The Log of Heavy-Tailed Gravity

Ruben Dewitte, and Samuel Standaert*

July, 2022

Abstract

The gravity equation has been the workhorse model for analyzing the determinants of bilateral trade flows for over 50 years. Predominantly, this gravity equation is estimated using Poisson Pseudo-Maximum-Likelihood (PPML). This paper demonstrates that PPML yields imprecise estimates if the underlying data is heavy-tailed, which is the case for trade data. Consequently, we argue that the gravity equation should be estimated using Quasi-Generalized Pseudo-Maximum-Likelihood estimators (QGPML). QGPML estimators exploit the variance structure of the data to deliver asymptotically uniformly better estimates than the PPML estimator. We demonstrate the superiority of QGPML over PPML through a Monte Carlo exercise and an application to bilateral trade data.

Keywords: Gravity, International Trade, Heavy-tailed Data, Pseudo-Maximum-Likelihood

JEL Codes: C13, C18, F14

*Dewitte (corresponding author): Ghent University, Tweekerkenstraat 2, 9000 Ghent, Belgium, rubenl.dewitte@ugent.be; Standaert: United Nations University and Ghent University, Potterierei 72, 8000 Brugge, sstandaert@cris.unu.edu. Ruben Dewitte gracefully acknowledges financial support from the Research Foundation Flanders (FWO) grant 12B8822N.

1 Introduction

The gravity equation provides a parsimonious and tractable representation of economic interaction in a many country world. It has been the workhorse model for analyzing the determinants of bilateral trade flows for over 50 years. This popularity exists despite the challenges bilateral trade flows pose for the applied researcher. The distribution of bilateral trade data is is non-negative with many zero values, heteroskedastic, positively skewed and heavy-tailed. The seminal paper of Silva and Tenreyro (2006) led to a predominant reliance on the Poisson Pseudo-Maximum-Likelihood (PPML) estimator to obtain efficient and consistent gravity estimates from non-negative, heteroskedastic trade flows with many zero values (see Yotov (2022) for an overview of gravity applications). Manning and Mullahy (2001) argue, however, that the PPML estimator can yield very imprecise estimates if, additionally, the underlying data is heavy-tailed.

This paper evaluates the performance of gravity estimation methodologies in non-negative trade data with many zero values that is heteroskedastic, positively skewed and heavy-tailed. We confirm the finding of Manning and Mullahy (2001) that pseudo-maximum-likelihood estimators such as PPML yield very imprecise estimates if trade data is heavy-tailed. Consequently, we argue that the gravity equation should be estimated with Quasi-Generalized Pseudo-Maximum-Likelihood estimators (QGPML). QGPML estimators exploit the variance structure of the data to deliver asymptotically uniformly better estimates than the PPML estimator (Gourieroux et al., 1984b), with a pronounced improvement in heavy-tailed data.

We extend the Monte Carlo exercise of Silva and Tenreyro (2006) to deliver heavy-tailed trade data, in line with empirical evidence. Simulations based on this extended Monte Carlo exercise demonstrate the difficult of the PPML to recover the true coefficients precisely. QGPML methods, on the other hand, are shown to mimic the PPML estimator in situations of low variance, and to efficiently recover the true coefficient estimates in cases of high variance.

An application to real data confirms the applicability of the QGPML estimators.

2 (Quasi-Generalized) Pseudo-Maximum-Likelihood estimators

The reliance on pseudo-maximum-likelihood estimators to estimate the gravity equation date back to Gourieroux et al. (1984a,b). Imagine bilateral export flows Y from country $i \in \{1, \ldots, I\}$ to country $j \in \{1, \ldots, J\}$ at time $t \in \{1, \ldots, T\}$ as a function of B (discrete and continuous) explanatory variables, gathered in the matrix $\boldsymbol{x}_{ijt}$ of dimension $IJT \times B$, and an i.i.d. disturbance term μ_{ijt}:

$$Y_{ijt} = e^{\boldsymbol{x}_{ijt}\boldsymbol{\beta}} \mu_{ijt}. \tag{1}$$

Gourieroux et al. (1984a,b) showed that consistent and asymptotically normal estimators of the parameter vector $\boldsymbol{\beta}$ can be obtained without specifying the probability density function (p.d.f.) of the disturbance. They consider PML estimators for $\boldsymbol{\beta}$ associated with linear exponential

Table 1: First-order conditions for the PML and QGPML estimators based of different linear exponential families.

Family	**PML**	**QGPML**
Normal	$2\boldsymbol{x}_{ijt}^{T}\left(Y_{ijt}-e^{\boldsymbol{x}_{ijt}\boldsymbol{\beta}}\right)e^{\boldsymbol{x}_{ijt}\boldsymbol{\beta}}$	$\frac{-2\boldsymbol{x}_{ijt}^{T}\left(Y_{ijt}-e^{\boldsymbol{x}_{ijt}\boldsymbol{\beta}}\right)e^{\boldsymbol{x}_{ijt}\boldsymbol{\beta}}}{e^{\boldsymbol{x}_{ijt}\hat{\boldsymbol{\beta}}}+\hat{\eta}^{2}e^{2\boldsymbol{x}_{ijt}\hat{\boldsymbol{\beta}}}}$
Poisson	$-\boldsymbol{x}_{ijt}^{T}\left(Y_{ijt}-e^{\boldsymbol{x}_{ijt}\boldsymbol{\beta}}\right)$	-
Gamma	$-\boldsymbol{x}_{ijt}^{T}\left(Y_{ijt}-e^{\boldsymbol{x}_{ijt}\boldsymbol{\beta}}\right)e^{-\boldsymbol{x}_{ijt}\boldsymbol{\beta}}$	$-\boldsymbol{x}_{ijt}^{T}\left(Y_{ijt}-e^{\boldsymbol{x}_{ijt}\boldsymbol{\beta}}\right)e^{-\boldsymbol{x}_{ijt}\boldsymbol{\beta}}\frac{e^{\boldsymbol{x}_{ijt}\hat{\boldsymbol{\beta}}}}{1+\hat{\eta}^{2}e^{\boldsymbol{x}_{ijt}\hat{\boldsymbol{\beta}}}}$
Negative binomial	-	$\frac{-\boldsymbol{x}_{ijt}^{T}\left(Y_{ijt}-e^{\boldsymbol{x}_{ijt}\boldsymbol{\beta}}\right)}{1+\hat{\eta}^{2}e^{\boldsymbol{x}_{ijt}\boldsymbol{\beta}}}$

Notes: PML stands for Pseudo-Maximum-Likelihood, and QGPML stands for Quasi-Generalized Pseudo-Maximum-Likelihood. $\hat{\boldsymbol{\beta}}$ and $\hat{\eta}$ indicate consistent estimates of the parameter vector $\boldsymbol{\beta}$ and the disturbance's standard deviation η^2 obtained in a preceding estimation stage.

families such as the Normal family with unit variance (i.e. the Nonlinear Least Squares (NLLS) estimator), the Poisson family, and the Gamma family by exploiting solely the first moment of the conditional distribution:

$$E\left[Y_{ijt}|\boldsymbol{x}_{ijt}\right]=e^{\boldsymbol{x}_{ijt}\boldsymbol{\beta}}; \tag{2}$$

where it is assumed there is a constant term in $\boldsymbol{x}_{ijt}\boldsymbol{\beta}$ such that $E\left[\mu_{ijt}\right]=1$. Table 1 provides the First-order conditions (FOCs) related to each specific family of PML estimators in the column PML.

As Gourieroux et al. (1984b) note, there does not exist one of the previous PML estimators that is uniformly better than the others, i.e. with a smaller asymptotic covariance matrix for all possible distributions of the disturbance term. If $\mu_{ijt}=1$, the PML estimator based on the Poisson family is asymptotically efficient and better than the other three estimators. If μ_{ijt} has a log-gamma distribution, the PML estimator based on the negative binomial family is asymptotically efficient and better than the others. Silva and Tenreyro (2006) argue that, in practice, the PPML estimator is the most robust estimator for different distributions of the distribution term. The PML estimator from the Normal family tend to provide too much weight to noisier observations, where $e^{\boldsymbol{x}_{ijt}\boldsymbol{\beta}}$ is large. The PML estimator for the Gamma family, on the other hand, downweights these larger observations and gives too much weight to observations of lesser quality, where $e^{\boldsymbol{x}_{ijt}\boldsymbol{\beta}}$ is small.

It is, however, possible to build estimators which are uniformly better than the previous PML estimators. Quasi-Generalized PML estimators exploit the second moment of the conditional distribution:

$$V\left[Y_{ijt}|\boldsymbol{x}_{ijt}\right] = e^{\boldsymbol{x}_{ijt}\boldsymbol{\beta}} + \eta^2 e^{2\boldsymbol{x}_{ijt}\boldsymbol{\beta}}, \tag{3}$$

where the variance of the disturbance term is defined as $V\left[\epsilon_{ijt}\right] \equiv \eta^2$, to construct more efficient estimators than the PML estimators (Gourieroux et al., 1984a,b).

These QGPML estimators require a consistent estimator $\hat{\eta}^2$ of η^2, however. As Bosquet and Boulhol (2014) note, the original estimator for η^2 proposed by (Gourieroux et al., 1984b) suffered from scale-dependence. Therefore, we follow Bosquet and Boulhol (2014) in the two-step approach to obtain a scale-independent estimate of η^2. In a first step, we rely on any previously mentioned PML estimator to obtain a consistent estimate for $\boldsymbol{\beta}$. In a second step, we estimate $\hat{\eta}^2 = \frac{\hat{b}}{\hat{a}}$ from:

$$\frac{\left(Y_{ijt} - e^{\boldsymbol{x}_{ijt}\hat{\boldsymbol{\beta}}}\right)^2}{e^{\boldsymbol{x}_{ijt}\hat{\boldsymbol{\beta}}}} = a + b e^{\boldsymbol{x}_{ijt}\hat{\boldsymbol{\beta}}} + \varepsilon_{ijt}. \tag{4}$$

The FOCs related to each specific family of QGPML estimators based on the consistent estimates $\hat{\boldsymbol{\beta}}$ and $\hat{\eta}^2$ are provided in Table 1 in the column QGPML. Notice that as $\hat{\eta}^2 \to 0$, the respective FOC of the QGPML estimators converge to the PPML estimator.

3 Monte Carlo Evidence

Our Monte Carlo evidence relies on a very simple extension of the original Monte Carlo exercise of Silva and Tenreyro (2006) which increases the variance of the independent continous variable and, as a result, increases the unconditional variance to reflect reality better. We simulate the multiplicative model

$$E\left[y_i \mid x\right] = \exp\left(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}\right), \qquad i = 1, \ldots, 1000, \tag{5}$$

where x_{1i} is drawn from a normal distribution with mean zero and standard deviation equal to two (relative to a standard normal distributed variable in Silva and Tenreyro (2006)), and x_2 is a binary dummy variable that equals 1 with a probability of 0.4 The two covariates are independent, and a new set of observations of all variables is generated in each replication using $\beta_0 = 0, \beta_1 = \beta_2 = 1$. Data on y are generated as

$$y_i = \mu\left(x_i\beta\right)\eta_i,$$

where η_i is a log normal random variable with mean 1 and variance σ_i^2.

To assess the performance of the estimators under different patterns of heteroskedasticity, we follow Silva and Tenreyro (2006) in considering the four following specifications of σ_i^2:

Case 1: $\sigma_i^2 = e^{-2\boldsymbol{x}_i\boldsymbol{\beta}}; V\left[y_i \mid x\right] = 1;$

Case 2: $\sigma_i^2 = e^{-\boldsymbol{x}_i\boldsymbol{\beta}}; V\left[y_i \mid x\right] = e^{-\boldsymbol{x}_i\boldsymbol{\beta}};$

Case 3: $\sigma_i^2 = 1; V[y_i \mid x] = e^{2\boldsymbol{x}_i\boldsymbol{\beta}}$;

Case 4: $\sigma_i^2 = e^{-\boldsymbol{x}_i\boldsymbol{\beta}} + \exp(x_{2i}); V[y_i \mid x] = e^{\boldsymbol{x}_i\boldsymbol{\beta}} + \exp(x_{2i})\, e^{2\boldsymbol{x}_i\boldsymbol{\beta}}$.

We perform 1.000 Monte Carlo simulations and estimate for each simulate the model using the PML and QGPML estimators discussed in section 2.

This finding is in line with Head and Mayer (2014) who find sizable bias for the Poisson due to the high coefficient of variation in their simulation, which is calibrated on real data.

Table 2: Wide (2) replication results Log of Gravity Table 1.

	β_1		β_2		$\hat{\eta}^2$	
	Bias	S.E.	Bias	S.E.	Est.	S.E.
Case 1: $V[y_i\|x] = 1$						
OLS	0.43075	0.030	0.40843	0.082		
NLLS	0.21991	5.391	-1.39153	33.773		
GPML	0.02864	0.068	0.00782	0.162		
PPML	-0.00008	0.004	-0.00047	0.007		
QGNLLS	-0.19157	6.031	0.02539	0.892	0.00000	0.000
G-QGPML	-0.00008	0.004	-0.00047	0.007	0.00000	0.000
NB-QGPML	-0.00008	0.004	-0.00047	0.007	0.00000	0.000
Case 2: $V[y_i\|x] = \mu(x_i\beta)$						
OLS	0.22119	0.020	0.21017	0.061		
NLLS	0.06709	1.710	-0.29479	6.164		
GPML	0.00796	0.039	-0.00283	0.097		
PPML	0.00003	0.005	-0.00117	0.019		
QGNLLS	0.00013	0.005	-0.00112	0.019	0.00057	0.001
G-QGPML	0.00015	0.005	-0.00111	0.019	0.00057	0.001
NB-QGPML	0.00014	0.005	-0.00111	0.019	0.00057	0.001
Case 3: $V[y_i\|x] = \mu(x_i\beta)^2$						
OLS	0.00046	0.013	-0.00301	0.054		
NLLS	1.37043	13.758	-3.55420	44.498		
GPML	0.00066	0.015	-0.00503	0.065		
PPML	-0.00703	0.092	-0.00907	0.251		
QGNLLS	-0.00430	0.044	-0.00428	0.132	0.29449	1.609
G-QGPML	0.00537	0.045	0.00138	0.132	0.29449	1.609
NB-QGPML	0.00417	0.045	0.00019	0.132	0.29449	1.609
Case 4: $V[y_i\|x] = \mu(x_i\beta) + exp(x_{2i})\mu(x_i\beta)^2$						
OLS	0.15796	0.023	-0.08077	0.083		
NLLS	1.69066	11.335	-3.90679	48.055		
GPML	0.00804	0.042	-0.00950	0.130		
PPML	-0.01221	0.122	-0.02756	0.330		
QGNLLS	-0.01161	0.056	-0.01360	0.167	0.19283	0.620
G-QGPML	0.00632	0.064	-0.00253	0.175	0.19283	0.620
NB-QGPML	0.00370	0.061	-0.00408	0.172	0.19283	0.620

4 Real-life Application

. . .

Table 3: Narrow replication results Log of Gravity Table 3.

Dependent Variable	NLLS	QGNLLS	GPML	QGGPML	QGNBPML	PPML	Q
Constant	-40.891	-39.971	-36.523	-29.511	-29.628	-32.326	
Log exporter's GDP	-59.998	-32.637	0.978	0.758	0.752	0.732	
Log importer's GDP	-58.987	-32.026	0.859	0.757	0.753	0.741	
Log exporter's GDP per capita	-21.582	-12.008	0.297	0.199	0.204	0.157	
Log importer's GDP per capita	-21.432	-11.774	0.254	0.176	0.180	0.135	
Log distance	-16.518	-10.163	-1.477	-0.827	-0.816	-0.784	
Contiguity dummy	-1.735	-0.368	0.204	0.191	0.224	0.193	
Common-language dummy	-0.463	0.381	0.692	0.870	0.870	0.746	
Colonial-tie dummy	-0.500	0.264	0.390	-0.070	-0.088	0.025	
Landlocked-exporter dummy	-0.516	-0.605	-0.589	-0.565	-0.543	-0.863	
Landlocked-importer dummy	-0.770	-0.806	-0.881	-0.678	-0.677	-0.696	
Exporter's remoteness	-17.708	-9.496	1.020	0.469	0.462	0.660	
Importer's remoteness	-18.215	-10.046	0.113	0.292	0.322	0.561	
Free-trade agreement dummy	-1.244	-0.100	1.549	0.004	-0.017	0.181	
Openess	-2.220	-1.278	-0.414	-0.380	-0.394	-0.107	
Nr. observations	18360	18360	18360	18360	18360	18360	
a	NA	1.1e-06	NA	1.1e-06	1.1e-06	NA	1

Table 4: Narrow replication results Log of Gravity Table 4.

Dependent Variable	NLLS	QGNLLS	GPML	QGGPML	QGNBPML	PPML	QI
Log distance	-484.365	-508.940	-1.933	-0.910	-0.870	-0.750	
Contiguity dummy	-62.981	-26.529	-0.457	0.381	0.411	0.370	
Common-language dummy	-23.964	-27.684	0.681	0.213	0.219	0.383	
Colonial-tie dummy	-2.426	-6.525	0.808	0.440	0.428	0.079	
Free-trade agreement dummy	-72.713	-54.657	1.472	0.113	0.142	0.376	
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	
Nr. observations	18360	18360	18360	18360	18360	18360	
a		1.1e-06		1.1e-06	1.1e-06		9.

5 Conclusion

...

References

Bosquet, C. and H. Boulhol (2014, October). Applying the GLM Variance Assumption to Overcome the Scale-Dependence of the Negative Binomial QGPML Estimator. *Econometric Reviews 33*(7), 772–784. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07474938.2013.806102.

Gourieroux, C., A. Monfort, and A. Trognon (1984a). Pseudo Maximum Likelihood Methods: Applications to Poisson Models. *Econometrica 52*(3), 701–720. Publisher: [Wiley, Econometric Society].

Gourieroux, C., A. Monfort, and A. Trognon (1984b). Pseudo Maximum Likelihood Methods: Theory. *Econometrica 52*(3), 681–700. Publisher: [Wiley, Econometric Society].

Head, K. and T. Mayer (2014, January). Chapter 3 - Gravity Equations: Workhorse,Toolkit, and Cookbook. In G. Gopinath, E. Helpman, and K. Rogoff (Eds.), *Handbook of International Economics*, Volume 4 of *Handbook of International Economics*, pp. 131–195. Elsevier.

Manning, W. G. and J. Mullahy (2001, July). Estimating log models: to transform or not to transform? *Journal of Health Economics 20*(4), 461–494.

Silva, J. M. C. S. and S. Tenreyro (2006, November). The Log of Gravity. *The Review of Economics and Statistics 88*(4), 641–658.

Yotov, Y. (2022). Gravity at Sixty: The Workhorse Model of Trade. *SSRN Electronic Journal*.

Online Appendix to "The Log of Heavy-Tailed Gravity"

Ruben Dewitte and Samuel Standaert*

July, 2022

*Dewitte (corresponding author): Ghent University, Tweekerkenstraat 2, 9000 Ghent, Belgium, rubenl.dewitte@ugent.be; Standaert: United Nations University and Ghent University, Potterierei 72, 8000 Brugge, sstandaert@cris.unu.edu. Ruben Dewitte gracefully acknowledges financial support from the Research Foundation Flanders (FWO) grant 12B8822N.

Appendix A Additional Figures and Tables

A.1 Figures

A.2 Tables

Appendix References