# ESTIMATION ISSUES IN SINGLE COMMODITY GRAVITY TRADE MODELS

Sören Prehn & Bernhard Brümmer

sprehn@gwdg.de

Georg August Universität Göttingen

2011

*Vortrag anlässlich der 51. Jahrestagung der GEWISOLA*
*„Unternehmerische Landwirtschaft zwischen Marktanforderungen und gesellschaftlichen Erwartungen"*
*Halle, 28. bis 30. September 2011*

## ESTIMATION ISSUES IN SINGLE COMMODITY GRAVITY TRADE MODELS

## SCHÄTZASPEKTE IN EIN-PRODUKT HANDELSGRAVITATIONSMODELLEN


### Abstract

Recently gravity trade models are applied to disaggregated trade data. Here many zeros are characteristic. In the presence of excess zeros usual Poisson Pseudo Maximum Likelihood (PPML) is still consistent, the variance covariance matrix however is invalid. Correct economic interpretation however requires also the last. So alternative estimators are looked for. STAUB & WINKELMANN (2010) argue that zero-inflated count data models (i.e. zero-inflated Poisson / Negative Binomial Pseudo Maximum Likelihood (ZIPPML / ZINBPML)) are no alternative since under model misspecification these estimators are inconsistent. Yet zero-inflated Poisson Quasi-Likelihood (PQL) is a reliable alternative. It is consistent even under model misspecifications and beyond that robust against unobserved heterogeneity. Another alternative is a log-skew-normal Two-Part Model (G2PM) which generalises the standard log-normal Two-Part Model (2PM). It is insofar advantageous as it adjusts for (negative) skewness and regression coefficients retain usual interpretations as in log-normal models. PQL is useful for multiplicative gravity model estimation and G2PM for log-linear gravity model estimation. Exemplarily the estimators are applied to intra-European piglet trade to assess their empirical performance and applicability for single commodity trade flow analysis. The empirical part favours PQL but G2PM is a reliable alternative for other trade flow analyses. PQL and G2PM should become standard tools for single commodity trade flow analysis.



### Keywords

Gravity Model, Excess Zeros, Poisson Quasi Likelihood, Generalised Two Part Model.



### Zusammenfassung

Seit Kurzem werden Handelsgravitationsmodelle auf disaggregierte Daten angewandt. Hierbei sind viele Nullen charakteristisch. Unter der Präsens von Exzess an Nullen ist die gewöhnliche Poisson Pseudo Maximum Likelihood (PPML) zwar konsistent, die Varianz Kovarianz Matrix ist allerdings nicht korrekt. Eine korrekte Interpretation erfordert allerdings auch letzteres. Von daher sind alternative Schätzer gesucht. Staub & Winkelmann (2010) argumentieren, dass nullinflationierte Zähldatenmodelle (d.h. nuninflationierte Poisson / Negativ Binomial Pseudo Maximum Likelihood (ZIPPML / ZINBPML)) keine Alternative darstellen, da diese Schätzer unter einer Modellmisspezifizierung inkonsistent sind. Doch nullinflationierte Poisson Quasi-Likelihood (PQL) ist eine mögliche Alternative. Diese ist konsistent sogar unter Modellmisspezifizierung und darüber hinaus noch robust gegen unbeobachtete Heterogenität. Eine andere Alternative ist ein log-normales Zwei Teile Modell (2PM). Dieses ist insofern von Vorteil, als es gegen (negative) Schiefe anpasst und die Regressionskoeffizienten die gewöhnlichen Interpretationen wie in log-normalen Modellen beibehalten. Exemplarisch werden die Schätzer auf den innereuropäischen Ferkelhandel angewandt, um deren empirische Güte und Anwendbarkeit für Ein-Produkt Handelsanalysen zu beurteilen. Der empirische Teil favorisiert PQL, G2PM sollte aber für andere Handelsanalyse eine mögliche Alternative sein. PQL und G2PM sollten Standardwerkzeuge der Ein-Produkt Handelsgravitationsmodellanalyse werden.



### Schlüsselbegriffe

Gravitationsmodell, Exzess an Nullen, Poisson Quasi Likelihood, Generalisiertes Zwei-Teile Modell.

## 1 Introduction

Recently research focus in gravity trade model analysis shifted from an aggregated macro-level to a disaggregated micro-level. This shift is not only statistically motivated but also empirically. Statistically it is more appropriate to disaggregate data and then to reaggregate the corresponding micro-level results than to do the whole analysis at an aggregated macro-level. The latter approach potentially suffers under an aggregation bias (ANDERSON & YOTOV, 2010).[1] Empirically the shift is motivated by the fact that policies often have commodity specific economic effects. These effects not only differ markedly between commodities but also the overall effect is often neglectable at macro-level (ANDERSON, 2011). Policy evaluation is better done at micro-level than at macro-level. Hence single commodity gravity trade models are now in the focus of interest in gravity trade model analysis.

Although there are applications of gravity trade models to disaggregated data there are only few papers which explicitly deal with estimation issues of single commodity gravity trade models[2]. An exemption is the paper by BURGER ET AL. (2009). Here the authors seize on the dispute of methods between multiplicative and log-linear gravity models. In their argumentation they partly follow SANTOS SILVA & TENREYRO (2006) that logarithmising leads to biased gravity estimates but they claim that for single commodity trade flow analysis Poisson Pseudo Maximum Likelihood (PPML) is inappropriate. PPML suffers under the problems of excess zeros and overdispersion; the former is a consequence of disaggregation which naturally increases the number of zero trade flows the latter is caused by unobserved heterogeneity.[3] Instead they recommend to apply zero-inflated Poisson Pseudo Maximum Likelihood / Negative Binomial Pseudo Maximum Likelihood (ZIPPML / ZINBPML). The latter approach not only deals with excess zeros but also with overdispersion.

In a recent paper STAUB & WINKELMANN (2010) however show that ZIPPML / ZINBPML are only consistent as long as the underlying model is not misspecified. If the model is misspecified the estimators are biased. The authors further show that PPML is still consistent even in the presence of excess zeros, but the variance covariance matrices are invalid. In their paper the authors instead recommend a zero-inflated Poisson Quasi-Likelihood (PQL). This estimator is not only consistent in the presence of excess zeros but also practically uneffected by unobserved heterogeneity. STAUB & WINKELMANN'S (2010) findings are important since they question BURGER ET AL.'S (2009) statements!

The problem of zero trade flows is also tackled in log-linear gravity models. Different estimators are applied here. In the presence of excess zeros Two-Part Models seem statistically to be the most reliable.[4] HILLBERRY (2002) proposes the use of standard log-normal Two-Part Models (2PM). Since logarithmising sometimes leads to (negatively) skewed distributions, a more general Two-Part Model is preferable. Recently CHAI & BAILEY (2008) develop a log-skew-normal 2PM (G2PM) which not only adjusts for asymmetries but also the regression coefficients still have the usual interpretations as in standard log-normal models.

Consistent estimation is always important for economic interpretation. This applies not only but especially for single commodity trade flow analysis where estimates often suffer beside under unobserved heterogeneity also under excess zeros. New estimators are looked for. This paper therefore proposes PQL and G2PM as reliable alternatives. Both estimators PQL and G2PM are applied to disaggregated data (i.e. intra-European piglet trade) to evaluate their

---

[1] ANDERSON & YOTOV (2010) indicate that ANDERSON & VAN WINCOOP'S (2003) results suffer under a downward bias if they are compared with results generated by disaggregated data.

[2] This paper focuses on estimation issues of single commodity gravity trade models its statements however should also apply for sector gravity trade models.

[3] For statistical details on excess zeros and overdispersion see below.

[4] Tobit estimators are biased in the presence of excess zeros (CHAI & BAILEY, 2008).

empirical performance and applicability. Further the paper recommends statistical testing procedures for model selection; overdispersion tests for Poisson Models and Likelihood Ratio (LR) tests for Two-Part Models. Both estimators and model specification tests should become standard tools for single commodity gravity trade flow analysis!

The focus of this paper is just methodological. So for the theoretical model homogeneous firms trade models (FELBERMAYR & KOHLER, 2010, EGGER & LARCH, 2011) are recommend, but not further discussed. Homogeneous firms trade models have the advantage that they share the same properties as heterogeneous firms trade models (HELPMAN ET AL., 2008), but assume a homogeneous firm structure. This often fits better with market characteristics especially for primary sectors (i.e. agricultural sectors). Theoretical aspects however are not further discussed here.

This paper is organised as follows. The first section deals with the appropriate specification of the theoretical model; homogeneous firms trade models vs. heterogeneous firms trade models and multiplicative gravity models vs. log-linear gravity models. In the second section methodological aspects of single commodity gravity estimation are discussed. The advantage of PQL and G2PM are sketched. The next section then applies both new estimators to intra-European piglet trade. The paper concludes with some recommendations for future single commodity gravity trade model estimation.

## 2 Model Specification

Starting point of every gravity trade model analysis is the specification of an appropriate gravity model. Here the ANDERSON & VAN WINCOOP (2003) (AvW) Model is standard for aggregated data. For disaggregated data however the AvW Model does not fit perfectly. The AvW Model neither deals with zero trade flows which are manifold at micro-level nor it deals with asymmetries in trade flows which are caused by different degrees of specialisation (HELPMAN ET AL., 2008). Helpman et al. propose a heterogeneous firms trade model which simutaneously deals with zero and asymmetric trade flows. A heterogeneous firm structure however does not

always describe market structure adequately. Especially for primary sector (i.e. agriculture) the firm structure is more homogeneous than heterogeneous. Usually farm sector is characterised by many small and rather homogeneous farms. Recently two papers one by FELBERMAYR & KOHLER (2010) and the other one by EGGER & LARCH (2011) develop homogeneous firms trade models.[5] These models have the advantage that they deal with zero and asymmetric trade flows but also restrict firm structure to be homogeneous.

The homogeneous firms trade model depends on two parts. The first part deals with the extensive trade margin whereas the second part deals with the intensive trade margin. Practically spoken the first part asks the question whether trade occurs yes or no and the second part discusses the question to which extent trade takes place. The two parts can either be estimated together in a multiplicative gravity model

$$X_{ij} = \exp(\beta_0 + \lambda_i + \chi_j + \gamma d_{ij}) * u_{ij}$$

or they can be estimated separately. Here the extensive trade margin is estimated via a limited dependent variable model (i.e. Probit or Logit) and the intensive trade margin usually via a log-linear gravity model[6]

[5] FELBERMAYR & KOHLER (2010) develop a Corner Solutions Model specification of the homogeneous firms trade model and EGGER & LARCH (2011) a Two-Part Model specification.

[6] Indeed EGGER & LARCH (2011) also propose to estimate the intensive trade margin in multiplicative form via PPML.

$$\ln(X_{ij}) = \beta_0 + \lambda_i + \chi_j + \gamma d_{ij} + \ln(u_{ij})$$

In both specifications $X_{ij}$ is the import value of importer j from exporter i and $\beta_0$ is a constant which captures the effect of total sector production $Y$. $\lambda_i = y_i + (\sigma - 1)\pi_i$ and $\chi_j = e_j + (\sigma - 1) p_j$ are exporter and importer fixed effects which capture the effects of exporter i's production $Y_i$ and outward multilateral resistance $\Pi_i$ and importer j's expenditure $E_j$ and inward multilateral resistance $P_j$ respectively (ANDERSON, 2010). $d_{ij}$ is the gravitational distance and $u_{ij}$ is an i.i.d. error term.[7]

**3 Implementation and Estimation Issues**

The second step in gravity trade model analysis is the econometric implementation of the theoretical model and its appropriate estimation. Here caution should be applied since not every functional form and estimation method is (always) appropriate. This especially applies for single commodity trade flow analysis where the problem of excess zeros and overdispersion is often existent. Excess zeros thereby means that there are more zeros in the data than predicted by the statistical model. Overdispersion is existent when the observed variance is higher than the variance of the statistical model. Non-consideration of both would lead to biased estimates.

In their seminal paper SANTOS SILVA & TENREYRO (2006) argue gravity models not to estimate in log-linear form but in multiplicative form. Logarithmising could lead to inconsistent estimates. If heteroskedasticity is present Jensen's Inequility would apply (i.e. $\ln(E(x)) \neq E(\ln(x))$ ) which then would render estimates inconsistent. They instead propose a Poisson Pseudo Maximum Likelihood (PPML), i.e. the endogenous $y$ is modeled by a Poisson Distribution Model

$$f^P(y \mid \lambda) = \frac{\exp(-\lambda)\lambda^y}{y!} \qquad , \lambda > 0$$

where the mean parameter is defined as $\lambda = \exp(x'\beta)$. PPML is consistent even in the presence of heteroskedasticity and it has the appeal to deal with zero trade flows.

BURGER ET AL. (2009) extend this framework for more disaggregated data. The authors partly follow SANTOS SILVA & TENREYRO (2006) that logarithmising leads to biased gravity estimates but they claim that for single commodity trade flow analysis PPML is not appropriate. PPML suffers under the problems of excess zeros and overdispersion. These problems have to be treated separately since they are caused by different reasons; excess zeros is caused by disaggregation which naturally increases the number of zero trade flows and overdispersion is caused by unobserved heterogeneity which ususually corresponds to an omitted variable problem. Following Burger et al. the problem of excess zeros alone can be tackled by a zero-inflated Poisson Pseudo Maximum Likelihood (ZIPPML). In the additional presence of overdispersion a zero-inflated Negative Binomal Pseudo Maximum Likelihood (ZINBPML) is appropriate. It is important to mention here that Burger et al. do not develop an asymptotic theory!

An asymptotic theory is just recently developed by STAUB & WINKELMANN (2010). Utilising the framework of GOURIEROUX ET AL.'s (1984a, 1984b) seminal papers on Pseudo Maximum Likelihood (PML) STAUB & WINKELMANN can show that ZIPPML / ZINBPML are only

[7] All small roman (greek) letters indicate logarithms and bold letters vectors.

consistent as long as the underlying model is not misspecified. If the model is misspecified the estimators are biased. The authors further show that PPML is still consistent even in the presence of excess zeros but the variance covariance matrices are invalid. In their paper the authors instead recommend a zero-inflated Poisson Quasi-Likelihood (PQL).[8]

PQL is a Poisson Model shifted by a constant zero-inflation parameter $\pi_i$. A comparison of the corresponding conditional expectation functions (CEF) exemplifies this statistical relation. Shifting the CEF of a Poisson Model $E(y|x) = \lambda = x'\beta$ by following constant term $\ln(1-\pi)$ yields the CEF of a PQL

$$E(y|x) = (1-\pi)\lambda = \exp(\ln(1-\pi) + x'\beta).$$

The zero-inflation parameter $\pi$ is not separately identifiable. It is only estimable in conjunction with the constant term $\beta_0$ of the Poisson Model, i.e. $\tilde{\beta}_0 = \ln(1-\pi) + \beta_0$. This is however of minor importance as the interpretations of the semi-elasticities $\partial(E(y|x)/E(y|x))/\partial x_k$ are not affected hereof (STAUB & WINKELMANN, 2010).[9] \\

PQL has the appeal that it is not only consistent in the presence of excess zeros but also practically uneffected by unobserved heterogeneity (STAUB & WINKELMANN, 2010). STAUB & WINKELMANN'S (2010) findings are important since they question BURGER ET AL.'s (2009) statements!

If the focus is on log-linear gravity model estimation then the importance of excess zeros also increases with the degree of disaggregation. For moderately disaggregated data usual 2-step estimators (i.e. Tobit estimators) are consistent as long as 'the sample proportion of zeros is roughly equivalent to the left tail area of the assumed [statistical model]' (CHAI & BAILEY, 2008, pg. 3644). If the tail-probability constraint is not fulfilled then 2-step estimators are biased. In the presence of excess zeros this is exactly the case which then renders 2-step estimators inappropriate for single commodity trade flow analysis. In the presence of excess zeros Two-Part Models (CRAGG, 1971) which relax the tail-probability contraint seem statistically to be more reliable.

HILLBERRY (2010) is the first proposing a Two-Part Model. HILLBERRY proposes a standard log-normal Two Part Model (2PM) however a more general Two-Part Model seems to be more preferable since logarithmising sometimes leads to (negatively) skewed distributions. Recently CHAI & BAILEY (2008) develop a log-skew-normal Two-Part Model (G2PM) which generalises 2PM. As usual the first part is estimated by a limited dependent variable model (i.e. Logit or Probit) and the second part is modeled by a Skew-Normal Distribution Model

$$f^{SN}(y|\mu,\sigma,\delta) = \frac{2}{\sqrt{\sigma^2+\delta^2}}\phi\left(\frac{y-\mu}{\sqrt{\sigma^2+\delta^2}}\right)\Phi\left(\frac{\delta}{\sigma}\frac{y-\mu}{\sqrt{\sigma^2+\delta^2}}\right)$$

where $\phi$ and $\Phi$ denotes the standard normal density and cumulative distribution function respectively (CHAI & BAILEY, 2008). $y$ is the logarithm of the endogenous Y, $\mu = x'\beta$, and $\sigma$ is the standard deviation. The terms $\sqrt{\sigma^2+\delta^2}$ and $\delta/\sigma$ are the scale and the shape

[8] In their paper STAUB & WINKELMANN (2010) develop two zero-inflated Poisson Quasi-Likelihood one with a constant zero-inflation parameter and the other one with a non-constant zero-inflation parameter. Here only the constant zero-inflated Poisson Quasi-Likelihood is applied given convergence problems of the other estimator.

[9] One important property of PQL is that its estimates are identical with those of PPML only the variance covariance matrices are different.

parameter of the Skew-Normal Distribution respectively. For $\delta / \sigma = 0$ the Skew-Normal Distribution coincides with the Normal Distribution otherwise negative (positive) skewness is given if $\delta / \sigma < 0$ $(\delta / \sigma > 0)$.

The advantage of G2PM over 2PM is that it not only adjusts for asymmetries but also regression coefficients still have the usual interpretations as in standard log-normal models. \\

Beside model specification and estimation model selection is another important task in gravity trade model analysis. So far there are only few papers which deal with testing procedures which allow to judge between different estimation methods. For Poisson Models overdispersion tests are recommended (BURGER ET AL., 2009) and for Two-Part Models Likelihood Ratio (LR) tests.[10]

**4 Application: Intra-European Piglet Trade**

The previous section discusses the statistical superiority of PQL over ZIPPML / ZINBPML and G2PM over 2PM respectively. PQL and G2PM are now applied to intra-European piglet trade to illustrate their empirical applicability.[11] The data set consists of roughly 80 % zero trade flows. Estimation results are presented in Table 1.[12] The homogeneous firms trade model is once estimated in multiplicative form via PPML and PQL and once in log-linear form via 2PM and G2PM. The benchmark model (i.e. traditional gravity model (ANDERSON & VAN WINCOOP, 2003)) is estimated via Ordinary Least Squares (OLS). A fixed effects structure with importer, exporter and time fixed effects is assume for each model.

Import data (i.e. CN8-Code 01039110) are extracted from the Statistical Office of the European Union (Eurostat); physical distance data from the Centre d'Etudes Prospectives et d'Informations Internationales (CEPII); all other data are self-constructed. The data frequency is annual, starting from 2000 to 2009. Following MARQUER (2010) the trade flow analysis concentrates on the eight most important exporters and importers i.e. Austria, Germany, Denmark, Spain, Italy, The Netherlands, Poland and Romania (see Figure 1).

Gravitational distance is approximated by physical distance (DIST), presence of a common border (CONTIG), presence of a common language (COMLANG), a binary indicator for trade between member states and new member states before EU enlargement (NMS_inter) and a binary indicator for trade between new member states before EU enlargement (NMS_intra).

The signs of all models are in accordance. Market characteristics and market developments are adaquately represented by estimation results. A common border favours piglet trade as well as a common language. The trade reversal indicated by the indicator NMS_inter fits well with market developments. After the EU enlargement large commercial hog producers as Smithfield opened up commercial pig plants in Easteurope. This decreases exports to member states of the EU15 as indicated. The other indicator NMS_intra indicates that piglet trade between new member states decreases after EU enlargement. This decrease can be explained by a large decrease in sow stocks in Easteuropean member states after 2005 (MARQUER, 2010). The time fixed effects also seem reasonably to capture global market developments. The time fixed effects follow hog price developments lagged by one year.[13] That piglet producers can not immediately adjust their production to hog price developments seems obvious. Also the 2007 price bubble is immanent. The time fixed effect of 2008 has even a negative sign. In consequence of the high feed prices in 2007 many piglet producers closed business which decreases supply.

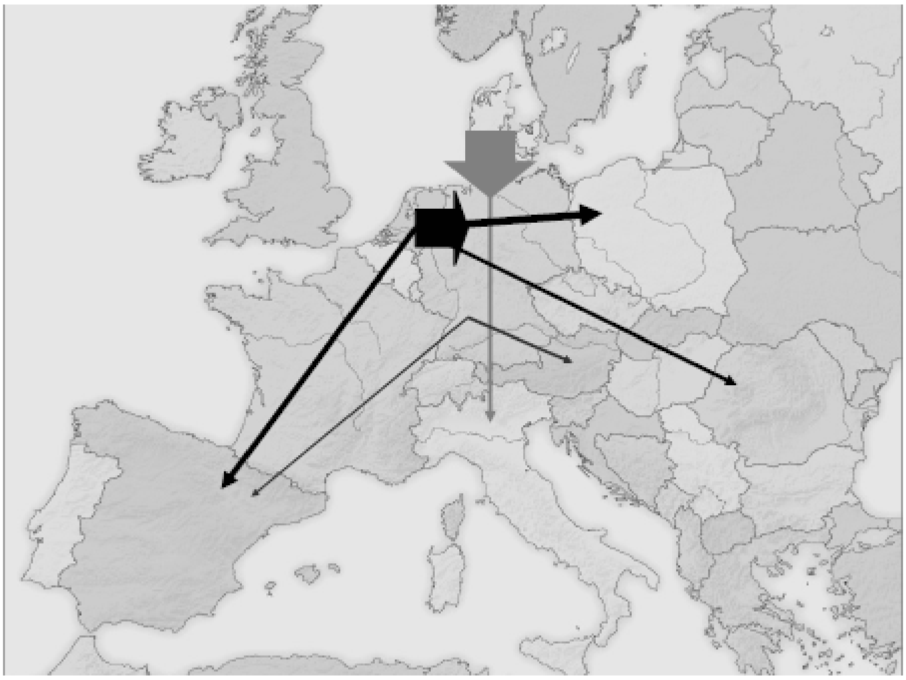
[10] General tests to decide between different model specifications (e.g. Poisson vs. Two-Part Model) are at the moment at the research frontier (SANTOS SILVA ET AL., 2010).

[11] For details on European pig farming and recent developments see MARQUER (2010).

[12] All estimations are done in R http://www.r-project.org/. Program code is available on request.

[13] Price charts can be found under http://www.bordbia.ie/industryservices/pig/pages/prices.aspx.

Figure: Intra-European Piglet Trade (Major Exporters / Importers)



Source: Own representation leant on Marquer (2010).

The only striking feature among the different models is the changing sign of the physical distance variable. This however is of minor importance as the physical distance variable is always insignificant. The major exporters (i.e. Denmark and The Netherlands) are located in the middle of Europa whereas the major importers (i.e. Spain, Italy, Romania) are located at periphery (see Figure 1). The german effect is captured mainly by the contiguity and the common border variable. This renders physical distance insignificant.

One question is still open the choice of the most appropriate model. Here an overdispersion test indicates overdispersion which then requires a PQL for the multiplicative gravity model. Overdispersion is also indicated by the change in standard errors between PPML and PQL. Also the change here is neglectable, this has not always to be the case. The other comparison is between the 2PM and the G2PM. Here a Likelihood Ratio (LR) test does not indicate negative skewness. So for the log-linear gravity model a 2PM estimation is appropriate. The final statistical question is then which model specification is to be favoured. Here a comparison between the significance levels of PQL and 2PM favours PQL. So intra-European piglet trade is best modeled by means of a gravity model in multiplicative form estimated by PQL.

## 5 Conclusions

Recently research focus in gravity trade model analysis shifted from an aggregated macro-level to a disaggregated micro-level. Gravity trade model analysis now focuses on single commodity trade flows. Consistent estimation however is as important as ever before. This paper therefore deals with estimation issues of single commodity gravity trade models.

This paper argues that in the presence of excess zeros usual PPML is not appropriate since variance covariance matrices are invalid. Zero-inflated count data models (i.e. ZIPPML / ZINBPML) are no alternative since under model misspefication these estimators are inconsistent. A reliable alternative is PQL. PQL is consistent even in the presence of model misspecifications and also robust against unobserved heterogeneity (STAUB & WINKELMANN, 2010). For log-linear gravity models G2PM is another alternative. G2PM is insofar

Table 1: **Overview Gravity Model Estimation Results**

| | LOGIT | OLS | PPML | PQL | 2PM | G2PM |
|---|---|---|---|---|---|---|
| INTERCEPT | −10.473*** | 11.915*** | −4.450 | −4.450 | 11.915*** | 10.724*** |
| | (3.470) | (4.329) | (5.499) | (5.428) | (4.329) | (4.205) |
| log(DIST) | 0.841** | −0.305 | 0.869 | 0.869 | −0.305 | −0.152 |
| | (0.375) | (0.566) | (0.709) | (0.702) | (0.566) | (0.547) |
| CONTIG | 0.844** | 0.393 | 2.377*** | 2.377*** | 0.393 | 0.372 |
| | (0.390) | (0.897) | (0.904) | (0.894) | (0.897) | (0.858) |
| COMLANG | −0.381 | 5.473*** | 2.774*** | 2.774*** | 5.473*** | 5.398*** |
| | (0.385) | (1.094) | (0.999) | (0.989) | (1.094) | (1.126) |
| NMS_inter | 2.382*** | 1.222 | 4.420*** | 4.420*** | 1.222 | 1.382 |
| | (0.807) | (1.297) | (1.032) | (1.030) | (1.297) | (1.301) |
| NMS_intra | 3.042*** | 0.238 | 5.529*** | 5.529*** | 0.238 | 0.410 |
| | (1.031) | (1.301) | (0.963) | (0.955) | (1.301) | (1.280) |
| FE_2000 | 1.041 | 0.965 | 3.215*** | 3.215*** | 0.965 | 1.110 |
| | (0.646) | (1.059) | (0.743) | (0.732) | (1.059) | (1.052) |
| FE_2001 | 1.322** | 2.965*** | 4.297*** | 4.297*** | 2.965*** | 2.964*** |
| | (0.590) | (1.112) | (0.671) | (0.647) | (1.112) | (1.088) |
| FE_2002 | 1.586*** | 1.688 | 6.480*** | 6.480*** | 1.688 | 1.839* |
| | (0.577) | (1.140) | (0.883) | (0.874) | (1.140) | (1.131) |
| FE_2003 | 0.249 | 1.112 | 1.771** | 1.771** | 1.112 | 1.049 |
| | (0.592) | (1.178) | (0.724) | (0.712) | (1.178) | (1.159) |
| FE_2004 | 0.164 | 0.764 | 1.102 | 1.102 | 0.764 | 0.784 |
| | (0.560) | (1.636) | (0.741) | (0.726) | (1.636) | (1.650) |
| FE_2005 | 0.612 | 1.844* | 3.536*** | 3.536*** | 1.844* | 1.908** |
| | (0.545) | (1.039) | (0.621) | (0.604) | (1.039) | (1.020) |
| FE_2006 | −0.018 | 0.339 | 1.949** | 1.949** | 0.339 | 0.389 |
| | (0.611) | (0.975) | (0.814) | (0.807) | (0.975) | (0.991) |
| FE_2007 | 0.857 | 4.549*** | 6.921*** | 6.921*** | 4.549*** | 4.809*** |
| | (0.698) | (1.237) | (0.767) | (0.757) | (1.237) | (1.255) |
| FE_2008 | −0.253 | −1.319 | −1.274 | −1.274 | −1.319 | −1.269 |
| | (0.698) | (1.088) | (0.921) | (0.892) | (1.088) | (1.071) |
| No. of Obs. | 560 | 114 | 560 | 560 | 114 | 114 |

*Notes:* Importer, exporter, and year fixed effects. Robust standard errors (clustering by country pair).
*Signif. levels:* 0 '***' 0.01 '**' 0.05 '*' 0.1 ' ' 1

advantageous as it adjusts for (negative) skewness and regression coefficients still retain usual interpretations as in standard log-normal models (CHAI & BAILEY, 2008).

The empirical application done in this paper favours PQL over G2PM. This however is not to be generalised rather research should always follow statistical testing procedures and exclude step by step different alternatives. Overdispersion tests should become standard tools for multiplicative gravity model analysis as well as LR tests for Two-Part Models. PQL and G2PM should become standard tools for single commodity gravity trade model estimation!

## References


ANDERSON, JAMES E. (2011): The Gravity Model. In: Annual Review of Economics 3, forthcoming.

ANDERSON, JAMES E. (2010): The incidence of gravity. In: Brakman, Steven and Bergeijk (eds.): The Gravity Model of International Trade: Advances and Applications. Cambridge University Press.

ANDERSON, JAMES E. AND VAN WINCOOP, ERIC (2003): Gravity with Gravitas: A Solution to the Border Puzzle. In: The American Economic Review 93 (1): 170-192.

ANDERSON, JAMES E. AND YOTOV, YOTO V. (2010): The Changing Incidence of Geography. In: The American Economic Review 100 (5): 2157-2186.

BURGER, MARTIJN, VAN OORT, FRANK AND LINDERS, GERT-JAN (2009): On the specification of the Gravity Model of Trade: Zeros, Excess Zeros and Zero-inflated Estimation. In: Spatial Economic Analysis 4 (2): 167-190.

CHAI, HIGH S. AND BAILEY, KENT R. (2008): Use of log-skew-normal distribution in analysis of continuous data with a discrete component at zero. In: Statistics in Medicine 27 (18): 3643-3655.

CRAGG, JOHN G. (1971): Some Statistical Model for Limited Dependent Variables with Application to the Demand for Durable Goods. In: Econometrica 39 (5): 829-844.

EGGER, PETER AND LARCH, MARIO (2011): An assessment of the Europe agreements' effects on bilateral trade, GDP, and welfare. In: European Economic Review 55 (2): 263-279.

FELBERMAYR, GABRIEL J. AND KOHLER, WILHELM (2010): Modelling the Extensive Margin of World Trade: New Evidence on GATT and WTO Membership. In: The World Economy 33 (11): 1430-1469.

GOURIEROUX, CHRISTIAN, MONFORT, ALAIN AND TROGNON, ALAIN (1984A): Pseudo Maximum Likelihood Methods: Theory. In: Econometrica 52 (3): 681-700.

GOURIEROUX, CHRISTIAN, MONFORT, ALAIN AND TROGNON, ALAIN (1984B): Pseudo Maximum Likelihood Methods: Applications to Poisson models. In: Econometrica 52 (3): 701-721.

HELPMAN, ELHANAN, MELITZ, MARC AND RUBINSTEIN, YONA (2008): Estimating Trade Flows: Trading Partners and Trading Volumes. In: Quarterly Journal of Economics 123 (2): 441-487.

HILLBERRY, RUSSELL H (2002): Aggregation Bias, Compositional Change, and the Border Effect. In: The Canadian Journal of Economics 35 (3): 517-530.

MARQUER, POL (2010): Pig farming in the EU, a changing sector. In: eurostat Statistics in focus 8/2010.

SANTOS SILVA, JOAO M. C. AND TENREYRO, SILVANA (2006): The Log of Gravity. In: The Review of Economics and Statistics 88 (4): 641-658.

SANTOS SILVA, JOAO M. C., TENREYRO, SILVANA AND WINDMEIJER, FRANK (2010): Is it different for zeros? Discriminating between models for non-negative data with many zeros. In: CEMMAP Working Paper CWP20/10.

STAUB, KEVIN E. AND WINKELMANN, RAINER (2010): Quasi-likelihood estimation of zero-inflated count model. In: Working Paper.