



## Volume 33, Issue 2

### The agglomeration by destination of U.S. state exports.

Andrew J Cassey  
*Washington State University*

Katherine N Schmeiser  
*Mount Holyoke College*

#### Abstract

Exporting firms in France, the Netherlands, and Russia cluster by destination beyond that expected by GDP or ports (Choquette and Meinen 2011; Koenig 2009). It is unknown if this also occurs in the United States. The difficulty of obtaining U.S. customs data is the reason this remains unknown. Using the aggregate reduced form equation based on a firm-level theory of exporter agglomeration in Cassey and Schmeiser (2012), we estimate the coefficient on an agglomeration variable as measured by aggregate export weight and test its statistical significance. We find a 1% increase in aggregate weight increases exports by 0.4%. This estimate is economically and statistically significant and robust.

---

We are indebted to Tom Holmes. We thank Pavan Dhanireddy for research assistance. Portions of this research supported by the Agricultural Research Center Project #0540 at Washington State University.

**Citation:** Andrew J Cassey and Katherine N Schmeiser, (2013) "The agglomeration by destination of U.S. state exports.", *Economics Bulletin*, Vol. 33 No. 2 pp. 1504-1510.

**Contact:** Andrew J Cassey - cassey@wsu.edu, Katherine N Schmeiser - kschmeis@mtholyoke.edu.

**Submitted:** June 03, 2013. **Published:** June 15, 2013.

## 1 Introduction

Exporters cluster by destination. This clustering of exporting firms is beyond that from gross domestic product (GDP), access to ports, or physical distance to destination. That exporters cluster by destination is shown empirically by Koenig (2009) for France, Choquette and Meinen (2011) for the Netherlands, and Cassey and Schmeiser (*forthcoming*) for Russia. In all three of these cases, the authors use transaction-level customs data to establish the stylized fact.

The three countries that have been studied are diverse enough to suggest that exporter clustering by destination is a general phenomena. Nonetheless, it is unknown if there is firm-level clustering in the United States. This is an important unknown for understanding the economic effects and impacts of changes to international trade and industrial policies. If exporter clustering by destination does exist for the United States, then current models likely underestimate the importance of informational barriers to trade. An underestimate in the model could have important policy ramifications if firms are not choosing export destinations based solely on product-destination matches, but are additionally concerned with the knowledge gained through the export experiences of their physical neighbors. Then the firm response to a country-specific export policy such as the 2012 free trade agreements with Colombia, Panama, and South Korea will differ from that estimated in current trade models. The implication is that export promotion policy can be improved if policy makers recognize the effect of the location of exporters on their policy design.

There are several ways to determine if there is clustering of exporting firms by destination in the United States. First, one could get access to customs data for the United States and perform the kinds of statistical tests done in Koenig, Choquette and Meinen, and Cassey and Schmeiser. But the difficulty in obtaining United States customs data makes this direct approach problematic. In a different approach, Lovely, Rosenthal, and Sharma (2005) use the location of the headquarters of U.S. exporters and non-exporters. They find that the level of clustering of headquarters of exporters increases with the difficulty of exporting to a country relative to the clustering of domestic-only headquarters. But that is not the same thing as showing U.S. exporting firms cluster by destination of their shipments. Finally, one could test for statistical significance of an aggregate reduced form clustering-by-destination variable derived from theory using readily obtainable state-level export data. This is the route we take.

Cassey and Schmeiser (*forthcoming*) is the first to develop a theory for exporter clustering by destination by positing a destination specific externality. This externality could be a spillover of product-independent information about how to export to a particular market or economies of scale in packing a containerized vessel, thus creating the potential for exporters to skip over an otherwise preferred foreign market to take advantage of reduced costs from clustering. Their firm-level model yields an aggregate export equation that is similar to the standard gravity equation commonly used in empirical trade work. The difference is the Cassey and Schmeiser equation has an agglomeration term missing from standard gravity. The usefulness of this equation for determining if U.S. exporters cluster by destination is that all of the variables are at the aggregate level and thus there is no need to work with

difficult-to-obtain customs data.

We use U.S. states as our level of aggregation and estimate and test the statistical significance of the agglomeration term for the United States. We estimate the agglomeration term to increase exports by 0.4% for a 1% increase in aggregate export weight. This estimate is statistically significant. We then compare our estimates to out-of-sample information for further robustness. We find a gravity-like equation with an agglomeration variable accounts for 40% more of the variation in the state export data than the benchmark gravity equation.

## 2 The Reduced Form Gravity Equation With Agglomeration Term

We use the Cassey and Schmeiser (*forthcoming*) theory of exporter agglomeration by destination. Building on the standard monopolistic competitive model (as in Chaney (2008)), firms choose which of  $N$  differentiated markets to export to *given* their location in space and the exports and export destinations of all other firms. Firm revenue depends on the quantity and price in the export market. Firm costs depend on idiosyncratic variable costs due to productivity differences drawn ex ante from a productivity distribution, state production characteristics such as the wage, and variable and fixed barriers to trade. The variable trade cost is

$$\tau_{ij} = D_{ij} \times A_{ij}^{-\eta} \quad (1)$$

where  $i$  indexes production regions such as U.S. states,  $j$  indexes destination countries,  $D_{ij}$  is physical distance between state  $i$  and country  $j$ , and  $A_{ij}$  is a measure of agglomeration of exports from state  $i$  selling in country  $j$ . The parameter  $\eta$  sets curvature. The agglomeration term  $A_{ij}$  is the hypothesized cost saving at the firm-level from a spillover, for example knowledge about the destination market port requirements or economies of scale in transportation. As agglomeration increases, the variable trade cost  $\tau_{ij}$  decreases. Note there cannot be much variation in economic distance variables such as same official language or colonial history in  $i$  and  $j$  if all of the exporting regions  $i$  are in the same country, such as states within the United States, and thus they are omitted.

Following Cassey and Schmeiser, we take the aggregate gravity equation as derived in Anderson and van Wincoop (2003) as the benchmark and then replace the standard trade barriers variable with (1):

$$\log \frac{X_{ij}}{Y_i Y_j} = \alpha + \beta \log D_{ij} - \eta \beta \log A_{ij} + \sum_i \kappa_i S_i + \sum_j \delta_j C_j + \varepsilon_{ij}. \quad (2)$$

In (2),  $X_{ij}$  is state  $i$  aggregate exports to country  $j$ ,  $Y$  is GDP,  $S_i$  is an indicator variable representing an exporting state specific characteristics such as wage, and  $C_j$  is an indicator variable representing importing country specific characteristics, sometimes called multilateral resistance, as well as exchange rate with the U.S. dollar or idiosyncratic demand for certain goods. Having  $Y$  on the left forces the coefficients on GDP to be one as derived in Anderson and van Wincoop.

The importance of (2) is that though it is derived from a deeper firm-level theory of agglomeration, it can be estimated at the aggregate level. If  $\eta = 0$  then (2) reduces to the

standard gravity equation and there is no evidence of agglomeration of exports by destination. We test for this formally below. Additionally, Cassey and Schmeiser show that the structural interpretation of  $\beta$  is the (negative of the) curvature parameter of the firm productivity distribution if that distribution is Pareto. Axtell (2001) estimates that indeed the firm productivity distribution is Pareto and  $-\beta \approx 1$ . We formally test if  $-\beta = 1$ .

### 3 U.S. State Export Data

Because U.S. customs data are private and difficult to obtain (one needs Special Sworn Status at the Census Bureau), we turn to publicly available U.S. state export data. The data set is the Origin of Movement (OM) export data for U.S. states (WISER). These data are the annual value of exports from each U.S. state to each country in the world. Additionally, the OM data gives the weight of these exports by state and country but *not* the number of exporting firms *by destination*. Cassey (2009) gives a detailed description of the OM data. Because of his findings on data quality, we limit our observations to manufactured exports only. We also restrict our sample to 2003. GDP information is from the World Economic Outlook Database (IMF 2006).

In principle, the agglomeration term  $A_{ij}$  could be measured in several ways. One way would be to count the number of other firms in state  $i$  exporting to country  $j$ . But that cannot be done without firm-level data. At the aggregate level, a more obtainable measure is weight of exports. There is, of course, a mechanical relationship between aggregate exports and aggregate export weight at the product or industry level as an increase in shipment weight must mean an increase in shipment value all else equal. This relationship creates an upward bias in estimation of the coefficient on the agglomeration term. However, the mechanical relationship between weight and value is strongest within narrowly defined product or industrial classification. Across industries the relationship is weaker because, for example, the value of an export shipment of steel and computers is not as related to weight as either of those shipments separately. In the results section below, we report several additional robustness checks using specifications that sever the mechanical relationship between export weight and value to document that any potential upwards bias in our estimates by using aggregate export weight as a measure of agglomeration is less distorting than the bias created by leaving out the agglomeration term entirely in the standard gravity equation.

### 4 Estimates and Discussion

We take (2) to our U.S. state export data and estimate using Ordinary Least Squares (OLS). The benchmark results without the agglomeration term ( $\eta = 0$  as a restriction) is reported in table I column (A) and the unrestricted model is reported in column (B). We estimate the coefficient on the agglomeration term  $\eta\beta = 0.433$ . This is economically significant in that an increase in aggregate weight of 1% increases exports by 0.4%. It is statistically significant with a p-value of less than 0.01.

The estimate on the physical distance coefficient  $\beta = -0.95$  and is statistically significant with a p-value of less than 0.01. From this, we estimate  $\eta = 0.454$ . We confidently reject

Table I. OLS estimates on agglomeration by destination for U.S. data

Var.	(A)	(B)	(C)	(D)	(E)
$D_{ij}$	-2.06* (.137)	-0.95* (.102)	-1.50* (.124)	-1.38* (.030)	-1.25* (.029)
$A_{ij}$		0.433* (.006)			
lag $A_{ij}$			0.243* (.010)		
$A_{ij,-k}$					0.066* (.029)
$N$	7061	7061	6617	167373	167373
$\hat{R}^2$	.541	.750	.600	.411	.413
$RMSE$	1.27	0.93	1.15	1.79	1.80
AIC	23646	19269	20863	670830	674238

Notes: The benchmark gravity equation in (A):  $\log \frac{X_{ij}}{Y_i Y_j} = \alpha - \beta \log D_{ij} + \sum_{i=2}^{50} \kappa_i S_i + \sum_{j=2}^{176} \delta_j T_j + \varepsilon_{ij}$ . Exporter and importer binary variables and a constant are estimated but not reported. Standard errors (in parentheses) are robust. Specification (B) adds the aggregate shipping weight. Specification (C) lags aggregate weight. Specifications (D) and (E) are at the state-country-product level and include a set of industry binary variables.  $A_{ij,-k}$  is aggregate weight less the weight of the industry exports.

\* indicate p-values less than .01.

that  $\eta = 0$  ( $F(1, 6835) = 81.87$ ). Therefore there is evidence of agglomeration of exports by destination.

We check the robustness of this result. First, notice the adjusted  $R^2$  of the unrestricted model is 0.75. It is 0.54 without the agglomeration term  $A_{ij}$ . Thus the agglomeration term accounts for 40% more of the variation in the data than without. Also, recall our estimate  $-\beta = 0.95$ , which is *not* statistically different from one and is close to the 0.99 to 1.10 point estimate range for U.S. data reported in Axtell (2001). In the restricted model,  $-\beta = 2.06$  which is too large to agree with Axtell. Finally we perform a Wald test between the restricted (A) and unrestricted models (B). The restriction harms the fit of the nesting model ( $p$ -value < 0.01). A comparison of the Akaike Information Criterion supports this result.

One concern with the results so far is that they are driven by the mechanical relationship between export value and weight. To break this mechanical relationship, we follow Koenig (2009) and reestimate (2) using the one year lag of state-country export weight instead of current year weight. This severs any mechanical relationship since past aggregate weight cannot affect current export value in any way other than a reduction of the transaction costs of trade. We lose more than a few observations compared to 2003 because many state-country pairs did not trade in 2002 that did in 2003.

The results are reported in table I column (C). In this case, we find a 1% increase in the previous year's aggregate shipping weight increases current year exports by 0.24%, which is both economically and statistically significant, though smaller than the previous specification. Also, we find that the model with agglomeration term accounts for 12% more variation in the data than the benchmark. Therefore, we find that the agglomeration term is important even if it is lagged one year.

Columns (D) and (E) report estimates run at the state-country-product level (4-digit NAICS). In these specifications, the agglomeration variable is aggregate state-country export weight less the state-country-industry weight. As with lagging weight, subtracting out industry-own weight from the agglomeration term breaks the mechanical relationship between industry exports and weight. Once again, the coefficient on the agglomeration term is both economically significant and statistically significant at the 1% level. A Wald test confirms the restriction of  $\eta = 0$  harms the fit of the nesting model ( $p - value < 0.01$ ). A direct test that  $\eta = 0$  fails ( $p - value < 0.01$ ). Thus the significance of the agglomeration term holds at the product level.

## 5 Conclusion

Increased access to customs data has allowed researchers to use firm-level stylized facts to document that exporting firms cluster by destination country of their shipments (Koenig 2009; Choquette and Meinen 2011; Cassey and Schmeiser (*forthcoming*)). Though this evidence suggests that exporter clustering by destination is a general phenomenon, it is unknown whether exporter clustering by destination exists in the United States. That this is unknown means that changes to international trade policy as well as state level industrial policy may have effects on industrial organization that are currently opaque and current models may underestimate the importance of informational barriers to trade on trade patterns. Export promotion policies targeting a specific country based on the standard gravity model may underestimate the impact of the policy if there are already many firms exporting to that destination or overestimate the impact if there are very few exporters shipping to that destination. Understanding that informational barriers affect firm export decisions can help policy makers better anticipate firm responses to export policy changes as well as provide a better understanding of domestic regional planning based on firm locations.

We find evidence of exporter clustering by destination in the United States by estimating and testing the statistical significance of an aggregate agglomeration term using data on U.S. state exports. This agglomeration term is theoretically derived in Cassey and Schmeiser (*forthcoming*) and has a structural interpretation that we use for a robustness check. Our finding of exporter clustering by destination for the United States is in line with the empirical results from transaction-level studies for other countries. But given the size and geographic isolation of the United States compared to European countries, our results speak to the generality of exporter clustering. Our evidence suggests that exporter clustering should be acknowledged and accounted for when designing international trade policy such as export promotion to specific destinations.

## References

- Anderson, J. E. and E. van Wincoop (2003) "Gravity with Gravitas: A Solution to the Border Puzzle" *American Economic Review* **93**, 170–192.
- Axtell, R. L. (2001) "Zipf Distribution of U.S. Firm Sizes" *Science* **293**, 1818–1820.
- Cassey, A. J. (2009) "State Export Data: Origin of Movement vs. Origin of Production" *Journal of Economic and Social Measurement* **34**, 241–268.
- Cassey, A. J. and K. N. Schmeiser (*forthcoming*) "The Agglomeration of Exporters By Destination" *Annals of Regional Science* doi: 10.1007/s00168-012-0538-9.
- Chaney T (2008) "Distorted gravity: The intensive and extensive margins of international trade" *American Economic Review* **98(4)**, 1707–1721.
- Choquette, E. and P. Meinen (2011) "Export Spillovers and the Extensive and Intensive Margins of Trade" URL <http://www.etnpconferences.net/sea/sea2011/PaperSubmissions/Submissions2011/S-F-22.pdf> unpublished.
- IMF (2006) *World Economic Outlook Database* International Monetary Fund Washington, DC <http://www.imf.org/external/pubs/ft/weo/2006/01/data/index.htm> (accessed December 15, 2006).
- Koenig, P. (2009) "Agglomeration and the Export Decisions of French Firms" *Journal of Urban Economics* **66**, 186–195.
- Lovely, M. E., S. S. Rosenthal, and S. Sharma (2005) "Information, Agglomeration, and the Headquarters of U.S. Exporters" *Regional Science and Urban Economics* **35**, 167–191.
- WISER (Various years) *Origin of Movement State Export Data* World Institute for Strategic Economic Research Holyoke, MA <http://www.wisertrade.org> (initially accessed Nov. 15, 2005).