Web Appendix for:

The Effect of Communication Costs on Trade in Headquarter Services

Anca D. Cristea*
University of Oregon

October 2014

This document presents data analyses that are not for publication, but complement the empirical estimations from the paper titled "The Effect of Communication Costs on Trade in Headquarter Services". All the references to regression equations, estimation coefficients or particular model specifications that are made in this document without being directly reported here correspond to content from the paper manuscript.

1 Detailed Data Appendix

There are four pieces of information that are essential for the econometric estimation: 1) data on the value of intra-firm knowledge transfers, 2) data on the average skill level of workers in the foreign market, 3) information on the costs and barriers of cross-border communication, and 4) information on the determinants of intra-firm exports of headquarter services to be used as control variables. Due to limited data availability, the estimation sample is based on U.S. data, and covers aggregate bilateral exports of headquarter services to 32 foreign countries over the period 1993-2008. In what follows, I provide details on the collection and preparation of the data to generate the sample used in all the estimation exercises.

Services Trade. Data on services trade is generally difficult to get, which is why we use U.S. data. The Bureau of Economic Analysis (BEA) collects and provides statistics on U.S. services, which consist of: "(1) cross-border trade in services [...] and (2) services supplied through the channel of direct investment by affiliates of multinational companies". For this paper, I use publicly available data provided on the BEA website under the header "detailed statistics for cross-border trade".

*Contact: Department of Economics, University of Oregon, 1285 University of Oregon, Eugene, OR 97403, USA. Tel: +1 (541) 346-1245. E-mail: cristea@uoregon.edu. Web: https://www.uoregon.edu/∼cristea.

Within this database, I focus on "other private services" (statistics by affiliation of transactors), which includes "business, professional and technical services" as the largest service class.

I measure the export volume of headquarters services as the "receipts [in million dollars] by U.S. parents from their foreign affiliates". I also measure the value of service exports to unaffiliated parties as "receipts [in million dollars] by unaffiliated parties". Based on these two variables, I construct the export share of headquarter services in the total exports of affiliated and unaffiliated services. In defining the denominator of this ratio as the sum of receipts by unaffiliated parties and by U.S. parents from their foreign affiliates, I essentially exclude the receipts by U.S. affiliates from their foreign parent groups.

In addition to affiliated trade data on services, I also collect data on the total sales of the majority owned foreign affiliates (MOFAs), distinguishing between all MOFAs and manufacturing MOFAs. This data is publicly available from the BEA based on annual survey records on the activity of non-bank U.S. multinational firms (U.S. Direct Investments Abroad) Table *III.E 3* provides data on sales by affiliates, country by industries.

Educational Attainment. Data on the average educational attainment by country and year come from the Barro and Lee (2010) dataset. I use information on the average years of education in the total population aged 25 years and above as a way to measure the knowledge and skill level of the foreign labor force. The original Barro-Lee data reports educational attainment variables at five-year intervals for a large number of countries. To get annual data, I follow the practice in the literature and apply a linear interpolation to generate values for in-between years (Blonigen et al., 2003).

Telecommunication. International telecommunication data are available for the U.S. from the U.S. Federal Communications Commission (FCC), Section 43.61 International Traffic Reports. I collect data on total traffic billed in the U.S. (i.e., outbound calls in minutes), on total traffic billed in the foreign countries and originating/terminating in the U.S. (inbound calls in minutes), and on the U.S. billed revenue per minute (i.e., outbound calling rate). The data applies to message telephone service (and not private line service).

Based on the available telecommunication traffic data, I construct a variable measuring the calls imbalance between the U.S. outbound and inbound minutes of telephone services. The imbalance in minutes of phone calls influences the negotiation of settlement rate between the U.S. and the foreign country, which is an important cost component for international calling rates. There are 9 countries with missing data on inbound telecommunication traffic (Argentina, Belgium, Chile, Indonesia, Malaysia, New Zealand, Norway, Sweden and Venezuela). Since data on outbound traffic is available for these countries, I impute their calls imbalance values by using information on the the fraction of calls imbalance in total outbound traffic, averaged across sample countries by year.

From available data I also construct a distance-weighted measure of neighboring countries' calling rates (similar in spirit to the market potential index) to use as instrument for a given country's calling rate. Based on the (population-weighted) bilateral distance data from CEPII and on the U.S. outbound billed revenue per minute of phone call, I construct the average calling rate index for country i as: $\sum_{j \neq i} \frac{Pcall_j}{dist_{ij}}$.

International Passenger Aviation. International aviation data is available from the *Databank 1B (DB1B) Origin and Destination Passenger Survey* provided by the U.S. Department of Transportation for the period 1993-2008. The DB1B represents a 10 percent sample of airline tickets drawn from airport-pair routes with at least one end-point in the U.S. Each airline ticket recorded in the data contains information on the complete trip itinerary including airports, air carriers marketing the ticket and operating each flight segment, the total air fare, distance traveled split by flight segments, ticket class type, as well as other segment level flight characteristics.

The original ticket level information is aggregated across all U.S. inbound and outbound itineraries between any origin-destination airport pairs connecting the U.S. to a foreign destination. The resulting sample includes data on the total number of passengers (business-class and economy-class combined), on the passenger-weighted average airfare for economy-class, and on the passenger-weighted average number of flight segments per one-direction of travel.

To complement the information on average airfares with cost-based measures, I add data on jet fuel price indices (in real terms) available from the International Air Transport Association (IATA). By interacting it with bilateral distances between countries, the variation in fuel prices over time will provide a good proxy for cost shocks to international aviation.

Internet penetration rates. Data on internet users per 100 people by country is available from the World Development Indicators database, provided by the World Bank.

Country Characteristics. Most of the country characteristics used as control variables in the estimation come from data sources that are standard in the trade literature. Population, real per-capita GDP (PPP converted, at 2005 constant prices), trade openness (defined as the sum of imports and exports divided by GDP, expressed at 2005 constant prices), and real exchange rates (LCU per US dollar) are available from the Penn World Tables (PWT) version 7.1. Data on factor endowments such as the size of active workforce and the capital stock (constant 2005 prices) are provided in the PWT version 8.0. I construct the capital to labor K/L ratio as the ratio of capital stock (in constant 2005 US dollars) to the employment level in a country.

Trade Costs. The main data source for bilateral trade costs is the gravity dataset provided by CEPII. This is a square matrix of all world pairs of countries, and for each country pair it reports information on contiguity, common official language, common colonial linkage, bilateral distance (population weighted) and time zone difference (in hours).

The market potential variable (Mkt Potent) is constructed from available data as a distance-weighted average of countries' incomes. Using the (population-weighted) bilateral distance data from CEPII, and the dat on real GDP from PWT, the market potential of country i is calculated as: $\sum_{j \neq i} \frac{rGDP_j}{dist_{ij}}$.

To measure the strength of ethnic networks, I use U.S. data on foreign-born population by origin of birth. This is provided in the 2000 Decennial US Census. As data on ethnic networks is quite persistent and not available on an annual basis, in the estimations I will interact the U.S. stock of foreign born population by country with a linear time trend.

International Trade Policy. Information on the top corporate income tax rates (in percentages) is available from the *World Tax Database* maintained by the University of Michigan. This is the statutory corporate tax rate established by a country at a given point in time.

The bilateral treaty information is codified as a 1/0 indicator variable. The data is compiled from various sources. The United Nations Conference on Trade and Development (UNCTAD) provides data on the bilateral investment treaties (BIT) signed between the U.S. and foreign countries. Information regarding the countries with which the U.S. has signed free trade agreements (FTA) is available from the gravity dataset provided by CEPII up until 2006, and from the WTO website for the remaining period 2007-2008 (see the "List of RTAs in force" available at www.wto.org). Finally, data on bilateral tax treaties (BTT) is provided by the Internal Revenue Service (see the "United States Income Tax Treaties A-Z" available at www.irs.gov).

2 Unreported Estimation Exercises

2.1 Sales of Majority-Owned Foreign Affiliates (MOFA)

One implication of the theory setting is that the size of the foreign affiliate (i.e., value of the cross-border production team) increases as a result of lower communication costs. This is because U.S. managers can now train larger teams of foreign workers, thus increasing their span of control. This implies that the production and sales of foreign affiliates should be expected to increase as a result of decreasing communication costs. A question of interest empirically is whether this effect is more pronounced at low levels of educational attainment. Since headquarter service exports respond differentially to changes in communication costs depending on the average skill of the foreign workforce, it can be expected that affiliate sales follow the same pattern.

The regression model given by equation (4) in the paper represents the baseline specification that is estimated in the data. I exploit an alternative yet indirect way to assess the extent to which communication costs hinder the supply of headquarter services by examining changes in the total sales of foreign affiliates. An implication of the theory framework is that the output produced by the foreign affiliate is a direct function of the size of the production team, which in turn depends on the availability of skilled workers at high levels of communication costs. So, it is expected that the value of foreign affiliate sales responds to the interaction between communication costs and skilled labor endowments in a similar way as intra-firm knowledge transfers. Therefore, I exploit data on the total sales of the majority owned foreign affiliates (MOFAs) and use the same regression specification as given by equation (4) in the paper to evaluate how well foreign talent substitutes for knowledge inputs from the headquarters when cross-border communication is difficult. If such substitution patterns operate in the data, then multinational production should be less affected and affiliate sales should be less responsive to communication costs. Thus, I expect to find that $\beta_1 < 0$ and, especially, that $\beta_2 > 0$.

[insert Table A1 here]

Table A1 reports the results from estimating the regression model given by equation (4), and using as dependent variable the total sales of majority-owned foreign affiliates (MOFA). Column 1 reports the OLS estimates obtained by using telephone rates as a measure of communication costs. Columns 2-4 report both stages of the instrumental variables estimation, which uses the bilateral volume of phone calls as proxy for the inverse of communication costs. In both cases, I find the same pattern of results as for the regressions explaining the share of headquarter services trade, a finding that is consistent with theory-based expectations. For robustness, columns 5-7 report the 2SLS estimates (first and second stages) using as measure of communication information on passenger aviation flows. Once again the pattern of results conforms with expectations.

The interpretation for these results is similar to prior explanations. Conditional on the average skill of a foreign worker, a reduction in communication costs implies that headquarter managers need to spend less resources to convey a certain amount of information per foreign worker, which enables them to increase their span of control and achieve higher levels of production output. However, when the average skill level of the foreign workforce is high, the amount of knowledge inputs supplied per worker is low, which means that the foreign production team commanded by the headquarters is already large. So, a marginal reduction in communication costs will only have a small impact on the team size, on the productivity of foreign labor, and thus on the total output of the foreign affiliate. Once factoring in possible inefficiencies from expanding too much the scale of foreign operations – which are outside of the theory framework – I can provide possible explanations for the lack of statistical significance in the marginal effect of communication on the total sales of foreign affiliates at high skill levels attained by the foreign workforce (which are reported at the bottom of the Table A1).

2.2 Alternative Specifications for Telecommunication and the Exports of Headquarter Services

Remove Calling Rates as Excluded Instrument. In estimating the effect of telecommunication on the exports of headquarter services, I have used international calling rates as an excluded instrumental variable when correcting for the endogeneity of the volume of international phone calls. This may create concerns if one thinks that the price and quantity of telecommunication are jointly determined in equilibrium, so the same mechanisms that create endogeneity in communication quantities must apply equally to prices. While I provide direct evidence that calling rates are indeed exogenous with respect to export shares of headquarter services, it is nevertheless worth investigating the sensitivity of our estimates to the elimination of calling rates from the set of excluded instrumental variables.

In an alternative specification mirroring the estimates reported in Table 4 in the paper, I estimate the model with the following set of excluded instruments for international phone call minutes: bilateral calls imbalance, weighted average of neighboring countries' calling rates, and the time zone difference interacted with a linear trend. Essentially, this is the same specification as reported in the paper, but where calling rates have been replaced with the variables used before as instruments *for* calling rates. Like before, interaction terms between these instruments and the skill level in the foreign country are used to instrument for the endogenous interaction terms between communication and skill. The OLS and 2SLS estimation results are reported in Table A2.

[insert Table A2 here]

For comparison purposes, columns 1 and 2 reproduce columns 1and 2 from Table 4 in the paper (which are based on calling rates as part of the instrument set). Columns 3-5 report the first and second stage estimates based on the modified set of excluded instruments (without calling rates). As expected, the pattern of results is unchanged between columns 2 and 3. This provides further evidence (besides the Hausman test for endogeneity, and the Hansen J statistic for over identifying restrictions) that calling rates are most likely exogenous to intra-firm trade in services, and that the findings in the paper are not affected by the inclusion or exclusion of calling rates as an instrument.

Estimations based on the Level of Headquarter Services. One potential concern with specifying the dependent variable as the exports share of headquarter services in total service exports comes from the fact that now unobservable factors that determine the boundaries of the firm and the composition of trade in terms of intra-firm vs. arm's length transactions may impose some challenges. For example, it could be the case that shocks to communication infrastructure affect related-party trade in services more than the overall volume of services trade, and that this systematic heterogeneity in responses to cost shocks may impact the validity of instrumental variables estimates.

To verify the sensitivity of the reported results to composition effects driven by the type of trade transactions, I experiment with a model specification mirroring the estimates reported in table 4 in the paper, but using as dependent variable the export level of headquarter services rather than the export share. The results are reported in Table A3.

[insert Table A3 here]

For comparison purposes, columns 1 and 2 reproduce columns 1 and 2 from Table 4 in the paper (which are based on export shares). Columns 3-6 report the OLS and IV estimates based on export volumes. Consistent with prior expectations, the pattern of results is unchanged between columns 1-2 and 3-4. This provides evidence that composition effects in terms of the nature of trade transactions are not impacting the main findings in the paper.

Table A1: Effect of Communication Flows on the Sales of Majority-Owned Foreign Affiliates

	Dependent Variable: Total Sales by MOFA (log)						
	Phone Calls				***Air Travel***		
			1st Stage			1st Stage	
	OLS	**2SLS**	*Calls*	*Calls×Skill*	**2SLS**	*Travel*	*Trav×Skill*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Communication Cost	-0.558**						
	[0.238]						
Communication Cost×Skill Level	0.224*						
	[0.115]						
Communication Volume		0.661***			1.276***		
		[0.133]			[0.265]		
Communication Volume × Skill Level		-0.224***			-0.672***		
		[0.066]			[0.137]		
Skill Level	-0.958	3.535***	-0.104	21.184***	6.065***	-4.019***	9.749***
	[0.678]	[1.224]	[0.632]	[1.174]	[1.315]	[1.426]	[2.653]
Time Zone Diff. × Skill Level			-0.032	-0.335**			
			[0.072]	[0.145]			
Calling Rate			-1.094***	-0.984***			
			[0.174]	[0.311]			
Calling Rate × Skill Level			0.275***	-0.054			
			[0.077]	[0.138]			
Air Fare						-1.981***	-1.915**
						[0.410]	[0.741]
Flight Segments per Trip						0.104	5.173**
						[1.134]	[2.084]
Fuel × Distance						0.000	-0.092
						[0.035]	[0.069]
Air Fare × Skill Level						0.733***	0.505
						[0.200]	[0.375]
Flight Segments × Skill Level						-0.656	-3.578***
						[0.498]	[0.926]
Country FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES
Observations	462	462	462	462	462	462	462
R-squared	0.912	0.915	0.870	0.968	0.917	0.809	0.950
First Stage Statistics:							
Partial F-stat			27.22	25.92		18.86	17.13
Hansen J stat		1.68 (0.20)			15.14 (0.00)		
Marginal effect of calling rates (evaluated at various levels of skill):							
at 1 std. below mean	-0.145***	0.247***			0.037		
	[0.053]	[0.065]			[0.116]		
at sample mean	-0.079	0.180***			-0.163		
	[0.050]	[0.070]			[0.128]		
at 1 std. above mean	-0.012	0.113			-0.364**		
	[0.068]	[0.079]			[0.150]		

*** p<0.01, ** p<0.05, * p<0.1, + p<0.15; Robust standard errors in brackets

Note: The results correspond to the regression equation (4), but with the dependent variable changed to the total sales of U.S. MOFAs. All continuous variables are expressed in logs. Both price-based (column 1) and quantity-based (columns 2 and 5) communication measures are used in the estimation, focusing on two communication proxies – telecommunication and air travel. The same exogenous variables as defined in Table 4 in the paper serve as instruments for phone calls and its interaction term. Similarly, air travel and its interaction term are instrumented for using the excluded variables from Table 5 in the paper. Columns 3 and 4, respectively 6 and 7 report the first stage results. All specifications include a complete set of control variables and fixed effects (identical to column 4 of Table 3 in the paper), which are omitted from the table but available upon request.

Table A2: Effect of Telecommunication on the Export Share of Headquarter Services: Alternative Instrumental Variables

Dependent Variable:	Share of HQ Services in Total Service Exports				
	Include Calling Rates		***Exclude Calling Rates***		
				1st Stage	
	OLS	**2SLS**	**2SLS**	*Calls*	*Calls×Skill*
	(1)	(2)	(3)	(4)	(5)
Phone Calls	1.158***	1.029***	0.984***		
	[0.249]	[0.321]	[0.321]		
Phone Calls × Skill Level	-0.531***	-0.364**	-0.409***		
	[0.124]	[0.158]	[0.149]		
Skill Level	9.057***	6.006**	6.867**	-2.630***	12.110***
	[2.501]	[2.958]	[2.793]	[0.782]	[1.529]
W. Avg. Neighbors' Calling Rates				-0.542**	0.406
				[0.251]	[0.506]
Calls Imbalance				0.316***	0.661***
				[0.063]	[0.132]
Time Zone Diff. × Year				-0.008***	-0.020***
				[0.002]	[0.005]
Neighbors' Calling Rates × Skill Level				0.155**	-0.391***
				[0.069]	[0.124]
Time Zone Diff. × Skill Level				0.436***	0.858***
				[0.107]	[0.211]
Country FE					
Year FE					
Observations	462	462	462	462	462
R-squared	0.425	0.408	0.421	0.917	0.980
First Stage Statistics:					
Hansen J stat		0.215	1.466		
Hansen J p-val		0.898	0.690		
F-stat				23.71	21.27

*** p<0.01, ** p<0.05, * p<0.1, + p<0.15; Robust standard errors in brackets

Note: The results correspond to the regression equation (4) in the paper. All continuous variables are expressed in logs. Columns 1 and 2 are the same as columns 1 and 2 from Table 4 in the paper. Columns 3-5 report instrumental variables estimates *without* calling rates as an excluded instrumental variable. In contrast to column 2, column 3 uses information on calls imbalance and on the weighted average of calling rates of neighboring countries as a replacement for calling rates. All specifications include a complete set of control variables and fixed effects (identical to column 4 of Table 3 in the paper), which are omitted from the table but available upon request.

Table A3: Effect of Telecommunication on the Export Volume of Headquarter Services

Dependent Variable:	**Exports of Headquarter Services:**					
	Export Share		***Export Volumes***			
					1st Stage	
	OLS	**2SLS**	**OLS**	**2SLS**	*Calls*	*Calls×Skill*
	(1)	(2)	(3)	(4)	(5)	(6)
Phone Calls	1.158***	1.029***	1.347***	1.022***		
	[0.249]	[0.321]	[0.398]	[0.330]		
Phone Calls × Skill Level	-0.531***	-0.364**	-0.622***	-0.489***		
	[0.124]	[0.158]	[0.194]	[0.163]		
Skill Level	9.057***	6.006**	10.311***	7.984***	-2.225**	15.829***
	[2.501]	[2.958]	[3.711]	[3.001]	[1.012]	[1.913]
Calling Rate					-1.085***	-0.961***
					[0.172]	[0.306]
Time Zone Difference × Year					-0.006***	-0.014***
					[0.002]	[0.004]
Calling Rate × Skill Level					0.274***	-0.057
					[0.075]	[0.135]
Time Zone Diff. × Skill Level					0.272**	0.430*
					[0.126]	[0.249]
Observations	462	462	462	462	462	462
R-squared	0.425	0.408	0.759	0.757	0.873	0.969
First Stage Statistics:						
Hansen J stat		0.215		2.966		
Hansen J p-val		0.898		0.227		
F-stat					21.35	21.74

*** p<0.01, ** p<0.05, * p<0.1, + p<0.15; Robust standard errors in brackets

Note: The results correspond to the regression equation (4) in the paper, but with the dependent variable changed to the volume of exports in headquarter services in columns 3-6. For comparison purposes, columns 1 and 2 are the same as columns 1 and 2 from Table 4 in the paper and use export shares as dependent variable. The IV (first and second stage) estimates from columns 3-5 are based on the same set of instruments for the volume of phone calls as used in the paper (i.e., calling rates, time zone differences, and each of their interactions with the skill level). All continuous variables are expressed in logs. All specifications include a complete set of control variables and fixed effects (identical to column 4 of Table 3 in the paper), which are omitted from the table but available upon request.