Economics Letters 125 (2014) 82–86

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Does distance reflect more than transport costs?

Chi-Young Choi [a,*,1], Horag Choi [b]

[a] University of Texas at Arlington, USA
[b] Monash University, Australia

HIGHLIGHTS

- We assess the effect of distance on intercity retail price dispersion in US cities for various products.
- We decompose the distance effect into two channels, transport costs and non-transport costs.
- Transport costs can explain the spatial dispersion of price mainly for tradable goods.
- Non-transport costs are relevant for both tradable and non-tradable prices.
- Distance contains more information than transport costs.

ARTICLE INFO

Article history:
Received 21 May 2014
Received in revised form
13 August 2014
Accepted 13 August 2014
Available online 23 August 2014

JEL classification:
E31
F15
R12

Keywords:
Distance
Geographic price differences
Transport costs
US cities

ABSTRACT

We decompose the effect of distance on intercity retail price dispersion in US into transport and non-transport cost components. We find that distance contains more information than transport costs. Care should be taken in interpreting distance effect as transport costs only.

1. Introduction

As a popular metric for transport costs, geographic distance has long been recognized as an important factor behind the price differences between locations (e.g., Engel and Rogers, 1996). Standard trade models, for example, typically assume that price difference between locations increases with distance and a large empirical literature has provided evidence that prices are more similar for locations which are geographically proximate (e.g., Crucini et al., 2012). Whereas the conventional literature has interpreted this distance effect as solely reflecting transport costs, distance may induce price wedges between locations via additional channels to transport costs in view of the growing evidence that other factors may also operate on the geographic distance (e.g., Atkin and Donaldson, 2013 and Gopinath et al., 2011). Local distribution costs, for instance, are likely to be more similar between nearby locations if distribution of goods is labor intensive and labor markets are geographically integrated (e.g., Anderson and van Wincoop, 2004 and Engel et al., 2003).

The current study investigates whether and to what extent distance contributes to the explanation of intercity price gaps. To this end, we use retail price data of US cities where transport costs are of central importance in determining price gaps in the absence of any formal barriers to trade. By utilizing the data on inter-spatial trade cost constructed by Allen and Arkolakis (forthcoming), we decompose the distance effect into two underlying sources in a similar spirit with Giri (2012): transport costs (hereafter, TC) versus the remainder that are influenced by distance but independent of TC, dubbed as non-transport costs (hereafter, NTC). The information on

* Correspondence to: Department of Economics, University of Texas at Arlington, Arlington, TX 76019, USA. Tel.: +1 817 272 3860; fax: +1 817 272 3145.
E-mail address: cychoi@uta.edu (C.-Y. Choi).
1 Any remaining errors are the authors'.

http://dx.doi.org/10.1016/j.econlet.2014.08.016
0165-1765/

NTC is extracted from regressing bilateral distance onto observable measures of TC. Given that NTC, including local distribution costs and mark-up rates, is known to constitute a large component of final consumer prices, distinction between the two channels would provide useful insights on understanding spatial price gaps.

Our regression analysis yields compelling evidence of distance effect on price differences for the entire panel of 42 products from 48 cities, after controlling for some key local characteristics like real income and population differences. Distance effect is significant in a broad product category, including services which are traditionally considered as non-tradables. Price differences are larger, more volatile, and disappear more slowly between cities that are farther apart. Both TC and NTC channels have positive effects on price differences, but their significances differ across product groups. While TC channel is important for tradables only, NTC channel is significant in both tradables and non-tradables, indicating that the distance effect found in service prices is driven mainly through the NTC channel. This argument is reinforced by our product level analysis in which we relate the marginal effect of each channel to distribution margin, an inverse measure of tradability constructed by Crucini and Shintani (2008). We note a greater impact of the NTC channel in the products with a higher distribution margin, and hence are less tradable, while the impact of TC channel is larger for more tradable products with a lower distribution margin.

2. Data

Our dataset comprises final retail prices of individual goods and services for selected US cities. The data are complied by the American Chamber of Commerce Researchers Association (ACCRA) and were also used by some other researchers (e.g., Crucini et al., 2012, Parsley and Wei, 1996 and Yazgan and Yilmazkuday, 2011).[2] Our sample covers 48 cities for 42 goods and services between 1985.Q1 and 2009.Q4 (100 quarters), resulting in 47,376 inter-city relative price series (1128 (=48 × 47/2) city pairs for 42 products). As presented in Table 1, 42 products are grouped into three categories: 14 perishable goods, 17 non-perishable goods, and 11 service products. Since our data are absolute prices for specific goods and services, they are well suited for the study of distance's effect on intercity price differences. They allow us to pin down the absolute size of price differences and the subsequent adjustment speed toward long-run equilibrium level. In the regression analysis, we also employ a dataset for city-level per capita real income and population retrieved from various sources.

3. Decomposition of distance and regression analysis

If distance contains more information than transport costs,[3] it can be regressed onto transport costs as follows,

$$\log(\mathrm{DISTANCE}_{ij}) = \alpha + \delta\tau_{ij} + \epsilon_{ij}, \qquad (1)$$

where $\mathrm{DISTANCE}_{ij}$ is the distance between cities i and j measured by the greater circle formula based on city's latitude and longitude data, τ_{ij} represents the marginal transport cost between cities i and j, and the residuals ($\hat{\epsilon}_{ij}$) captures the information on NTC that are influenced by distance but are independent of TC. The intuition behind this orthogonal decomposition is to extract the information on unobservable NTC embedded in distance from observable TC.[4] A related challenge is to measure and infer the value of TC (τ_{ij}) that have previously rendered researchers rely on distance as its proxy. Here we utilize a novel dataset on iceberg trade costs among US counties recently constructed by Allen and Arkolakis (forthcoming). We then extract the information on NTC ($\hat{\epsilon}_{ij}$) by regressing bilateral distance onto their estimates of symmetric iceberg trade costs as $\tau_{ij} = \tau_{ji}$ in (1). We note a high correlation between log distance and τ_{ij} with the simple correlation coefficient exceeding 0.85.

We evaluate the effects of distance and its two components on various measures of intercity price differentials using the following equations,

$$y^k_{ij} = \rho \log(\mathrm{DISTANCE}_{ij}) + X\beta + \varepsilon_{ij}, \qquad \text{(Regression 1)} \qquad (2)$$

$$y^k_{ij} = \alpha_1\hat{\epsilon}_{ij} + \alpha_2\tau_{ij} + \alpha_3\tau^2_{ij} + X\beta + \varepsilon_{ij}, \qquad \text{(Regression 2)} \qquad (3)$$

where y^k_{ij} denotes price differentials for city-pair i and j for product k for which we consider three different measures: (i) long-run average; (ii) persistence; and (iii) volatility.[5] To control for additional potential determinants of intercity price differences, we include a set of explanatory variables, $X = \{RINCOME_{ij}, POP_{ij}, SameState_{ij}, D^P_k, D^C_i, D^C_j\}$, where 'RINCOME', and 'POP' respectively denote city-pair differences in real per capita income and population computed by $[\max(z_i, z_j) - \min(z_i, z_j)]/\max(z_i, z_j)$ in which z_k denotes the corresponding variable for city k. Real income is known to have a positive effect on price levels as stipulated in the notion of pricing-to-market. Population difference is to capture relative city-size which is a significant determinant of the intercity price gaps. '*SameState*' represents an intra-state dummy variable which takes on the value of one if two cities are in the same state and zero otherwise. It controls for state-specific characteristics like state-tax and policy environment and hence it is expected to enter with a negative sign because cities in the same state are likely to have similar price levels with more homogeneous tax schemes and economic environments (e.g., industrial structure). D^P_k denotes product-specific dummies and D^C_h denotes city-specific dummies which are to capture any idiosyncratic aspects of the price of a given city.[6] Notice that the regressor of $\hat{\epsilon}_{ij}$ in Regression 2 is a residual and thus it is susceptible to the so-called generated regressor problem that OLS-based standard errors are invalid (e.g., Yilmazkuday, 2012). As noted by Pagan (1984, p. 242), however, standard inference is still valid if unlagged residuals are used in a regression as in our case. It is also worth noting that the squared term of TC (τ^2_{ij}) is included in Regression 2 to capture a nonlinear quadratic effect of transport costs on price dispersion along the lines of Engel and Rogers (1996).

Table 2 presents the regression results for the full sample and three sub-samples of product categories. As reported in the left-hand panel (columns 2–5), the results of Regression 1 clearly indicate that distance is highly significant in all regressions with an expected positive sign, after controlling for real income, population, and state border effect. This confirms the common wisdom that prices are more disparate, and price differences are more volatile and more persistent for the city-pairs farther apart. The size of distance effect, however, differs across product categories, with the largest effect in perishables and the smallest in services.

2 Because the price data have occasional missing observations due to frequent revisions in the coverage of cities and products, we drop any series that have missing observations for more than two consecutive quarters.

3 We maintain this on a couple of empirical grounds. First, TC is highly significant when it is run against price gaps, but its explanatory power disappears once we control for distance. Second, the residuals obtained from running TC onto distance turn out to have little explanatory power on spatial price differences.

4 This decomposition approach shares a similar spirit with Giri (2012) who shows that cross-country dispersion in prices of goods can be explained by two sources: (i) trade costs and (ii) non-traded input costs of distribution.

5 Long-term average ($\alpha/(1-\rho)$) and persistence (ρ) of price differences are estimated in a linear AR(p) model, $q_{ij,t} = \alpha + \rho q_{ij,t-1} + \sum_{h=1} k\Delta q_{ij,t-h} + \varepsilon_{ij,t}$ where q_{it} denotes the (log) price differential between cities i and j at time t.

6 The city-specific dummies cannot fully capture the segmented labor markets effect when goods are heterogeneous as is often the case because they simply take out the mean effects across goods only.