frac{p}{\tau}}^{i}\left(p_i\right)\right)^{r} \int_{\varphi_i^*\left(p_i\right)}^{\varphi_i^*\left(\frac{p_j}{\tau_{ij}}\right)} \varphi dF_{\varphi}^{i}\left(\varphi\right), A_i M_i s_{ij} \left(F_{\frac{p}{\tau}}^{i}\left(\frac{p_j}{\tau_{ij}}\right)\right)^{r} \int_{\varphi_i^*\left(p_i\right)}^{\varphi_i^*\left(\frac{p_j}{\tau_{ij}}\right)} \varphi dF_{\varphi}^{i}\left(\varphi\right)\right]$. By summing over all $r$, the result follows immediately.

Without loss of generality, index the prices (net of transportation costs) in other potential destinations that are between $p_i$ and $\frac{p_j}{\tau_{ij}}$ as $p_k^{ij}$, where $p_0^{ij} \equiv p_i$, $p_K^{ij} = \frac{p_j}{\tau_{ij}}$, and $p_{k-1}^{ij} \leq p_k^{ij}$ $\forall k \in \{1, ..., K\}$. Note that $F_{\frac{p}{\tau}}^i \left(p_k^{ij}\right) - s_{ik} = F_{\frac{p}{\tau}}^i \left(p_{k-1}^{ij}\right)$ $\forall k \in \{1, ..., K\}$, where $s_{ik}$ is the probability that region $i$ searches the region with price $p_k^{ij}$. Hence:

$$\sum_{r=0}^{\infty} \int_{p_i}^{\frac{p_j}{\tau_{ij}}} K_i(\tilde{p})^{\theta_i - 1} d\left(\left[F_{\frac{p}{\tau}}^i(\tilde{p})\right]^r\right) = \sum_{k=1}^{K} K_i\left(p_k^{ij}\right)^{\theta_i-1}\left(\frac{1}{1 - F_{\frac{p}{\tau}}^i\left(p_k^{ij}\right)} - \frac{1}{1 - F_{\frac{p}{\tau}}^i\left(p_{k-1}^{ij}\right)}\right).$$

Rearranging the sum yields:

$$\sum_{r=0}^{\infty} \int_{p_i}^{\frac{p_j}{\tau_{ij}}} K_i(\tilde{p})^{\theta_i - 1} d\left(\left[F_{\frac{p}{\tau}}^i(\tilde{p})\right]^r\right) = -\frac{K\left(p_1^{ij}\right)^{\theta_i-1}}{1 - F_{\frac{p}{\tau}}^i(p_i)} + \frac{K_i\left(\frac{p_j}{\tau_{ij}}\right)^{\theta_i-1}}{1 - F_{\frac{p}{\tau}}^i\left(\frac{p_j}{\tau_{ij}}\right)} + \sum_{k=1}^{K-1} \frac{K_i\left(p_k^{ij}\right)^{\theta_i-1} - K_i\left(p_{k+1}^{ij}\right)^{\theta_i-1}}{1 - F_{\frac{p}{\tau}}^i\left(p_k^{ij}\right)}. \tag{29}$$

Substituting equation (29) into equation (28) yields:

$$Q_{ij} = \frac{\theta_i}{\theta_i - 1} f_i^{1-\theta_i} A_i M_i s_{ij} \sum_{k=1}^{K} \frac{K_i\left(p_{k-1}^{ij}\right)^{\theta_i-1} - K_i\left(p_k^{ij}\right)^{\theta_i-1}}{1 - F_{\frac{p}{\tau}}^i\left(p_{k-1}^{ij}\right)},$$

as required.

To derive the approximation in equation (6), note that since $\frac{\partial}{\partial p} K_i(p) = -\left(1 - F_{\frac{p}{\tau}}^i(p)\right)$, I can write:a first order Taylor approximation yields:

$$\sum_{l=1}^{L} \frac{K_i\left(p_{l-1}^{ij}\right)^{\theta_i-1} - K_i\left(p_l^{ij}\right)^{\theta_i-1}}{1 - F_{\frac{p}{\tau}}^i\left(p_{l-1}^{ij}\right)} = (\theta_i - 1) \sum_{l=1}^{L} \frac{K_i\left(p_{l-1}^{ij} + \Delta p_l\right)^{\theta_i-1} - K_i\left(p_{l-1}^{ij}\right)^{\theta_i-1}}{\frac{\partial}{\partial p} K_i\left(p_{l-1}^{ij}\right)} K_i\left(p_{l-1}^{ij}\right)^{\theta_i-2},$$

where $\Delta p_l \equiv p_l^{ij} - p_{l-1}^{ij}$, so that by the definitions of derivatives and Riemann integrals, we have:

$$\lim_{\max_l(\Delta p_l)\to 0} \sum_{l=1}^{L} \frac{K_i\left(p_{l-1}^{ij}\right)^{\theta_i-1} - K_i\left(p_l^{ij}\right)^{\theta_i-1}}{1 - F_{\frac{p}{\tau}}^i\left(p_{l-1}^{ij}\right)} = (\theta_i - 1) \int_{p_i}^{\frac{p_j}{\tau_{ij}}} K_i(p)^{\theta_i - 2} dp,$$

as required.

## A.4 Existence and uniqueness of equilibrium

In this subsection, I prove the existence and uniqueness of equilibrium prices. I first prove existence and then prove uniqueness.

### A.4.1 Existence

The proof of existence is comprised of 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 Brouwer's fixed point theorem to prove existence.

**Step 1:** In this step, I characterize 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:

$$Q_{jj} = \max\left\{\theta_j A_j M_j \int_1^{\varphi^*(p_j)} \varphi^{-\theta_j} d\varphi, 0\right\} = \max\left\{\frac{\theta_j}{\theta_j - 1} A_j M_j \left(1 - \left(\frac{K_j(p_j)}{f_j}\right)^{\theta_j - 1}\right), 0\right\}. \tag{30}$$

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 the home price, the lower 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).

Total imports $I_j$ requires summing over all exporters to $j$ using equations (4) and (5):

$$I_j = \sum_{i \neq j} \mathbf{1}\left\{\frac{p_j}{\tau_{ij}} > p_i\right\} \frac{\theta_i}{\theta_i - 1} f_i^{1-\theta_i} A_i M_i s_{ij} \sum_{l=1}^{L} \frac{K_i\left(p_{l-1}^{ij}\right)^{\theta_i - 1} - K_i\left(p_l^{ij}\right)^{\theta_i - 1}}{1 - F_{\frac{p}{\tau}}^i\left(p_{l-1}^{ij}\right)} \geq 0. \tag{31}$$

Recall that $p_L^{ij} \equiv \frac{p_j}{\tau_{ij}}$, so that $I_j$ is also 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 $i$ and 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 $j$).[61]

Define $r_j(\mathbf{p}) : \mathbb{R}_+^N \to \mathbb{R}_+ \equiv Q_{jj} + I_j$ to be the total supply of rice to region $j$ as a function of all prices in the world $\mathbf{p} \equiv [p_1, ..., p_j, ...p_N]'$. From above, $r_j(\mathbf{p})$ is continuous in all elements of $p$, strictly increasing in its $j^{th}$ element, and weakly decreasing in all other elements (strictly decreasing as long as $i$ exports to $j$ or $j$ exports to $i$). Since $r_j(\mathbf{p})$ is constructed using equations (4) and (5), it satisfies the first property of equilibrium in Section 2.4.

**Step 2:** In this step, I characterize a function whose fixed point is the set of equilibrium prices. Define $G(\mathbf{p}) : \mathbb{R}_+^N \to \mathbb{R}_+^N \equiv [D_1(r_1(\mathbf{p})), ..., D_j(r_j(\mathbf{p})), ..., D_N(r_N(\mathbf{p}))]'$ to be the set of prices that would result from the inverse demand function in each region when each region is supplied with $r_j(\mathbf{p})$. Since $D_j(\cdot)$ is continuous for all $r$ and $r_j(\cdot)$ is continuous for all elements of $\mathbf{p}$, $G$ is continuous in all elements of $\mathbf{p}$ as well. Since $r_j(\mathbf{p})$ satisfies the first property of equilibrium, a set of prices $\mathbf{p}^*$ are a set of equilibrium prices if and only if $G(\mathbf{p}^*) = \mathbf{p}^*$, i.e. the prices resulting from the inverse demand function given supply $r_j(\mathbf{p}^*)$ in all $j$ 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 Brouwer's fixed point theorem. I

[61] Continuity is not affected by the presence of the indicator function, as when $\frac{p_j}{\tau_{ij}} = p_i$, $\sum_{l=1}^{L} \frac{K_i\left(p_{l-1}^{ij}\right)^{\theta_i-1} - K_i\left(p_l^{ij}\right)^{\theta_i-1}}{1-F_{\frac{p}{\tau}}^i\left(p_{l-1}^{ij}\right)} = 0$. 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 $j$; if there are a sufficient number of regions in the world, however, these indirect effects can be safely ignored.

do so in the following lemma:

**Lemma 2** *Define* $\tilde{p}^{\max} \equiv \max_{j\in\{1,...,N\}} D_j\left(\frac{\theta_j}{\theta_j-1}A_jM_j\right)$ *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* $j \in \{1,..,N\}$ *and* $\mathbf{p} \in \left\{[0,b]^N \,|p_j = b\right\}$, $D_j\left(r_j\left(\mathbf{p}\right)\right) < b$.

**Proof.** I prove the lemma by contradiction. Suppose not, i.e. there exists a $j \in \{1,..,N\}$ and $\mathbf{p} \in \left\{[0,b]^N \,|p_j = b\right\}$ such that $D_j\left(r_j\left(\mathbf{p}\right)\right) \geq b$. By the definition of $\tilde{p}^{\max}$, we have:

$$D_j\left(r_j\left(\mathbf{p}\right)\right) \geq b > \tilde{p}^{\max} \geq D_j\left(\frac{\theta_j}{\theta_j-1}A_jM_j\right) \Rightarrow D_j\left(r_j\left(\mathbf{p}\right)\right) > D_j\left(\frac{\theta_j}{\theta_j-1}A_jM_j\right). \tag{32}$$

Since $D_j\left(\cdot\right)$ is a strictly decreasing function, equation (32) implies:

$$r_j\left(\mathbf{p}\right) < \frac{\theta_j}{\theta_j-1}A_jM_j. \tag{33}$$

Substituting $r_j\left(\mathbf{p}\right) = \frac{\theta_j}{\theta_j-1}A_jM_j\left(1-\left(\frac{K_j(p_j)}{f_j}\right)^{\theta_j-1}\right) + I_j\left(\mathbf{p}\right)$ into equation (33) yields:

$$I_j\left(\mathbf{p}\right) < \frac{\theta_j}{\theta_j-1}A_jM_j^{\theta_j-1}\left(\frac{K_j\left(p_j\right)}{f_j}\right)^{\theta_j-1}. \tag{34}$$

Equation (34) is intuitive: for $j$ to have a price at least as great as $b$, it must be the 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 $p_i \in [0,b]$ $\forall i \in \{1,...,N\}$, $p_j \geq \frac{p_i}{\tau_{ji}}$ $\forall i \in \{1,...,N\}$, i.e. 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 (3) this implies that $K_j\left(p_j\right) = 0$, i.e. 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, i.e. $I_j\left(\mathbf{p}\right) < 0$. This, of course, is impossible since from equation (31) imports are non-negative, i.e. $I_j\left(\mathbf{p}\right) \geq 0$. As a result, there is contradiction. ■

**Step 4:** In this step, I show that Brouwer's fixed point theorem guarantees the existence of an equilibrium set of prices. From Step 2, $G\left(\mathbf{p}\right)$ is a continuous function and a vector of prices $\mathbf{p}^*$ is an equilibrium vector of prices if and only if $G\left(\mathbf{p}^*\right) = \mathbf{p}^*$. Let $b > \tilde{p}^{\max}$ be a scalar greater than the maximum autarkic price and define $\tilde{G} : [0,b]^N \to [0,b]^N \equiv [\min\{b, D_1\left(r_1\left(\mathbf{p}\right)\right)\},...,\min\{b, D_j\left(r_j\left(\mathbf{p}\right)\right)\},...,\min\{b, D_N\left(r_N\left(\mathbf{p}\right)\right)\}]$. Since $\tilde{G}$ is a continuous function on a compact and convex set, by Brouwer's fixed point theorem there exists a $\mathbf{p}^* \in [0,b]^N$ such that $\tilde{G}\left(\mathbf{p}^*\right) = \mathbf{p}^*$.

It remains to show that $G\left(\mathbf{p}^*\right) = \mathbf{p}^*$. Since $\tilde{G}\left(\mathbf{p}^*\right) = \mathbf{p}^*$, from the definition of $\tilde{G}$, we have $p_j^* = \min\left\{b, D_j\left(r_j\left(\mathbf{p}^*\right)\right)\right\}$ for all elements $j \in \{1,...,N\}$. From Lemma 2, It cannot be the case that $p_j^* = b$ since Lemma 2 guarantees $D_j\left(r_j\left(\mathbf{p}^*\right)\right) < b$. Hence it must be the case that $p_j^* = D_j\left(r_j\left(\mathbf{p}^*\right)\right)$ $\forall j \in \{1,...,N\}$, or equivalently, $G\left(\mathbf{p}^*\right) = \mathbf{p}^*$, as required.

### A.4.2 Uniqueness

The proof of uniqueness is comprised of four steps. In the first step, I characterize the excess demand function. In the second step, I define the concept of a connected trade network. In the third step, I show that normalizing the prices have no effect on trade flows. In the fourth step, I prove uniqueness by extending the proof presented on p.613 of Mas-Colell, Whinston, and Green (1995) to the case where each region does not necessarily trade with all other regions.

**Step 1:** Define $Z(\mathbf{p}): \mathbb{R}_+^N \to \mathbb{R}^N \equiv [D_1^{-1}(p_1) - r_1(\mathbf{p}), ..., D_n^{-1}(p_n) - r_n(\mathbf{p}), ..., D_N^{-1}(p_N) - r_N(\mathbf{p})]'$ to be the function that yields the excess demand in each region as a function of the prices in all regions, where $r_n(\mathbf{p})$ is defined in Appendix A.4.1 and $D_n^{-1}(p_n)$ is the consumer demand function, which is monotonically decreasing by the inverse function theorem. There are two important characteristics of $Z(\mathbf{p})$. First, a vector of prices $\mathbf{p}^*$ are equilibrium prices if and only if $Z(\mathbf{p}^*) = 0$, i.e. there is no excess demand in any region. This is immediately evident by applying the consumer demand function element by element to both sides of the equilibrium condition $G(\mathbf{p}^*) = \mathbf{p}^*$ from Appendix A.4.1. Second, from Appendix A.4.1, $\frac{\partial}{\partial p_j} r_i(\mathbf{p}) \leq 0$ for all $j \neq i$ with the equality strict if $j$ exports to $i$ or $i$ exports to $j$. From the definition of $Z(\mathbf{p})$, this implies that for the $n^{th}$ element of $Z(\mathbf{p})$, $Z_n(\mathbf{p})$, we have $\frac{\partial}{\partial p_j} Z_n(\mathbf{p}) \geq 0$ for all $j \neq n$ with the equality strict if trade between $i$ and $j$ (in either direction) occurs.

**Step 2:** In this step, I define the concept of a connected trade network. A group of regions $G$ is connected if, for any two regions $i \in G$ and $j \in G$, there exists a set of regions (which I refer to as a "path") $K_{ij} \equiv \{k_1^{ij}, ..., k_P^{ij}\}$ such that $K_{ij} \subseteq G$, $k_1^{ij} = i$, $k_P^{ij} = j$, and for all $n \in \{1, ..., P-1\}$, region $k_n^{ij}$ either exports or imports from region $k_{n+1}^{ij}$.

**Step 3:** In this step, I show that normalizing the prices by a scalar does not affect the equilibrium. Formally, if $Z(\mathbf{p}) = 0$ and $\alpha \in \mathbb{R}_{++}$, then $Z(\alpha\mathbf{p}) = 0$, i.e. $Z(\cdot)$ is homogeneous of degree zero. To see this, note that from equation (3), $K(\cdot)$ is homogeneous of degree zero. Since scaling prices by $\alpha$ also entails scaling the fixed costs of search by alpha, equations (30) and (31) immediately imply that $r(\mathbf{p})$ is homogenous of degree zero. Since the demand function is also homogeneous of degree zero (since multiplying all prices by a constant along with wealth does not change the utility maximization problem), $Z(\cdot)$ is homogeneous of degree zero.

**Step 4:** In this step, I prove uniqueness. Formally, the statement to be proven is: if $Z(\mathbf{p}) = 0$, then for all $\mathbf{p}'$ such that $Z(\mathbf{p}') = 0$ and connected groups of regions $G$, there exists a scalar $\alpha \in \mathbb{R}_{++}$ such that $p_l = \alpha p_l'\ \forall l \in G$, i.e. the equilibrium prices within any connected group of regions are unique up to a normalization.

I prove this by contradiction. Suppose not. Then there exists a $\mathbf{p}'$ such that $Z(\mathbf{p}') = 0$ and there does not exist an $\alpha \in \mathbb{R}_{++}$ such that $p_l = \alpha p_l'\ \forall l \in G$. Let $i \equiv \arg\max_{j \in G} \left(\frac{p_j'}{p_j}\right)$ be the region with the greatest ratio of prices between $\mathbf{p}'$ and $\mathbf{p}$. Since $Z(\cdot)$ is homogeneous of degree zero, without loss of generality, I can normalize $\mathbf{p}'$ by $\frac{p_i'}{p_i}$, so that $p_i = p_i'$ and $p_j' \geq p_j$ for all $j \in G$. Furthermore, since $\mathbf{p}$ and $\mathbf{p}'$ are not collinear, there exists a region $l \in G$ such that $p_l' > p_l$.

Since the trade network is connected, there exists a path $K_{il} \subseteq G \equiv \{k_1^{il}, ..., k_P^{il}\}$ from region $i$ to region $l$. Consider the effect of a change in price from $\mathbf{p}$ to $\mathbf{p}'$ (element by element) on the excess demand in region $k_1^{il}$ (i.e. region $i$). Since $p_{k_1^{il}} = p'_{k_1^{il}}$, $p_j' \geq p_j$ for all $j \in \{1, ..., N\}$, and $\frac{\partial}{\partial p_j} Z_{k_1^{il}}(\mathbf{p}) \geq 0$ for all $j \neq k_1^{il}$, $Z_{k_1^{il}}(\mathbf{p}') \geq Z_{k_1^{il}}(\mathbf{p})$, i.e. the excess demand in region $k_1^{il}$ will not fall. Furthermore, since $p_j' \geq p_j$ for all $j \in \{1, ..., N\}$, either $p'_{k_2^{il}} > p_{k_2^{il}}$ or $p'_{k_2^{il}} = p_{k_2^{il}}$. If $p'_{k_2^{il}} > p_{k_2^{il}}$, then since $k_1^{ij}$ and $k_2^{ij}$ are trading partners, $\frac{\partial}{\partial p_{k_2^{ij}}} Z_{k_1^{il}}(\mathbf{p}) > 0$, so that $Z_{k_1^{il}}(\mathbf{p}') > Z_{k_1^{il}}(\mathbf{p})$, a contradiction. That is, if the price in region $k_2^{il}$ increases, the total quantity supplied to its trading partner $k_1^{il}$ will fall, increasing the excess demand in region $k_1^{il}$, contradicting the fact that the excess demand in region $k_1^{il}$ must remain at zero in equilibrium. Hence, it must be that $p'_{k_2^{il}} = p_{k_2^{il}}$. Consider now the effect of a change in price from $\mathbf{p}$ to $\mathbf{p}'$ (element by element) on the excess demand in region $k_2^{il}$. Since $p'_{k_2^{il}} = p_{k_2^{il}}$, from the exact same argument used for region $k_2^{il}$, it must be that $p'_{k_3^{il}} = p_{k_3^{il}}$, otherwise $Z_{k_2^{il}}(\mathbf{p}') > Z_{k_2^{il}}(\mathbf{p})$. Proceeding iteratively along the path $K_{il}$, this implies that $p'_{k_P^{il}} = p_{k_P^{il}}$. Since region $k_P^{il}$ is region $l$ by definition, this implies that $p'_l = p_l$, which is a contradiction, since $\mathbf{p}$ and $\mathbf{p}'$ are not collinear.

Intuitively, the reason that a connected network is required for there to be a unique set of prices is that in the absence of connected network, changes to the normalization of prices in one trading bloc (i.e. connected subset of the graph) will not necessarily induce trade to another unconnected trading bloc. Put another way, within any connected subset of the trade network, the equilibrium prices are unique up to a normalization, but disconnected subsets of the trade network can have different normalizations.

## A.5 Limiting case when fixed costs approach zero

A standard complete information trade model assumes that consumers have infinite marginal utility at zero consumption and transportation costs satisfy the triangle inequality, i.e. $\tau_{ij} \leq \tau_{ik}\tau_{kj}\ \forall i, j, k$. The equilibrium of this model has the following two characteristics: 1) all producers sell to the destination with the highest price net of transportation costs; and 2) if region $i$ exports to region $j$, then $\frac{p_j}{p_i} = \tau_{ij}$.[62] I show that in the limit where $f_i$ approaches zero, the incomplete information model satisfies both characteristics.

One major difference is that in the presence of information frictions, arbitrage opportunities exist in equilibrium. In a standard trade model with iceberg transportation costs,[63] all producers sell their produce to the destination with the greatest price net of transportation costs. Since some producers always sell locally,[64] any exporting producers must be indifferent between selling locally and exporting. This implies the familiar no arbitrage equation: if $Q_{ij} > 0$, then $\frac{p_j}{p_i} = \tau_{ij}$, i.e. the price ratio is exactly equal to the transportation cost. With information frictions, however, equation (4) implies that the $\frac{p_j}{p_i} \geq \tau_{ij}$, i.e. the price ratio is at least as great as the transportation cost.

That all producers sell to the destination with the highest price net of transportation costs in the limit where $f_i$ approaches zero follows immediately from equation (2) since $\lim_{f_i \to 0} \varphi_i^*(p) = \lim_{f_i \to 0} \frac{f_i}{K_i(p)} = 0\ \forall p < p^{\max}$, i.e. for all prices less than $p^{\max}$, producers of all sizes will find it

[62] Since consumer preferences have infinite marginal utility at zero consumption, all regions must consume some rice for prices to remain finite. Since producers selling locally face no transportation costs and producers sell to the market with the highest price net of transportation costs, the triangle inequality guarantees that some producers always sell locally. See section A.5 in the appendix for details.

[63] By "standard" I mean that: (1) consumers have infinite marginal utility at zero consumption; and (2) transportation costs satisfy the triangle inequality, i.e. $\tau_{ij} \leq \tau_{ik}\tau_{kj}\ \forall i, j, k$.

[64] Since consumer preferences have infinite marginal utility at zero consumption, all regions must consume some rice for prices to remain finite. Since producers selling locally face no transportation costs and producers sell to the market with the highest price net of transportation costs, the triangle inequality guarantees that some producers always sell locally. See section A.5 in the appendix for details.

optimal to continue to search. Indeed, for equation (1) to be satisfied when $f_i = 0$, it must be the case that $K_i(p) = 0$, which only occurs when $p = p^{\max}$, i.e. when the producer has found the region with the highest price net of transportation costs.

I show that the limiting case satisfies the second characteristic by contradiction. Suppose that region $i$ exports to region $j$ but $\frac{p_j}{p_i} \neq \tau_{ij}$. Since producers sell to the destination with the highest price net of transportation costs, it must be that $\frac{p_j}{\tau_{ij}} > p_i$ and no producers in region $i$ sell locally. By the triangle inequality, this implies that no producers sell to region $i$.[65] Since consumers have infinite marginal utility at zero consumption, however, this implies that the price in region $i$ is infinitely high, which contradicts the fact that $\frac{p_j}{\tau_{ij}} > p_i$. Hence, it must be the case that if region $i$ exports to region $j$, then $\frac{p_j}{p_i} = \tau_{ij}$.

# B Data Description

This section describes in detail the data used in Section 3 through Section 6.

## 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 of which more than 99% of both the value and 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.[66] For every commodity (classified at the SITC 5-digit level), this 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 the create a data set of province-province bilateral trade flows. To create a measure of the bilateral freight costs in the standard iceberg cost form, I calculated the mean (non-missing) freight cost as a fraction of the total value of shipments of a commodity in a origin-destination province pair in a particular year.[67]

I then identified 51 agricultural commodities using the SITC classification codes, which constitute the sample of analysis for Figure 2 . Of these commodities, 10 could be matched

[65]To see this, suppose that there exists some region $k$ that exports to region $i$. Since exporters in region $k$ only export to the destination with the highest price net of transportation costs, it must be the case that $\frac{p_i}{\tau_{ki}} \geq \max_l \frac{p_l}{\tau_{kl}}$. Since $\frac{p_j}{\tau_{ij}} > p_i$, this implies that $\frac{p_j}{\tau_{ki}\tau_{ij}} > \frac{p_i}{\tau_{ki}}$. Since the triangle inequality implies that $\tau_{ki}\tau_{ij} \geq \tau_{kj}$, this means that $\frac{p_j}{\tau_{kj}} > \frac{p_i}{\tau_{ki}}$, which is a contradiction.

[66]In an apparent error, the data also included information on trade flows in the 4th quarter alone, which I used to construct Figure 2.

[67]Unfortunately, the freight data is missing for a large (38 percent) 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 indicative of the magnitude of overall transportation costs.

with wholesale price and production data. Statistics about these 10 commodities are presented in Table 1. As is evident, these commodities constitute a large majority of the Philippines agricultural sector, comprising 65 percent the total value of agricultural output and 60 percent the total agricultural area.[68] These commodities constitute the primary sample of analysis.

In total, I observe 4,337 non-zero province origin-province destination-year-commodity trade flows (of 32,922 potential trade pairs) spanning 1995 and 2009 where the wholesale price is observed in both the origin and destination province and production data is available for the origin province.[69] 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 47 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-year 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 traders will receive when exporting 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 survey 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 $3 - 5$ 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-year level 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 non-subsistence farmers

[68]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).

[69]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.

(defined as farmers with more than 1 hectare of land) using the 1991 agricultural census, where the fraction is calculated at the commodity-province level.

## B.4 Household income and expenditure

Data on expenditure on agricultural commodities used in Section 3.4 comes from the Family Income and Expenditure Survey (FIES). The FIES is administered by the Philippines National Statistics Office (NSO) every three years to a random sample of households throughout the Philippines. I construct the expenditure on agricultural products in province $i$ in year $t$, $X_{it}$, by taking the mean observed total food expenditures in each province in each survey year (1994, 1997, 2000, 2003, and 2006) and multiplying it by the population of that province. For years when the survey was not administered, I estimate the expenditure by extrapolating from the observed expenditure using a province-specific quadratic time trend.

For each household, the FIES also reports whether the primary source of income was from agricultural activities. I use this fraction to determine the relative number of farmers and consumers in each province (the absolute number of farmers are calculated from production data given the landholding distribution).

## B.5 Rainfall

The Philippine Atmospheric, Geophysical & Astronomical Services Administration (PAGASA) collects daily rainfall data from 47 weather stations located throughout the Philippines (the location of the stations is presented in the online appendix). Using this data, I constructed provincial level daily rainfall measures using an distance weighting technique suggested by Dirks, Hay, Stow, and Harris (1998), i.e. $r_{it} = \sum_s w_{is} r_{st}$, 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 $\alpha$ 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.

After constructing daily province-level rainfall measures, I created variables measuring the mean and standard deviation of rainfall for each month in each year. These monthly measures of rainfall, appropriately de-meaned, provided the instruments to isolate idiosyncratic price shocks in Sections 3.3 and 3.4.

## B.6 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 this data, I construct a measure of the number of civilian cell phone towers in operation in every province in every month between 1990 to 2009.[70]

## B.7 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 hectares 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 observe the total land area under his/her cultivation used to cultivate that commodity $L_{fic}$. To estimate the Pareto distribution shape parameter $\theta_{ic}$, I first restrict the sample to only those farmers cultivated at least one hectare of land in order to exclude subsistence farmers who do not sell their produce to the market.[71] I then define $\theta_{ic}$ to parameter that maximizes the likelihood of observing $\{L_{fic}\}_{f=1}^{F}$ under a Pareto distribution:

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

Hence, $\theta_{ic}$ 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 the online appendix for a map of the distribution of shape parameters). In the online appendix, I also present a figure that depicts the relationship between the observed landholding distribution and the Pareto distribution with estimated $\theta_{ic}$ for two provinces.[72] 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 $\theta_{ic}$ (in this case, for Bohol) is associated with a greater concentration of land amongst smaller farmers.

## B.8 Farmer sales

Data on individual farmer-trader transactions used in Section 6 comes 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

[70] I focus only on 900 MHz, 1800 MHz, 2G, and 3G towers, as these are the frequencies used for standard GSM mobile phones. Towers broadcasting at other frequencies are used primarily for non-civilian purposes (e.g. military, ship communications, etc).

[71] 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.

[72] Other provinces generate similar figures.

five producing municipalities in the province are interviewed about the quantity of the commodity they sold in the past month, they 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.[73] These data include information on roughly 2.3 million unique farmer-trader transactions in roughly 134,000 province-commodity-year-month markets.

## B.9 United States Data

Bilateral state commodity flows come from the 2007 Commodity Flow Survey (http://www.bts.gov/). The publicly available data is at the 2-digit level, allowing the comparison of only two commodities: corn grain (corresponding to category "Cereal grains") and hay (corresponding to category "Animal feed and products of animal origin, nec"). Annual prices of corn grain and hay come from the National Agricultural Statistics Service (http://www.nass.usda.gov/). Monthly rainfall data at the weather station level comes from the Global Historical Climatology Network (http://www.ncdc.noaa.gov/ghcnm/), which is aggregated to the state level using the inverse distance weighting formula discussed in Section B.5.

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