Economics Bulletin

Volume 40, Issue 2

Conditional GMM estimation for gravity models

Masaya Nishihata
Mitsubishi UFJ Research and Consulting Co., Ltd.

Taisuke Otsu
London School of Economics and Keio Economic Observatory

Abstract

This paper studies finite sample performances of the conditional GMM estimators for a particular conditional moment restriction model, which is commonly applied in economic analysis using gravity models of international trade. We consider the GMM estimator with growing moments and Dominguez and Lobato's (2004) process-based GMM estimator. Under the simulation designs by Santos Silva and Tenreyro (2006, 2011), we find that Dominguez and Lobato's (2004) estimator is favorably comparable with the Poisson pseudo maximum likelihood estimator, and outperforms other estimators.

Thanks: The authors would like to thank anonymous referees for helpful comments.
Citation: Masaya Nishihata and Taisuke Otsu, (2020) "Conditional GMM estimation for gravity models", *Economics Bulletin*, Volume 40, Issue 2, pages 1106-1111
Contact: Masaya Nishihata - nishihata@murc.jp, Taisuke Otsu - t.otsu@lse.ac.uk.
Submitted: November 03, 2019. **Published:** April 29, 2020.

1. Setup and estimators

This note is concerned with estimation of the conditional moment restriction model

$$\mathrm{E}[Y|X] = \exp(X'\beta), \tag{1}$$

almost surely, where Y is a scalar dependent variable, X is a k-dimensional vector of covariates, and β is a k-dimensional vector of parameters. This model can be considered as an example of the nonlinear regression model for a continuous Y or the Poisson regression model for a non-negative integer Y. This particular model has been extensively applied and studied in economic analysis using gravity models of international trade. See, e.g., Eaton and Kortum (2002), Anderson and van Wincoop (2003), Santos Silva and Tenreyro (2006), among others.

Based on a random sample $\{Y_i, X_i\}_{i=1}^n$, popular estimators for β are the nonlinear least squares (NLS) estimator $\hat{\beta}_{NLS} = \arg\min_\beta n^{-1}\sum_{i=1}^n \{Y_i - \exp(X_i'\beta)\}^2$ whose first-order condition is

$$\frac{1}{n}\sum_{i=1}^n \{Y_i - \exp(X_i'\hat{\beta}_{NLS})\}\exp(X_i'\hat{\beta}_{NLS})X_i = 0, \tag{2}$$

and the Poisson pseudo maximum likelihood (PPML) estimator whose first-order condition is

$$\frac{1}{n}\sum_{i=1}^n \{Y_i - \exp(X_i'\hat{\beta}_{PPML})\}X_i = 0. \tag{3}$$

In an influential paper, Santos Silva and Tenreyro (2006) argued the inconsistency problem of the OLS estimator for the log-linear model under heteroskedastic normal errors, and investigated the NLS and PPML estimators. In particular, Santos Silva and Tenreyro (2006) advocated the use of the PPML estimator under heteroskedastic errors rather than the NLS estimator. Their argument is that the NLS estimator tends to give more weights on the observations where $\exp(X_i'\hat{\beta}_{NLS})$ is large and generally noisier, and the NLS estimator tends to be less efficient than the PPML estimator. A simulation study by Santos Silva and Tenreyro (2006) endorsed the excellent performance of the PPML estimator.

In this note, we examine the finite sample performance of the conditional GMM estimator for the model in (1). By the law of iterated expectations, the conditional moment restriction (1) implies unconditional moment restrictions

$$\mathrm{E}[\{Y - \exp(X'\beta)\}h(X)] = 0, \tag{4}$$

for any function $h(\cdot)$ (as far as the above expectation is well-defined). Thus, both the NLS estimator (which specifies $h(X) = \exp(X'\beta)X$) and PPML estimator (which specifies $h(X) = X$) are consistent and also asymptotically normal under suitable regularity conditions.

In the context of estimation of the conditional moment restriction models, there are two substantial issues for the choice of $h(\cdot)$. First, the conditional moment restriction in (1) implies infinitely many unconditional moment restrictions in the form of (4). Thus, generally neither the NLS nor PPML estimator achieves the semiparametric efficiency bound to estimate β in the model (1). Currently several efficient estimation methods are available, such as the optimal instrumental variable estimator, and growing moment-based estimator (see, Chapter 7 of Hall (2005) for a survey). In our simulation study below, we consider the GMM estimator with growing moments (Donald, Imbens and Newey, 2003):

$$\hat{\beta}_{GMM} = \arg\min_{\beta} \left(\frac{1}{n} \sum_{i=1}^{n} g_{ni}(\beta) \right)' \left[\frac{1}{n} \sum_{i=1}^{n} g_{ni}(\hat{\beta}) g_{ni}(\hat{\beta})' \right]^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} g_{ni}(\beta) \right),$$

where $\hat{\beta}$ is a preliminary estimator, and $g_i(\beta) = \{Y_i - \exp(X_i'\beta)\}h_{ni}$ with a vector of basis functions $h_{ni} = (p_1(X_i), \ldots p_{k_n}(X_i))'$ for $k_n \to \infty$ as $n \to \infty$. A common drawback of efficient estimation methods for the conditional moment restrictions is that they typically involve some tuning parameters, such as the series lengths and bandwidths, to be chosen by the researcher.

The second issue is on consistency of point estimators. In an insightful paper, Dominguez and Lobato (2004) argued that even though the conditional moment restriction (1) uniquely identifies the parameters β, the implied unconditional moment restrictions (4) with finite dimensional $h(\cdot)$ may not fully exploit information contained in (1) and identification of β may not be guaranteed. In this case, the GMM estimator is typically inconsistent. To address this issue, Dominguez and Lobato (2004) observed that the conditional moment restriction (1) is equivalent to the continuum of the unconditional moment restrictions $\mathrm{E}[\{Y - \exp(X'\beta)\}I(X \leq x)] = 0$ for all x, and proposed the following estimator[1]

$$\hat{\beta}_{DL} = \arg\min_{\beta} \sum_{l=1}^{n} \left[\sum_{i=1}^{n} \{Y_i - \exp(X_i'\beta)\} I(X_i \leq X_l) \right]^2. \tag{5}$$

Dominguez and Lobato (2004) showed the consistency and asymptotic normality of this estimator under mild regularity conditions. Although $\hat{\beta}_{DL}$ does not achieve the semiparametric efficiency bound, it does not involve any tuning parameters.[2]

In the next section, we evaluate the finite sample properties of $\hat{\beta}_{GMM}$ and $\hat{\beta}_{DL}$ based on the simulation designs motivated by gravity models.

2. Simulation

We now assess the finite sample performances of the conditional GMM estimators and other estimators by Monte Carlo simulations. We first adopt simulation designs by Santos Silva and Tenreyro (2006). The dependent variable is generated by

$$Y_i = \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i})\eta_i, \tag{6}$$

for $i = 1, \ldots, 1000$, where X_{1i} follows the standard normal distribution, X_{2i} is a dummy variable that takes 1 with probability 0.4 and 0 otherwise, η_i is a log-normal random variable with mean 1 and variance σ_i^2, and $\beta = (\beta_0, \beta_1, \beta_2)' = (0, 1, 1)'$. The covariates X_{1i} and X_{2i} are independent. As in Santos Silva and Tenreyro (2006), we consider the following specifications of the conditional variance σ_i^2:

Case 1: $\sigma_i^2 = \exp(-2X_i'\beta)$; $\mathrm{Var}(Y_i|X_i) = 1$,

[1] For k-dimensional vectors a and b, let $I(a \leq b)$ be the element-by-element indicator, which takes 1 if $a_j \leq b_j$ for all $j = 1, \ldots, k$, and 0 otherwise.

[2] Although it is beyond the scope of this paper, it is interesting to extend our analysis for a bilateral setup to incorporate country-specific fixed effects. First of all, the asymptotic property of $\hat{\beta}_{DL}$ under the bilateral setup is an open question. Second, an efficient algorithm to implement $\hat{\beta}_{DL}$ for a large number of parameters needs to be developed.

Case 2: $\sigma_i^2 = \exp(-X_i'\beta)$; $\mathrm{Var}(Y_i|X_i) = \exp(X_i'\beta)$,

Case 3: $\sigma_i^2 = 1$; $\mathrm{Var}(Y_i|X_i) = \exp(2X_i'\beta)$,

Case 4: $\sigma_i^2 = \exp(-X_i'\beta) + \exp(X_{2i})$; $\mathrm{Var}(Y_i|X_i) = \exp(X_i'\beta) + \exp(X_{2i})\exp(2X_i'\beta)$.

However, these simulation designs may not imitate real trade data sufficiently. Typical trade data are rounded and include a large number of zeros. Therefore, we also conduct simulations with rounding errors in the dependent variable for each case. See Santos Silva and Tenreyro (2006) for detailed descriptions.

For this model, we consider six estimation methods: (i) DL, (ii) GMM, (iii) PPML, (iv) GPML, (v) NLS, and (vi) OLS.[3]

Table 1 presents estimation biases and MSEs for β_1 and β_2 based on 10,000 Monte Carlo replications. As shown in Santos Silva and Tenreyro (2006), PPML performs very well for all cases. In each case, PPML has a small bias and is relatively robust to rounding errors in the dependent variable. GMM is more robust to rounding errors than PPML. Similar to NLS, however, GMM is somewhat biased in the cases where heteroskedasticity is severe. Among the methods we consider, the performance of DL is the best. The biases of DL are small in various situations and outperforms PPML in terms of MSE in the cases where heteroskedasticity is severe (Cases 3 and 4).[4] This outperformance of DL is maintained even when the rounding errors are present, which implies that DL may outperform PPML in a real-world setting because the simulation with rounding errors has in common with a typical trade data in having a large number of zeros.

We next consider more realistic simulation designs adopted in Santos Silva and Tenreyro (2011). The dependent variable is generated by $Y_i = \sum_{j=1}^{m_i} Z_{ij}$ for $i = 1, \ldots, 1000$, where Z_{ij} follows a χ_1^2 distribution, and m_i is independent of Z_{ij}'s and follows a negative-binomial distribution with the conditional mean and variance specified below. In this setup, m_i and Z_{ij} can be interpreted as the number of exporters and quantity exported by firm j, respectively. The covariates $X_i = (X_{1i}, X_{2i})'$ and slope parameters $\beta = (\beta_0, \beta_1, \beta_2)'$ are same as in the first simulations in (6), and we set $\mathrm{E}[m_i|X_i] = \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i})$ and $\mathrm{Var}(m_i|X_i) = a\mathrm{E}[m_i|X_i] + b\mathrm{E}[m_i|X_i]^2$, where

Case 1: $(a, b) = (10, 0)$; $\Pr(Y_i = 0) = 0.62$,

Case 2: $(a, b) = (50, 0)$; $\Pr(Y_i = 0) = 0.83$,

Case 3: $(a, b) = (1, 5)$; $\Pr(Y_i = 0) = 0.65$,

Case 4: $(a, b) = (1, 15)$; $\Pr(Y_i = 0) = 0.81$.

See Santos Silva and Tenreyro (2011) for detailed descriptions. In this setup, the conditional expectation of Y_i is specified as

$$\mathrm{E}[Y_i|X_i] = \mathrm{E}[m_i|X_i] = \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i}).$$

[3]For GMM, we set the initial estimator $\hat{\beta}$ as the PPML estimator and $h_{ni} = (1, X_{1i}, X_{2i}, X_{1i}^2, X_{1i}X_{2i})'$. Our preliminary simulation suggests that the results are less sensitive to the choice of h_{ni}.

[4]As pointed out by Dominguez and Lobato (2004, p. 1605), DL is considered as an adaptation of the minimum distance estimator to the conditional moment restriction models. For nonlinear regression models, Koul (2002, Ch. 5) provided certain robustness properties for the minimum distance estimator against heteroskedastic errors. Although it is beyond the scope of this paper, it is interesting to see whether such robustness properties continue to hold for the current setup to explain the favorable finite sample performances of DL in these cases.

Table 2 presents biases and MSEs for estimating β_1 and β_2 based on 10,000 Monte Carlo replications.[5][6] Similar to the first simulations, the results show that DL performs well for all cases. In particular, when the conditional variance of Y_i is quadratic (Cases 3 and 4), the MSEs of DL are smaller than those of PPML.

Overall, our simulation results suggest that DL compares favorably with PPML and is better than other estimation methods.

References

[1] Anderson J. E. and E. van Wincoop (2003) "Gravity with gravitas: a solution to the border puzzle" *American Economic Review* **93**, 170-192.

[2] Dominguez, M. A. and I. N. Lobato (2004) "Consistent estimation of models defined by conditional moment restrictions" *Econometrica* **72**, 1601-1615.

[3] Donald, S. G., Imbens, G. W. and W. K. Newey (2003) "Empirical likelihood estimation and consistent tests with conditional moment restrictions" *Journal of Econometrics* **117**, 55-93.

[4] Eaton J. and S. S. Kortum (2002) "Technology, geography, and trade" *Econometrica* **70**, 1741-1779.

[5] Hall, A. R. (2005) *Generalized Method of Moments*, Oxford University Press.

[6] Koul, H. L. (2002) *Weighted Empirical Processes in Dynamic Nonlinear Models*, New York: Springer-Verlag.

[7] Santos Silva, J. M. C. and S. Tenreyro (2006) "The log of gravity" *Review of Economics and Statistics* **88**, 641-658.

[8] Santos Silva, J. M. C. and S. Tenreyro (2011) "Further simulation evidence on the performance of the Poisson pseudo-maximum likelihood estimator" *Economics Letters* **112**, 220-222.

[5]The results of NLS and OLS are not presented here because Santos Silva and Tenreyro (2006, 2011) reported these estimation methods are biased and inefficient.

[6]We also analyzed the performance of these estimation methods in cases with rounding errors. Since the results are overall similar to those without rounding errors, we omit them.

Table 1. Simulation Results for Designs in Santos Silva and Tenreyro's (2006)

	Without rounding errors				With rounding errors			
	β_1		β_2		β_1		β_2	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
	Case 1: $\text{Var}(Y_i \vert X_i) = 1$							
DL	-0.00056	0.00060	0.00088	0.00215	0.02271	0.00119	0.04265	0.00438
GMM	-0.00125	0.00035	-0.00260	0.00223	0.00205	0.00018	0.02303	0.00269
PPML	0.00000	0.00027	0.00032	0.00075	0.01905	0.00068	0.02075	0.00130
GPML	0.01318	0.00494	0.00787	0.00708	0.11029	0.02159	0.09417	0.02097
NLS	-0.00001	0.00006	0.00006	0.00030	0.00205	0.00007	0.00285	0.00033
OLS	0.39001	0.15363	0.35675	0.13021	—	—	—	—
	Case 2: $\text{Var}(Y_i \vert X_i) = \exp(X_i'\beta)$							
DL	-0.00018	0.00076	-0.00012	0.00231	0.02610	0.00155	0.04791	0.00506
GMM	-0.00063	0.00055	0.00088	0.00322	0.00147	0.00056	0.02925	0.00435
PPML	-0.00023	0.00038	-0.00005	0.00158	0.02187	0.00091	0.02327	0.00227
GPML	0.00435	0.00183	0.00142	0.00390	0.13350	0.02306	0.11279	0.02041
NLS	0.00028	0.00112	0.00109	0.00330	0.00246	0.00112	0.00405	0.00335
OLS	0.21064	0.04522	0.19972	0.04229	—	—	—	—
	Case 3: $\text{Var}(Y_i \vert X_i) = \exp(2X_i'\beta)$							
DL	-0.00067	0.00284	0.00006	0.00508	0.03052	0.00390	0.05772	0.00904
GMM	-0.00863	0.01250	0.01335	0.03557	-0.00763	0.01232	0.04831	0.03976
PPML	-0.00328	0.00527	-0.00079	0.01034	0.02383	0.00587	0.02745	0.01149
GPML	-0.00028	0.00099	0.00002	0.00415	0.19717	0.04249	0.16435	0.03452
NLS	0.14259	11.04195	0.18036	26.31483	0.14472	10.84810	0.18099	26.12356
OLS	-0.00037	0.00071	0.00011	0.00290	—	—	—	—
	Case 4: $\text{Var}(Y_i \vert X_i) = \exp(X_i'\beta) + \exp(X_{2i})\exp(2X_i'\beta)$							
DL	-0.00123	0.00803	-0.00219	0.01237	0.03444	0.00953	0.04744	0.01568
GMM	-0.02431	0.02052	0.01512	0.06500	-0.01632	0.01987	0.04956	0.07079
PPML	-0.00934	0.01035	-0.00817	0.02101	0.01800	0.01071	0.01694	0.02186
GPML	0.00361	0.00330	-0.00304	0.01196	0.12920	0.02391	0.10101	0.02701
NLS	0.37629	39.34492	0.75528	1190.303	0.38910	40.00869	0.73215	1197.705
OLS	0.13231	0.01898	-0.12586	0.02145	—	—	—	—

TABLE 2. Simulation Results for Designs in Santos Silva and Tenreyro's (2011)

	β_1		β_2	
	Bias	MSE	Bias	MSE
Case 1: $\text{Var}(m_i\|X_i) = 10\exp(X_i'\beta)$				
DL	-0.00018	0.00912	0.00528	0.02793
GMM	-0.00308	0.00592	0.00412	0.04001
PPML	0.00128	0.00449	0.00205	0.01901
GPML	0.05039	0.02701	0.02274	0.05165
Case 2: $\text{Var}(m_i\|X_i) = 50\exp(X_i'\beta)$				
DL	-0.00386	0.04053	0.02201	0.12262
GMM	-0.01223	0.02615	0.03151	0.31451
PPML	0.00302	0.01953	0.01294	0.08325
GPML	0.16546	0.13510	0.08546	0.23551
Case 3: $\text{Var}(m_i\|X_i) = \exp(X_i'\beta) + 5\exp(2X_i'\beta)$				
DL	-0.00022	0.01604	0.00476	0.03312
GMM	-0.03381	0.04918	0.07952	0.22585
PPML	-0.01323	0.02459	0.00259	0.05650
GPML	0.01467	0.01266	0.00747	0.03476
Case 4: $\text{Var}(m_i\|X_i) = \exp(X_i'\beta) + 15\exp(2X_i'\beta)$				
DL	-0.00581	0.04352	-0.00470	0.08677
GMM	-0.08219	0.09929	0.27926	1.99520
PPML	-0.03660	0.06107	-0.01439	0.15400
GPML	0.01249	0.02447	-0.00242	0.08095

ECONOMIC POLICY DEPARTMENT, MITSUBISHI UFJ RESEARCH AND CONSULTING CO., LTD., 5-11-2 TORANOMON, MINATO-KU, TOKYO 105-8501, JAPAN.
Email address: nishihata@murc.jp

DEPARTMENT OF ECONOMICS, LONDON SCHOOL OF ECONOMICS, HOUGHTON STREET, LONDON, WC2A 2AE, UK, AND KEIO ECONOMIC OBSERVATORY (KEO), 2-15-45 MITA, MINATO-KU, TOKYO 108-8345, JAPAN.
Email address: t.otsu@lse.ac.uk