Collinearity in growth regressions: the example of worker remittances

Working Paper Series

#2009-020

Collinearity in growth regressions: The example of worker remittances

Thomas H.W. Ziesemer

United Nations University - Maastricht Economic and social Research and training centre on Innovation and Technology
Keizer Karelplein 19, 6211 TC Maastricht, The Netherlands
Tel: (31) (43) 388 4400, Fax: (31) (43) 388 4499, e-mail: info@merit.unu.edu, URL: http://www.merit.unu.edu

Collinearity in growth regressions: The example of worker remittances

Thomas H.W. Ziesemer, Department of Economics and UNU-MERIT, University of Maastricht, P.O.Box 616, 6200 MD Maastricht, The Netherlands. Fax: ++31-43-3884150. Phone: ++31-43-3383872. T.Ziesemer@algec.unimaas.nl.*

Abstract. The sign of worker remittances in growth regressions is heavily disputed in the literature. Comparing two growth regressions with different signs for the remittance variable we show that collinearity with the lagged dependent variable might indicate that collinearity should be investigated comprehensively and might lead to a change in specifications which differ in the variance inflation factors (VIF). In our case the variance inflation factor for remittances depends on the use of a five or one-year lag of the lagged dependent. In the regression with a VIF below ten, the standard critical value, the sign of remittances is positive.
Keywords: Growth, remittances. JEL-code: F24, O11,15,40.

UNU-MERIT Working Papers
ISSN 1871-9872

Maastricht Economic and social Research and training centre on Innovation and Technology, UNU-MERIT

UNU-MERIT Working Papers intend to disseminate preliminary results of research carried out at the Centre to stimulate discussion on the issues raised.

Introduction

Every researcher who has suggested one of the about 140 regressors in growth regressions (Durlauf et al. 2005) had to respond to the question of reversed causality. It is much less clear though that anyone had to respond to the question whether or not the sign of a suggested regressor may be turned around by approximate collinearity (see Davidson McKinnon 2004 on the basics) with the lagged dependent variable although there were some debates on the right signs in growth regression. In general, the issue is important because each chapter in any textbook on development economics suggests the relevance for income effects. If income effects are important collinearity with the lagged dependent variable can be a major issue in any growth regression. Examples are controversies in regard to applications for development aid (see Doucouliagos, H. and M**.** Paldam (2008) and on the impact of worker remittances on growth found by Chami et al. (2005). It is this latter case in which we are interested in relation to the collinearity issue.

Chami et al. (2005) have argued that remittances provide an incentive to reduce effort thereby making weak economic performances more likely. They find negative impacts of remittances on growth in a cross-section regression. In Lucas (2005) and IMF (2005) this result is attributed intuitively to weak or inadequate instruments and in the latter no growth effect is found. Catrinescu et al. (2006) extend the approach of Chami et al. to include policy and institutional variables and estimate a panel using the Anderson-Hsiao estimator. They find some significantly positive results for the impact of remittances on growth, but these are reported to be not very robust. Giuliano and Ruiz-Arranz (2005) add remittances multiplied to financial variables as a regressor and find positive growth effects for financially less developed countries. In summary, these papers see the reason for the negative sign found by Chami et al. (2005) in inadequate instruments, omitted variables and inadequate estimation methods. We advance another possibility that might be useful for future research: approximate collinearity with the lagged dependent variable.

Methodology

Growth regressions can be written as follows (Durlauf et al. 2005).

$$Log(y_t) = \alpha_i + (\beta+1)log(y_{t-1}) + \gamma x_1 + \eta x_2 + u_{it}$$

'log' indicates a natural logarithm, y is GDP per capita or per worker, x_1 denotes regressors used in mathematically formulated growth models, in particular the augmented Solow model (see Mankiw et al. 1992) and x_2 denotes other regressors, which are added although they are not included in a growth model. Examples for the latter are official development aid or worker remittances. Such effects are normally interpreted to mirror the impact of a variable on the total factor productivity (see Rogriguez (2006)), which can be considered to be a weighted average of sectoral productivities. These variables then either affect the weights of the sectors through the shift of demand and factor inputs or they have an impact on the sectoral technical progress (see Timmer and Szirmai (2000)). The expected signs for the coefficient of the lagged dependent variable normally obtained in growth regressions are $\beta < 0 < \beta+1$.

We will estimate such a growth regression for more than 40 countries with per capita income above \$1200 in prices of the year 2000.[1] In the first instance we obtain the result that the impact

[1] In related work on countries with per capit income below \$1200 we found no ambiguity in the coefficients for aid and remittances.

of worker remittances on growth is negative under some additional assumptions. One of these assumptions is the use of a lagged dependent variable with a five years lag that is significantly correlated with the remittance variable. However, a one-year lagged dependent variable is much less significantly correlated with remittances, depending on the set of controls used though. Using a one-year lagged dependent variable and reworking the regression towards having only significant variables the sign for remittances changes into a positive one. Finally, the variance inflation factors (see Kennedy (2003) for an extensive treatment) all regressors in both equations are calculated indicating that the remittance variable in the second equation is much less correlated with other regressors than in the first equation. Due to other multi-collinearities it remains an open question, which of the changes is actually turning the sign around, but it is the one with the lagged dependent variable which is economically plausible, therefore checked first and easily tested before the ultimate plausibility comes from the comparison of variance inflation factors indicating the strength of the multicollinearity.

All data are taken from the World Development Indicators. We use the fixed effects method, which is known to have a downward bias for the lagged dependent variable of an order of magnitude of *1/T*, if we have more than thirty observations as we do in the first regression. If we have less than thirty observations we use the systems GMM method of Arellano-Bover (1995) because fixed effects estimation is then underestimating the coefficient of the lagged dependent variable. With this method we can use instruments to correct for the endogeneity of the lagged dependent variable and other regressors. In our case the coefficient of the lagged dependent variable is slightly larger than that of the fixed effects regression and the test for the validity of the instruments and not having too many of them (see Roodman 2007) is also passed.

Results

The growth regression for the log of the GDP per capita, *log(gdppc)*, we would have defended in the first instance is as follows (p-values in parantheses)[2].

$$
\begin{aligned}
\mathrm{Log}(gdppc) - \log(gdppc(-5)) = &\underset{(0.0003)}{-4.66}\underset{(0)}{-0.14}\log(gdppc(-5)) + \underset{(0)}{0.11}\log(gfcfgdp) \\
&\underset{(0.05)}{-.0245}\log(gfcfgdp(-5)) + \underset{(0.092)}{0.0014}Lit(-1) + \underset{(0.0125)}{1.86}(1/t) + \underset{(0.0003)}{6.54}(wr/gdp)^2 \underset{(0.0004)}{-1.57}wr/gdp \\
&+ \underset{(0.015)}{1.08}wr(-1)/gdp(-1) - \underset{(0.004)}{3.76}(wr(-1)/gdp(-1))^2 + \underset{(0.03)}{0.52}oda/gdp \underset{(0.027)}{-2.78}(oda/gdp)^2 + \\
&+ \underset{(0)}{0.2}\log(wld) \underset{(0.046)}{-0.057}\log(l)
\end{aligned}
\qquad (1)
$$

Periods: 34 (1971 2005). Countries: 45. Obs.: 634. Adj.R^2 =0.996; DW=1.77

[2] A value of (0) indicates zeros for four digits. Three lagged growth rates are employed as serial correlation correction.

The lagged dependent variable has a sign and size of the coefficient in accordance with the expectation given above. The sum of the coefficients of the investment variables, *gfcfgdp*, is positive. Literacy, *Lit*, also has a positive sign and the growth of the GDP of the world, *wld*, as an income argument in the export demand function stemming from the idea of growth modeled with imported inputs in Bardhan and Lewis (1970) has a positive sign. The natural logarithm of the labour force, *log(l)*, has a negative sign of approximately the same order of magnitude as the world income variable.[3] The squared values for remittances, *wr/GDP*, and aid, *oda/GDP*, are very small. Therefore the linear ones dominate. Under the assumption that variables and their lags are of similar size remittances have a negative impact and aid has a positive one. However, a look at Table 1 shows that the regressors with the exception of the lagged investment variables are pairwise significantly correlated with the lagged dependent variable. The result may therefore stem from collinearity, which may have an impact on the sign of regressors. Table 2 shows results from regressing the remittance variables on the GDP per capita and its one and five year lags. The correlation is most strong for the five year lag used in the above regression. One may therefore want to avoid five year lags. Changing them into one-year lags and eliminating the most highly insignificant regressors lead us to the following result.

$$\underset{}{\text{Log(gdppc)}} = c - \underset{(0)}{0.09\log(\text{gdppc}(-1))} + \underset{(0)}{0.123\log(\text{gfcfgdp})} - \underset{(0)}{0.09\log(\text{gfcfgdp}(-1))} \qquad (2)$$

$$- \underset{(0.015)}{0.19d(\log(L))} + \underset{(t=1.67))}{0.00146\text{sum(Lit)}} + \underset{(0.073)}{1.06(\text{wr}(-1)/\text{gdp}(-1))^2} + \underset{(0.033)}{0.3\text{oda}(-1)/\text{gdp}(-1)}$$

$$- \underset{(0.052)}{0.52(\text{oda}(-1)/\text{gdp}(-1))^2} + \underset{(0.0001)}{0.114\log(\text{wld})} - \underset{(0.0003)}{0.099\log(l)}$$

Per.: 23 (1981 2005). Countr.: 42. Obs.: 558. s.e.e.: 0.037. J=267. Instr.rank:257. p(J) = 0.168.

For the literacy variable we now use a polynomial distributed lag of the first degree with 10 lags, which has negative growth effects for the first five lags and but significantly positive effects thereafter. These lags cost us some observations and therefore the adequate method is that of Arellano-Bover (1995). Moreover, the aid variables are used now with a one year lag. The major difference though is that the remittance variable now has a positive effect, which it did not when using the five-year lag for the lagged dependent variable. Moreover, only the squared lag of remittances is significant.

The econometric literature on multicollinearity emphasizes the variance inflation factor, $1/(1-R_i)$, where R_i is the coefficient of determination for the regression of regressor i on all the other regressors. In Table 3 we provide the values for R_i and the variance inflation factors for both

[3] Using the formulas in Mutz and Ziesemer 2008 and assuming an elasticity of production for capital of 0.33, we get a price elasticity of export demand of (-6.9) in regression (1) and (-4) in regression (2). Both values seem quite reasonable. Again according to these growth rate formulas we can obtain the income elasticity of export demand as the ratio of the coefficients for the world income and the labour variable. This coefficient is 3.5 and therefore far too high for the first regression and slightly above unity for the second regression.

regression.[4] The worker remittance variables have a high collinearity in the first regression but a much lower one in the second.

Conclusion

In both regressions the aid variable has a significantly positive sign.[5] Switching from the five-year lag to the one-year lag because of the collinearity with the lagged dependent variable in the first regression, ultimately changes the sign of the remittance variable after other adjustments are made. The variance inflation factors indicate that the sign and significance of the remittance variable are based on correlation with the other regressors in the first equation but much less so in the second, where the variance inflation factor is below the standard critical value of 10 (see Kennedy 2003). Therefore we cautiously suggest that the positive sign for remittances is more convincing for our sample. As a tentative interpretation, remittances and aid are unlikely to contribute to total factor productivity growth (tfp) via technical change; but rather remittances and aid seemingly are spent in sectors with above average tfp and thereby shift more weight to them and generate higher aggregate tfp levels.

[4] For this purpose we use the fixed effects version of equation (2) because we do not have lagged dependent variables in most cases and we need an R-squared value and therefore a constant; both are not calculated in the Arellano-Bover method.

[5] The squared term generates a function with peaks at 9.35% and 28.8% of GDP only.

References

Catrinescu, N., M. Leon-Ledesma, M Piracha and B. Quillin (2006). Remittances, Institutions and Economic Growth. *IZA DP* No. 2139, May.

Chami, R., C.Fullenkamp and S. Jahjah (2003). Are Immigrant Remittance Flows a Source of Capital for Development? *IMF Staff Papers*, 2005, v. 52, iss. 1, pp. 55-81.

Davidson, R. and J.G. MacKinnon (2004) *Econometric Theory and Methods*. Oxford University Press. New York et al.

Doucouliagos, H. and M**.** Paldam (2008), Aid Effectiveness on Growth: A Meta Study; *European Journal of Political Economy*, March, v. 24, iss. 1, pp. 1-24.

Durlauf, S.N., P.A. Johnson, J.R.W. Temple (2005). Growth Econometrics. In: *Handbook of Economic Growth*. Eds.: P. Aghion and S. Durlauf. Amsterdam, North-Holland. 555-677.

Giuliano, P. and M. Ruiz-Arranz (2005). Remittances, Financial Development, and Growth. *IMF Working Paper* 05/234. December.

IMF (2005). *World Economic Outlook*. Washington , Chap.2, 69-107.

Judson, R.A. and A.L. Owen (1999). Estimating dynamic panel data models: a guide for macroeconomists. *Economics Letters* 65, 9–15.

Kennedy, P. (2003), A Guide to Econometrics, 5th edition, MIT Press and WileyBlackwell.

Lucas, R.E.B. (2005). International Migration to the High-income Countries: Some consequences for Economic Development in the Sending Countries. In Blanpain, R. (ed.): *Confronting Globalization*. Kluwer, The Hague, 157-190.

Mankiw, N.G., D. Romer and D. N.Weil (1992). A Contribution to the Empirics of Growth. *Quarterly Journal of Economics*, 107, 407-437.

Mutz, C. and T.Ziesemer (2008). Simultaneous Estimation of Income and Price Elasticities of Export Demand, Scale Economies and Total Factor Productivity Growth for Brazil, *Applied Economics*, Volume 40, Issue 22, November 2008, pages 2921 - 2937.

Rodriguez, F. (2006). Cleaning Up the Kitchen Sink: On the Consequences of the Linearity Assumption for Cross-Country Growth Empirics. *Wesleyan Economic* WP 2006-004.

Roodman, D. (2007). A Short Note on the Theme of Too Many Instruments. *CGDEV* WP 125, August.

Timmer, M.P. and A.Szirmai (2000). Producitivity growth in Asian manufacturing: the structural bonus hypothesis examined. *Structural Change and Economic Dynamics* 11, 371-392.

Appendix 1: Tables

Table 1: Uncontrolled correlation matrix and marginal significance levels

Covariance Analysis: Ordinary
Sample (adjusted): 1971 2005
Included observations: 650 after adjustments
Balanced sample (listwise missing value deletion)

Correlation Probability	1	2	3	4	5	6	7	8	9	10	11	12	13
1.LOG(GDPPC)	1												

2.LOG(GDPPC(-1))	1.00	1.00											
	0.00	-----											
3.LOG(GDPPC(-5))	0.98	0.98	1.00										
	0.00	0.00	-----										
4.LOG(GFCFGDP)	0.00	-0.02	-0.10	1.00									
	0.99	0.66	0.01	-----									
5.LOG(GFCFGDP(-5))	0.04	0.04	0.03	0.44	1.00								
	0.27	0.32	0.40	0.00	-----								
6.LIT(-1)	0.56	0.56	0.54	-0.12	-0.12	1.00							
	0.00	0.00	0.00	0.00	0.00	-----							
7.D(LOG(L))	-0.09	-0.08	-0.07	-0.05	0.01	-0.15	1.00						
	0.02	0.04	0.09	0.22	0.75	0.00	-----						
8.(WR/GDP)^2	-0.19	-0.19	-0.19	0.18	0.21	-0.10	0.09	1.00					
	0.00	0.00	0.00	0.00	0.00	0.01	0.02	-----					
9.WR/GDP	-0.27	-0.27	-0.26	0.16	0.17	-0.23	0.08	0.94	1.00				
	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	-----				
10.WR(-1)/GDP(-1)	-0.26	-0.26	-0.26	0.15	0.19	-0.23	0.08	0.92	0.98	1.00			
	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	-----			
11.(WR(-1)/GDP(-1))^2	-0.19	-0.19	-0.19	0.15	0.22	-0.09	0.08	0.95	0.91	0.94	1.00		
	0.00	0.00	0.00	0.00	0.00	0.02	0.04	0.00	0.00	0.00	-----		
12.ODA/GDP	-0.41	-0.41	-0.39	0.06	0.09	-0.30	0.09	0.62	0.62	0.62	0.62	1.00	
	0.00	0.00	0.00	0.11	0.02	0.00	0.02	0.00	0.00	0.00	0.00	-----	
13. (ODA/GDP)^2	-0.28	-0.28	-0.26	0.12	0.13	-0.17	0.08	0.56	0.52	0.52	0.55	0.91	1.00
	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	-----

Table 2: Collinearity of remittance and GDP per capita

Regressors	*Coefficient*	*Std. Error*	*t-Statistic*	*Prob.*
Dependent Variable: WR/GDP				
C	0.034	0.085	0.393	0.694
LOG(GDPPC)	0.038	0.027	1.437	0.151
LOG(GDPPC(-1))	-0.018	0.024	-0.743	0.458
LOG(GDPPC(-5))	-0.020	0.012	-1.690	0.091
Dependent Variable: $(WR/GDP)^2$				
C	0.000	0.019	0.003	0.998
LOG(GDPPC)	0.005	0.010	0.467	0.641
LOG(GDPPC(-1))	0.001	0.009	0.123	0.902
LOG(GDPPC(-5))	-0.005	0.004	-1.388	0.166
Dependent Variable: WR(-1)/GDP(-1)				
C	0.037	0.088	0.415	0.679
LOG(GDPPC)	0.038	0.028	1.363	0.173
LOG(GDPPC(-1))	-0.012	0.025	-0.468	0.640
LOG(GDPPC(-5))	-0.026	0.012	-2.166	0.031
Dependent Variable: $(WR(-1)/GDP(-1))^2$				
C	0.004	0.021	0.181	0.857
LOG(GDPPC)	-0.002	0.010	-0.181	0.857
LOG(GDPPC(-1))	0.010	0.009	1.069	0.285
LOG(GDPPC(-5))	-0.008	0.004	-1.863	0.063

Table 3: Variance Inflation Factors

	R-sq. Regr.1	R-sq. Regr.2	*VIF Regr. 1*	*VIF Regr.2*
LOG(GDPPC(-1))	-	0.984	-	63.0
LOG(GDPPC(-5))	0.978	-	46.4	-
LOG(GFCFGDP)	0.657	0.784	2.9	4.6
LOG(GFCFGDP(-1))	-	0.805	-	5.1
LOG(GFCFGDP(-5))	0.656	-	2.9	-
LIT(-1)	0.981	-	52.7	-
1/(@trend)	0.922	-	12.9	-
LOG(GDPPC(-1))-LOG(GDPPC(-6))	0.898	-	9.8	-
LOG(GDPPC(-2))-LOG(GDPPC(-7))	0.938	-	16.0	-
LOG(GDPPC(-3))-LOG(GDPPC(-8))	0.881	-	8.4	-
(WR/GDP)^2	0.985	-	65.2	-
WR/GDP	0.992	-	128.8	-
WR(-1)/GDP(-1)	0.992	-	130.0	-
(WR(-1)/GDP(-1))^2	**0.985**	**0.849**	**66.4**	**6.6**
ODA/GDP	0.960	0.962	24.9	26.5
(ODA/GDP)^2	0.942	-	17.3	-
ODA(-1)/GDP(-1)	-	0.962	-	26.5
(ODA(-1)/GDP(-1))^2	-	0.914	-	11.7
LOG(WLD)	0.964	0.906	28.1	10.6
LOG(L)	0.998	0.999	604.6	869.6
d(log(L))	-	0.389	-	1.6

The UNU-MERIT WORKING Paper Series

2009-01 *Effectiveness of R&D Tax Incentives in Small and Large Enterprises in Québec* by Rufin Baghana and Pierre Mohnen
2009-02 *Bridges in social capital: A review of the definitions and the social capital of social capital researchers* by Semih Akçomak
2009-03 *The Role of Firms in Energy Transformation* by Radhika Perrot
2009-04 *Standards as a platform for innovation and learning in the global economy: a case study of Chilean salmon farming industry*
2009-05 *Consumer behaviour: evolution of preferences and the search for novelty* by M. Abraham Garcia-Torres
2009-06 *The role of consumption and the financing of health investment under epidemic shocks* by Théophile T. Azomahou, Bity Diene and Luc Soete
2009-07 *Remittances, lagged dependent variables and migration stocks as determinants of migration from developing countries* by Thomas H.W. Ziesemer
2009-08 *Thinking locally: Exploring the importance of a subsidiary-centered model of FDI-related spillovers in Brazil* by Anabel Marin and Ionara Costa
2009-09 *Are International Market Demands Compatible with Serving Domestic Social Needs? Challenges in Strengthening Innovation Capacity in Kenya's Horticulture Industry* by Mirjam Steglich, Ekin Keskin, Andy Hall and Jeroen Dijkman
2009-10 *Industrialisation as an engine of growth in developing countries* by Adam Szirmai
2009-11 *The motivations, organisation and outcomes of university-industry interaction in the Netherlands* by Isabel Maria Bodas Freitas and Bart Verspagen
2009-12 *Habit Formation, Demand and Growth through product innovation* by M. Abraham Garcia-Torres
2009-13 *The Diffusion of Informal Knowledge and Innovation Performance: A sectoral approach* by M. Abraham Garcia-Torres and Hugo Hollanders
2009-14 *What does it take for an R&D tax incentive policy to be effective?* by Pierre Mohnen and Boris Lokshin
2009-15 *Knowledge Base Determinants of Technology Sourcing in the Clean Development Mechanism Projects* by Asel Doranova, Ionara Costa and Geert Duysters
2009-16 *Stochastic environmental effects, demographic variation, and economic growth* by Théophile T. Azomahou and Tapas Mishra
2009-17 *Measuring eco-innovation* by Anthony Arundel and René Kemp
2009-18 *Learning How to Consume and Returns to Product Promotion* by Zakaria Babutsidze
2009-19 *Strengthening Agricultural Innovation Capacity: Are Innovation Brokers the Answer?* by Laurens Klerkx, Andy Hall and Cees Leeuwis
2009-20 *Collinearity in growth regressions: The example of worker remittances* by Thomas H.W. Ziesemer

[*] I am grateful to Bertrand Candelon, Femke Cramer, Huub Meijers, Pierre Mohnen, Eddy Szirmai, Bart Verspagen for useful comments and participants of a UNU-MERIT seminar and the research day of the Netherlands network in Economic (NAKE) for interesting discussions.