Economics Bulletin

Volume 44, Issue 4

Distance as a determinant of trade costs: a different type of distance puzzle?

Parul Khokher
San Bernardino Valley College

David Buehler
Penn State University

Roger White
Embry-Riddle Aeronautical University

Abstract

Employing a gravity model approach to examine the determinants of annual bilateral trade costs for the period 1995-2018, we present evidence of a novel "trade cost-distance puzzle." Previous studies have identified the existence of a traditional "distance puzzle", sometimes referred to as the "missing globalization puzzle", that notes an increasingly negative influence of distance on bilateral trade flows across time periods. This paper extends the traditional "distance puzzle" from the trade costs perspective. After controlling for variables widely considered as determinants of trade costs, we demonstrate that while distance, as expected, is positively related to trade costs, the influence of distance on trade costs increases over time. This finding is contrary to the intuition regarding the effects of globalization on international trade. As a second contribution, we propose a partial solution to the puzzle. Applying the Poisson pseudo-maximum likelihood (PPML) estimation technique to balanced panel data sets resolves the puzzle for total trade costs and manufacturing trade costs; however, the puzzle persists for agricultural trade costs.

Citation: Parul Khokher and David Buehler and Roger White, (2024) "Distance as a determinant of trade costs: a different type of distance puzzle?", *Economics Bulletin*, Volume 44, Issue 4, pages 1560-1572
Contact: Parul Khokher - pkhokher@valleycollege.edu, David Buehler - dlb74@psu.edu, Roger White - rwhite@erau.edu.
Submitted: July 09, 2024. **Published:** December 30, 2024.

Economics Bulletin

Volume 44, Issue 4

Distance as a determinant of trade costs: a different type of distance puzzle?

Parul Khokher
San Bernardino Valley College

David Buehler
Penn State University

Roger White
Embry-Riddle Aeronautical University

Abstract

Employing a gravity model approach to examine the determinants of annual bilateral trade costs for the period 1995-2018, we present evidence of a novel "trade cost-distance puzzle." Previous studies have identified the existence of a traditional "distance puzzle", sometimes referred to as the "missing globalization puzzle", that notes an increasingly negative influence of distance on bilateral trade flows across time periods. This paper extends the traditional "distance puzzle" from the trade costs perspective. After controlling for variables widely considered as determinants of trade costs, we demonstrate that while distance, as expected, is positively related to trade costs, the influence of distance on trade costs increases over time. This finding is contrary to the intuition regarding the effects of globalization on international trade. As a second contribution, we propose a partial solution to the puzzle. Applying the Poisson pseudo-maximum likelihood (PPML) estimation technique to balanced panel data sets resolves the puzzle for total trade costs and manufacturing trade costs; however, the puzzle persists for agricultural trade costs.

Citation: Parul Khokher and David Buehler and Roger White, (2024) "Distance as a determinant of trade costs: a different type of distance puzzle?", *Economics Bulletin*, Volume 44, Issue 4, pages 1560-1572
Contact: Parul Khokher - pkhokher@valleycollege.edu, David Buehler - dlb74@psu.edu, Roger White - rwhite@erau.edu.
Submitted: July 09, 2024. **Published:** December 30, 2024.

1. Introduction

It is well-established in the international trade literature that physical distance is a robust proxy for transportation costs and, more generally, for international trade costs (Anderson and van Wincoop, 2004). Although the phenomenon of globalization has long been expected to diminish the extent to which distance impedes international trade, several studies that employ the empirical gravity equation have documented unexpected, increasingly negative influences of geographical distance on trade flows over time. This finding has been labeled the “distance puzzle” (Disdier and Head, 2008) and the “missing globalization puzzle” (Coe, et al., 2002). Since the puzzle was first identified, several works have posited solutions, achieving varying degrees of success.

To address the traditional distance puzzle, Yotov (2012) employs a two-step approach that integrates international and domestic trade within conventional gravity estimations and reveals declining differences, over time, between international and domestic distance elasticity estimates. Building on this, Bergstrand et al. (2013) employs a structural gravity specification that incorporates time-varying direction-specific country fixed effects in both domestic and international panel trade data. More recent works that also offer potential solutions to the puzzle include Buehler and White (2015), Borchert and Yotov (2017), Yilmazkuday (2017), and Brei and von Peter (2018). Collectively, these works shed light on the complex interplay between globalization, trade flows, and geographical distance.

This paper closely relates to the literature on the “distance puzzle” and trade costs. The determinants of trade costs have been a popular topic, with Chen and Novy (2009), Arvis, et al. (2013), and Gervais (2019), among others, examining the determinants of trade costs. Kaminchia (2019), for instance, examines sector- and industry-level data to assess the relationship between bilateral trade costs and economic integration in East Africa while considering the geodesic distance between trading partners, among other traditional control variables. Marti and Puertas (2019) consider the determinants of trade costs for eight European exporters, considering geodesic distance while also including entry costs, bilateral exchange rates, and shared border, among other explanatory variables. Hou et al. (2020) find that institutional quality significantly lowers trade costs. Ali and Milner (2022) examine the relationship between trade costs and the composition of trade among partners. In all instances, positive and statistically significant coefficient estimates are reported for the distance variables, yet variation in the effects of distance on trade costs over time has not been considered.

Analyzing potential variation in the trade costs-distance relationship is one of the main contributions of this paper. The “distance puzzle” and trade costs both are important issues in international trade and yet, to our best knowledge, no previous study has made a close connection between them. This paper extends the traditional “distance puzzle” from the trade costs perspective. To this end, our dependent variable series are measures of three categories of bilateral trade costs rather than bilateral trade flows.

Offering an alternative to geographical distance as a measure of trade costs, Novy (2009, 2013) estimates bilateral international trade costs using the inverse gravity model. Measured as an *ad valorem* tariff-equivalent increase in international trade costs relative to domestic, or internal, trade costs, the measure includes the impact of geographical distance on international trade flows.[1] Because international trade costs represent all costs associated with the exchange of goods across national borders, it is intuitive that, over time, the influence of geographical distance on bilateral trade costs may decline or, perhaps, remain constant.

Since understanding the determinants of trade costs may help to inform public policy actions that enhance welfare by reducing such costs and increasing trade flows, accurate measurement of the influence of distance on trade costs is essential. We investigate whether a trade cost-distance puzzle exists, providing a detailed analysis of the evolving influence of distance on three categories of trade costs: total, agricultural, and manufacturing. More specifically, we examine the determinants of annual bilateral trade costs during the period from 1995-2018.

Our empirical approach involves regressing our three measures of bilateral trade costs, in turn, on a set of variables commonly used as proxies for trade costs in gravity models of trade. This permits the isolation of the influence of geographical distance on trade costs while allowing for variation in the effect over time. Following each of these estimations, the year-specific distance coefficients are regressed on a time variable to determine whether the influence of geodesic distance changes significantly over our reference period. These initial estimations allow us to determine whether a trade costs distance puzzle exists.

The contribution of this paper is twofold. First, through the application of the Ordinary Least Squares (OLS) technique, we demonstrate the existence of a novel trade cost-distance puzzle. More specifically, we find that, even though trade costs have generally fallen over time (Novy, 2013), the magnitudes of year-specific coefficients reveal an increasing influence of geographic distance on each trade cost measure considered. Second, following Buehler and White (2015), we use the Poisson pseudo-maximum likelihood (PPML) estimation technique as a potential solution to the puzzle.[2] We find that the PPML technique, when used in conjunction with balanced data sets (i.e., all country pairs for which trade costs data are available for all years in our reference period), resolves the puzzle for total trade costs and manufacturing trade costs.[3] Even so, the puzzle remains unresolved for agricultural trade costs.

2. Empirical Model, Descriptive Statistics, and Estimation Strategy

2.1 Econometric Specification

To examine whether a trade cost-distance puzzle exists, as noted, we estimate equation (1) while considering three dependent variable series: total trade costs, agricultural trade costs, and

---

[1] The trade cost measure, as described by Novy (2013) "is a function of observable trade data and can therefore be calculated easily with time series and panel data to track the changes of trade costs over time" (pg. 116).

[2] While trade cost data are generally available for 195 countries, estimates of total, agricultural, and manufacturing trade costs are not available for all country pairs and all years. The unbalanced datasets we examine include all country pairs for which trade costs are available for any of the years 1995-2018. The balanced datasets we examine include all country pairs for which trade costs are available for all years during the reference period.

[3] Results presented in this paper are generated from examinations of balanced panel data sets. Repeating our analysis using unbalanced data sets also results in the identification of a trade cost-distance puzzle. While, due to space limitations, we do not report the "unbalanced results" here, these additional results are included in an online appendix (i.e., as Supplemental Material).

manufacturing trade costs.[4] The dependent variable, TC_{ijt} is an *ad valorem* tariff-equivalent comprehensive measure of bilateral trade costs derived using the inverse gravity framework (Chen and Novy, 2009). The measure includes all trade costs, broadly defined to represent international transport costs and tariffs, along with other trade cost components discussed in Anderson and van Wincoop (2004) (e.g., direct and indirect costs associated with language differences, currencies, and import/export procedures). Trade costs data are from UNESCAP via the World Bank (2014).

$$TC_{ijt} = \alpha_0 + \beta_{\Phi_t}(lnD_{ij} \times \Phi_t) - \beta_1 BORDER_{ij} - \beta_2 COMLANG_{ij} - \beta_3 COMCOL_{ij} - \beta_4 COLONY_{ij} - \beta_5 RTA_{ijt} + \beta_{\Upsilon_{it}} \Upsilon_{it} + \beta_{\Psi_{jt}} \Psi_{jt} + \varepsilon_{ijt} \quad (1)$$

Our primary variables of interest are those that interact geodesic distance between trading partners (lnD_{ij}) with year-specific dummy variables (Φ_t).[5] Our *a priori* expectation is that the corresponding coefficients (i.e., the β_{Φ_t} values) will be positive and statistically significant; however, given the topic at hand, our interest lies in whether the distance coefficients vary and, specifically, whether they increase over time. A pattern of increasing magnitudes for the year-specific distance coefficients will be considered evidence that supports the existence of a trade cost-distance puzzle.

In equation (1), we include controls that have commonly been employed in prior studies to represent aspects of trade costs. These variables are all from CEPII and includes dummies that indicate whether countries are adjacent $(BORDER_{ij})$, share a common official language(s) $(COMLANG_{ij})$, were colonized by the same country $(COMCOL_{ij})$, were both colonized but by different countries $(COLONY_{ij})$, or are members of one or more of the same multilateral trade agreements (RTA_{ijt}). For each of these variables, the *a priori* expected signs are indicated in the equation.

Consistent with the results demonstrated by Head and Mayer (2014) regarding the need for multilateral resistance terms in the traditional gravity setting, inclusion of such terms is also necessary in this model with trade costs as the dependent variable.[6] The fixed effects absorb time-varying country-specific determinants of bilateral trade costs; thus, all explanatory variables that are explicit in equation (1) are country pair-specific. The subscripts *i* and *j* represent the reporter and partner countries, respectively, *t* represents time (i.e., year), and ε_{ijt} is an assumed i.i.d. error term.

Figure 1 illustrates average annual values for our dependent variable series. We see the average annual value of each trade cost measure declines, rather consistently, across the

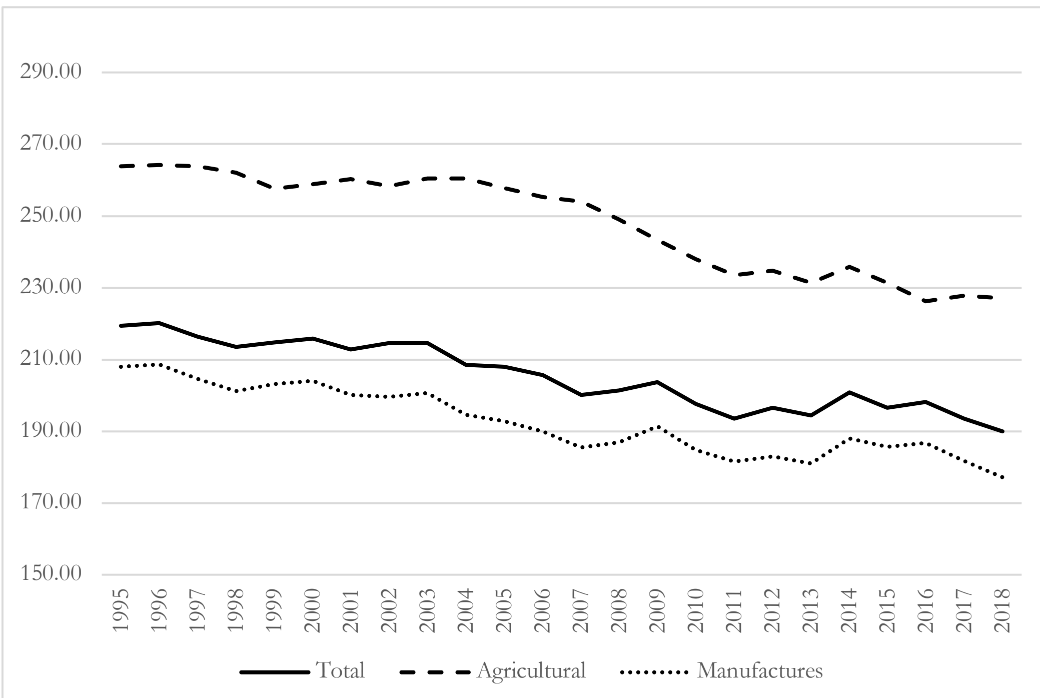
[4] In Equation (1), we present the dependent variable as TC_{ijt}, which includes our three measures of trade costs. However, when performing our estimations using OLS, our dependent variable series are log-transformed. This transformation is not performed when we employ the PPML estimation technique.

[5] The measure of geodesic distance we employ is the *DIST* variable from the CEPII Gravity Database (2021). It is calculated following the great circle formula, using the latitudes and longitudes of the most important cities/agglomerations (in terms of population).

[6] To control for multilateral resistances and for consistent estimation of the distance coefficient, we include time-varying fixed effects for both reporter (Υ_{it}) and partner (Ψ_{jt}) countries (Baldwin and Taglioni (2006) and Head and Mayer (2014)).

reference period. We also see that agricultural trade costs are considerably higher than total trade costs or manufacturing trade costs. For instance, the average agricultural trade cost value in 2018 is equal to 227.11, while the corresponding values for total and manufacturing trade costs are 189.95 and 177.17, respectively. This means that in 2018 in the agriculture sector, international trade is 2.27 times more costly than domestic trade while in total and in the manufacturing sector, international trade is 1.89 and 1.77, respectively, times more costly than domestic trade.[7] It is also evident in the figure that the mean total trade cost and manufacturing trade cost series are highly correlated. In fact, the correlation coefficient for the total trade cost and manufacturing trade costs series is 0.99.

Figure 1: Average Trade Costs, 1995-2018



2.2 Descriptive Statistics

Descriptive statistics are presented in Table 1. In column (a), we see the mean trade cost value is equal to 205.43. This indicates that, on average during our reference period, international trade was about 2.05 times more costly than domestic trade. The values in Figure 1 illustrate the time paths of our three trade costs measures. Given the high pairwise correlation between total trade costs and manufacturing trade costs, it is perhaps unsurprising that the mean values for these samples are less dissimilar as compared to the total and agricultural trade costs samples.

[7] Trade costs for agricultural goods are about four-fifths higher than for manufacturing goods. Agricultural trade costs declined less than trade costs in manufacturing over the period 1995 to 2018, in part because of slower progress in tariff reductions and narrower coverage of trade agreements (World Bank, 2021).

Turning attention to the explanatory variables, we find considerable differences between mean values across samples; however, these differences are the result of variation in the compositions of the samples for which trade cost data are available. Evident in the number of observations (n), total trade costs and manufactures trade costs data are available for larger (and different) numbers of countries as compared to agricultural trade costs. Therefore, inferences made from interpretation of the differences in mean values has significant limitations.

Table I: Descriptive Statistics

	Total	Agricultural	Manufacturing
n =	67,104	46,848	59,424
	(a)	(b)	(c)
Trade Costs$_{ijt}$	205.43	247.9	192.47
	(107.7)	(117.63)	(98.79)
Distance$_{ij}$	7,128.41	6,087.40	7,265.12
	(4,522.54)	(4,896.56)	(4,563.46)
Common Border$_{ij}$	0.0401	0.0686	0.042
	(0.3365)	(0.2529)	(0.2006)
Common Language$_{ij}$	0.1302	0.1363	0.1284
	(0.3365)	(0.3431)	(0.3346)
Common Colonizer Post-1945$_{ij}$	0.0286	0.0287	0.0258
	(0.1667)	(0.1669)	(0.1587)
Colonial Relationship$_{ij}$	0.0343	0.0523	0.0331
	(0.1821)	(0.2225)	(0.1789)
Regional Trade Agreement$_{ij}$	0.2732	0.4227	0.2897
	(0.4456)	(0.0494)	(0.4536)

Standard deviations in parentheses.

2.3 Estimation Strategy

We first estimate equation (1) using the OLS estimation technique and each of our three measures of bilateral trade costs in turn. To determine whether the influence of geodesic distance changes significantly over our reference period, following each estimation, we regress the year-specific distance coefficients on a time variable. These estimations allow us to determine whether a trade cost-distance puzzle exists.

We continue our analysis by repeating the steps described thus far with the sole difference being that we use the PPML estimation technique rather than OLS. Following each estimation, we again regress the year-specific distance coefficients on a time variable to discern whether a trade cost-distance puzzle exists. Thus, we consider, for each measure of trade costs, the possibility of a trade cost-distance puzzle while using the OLS and PPML estimation techniques.

3. Results

We begin our analysis by demonstrating the existence of the novel trade costs-distance puzzle. Results obtained when the OLS regression technique is used to estimate equation (1) are presented in columns (a) through (c) of Table II. For each dependent variable series, the coefficient estimates of the year-specific distance variables coefficient values are positive and individually statistically significant from zero. More importantly for our topic, we find that the coefficient magnitudes generally increase over our reference period. For instance, the magnitudes of distance coefficients presented in column (a) (i.e., when OLS regression is applied to total trade costs) increase from 0.2576 in 1995 to 0.3094 in 2018. As anticipated, trade costs are lower among country pairs that share a common border, have a common official language, were colonized by the same country, were both colonized but by different countries, or are parties to one or more common preferential trade agreements.

Table II: OLS and PPML Estimations

	OLS Estimations			*PPML Estimations*		
	Total	Agricult.	Manufact.	Total	Agricult.	Manufact.
	(a)	(b)	(c)	(d)	(e)	(f)
ln $Distance_{ij}$ x 1995	0.2576***	0.2513***	0.2866***	0.27***	0.2578***	0.2994***
	(0.0078)	(0.0096)	(0.008)	(0.0083)	(0.0101)	(0.0081)
ln $Distance_{ij}$ x 1996	0.259***	0.2533***	0.2874***	0.2679***	0.2558***	0.2939***
	(0.0077)	(0.009)	(0.0085)	(0.0081)	(0.0093)	(0.0081)
ln $Distance_{ij}$ x 1997	0.2664***	0.2645***	0.2886***	0.2801***	0.2638***	0.295***
	(0.0077)	(0.0092)	(0.0084)	(0.0082)	(0.0098)	(0.0083)
ln $Distance_{ij}$ x 1998	0.2726***	0.2743***	0.2924***	0.2822***	0.2769***	0.2956***
	(0.0075)	(0.0092)	(0.0082)	(0.0082)	(0.0098)	(0.0084)
ln $Distance_{ij}$ x 1999	0.2738***	0.2597***	0.2969***	0.2835***	0.259***	0.3021***
	(0.0075)	(0.009)	(0.008)	(0.0085)	(0.0094)	(0.0085)
ln $Distance_{ij}$ x 2000	0.285***	0.267***	0.3019***	0.2946***	0.2686***	0.3057***
	(0.0076)	(0.0095)	(0.0079)	(0.0086)	(0.01)	(0.0087)
ln $Distance_{ij}$ x 2001	0.2801***	0.2673***	0.2986***	0.2837***	0.2713***	0.2979***
	(0.0078)	(0.0096)	(0.0081)	(0.0088)	(0.0103)	(0.0088)
ln $Distance_{ij}$ x 2002	0.2802***	0.2676***	0.2962***	0.2825***	0.2683***	0.2901***
	(0.0082)	(0.0097)	(0.0085)	(0.0106)	(0.0098)	(0.01)
ln $Distance_{ij}$ x 2003	0.2818***	0.2756***	0.3001***	0.2801***	0.279***	0.2947***
	(0.0081)	(0.0098)	(0.0084)	(0.009)	(0.0105)	(0.0089)
ln $Distance_{ij}$ x 2004	0.2879***	0.2863***	0.3034***	0.2885***	0.2888***	0.2961***
	(0.0081)	(0.0094)	(0.0085)	(0.009)	(0.0101)	(0.009)
ln $Distance_{ij}$ x 2005	0.2892***	0.291***	0.3011***	0.2884***	0.2987***	0.29***
	(0.0082)	(0.0102)	(0.0085)	(0.0093)	(0.0112)	(0.0093)
ln $Distance_{ij}$ x 2006	0.2916***	0.2969***	0.3055***	0.2928***	0.3032***	0.3008***
	(0.0078)	(0.01)	(0.0079)	(0.009)	(0.011)	(0.0082)

ln $Distance_{ij}$ x 2007	0.2849***	0.2996***	0.297***	0.285***	0.3002***	0.2882***
	(0.0079)	(0.0101)	(0.0082)	(0.0089)	(0.0123)	(0.0084)
ln $Distance_{ij}$ x 2008	0.2851***	0.3046***	0.2974***	0.2833***	0.3123***	0.2897***
	(0.008)	(0.0098)	(0.0083)	(0.0092)	(0.0104)	(0.0087)
ln $Distance_{ij}$ x 2009	0.2825***	0.3278***	0.2964***	0.2854***	0.3306***	0.2965***
	(0.008)	(0.0104)	(0.0083)	(0.0091)	(0.0107)	(0.0087)
ln $Distance_{ij}$ x 2010	0.284***	0.3217***	0.2981***	0.2837***	0.3275***	0.2943***
	(0.008)	(0.0111)	(0.0083)	(0.0091)	(0.0116)	(0.0087)
ln $Distance_{ij}$ x 2011	0.2794***	0.3216***	0.296***	0.2707***	0.3223***	0.2821***
	(0.0082)	(0.0108)	(0.0086)	(0.0099)	(0.011)	(0.0101)
ln $Distance_{ij}$ x 2012	0.2824***	0.3357***	0.2996***	0.2776***	0.3393***	0.2949***
	(0.0081)	(0.0113)	(0.0084)	(0.0094)	(0.0116)	(0.0086)
ln $Distance_{ij}$ x 2013	0.2845***	0.3329***	0.3029***	0.2707***	0.3398***	0.2874***
	(0.0083)	(0.0117)	(0.0085)	(0.0097)	(0.0123)	(0.0085)
ln $Distance_{ij}$ x 2014	0.2824***	0.3275***	0.3009***	0.272***	0.3341***	0.2884***
	(0.0082)	(0.0109)	(0.0083)	(0.0093)	(0.011)	(0.0085)
ln $Distance_{ij}$ x 2015	0.28***	0.3265***	0.3016***	0.2696***	0.3332***	0.2926***
	(0.0087)	(0.0112)	(0.0089)	(0.0091)	(0.0119)	(0.009)
ln $Distance_{ij}$ x 2016	0.2797***	0.324***	0.2966***	0.2673***	0.3238***	0.2892***
	(0.0087)	(0.0107)	(0.0089)	(0.0088)	(0.0111)	(0.0091)
ln $Distance_{ij}$ x 2017	0.29***	0.3175***	0.3083***	0.2752***	0.3166***	0.296***
	(0.0087)	(0.0102)	(0.0089)	(0.0089)	(0.0108)	(0.0088)
ln $Distance_{ij}$ x 2018	0.3094***	0.328***	0.3405***	0.2912***	0.3257***	0.3204***
	(0.0094)	(0.0101)	(0.0107)	(0.0088)	(0.0105)	(0.0096)
Common $Border_{ij}$	-0.21***	-0.2213***	-0.1608***	-0.2784***	-0.2303***	-0.2208***
	(0.0077)	(0.0076)	(0.0079)	(0.009)	(0.0083)	(0.0093)
Common $Language_{ij}$	-0.1551***	-0.0707***	-0.1852***	-0.1537***	-0.0574***	-0.1995***
	(0.0042)	(0.0057)	(0.0043)	(0.0048)	(0.0065)	(0.0049)
Com. Colon. $Post\text{-}1945_{ij}$	-0.1418***	-0.0998***	-0.1004***	-0.2056***	-0.1256***	-0.1567***
	(0.0081)	(0.0108)	(0.0094)	(0.0086)	(0.0118)	(0.0103)
Colonial $Relationship_{ij}$	-0.2001***	-0.2539***	-0.1927***	-0.2152***	-0.2987***	-0.1845***
	(0.0054)	(0.0072)	(0.0065)	(0.0059)	(0.0079)	(0.0067)
RTA_{ij}	-0.1103***	-0.0578***	-0.117***	-0.1133***	-0.0684***	-0.1167***
	(0.0036)	(0.0047)	(0.0037)	(0.0042)	(0.0052)	(0.0043)
Constant	2.8558***	3.0228***	2.6303***	3.0227***	3.1009***	2.8206***
	(0.02)	(0.0247)	(0.0204)	(0.0232)	(0.0274)	(0.0226)
N	67,104	46,848	59,424	67,104	46,848	59,424
Adjusted R^2	0.7921	0.6565	0.8121	.	.	.
Pseudo R^2	.	.	.	0.673	0.5586	0.6852

Standard errors in parentheses. Each estimation includes year-varying country fixed effects. "***", "**", and "*" denote statistical significance from zero at the 1%, 5%, and 10% levels, respectively.

As noted, we examine whether the observed changes in distance coefficients across time are statistically significant by regressing the year-specific distance coefficients on a time variable. Results are presented in Table III. For all three trade cost categories, the time trend is positive and statistically significant from zero. More specifically, the slope coefficients that correspond with bilateral total, agricultural, and manufacturing trade costs are 0.001 ($p < 0.01$), 0.0038 ($p < 0.01$), and 0.0009 ($p < 0.01$), respectively. This confirms the existence of a trade costs distance puzzle.

Table III: Distance Coefficients Time Trends, OLS Estimations

	Total (a)	Agriculture (b)	Manufactures (c)
Time	0.001***	0.0038***	0.0009***
	(0.0002)	(0.0003)	(0.0002)
Constant	-1.7513***	-7.3942***	-1.4427***
	(0.4689)	(0.5983)	(0.4895)
N	24	24	24
Adjusted R^2	0.4361	0.8771	0.3366

Standard errors in parentheses. "***", "**", and "*" denote statistical significance from zero at the 1%, 5%, and 10% levels, respectively.

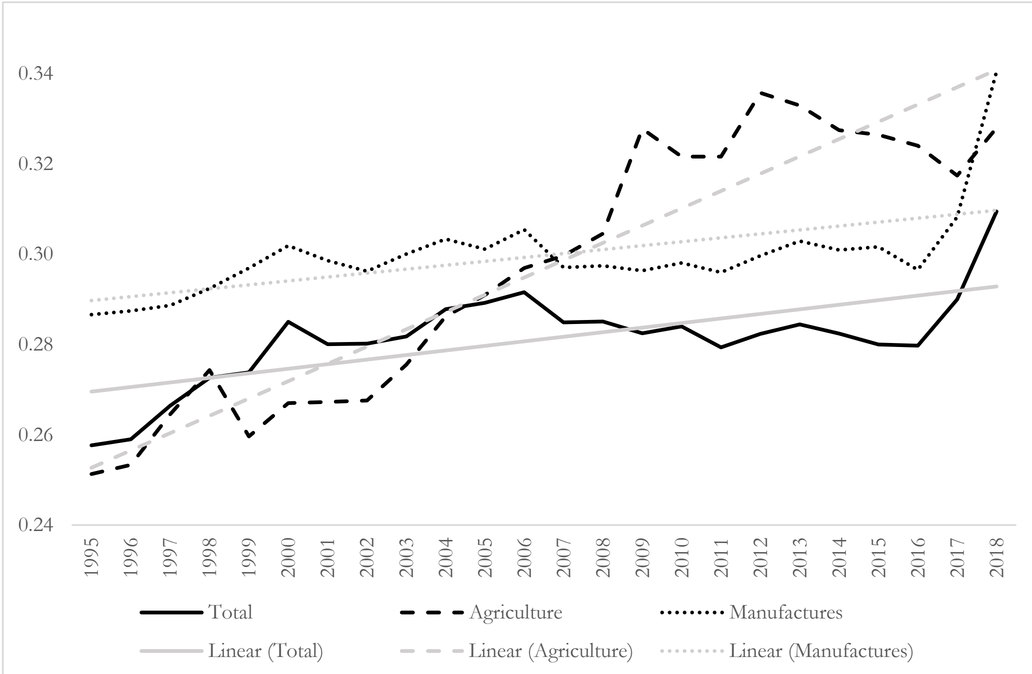
Figure 2 illustrates the trade costs-distance puzzle, providing the time paths of the estimated distance coefficients. The figure shows the relationship between distance and each of our measures of trade costs, with the points corresponding to the estimated coefficients that are presented in Table II. A linear trend line is added to each set of coefficient estimates.

Attempting to resolve the trade cost-distance puzzle, we follow Buehler and White (2015) and apply the PPML estimation technique. Estimation of equation (1) using PPML produces the sets of distance coefficients that are presented in columns (d) through (f) of Table II. Here we find something interesting and contrary to our expectations: the puzzle is resolved for the total trade costs and manufacturing trade costs, yet it is not resolved for agricultural trade costs.[8] This is verified in Table IV. Regressing the year-specific distance coefficients presented in Table II on the time variable produces statistically insignificant slope coefficients for the total trade costs and manufacturing trade costs series (i.e., -0.0002 ($p > 0.10$) and -0.0001 ($p > 0.10$), respectively).[9]

[8] PPML estimation is recommended by Silva and Tenreyro (2006), and the technique is used in Borchert and Yotov (2017) to resolve the traditional "distance puzzle." As indicated in these articles, PPML estimation includes the advantages of accounting for heteroskedasticity in the data and information contained in zero trade flows.

[9] To test the robustness of our results, we repeated our analysis using three alternative measures of distance between trading partners – namely, the geodesic distance, calculated following the great circle formula, which uses latitudes and longitudes of capital cities (*DISTCAP* in the CEPII Gravity Database) and two weighted measures of distance that consider the population distributions within countries (*DISTW* and *DISTWCES*). In all cases, the results presented here were found to be robust. These results are available in an online appendix (i.e., as Supplemental Material).

Figure 2: Time Plots of Distance Coefficients, OLS Estimations



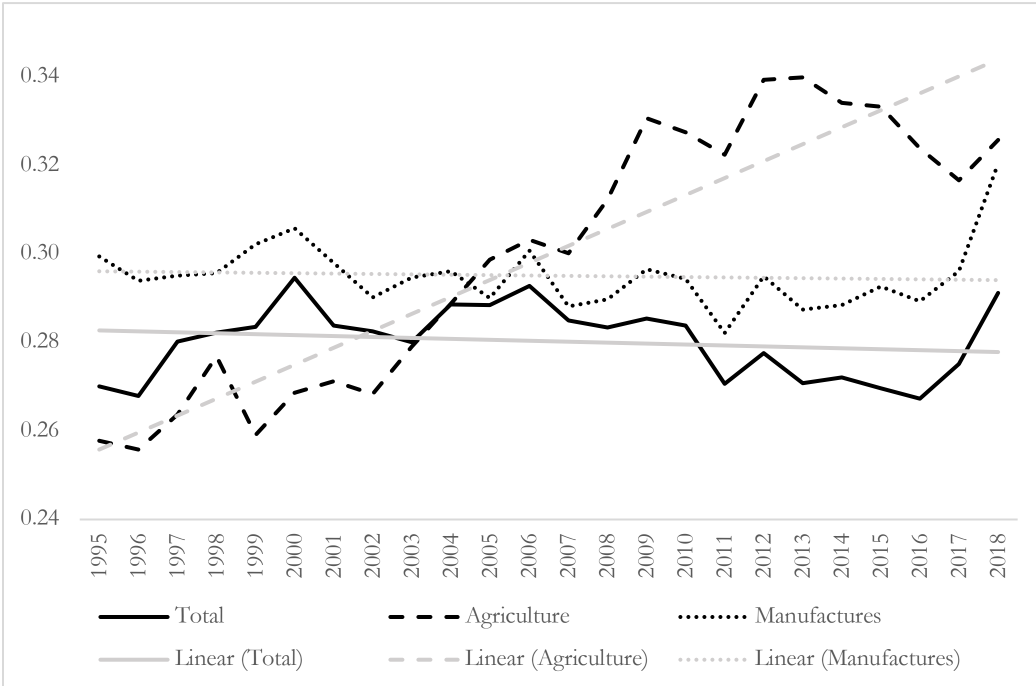
Our results indicate that the application of the PPML estimation technique resolves the puzzle for total trade costs and for manufacturing trade costs. Figure 3 demonstrates the persistence of the puzzle for agricultural trade costs. Increasing magnitudes for the year-specific distance coefficients in the agriculture sector suggests that globalization has not yet diminished the extent to which distance impedes international trade of such goods. We consider the puzzle to be unresolved if the time series estimation of the distance coefficients results in a positive and significant estimated coefficient. This is a more robust test than comparing the distance coefficient only at the beginning and end of the sample.

Table IV: Distance Coefficients Time Trends, PPML Estimations

	Total (a)	Agriculture (b)	Manufactures (c)
Time	-0.0002	0.0038***	-0.0001
	(0.0002)	(0.0003)	(0.0002)
Constant	0.7065	-7.3869***	0.4691
	(0.487)	(0.6934)	(0.4549)
N	24	24	24
Adjusted R^2	-0.0103	0.8413	-0.0385

See Table II notes.

Figure 3: Time Plots of Distance Coefficients, PPML Estimations



4. Conclusions

Examining data for total, agricultural, and manufacturing trade costs over the period 1995-2018, we have demonstrated the existence of a novel trade cost-distance puzzle that is related to the "distance puzzle" more commonly found in the international trade literature. Similar to the traditional distance puzzle, one would expect, as a result of increased globalization, distance to have a falling or, perhaps, constant influence on trade costs over time. However, when using the standard OLS estimation technique to examine both unbalanced and balanced samples, we find the influence of distance on trade costs has significantly *increased* over time. These results hold whether we examine total trade costs, agricultural trade costs, or manufacturing trade costs. Thus, we confirm the existence of the trade cost-distance puzzle.

A second contribution of this paper involves proposing a solution that partially resolves the puzzle. When we employ the PPML estimation technique to examine balanced samples, we resolve the puzzle for total trade costs and for manufacturing trade costs. The puzzle persists, however, for agricultural trade costs. Accordingly, we propose the application of the PPML estimation technique in conjunction with balanced data sets as a partial solution to the novel trade cost-distance puzzle.

Lastly, it is important to remember the evidence reported here is a starting point. Understanding the determinants of trade costs may assist in the formulation of public policy actions that reduce trade costs and, thus, increase trade flows. Given the implications for enhancing welfare, additional research is needed to fully resolve the puzzle.

References

Ali, Salamat and Chris Milner “Trade costs and the composition of developing countries' exports.” Review of Development Economics (2022): n. pag.

Anderson, James, and Eric Van Wincoop (2004) “Trade Costs”, *Journal of Economic Literature* 42(3): 691-751.

Arvis, Jean-François, Yann Duval, Ben Shepherd, and Chorthip Utoktham (2013) “Trade Costs in the Developing World: 1995-2010”, Policy Research Working Paper; No. 6309. World Bank, Washington, DC.

Baldwin, Richard and Daria Taglioni (2006) "Gravity for Dummies and Dummies for Gravity Equations," NBER Working Papers 12516, National Bureau of Economic Research, Inc.

Bergstrand, Jeffrey H., Mario Larch, Yoto V. Yotov (2013) “Economic Integation Agreements, Border Effects, and Distance Elasticities in the Gravity Equation”, CESifo Working Paper, No. 4502, Center for Economic Studies and ifo Institute (CESifo), Munich

Buehler, David and Roger White (2015) “A Simple Solution to the Distance Puzzle: Balanced Data and Poisson Estimation”, *Applied Economics Letters* 22(7): 587-592.

Borchert, Ingo and Yoto V. Yotov (2017) “Distance, Globalization, and International Trade,” *Economics Letters* 153(C): 32-38.

Brei, Michael and Goetz von Peter (2018) “The Distance Effect in Banking and Trade”, *Journal of International Money and Finance* 81(C): 116-137.

Coe, David T., Arvind Subramanian, Natalia T. Tamirisa, and Rikhil R. Bhavnani (2002) “The Missing Globalization Puzzle”, IMF Working Paper No. 02/171.

Chen, Natalie and Dennis Novy (2009) “International Trade Integration: A Disaggregated Approach”, *CESifo Working Paper No. 2595*. March.

Centre d’Etudes Prospectives et d’Informations Internationales (CEPII) (2021) Gravity Dataset. Online. Available at: http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id=8. Accessed February 2021.

Disdier, Anne-Célia, and Keith Head (2008) “The Puzzling Persistence of the Distance Effect on Bilateral Trade.” The Review of Economics and Statistics, vol. 90, no. 1, pp. 37–48.

ESCAP-World Bank Trade Cost Database (2021). Online. Available at: https://databank.worldbank.org/source/escap-world-bank-international-trade-costs#. Accessed February 2021.

Gervais, Antoine (2019) "Indirect measures of trade costs: Limitations and caveats", *Economics Letters*, Volume 176, Pages 96-102.

Head, Keith and Thierry Mayer (2014). "Chapter 3 - Gravity Equations: Workhorse, Toolkit, and Cookbook" Editor(s): Gita Gopinath, Elhanan Helpman, Kenneth Rogoff, Handbook of International Economics, Elsevier, Volume 4, Pages 131-195.

Hou, Yulin, Yun Wang, and Wenjun Xue (2020) "What Explains Trade Costs? Institutional Quality and Other Determinants", *Review of Development Economics* 25(1): 478-499.

Kaminchia, Sheila Mwendwa (2019) "The Determinants of Trade Costs in the East African Community", *Journal of Economic Integration* 34(1): 38-85.

Marti, Luisa and Rosa Puertas (2019) "Factors Determining the Trade Costs of Major European Exporters", *Maritime Economics & Logistics* 21: 324–333.

Novy, Dennis (2013) "Gravity Redux: Measuring International Trade Costs with Panel Data", *Economic Inquiry* 51(1): 101–21.

Silva, J. M. C. Santos and Silvana Tenreyro (2006) "The Log of Gravity", The Review of Economics and Statistics; 88 (4): 641–658.

World Bank (2021). "Global Recovery Strong but Uneven as Many Developing Countries Struggle with the Pandemic's Lasting Effects", Online: https://www.worldbank.org/en/news/press-release/2021/06/08/world-bank-global-economic-prospects-2021.

Yotov, Yoto V. (2012) "A Simple Solution to the Distance Puzzle in International Trade", *Economics Letters* 117(3): 794-798.

Yilmazkuday, Hakan (2017) "A Solution to the Missing Globalization Puzzle by Non-CES Preferences", *Review of International Economics* 25(3): 649-676.