Irrelevant but highly persistent instruments in stationary regressions with endogenous variables containing near-to-unit roots.

Nicolaos Kourogenis* Ekaterini Panopoulou† Nikitas Pittis‡

November 2005

Abstract

This paper suggests that IV estimators, utilizing irrelevant but persistent instruments may produce reliable inferences, in small samples, in cases where the endogenous variables contain autoregressive roots near unity. In such cases, these estimators appear to outperform IV estimators with strong instruments as well as some asymptotically efficient cointegration estimators.

JEL classification: *C12, C13, C22*
Keywords: *Instrumental variables estimator, persistent instruments, near-to-unit roots.*

Acknowledgements: We acknowledge financial support from the Greek Ministry of Education and the European Union under "Hrakleitos" grant. We are grateful to Guido Kuresteiner and Aris Spanos for helpful comments. The usual disclaimer applies.

*Department of Banking and Financial Management, University of Piraeus.

†Department of Banking and Financial Management, University of Piraeus, Greece and Department of Economics, National University of Ireland Maynooth.

‡Department of Banking and Financial Management, University of Piraeus. Correspondence to: Nikitas Pittis, Department of Banking and Financial Management, University of Piraeus, 80 M.Karaoli and A. Dimitriou str. 18534 Piraeus, Greece. E-mail: npittis@unipi.gr

1 Introduction

Selecting appropriate instruments in the context of an Instrumental Variables (IV) procedure is of paramount importance for producing reliable inferences on the structural parameters of interest. It is now well understood that if the instruments are only weakly correlated to the endogeneous variables, then IV estimators are likely to fare no better than the Ordinary Least Squares (OLS) estimator (see Nelson and Startz 1990a, 1990b, Buse 1992, Bekker 1994, Bound, *et. al.* 1995, Dufour 1997, Staiger and Stock 1997 and Wang and Zivot 1998).

The literature on the 'weak instruments' issue implicitly refers to cases where the regressor is either serially uncorrelated or exhibits a very low degree of persistence. This is due to the fact that a persistent regressor is always accompanied by strong instruments, namely its own lagged values. If, for example, the regressor, x_t, follows an AR(1) process, with coefficient ρ_x, then the lagged value of the regressor, x_{t-1}, is readily available as an instrument for x_t. In such a case, the 'weak instrument' problem is not an issue, unless ρ_x is close to zero. Moreover, the higher is the value of ρ_x, the stronger is x_{t-1} as an instrument for x_t. However, this is true only for values of ρ_x less than one. If $\rho_x = 1$, the regressor is an I(1) process, participating in a cointegrating regression. In such a case, the OLS estimator is super-consistent, which in turn implies that 'first-order' asymptotic bias effects dissapear. In such a case, an IV procedure, such as the two-stages least squares (TSLS) estimator, is inappropriate since it is designed to deal with a problem that no longer exists. The asymptotic problems in the cointegration case are of different nature, usually referred to as 'second-order' effects (see, for example, Phillips 1988, Park and Phillips 1988, Phillips and Loretan 1991). To deal with these problems, one has to employ an asymptotically efficient cointegration estimator, rather than a standard IV one. If one insists on using IV procedures in the case of cointegration, then she ends up with an estimator whose asymptotic distribution suffers from nuisance parameter dependencies (second-order effects) arising not only from the correlation between the regression error and the regressor, but also from the correlation between the instrument and the regressor! In other words, the problem of 'weak instruments' is reversed. In the case of cointegration, a weak, or even more so, an irrelevant instrument is beneficial, since it simplifies the nuisance parameter dependencies in the asymptotic distribution of the IV estimator without affecting the consistency of this estimator. This may be thought of as a beneficial artifact of the spurious regression theory (see Phillips and Hansen 1990).

The preceding discussion implies that the issue of 'weak instruments' should be examined in conjuction with the time series properties of the data in hand. It is true that a weak instrument is likely to be a problem in a low-persistence environment, but it is also true that a strong instrument may create more problems than it solves in a 'high-persistence' or 'near-to-unit-root' framework. As ρ_x moves from the stationary to the unit-root region, first-order effects are declining but second-order effects are emerging. Although the asymptotic theory has provided clear answers on the properties of IV estimators for the two polar cases $|\rho_x| < 1$ and $\rho_x = 1$, it is of little help to suggest the optimal estimation procedure, in finite samples, for the cases that ρ_x is less than but close to unity. To put it differently, it is not clear whether first or second order effects are predominant in the case that ρ_x is in the viscinity of unity. This paper examines these issues in some detail. Specifically, we address the following questions: What is the optimal way to estimate the structural parameter of interest, for samples of typical sizes, when the regressor is a stationary but highly persistent process, correlated with the regression error? Is it still optimal to employ an IV procedure that utilizes the strongest available instrument(s), as the relevant asymptotic theory suggests? Or is it better to treat the regression as a nearly-cointegrated one and employ an asymptotically efficient cointegration estimator?[1] This paper offers simulation evidence against these options. Both methods are outperformed by a TSLS estimator that utilizes irrelevant but highly persistent instruments.

The paper is organized as follows. Section 2 introduces the DGP and briefly reviews the relevant theory. Section 3 reports the simulation findings and Section 4 concludes the

[1] Elliot (1998) examines the problems with employing standard cointegration estimators in cases where the series involved in the regression contain near-to-unit roots. He demonstrates that commonly applied hypothesis tests on the parameters of interest suffer from severe size distortions, when slowly mean reverting processes are approximated by ones with unit roots. He also shows that using lags of the regressors as instruments is inappropriate in this case.

paper.

2 The Model, and Some Background Theory

Consider the regression equation:

$$y_t = \theta x_t + u_{1t} \tag{1}$$

where the regressor is generated via an AR(1) process:

$$x_t = \rho_x x_{t-1} + u_{2t} \tag{2}$$

We also assume the presence of a third variable, z_t, that might serve as an instrument for identifying θ, which also follows an AR(1) process,

$$z_t = \rho_z z_{t-1} + u_{3t} \tag{3}$$

The error vector $\mathbf{u}_t = [u_{1t}, u_{2t}, u_{3t}]^\intercal$ is assumed to be normal, independent and identically distributed with zero mean and covariance matrix Σ. Specifically,

$$\begin{bmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{bmatrix} \tilde{} NIID \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \tag{4}$$

Let us first review some useful results from the existing literature for the stationary and cointegrating regression cases, defined by $|\rho_x| < 1$ and $\rho_x = 1$, respectively, starting from the former.

Stationary Regression

We first assume that $|\rho_x|$ and $|\rho_z|$ are less than one, which means that the instrument and the regressor are I(0) processes. If $\sigma_{12} \neq 0$, the OLS estimator, $\widehat{\theta}_{LS}$, results in asymptotic bias given by the following expression:

$$p\lim\left(\widehat{\theta}_{LS}-\theta\right)=\frac{\sigma_{12}}{\sigma_{22}}(1-\rho_x^2) \tag{5}$$

It can be seen that the asymptotic bias of $\widehat{\theta}_{LS}$ is proportional to the degree of correlation between the regression error and the error that drives the regressor, and inversely proportional to the degree of persistence of the regressor.

Next, assume that $\sigma_{12} \neq 0$ and $\sigma_{13} = 0$. In such a case, θ can be consistently estimated by TSLS. The set of available instruments can be identified by considering the first-stage regression, implied by the DGP under study. This can be obtained by first noting that $u_{2t} = \frac{\sigma_{23}}{\sigma_{33}} u_{3t} + \nu_t$ and then substituting this expression back into equation (2), to obtain,

$$x_t = \rho_x x_{t-1} + \frac{\sigma_{23}}{\sigma_{33}} z_t - \rho_z \frac{\sigma_{23}}{\sigma_{33}} z_{t-1} + \nu_t \tag{6}$$

The first-stage regression implies that there are three available instruments, namely x_{t-1}, z_t and z_{t-1}. In the case that $\rho_x = \sigma_{23} = 0$, the mean of the TSLS estimator employing all the three available, but irrelevant, instruments is the probability limit of the OLS estimator.

Cointegrating Regression

Let us now focus attention on the case $\rho_x = \rho_z = 1$. Equations (1) - (2) form a triangular cointegration system, put forward by Phillips (1988). In such a case, $\widehat{\theta}_{LS}$ is T-consistent, even if $\sigma_{12} \neq 0$. However, if $\sigma_{12} \neq 0$, 'long-run endogeneity' problems (second-order effects) are still encountered within the OLS estimation method. Standard IV procedures are not designed to deal with such effects. Instead, an asymptotically efficient cointegration estimator, such as the Fully Modified Least Squares (FMLS), or the Dynamic OLS (DOLS) estimator should be used. (see Phillips and Hansen 1990, Stock and Watson 1993). Phillips and Hansen (1990) examine the behaviour of IV estimators in a cointegration framework, and show that, due to the non-diagonality of Σ, the presence of relevant instruments makes the asymptotic dependence of the IV estimator on nuisance parameters more complicated than that of the OLS estimator. If, however, the instrument and the regressor error are stochastically independent, that is when $\sigma_{23} = 0$, the nuisance parameter dependencies are reduced. In other words, asymptotic theory suggests that irrelevant instruments are preferable to strong ones, in the case that IV procedures are applied on a cointegrating regression.

Stationary Regression with near-to-unit Roots

Finally, let us assume that ρ_x is close to but less than unity, for example $\rho_x = 0.95$. What is the optimal procedure for estimating θ in this case? Asymptotically, the problem falls into the category of regressions with stationary variables, where only first-order effects, arising from $\sigma_{12} \neq 0$, are present. In finite samples, however, second-order effects, arising from the fact that the regressor resembles a unit-root process are also likely to appear. The presence of both first and second order effects suggests the adoption of an IV estimator with irrelevant but very persistent instruments. Such instruments may be spuriously correlated with the regressor, thus (pseudo) dealing with the first-order effects and, at the same time, minimizing the second-order effects.

3 Monte Carlo Results

The sets of instruments, used in the first-stage regression, are $\{z_t\}$, $\{x_{t-1}\}$, $\{z_t, z_{t-1}, z_{t-2}\}$ and $\{z_t, x_{t-1}, z_{t-1}\}$, resulting in the IVZ, IVX, IVZZ and IVZX estimators, respectively. We also include the OLS estimator for comparison purposes, and two asymptotically efficient cointegration estimators, namely FMLS and DOLS that are expected to perform best in the exact cointegration case ($\rho_x = 1$). The autoregressive parameters, ρ_x and ρ_z, take values in the intervals [0, 0.8] and (0.8, 1], by steps of 0.1 and 0.02, respectively. In the first set of experiments we assume that $\rho_x = \rho_z$. For each value of $\rho_x (= \rho_z)$, we generate 2000 series of length 150 (350) starting with $u_{10} = u_{20} = 0$, and then discard the initial 50 observations, thus generating a sample size of 100 (300). The accuracy of the seven estimators, introduced above, is assessed by means of the median bias, since for IVZ and IVX the unconditional mean does not exist. To examine the effects of persistent instruments on hypothesis testing on θ, we also report the mean, standard deviation, skewness and kurtosis coefficients of the estimators' t-statistics. The performance of these tests is assessed by comparing the 2.5% ($t_{0.025}$) and the 97.5% ($t_{0.975}$) points in the empirical distributions of the relevant t-statistics with those from the standard N(0,1). Finally, we report the (average) F-statistics from the first-stage regressions. As for the rest of the parameters, we set $\theta = 1$, $\sigma_{11} = \sigma_{22} = 1$, $\sigma_{12} = 0.7$ and $\sigma_{13} = 0$, that is, we introduce a rather strong 'endogeneity' effect and maintain the orthogonality condition for z_t. Finally, the key parameter, σ_{23}, is set, throughout, equal to zero. This means that IVZ and IVZZ utilize solely irrelevant instruments for all the values of ρ_x and ρ_z.

For brevity, we do not report the full set of results. Instead, we present the results for the cases $\rho_x = \rho_z = 0$, $\rho_x = \rho_z = 0.5$, $\rho_x = \rho_z = 0.96$ and $\rho_x = \rho_z = 1$ for a sample size equal to 100, in Tables 1A to 1D, respectively. The results may be summarized as follows:

(i) When the regressor and the instrument exhibit zero degree of persistence, that is when $\rho_x = \rho_z = 0$, all the IV estimators employ irrelevant (and serially uncorrelated) instruments and the results are similar to those obtained in the standard 'weak instruments' literature: The F-statistics from the first-stage regressions are very close to unity, and the median bias of each of these estimators is almost identical to the OLS one. The empirical distributions of the associated t-statistics are skewed and shifted to the right, meaning that the t-ratio is expected to be large even if the null hypothesis is true. For example, the 5% empirical sizes of IVZZ and IVZX are 27.4% and 28.1%, respectively.

(ii) When the regressor and the instrument exhibit a moderate degree of persistence, that is when $\rho_x = \rho_z = 0.5$, the results are, to a large extent, consistent with the relevant theory. The best performing estimator is IVZX, whose median bias is smaller than that of OLS by a factor of twenty, followed by IVX. For this level of persistence, IVZ and IVZZ still follow, to a large extent, the behaviour of OLS. However, some small but important differences between this and the previous case are visible: First, the F-statistics for IVZ and IVZZ have increased from 0.98 to 1.70 and from 0.99 to 1.28, respectively, despite the fact that their population analogues, remain fixed to zero. Second, the median bias of IVZ as a ratio to that of OLS has decreased from 1.004, in the zero persistence case, to 0.87 in the present case. Third, the distributional divergencies of the IVZ and IVZZ t-statistics from the standard normal, have slightly decreased.

(iii) As the degree of persistence rises, the performance of IVZ and IVZZ improves monotonically. For $\rho_x = \rho_z = 0.96$, the instruments, employed by these estimators, do not appear to be irrelevant at all! The corresponding F-statistics are now as large as 20.94 and 8.28, respectively, thus heavily over-estimating their population analogues, which remain equal to zero. This means that 'spurious' regression effects in the first-stage regressions are clearly in place, despite the fact that the series involved are still I(0). However, these effects turn out to be quite beneficial as far as statistical inferences on θ are concerned. The median bias of IVZ (IVZZ), as a ratio to the median bias of OLS, is as low as 0.42 (0.57). Moreover, the distribution of the IVZ t-statistic is located close to zero (around 0.307) as opposed to that of OLS, located around 2.16. In fact, IVZ produces the best-centered t-statistic of all the estimators under consideration. For example, the mean value of the IVZ t-statistic is closer to zero than that of the IVX t-statistic, which reaches the value of -0.592. In other words, the empirical distribution of the t-statistic produced by an IV estimator utilizing an irrelevant instrument is better centered than that of an IV estimator, employing an extremely strong instrument. Moreover, the IVZ t-statistic is, in general, better approximated by a standard N(0,1), than any other estimator's t-statistic. For example, the $t_{0.025}$ and $t_{0.975}$ points for IVZ are -1.21 and 2.09 respectively, thus resulting in an empirical size of 3.6%. On the other hand, the corresponding pairs for OLS, IVX, DOLS and FMLS are (0.588, 3.776), $(-2.37, 1.332)$, $(-1.219, 3.773)$, and $(-0.679, 3.773)$, resulting in empirical sizes of 59.1%, 7.05%, 25.5% and 35.95%, respectively. This in turn implies that IVZ outperforms not only IVX, but also FMLS and DOLS, as far as hypothesis testing on θ is concerned.

(iv) In the extreme case $\rho_x = \rho_z = 1$, IVZZ and, especially, IVZ continue to perform surprisingly well. In this case the dominance of IVZ over IVX is clear in all aspects of statistical inference. For example, the mean values of the IVZ and IVX t-statistics are 0.040 and -1.019, respectively and the $(t_{0.025}, t_{0.975})$ pairs are $(-1.603, 1.739)$, and $(-2.645, 0.840)$, respectively. It is interesting to note that the performance of IVZ is comparable even to that of the cointegration estimators, FMLS and DOLS, which now operate in their natural environment.

The effects described above are summarized in Figures 1 and 2, that describe the median bias of IVZ and IVX, respectively relative to that of OLS, for sample sizes of 100 and 300. It can be seen that the relative bias of IVZ, as opposed to that of IVX, tends to zero as $\rho_x (= \rho_z)$ tends to one. It can also be seen that as the sample size increases, and the relevant asymptotic theory of stationary regressions becomes more relevant, the 'irrelevant instruments' effect weakens. However, the rate at which this effect declines appears to be extremely slow.

In all the experiments, so far, we have retained the assumption $\rho_x = \rho_z$, that is, the instruments and the regressor exhibit the same degree of persistence. How many of the above results remain valid when $\rho_x \neq \rho_z$? To answer this question, we run another set of experiments, where the value of ρ_x is kept fixed to a particular value from the set $I = \{0, 0.1, ...1\}$. For this value of ρ_x, ρ_z takes sequentially all the values of I. We repeat the same procedure until all the values of $\rho_x \in I$ are exhausted. Overall, we run 121 simulations, plus some additional, more specific ones, for ρ_x in the neighborhood of unity. The results (not reported) suggest that the general picture, described above, remains the same for the cases that the instruments and the regressor exhibit different degrees of persistence, provided that the difference $|\rho_x - \rho_z|$ is not very large. For example, when $\rho_x = 0.96$, then IVZ performs satisfactorily well for a value of ρ_z as low as 0.8 (and, of course, as large as unity).

4 Conclusions

Our conclusions from the investigation of the behaviour of the TSLS procedure, under alternative degrees of persistence of the regressor and the instruments used, are the following: First, the performance of the estimator, utilizing solely irrelevant instruments, improves monotonically, as the degree of persistence of the regressor and that of the instruments, increases. Second, in the case where the regressor and the instruments are near-to-unit root processes, the estimator that utilizes a single irrelevant instrument, outperforms IV estimators with strong instruments, as well as asymptotically efficient cointegration estimators.

References

[1]Bekker, P.A. (1994). Alternative Approximations to the Distributions of Instrumental Variables Estimators, *Econometrica*, 62, 657-681.

[2]Bound, J., D. Jaeger, and R. Baker (1995): Problems with Instrumental Variable estimation When the Correlation between the Instruments and the Endogenous Explanatory Variable is Weak, *Journal of the American Statistical Association*, 90, 443-450.

[3]Buse, A. (1992) The Bias of Instrumental Variable Estimators, *Econometrica*, 60, 173-180.

[4]Dufour, J.M. (1997), Some Impossibility Theorems in Econometrics with Application to Structural and Dynamic Models, *Econometrica*, 65, 1365-1388.

[5]Elliot, G. (1998). The Robustness of Cointegration Methods when Regressors Almost Have Unit Roots, *Econometrica*, 66, 149-158.

[6]Nelson, C.R. and R.Startz (1990a): The Distribution of the Instrumental Variable Estimator and its t-ratio when the Instrument is a Poor One, *Journal of Business, 63, S125-S140.*

[7]Nelson, C.R. and R.Startz (1990b): Some Further Results on the Exact Small Sample properties of the Instrumental Variable Estimator, *Econometrica,* 58, 967-976.

[8]Park, J.Y. and P.C.B. Phillips (1988), Statistical Inference in Regressions with Integrated Processes: Part 1, *Econometric Theory*, 4, 468-498.

[9]Phillips, P.C.B. (1988), Reflections on Econometric Methodology, *Economic Record*, 64, 344-359.

[10]Phillips, P.C.B. and E.J. Hansen (1990), Statistical Inference in Instrumental Variables Regressions with I(1) processes, *Review of Economic Studies*, 57, 99-125.

[11]Phillips, P.C.B. and M. Loretan (1991), Estimating Long-run Economic Equilibria, *Review of Economic Studies,* 58, 407-436.

[12]Staiger, D. and J.H. Stock (1997), IV Regression with Weak Instruments, *Econometrica*, 65, 557-586.

[13]Stock, J.H. and M.W. Watson (1993), A Simple Estimator of Cointegrating Vectors in Higher-order Integrated Systems, *Econometrica*, 61, 783-820.

[14]Wang, J. and E. Zivot (1998), Inference on Structural Parameters in Instrumental Variables Regression with Weak Instruments, *Econometrica,* 66,1389-1404.

Table 1

Small Sample Performance of Alternative Estimators

(Sample Size = 100)

Estimator	*Median bias*	*Mean(t)*	*Standard deviation (t)*	*Skewness*	*Kurtosis*	$t_{0.025}$	$t_{0.975}$	*F-stat*
Panel A				$\rho_x=\rho_z=0$				
OLS	0.698	9.809	1.447	0.159	3.107	7.054	12.771	---
IVZ	0.701	0.796	0.783	0.848	3.993	-0.378	2.615	0.98
IVX	0.737	0.856	0.815	0.657	3.094	-0.347	2.672	1.03
IVZZ	0.681	1.371	1.018	0.391	3.031	-0.340	3.511	0.99
IVZX	0.686	1.386	1.019	0.321	2.777	-0.317	3.464	0.99
DOLS	0.694	7.326	1.836	0.555	3.413	4.232	11.493	---
FMLS	0.698	10.155	1.887	0.376	3.602	6.707	14.170	---
Panel B				$\rho_x=\rho_z=0.5$				
OLS	0.531	7.601	1.114	0.113	2.950	5.486	9.789	---
IVZ	0.463	0.693	0.793	0.576	3.008	-0.532	2.494	1.70
IVX	-0.026	-0.016	0.925	0.736	3.400	-1.404	2.105	31.51
IVZZ	0.509	1.149	0.895	0.200	2.719	-0.424	2.962	1.28
IVZX	0.020	0.243	0.971	0.687	3.452	-1.286	2.391	11.20
DOLS	0.351	5.215	1.515	0.511	3.401	2.676	8.478	---
FMLS	0.471	7.334	1.510	0.368	3.533	4.595	10.559	---
Panel C				$\rho_x=\rho_z=0.96$				
OLS	0.077	2.155	0.825	0.048	3.136	0.588	3.776	---
IVZ	0.032	0.307	0.844	0.302	2.841	-1.210	2.092	20.94
IVX	-0.025	-0.592	0.936	0.176	3.100	-2.370	1.332	714.59
IVZZ	0.044	0.404	0.837	0.193	2.826	-1.123	2.138	8.28
IVZX	-0.022	-0.516	0.933	0.167	3.109	-2.285	1.404	240.99
DOLS	0.028	1.161	1.243	0.107	3.335	-1.219	3.773	---
FMLS	0.032	1.567	1.135	0.061	3.174	-0.679	3.848	---
Panel D				$\rho_x=\rho_z=1$				
OLS	0.030	1.080	0.936	-0.081	3.136	-0.816	2.886	---
IVZ	0.001	0.040	0.856	0.074	2.802	-1.603	1.739	54.84
IVX	-0.030	-1.019	0.897	0.237	3.203	-2.645	0.840	1558.69
IVZZ	0.006	0.102	0.877	0.046	2.784	-1.593	1.781	20.37
IVZX	-0.029	-0.958	0.897	0.212	3.190	-2.599	0.909	525.99
DOLS	0.000	0.004	1.143	0.029	3.423	-2.234	2.281	---
FMLS	0.001	0.097	1.077	0.072	3.136	-1.952	2.293	---

Figure 1

Median Bias of IVZ relative to that of OLS

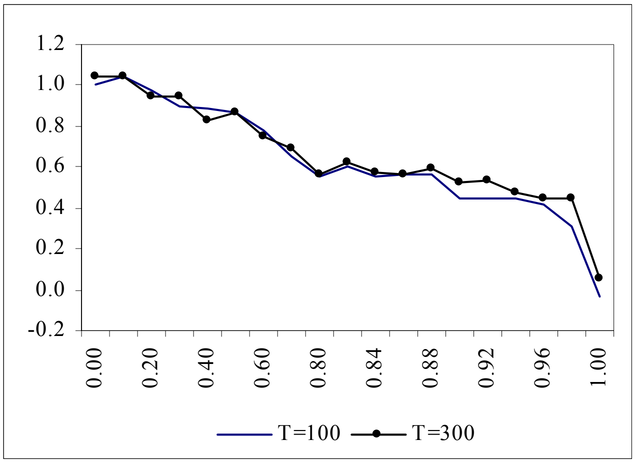

Figure 2

Median Bias of IVX relative to that of OLS