The Price of Distance: Producer Heterogeneity, Pricing to Market, and Geographic Barriers*

Kazuko Kano,[†] Takashi Kano,[‡] and Kazutaka Takechi[§]

May 18, 2012

*Preliminary draft. We would like to thank Andrew Bernard for useful discussions and a suggested title. We are also grateful to Alexandre Skiba for helpful discussions, and to seminar participants at Hosei University for insightful comments. Financial support for this research was provided by a Grant-in-Aid for Scientific Research (No. 24530270) and Seimeikai Foundation.

[†]Graduate School of Economics The University of Tokyo 7-3-1 Hongo, Bunkyo-ku Tokyo, 113-0033, Japan. Email: kkano@e.u-tokyo.ac.jp

[‡]Graduate School of Economics Hitotsubashi University 2-1 Naka, Kunitachi Tokyo, 186-8601, Japan. Email: tkano@econ.hit-u.ac.jp

[§]Faculty of Economics, Hosei University, 4342 Aihara-machi, Machida-shi, Tokyo, 194-0298, Japan. Email: ktakechi@hosei.ac.jp

Abstract

This study investigates the distance effect on price differences across regions. In order to identify the distance effect, we need to incorporate producer heterogeneity and pricing-to-market behavior. Because geographic barriers alter the threshold levels of productivity to set a positive price across markets, the effect of distance on price differentials can be underestimated without accounting for heterogeneity and pricing to market. By incorporating these factors, empirical analysis using micro-level data reveals that the distance effect is found to be significantly large, suggesting that the price of geographical barriers is still high for regional transportation.

Keywords: law of one price; transportation costs; geographic barriers; producer heterogeneity; pricing to market.

JEL classification code: F11; F14; F41.

1 Introduction

The distance effect on price differences across regions is a key factor for understanding the significance of geographic barriers. Geographical separation creates price differences across regions even without any institutional differences (such as tariffs, tax differences, and national borders). The previous law of one price (LOP) studies report a small effect of distance (the distance elasticity parameter is normally less than 0.01), while the distance effect is considered to include trade costs other than transportation costs (Engel and Rogers 1996; Parsley and Wei 1996; Parsley and Wei 2001; Crucini, Shintani, and Tsuruga 2010). Hence, it is not clear why the distance effect is so small and to what extent geographical barriers (transportation costs) explain price differentials. Because the distance elasticity of transportation costs is a key parameter when assessing the impact of geographical barriers, the trade literature, such as Hummels (2007), Crozet and Koening (2010), and Balistreri, Hillberry, and Rutherford (2011), has tried to identify and estimate this parameter (the distance elasticity is more than 0.15). However, the price differential effect of geographic barriers (distance) has not been examined extensively.

This study addresses the issue of measuring the impact of transportation costs using price differential data. In order to measure transportation costs properly using price data, as Anderson and van Wincoop (2004) point out, the difference between market prices and prices in production place must be used, not just the two market prices. In addition, because when distance gets longer, it will not only increase price differentials, but also decrease product delivery propensity, distance causes selection biases. Thus, delivery choice to other regions (export decision) should be accounted for to control for sample selection biases as in Helpman, Meliz, and Rubinstein (2008). The recent studies, Donaldson (2010) and Kano, Kano, and Takechi (2010) (hereafter KKT (2010)), follow Anderson and van Wincoop's (2004) suggestion of using the price in the source region of production. Donaldson (2010) identifies the source region of salt production in India and utilizes the information to measure the transportation costs using market prices. KKT (2010) use agricultural wholesale price data in Japan, where the source and market prices are available. Furthermore, KKT (2010) propose an estimation procedure to take into account selectivity bias following Helpman et al. (2008). KKT (2010) demonstrate that without controlling for this bias, the distance effect is quite weak. However, after controlling for these problems, there is a significant impact of distance on price differences.

While these studies reveal the important bias in the distance effect estimations, there still remain the possible causes of biases: producer heterogeneity and pricing to market. Because producer heterogeneity and pricing-to-market behavior cause different pricing across markets, price differentials may not simply represent transportation costs. In KKT (2010), while markets are monopolistically competitive, producers set invariant markups and there is no producer heterogeneity. Whereas Donaldson (2010) uses the Eaton and Kortum (2002) model in which producer productivity is dispersed, the market is perfectly competitive. Therefore, different pricing behavior across markets is not taken into account.

In this paper, we show that the estimates of the distance effect are underbiased if producer heterogeneity and pricing to market are ignored. We incorporate producer heterogeneity and nonhomothetic preferences in a monopolistic competition model. KKT (2010) use a CES utility function and monopolistic competition, and the price difference is a function of transportation cost only. On the other hand, in a nonhomothetic preference framework, because an individual firm's pricing depends on local market characteristics (as shown by, for example, Melitz and Ottaviano, 2008), price differences do not simply indicate transportation costs, but include market structure (the number of products) and the productivity threshold value. Because transportation costs reduce profitability in a remote market, the productivity threshold level to set a positive price depends on transportation costs. In particular, as the productivity threshold increases, only highly productive and thus low-price-setting firms produce. Hence, ignoring producer heterogeneity creates omitted variable bias, which causes underestimation of the distance effect. Thus, we contribute to the literature by estimating the distance effect while controlling for heterogeneity and pricing to market.

By estimating the price difference equation with sample selection, producer heterogeneity, and pricing-to-market behavior taken into account, a large distance effect on price differentials is found. In the previous literature, using information on source regions, Donaldson (2010) and KKT (2010) find a significant and reasonable distance effect coefficient: in Donaldson (2010) it is 0.24 and in KKT (2010) it is 0.21 to 0.325. In this study, the coefficients of distance effect range from 0.973 to 1.301. These estimates seem to be large, but are consistent with the results in the economic geography literature. In particular, large distance effects are found when investigating truck transportation. Because truck transportation is a major transport type in our analysis, the results are close to those of Combes and Lafourcade (2005), who use data on trade shipped by truck and estimate the distance elasticity to be 0.8. This estimate may suffer from selection bias, suggesting that our findings are corroborated by the fact of truck transportation and the underbias of the distance elasticity caused by self-selection. Therefore, we conclude that there is a substantially large bias without incorporating producer heterogeneity and pricing-to-market behavior. The price of geographic barriers (distance) is still high for regional transportation, even in a country with highly developed transportation infrastructure such as Japan.

The introduction of nonhomothetic preferences is essential to investigate the distance effect on price differentials with producer heterogeneity. If a CES utility function is used and thus monopolistically competitive firms set constant markup prices, the heterogeneity term will be cancelled out in the price difference equation and the price difference depends on transportation costs only. If the focus is not on price differences, then important implications are obtained for price levels under firm heterogeneity using CES, as Ghironi and Melitz (2005) and Bergin, Glick, and Taylor (2006) show the emergence of Ballasa–Samuelson effects. Here, because we study price differentials, there is no room for producer heterogeneity in a standard CES framework. Nonhomothetic preferences lead firms to set different prices across markets, and these prices depend on a heterogeneous threshold, so heterogeneity plays an important role here.

Even in a CES utility framework, pricing-to-market behavior may be taken into account. If the utility function parameter is different across markets, price differences are a function of not only transportation costs, but also the elasticity parameter in each market. Thus, the price in each market is different. However, these elasticity parameters are not directly related to transportation costs, therefore market-specific effects can control for these. In a nonhomothetic preference model, the thresholds are related to transportation costs, hence the distance effect will be biased unless it is controlled for. We estimate both nonhomothetic and CES models to capture the magnitude of the bias.

This paper is organized as follows. In Section 2, we develop our nonhomothetic preference model with producer heterogeneity. For comparison, we also develop a CES model. Then, in Section 3 the empirical framework is derived and Section 4 reports the estimation results. The final section concludes.

2 Model

In this section, we develop a model of pricing and delivery pattern. Consumers purchase a variety of products delivered from their own and other regions. Each product is produced by a single producer. These producers are heterogeneous in terms of productivity and engage in monopolistic competition. Because one of main purposes in this paper is to demonstrate the differences between nonhomothetic and CES preference cases, we first introduce the nonhomothetic model. Then, we consider a CES utility model for comparison.

2.1 Consumers

Consumer preferences are expressed by a nonhomothetic utility function. Nonhomothetic preferences have been introduced to account for pricing to market (Melitz and Ottaviano, 2008; Simonovska, 2010).

We employ the Simonovska (2010) framework, which is suitable for our analysis, because it enable us to compare with the CES model easily. Our derivations rely on Simonovska (2010), in which she focuses on trade volumes and price levels, while ours is on individual pricing across markets.[1]

Consumer nonhomothetic preferences in region i are expressed by:

$$u_i = \int_{\omega \in \Omega} \ln(q_i(\omega) + \bar{q}) d\omega, \tag{1}$$

where ω is a variety index, Ω is the set of products available in market i, and $q_i(\omega)$ is consumption of variety ω. The presence of $\bar{q}$ makes these preferences nonhomothetic. If $\bar{q} = 0$, the utility function is a typical homothetic function. The size of $\bar{q}$ can be changed, so this can be normalized to 1 as in Young (1991). Each consumer is assumed to supply one unit of labor. Thus, income is equal to wage, w_i. The budget constraint is:

$$w_i = \int_{\omega \in \Omega} p_i(\omega) q_i(\omega) d\omega. \tag{2}$$

Then, from utility maximization, the demand function is obtained by:

$$q_i(\omega) = \frac{w_i + \bar{q} P_i}{N_i p_i(\omega)} - \bar{q}, \tag{3}$$

where $P_i = \int_{\omega \in \Omega} p_i(\omega)$ is the price index and $N_i = \int_{\omega \in \Omega} d\omega$ is the number of products in market i. This demand function has regular characteristics such that demand is decreasing in prices and increasing in income (wage). If the number of products supplied to this market rises, the demand for each product will fall. This will in turn affect pricing behavior of producers.

2.2 Producers

Consider a producer located in region j. The number of potential producers is assumed to be fixed, so firms decide whether to produce or shut down. This is similar to the short-run equilibrium case in Melitz

[1] Simonovska (2010) demonstrates how the nonhomothetic model works in general equilibrium and also compares it with the CES model.

and Ottaviano (2008). The timing of delivery decision is made as follows. Producer productivity, ϕ is assumed to follow a random distribution, $G(\phi)$. Producers incur a fixed cost before their productivity level is realized and based on realized productivity, they decide whether to deliver and set their optimal prices. This enables us to establish a similar delivery choice decision problem as in the CES case.

The producer profit maximization problem is to maximize profits, π_{ij}:

$$\max_{p_{ij}} \pi_{ij} = p_{ij}q_{ij} - \frac{\tau_{ij}w_j}{\phi}q_{ij}, \tag{4}$$

where p_{ij} is the price in region i from region j, q_{ij} is the quantity sold in region i from region j, and τ_{ij} is the iceberg-type transportation cost, $\tau_{ij} > 1$ for $i \neq j$ and $\tau_{ij} = 1$ for $i = j$. Because labor is assumed to be the only input, the wage rate, w_j, indicates unit cost and ϕ is a measure of productivity. This productivity parameter differs among producers (firm heterogeneity). Because each product is produced by a single producer, the number of varieties is equal to the number of producers. The optimal price set by a producer with productivity ϕ is:

$$p_{ij}(\phi) = (\frac{\tau_{ij}w_j(w_i + \bar{q}P_i)}{\phi N_i \bar{q}})^{1/2}. \tag{5}$$

The optimal price depends on not only transportation costs, but also local market characteristics. If income in markets (w_i) is high, producers can charge high prices. The existence of the large number of competitors means large Ni, which induces low prices because of tough competition.

Unlike the CES preference case, if the price is sufficiently high, demand will be zero. Then, the profit for the firm in region j from supplying this product to region i will also be zero. We denote the productivity of this firm ϕ^*_{ij}. Then, this threshold value is expressed by:

$$\phi^*_{ij} = \frac{\tau_{ij}w_jN_i\bar{q}}{w_i + \bar{q}P_i}. \tag{6}$$

The threshold value, ϕ^*_{ij}, is increasing in transportation costs, τ_{ij}: only high productive firms can overcome trade barriers. This property holds in the CES preference case too. On the other hand, market structure measured by the number of firms, N_i, influences the threshold value, while it has no effect in the CES case. This is caused by variable markups in the nonhomothetic model.

Furthermore, in the CES model, firms charge a constant markup over marginal cost. The optimal price in the nonhomothetic case depends on market structure through ϕ^*_{ij}, which means that productivity threshold itself matters for each producer's price. In other words, aggregate producer characteristics affect individual pricing behavior. Thus, this requires us to account for heterogeneity and pricing to market to identify transportation costs using regional price differential data. Because of the assumption of monopolistic competition, the product index can be expressed by the producer's productivity measure: $P_i = \sum_\nu \int_{\phi^*_{i\nu}} p_{i\nu}(\phi)\mu(\phi)d\phi$ and $N_i = \sum_\nu N_{i\nu} = \sum_\nu \int_{\phi^*_{i\nu}} \mu(\phi)d\phi$, where μ is a conditional density function of ϕ conditional on delivery.

The relationship between the optimal price and the threshold value in this case is similar to that in the Melitz and Ottaviano (2008) case. Melitz and Ottaviano (2008) use a quadratic utility function and show how market size affects the key features in a model with firm heterogeneity. The optimal price is increasing in the threshold level of productivity and the number of firms is negatively related to the threshold value. Thus, many properties derived here are shared in the nonhomothetic models.

Assuming that productivity follows a Pareto distribution ($G(\phi) = 1 - b^\theta/\phi^\theta$, $\theta > 0$), the expected profit will be:

$$E\pi_{ij} = (1 - G(\phi^*_{ij})) \int \pi_{ij}\mu d\phi, \tag{7}$$

where $\mu = g/(1 - G(\phi^*)) = \phi^{*\theta}/\phi^{\theta+1}$. This is the conditional density where the productivity is above ϕ^*_{ij}. Then, expected profit is calculated as follows:

$$(1 - G(\phi^*_{ij})) \int \pi_{ij}\mu d\phi = \frac{b^\theta \tau_{ij} w_i \overline{q}}{(2\theta + 1)(\theta + 1)\phi^{*\theta+1}}. \tag{8}$$

Firms decide whether to deliver their product to region i depending on the above profit measure and fixed entry costs. This captures the self-selection problem in delivery patterns. The productivity threshold, ϕ^*, affects pricing behavior and delivery choice. In our setting, even though productivity is higher than the threshold level, such a firm still may not deliver their product because of negative expected profits. Thus, this threshold parameter does not directly separate producers who deliver and those who do not. Rather, the threshold directly influences prices across markets.

2.3 CES Case

We intend to compare our results with that of the CES utility function case. As we will see, the same implications for price differentials are derived with or without producer heterogeneity. Thus, we consider the CES model without heterogeneity to compare with the results in KKT (2010).

We briefly specify a consumer's preferences by a simple CES model as follows:

$$u_i = [\int_{\omega\in\Omega} x_i(\omega)^{\alpha} d\omega]^{1/\alpha}.$$

Then, maximizing this utility with the budget constraint ($w_i = \int p_i(\omega)q_i(\omega)d\omega$) yields the demand function:

$$x_i = \frac{p_i(\omega)^{-\epsilon}}{P_i^{1-\epsilon}} w_i,$$

where ϵ is the elasticity of substitution, $\epsilon = 1/(1-\alpha)$, and $P_i = [\int_{\omega\in\Omega} p_i(\omega)^{1-\epsilon} d\omega]^{1/(1-\epsilon)}$.

We consider a homogeneous firm in a monopolistically competitive market. The firm's profits are:

$$\pi_{ij} = p_j(\omega)x_j(\omega) - \frac{\tau_{ij} w_j x_j(\omega)}{\phi}.$$

Then, by profit maximization, the optimal price is obtained using constant markup pricing as follows:

$$p_{ij}(\phi) = \frac{\tau_{ij} w_j}{\phi\alpha}.$$

Substituting this into the profit function yields:

$$\pi_{ij}(\phi) = (1-\alpha)(\frac{\tau_{ij} w_j}{\alpha P_i \phi})^{1-\epsilon} w_i.$$

Because firms are assumed to be homogeneous, their decision to deliver does not depend on a randomly selected level of productivity. The choice is based on the comparison of profits and fixed cost of delivery. If $\pi_{ij}/f_{ij} > 1$, then firms in region j will deliver their products to region i. Thus, similar to the heterogeneous firm case, delivery data are truncated because of self-selection by the producers.

This break-even productivity level depends on transportation costs. If transportation costs, τ_{ij}, are high, firms that are productive enough are able to make positive profits: ϕ_{ij} is increasing in τ_{ij}. However, as we mentioned, market structure does not affect ϕ_{ij} directly, but only through the price index, P_i.

2.4 Price Differentials

Our approach of taking the difference between the price in markets and source regions allows us to measure transportation costs in an appropriate way. Because retail prices do not have information about source, taking the difference between two market prices does not necessarily measure transportation costs. However, if the source price and the market price with information about source are available, the difference between these prices captures the costs of transportation.

Using the optimal prices set by firms, the price differential between market and source is:

$$p_{ij}/p_{ii} = \tau_{ij}{\phi_{ii}^{*}}^{1/2}/{\phi_{ij}^{*}}^{1/2}. \tag{9}$$

Because the threshold value, ϕ_{ij}^{*}, depends on transportation costs, ignoring producer heterogeneity causes biases in identifying the relationship between the price differential and transportation costs. If τ_{ij} increases, ϕ_{ij}^{*} will increase. Because ϕ_{ii}^{*} does not depend on τ_{ij}, a larger ϕ_{ij}^{*} induces a smaller price differential. Thus, heterogeneity reduces the price differential. This omitting variable bias may cause underestimation of the effect of transportation costs.

In addition, ϕ_{ij}^{*} also depends on the number of firms, N_i. This is a function of the threshold value itself and thus affected by transportation costs. Hence, the changes in τ_{ij} are associated with the changes in market structure. This implies that market prices are charged depending on market structure and therefore the number of firms across markets is a determinant of price differentials. Without controlling for this type of pricing-to-market behavior, the estimates of transportation costs will be biased.

In a CES utility framework, the price difference is:

$$p_{ij}/p_{ii} = \tau_{ij}.$$

As we mentioned, one of our objectives in this paper is to highlight the changes by incorporating firm heterogeneity. As a matter of fact, this equation holds with and without producer heterogeneity. This is because even if firms' productivity is heterogeneous, optimal pricing does not depend on the threshold value of productivity, which is a key factor of heterogeneity. Besides, each firm's productivity is cancelled out when considering price differentials. Hence, producer heterogeneity does not play an important role in the link between price differences and transportation costs in a CES model. Producer heterogeneity matters for the link between price differences and distance not when preferences are CES, but when they are nonhomothetic. If we introduce nonhomothetic preferences, firms set variable markups across markets as the optimal prices and thus we deal with pricing-to-market behavior. Therefore, the bias caused by producer heterogeneity is indispensable for pricing to market.

With regards to market characteristics, compared with the nonhomothetic case, price differentials are independent of these characteristics. This is because in the CES case, again, productivity threshold level does not affect individual pricing. The thresholds are derived from the zero profit conditions, which determines the selection of producers who deliver, not prices. As a result, when obtaining price differentials, market characteristic variables are cancelled out. Hence, market characteristics such as income and the number of competitors can be taken into account in the nonhomothetic case only.

As mentioned in the Introduction, even in a CES model, pricing-to-market behavior can be taken into account by introducing demand parameter heterogeneity across markets. Namely, if the elasticity of substitution parameter is different among markets, the price difference will be:

$$p_{ij}/p_{ii} = \tau_{ij}(\alpha_i/\alpha_j),$$

where α_i and α_j are demand parameters in the CES utility function. However, these can be controlled for by using region-specific effects and have nothing to do with transportation cost, τ_{ij}. Hence, this does not cause any biases in the estimation of the distance effect.

By using the formula for the threshold value in the nonhomothetic model, ϕ^*_{ij}, we are able to express the price differential as follows:

$$p_{ij}/p_{ii} = \tau_{ij}^{1/2}\frac{(w_j + \bar{q}P_j)^{1/2}}{(w_i + \bar{q}P_i)^{1/2}}(\frac{N_i}{N_j})^{1/2}. \tag{10}$$

The heterogeneity effect reduces the direct impact of transportation costs from τ_{ij} to $\tau_{ij}^{1/2}$ in our nonhomothetic specification. In general, the transportation costs effect will be also weakened. This is because the effect of a transportation cost increase on price differentials is mitigated by the producer selection. With high transportation costs, only high productive firms are able to ship their products. Such firms set the prices at the low level. Thus, the further markets are apart, the lower is the magnitude of the increase in prices. This mechanism creates the underbias of distance elasticity using the price differential data only.

This selection mechanism is in effect at the individual pricing level. This mechnism also influences average price changes associated with general productivity shocks, as Ghironi and Melitz (2005) and Atkeson and Burstein (2008) show. If only high productive firms can export due to negative shocks, then because they charge the price at the low level, the average prices will be low. If free entry is assumed, firm exit because of negative shocks will cause labor demand decrease and thus labor costs decrease. This induces that low productive firms can export, implying the increase in the average export price. Thus, depending on the assumption of entry conditions, the average prices either increases or decreases.

In our study, because we do not consider free entry, negative shocks will decrease individual prices set at a market.

The price differentials are also affected by source and market characteristics and market structure affect price differences. If these factors are correlated with transportation costs, omitted variable biases occur. Taking the log of the above equation yields:

$$\ln p_{ij} - \ln p_{ii} = (1/2)\ln \tau_{ij} + (1/2)\ln N_i - (1/2)\ln N_j + (1/2)\ln(w_j + \bar{q}P_j) - (1/2)\ln(w_i + \bar{q}P_i). \quad (11)$$

The price differential depends on not only transportation costs, but also market characteristics, such as the number of products and price indices. This property reflects the behavior of pricing to market. Because the optimal price depends on local market characteristics, the price differentials reflect market structure. To be able to capture this element in the nonhomothetic model is an advantage over the CES framework.

So far, we have not imposed any functional form on transportation costs. We adopt a conventional specification:

$$\tau_{ij} = D_{ij}^{\gamma} e^{\mu + u_{ij}},$$

where D_{ij} is the distance between two regions. That is, if $\gamma > 0$, as distance increases, transportation costs increase. The constant term μ corresponds to the uniform transportation costs component and u_{ij} is unobservable transportation costs, $u_{ij} \sim N(0, \sigma_u)$. The log form is:

$$\ln \tau_{ij} = \gamma \ln D_{ij} + \mu + u_{ij}.$$

The estimation of distance elasticity, γ, is our main parameter. It is important to account for delivery choice, producer heterogeneity, and pricing to market to identify this.

We make one remark on the threshold value used here (ϕ_{ij}^*). Even if $\phi_{ij}^* < \phi$, the producer with productivity ϕ may not deliver its product because demand is too low to cover fixed costs. Thus, producer heterogeneity (threshold value ϕ^*) matters mainly for the individual pricing decision, not the delivery decision. We take the delivery decision into account by considering the sample selection problem caused by the positive profit condition.

2.5 Delivery Choice

The price differential is observed only when there is actual delivery. Thus, there will be data truncation problem. The delivery choice is made based on profitability, so we consider producer's decision of delivery. Because producers pay f_{ij}, the delivery decision is summarized by the variable, Z_{ij}:

$$Z_{ij} = \frac{\frac{b^{\theta}\tau_{ij}w_i\bar{q}}{(2\theta+1)(\theta+1){\phi^*_{ij}}^{\theta+1}}}{f_{ij}}.$$

Thus, if Z_{ij} is greater than 1, firms in region j choose to deliver the product in region i. By taking logs, we have the following delivery choice equation.

$$\ln Z_{ij} = \theta \ln b + \ln \tau_{ij} + \ln w_i + \ln \bar{q} - \ln(2\theta + 1)(\theta + 1) - (\theta + 1)\ln \phi^* - \ln f_{ij}$$

$$= \theta \ln b - \theta \ln \tau_{ij} - \theta \ln w_i - \theta \ln \bar{q} - \ln(2\theta + 1)(\theta + 1) - (\theta + 1)\ln N_j + (\theta + 1)\ln(w_j + \bar{q}P_j) - \ln f_{ij}.$$

If $\ln Z_{ij} > 0$, then there will be delivery from j to i. Because the price differential is observed only when $\ln Z_{ij} > 0$, we take into account this selectivity bias to estimate the price difference equation. We jointly estimate the price differential and delivery choice equations.

Similarly, in the CES framework, the delivery choice is expressed by Z_{ij}:

$$Z_{ij} = \frac{(1-\alpha)[\frac{\tau_{ij}w_j}{\alpha P_i \phi}]^{1-\epsilon} w_i}{f_{ij}}.$$

Thus, taking logs yields a similar expression for delivery choice:

$$\ln Z_{ij} = z_{ij} = \ln(1-\alpha) + (1-\epsilon)\ln \tau_{ij} + (1-\epsilon)\ln w_j - (1-\epsilon)\ln \alpha - (1-\epsilon)\ln P_i - (1-\epsilon)\ln \phi + \ln w_i - \ln f_{ij}.$$

Unlike Helpman et al. (2008), our focus is on individual firm's choice of prices, not on trade volume. Thus, it is not necessary to control for the effect of heterogeneity on aggregate variables. What we need to account for are the impact of heterogeneity on individual firm's pricing across markets and its delivery choice according to this selection mechanism.

Similar to the nonhomothetic preference case, we estimate the price differential equation with this selection bias accounted for. While the KKT (2010) model is estimated using an instrumental variable approach, controlling for selection bias is based on this delivery choice equation. Thus, in this paper we simply estimate the price differential and delivery choice equations by maximum likelihood. We use regional dummies to control for market specific effects such as price indices (Anderson and van Wincoop, 2003; Helpman et al., 2008).

2.6 Empirical Specifications

For estimation, we need to parameterize the price differential and delivery choice equations. As in Helpman et al. (2008), fixed costs have the following specification: $f_{ij} = \exp(\lambda_i + \lambda_j - \nu_{ij})$. The estimating equations are expressed as follows:

$$\begin{aligned} z_{ij} =& -\ln f_{ij} + \theta(\ln b - \overline{q}) - \theta\mu - \theta u_{ij} - \ln(2\theta+1)(\theta+1) \\ & - \theta\gamma \ln D_{ij} - \theta \ln w_i - (\theta+1)\ln N_j + (\theta+1)\ln(w_j + \overline{q}P_j) \\ =& c_0 + \theta c_1 - \theta\gamma \ln D_{ij} - \theta \ln w_i - (\theta+1)\ln N_j + (\theta+1)c_2 dum_j + c_3 dum_i + \eta_{ij}, \end{aligned} \tag{12}$$

where $c_0 = -\theta\mu - \ln(2\theta+1)(\theta+1)$, $c_1 = \ln b - \overline{q}$, $\ln(w_j + \overline{q}P_j) - \lambda_j$ is a function of region j's specific effect, therefore $\ln(w_j + \overline{q}P_j) - \lambda_j = c_2 dum_j$, and dum_j is j's specific effect. The error term is $\eta_{ij} = -\theta u_{ij} + \nu_{ij} \sim N(0, \theta^2\sigma_u^2 + \sigma_\nu^2)$.

Similarly, the price difference equation is:

$$\begin{aligned} q_{ij} &= \ln p_{ij} - \ln p_{ii} \\ &= (1/2)\mu + (1/2)\gamma \ln D_{ij} + (1/2)\ln(1+N_i) - (1/2)\ln(1+N_j) + c_4 dum_j - c_5 dum_i + (1/2)u_{ij}, \end{aligned} \tag{13}$$

where dum_j controls for region-specific effects including wages and price indices as in the delivery choice equation.

As in KKT (2010), with regards to the identification of the distance elasticity, γ, the price difference and product delivery equations reveal an important result. Simply estimating the price difference equation only may lead to underestimation of γ. This is because the errors in these equations are correlated: because $\eta_{ij} = \theta u_{ij} + \nu_{ij}$, the error terms, η_{ij} and u_{ij}, are correlated. Taking the conditional expectation of q_{ij} yields: $E[q_{ij}|X] = (1/2)\mu + (1/2)\gamma \ln D_{ij} + (1/2)\ln(1 + N_i) - (1/2)\ln(1 + N_j) + c_6 dum_j - c_7 dum_i + (1/2)E[u_{ij}|X]$, where X is a vector of observables. Because $E[u_{ij}|X] = \rho\frac{\sigma_u}{\sigma_\eta}E[\eta_{ij}|X]$, if we ignore this correlation, there will be bias in the estimate of the distance effect. This bias term is expressed as an inverse Mill's ratio: $E[\eta_{ij}|X] = \phi(\hat{z}_{ij})/\Phi(\hat{z}_{ij})$. Hence, to obtain consistent estimates, we need to account for the correlation between the price difference and delivery choice equations, and the significance of sample selection relies on this correlation parameter, ρ (Helpman et al., 2008).

To take into consideration this selection effect, we employ a full information maximum likelihood (FIML) approach. We assume that the distribution of the errors is joint normal. The log-likelihood function is:

$$L = \sum_{i,j}(1 - T_{ij})\ln[\Phi(-W_{1ij})] + \sum_{i,j} T_{ij}\ln[\Phi(\frac{W_{1ij} + 2\rho\sigma_u^{-1}(W_{2ij})}{(1 - \rho^2)^{1/2}})]$$
$$+ \sum_{i,j} T_{ij}\ln\phi(\frac{W_{2ij}}{(\sigma_u/2)}) - \sum_{i,j} T_{ij}\ln(\sigma_u/2),$$

where $W_{1ij} = c_0 + \theta c_1 + c_3 + \theta\gamma \ln D_{ij} + \theta \ln w_i + (\theta + 1)\ln N_j + (\theta + 1)c_4 dum_j + c_5 dum_i$ and $W_{2ij} = q_{ij} - (1/2)\mu + (1/2)\gamma \ln D_{ij} + (1/2)\ln(1 + N_i) - (1/2)\ln(1 + N_j) + c_6 dum_j - c_7 dum_i$. Using FIML has several advantages: it is efficient, allows us to examine delivery choice, and it can detect unobservable factors driving self-selection bias in an explicit way. Therefore, our approach has the disadvantage of possible misspecification. However, we address this misspecification issue by doing diagnosis checks.

In the case of CES utility without producer heterogeneity, the estimating equation is:

$$z_{ij} = \beta - (\epsilon - 1)\gamma d_{ji} + \epsilon \ln P_i + \xi_j + \omega_l + \lambda_i + \eta_{ij},$$

where $\beta = \ln(1-\alpha) - (1-\epsilon)\ln\alpha - (1-\epsilon)\mu$, $\omega_l = (1-\epsilon)\phi$, and $\eta_{ij} = (1-\epsilon)u_{ij} + \nu_{ij}$. The price difference equation is:

$$q_{ij} = \mu + \gamma d_{ij} + u_{ij}.$$

Then, the log-likelihood function is as follows:

$$L = \sum_{i,j}(1-T_{ij})\ln[\Phi(-W_{3ij})] + \sum_{i,j} T_{ij}\ln[\Phi(\frac{W_{3ij} + \rho\sigma_u^{-1}(W_{4ij})}{(1-\rho^2)^{1/2}})]$$
$$+ \sum_{i,j} T_{ij}\ln\phi(\frac{W_{4ij}}{\sigma_u}) - \sum_{i,j} T_{ij}\ln\sigma_u,$$

where $W_{3ij} = \beta - (\epsilon - 1)\gamma d_{ji} + \epsilon \ln P_i + \xi_j + \omega_l + \lambda_i$ and $W_{4ij} = q_{ij} - \mu - \gamma d_{ij}$. We use the consumer price index as the price index, and using region-specific effects controls for other region-specific factors.

These two empirical models, one being the nonhomothetic model and the other the CES model, account for the data truncation problem caused by the self-selection of producers. The main difference between these approaches is in the price differential equation. In the CES case, it is simply a function of distance. On the other hand, in the nonhomothetic case, the effect of distance is different, and there are local market characteristics, which reflect producer heterogeneity and pricing-to-market behavior. We apply our model to the price and delivery data to find the distance elasticity.

3 Data

We apply our approach to data on wholesale prices of individual goods and delivery patterns across regions. Using wholesale prices enables us to focus on transportation costs because retail prices include local distribution costs. The individual goods are agricultural products in Japan. The wholesale prices of agricultural products in source regions and markets are available, thus the price difference between market and source prices can measure transportation costs properly.

The data source for wholesale prices is the Daily Wholesale Market Information of Fresh Vegetables and Fruits ("Seikabutsu Hinmokubetsu Shikyo Joho" in Japanese). The data set is collected by the Center of Fresh Food Market Information Service ("Zenkoku Seisen Syokuryohin Ryutsu Joho Senta" with the URL http://www2s.biglobe.ne.jp/fains/index.html), recording almost all transactions at 55 wholesale markets officially opened and operated in the 47 prefectures in Japan on a daily basis. This daily market survey covers the wholesale prices of 120 different fruits and vegetables.

Each agricultural product is further categorized by different varieties, sizes, and grades, as well as producing prefectures. Hence, for example, the data set reports the wholesale prices of potato at six different wholesale markets of the "Dansyaku (Irish Cobbler equivalent)" variety of size "L" with grade "Syu (excellent)" that was produced in "Hokkaido" prefecture on September 7, 2007. Because prices depend on characteristics, each combination of characteristics is identified as the same product. This high degree of categorization is important because the LOP exercises require analysis of identical goods for comparing prices to infer transportation costs precisely. We focus on eight selected vegetables: cabbage, carrot, Chinese cabbage (c-cabbage, hereafter), lettuce, Shiitake mushroom (s-mushroom, hereafter), spinach, potato, and welsh onion. In this paper, we examine the 2007 survey that reports the market transactions for 274 days.

The price reported in each market has three forms: the highest price, the mode price, and the lowest price. Most markets record all three prices, but several markets report only the highest and the lowest prices or only the mode price. Thus, we construct our price variable by averaging these price variables. We use the mode price when this is the only price available. The transaction unit of each product is also reported. To obtain the same unit of measurement for each product, we divide the price by the number of transaction units. Table 1 summarizes several descriptive statistics for these products. The first row reports the average price per kilogram (1 kilogram = approximately 2.2 pounds). S-mushroom is the most expensive, at 1113.627 yen (approximately 13 US dollars), and the cheapest is c-cabbage, which is 61.628 yen (approximately 0.9 US dollars).

Table 1 also shows that each product is highly categorized by product variety, size, and grade. The number of distinct products is large; 1,207 for cabbage; 1,186 for carrot; 1,001 for c-cabbage; 903 for lettuce; 1,423 for potato; 909 for s-mushroom; 551 for spinach; and 1,115 for welsh onion. For each product entry l, we count the number of deliveries as $T_{ij}(l) = 1$ and nondelivery as $T_{ij}(l) = 0$ only for the dates on which the product is traded at the wholesale market in producing prefecture j. We identify product delivery $T_{ij}(l) = 1$ if the data reports that the source prefecture of product entry l sold in consuming region i is region j. The price differential is constructed by subtracting the wholesale price in producing prefecture j, $p_j(\omega)$, from that in consuming prefecture i, $p_i(\omega)$. If the sample of $q_{ij}(\omega)$ is available, this means that $T_{ij}(\omega) = 1$ for pair (i, j).

The bottom part of Table 1 reports that the total number of both delivery and nondelivery observations across all products is greater than 190,000 for each vegetable. This is the number of observations used in our FIML estimation. Out of the total number of delivery and nondelivery cases, the number of delivery cases is relatively small: it is approximately 10,000 for each vegetable. Our data set, hence, indicates that product delivery is quite limited. The data truncation issue is quite important in this sample.

The other data we use in this paper are obtained as follows. The geographical distance between prefectural pair (i, j) is approximated by that between the prefectural head offices placed in the prefectural capital cities. The distance data are provided by the Geographical Survey Institute (GSI) of the Government of Japan. The data are publicly available at the GSI website.[2] We use daily temperature to control for supply and demand shocks. The daily temperature data are reported by the Japan Meteorological

[2]http://www.gsi.go.jp/kokujyoho/kenchokan.html

Agency. We download the data from their website.[3] Finally, we use the monthly data of the scheduled cash earnings for wages, which is reported in the Monthly Labour Survey ("Maitsuki Kinrou Tokei Chosa") conducted by the Ministry of Health, Labour, and Welfare. The data are available at the URL: http://www.mhlw.go.jp/toukei/list/30-1.html.

Table 1: Summary statistics

	Cabbage	Carrot	C-Cabbage	Lettuce	Potato	S-Mushroom	Spinach	Welsh Onion
Average price (yen per kg)	77.833	101.25	61.628	183.909	79.565	1113.627	496.372	382.099
Product entry								
No. of varieties	3	10	4	7	10	1	4	11
No. of size categories	63	62	50	71	50	74	17	103
No. of grade categories	34	66	50	46	93	55	85	58
No. of producing prefectures	47	46	46	43	47	44	47	46
No. of wholesale markets	47	47	47	47	47	47	47	47
No. of distinct product entries	1,207	1,186	1,001	903	1,423	909	551	1,115
Data truncation								
No. of $T_{ij}(\omega) = 0$ or 1	369,343	198,129	241,871	239,703	264,280	476,919	466,337	547,272
No. of $T_{ij}(\omega) = 1$	15,841	8,395	10,803	11,565	10,921	11,845	15,977	14,874

4 Estimation Results

Table 2 reports estimation results. The top half of Table 2 reports our main results. The results using the CES utility function and the simple regression results are also reported for comparison in the bottom half of Table 2. The distance elasticity in the nonhomothetic framework ranges from 0.973 (cabbage) to 1.301 (s-mushroom). This indicates that when the shipment distance from origin to destination increases by 1 percent, the price differential also increases by about 1 percent. These are larger than those in previous studies. Thus, our estimations imply an underbias of the distance elasticity in previous studies.

As in previous studies, if we use two market prices to construct price differentials and regress these on distance, then the distance effect coefficient is at most 0.05. That is, even if the distance is doubled, the price difference increases by only 5 percent. Thus, even using our data, regressing only a price difference

[3] http://www.data.jma.go.jp/obd/stats/etrn/index.php

(not identifying source region) on distance, which is the conventional method in the literature, yields similar results. The results of the CES utility function is similar to KKT (2010). As in KKT (2010), the price differential measure is the difference between market price and the price at production prefecture following Anderson and van Wincoop (2004). Delivery choice is explicitly modeled to control for sample selection. One difference from KKT (2010) is that KKT (2010) propose an instrumental variable estimation for structural estimation. While the results of the CES framework indicate significantly large distance effects, 0.301 to 0.522, these are smaller than those of the nonhomothetic model.

Comparing our results with those of a simple regression and the CES framework, our results indicate a much larger distance effect when incorporating producer heterogeneity and pricing to market. The CES results show that there is a large distance effect compared with the conventional OLS results. The results under nonhomothetic preferences are found to be even larger than the CES case. This is consistent with our argument, in which producer heterogeneity affects the pricing decision in each market and thus causes underbias in the distance elasticity estimates. This is because transportation costs induce only productive firms to deliver products, and these firms can charge a low price. Large distance elasticity estimates also imply that geographic barriers influence delivery choice. The probability of delivery will be reduced by the increase in transportation costs. Thus, large distance effects after accounting for producer heterogeneity suggest that the price of geographical barriers is still high for regional transportation.

Another important parameter in our estimations is the heterogeneity parameter, θ. Our estimates are from 0.634 to 1.373. A small θ means that there is a large dispersion in productivity. These estimates can be considered to be small (producer heterogeneity is largely dispersed). However, this may be because farmers in Japan are quite heterogeneous. In particular, small farms operated by elderly people in suburban areas produce agricultural products. On the other hand, in agricultural prefectures such as Hokkaido, there exist large-scale farms. In 2009, the average area under cultivation was 20.50 hectares (approximately 50.66 acres) in the Hokkaido prefecture, while that in the other prefectures was 1.41

hectares (approximately 3.48 acres). These farms may deliver their products to the same market. In our framework, all prefectures have the same productivity distribution, so the low value of θ may reflect this dispersion among farms.

In the trade literature, the heterogeneity parameter, θ, is investigated extensively. In the Eaton and Kortum (2002) framework, this is the elasticity of the trade parameter, which is a crucial parameter for welfare gain analysis from trade (Arkolakis, Costinot, and Rodriguez-Clare, 2011). Eaton and Kortum (2002) estimate this parameter to be 8.28, Bernad, Jensen, Eaton, and Kortum (2003) estimate it to be 3.6, Crozet and Koenig (2010) estimate from 1.65 to 7.31, Simonovska and Waugh (2010) use the simulated method of moments to obtain estimates from 3.57 to 4.46, and Balistreri et al. (2011) estimate from 3.924 to 5.171. Donaldson (2010) uses the Eaton and Kortum (2002) model to estimate the productivity variability parameter, and estimates a value of 3.8 on average. As in Donaldson (2010), we use price data to estimate two crucial parameters in the producer heterogeneity model. Our estimates are lower than these studies. This may be because the more disaggregated the product level is, the higher the dispersion of heterogeneity. Our sample is disaggregated product-level data, and has a quite fine categorization, thus our estimates report a small θ.

The correlation parameter ρ is also important for the significance of sample selection. These estimates are from -0.12 to -0.313. All results are negative and statistically significant. Hence, in order to identify the true parameter, controlling for selectivity bias is crucial. When there is a positive shock that increases the price differentials caused by transportation costs (for example, a fuel price increase), then this same shock decreases the probability of delivery. Without controlling for this negative correlation caused by unobservable shocks, as we have seen, the distance effects are found to be small. We detect the existence of such a negative effect.

The relevance of the estimates depends on the empirical validity of our model. For model validation purposes, we conduct diagnosis checks of our model with respect to two important aspects of the actual

Table 2: Estimation results

	Cabbage	Carrot	C-cabbage	Lettuce	Potato	S-mushroom	Spinach	Welsh onion
Nonhomothetic								
γ	0.973	1.165	1.181	1.058	1.226	1.301	1.081	0.975
	(0.006)	(0.01)	(0.009)	(0.008)	(0.009)	(0.011)	(0.007)	(0.006)
θ	1.061	0.796	1.009	0.908	0.634	1.292	1.172	1.373
	(0.006)	(0.006)	(0.007)	(0.006)	(0.004)	(0.013)	(0.007)	(0.008)
ρ	-0.257	-0.313	-0.251	-0.289	-0.164	-0.12	-0.279	-0.286
	(0.003)	(0.004)	(0.003)	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)
log-likelihood	-48193.573	-35516.506	-30662.217	-39012.544	-45472.895	-11895.094	-40103.727	-33217.413
CES								
γ	0.301	0.362	0.412	0.426	0.348	0.522	0.433	0.384
	(0.002)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)
ϵ	4.258	3.082	3.702	3.081	3.338	4.572	3.739	4.186
	(0.019)	(0.015)	(0.014)	(0.014)	(0.016)	(0.024)	(0.015)	(0.017)
ρ	-0.847	-0.866	-0.826	-0.863	-0.771	-0.543	-0.835	-0.844
	(0.002)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)
log-likelihood	-26486.776	-20688.309	-17725.484	-25224.822	-27287.880	-5636.887	-25461.829	-20522.294
OLS								
γ	0.033	0.051	0.042	0.022	0.037	0.007	0.044	0.033
N	369,343	198,129	241,871	239,703	264,280	476,919	466,337	547,272
PCP for T_{ij}	0.966	0.964	0.961	0.961	0.966	0.994	0.979	0.988

The numbers in parentheses are standard errors. All estimations include origin and destination dummies, origin and destination daily temperature, and the number of products in both equations, and wages for the selection equation.

data: the pattern of product delivery and the association of price differentials with delivery distances. First, we calculate the percentage correctly predicted (PCPs) measures for $T_{ij}(l) = 0$ or 1. To construct the PCPs, we calculate the predicted conditional probabilities and the predicted delivery index where the predicted probabilities are greater than 0.5. The results are reported on the bottom line in Table 2. The PCPs are all greater than 0.96, which suggests that our model successfully predicts actual delivery patterns.

The second diagnosis is about price differentials with respect to delivery distances. The question is whether our sample-selection model predicts the actual price differentials. To do this diagnosis check, we derive the prediction of the model for price differentials with selectivity bias controlled for. Each window of Figure 1 plots the resulting predicted price differentials (dots) as well as the data counterparts (crosses) against the corresponding log distances for each vegetable. The dots are distributed inside the cloud formed by the crosses in all windows. This means that our model successfully predicts the relationship between the price differentials and distances overall.

One issue remaining when comparing the results of the nonhomothetic and CES models is the elasticity of substitution parameter, ϵ. In the nonhomothetic preference model, the utility function is in log form to obtain an explicit solution for the optimal price. Because the coefficient of distance in the selection equation is $\theta\gamma$ in the nonhomothetic case and $(\epsilon - 1)\gamma$ in the CES model, ignoring the elasticity of substitution may cause small estimates of θ and large estimates of γ. If this composite remains constant, a small elasticity of substitution may imply a large distance effect. However, there is no such positive relationship found in our empirical results of the CES model. Thus, the direction of bias (if it exists) is not clear, and this is a limitation of our study.

5 Concluding Remarks

We have investigated the impact of producer heterogeneity and pricing-to-market behavior on the distance elasticity in regional price differentials. Because, in a conventional CES utility framework, producer heterogeneity does not play a crucial role in the identification of the distance effect, we developed a nonhomothetic preference model, and therefore incorporate pricing-to-market behavior.

Our empirical analysis shows that ignoring these factors causes underestimation in the CES utility framework. The distance effect is significantly large for regional price differences. These results suggest that the price of geographical barriers is still high for regional transportation. Even though Japan is considered to have a well-established infrastructure and sophisticated logistics system, geographic barriers are large enough to create substantial price differences. Thus, in a country with lower transportation facilities and services, regional differences may be large, and markets can be geographically segmented.

References

[1] Anderson J., E. van Wincoop, 2004 Trade costs, Journal of Economic Literature 42, 691-751.

[2] Arkolakis, C., Costinot, A., A. Rodrigeuz-Clare, 2011 New trade models, same old gains? American Economic Review, forthcoming.

[3] Atkeson, A., A. Burstein 2008 Pricing-to-market, trade costs, and international relative prices, American Economic Review 98, 1998-2031.

[4] Balistreri, E. J., Hillberry, R. H., T. F. Rutherford, 2011 Structural estimation and solution of international trade models with heterogeneous firms, Journal of International Economics 83, 95–108.

[5] Bergin, P. R., Glick, R., A. M. Taylor, 2006 Productivity, tradability, and the long-run price puzzle, Journal of Monetary Economics 53, 2041–2066.

[6] Combes, P.-P., M. Lafourcade, 2005 Transport costs: measures, determinants, and regional policy implications for France, Journal of Economic Geography 5, 319–349.

[7] Crozet, M., P. Koenig, 2010 Structural gravity equations with intensive and extensive margins, Canadian Journal of Economics 43, 41–62.

[8] Crucini, M. J., Shintani, M., T. Tsuruga, 2010 The law of one price without the border: the role of distance versus sticky prices, Economic Journal 120, 462–480.

[9] Donaldson, D, 2010 Railroads of the Raj: estimating the impact of transportation infrastructure, NBER Working paper 16487.

[10] Eaton, J., S. Kortum 2002 Technology, geography, and trade, Econometrica 70, 1741-1779.

[11] Engel, C., J. H. Rogers 1996 How wide is the border? American Economic Review 86, 1112–25.

[12] Ghironi, F., M. J. Melitz, 2005 International trade and macroeconomic dynamics with heterogeneous firms. Quarterly Journal of Economics 120, 865-915.

[13] Helpman, E., Melitz, M. J., Y. Rubinstein, 2008 Estimating trade flows: trading partners and trading volumes, Quarterly Journal of Economics 123, 441-487.

[14] Hummels, D., 2007 Transportation costs and international trade in the second era of globalization, Journal of Economic Perspectives 21, 131–154.

[15] Kano, K., Kano, T., K. Takechi, 2010 Exaggerated death of distance: revisiting distance effects on regional price dispersions, mimeo.

[16] Melitz, M., G. Ottaviano, 2008 Market size, trade, and productivity, Review of Economic Studies, 75, 295–316.

[17] Parsley, D. C. and S-J. Wei 1996 Convergence to the law of one price without trade barriers or currency fluctuation, Quarterly Journal of Economics 111, 1211–1236.

[18] Parsley, D. C. and S-J. Wei 2001 Explaining the border effect: the role of exchange rate variability, shipping costs, and geography, Journal of International Economics 55, 87–105.

[19] Simonovska, I., 2010 Income differences and prices of tradables, NBER Working Paper 16233.

[20] Simonovska, I. and M. E. Waugh, 2010 The elasticity of trade: estimates and evidence, mimeo.

[21] Young, A., 1991 Learning by doing and the dynamic effects of international trade, Quarterly Journal of Economics 106, 369–405.

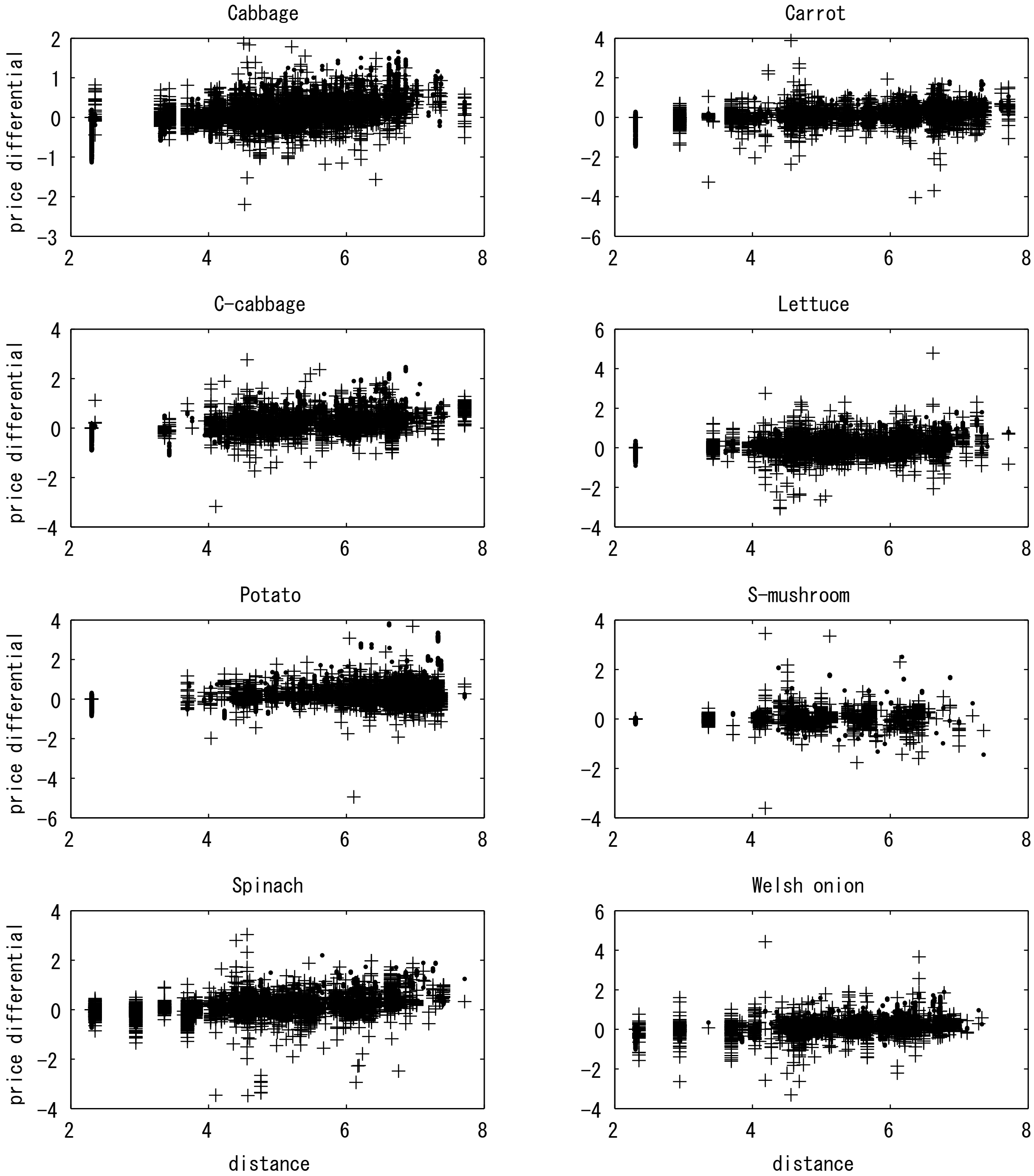

Figure 1: Actual (+) and Predicted (.) Values