Journal of International Economics 100 (2016) 174–193

Contents lists available at ScienceDirect

Journal of International Economics

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

Gravity with scale effects☆

James E. Anderson[a,b,*], Mykyta Vesselovsky[c], Yoto V. Yotov[d,e,f]

[a] Department of Economics, Boston College, Chestnut Hill, MA 02467, USA
[b] NBER, USA
[c] Global Affairs Canada, Ottawa, ON K1A OG2, Canada
[d] School of Economics, Drexel University, Philadelphia, PA 19104, USA
[e] CESifo Research Network, Munich, Germany
[f] Economic Research Institute, Bulgarian Academy of Sciences, Sofia, Bulgaria

ARTICLE INFO

Article history:
Received 10 March 2014
Received in revised form 9 January 2016
Accepted 7 March 2016
Available online 19 March 2016

JEL classification:
F14
F15
F16

Keywords:
Gravity
Exchange rates
Goods and services

ABSTRACT

This paper extends the structural gravity model to incorporate scale effects and exchange rate passthrough. Bilateral scale effects in cross-border trade are inferred from the difference in distance elasticities between cross border and inter-provincial bilateral trade in a majority of 28 goods and services sectors for Canada's provinces. Bilateral-specific relationship investment is a possible explanation. Incomplete passthrough of large exchange rate changes from 1997 to 2007, amplified by scale effects, produces direct effects on bilateral trade for 12 of 19 goods sectors but none of 9 services sectors.

1. Introduction

We find evidence of scale effects in cross-border bilateral trade in this paper. Differences between cross-border and domestic gravity elasticities are interpreted to reflect scale effects in an extended structural gravity model applied to Canadian provincial bilateral trade flow data for 28 goods and services sectors over the decade 1997–2007.Wesuggest an explanation based on bilateral relationship-specific investment, but in the absence of direct evidence on such investment the scale effect remains a black box. Estimated scale effects are economically substantial. The absolute value of the distance elasticity of cross border bilateral trade exceeds that of domestic bilateral trade statistically and quantitatively significantly in a majority of 28 sectors for Canada's provinces (23 of 28 cases for inward trade and 13 of 28 cases for outward trade). In aggregate goods trade a 100% rise in imports lowers Canadian trade costs by 12.3% and lowers US trade costs by 6.1%, assuming an elasticity of substitution equal to 6.13 (based on Head and Mayer, 2014). For aggregate services imports of Canada, the corresponding reduction is 9.3% while for the US the estimated elasticity is not significantly different from 0. The lower US destination scale effects satisfy the intuition that the order of magnitude larger US market tends to exhaust scale effects in cross-border trade.

Our black box empirical model of scale effects can be interpreted in terms of the recent literature on network links in export dynamics. Chaney (2014) is a prime example. Bilateral links are specific capacities formed and maintained by specific buyer–selleragent interactions. These occur for example between employees of firms' purchasing and marketing departments or via intermediaries' services purchased by buyers and sellers. The total number of such individual bilateral links is an aggregate bilateral capacity that naturally

☆ We thank Mario Larch for useful comments and discussions. The paper has also benefitted from comments at seminars at Oxford and Pompeu Fabra Universities, and at the 2014 Meetings of the European Economic Association. We are indebted to Emily Yu for invaluable help with the data. The research was supported by the Department of Foreign Affairs, Trade and Development, Canada. All errors are our own. An earlier version circulated as "Gravity, Scale and Exchange Rates", NBER WP No. 18807.
* Corresponding author at: Department of Economics Boston College 140 Commonwealth Avenue Chestnut Hill, MA 02467, USA.

http://dx.doi.org/10.1016/j.jinteco.2016.03.003
0022-1996/

suggests bilateral scale effects on trade costs.[1] The volume changes suggested by this mechanism can be on either the intensive or extensive margin. Lacking firm level data, we are unable to discriminate between these. Alternatively, bilateral scale effects measured by the model could reflect fixed costs associated with designing products tailored to different destinations, e.g. Manova and Zhang (2012), fixed costs that are increasing with distance, e.g. Krautheim (2012), or firm search costs for buyers, e.g. Eslava et al. (2015). Yi (2010) influentially argues that fragmentation can explain nonlinear effects of trade costs on trade. Dis-agglomeration in the sense of Jones et al. (2005) emphasizes fragmentation of vertically integrated production driven by external increasing returns to scale in transportation and communication. While each of the above channels could potentially contribute to explaining our findings,the scale effects of this paper remain 'dark' like all gravity costs in Head and Mayer's (2013) cosmological metaphor. Stepping back from the model, our empirical results pose a distance/border puzzle in addition to the time series non-declining distance puzzle emphasized by Disdier and Head (2008).[2]Head and Mayer (2014) suggest that distance elasticities may vary bilaterally with market size, but we find little support for this pattern.

The starting point of the model is to allow trade costs to vary tractably with bilateral volume and do so differently in cross-border than in domestic trade. Invariance to volume is assumed in the standard gravity literature, a limiting case consistent with long run equilibrium investment of various sorts including that by atomistic agents in bilateral relationships. Cross-border trade, size-adjusted, is much smaller than domestic trade and tends to have a much shorter history. This difference suggests our specification that identifies a cross-border scale effect normalized by any effective domestic scale effect, i.e. the scale effects in our theory (and empirics) are defined and should be interpreted as *relative* to domestic scale effects in intra-national trade.[3] Other potential reasons for different gravity elasticities within and across borders are plausibly neutralized for US–Canadatrade, where mode choice plays little role. Most goods trade moves by road or rail, both within and across borders. Mode choice in tourism is dominated by distance, equally within and across borders. In contrast, mode choice plays an important role in other interregional and international trade, as emphasized by Hummels (2007) and Hillberry and Hummels (2008) among others.

Our data from 1997 to 2007 contains dramatic Canada–US exchange rate variation: an 11% depreciation followed by a 45% appreciation.[4] To control for this source of volume changes we develop a treatment of the effect of exchange rate changes with incomplete passthrough in combination with scale effects acting on the structural gravity model. Previous gravity applications could not simultaneously measure exchange rate effects and control for multilateral resistance with importer-time and exporter-time fixed effects. The availability of inter-provincial and cross-border trade applied to our extended model resolves this indeterminacy.[5] We concentrate here on bilateral trade between the US and Canada, suppressing relationships with the Rest of the World (ROW).[6] Sensitivity experiments include bilateral trade between Canada and Mexico. Our estimates of scale elasticities and other trade cost parameters turn out to be insignificantly affected by our treatment of exchange rate changes, though exchange rates separately contribute significantly to explaining variation in trade flows.

The parametric scale elasticity assumed in this paper is a simplification. Support for the simplification is provided by our finding that the inferred elasticities are constant over time in a decade in which cross-border volume varies substantially. Further support comes from results of estimating simple alternative specifications. No universal constancy is suggested, because we find directional asymmetry in the bilateral cross-border scale effects and sectoral variation of scale effects.

An obvious caveat is that the scale and passthrough elasticities are black box parameters. The variation of estimates across sectors suggests a payoff to opening the boxes. For scale elasticities, the caveat applies especially to a few sectors where the results suggest a mis-specified trade cost equation. As for exchange rate passthrough elasticities, there is ample evidence that exchange rate passthrough is incomplete over horizons of several years (Goldberg and Knetter, 1997) but it is unlikely to be constant.[7] Following much of the literature, we do not model incomplete exchange rate passthrough in this paper,[8] nor the exchange rate itself.

The modeling innovations of this paper may be useful in other settings. The data must contain both international and intra-national bilateral trade flows in order to identify scale elasticities on cross-border trade. A time series dimension to the preceding intra-national and international dimensions is required to examine passthrough implications of relative price changes at the border andtheir interaction with scale effects. The same modeling treatment applies to any change in cross-border frictions such as tariff reforms (potentially incompletely passed through) or free trade agreements. These are absent from Canada–US trade in 1997–2007 but would be relevant for gravity applications to other data sets.

Section 2 sets out the theoretical foundation. Section 3 develops the econometric specification and describes the data. Section 4

[1] Bilateral trade scale effects are distinct from scale effects in supply to all destinations (classic external scale economies as in Antweiler and Trefler, 2002) or demand from all origins (non-homothetic preferences across sectors as in Fieler, 2011). Any such aggregate scale effects are captured in the empirical application by country–sector–time fixed effects.

[2] Yotov (2012) notes that the non-declining distance puzzle goes away when cross-border and domestic distance elasticities are allowed to differ, as in this paper.

[3] In the sensitivity analysis we capitalize on a unique feature of our data, which enables us to distinguish between intra-provincial and inter-provincial trade. This offers an opportunity for us to actually test for possible scale effects within Canada. We do not find such effects for goods trade, but they are present for services.

[4] In 1997 the exchange rate stood at 0.72, then it fell to 0.64 in 2003, and in 2007 it was at 0.93.

[5] Some previous empirical gravity models have inserted real exchange rates into gravity equations without a theoretical foundation. The standard practice in these studies is to include a real exchange rate variable in a traditional version of the empirical gravity model, with no country-time fixed effects to control for multilateral resistance and with country mass variables represented by GDP and population. See for example Griffoli (2006), Kim et al. (2003) and Martínez-Zarzoso and Nowak-Lehmann (2003). A prominent but tangentially related literature considers the effect of exchange rate *regimes* such as currency unions on bilateral trade patterns. See Baldwin (2006) for a review of the literature on the effects of exchange rate *regimes*.

[6] We suppress ROW for three reasons. First, the trade cost function we develop below is unlikely to plausibly approximate such a heterogeneous aggregate region. Second, aggregation may bias our inferences regarding scale effects and controlling for exchange rate effects on trade with such a large region. Third, introducing ROW data does not have any effect on our model of bilateral trade or its estimated results due to the separable fixed effects estimation structure that we use.

[7] Goldberg and Knetter conclude that "While the response varies by industry, a price response equal to one-half the exchange rate change would be near the middle of the distribution of estimated responses for shipments to US" (p. 3). We abstract from explaining high frequency trade movements (within a year) because these may reflect random shocks and dynamic adjustment that have yet to be integrated with the gravity model. Differences in currency invoicing practices and length of contract terms affect high frequency price responses to exchange rate changes. It is possible that such differences across sectors may induce differing passthrough rates that persist in the medium run. In that case differing invoicing and contracting practices may help explain part of the differences in results we report across sectors.

[8] A search for evidence of pricing-to-market using our industry level data produced no informative results. One reason for this is the insensitivity of markups to exchange rates under CES preferences, as explained in Technical Appendix B, which is available by request.