

A PRACTICAL APPROACH TO ESTIMATING
SECTOR-LEVEL SUBSTITUTION ELASTICITIES
WITH PPML

Samantha Schreiber

ECONOMICS WORKING PAPER SERIES
Working Paper 2022–03–A

U.S. INTERNATIONAL TRADE COMMISSION
500 E Street SW
Washington, DC 20436

March 2022

Office of Economics working papers are the result of ongoing professional research of USITC Staff and are solely meant to represent the opinions and professional research of individual authors. These papers are not meant to represent in any way the views of the U.S. International Trade Commission or any of its individual Commissioners. The author thanks David Riker, Saad Ahmad, and Peter Herman for very helpful comments.

A Practical Approach to Estimating Sector-Level Substitution Elasticities with PPML
Samantha Schreiber
March 2022

Abstract

In this paper, we provide a practical approach to estimating the elasticity of substitution by sector using a PPML estimator, building on the methodology described in Riker (2020). The method is illustrated by estimating elasticities for NAICS 3-digit and 6-digit manufacturing sectors. The PPML estimates are compared with estimates produced using OLS and across different levels of aggregation in the data. We find that the PPML estimator tends to produce elasticity estimates that are larger and have greater variability than the OLS estimates.

Samantha Schreiber
Research Division, Office of Economics
samantha.schreiber@usitc.gov

1 Introduction

The elasticity of substitution across varieties of a product is an important parameter in trade policy analysis. The parameter describes the willingness of consumers to shift sourcing after a relative price change. Considerable literature exists to quantify this parameter value using a variety of econometric methods (e.g. Feenstra 1994; Soderbery 2015; Caliendo and Parro 2015; and many others).¹ Riker (2020) provided an econometric method for estimating substitution elasticities by sector using variation in international trade costs with very few data requirements necessary to produce an estimate. We build on their methodology, which employed an Ordinary Least Squares (OLS) estimator, by introducing a Pseudo-Poisson Maximum Likelihood (PPML) estimator as an extension. We estimate substitution elasticities for NAICS 3-digit and 6-digit manufacturing sectors as an illustration of the methodology. PPML estimates are compared with OLS estimates across specifications and different levels of data aggregation. We find that the PPML estimator tends to produce elasticity estimates that are larger and have greater variability than the OLS estimates.

The sections of this paper are organized as follows: section 2 describes the econometric specification used in the paper. Section 3 applies the methodology to NAICS 3-digit and 6-digit manufacturing sectors. This section also draws comparison between methods and levels of aggregation. Finally, in section 4, we conclude and offer areas where the research may be expanded.

2 Methodology

In the following section, we first describe the log-linear trade cost estimator in Riker (2020) and then present a PPML estimator as an extension. Both methods of estimating

¹Ahmad et al (2021) provides a review of the elasticity estimation research by summarizing and comparing sector-level estimates across papers.

the substitution elasticities use a gravity model with variation in international trade costs to identify the parameter, controlling for other demand and supply-side factors with a set of fixed effects. Constant elasticity of substitution (CES) demand for a specific product j from country c into custom's district d in year t is defined as:

$$v_{jcdit} = k_{jct} E_{jit} (P_{jit})^{\sigma_j - 1} (p_{jct} f_{jcdt})^{1 - \sigma_j} (s_{jdit})^{-\sigma_j} \quad (1)$$

where v_{jcdit} is the landed duty-paid value of individual i 's expenditures, k_{jct} is a demand factor that reflects import quality, E_{jit} are total expenditures of individual i on product j from all sources, P_{jit} is the price index, p_{jct} is the producer price of product j from country c , f_{jcdt} are trade costs, and s_{jdit} are domestic shipping costs from district d to the individual. The elasticity of substitution for product j is denoted σ_j . To arrive at the estimating equation, we aggregate across individuals and take logs:

$$\ln v_{jcdt} = \alpha_{jct} + \gamma_{jdt} + \beta_j \ln f_{jcdt} + \epsilon_{jcdt} \quad (2)$$

where $\beta_j = (1 - \sigma_j)$ and ϵ_{jcdt} is the error term. Expressions for country-year (α_{jct}) and district-year (γ_{jdt}) fixed effects can be derived from equation (1):

$$\alpha_{jct} = \ln [k_{jct} (p_{jct})^{1 - \sigma_j}] \quad (3)$$

$$\gamma_{jdt} = \ln \left[\sum_i E_{jit} (P_{jit})^{\sigma_j - 1} (s_{jdit})^{-\sigma_j} \right] \quad (4)$$

Country-year and district-year fixed effects reduce the data requirements in the econometric specification, controlling for variables where data would be difficult to obtain. The country-year fixed effects control for supply-side factors like producer price changes and changes in import quality. The district-year fixed effects control for factors at the destination such as changes in the price index and total expenditures terms.

Next we define the PPML estimating equation. This method was introduced in the gravity literature by Santos Silva and Tenreyro (2006) and has been recommended throughout the literature including in Yotov et al (2016). Taking the exponential of equation 2, the PPML estimating equation can be written as:

$$v_{jcdt} = \exp[\alpha_{jct} + \gamma_{jdt} + \beta_j \ln f_{jcdt}] + \delta_{jcdt} \quad (5)$$

where δ_{jcdt} is the error term for the PPML specification.

There are benefits to using PPML when estimating the substitution elasticities. First, one drawback of OLS is that it cannot take into account the information contained in the zero trade flows. These observations are not included when the dependent variable is logarithmic, as is the case with the OLS estimator, but are included when the model is in multiplicative form. Santos Silva and Tenreyro (2011) show that the PPML estimator is generally well-behaved, even when the proportion of zeros in the estimation data set is large.² Another benefit of using PPML to estimate gravity is the ability to produce consistent estimates when heteroskedasticity is present in the trade data. Santos Silva and Tenreyro (2006) point out that log-linearized models estimated by OLS can lead to misleading conclusions if the data is heteroskedastic. Equation (2) critically relies on the assumption that ϵ_{jcdt} is independent of the regressors. But by Jensen's inequality, the expected value of the log of a random variable does not equal the log of its expected value. If the variance of ϵ_{jcdt} depends on the regressors, it violates the condition of consistency for OLS.

The two methods described above can be implemented with very few data requirements,

²Yotov et al (2016) discuss different solutions to the zero trade flows issue that have been used in the literature. One option is to add very small values to the zeroes so that they are not dropped. But as Head and Mayer (2014) points out, this should be avoided because the amount added to the zeroes is arbitrary and depends on the units of measurement, changing the interpretation of the coefficients. Also, Helpman et al (2008) propose a two-step selection process where the first stage determines the probability to export and the second stage use an OLS estimation that accounts for the selection into exporting due to fixed costs. But this approach can be challenging to implement and it may be difficult to find exclusion restrictions for the first stage. Yotov et al (2016) describe PPML as a convenient solution to the zero trade flows issue.

requiring only a panel of U.S. imports data disaggregated by sector, source country, district of entry and year. We use a practical measure of trade costs (f_{jcdt}): the ratio of landed duty-paid import values and customs values, which includes international freight costs, import charges, and import duties. Using the ratio of the landed duty-paid imports over customs values for trade costs is a convenient method but has a few limitations. First, it has been pointed out in the literature that using ratios as proxies for trade costs may be problematic due to measurement error (Bergstrand et al, 2013). Second, as described above, one of the benefits of PPML is that it can capture information implicit in the zero trade flows. However, these trade flows will not be included in our PPML estimation in the section below because the trade cost measure will also be zero. An area of future research is to use observable determinants like distance as a proxy for trade costs.³

3 Estimation

Next, we illustrate the differences between methods by estimating the substitution elasticities for NAICS 3-digit and 6-digit manufacturing sectors.⁴⁵ A five year panel (2016–20) of U.S. imports data was obtained from USITC’s DataWeb, disaggregated by NAICS industry code, origin country, U.S. district of entry, and year. Elasticities were estimated using country-year district-year fixed effects denoted as directional fixed effects in the tables

³The estimated coefficient of international distance would reflect the trade elasticity as well as the change in the cost of distance. The latter component could be estimated by regressing the trade cost ratio on international distance, and removed from the coefficient to isolate the trade elasticity.

⁴Manufacturing codes were chosen as an example, but this method could be applied to any of the NAICS sectors.

⁵Redding and Weinstein (2019) discuss the implications of aggregating trade data in the gravity model. Data aggregation involves summing the level of trade rather than the log of trade, the left hand side of the gravity equation. This implies that estimating gravity at another level of aggregation may be interpreted as a log-linear approximation of the true gravity relationship. The method described in this paper of estimating the elasticity of substitution may, therefore, be best for more disaggregated sectors. We present both 3-digit and 6-digit estimates as an illustration.

below.⁶

We use the *reghdfe* Stata command for OLS regressions and the *ppmlhdfe* Stata command for PPML regressions.⁷ The dependent variable used in the OLS regression is the log of trade flows and the dependent variable used in the PPML regressions is trade flows in levels. The magnitude of the dependent variable affects the convergence of the PPML estimator, so we re-scaled the PPML dependent variable to be in millions of dollars.

Table 1: Median Elasticity Estimates by Method

	Median OLS estimate, directional fixed effects	Median PPML estimate, directional fixed effects
NAICS 3-digit	5.58	13.23
NAICS 6-digit	5.45	9.08

Note: This table presents median elasticity of substitution estimates for NAICS 3-digit and 6-digit manufacturing sectors. The full set of estimates are reported in the appendix. Directional fixed effects refer to country-year and district-year fixed effects.

Median elasticity estimates are summarized in table 1, with the full set of estimates by NAICS code reported in the appendix. This table illustrates that the estimator choice can lead to significantly different outcomes. Median PPML estimates are larger than median OLS estimates. The pattern is also true at the product level for nearly all of the 3-digit and 6-digit NAICS codes in tables 2 and 3. The distribution of estimates by model and level of disaggregation is shown in figure 1. For both NAICS 3-digit and 6-digit classifications, the distribution of OLS estimates have a lower mean value and are less dispersed than the PPML distributions. To measure variability, we calculated the standard deviation of the estimates. The OLS standard deviation (1.94) was significantly lower than the PPML standard deviation (8.42). The OLS estimates also had a smaller spread than the PPML estimates.

⁶Country-district pair fixed effects could also be included in equations (2) and (5) to control for any time-invariant factors between locations. However, we do not use pair fixed effects because the inclusion of pair fixed effects would remove most of the variation in our trade cost measure.

⁷The *ppmlhdfe* Stata command by Correia, Guimaraes, and Zylkin was used instead of the *ppml* command because of the multiple types of fixed effects. The *glm, family(poisson)* Stata command could also be used.

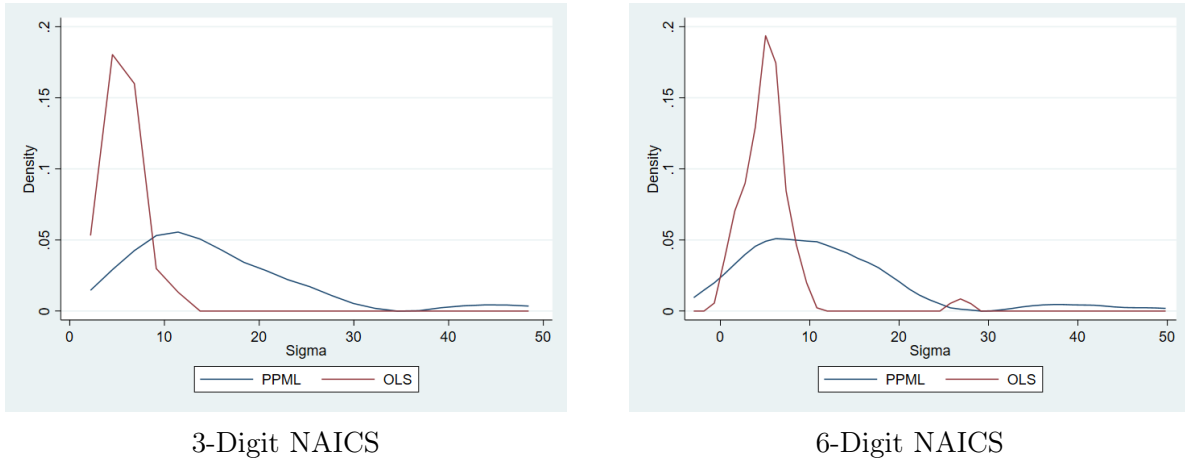


Figure 1: Distribution of Estimates by Method and Aggregation Level

When estimating elasticities over a wide range of sectors, the number of outlier or negative estimates can be a useful measure of plausibility. For the NAICS 3-digit estimates, zero of the OLS estimates and 14 of the PPML estimates were above 10. At the 6-digit level, there were two OLS estimates and 17 PPML estimates either below one or larger than 10.⁸ While the PPML estimates are larger, with many estimates above 10, it is interesting to note that their relative ranking based on magnitude of the estimate is similar to the OLS ranking. Under both methods, NAICS 327 (Nonmetallic Mineral Processing) had the lowest ranked estimate and NAICS 334 (Computer and Electronic Product Manufacturing) had the highest ranked estimate. NAICS 321 (Wood Product Manufacturing) has estimates that were both ranked fourth. 12 of the 21 estimates at the 3-digit level were within three ranks of each other. So despite differences in magnitude, the estimate ranks were roughly preserved.

Comparing estimates with the literature, Fontagne et al (2019) estimated product-level elasticities at the 6-digit Harmonized Tariff Schedule (HTS) level using a PPML estimator. Noting that they estimated all 5,000 product codes and we only have a small sub-sample

⁸It is possible that some product codes may have too little variation in the data to produce a suitable estimate. The product code could have too few country-district observations or little variation in trade costs over the years used in the panel. We checked the number of observations for each NAICS code in the estimation ex ante to eliminate products with few observations. At the more aggregated NAICS 3-digit level, this specific issue is not a concern.

of estimates in this paper, our median and mean estimates are roughly similar to their estimates. The mean substitution elasticity estimate at the HS6 level in Fontagne et al (2019) was 10.8 and the median estimate was 8.3. The mean NAICS 6-digit PPML estimate in this paper is 12.1 and the median estimate is 9.1. Comparing to other estimates in the literature, Broda and Weinstein (2006) report a mean of 6.6 at the SITC 5-digit level and 12.6 at the tariff-line level, and Romalis (2007) reports estimates between 6.2 and 10.9 at the HTS 6-digit level. Other papers have lower averages: for example, Simonovska and Waugh (2014) report a mean of 4.12 and Giri et al (2020) report a median elasticity of 4.38.

4 Conclusion

This paper describes a practical approach to estimating the elasticity of substitution at the sector level with a PPML estimator, building on the methodology described in Riker (2020). The method is practical because there are few data requirements—a panel of import data and a measure of trade costs—with unobservable factors absorbed into the directional and pair fixed effects. As an illustration, the method is applied to NAICS 3-digit and 6-digit manufacturing sectors. Substitution elasticity estimates using PPML are compared to estimates produced with OLS methods and across different levels of aggregation. We find that PPML estimates tend to be larger and have greater variability than the OLS estimates.

Further research in this area could estimate substitution elasticities over a wider variety of NAICS codes to see if the patterns described in this paper continue to hold. Also, this paper identifies patterns in a sample of elasticity estimates, but does not attempt to identify the superiority of one method over another. Additional work is needed to understand which method may be better. Future versions of this paper could also replace the trade cost measure, calculated using the ratio of landed duty-paid import values and customs values, with a proxy based on observable determinants. The methodology could also be extended

further to include domestic trade flows to follow recommendations provided in Yotov et al (2016).

References

- Ahmad, S., Montgomery, C. and Schreiber, S. (2021). A Comparison of Sectoral Armington Elasticity Estimates in the Trade Literature, *Journal of International Commerce and Economics* .
- Anderson, J. E. and van Wincoop, E. (2003). Gravity with Gravitas: A Solution to the Border Puzzle, *American Economic Review* **93**(1): 170–192.
- Bergstrand, J. H., Egger, P. and Larch, M. (2013). Gravity Redux: Estimation of Gravity-Equation Coefficients, Elasticities of Substitution, and General Equilibrium Comparative Statics under Asymmetric Bilateral Trade Costs, *Journal of International Economics* **89**(1): 110–121.
- Broda, C. and Weinstein, D. E. (2006). Globalization and the Gains from Variety, *The Quarterly Journal of Economics* **121**(2): 541–585.
- Caliendo, L. and Parro, F. (2015). Estimates of the Trade and Welfare Effects of NAFTA, *The Review of Economic Studies* **82**: 1–44.
- Feenstra, R. C., Luck, P., Obstfeld, M. and Russ, K. N. (2018). In Search of the Armington Elasticity, *The Review of Economics and Statistics* **100**(1): 135–150.
- Fontagne, L., Guimbar, H. and Orefice, G. (2019). Product-Level Trade Elasticities: Worth Weighting For, *CEPII Working Paper* (2019-17).
- Giri, R., Yi, K.-M. and Yilmazkuday, H. (2021). Gains from Trade: Does Sectoral Heterogeneity Matter?, *Journal of International Economics* **129**: 1–27.
- Head, K. and Mayer, T. (2014). Chapter 3 - Gravity Equations: Workhorse, Toolkit, and Cookbook, *Handbook of International Economics* **4**: 131–195.

- Hertel, T., Hummels, D., Ivanic, M. and Keeney, R. (2007). How Confident can we be of CGE-based Assessments of Free Trade Agreements?, *Economic Modelling* **24**(4): 611–635.
- Redding, S. J. and Weinstein, D. E. (2019). Aggregation and the gravity equation, *AEA Papers and Proceedings* **109**: 1–6.
- Riker, D. (2020). A Trade Cost Approach to Estimating the Elasticity of Substitution, *USITC Economics Working Paper Series 2020-07-D* .
- Romalis, J. (2007). NAFTA’s and CUSFTA’s Impact on International Trade, *The Review of Economics and Statistics* **89**(3): 416–435.
- Santos Silva, J. M. C. and Tenreyro, S. (2006). The Log of Gravity, *The Review of Economics and Statistics* **88**(4): 641–658.
- Santos Silva, J. M. C. and Tenreyro, S. (2011). Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator, *Economic Letters* **112**(2): 220–222.
- Simonovska, I. and Waugh, M. E. (2014). The Elasticity of Trade: Estimates and Evidence, *Journal of International Economics* **92**(1): 34–50.
- Soderbery, A. (2015). Estimating Import Supply and Demand Elasticities: Analysis and Implications, *Journal of International Economics* **96**: 1–17.
- Yotov, Y. V., Piermartini, R., Monterio, J.-A. and Larch, M. (2016). *An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model*, WTO Publications.

Appendix

Table 2: Elasticity of Substitution Estimates at the NAICS 3 digit Level

Naics	OLS estimate, directional fixed effects	PPML estimate, directional fixed effects
311	5.576 (0.244)	14.29 (1.571)
312	3.368 (0.399)	6.472 (1.296)
313	5.219 (0.239)	13.66 (1.099)
314	5.372 (0.222)	20.60 (1.910)
315	6.119 (0.292)	13.23 (1.598)
321	3.437 (0.217)	8.229 (1.091)
322	4.287 (0.289)	11.16 (1.457)
323	3.243 (0.371)	11.15 (1.948)
324	9.290 (0.617)	9.619 (1.530)
325	7.134 (0.355)	26.05 (1.299)
316	5.616 (0.270)	7.204 (1.143)
326	4.071 (0.207)	22.97 (2.567)
327	2.870 (0.196)	5.846 (0.951)
331	6.012 (0.345)	8.954 (0.895)
332	5.073 (0.286)	16.72 (2.174)
333	6.450 (0.439)	15.13 (3.567)
334	10.41 (0.601)	44.80 (5.600)
335	6.707 (0.323)	22.52 (2.885)
336	7.245 (0.555)	10.74 (2.142)
337	4.293 (0.171)	9.304 (1.593)
339	7.110 (0.345)	19.37 (2.291)

Note: This table presents elasticity of substitution estimates at the NAICS 3-digit level for manufacturing products. Point estimates are listed with standard errors in parentheses. Directional fixed effects refer to country-year and district-year fixed effects. Median estimates are reported in the body of the paper.

Table 3: Elasticity of Substitution Estimates at the NAICS 6 digit Level

Naics	OLS estimate, directional fixed effects	PPML estimate, directional fixed effects
311111	7.556 (0.759)	19.46 (3.212)
311119	5.028 (0.570)	8.828 (1.017)
311211	3.339 (0.351)	5.410 (1.050)
311212	4.822 (0.825)	11.94 (1.444)
311213	1.874 (1.120)	1.401 (0.758)
311221	3.203 (0.437)	2.848 (0.639)
311224	7.346 (0.564)	18.42 (1.854)
311225	5.436 (0.754)	9.903 (1.843)
311230	2.484 (0.769)	2.857 (0.929)
31131X	4.050 (0.544)	7.327 (0.874)
311340	5.449 (0.609)	22.32 (2.254)
31135X	8.716 (0.551)	36.86 (2.893)
311411	3.386 (0.585)	0.065 (1.097)
311412	4.489 (2.910)	6.517 (2.935)
311421	4.376 (0.391)	3.259 (0.806)
311422	5.793 (1.181)	15.65 (2.156)
311423	5.949 (0.537)	8.594 (1.153)
311511	1.630 (0.668)	1.168 (0.724)
311512	6.347 (1.253)	5.502 (2.639)
311513	5.663 (1.045)	8.611 (1.485)
311514	6.090 (0.516)	11.94 (1.506)
311520	5.773 (1.368)	6.993 (1.238)
311611	9.194 (0.883)	14.30 (2.331)
311612	7.188 (2.603)	18.82 (8.459)
311613	0.643 (0.596)	1.009 (0.675)
311615	26.85 (4.421)	46.77 (8.884)
311710	8.522 (0.696)	9.076 (1.255)
31181X	5.067 (0.475)	19.12 (3.461)
311824	3.081 (0.641)	16.49 (2.354)
311911	6.027 (0.639)	7.281 (1.584)
311919	2.050 (0.539)	5.830 (1.415)
311920	6.637 (0.394)	13.41 (1.435)
311930	5.488 (0.933)	38.79 (5.940)
311941	5.412 (0.556)	9.090 (1.629)
311942	5.296 (0.396)	10.42 (1.423)
311991	6.678 (1.253)	6.824 (1.594)
311999	7.447 (0.506)	14.29 (1.393)

Note: This table presents elasticity of substitution estimates at the NAICS 6-digit level for manufacturing products under the 3-digit subheading 311. Point estimates are listed with standard errors in parentheses. Directional fixed effects refer to country-year and district-year fixed effects.