



Does distance reflect more than transport costs?



Chi-Young Choi^{a,*}, 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.

© 2014 Elsevier B.V. All rights reserved.

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'.

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 compiled 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,

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

where 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 (ϵ_{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,

$$y_{ij}^k = \rho \log(\text{DISTANCE}_{ij}) + X\beta + \epsilon_{ij}, \quad (\text{Regression 1}) \quad (2)$$

$$y_{ij}^k = \alpha_1 \hat{\epsilon}_{ij} + \alpha_2 \tau_{ij} + \alpha_3 \tau_{ij}^2 + X\beta + \epsilon_{ij}, \quad (\text{Regression 2}) \quad (3)$$

where y_{ij}^k 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, $X = \{\text{RINCOME}_{ij}, \text{POP}_{ij}, \text{SameState}_{ij}, D_k^p, D_i^c, D_j^c\}$, 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_k^p denotes product-specific dummies and D_h^c 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_{ij}^2) 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, $q_{ij,t} = \alpha + \rho q_{ij,t-1} + \sum_{h=1}^k k \Delta q_{ij,t-h} + \epsilon_{ij,t}$ 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.

Table 1
Product categories.

Category	Products
Perishables (14)	Steak, Ground beef, Whole chicken, Milk, Eggs, Margarine, Cheese, Potatoes, Bananas, Lettuce, Bread, McDonald Hamburger, Pizza, Fried chicken
Non-perishables (17)	Canned tuna, Coffee, Sugar, Corn flakes, Canned peas, Canned peaches, Tissue, Detergent, Shortening, Frozen corn, Soft drink, Gas, Toothpaste, Man's shirt, Tennis balls, Beer, Wine
Services (11)	Auto maintenance, Doctor visit, Dentist visit, Man's haircut, Beauty salon, Dry cleaning, Newspaper, Appliance repair, Movie, Bowling, Telephone

Table 2
Distance effect on intercity price differences.

Sample	Regression 1				Regression 2			
	Regressor	Gap	Pers	Vol	Regressor	Gap	Pers	Vol
Full {47,376}	log(DISTANCE)	0.005***	0.017***	0.005***	NTC ($\hat{\epsilon}_{ij}$)	0.003***	0.016***	0.004***
					TC (τ_{ij})	0.055***	0.447***	0.089***
					TC-SQUARED (τ_{ij}^2)	-0.011***	-0.118***	-0.022***
	RINCOME	0.033*	0.087	0.019	RINCOME	0.036*	0.092	0.020
	CITY SIZE	0.018***	0.086***	0.014***	CITY SIZE	0.017***	0.084***	0.013***
	SAME STATE	-0.003***	-0.058***	-0.013***	SAME STATE	-0.010**	-0.011	-0.004*
	Adj-R ²	0.184	0.227	0.542	Adj-R ²	0.185	0.227	0.543
Perishable {15,792}	log(DISTANCE)	0.009***	0.023***	0.006***	NTC ($\hat{\epsilon}_{ij}$)	0.005***	0.021***	0.005***
					TC (τ_{ij})	0.093***	0.268*	0.075***
					TC-SQUARED (τ_{ij}^2)	-0.017**	-0.060	-0.016***
	RINCOME	0.048	-0.001	0.021	RINCOME	0.054	0.003	0.023
	CITY SIZE	0.002*	0.135***	0.015*	CITY SIZE	0.000	0.133***	0.015*
	SAME STATE	-0.003***	-0.076***	-0.018***	SAME STATE	-0.016***	-0.050***	-0.010***
	Adj-R ²	0.236	0.186	0.707	Adj-R ²	0.238	0.186	0.707
Non-Perishable {19,176}	log(DISTANCE)	0.004***	0.021***	0.005***	NTC ($\hat{\epsilon}_{ij}$)	0.003***	0.021***	0.004***
					TC (τ_{ij})	0.014*	0.778***	0.131***
					TC-SQUARED (τ_{ij}^2)	0.000	-0.213***	-0.035***
	RINCOME	0.018	0.257	0.024	RINCOME	0.017	0.263	0.025
	CITY SIZE	0.013***	0.053	0.010	CITY SIZE	0.013***	0.051	0.010
	SAME STATE	-0.005***	-0.086***	-0.016***	SAME STATE	-0.007***	-0.004	-0.002
	Adj-R ²	0.150	0.138	0.550	Adj-R ²	0.150	0.139	0.551
Services {12,408}	log(DISTANCE)	0.002***	0.006***	0.002***	NTC ($\hat{\epsilon}_{ij}$)	0.001**	0.004*	0.002**
					TC (τ_{ij})	0.046	0.234	0.046
					TC-SQUARED (τ_{ij}^2)	-0.012	-0.063	-0.011
	RINCOME	0.096***	-0.121	0.007	RINCOME	0.097***	-0.116	0.008
	CITY SIZE	0.039***	0.090***	0.021*	CITY SIZE	0.039***	0.089***	0.021*
	SAME STATE	0.000	-0.010	-0.004	SAME STATE	0.005	0.018	0.001
	Adj-R ²	0.089	0.111	0.349	Adj-R ²	0.089	0.111	0.349

Note: See Eqs. (2)–(3) for regression equations. 'Gap' denotes long-term average city-pair price differentials, 'Pers' is persistence of log price differences, and 'Vol' represents temporal volatility of price difference measured by standard deviation. Numbers in the curved bracket represent the number of observations in each regression.

* Indicate statistical significance at the 10% error levels and heteroskedasticity robust standard errors are used.

** Indicate statistical significance at the 5% error levels and heteroskedasticity robust standard errors are used.

*** Indicate statistical significance at the 1% error levels and heteroskedasticity robust standard errors are used.

The marginal effect of distance on the volatility of intercity relative price, for example, is in the range of 0.002 (services) and 0.006 (perishables), implying that a one percent increase in distance between cities leads to an increment of volatility by 0.002–0.006 on average. Interestingly, distance effect is significant for service products which are conventionally recognized as non-tradables.

When we break down distance into TC and NTC components, our analysis produces an intriguing result. As reported in the last four columns of Table 2, while both TC and NTC are significant for tradable products, only NTC is significant for service products. This may indicate that the significant distance effect observed in service price differences is driven through the NTC channel. Although TC is indeed at the root of the distance effect in tradable products as has long been viewed in the literature, it cannot account for the 'large and persistent' price dispersion observed in service products. The fact that only NTC is a significant explainer of price differences for services most likely can be accounted for by geographic segmentation of labor market. Topel (1994), for example, maintains that the extent of labor markets is limited by geography and relative wages are greatly affected by local factors. In a similar context, Engel et al.

(2003) document that local distribution costs are liable to be more similar between nearby locations when distribution of goods is labor intensive.

As a further investigation on this issue, we carry out a product level analysis by relating the marginal effects of TC and NTC on price gaps to *distribution margin* which was featured in the literature exploring relative price dynamics (e.g., Crucini and Shintani, 2008).⁷ The basic idea is that products with different distribution margins may have different responses to TC and NTC. It is conceivable that products with a higher distribution margin, and hence are less tradable, would respond more strongly to NTC channel, but less to TC channel, and vice versa. The result is illustrated in Fig. 1, which plots the degree of distribution margins (on the horizontal axis) against the marginal effects of TC and NTC (on the vertical axis) on long-term average (top row), persistence (middle

⁷ As a proxy for the relative cost of non-traded inputs faced by retailers between locations, distribution margin measures the gap between retail and producer prices, including NTC such as wholesale and retail margins. It is also interpreted as a degree of non-tradability, ranging from 0 (fully tradable) and 1 (non-tradable).

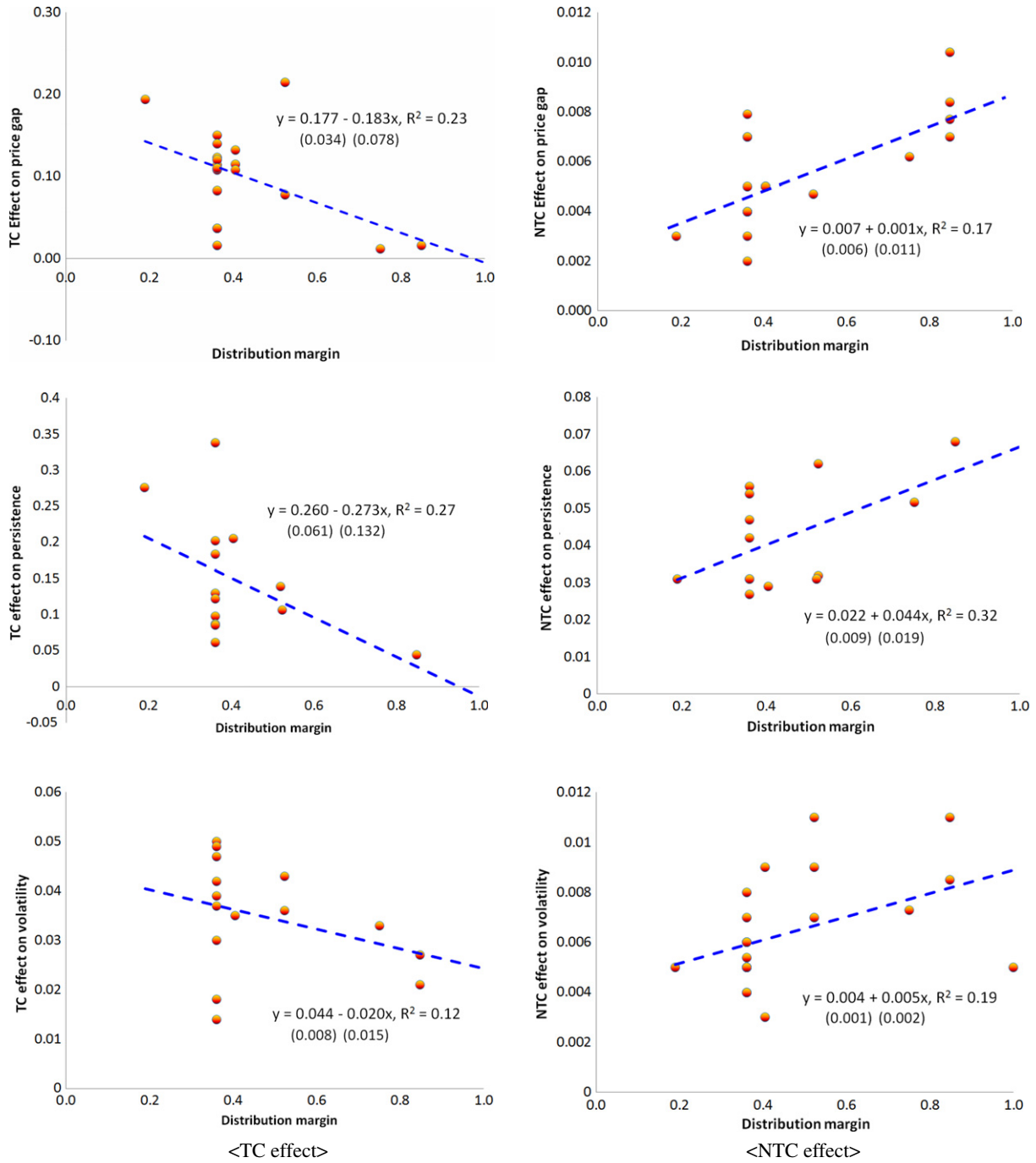


Fig. 1. Marginal effect (on the vertical axis) and distribution margin (on the horizontal axis).

row), and volatility (bottom row) of intercity price differences.⁸ A visual inspection of the scatterplots reveals a clear inverse association of distribution margin with TC, but a positive relationship with NTC. That is, products with a higher distribution margins, which are less tradable, experience a weaker effect of TC, but a stronger effect of NTC, on the long-term price disparities, and persistence and volatility of intercity price gaps, consistent with our prior intuition. Overall, our findings highlight the heterogeneous effects of TC and NTC components of distance on cross-city price differentials and support our view that distance reflects more than TC.

⁸ In this exercise, we consider only the products in which marginal effects are statistically significant.

4. Concluding remarks

Although distance is generally thought to affect price differences mainly through transport cost channel, there could be other channels through which it operates on spatial price differences in view of the growing body of evidence that mark-ups and other related factors also systematically vary with distance (e.g., Atkin and Donaldson, 2013). In this study, we decompose distance effect into TC and NTC by utilizing a novel dataset of trade cost constructed by Allen and Arkolakis (forthcoming). Our empirical analysis using US retail prices convincingly suggests that distance affects intercity price differences not just through TC channel but also through NTC channel. In fact, distance induces price wedges in service products mainly through the NTC channel, probably via local distribution costs and labor market segmentation. Distance effect therefore

should not be solely credited to transport costs as often assumed in the literature.

References

- Allen, T., Arkolakis, C., 2014. Trade and the topography of the spatial economy. *Quart. J. Econ.* (forthcoming).
- Anderson, J.E., van Wincoop, E., 2004. Trade costs. *J. Econom. Lit.* XLII, 691–751.
- Atkin, D., Donaldson, D., 2013. Who's getting globalized? The size and nature of Intranational trade costs. Mimeo, Yale University.
- Crucini, M.J., Shintani, M., 2008. Persistence in law of one price deviations: evidence from micro-data. *J. Monetary Econ.* 55, 629–644.
- Crucini, M.J., Shintani, M., Tsuruga, T., 2012. Noisy information, distance and law of one price dynamics across US cities. NBER Working Papers No. 17815, National Bureau of Economic Research.
- Engel, C., Rogers, J.H., 1996. How wide is the border? *Amer. Econ. Rev.* 86 (5), 1112–1125.
- Engel, C., Rogers, J.H., Wang, S.-Y., 2003. Revisiting the border: an assessment of the law of one price using very disaggregated consumer price data. *International Finance Discussion Papers No. 777*, Board of Governors of the Federal Reserve System.
- Giri, R., 2012. Local costs of distribution, international trade costs and micro evidence on the law of one price. *J. Int. Econ.* 86, 82–100.
- Gopinath, G., Gourinchas, P.-O., Hsieh, C.-T., Li, N., 2011. International prices, costs, and markup differences. *Amer. Econ. Rev.* 101, 2450–2486.
- Pagan, A., 1984. Econometric issues in the analysis of regressions with generated regressors. *Internat. Econom. Rev.* 25 (1), 221–247.
- Parsley, D.C., Wei, S.-J., 1996. Convergence to the law of one price without trade barriers or currency fluctuations. *Quart. J. Econ.* 111, 1211–1236.
- Topel, R.H., 1994. Regional labor markets and the determinants of wage inequality. *Amer. Econ. Rev.* 84 (2), 17–22.
- Yazgan, M.E., Yilmazkuday, H., 2011. Price-level convergence: new evidence from US cities? *Econom. Lett.* 110, 76–78.
- Yilmazkuday, H., 2012. Understanding interstate trade patterns. *J. Int. Econ.* 86, 158–166.