

# Bias in Estimating Border- and Distance-related Trade Costs: Insights from an Oligopoly Model

A. Kerem Coşar  
University of Chicago  
Booth School of Business

Paul L. E. Grieco  
Pennsylvania State University

Felix Tintelnot  
University of Chicago

October 3, 2014

## Abstract

Regressions of price differences between locations in different countries without controlling for the local market structure and the location of origin of the product will lead to a biased estimate of the impact of national boundaries. We demonstrate that non-classical measurement error in distance and unaccounted mark-up differences across countries are responsible for these biases. In a quantitative exercise based on our previous work (Coşar et al., 2014), we show that the estimated border effect with price difference regressions overstates the true border effect by a factor of two or more.

**JEL Codes: F10, L13.**

## 1 Introduction

This note illustrates a potential bias in using data on prices and distances between sales locations to estimate border and distance related trade costs. We compare a structural equation for price differences to a reduced form specification to illustrate biases related to mis-measurement of distance and endogeneity due to variable markups. A simulation exercise where we generate price differences from a structural model and use these to estimate the reduced form specification illustrates that these biases are substantial. The result is that the border effect is significantly over-estimated by the reduced form approach.

The conceptual setup is deliberately stylized. Consider an oligopolistic industry in a two country setting. There is a discrete, finite set of firms and buyers with exogenously given locations. Each firm has a single location in space and at least some firms compete in both countries. Let  $\phi_\ell$  be the marginal cost of production for firm  $\ell$ . In addition, the firm faces a cost of  $\beta_d \cdot d$  to sell over a distance of  $d$ , and a border cost of  $\beta_b$  if it is exporting.

The total cost for firm  $\ell$  of supplying buyer  $i$  is,

$$c_{\ell i} = \phi_{\ell} + \beta_d \cdot d_{\ell i} + \beta_b \cdot b_{\ell i}, \quad (1)$$

where the border dummy  $b_{\ell i}$  equals one if  $\ell, i$  are in different countries, and zero otherwise. The term  $d_{\ell i}$  denotes the distance between the firm and buyer.

For the basic results we want to highlight in this paper, the details of the market structure do not matter. Products supplied by different firms could be homogeneous or differentiated. Firms could engage in quantity (Cournot) or price (Bertrand) competition.<sup>1</sup> Buyers could be downstream firms or final consumers. Different subsets of firms may be competing in the two markets. We thus focus on a generic Nash equilibrium with mark-up pricing in which firm  $\ell$  sets its price at location  $i$  as,

$$p_{\ell i} = c_{\ell i} + m_{\ell i}. \quad (2)$$

## 2 Estimating trade and border costs from price differentials

A researcher interested in estimating distance- and border-related costs of trade rarely has data about the exact location of firms. Rather, the researcher typically observes prices in various consumption locations. Endowed with such data, many scholars used price dispersion between markets to infer the primitives of trade costs, i.e.  $\beta_d$  and  $\beta_b$ , by estimating reduced form models.

In a typical exercise, the researcher starts with price data for identical, tradable goods indexed by  $\ell$  observed in locations  $(i, j)$  and estimates the following relationship,

$$|p_{\ell i} - p_{\ell j}| = \gamma_{\ell} + \gamma_i + \gamma_j + \delta_d \cdot d_{ij} + \delta_b \cdot b_{ij} + \varepsilon_{\ell ij}. \quad (3)$$

Note that our econometrician regresses price differentials on the distances  $d_{ij}$  between buyers—but not the distances between buyers and producers—a border dummy  $b_{ij}$  denoting whether or not the two buyers are on opposite sides of a national boundary, and dummy variables representing market and firm fixed effects to control for differences that impact price levels.<sup>2</sup> Estimates of  $(\hat{\delta}_d, \hat{\delta}_b)$  are then used to infer  $(\beta_d, \beta_b)$ , respectively. Following Engel and Rogers (1996), particular attention has been paid to the ratio  $\hat{\delta}_b/\hat{\delta}_d$ , i.e. the “border effect,” showing the equivalent number of miles that would generate the same price dispersion as the border dummy coefficient.<sup>3</sup>

<sup>1</sup>Brander (1981) and Brander and Krugman (1983) analyze international trade under Cournot competition in identical products. Friberg and Ganslandt (2008) analyze the case for Bertrand competition in differentiated products.

<sup>2</sup>Authors have estimated several variations of this specification. For example, one could use price covariances over time instead of absolute price differences (Engel and Rogers, 1996), or not include market fixed effects (Broda and Weinstein, 2008).

<sup>3</sup>Sometimes log distance is used as a regressor, in which case the border effect is represented as  $\exp(\hat{\delta}_b/\hat{\delta}_d)$ .

The reduced form specification (3) can be compared to a structural model of price differences derived from (1) and (2),

$$|p_{\ell i} - p_{\ell j}| = |\beta_d \cdot (d_{\ell i} - d_{\ell j}) + \beta_b \cdot (b_{\ell i} - b_{\ell j}) + (m_{\ell i} - m_{\ell j})|, \quad (4)$$

Clearly, the reduced form analysis will be subject to specification error. The remainder of this section details two specific biases which will tend to result in the reduced form over estimating the border effect.

## 2.1 Measurement error in distance

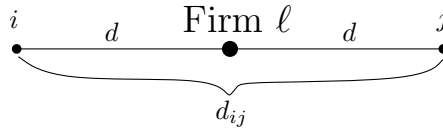
For simplicity in order to isolate the role of measurement error, assume for this subsection that the firm charges a constant absolute markup, all three points are in the same country, and distance costs are non-negative. Then the structural equation for price differences simplifies to,

$$|p_{\ell i} - p_{\ell j}| = \beta_d \cdot |d_{\ell i} - d_{\ell j}|.$$

The reduced form estimating equation (3) replaces  $|d_{\ell i} - d_{\ell j}|$ , the differences in producer-to-consumer distances, with consumer-to-consumer distances  $d_{ij}$ . The reverse triangle inequality implies that for any configuration of production and consumption locations  $d_{ij} \geq |d_{\ell i} - d_{\ell j}|$ . Since the distance between consumers is necessarily larger than the difference in the firm-consumer distances, this tends to bias  $\hat{\delta}_d$  towards zero relative to the true distance parameter  $\beta_d$ .

To illustrate, consider a pair of consumption points  $i, j$  that are  $d = d_{\ell i} = d_{\ell j}$  miles from firm  $\ell$  in opposite directions (figure 1). In this case  $d_{ij} = 2d$  and  $|d_{\ell i} - d_{\ell j}| = 0$ . Then for any  $d$ ,  $|p_{\ell i} - p_{\ell j}| = 0$ . It is clear that a dataset of such pairs in which  $d$  varies cannot be used to estimate distance costs as it does not provide variation in  $|d_{\ell i} - d_{\ell j}|$ . However under the reduced form specification, the variation in  $d$  would be used to estimate  $\delta_d = 0$  for *any* value of  $\beta_d$ .

Figure 1: Measurement Error in Distance



This intuition will carry over to the general case. The use of destination-to-destination distances to proxy for origin-to-destination distances will result in non-classical measurement error and biased estimates of the distance cost. The triangle inequality strongly suggests that this bias will be downward. As a result, we would expect the “border effect” estimate to be biased upward relative to  $\beta_d/\beta_d$ . This intuition resonates with Atkin and Donaldson (2014) who show that when using origin-destination location pairs for price

difference regressions the estimated trade costs are significantly larger than when using all pairs of locations.

## 2.2 Omitted variable bias from missing mark-ups

A second key difference between (3) and (4) is that the reduced form does not account for differences in markups, which are typically unobserved and would be transmitted to the error term. If markups are correlated with either distance or destination country, this will introduce an endogeneity bias into the estimation of (3). While the inclusion of origin and destination fixed effects may partially control for markup levels, they cannot be successful in controlling for markup differences between *pairs* of locations, which are they key factor in (4).

It is natural to assume that markups will be correlated with distance and export status. For example, greater distance implies higher cost, which will in turn imply lower markups in a typical differentiated Bertrand or Cournot setting. Moreover, if the number of competing firms changes across countries this is likely to cause a discontinuous change in markups at the border which will necessarily be correlated with the border dummy. In principle, this bias could be positive or negative depending on whether markups are higher in the firms home or export country.

This result resonates with Gorodnichenko and Tesar (2009) who argue that in the presence of location heterogeneity, model-free border-effect regressions based on price differentials fail to identify border frictions. In our case, this heterogeneity is markup variation across space.

## 3 Quantitative importance

In Coşar et al. (2014), we estimate a model of spatial oligopolistic competition in the wind turbine industry using data from Germany and Denmark. Our data inform us about the locations of wind farm and producers, as well as winning suppliers. In the model, a finite number of wind turbine producers engage in Bertrand competition in bidding for a number of potential wind farms. In particular, five Danish firms and a German firm have sales in both countries. While the structural estimation there is based on a more elaborate discrete choice model, the basic setup described above captures its key features. Essentially, the use of source-to-destination distances together with a structural model to control for markups enables us to directly estimate the structural parameters cost function (1).

To illustrate the bias in the reduced form specification, we use our model to directly compute price differences for a given parameterization of the cost function via (4). We then use these price differences to estimate (3) for each firm. We can then compare the “true” distance and border parameters to the reduced form estimates.<sup>4</sup>

---

<sup>4</sup>For this exercise, we use parameter estimates from the working paper version of Coşar et al. (2014) dated December 2012.

Table 1 reports the border and distance coefficients, as well as the implied width of the border. All coefficients are significant at the 1 percent level. The reduced form estimates are also statistically different from the structural parameters at the 1 percent level. With distance measured in 100 kms, our cost specification (1) implies a border effect of  $100 \times 0.873/0.2 = 435.5$  km. For five out of six firms, the reduced-form border effect is around twice the width of the border implied by the structural coefficients used in simulating the prices, suggesting that problems we identified can lead to substantial bias in such estimations.

Table 1: Structural vs. Reduced-Form Border Effect Estimates

	<b>Border Coefficient</b>	<b>Distance Coefficient</b>	<b>Border Width in km (<math>100 \cdot border/distance</math>)</b>
Structural Parameters	0.873	0.200	435.5
Reduced-form Estimates			
Firm 1 (DNK)	1.049	0.124	844.3
Firm 2 (DNK)	1.099	0.124	886.5
Firm 3 (DNK)	1.014	0.124	816.6
Firm 4 (DNK)	0.844	0.119	709.1
Firm 5 (DNK)	1.194	0.134	888.9
Firm 6 (DEU)	0.489	0.015	3253.8

*Notes:* Reduced form estimates are based on pairs of 1,225 locations. The total number of observations is thus 1,499,400.

## 4 Conclusion

Using a general setup with mark-up pricing, in which prices are a function of the distance between producers and consumers, border frictions and the general level of competition, we argued that reduced form estimates based on price differentials between consumption locations lead to biased estimates of distance- and border-related costs. In a quantitative application based on the results of our related work, we found that these bias can be substantial. These insights shows the importance of using disaggregated data—in our case the knowledge of firm locations—to properly control for variation in distance costs and markups. These issues apply to a large range of industries in which specific producers operate in only a few locations and the set of active firms differs across national boundaries.

## References

- Atkin, D., and Donaldson D. (2014) “Who’s Getting Globalized? The Size and Nature of Intranational Trade Costs.” Mimeo.
- Brander, J.A. (1981) “Intra-industry Trade in Identical Commodities.” *Journal of International Economics*, Vol. 11:1.
- Brander, J., and Krugman P. (1983) “A Reciprocal Dumping Model of International Trade.” *Journal of International Economics*, Vol. 15:3.
- Broda, C., and Weinstein D.E. (2008) “Understanding International Price Differences using Barcode Data.” National Bureau of Economic Research working paper w14017.
- Coşar, A. K., Grieco, P. L., and Tintelnot, F. (2014) “Borders, Geography, and Oligopoly: Evidence from the Wind Turbine Industry.” Forthcoming in the *Review of Economics and Statistics*.
- Engel C., Rogers J.H. (1996) “How Wide Is the Border?” *The American Economic Review*, Vol. 86:5.
- Friberg, R., and Ganslandt M. (2008) “Reciprocal Dumping with Product Differentiation.” *Review of International Economics*, Vol. 16:5.
- Gorodnichenko Y., and Tesar L. (2009) “Border Effect or Country Effect? Seattle May Not Be so Far from Vancouver After All.” *American Economic Journal: Macroeconomics*, Vol. 1:1.