Comments on "The log of gravity revised"

J.M.C. Santos Silva* Silvana Tenreyro†

December 29, 2008

1. INTRODUCTION

In a recent paper, Martínez-Zarzoso, Nowak-Lehmann and Vollmer (2007) revisit the results we presented in Santos Silva and Tenreyro (2006) and claim that their findings contradict our conclusions. We welcome the interest of these researchers in our work but, unfortunately, their conclusions are grounded on a mixture of misunderstandings and technical errors and consequently are very much untenable. Because their results can mislead some readers, in this note we try to clarify the main issues raised by the work of Martínez-Zarzoso, Nowak-Lehmann and Vollmer (2007).[1]

2. THE ADEQUACY OF THE PPML ESTIMATOR

In the abstract of their paper, Martínez-Zarzoso, Nowak-Lehmann and Vollmer (2007), hereinafter MZNLV, state that "Contrary to Santos Silva and Tenreyro (2006),

*University of Essex and CEMAPRE. Wivenhoe Park, Colchester CO1 1ED, United Kingdom. Fax: +44 (0)1206 872769. E-mail: jmcss@essex.ac.uk.

†London School of Economics, CEP, and CEPR. Department of Economics, s.600. St. Clement's Building. Houghton St., London WC2A 2AE, United Kingdom. Fax: +44 (0)20 78311840. E-mail: s.tenreyro@lse.ac.uk.

[1] It is worth pointing out that one of the authors of the present note transmited the essence of these comments to Martínez-Zarzoso, Nowak-Lehmann and Vollmer in December 2006, well before the working paper version of their work was made public.

the results of a simulation study indicate that the Pseudo Poisson Maximum Likelihood estimator (PPML) is not always the best estimator." This is, of course, a gross misinterpretation of our results as it is obvious that no estimator can always outperform its competitors. What we have claimed, and maintain, is that the PPML estimator is consistent and is likely to do well in a variety of circumstances. Indeed, in contradistinction to what is stated by MZNLV, their results actually confirm our conclusions.

Still in the abstract, MZNLV report that "We find that the PPML assumption concerning the pattern of heteroskedasticity is, in most cases, rejected by the data". However, the results presented in Table 6 of the MZNLV paper shows that this hypothesis is actually accepted at the 1% level in all data sets that were tested. It is important to keep in mind, however, that the hypothesis that the variance of the error term is proportional to the conditional mean is needed for the efficiency of the PPML estimator, but not for its consistency. Therefore, even if this hypothesis were to be rejected, that would not necessarily imply that the PPML would be inappropriate.

The simulation results reported in tables 1 and 2 of the MZNLV paper confirm that the PPML estimator has a good performance in all the cases that were considered, confirming our assertion that this estimator has all the characteristics needed to make it a promising workhorse for the estimation of gravity equations.

Not surprisingly, MZNLV find that the assumption that the data actually follows a Poisson distribution is strongly rejected (see the results of the "poisgof" test, labelled "Estat gof chi2" in tables 3 and 4 and the comments on page 15). Actually, no test was needed to reach this conclusion because the dependent variable is not even restricted to have integer values and therefore it is not clear why this test is performed at all. Recall that for the PPML estimator to be consistent the data does not have to follow a Poisson distribution, that is why the estimator is said to be a pseudo-maximum likelihood estimator rather than a maximum likelihood estimator.

3. PPML IS OUTPERFORMED IN OUT-OF-SAMPLE FORECASTS

MZNLV perform some out-of-sample forecast experiments to compare the predictive ability of models estimated in their multiplicative form (by PPML) and models estimated in the log-linearized form. However, the results obtained in these experiments are impossible to interpret because the authors use the log-linearized models to predict the log of trade, whereas the models estimated in the multiplicative form are used to predict trade. Because the variables that are predicted by the competing models are not the same, the results are obviously uninterpretable. Therefore, the claim that the models estimated by PPML are outperformed in out-of-sample forecasts is completely unfounded and very misleading. If anything, the out-of-sample forecast experiments performed by MZNLV clearly illustrate the problems with the use of the log-linearized models. Indeed, because of Jensen's inequality, in general these models cannot be used to obtain predictions for the value of trade itself.

The fact that the models used by MZNLV have different dependent variables also renders invalid all comparisons based on the value of the log-likelihood function, the AIC, the SJ goodness of fit statistic, and the RMSE (see tables 3, 4 and 5). In short, the empirical applications presented by MZNLV provide no basis to compare the competing approaches to the estimation of the gravity equations.

4. PPML IS OUTPERFORMED IN TERMS OF EXPECTED LOSS

When describing their simulation experiments, MZNLV (page 7) explain that "Deviating from Santos Silva and Tenreyro (2006), we will evaluate the performance of the estimators not only by looking at the bias of the estimates but also by computing their expected loss." Evaluating the performance of the estimators using a loss function that takes into account inconsistency and efficiency is of course potentially interesting. However, for large enough samples, the inconsistency will have a dominating effect as the variance of standard estimators shrinks to zero. That is, asymptotically, a consistent estimator will have smaller loss than any inconsistent estimator, no matter how efficient they are! Therefore, because the samples used in the estimation of gravity equations are typically very large, there are good grounds to focus on the inconsistency as a measure of performance.

It is also important to note that the results presented by MZNLV are based on samples of size 1000, which are much smaller than the samples typically used in the estimation of gravity equations. Therefore, their results provide no guidance about the expected loss of the different estimators for realistic empirical applications.

5. LOG-LINEARIZED MODELS SHOULD BE ESTIMATED BY GLS

MZNLV, see equation (1) in page 4, assume that the model to be estimated has the form

$$y_i = \exp(x_i\beta)\,\varepsilon_i,$$

where $\mathrm{E}(\varepsilon_i|x_i) = 1$. From here, assuming that $y_i > 0$, the following log linear model is obtained

$$\ln(y_i) = x_i\beta + \ln(\varepsilon_i)\,. \tag{1}$$

MZNLV are right to point out (as Santos-Silva and Tenreyro (2006) did) that, due to Jensen's inequality, $\mathrm{E}(\ln(\varepsilon_i)|x_i) \neq 0$. However, they state that this will only affect the estimation of the intercept, therefore implicitly assuming that $\mathrm{E}(\ln(\varepsilon_i)|x_i)$ is a constant.

The main message of Santos Silva and Tenreyro (2006), which apparently was missed by MZNLV, is that only under very strong assumptions on the distribution of ε_i will $\mathrm{E}(\ln(\varepsilon_i)|x_i)$ be a constant. Therefore, unless this extremely strong assumption is valid, $\mathrm{E}(\ln(\varepsilon_i)|x_i)$ will be a function of x_i and the estimation of the log-linearized model will be inconsistent for all the elements in β. (Concretely, if ε_i is heteroskedastic, $\mathrm{E}(\ln(\varepsilon_i)|x_i)$ will be a function of x_i.) That is why Santos Silva and Tenreyro (2006) recommend that the gravity model should be estimated in its multiplicative form and warn against the danger of using the log-linearization.

MZNLV state that if $\ln(\varepsilon_i)$ is heteroskedastic, estimation of (1) by OLS will no longer be efficient. However, the problem is much more serious that this: if $\ln(\varepsilon_i)$ is heteroskedastic, estimation of (1) by OLS will generally be **inconsistent** for β. MZNLV (page 5) go on to claim that the efficiency issue can be addressed by using GLS or FGLS. However, the authors fail to realize that they are addressing the wrong problem because the issue is not the efficiency of the estimator but its consistency. Like OLS, the FGLS estimator of the log-linearized model proposed by MZNLV is inconsistent for β.

It is interesting to notice that MZNLV test the assumption that the log-linear model is valid. The results of the test are presented in Table 6 under the label "Park TEST LOG LOG" and the test is described in pages 14 and 15. In page 15 MZNLV clearly state that the null of this test is rejected in all cases (p-values smaller than 0.01), but they fail to acknowledge that this implies that any approach based on the log-linear model is likely to be invalid.

Finally, it is puzzling to notice that although MZNLV suggest the estimation of β by estimating the log-linearized model using FGLS, they never actually explain how to construct the matrix $\hat{\Omega}$ that is used in their equation (4). The situation is even more puzzling because the reference they provide (Greene, 2000, p. 618) refers to the estimation of Ω in systems of regression equations, which is not at all helpful in this context. Finally, it is also unclear how the FGLS estimator is used in the simulations where y_i can be zero. Although we presume that the the zero observations are simply dropped, MZNLV are mute about this issue.

6. GENERATING ZEROS IN THE SIMULATIONS

In the simulations performed by Santos Silva and Tenreyro (2006) the data were generated using a log-normal model, which never generates zero values for the dependent variable. To check the robustness of the PPML estimator, simulations were also performed rounding the dependent variable to the nearest integer, thereby creating a sizable proportion of zeros in the dependent variable. As explained by Santos Silva and Tenreyro (2006), this rounding of the dependent variable makes all estimators inconsistent and therefore the results should be viewed as a way to assess the sensitivity of the competing estimators to a kind of measurement error that is likely to occur in official statistics.

MZNLV (page 8) criticize these experiments stating that, because the new dependent variable only assumes integer values, the experiments favour the PPML estimator. This, of course, is not true because, as stated above, nothing in the PPML estimator depends on y_i assuming only integer values.

To avoid this alleged favouring of the PPML estimator, MZNLV use two different mechanisms to generate the zeros. In one case, 15% of the observations of y_i are randomly selected and set to zero. If the original data is generated with conditional mean $\exp(x_i\beta)$, the modified data will have conditional mean $\exp(\ln(0.85) + x_i\beta)$ and therefore, apart from adding noise to the experiments, only the intercept of the model will be affected by this way of generating the zeros. Not surprisingly, the results in this case are very close to those obtained with the data without zeros.

In a second set of experiments, the zeros introduced by MZNLV are linked to the value of the covariates. Although the authors do not give details on how exactly the zeros are generated, it is clear that the model will now be misspecified and none of the estimators considered is expected to be consistent. This situation, therefore, parallels the one in Santos Silva and Tenreyro (2006). We find it encouraging that even in this misspecified model the bias of the PPML estimator is never above 6.7%, whereas all other estimators have biases that can be much larger that this.

In any case, only 15% of the observations are set to zero in the MZNLV experiments. Given that most data sets used for the estimation of gravity equations have more that 50% of zeros, the results presented by MZNLV can be very misleading because they downplay the importance of the problems created by the presence of the observations for which it is not possible to compute the log of the dependent variable.

7. CONCLUDING REMARKS

We believe that it is natural that some researchers find it difficult to accept that the long tradition of estimating gravity equations in their log-linear form is based on untenable assumptions. Therefore, we are not surprised to see that some authors try to argue in favour of its use and try to find arguments against the PPML estimator proposed by Santos Silva and Tenreyro (2006).

Of course, we are fully aware that the PPML estimator can certainly be outperformed in some situations, and we very much welcome the scrutiny of our results. However, we are also fully convinced that the PPML estimator is generally appropriate in this context and we believe that it should be the benchmark against which potential alternatives have to be judged. This comparison, however, has to be made using fair econometric methods with solid theoretical justification.

REFERENCES

Greene, W.H. (2000). *Econometric Analysis* 4th Ed., Upper Saddle River (NJ): Prentice Hall International, Inc.

Martínez-Zarzoso, I., Nowak-Lehmann D., F. and Vollmer, S. (2007). "The Log of Gravity Revisited," CeGE Discussion Papers 64, University of Goettingen, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=999908.

Santos Silva, J.M.C. and Tenreyro, Silvana (2006). "The Log of Gravity," *The Review of Economics and Statistics*, 88, 641-658.