Appendices for "The Price of Distance: Pricing-to-Market and Geographic Barriers", by Kano, Kano, and Takechi (2021)

Appendix A: A general demand model

The general demand function in the paper is developed by Arkolakis et al. (2019). In this section, we show the properties of the general demand function. The demand function for a differentiated good, ω, is represented by the following demand function (Arkolakis et al. 2019):

$$q_i(\omega) = Q(p, w_i)D(p(\omega)/P(p, w_i)),$$

where $p = \{p(\omega)\}$ is a price schedule, Q is an aggregate demand shifter, D is a demand function, and P is a price index.

This general demand has important implications. Because the demand elasticity, $\epsilon_D = -(1/P)(\partial D/\partial p(\omega))(p(\omega)/D)$, may vary with $p(\omega)$, variable markups occur under monopolistic competition. The effect of other goods price, p, is, in effect, through the aggregate shifters, Q and P. This property is suitable for the industry that we investigate because we do not have strategic interactions among producers. By assuming symmetry across goods, the demand in region i can be expressed by $q_i(\omega) = Q_i D(p(\omega)/P_i)$, where Q_i and P_i are aggregate variables in region i.

The producer profit-maximization problem in a monopolistically competitive market is to maximize profits by choosing a market-specific price:

$$\max_{p_{ij}} \pi_{ij} = \max_{p_{ij}} p_{ij} Q_i D(p_{ij}/P_i) - \frac{\tau_{ij} w_j}{\psi} Q_i D(p_{ij}/P_i),$$

where p_{ij} is the price in region i for products from region j.

The first-order condition can be expressed by $(p_{ij} - c_{ij})/p_{ij} = 1/\epsilon_D$, where $c_{ij} = w_j \tau_{ij}/\psi$ and $\epsilon_D = -(1/P_i)(\partial D/\partial p_{ij})(p_{ij}/D)$. Let us designate $m = p_{ij}/c_{ij}$ as a measure of markups and $v = P_i/c_{ij}$ as a producer efficiency measure. The optimal condition is expressed as follows: $m = \epsilon_D(m/v)/(\epsilon_D(m/v) - 1)$. Then, there exists a unique function, $m = \mu(v)$ for $v > 0$, in which $\epsilon'_D > 0$ is a sufficient condition for existence and uniqueness. Thus, the markup m associated with the behavior of μ function varies. This μ function differentiates between constant and variable markups. The optimal price is obtained as follows:

$$p_{ij} = \mu(\frac{P_i \psi}{\tau_{ij} w_j}) \frac{\tau_{ij} w_j}{\psi}.$$

Appendix B: Identification issues

In this section, we discuss the bias involved in trade cost estimation in the literature. As shown by Cosar et al. (2015), when comparing prices between markets, not prices between market and source, there is an underbias in the distance effect. For expositional simplicity, without variable markups, the price differentials between market h and i from source j can be expressed as follows: $\ln p_{hj} - \ln p_{ij} = \gamma(\ln d_{hj} - \ln d_{ij})$. However, when comparing the prices between markets, h and

i, the distance between h and i is used. Thus, it is common for the following equation to be estimated: $\ln p_{hj} - \ln p_{ij} = \gamma(\ln d_{hi})$. The triangle inequality suggests that $d_{hi} \geq |d_{hj} - d_{ij}|$ and, hence, the resulting estimates of γ tend to be underbiased (Cosar et al. (2015)). This finding reveals that it is necessary to use source information to properly measure the distance effect. In addition, as considered by Eaton and Kortum (2002) and Simonovska and Waugh (2014), the relationship between market price differentials and trade costs is obtained when the maximum of relative price satisfies the following inequality: $\max(p_{hj}/p_{ij}) \leq \tau_{hi}$. However, if the source price is available, then the relative price captures trade costs precisely: $p_{hi}/p_{ii} = \tau_{hi}$. As discussed, we have $\ln p_{hi} - \ln p_{ii} = \ln \tau_{hi} \geq \ln p_{hj} - \ln p_{ij}$. Hence, the price differentials used in the literature, $\ln p_{hj} - \ln p_{ij}$, are lower than the true trade costs using the source price, which leads to underbiased results for the distance effects.

Another important identification issue concerns the joint distribution of the dependent variable (price differentials, r_{ijt}^{NHOM}) and the main explanatory variable (distance, D_{ij}), as in Helpman et al. (2008), Johnson (2012), and Kano et al. (2013). When estimating the distance elasticity, using only the price differential equation (24) may lead to a bias toward underestimation in the results. This is because observable price differentials are selected by the delivery choice rule (equation (23)). The magnitude of the distance elasticity is captured by the fact that larger distances widen price differentials because of larger trade costs. At the same time, in more distant markets, these possibly wider price differentials are not observed because it is too costly for producers to deliver their products. Thus, if we restrict our attention to the observed price differentials, we may end up using only the smaller price differential data. Formally, the conditional expectation of price differentials is expressed as: $E[r_{ijt}^{NHOM}|z_{ijt}^{NHOM} \geq 0] = (1/2)\mu + (1/2)\gamma \ln D_{ij} + (1/2)\chi_1 \ln N_{jt} - (1/2)\chi_2 \ln N_{it} + c_7 dum_j - c_8 dum_i + c_9 temp_{it} + c_{10} temp_{ji} + c_{11} month_t + (1/2)E[u_{ijt}|z_{ijt}^{NHOM} \geq 0]$. Because $E[u_{ijt}|z_{ijt}^{NHOM} \geq 0] = \rho\frac{\sigma_u}{\sigma_\eta}E[\eta_{ijt}| \geq 0]$, the error term in the price differential equation is correlated with that in the selection equation. This bias term is expressed as an inverse Mills ratio: $E[\eta_{ijt}|z_{ijt}^{NHOM} \geq 0] = \phi(\hat{z}_{ijt}^{NHOM})/\Phi(\hat{z}_{ijt}^{NHOM})$ (Helpman et al. 2008). Hence, to obtain consistent estimates, we must consider the correlation between the price differential and delivery choice equations; the significance of the sample selection relies on this correlation parameter, ρ.

We also carry out Monte Carlo experiments based on the model to illustrate the following issues: 1) how the exercises in the previous literature lead to underbiased estimates; 2) how the selection problem causes a bias toward underestimation in estimating the distance effect; and 3) how pricing-to-market behavior makes price differentials less than trade costs, given the distance to market for the product delivered. We create an economy based on a straight line. There are 47 regions for Japan's 47 prefectures, which are placed on the integer line between 1 and 47. Each point represents a prefecture. The left end of the line represents 1 and the right end 47. Thus, the distance between two regions (i, j) is defined by $|i - j|$, and the internal distance of the region is assumed to be 0.376. This is obtained from the internal distance formula, which is 0.376 times the land area of the prefecture (Head and Mayer 2000). We assume that the area is equal to 1. In each region, there are monopolistically competitive producers, potentially supplying their goods to all regions. If the gross profit from supplying to a market minus the fixed costs is not negative, the producers deliver their products to this market.

Consumers reside in each region and are assumed to have nonhomothetic preferences. The nonhomothetic preferences are shown in equation (13). The producers face the demand function in equation (14) and maximize their profits by choosing the price, as in the CES case. The resulting optimal price is given by equation (15) and the price differential is shown by equation (17). We set the wage rate at one for simplicity. The producer dispersion parameter θ is set to 2.5. The parameter on which we focus is the distance elasticity parameter γ, which is set to 0.5. We allow for idiosyncratic random variations in trade costs and fixed costs. The standard deviation for trade costs, σ_u, is 0.35, and that for fixed costs, σ_μ, is 0.5. The fixed trade costs are set to 0.5. The measure of the mass of products, measured by the number of products in each region, is drawn from a multinomial distribution.

We first draw 500 sets of a random variable of productivity from a Pareto distribution. Next, we calculate trade costs, price differentials, and latent variables of delivery choice, following equations (5), (17), and (19), respectively. We also conduct a common exercise in the literature, which involves picking up two market prices and relating these to the distance between markets. Figure B1 depicts the trade costs and price differentials simulated from the model. The left side of Figure B1 shows the relationship between trade costs and distance, and the absolute value of the price differentials between markets and the distance between markets, not the destination–source relationship. As shown, the effect of distance is weaker for market price differentials than for trade costs. The slope of the true trade cost–distance relationship is 0.51 ($*$), whereas that for market price differentials and distance is 0.025 (squares). This small size of the distance effect on price differentials is quite similar to that in the previous LOP studies, even though our data are artificial. The middle of Figure B1 plots the trade costs and reveals the difference between untruncated ($*$) and truncated (dots) data caused by the self-selection bias. Because delivery patterns depend on profitability, subject to trade costs, a large part of the trade cost data for distant markets is truncated. As in the left panel of the figure, the slope coefficient of the untruncated data is 0.51 ($*$), compared with 0.381 for the truncated data (dots). Thus, the correlation between trade costs and distance leads to underestimation when only truncated data are used.

Another cause of the bias toward underestimation is a result of pricing-to-market behavior. The right side of Figure B1 shows that there is a stronger relationship between trade costs (dots) and distance than between observable price differentials (circles) and distance. The slope of the observable price differentials–distance line is 0.202. This is because some portion of the trade cost increase associated with distance is absorbed by the producers. The producers do not pass through trade cost increases in the form of price increases. Hence, price differentials between regions do not correctly measure trade costs. The relationship between observed price differentials and distance is weaker than that between trade costs and distance.

As seen in these figures, there is a bias toward underestimation that occurs when using only observed price differential data. Self-selection and pricing-to-market means that price differential observations do not represent trade costs. To conduct our analysis, we must formulate the model to correct the biases in identifying the distance effect, which is the purpose of this study.

Appendix C: Estimation results

Table C1 shows the estimation results of the variable markup model. It reports the coefficients of all market specific variables for the variable markup term.

References

Arkolakis, C., Costinot, A., Donaldson, D., Rodriguez-Clare, A. (2019). The elusive pro-competitive effects of trade, Review of Economic Studies 86, 46–80.

Cosar, A. K., Grieco, P. L. E., Tintelnot, F. (2015) Bias in estimating border- and distance-related trade costs: insights from an oligopoly model, Economics Letters 126, 147–149.

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

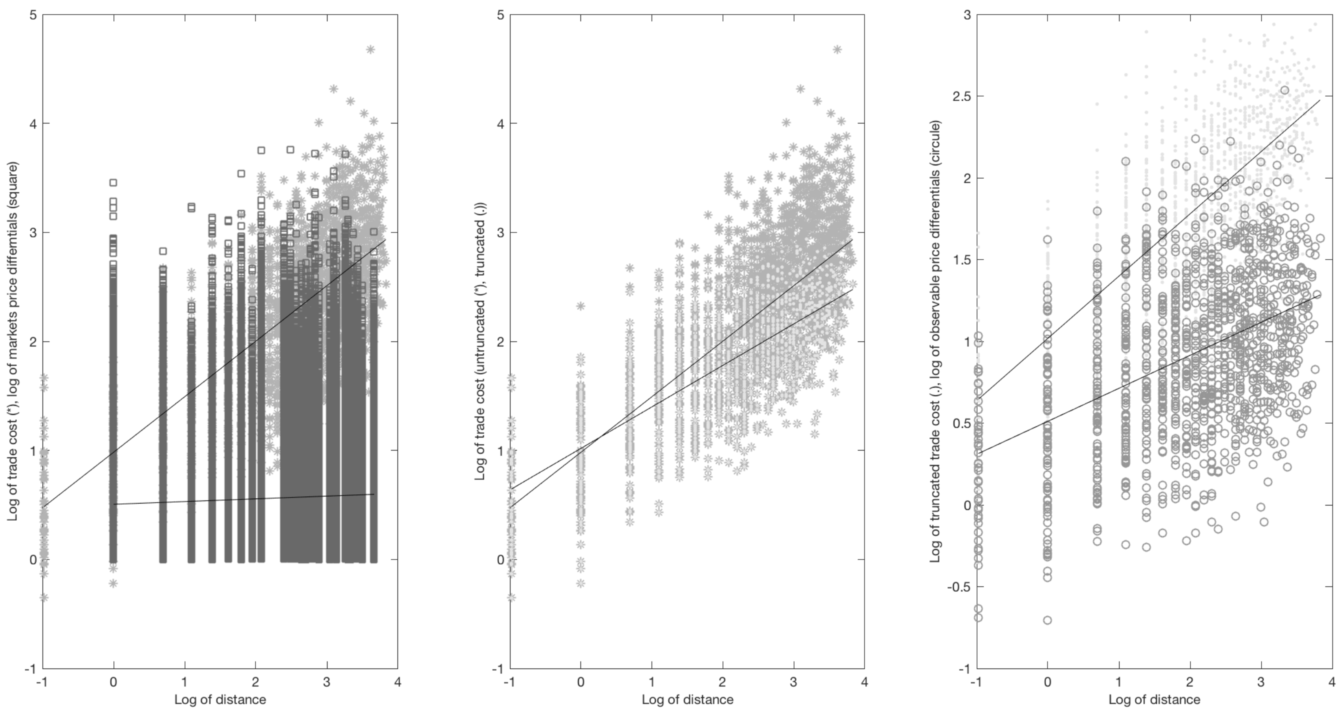
Head, K., Mayer, T. (2000). Non-Europe: the magnitude and causes of market fragmentation in the EU, Weltwirtschaftliches Archiv 136, 284–314.

Helpman, E., Melitz, M. J., Rubinstein, Y. (2008). Estimating trade flows: trading partners and trading volumes, Quarterly Journal of Economics 123, 441–487.

Johnson, R. C. (2012). Trade and prices with heterogeneous firms, Journal of International Economics 86, 43–56.

Kano, K., Kano, T., Takechi, K. (2013). Exaggerated death of distance: revisiting distance effects on regional price dispersions, Journal of International Economics 90, 403–413.

Simonovska, I., Waugh, M. E. (2014). The elasticity of trade: estimates and evidence, Journal of International Economics 92, 34–50.



Note: The data are generated by Monte Carlo simulation. The horizontal axis is the log distance between markets in all figures. Left figure: vertical axis is log of trade costs (star) and price differentials between markets (not destination–origin pairs) (square). Middle figure: vertical axis is log of trade costs (star) and price differentials between destination and source when truncated (dots). Right figure: vertical axis is price differentials between destination and source when truncated (dots) and observable price differentials between destination and source (circle).

Figure B1: Monte Carlo experiment: Logs of distance and price differentials

Table C1: Estimation Results

ACDR	Cabbage	Carrot	C-Cabbage	Lettuce	Potato	S-Mushrooms	Spinach	WelshOnions
γ	1.049	1.525	2.372	2.471	2.525	3.174	2.347	4.868
	(0.028)	(0.058)	(0.074)	(0.072)	(0.043)	(0.101)	(0.048)	(0.07)
γ_{d^2}	-0.241	-0.329	-0.539	-0.561	-0.523	-0.713	-0.557	-1.102
	(0.007)	(0.013)	(0.017)	(0.016)	(0.01)	(0.024)	(0.011)	(0.016)
γ_{d^3}	0.019	0.023	0.04	0.041	0.035	0.05	0.043	0.079
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
γ_{w_i}	0.008	-1.478	-0.339	-0.443	-1.443	-0.443	1.011	0.943
	(0.037)	(0.171)	(0.077)	(0.069)	(0.037)	(0.147)	(0.047)	(0.227)
γ_{w_j}	0.179	0.212	-2.629	-1.962	2.33	0.244	-0.776	-1.221
	(0.017)	(0.155)	(0.199)	(0.156)	(0.091)	(0.117)	(0.106)	(0.175)
γ_{p_i}	0.092	-0.092	0.089	0.042	0.291	0.048	-0.009	-0.033
	(0.009)	(0.012)	(0.013)	(0.012)	(0.012)	(0.015)	(0.01)	(0.012)
γ_{p_j}	-0.189	0.144	-0.134	-0.028	-0.709	-0.053	-0.008	0.004
	(0.009)	(0.014)	(0.011)	(0.012)	(0.012)	(0.016)	(0.011)	(0.012)
γ_{N_i}	0.057	-0.153	-0.067	-0.069	-0.128	-0.279	-0.134	-0.313
	(0.009)	(0.015)	(0.022)	(0.019)	(0.028)	(0.017)	(0.013)	(0.014)
γ_{N_j}	-0.032	0.17	0.15	0.15	0.067	0.304	0.209	0.331
	(0.009)	(0.015)	(0.024)	(0.021)	(0.03)	(0.018)	(0.013)	(0.015)
γ_{dN_i}	-0.019	0.019	0.011	0.01	0.018	0.057	0.017	0.061
	(0.002)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)
γ_{dN_j}	0.013	-0.034	-0.032	-0.032	-0.014	-0.065	-0.059	-0.06
	(0.002)	(0.003)	(0.005)	(0.004)	(0.005)	(0.005)	(0.003)	(0.003)
γ'_{w_i}	-5.359	-0.921	-0.864	-0.702	-3.423	3.019	2.677	-5.922
	(0.24)	(0.468)	(0.321)	(0.331)	(0.272)	(0.085)	(0.088)	(0.398)
γ'_{w_j}	6.589	1.265	-0.024	0.138	7.299	0.785	0.916	10.381
	(0.27)	(0.383)	(0.158)	(0.127)	(0.377)	(0.069)	(0.217)	(0.53)
γ'_{p_i}	0.557	0.407	0.2	0.471	0.173	-0.336	0.445	0.462
	(0.017)	(0.047)	(0.03)	(0.032)	(0.029)	(0.029)	(0.031)	(0.042)
γ'_{N_i}	0.282	0.407	1.156	1.156	0.633	0.432	0.082	0.219
	(0.03)	(0.038)	(0.054)	(0.043)	(0.026)	(0.051)	(0.027)	(0.031)
ρ	-0.203	-0.323	-0.307	-0.308	-0.339	-0.178	-0.259	-0.245
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
loglik	-39674.026	-33483.064	-40446.776	-40428.039	-51563.9	-11461.099	-35768.155	-28801.848
N	385,776	211,101	253,287	252,810	274,893	480,307	478,726	553,854

Note: The numbers in parentheses are standard errors calculated using the inverse Hessian matrix of the log-likelihood function. All estimations include dummy variables (month, origin fixed effects, and market fixed effects).