ISSN 1503-299X

WORKING PAPER SERIES

No. 3/2006

A GAUSSIAN IV ESTIMATOR OF COINTEGRATING RELATIONS

Gunnar Bårdsen
Niels Haldrup

Department of Economics

Norwegian University of Science and Technology
N-7491 Trondheim, Norway
www.svt.ntnu.no/iso/wp/wp.htm

A Gaussian IV estimator of cointegrating relations

Gunnar Bårdsen and Niels Haldrup*

13 February 2006.

Abstract. In static single equation cointegration regression models the OLS estimator will have a non-standard distribution unless regressors are strictly exogenous. In the literature a number of estimators have been suggested to deal with this problem, especially by the use of semi-nonparametric estimators. Theoretically *ideal* instruments can be defined to ensure a limiting Gaussian distribution of IV estimators, but unfortunately such instruments are unlikely to be found in real data. In the present paper we suggest an IV estimator where the Hodrick-Prescott filtered trends are used as instruments for the regressors in cointegrating regressions. These instruments are *almost ideal* and simulations show that the IV estimator using such instruments alleviate the endogeneity problem extremely well in both finite and large samples.

Keywords: Cointegration, Instrumental variables, Mixed Gaussianity.
JEL Classification: C2, C22, C32

1. Introduction

It is well known that non-stationary, integrated, time series generally give rise to estimators and test statistics having non-standard distributions, that is, distributions that will not be limiting Gaussian distributions. Estimating cointegrating regressions, in particular, has generated considerable interest ever since the seminal contributions of Engle and Granger (1987). Even though their static estimator is super-consistent asymptotically, it was early noted (Banerjee et al., 1986) that it could be biased in finite samples. Furthermore, the distribution of the parameters are non-Gaussian unless the regressors are strictly exogenous, invalidating hypothesis testing in the usual manner by comparing t-ratios with the Gaussian distribution.

Many alternative estimators have been proposed as solutions to these problems. They generally fall into two categories: single equation methods with parametric or non-parametric corrections, or system estimators modelling the endogeneity parametrically. Examples of the former category include Phillips and Hansen (1990), Phillips and Loretan (1991), Saikkonen (1991), Stock (1987), Stock and Watson (1993), Park

*We are grateful for comments from Arvid Raknerud, Anders Rygh Swensen and participants at the Third Annual Conference of the European Social Science Network on Econometric Methods for the Modelling of Nonstationary Data, Policy Analysis and Forecasting (EMM), Sassari, Italy, 29 September - 2 October 2004. Affiliations: Gunnar Bårdsen: Department of Economics, Norwegian University of Science and Technology, 7491 Trondheim, Norway, and Bank of Norway, e-mail: gunnar.bardsen@svt.ntnu.no. Niels Haldrup: Department of Economics, University of Aarhus, Building 1322, DK-8000 Aarhus C, Denmark. e-mail: nhaldrup@econ.au.dk

(1992), and Phillips (2006). Examples of the latter category include Stock and Watson (1988) and the Johansen (1988) full information maximum likelihood method.

In this paper we propose yet another estimator based on instrumental variables. The motivation for the estimator is twofold: first, existing Monte-Carlo studies (Banerjee et al., 1993; Stock and Watson, 1993; Gonzalo, 1994; Haug, 1996) suggest that none of the existing single-equation methods seem to be superior—not even satisfactory. Second, even though the full information method of Johansen is appealing both from Monte-Carlo evidence as well as from a methodological perspective, the empirical implementation may entail difficulties (Stock and Watson, 1993). Depending upon the problem at hand, obtaining a satisfactory representation for all endogenous variables can be challenging.

An alternative would therefore be to turn to limited information methods like instrumental variable estimators, recognizing that only some of the cointegrating vectors are likely to be modelled with any success. This is the approach taken by Phillips and Hansen (1990). However, this brings up the question of finding good instruments. Any instrument has got to be cointegrated with the variable in question in the present case, in addition to being uncorrelated with the system variables. This seems to bring the argument full circle: one is back to the uncertain task of modelling a system of cointegrating relationships of varying quality, unless one can ensure the quality of the instruments. In their follow up simulation study Hansen and Phillips (1990) recognize this problem by using spurious instruments in addition to *ideal* instruments which unfortunately are unlikely to be found in real data. The use of spurious or irrelevant instruments has recently attracted some attention, see Phillips (2006), who shows how (irrelevant) deterministically trending instruments can be used to produce asymptotically efficient estimates of a cointegrated system.

The approach we take in this paper is instead one of generating *relevant* instruments, i.e. instruments that almost fulfill both the necessary conditions, and then use a standard instrumental variable estimator to obtain the cointegrating relationships of interest. Trends generated from trend-cycle decompositions based on e.g. the Hodrick-Prescott filter serves the role of being valid instruments. Monte-Carlo evidence suggests that the approach works surprisingly well in finite samples. In particular, the finite sample distributions of the t-ratios in most cases are very close to the standard normal distribution, thereby allowing standard inference.

2. Some preliminaries

Following Phillips and Durlauf (1986) and Davidson (2000), consider the model

$$\begin{aligned} y_{1t} &= \gamma' \mathbf{y}_{2t} + u_{1t} \\ \Delta \mathbf{y}_{2t} &= \mathbf{u}_{2t} \end{aligned}$$

where u_{1t} is a scalar residual and $\mathbf{v}_t$ is of dimension $(m-1) \times 1$. Conformably with this, define $\eta_t = (u_{1t}, \mathbf{u}_{2t}')'$ of dimension $m \times 1$ and consider the sequential sum

$\mathbf{S}_t = \sum_{j=1}^{t} \eta_j$ with $\mathbf{S}_0 = \mathbf{0}$. Also define

$$\mathbf{\Sigma} = \lim_{\mathbf{n}\to\infty} \frac{1}{n} \sum_{t=1}^{n} E(\eta_t \eta_t')$$

$$\mathbf{\Lambda} = \lim_{\mathbf{n}\to\infty} \frac{1}{n} \sum_{t=1}^{n} E(\mathbf{S}_t \eta_t') = \lim_{n\to\infty} \frac{1}{n} \sum_{t=2}^{n} \sum_{j=1}^{t-1} E(\eta_j \eta_t')$$

$$\lim_{n\to\infty} \frac{1}{n} \sum E(\mathbf{S}_n \mathbf{S}_n') = \mathbf{\Omega} = \mathbf{\Sigma} + \mathbf{\Lambda} + \mathbf{\Lambda}' < \infty$$

where

$$\mathbf{\Omega} = \begin{bmatrix} \omega_{11} & \mathbf{\Omega}_{12} \\ \mathbf{\Omega}_{21} & \mathbf{\Omega}_{22} \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{21} & \mathbf{\Sigma}_{22} \end{bmatrix} + \begin{bmatrix} \lambda_{11} & \mathbf{\Lambda}_{12} \\ \mathbf{\Lambda}_{21} & \mathbf{\Lambda}_{22} \end{bmatrix} + \begin{bmatrix} \lambda_{11} & \mathbf{\Lambda}_{12}' \\ \mathbf{\Lambda}_{21}' & \mathbf{\Lambda}_{22}' \end{bmatrix}$$

Under regularity conditions we have that

$$\frac{1}{\sqrt{n}} \mathbf{S}_{[nr]} \xrightarrow{d} \mathbf{B}(r) = \begin{pmatrix} B_1(r) \\ \mathbf{B}_2(r) \end{pmatrix}$$

i.e. a vector Brownian motion process defined on the unit interval $[0, 1]$.

The continuous mapping theorem states that

$$\frac{1}{n^{3/2}} \sum_{t=1}^{n} \mathbf{S}_t \xrightarrow{d} \int_0^1 \mathbf{B} dr$$

$$\frac{1}{n^2} \sum_{t=1}^{n} \mathbf{y}_{2t} \mathbf{y}_{2t}' \xrightarrow{d} \int_0^1 \mathbf{B}_2 \mathbf{B}_2' dr$$

$$\frac{1}{n} \sum_{t=1}^{n} \mathbf{y}_{2t} u_{1t} \xrightarrow{d} \int_0^1 \mathbf{B}_2 dB_1 + \mathbf{\Sigma}_{21} + \mathbf{\Lambda}_{21}$$

Note that

$$\mathbf{\Sigma}_{21} = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E(\mathbf{u}_{2t} u_{1t)}$$

$$\mathbf{\Lambda}_{21} = \lim_{n\to\infty} \frac{1}{n} \sum_{t=2}^{n} \sum_{j=1}^{t-1} E(\mathbf{u}_{2t-j} u_{1t})$$

ESTIMATION BY OLS:

Consider the static least squares regression

$$y_{1t} = \widehat{\gamma}' \mathbf{y}_{2t} + \widehat{u}_{1t}$$

Given the above results and the continuous mapping theorem it follows that

$$n(\widehat{\gamma}-\gamma) = \left(\frac{1}{n^2}\sum_{t=1}^{n}\mathbf{y}_{2t}\mathbf{y}_{2t}'\right)^{-1}\frac{1}{n}\sum_{t=1}^{n}\mathbf{y}_{2t}u_{1t}$$
$$\xrightarrow{d}\left(\int_0^1 \mathbf{B}_2\mathbf{B}_2' dr\right)^{-1}\left(\int_0^1 \mathbf{B}_2 dB_1 + \mathbf{\Sigma}_{21} + \mathbf{\Lambda}_{21}\right)$$

This distribution is *not* a standard Gaussian distribution and hence we cannot conduct standard inference based on this.

Mixed Gaussianity.

If it occurs that y_{2t} is strictly exogenous, then $\Sigma_{21} = \Lambda_{21} = 0$ and B_1 and B_2 are uncorrelated. In this case the above result is modified and the relevant limiting distribution is given by (see e.g. Davidson, 2000)

$$n(\widehat{\gamma}-\gamma) \xrightarrow{d} \int_0^1 \mathbf{N}(\mathbf{0},\omega_{11}\mathbf{G}^{-1})dP(\mathbf{G})$$

where

$$\mathbf{G} = \int_0^1 \mathbf{B}_2\mathbf{B}_2' dr$$

This distribution is known as a mixture of normals, or mixed Gaussian distribution, that is, conditional on $\mathbf{B}_2$ the resulting distribution is normal. For a fixed vector $\mathbf{a} \neq \mathbf{0}$ (for instance the i'th column of the identity matrix) it holds that

$$n\frac{\mathbf{a}'(\widehat{\gamma}-\gamma)}{\sqrt{\omega_{11}\mathbf{a}'\mathbf{G}^{-1}\mathbf{a}}} \xrightarrow{d} \mathrm{N}(0,1)$$

Note however, that this assumes ω_{11} to be known. If u_{1t} is serially uncorrelated there are no problems and $\omega_{11} = \sigma_{11}$ can be estimated in a standard fashion. However, in the presence of autocorrelation a Newey-West based estimator can be adopted in the estimation of ω_{11}.

The main requirement for the mixed gaussianity result to apply is that the parameters Σ_{21} and Λ_{21} vanish and that B_1 and $\mathbf{B}_2$ are uncorrelated. The motivation underlying the present paper is to design the statistical model in such a way that these nuisance parameters are annihilated and hence subsequently an approximate limiting mixed Gaussian distribution can be achieved. Our idea is to consider an instrumental variables estimator of the cointegrating relation where the instruments are chosen as appropriately filtered series.

3. An IV estimator of cointegrating relations

Phillips and Hansen (1990) and Hansen and Phillips (1990) examine the use of instrumental variables in the estimation of parameters of cointegrating relations. A number of cases are considered, e.g. standard IV estimation using instruments that are cointegrated with the variables they are intended to instrument as well as spurious stochastic and deterministic instruments being spurious in the sense that they have no structural relationship to the variables being instrumented. Interestingly, the IV estimator will be consistent irrespective of the properties of the instruments being spurious or non-spurious, a beneficial artifact of spurious regression theory[1]. Nonetheless, standard Gaussian limiting distributions when testing hypotheses on model parameters do not generally apply unless some bias correction or fully modified correction is made of the estimators. However, one particular case is an exception namely when instruments are chosen to be cointegrated with the explanatory variables *and* are uncorrelated with the system variables. In this case instruments are *ideal* and a mixed Gaussian distribution results.

To see this, assume that instruments can be found for each element in the $\mathbf{y}_{2t}$-vector and denote these $\mathbf{z}_{2t}$. The instruments should be chosen such that the single rows in $\mathbf{y}_{2t}$ cointegrate with the associated row in $\mathbf{z}_{2t}$ and such that the instruments chosen are uncorrelated with the error term u_{1t}. Hence we have as many instruments as we have variables to instrument.

The model to be estimated reads

$$y_{1t} = \widetilde{\gamma}\mathbf{y}_{2t} + \widetilde{u}_{1t}$$

where "$\sim$" signifies IV estimates.

The IV estimator can be written

$$\widetilde{\gamma} = \left(\sum_{t=1}^{n} \mathbf{z}_{2t}\mathbf{y}_{2t}'\right)^{-1} \sum_{t=1}^{n} \mathbf{z}_{2t} y_{1t}.$$

from which it follows that

$$n\left(\widetilde{\gamma} - \gamma\right) = \left(\frac{1}{n^2}\sum_{t=1}^{n} \mathbf{z}_{2t}\mathbf{y}_{2t}'\right)^{-1} \frac{1}{n}\sum_{t=1}^{n} \mathbf{z}_{2t} u_{1t}$$

Now, define

$$\mathbf{z}_{2t} = \mathbf{y}_{2t} - \mathbf{v}_{2t}, \tag{1}$$

where $\mathbf{v}_{2t}$ is a stationary (cyclical) component of the $\mathbf{y}_{2t}$ series and $\mathbf{v}_{2t}$ is orthogonal to the increments of $\mathbf{z}_{2t}$ and u_{1t}.

[1] In a recent paper Phillips (2006) exploits this feature to develop an asymptotically efficient estimator of cointegrated systems based on *irrelevant* deterministically trending instruments.

It follows that

$$\frac{1}{n^2}\sum_{t=1}^{n}\mathbf{z}_{2t}\mathbf{y}_{2t}' = \frac{1}{n^2}\sum_{t=1}^{n}(\mathbf{y}_{2t}-\mathbf{v}_{2t})\mathbf{y}_{2t}' + o_p(1)$$

$$\xrightarrow{d} \int_0^1 \mathbf{B}_2\mathbf{B}_2' dr \tag{2}$$

$$\frac{1}{n}\sum_{t=1}^{n}\mathbf{z}_{2t}u_{1t} = \frac{1}{n}\sum_{t=1}^{n}(\mathbf{y}_{2t}-\mathbf{v}_{2t})u_{1t}$$

$$\xrightarrow{d} \int_0^1 \mathbf{B}_2 dB_1 \tag{3}$$

And from the continous mapping theorem we have that

$$n\left(\widetilde{\gamma}-\gamma\right) = \left(\frac{1}{n^2}\sum_{t=1}^{n}\mathbf{z}_{2t}\mathbf{y}_{2t}'\right)^{-1}\frac{1}{n}\sum_{t=1}^{n}\mathbf{z}_{2t}u_{1t}$$

$$\xrightarrow{d} \int_0^1 \mathbf{N}(\mathbf{0},\omega_{11}\mathbf{G}^{-1})dP(\mathbf{G})$$

which mimics the case of strictly exogenous regressors and mixed Gaussianity.

It also follows straightforwardly that

$$n\frac{\mathbf{a}'(\widehat{\gamma}-\gamma)}{\sqrt{\omega_{11}\mathbf{a}'\mathbf{G}^{-1}\mathbf{a}}} \xrightarrow{d} N(0,1)$$

When $\mathbf{a}$ is a column from the identify matrix, this is just a $t-$test, given, of course, that the parameter ω_{11} is estimated. If u_{1t} has no autocorrelation ω_{11} is estimated using the estimator $s^2 = \frac{1}{n}\sum_{t=1}^{n}\widetilde{u}_{1t}^2$. If there is autocorrelation a Newey-West type estimator is needed.

Hansen and Phillips (1990) argue that in the above case $\mathbf{z}_{2t}$ is indeed an *ideal* instrument for $\mathbf{y}_{2t}$ because these cointegrate and yet $\mathbf{z}_{2t}$ is not contaminated with u_{1t} as is $\mathbf{y}_{2t}$. They also argue that this case unfortunately is idealized in actual applications because it can be hard to find instruments $\mathbf{z}_{2t}$ satisfying the required properties. However, we will argue that indeed this criticism is perhaps not too big a practical problem since filtered $\mathbf{y}_{2t}$ series, using the decomposition $\mathbf{y}_{2t} = \mathbf{z}_{2t} + \mathbf{v}_{2t}$ can be easily constructed. For instance, structural time series models, see e.g. Harvey (1989), are founded on such decompositions of time series where in the present case the multivariate random walk component $\mathbf{z}_{2t}$ will serve as instruments. But the series can also be filtered otherwise by the Hodrick-Prescott filter say where the trend components will serve as instruments and can be straightforwardly extracted. Instruments created in this fashion are *almost ideal* in the sense that the instruments $\mathbf{z}_{2t}$ cointegrate with $\mathbf{y}_{2t}$ and are approximately uncorrelated with both $\mathbf{v}_{2t}$ and u_{1t}. Nothing guarantees exact uncorrelatedness of these terms. However, the filtered trend series, i.e. the instruments, will clearly play the role of reducing the correlation compared to the non-filtered series and hence will annihilate the endogeneity which was the source of the problem.

3.1. Choice of Instruments. In the previous section we suggested to use instruments for the elements in $\mathbf{y}_{2t}$ that cointegrate with $\mathbf{y}_{2t}$ and which are serially independent of u_{1t}. We further required that instruments $\mathbf{z}_{2t}$ satisfy $\mathbf{y}_{2t} = \mathbf{z}_{2t} + \mathbf{v}_{2t}$ with $\Delta\mathbf{z}_{2t}$ being *almost* orthogonal to $\mathbf{v}_{2t}$ and u_{1t}. Such decompositions can be undertaken in a number of ways. Here we advocate for using the Hodrick-Prescott filtered $\mathbf{y}_{2t}$ series as instruments. In so doing, $\mathbf{z}_{2t}$ becomes the non-stationary secular component of the series which $\mathbf{v}_{2t}$ is the stationary cyclical component. The HP detrending method is based on the assumption that the row elements of $\mathbf{z}_{2t}$ and $\mathbf{v}_{2t}$ are statistically independent. An estimate of the filtered series can be found as the solution to the convex optimization problem

$$\min_{\{z_{2it}\}_{t=1}^{n}} \left\{ \sum_{t=1}^{n} (y_{2it} - z_{2it})^2 + \lambda \sum_{t=2}^{n-1} \left[(z_{2it+1} - z_{2it}) - (z_{2it} - z_{2it-1}) \right]^2 \right\}$$

where λ is a smoothing parameter which regulates the trade-off between the goodness of fit and the smoothness of the HP-trend series. Typically, $\lambda = 1600$ is used for quarterly data whereas $\lambda = 129{,}600$ and 6.25 for monthly and annual data, respectively.

4. Simulations

The simulations reported in this section will investigate the performance of our proposed estimator in small and larger samples, consisting of 50 and 500 observations. In all cases the first generated 100 observations were discarded, to avoid dependence upon starting values. The data-generating process we use is taken from Phillips and Loretan (1991), providing a convenient benchmark. The model is

$$\begin{aligned} y_{1t} &= \alpha + \gamma y_{2t} + u_{1t} \\ \Delta y_{2t} &= u_{2t}, \quad t = 1, \ldots, T \\ \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} &= u_t = \varepsilon_t + \theta \varepsilon_{t-1}, \quad \varepsilon_t \equiv iid\ N(0, \Sigma), \end{aligned}$$

where

$$\begin{aligned} \alpha &= 0, \qquad \gamma = 2, \qquad T = 50, 500 \\ \theta &= \begin{bmatrix} 0.3 & 0.4 \\ \theta_{21} & 0.6 \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 1 & \sigma_{21} \\ \sigma_{21} & 1 \end{bmatrix}, \end{aligned}$$

and where θ_{21} and σ_{21} are varied over the values

$$\begin{aligned} \theta_{21} &= \{0.8, 0.4, 0.0, -0.8\} \\ \sigma_{21} &= \{-0.85, -0.5, 0.5\} \end{aligned}$$

The estimators considered are OLS

$$y_{1t} = \hat{\alpha} + \hat{\gamma} y_{2t} + \hat{u}_{1t}$$

and filtered IV

$$y_{1t} = \tilde{\alpha} + \tilde{\gamma} y_{2t} + \tilde{u}_{1t},$$

where $z_{3t} = HP(y_{2t})$, the Hodrick-Prescott filtered trend of y_{2t}, is used as an instrument for y_{2t}. The IV-estimators are labelled $\{IV1, IV2, IV3\}$ corresponding to smoothing parameter $\lambda = \{6.25, 1600, 129600\}$. The computed estimates are, with variables expressed in deviations from means, for OLS:

$$\hat{\gamma} = \frac{\sum_{t=1}^{T} y_{1t} y_{2t}}{\sum_{t=1}^{T} y_{2t}^2}$$

$$\hat{t} = \frac{\hat{\gamma} - \gamma}{\widehat{se(\hat{\gamma})}},$$

$$\text{where } \widehat{se(\hat{\gamma})} = \sqrt{\frac{\sum_{t=1}^{T} (y_{1t} - \hat{\gamma} y_{2t})^2}{(T-2) \sum_{t=1}^{T} y_{2t}^2}}.$$

and for IV:

$$\tilde{\gamma} = \frac{\sum_{t=1}^{T} z_{2t} y_{1t}}{\sum_{t=1}^{T} z_{2t} y_{2t}}$$

$$\tilde{t} = \frac{\tilde{\gamma} - \gamma}{\widehat{se(\tilde{\gamma})}},$$

$$\text{where } \widehat{se(\tilde{\gamma})} = \sqrt{\frac{\sum_{t=1}^{T} (y_{1t} - \tilde{\gamma} y_{2t})^2 \sum_{t=1}^{T} z_{2t}^2}{T \left(\sum_{t=1}^{T} z_{2t} y_{2t} \right)^2}}.$$

In the simulations reported, the standard errors for both estimators are adjusted using a Newey and West (1987) autocovariance correction. However, simulations without corrections produced very similar results, providing support for the asymptotic theory.

5. Distribution of t-ratios

To save space, we will focus on the distributions of the t-ratios. The qualitative results for the parameter biaseses are similar and are available upon request.

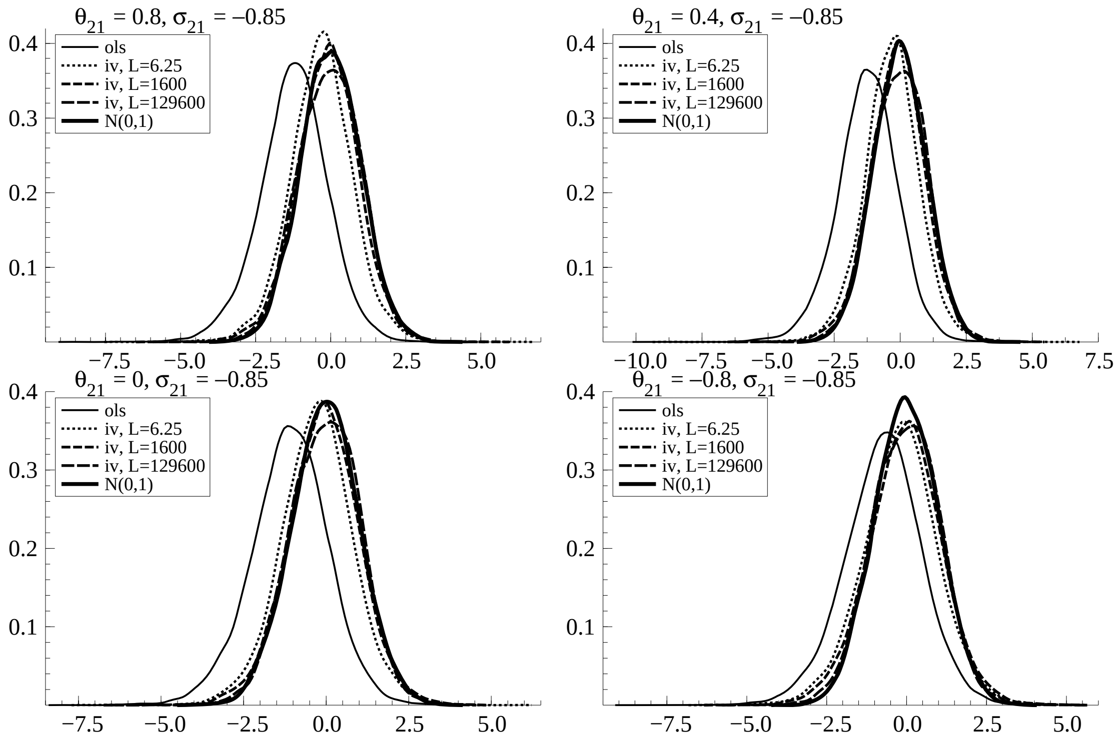

Figure 1: Distributions of t-ratios with $\sigma_{21} = -0.85$, using 50 observations.

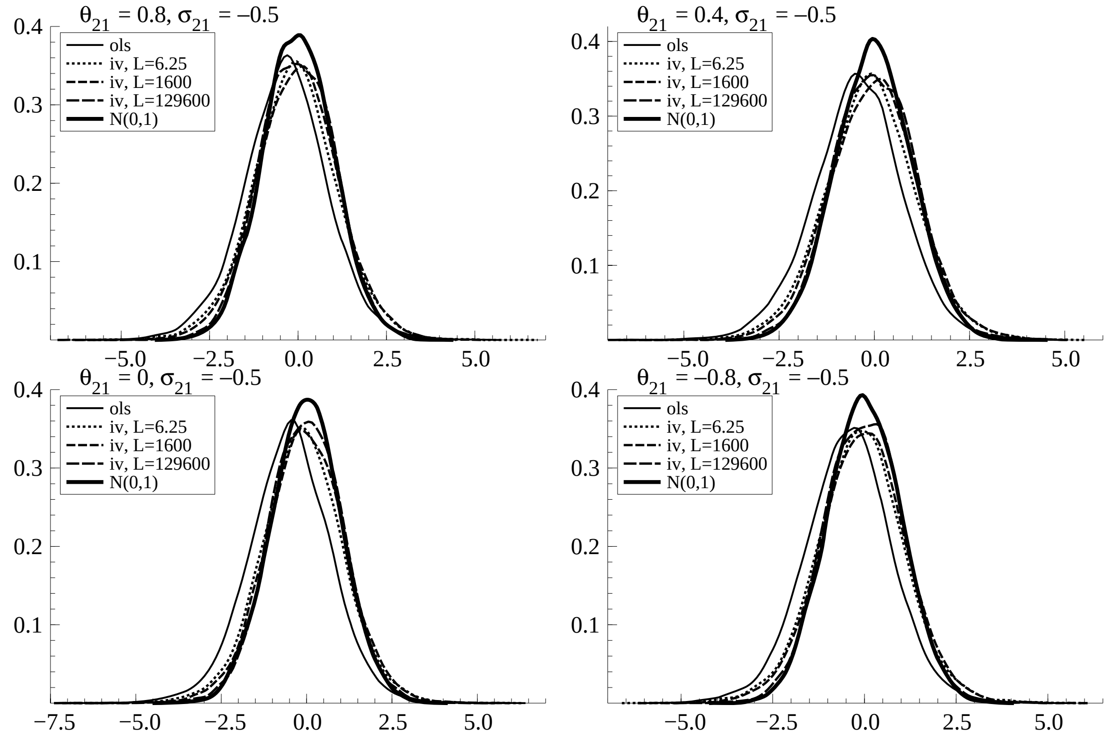

Figure 2: Distributions of t-ratios with $\sigma_{21} = -0.5$, using 50 observations.

5.1. Small sample properties: 50 observations. The results using 50 observations are illustrated in Figures 1–3, containing the estimated densities, and summarized in the percentile table 1. The results for OLS are comparable to the ones reported by Phillips and Loretan (1991), with negative biases in most cases. The problems of OLS are most substantial in the case of negative covariance with positive moving average errors. Turning to our proposed IV estimator, the simulations are very encouraging. The estimator provides very good corrections. As demonstrated in the graphs, the small-sample distributions are quite similar to the predicted asymptotic normality. As regards the relative performance of the different instruments, examination of table 1 reveals a pervasive pattern of the results improving with the degree of smoothing. The case of $\lambda = 129600$, yields close to normal critical values in nearly all instances. In general, the simulations indicate that this simple estimator performs quite well, especially in comparison to supposedly more optimal procedures, as reported in Phillips and Hansen (1990); Phillips and Loretan (1991); Stock and Watson (1993); Haug (1996).

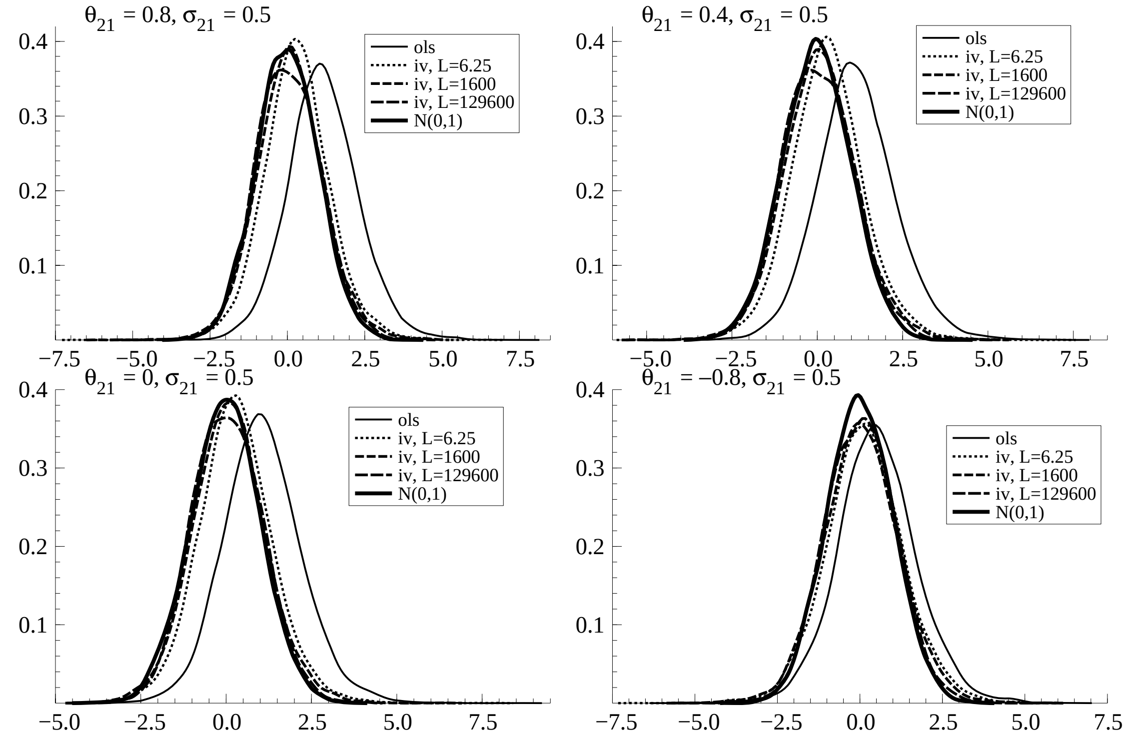

Figure 3: Distributions of t-ratios with $\sigma_{21} = 0.5$, using 50 observations.

However, the sample can also become so small that the simulations become misleading. When doing simulations there is the possibility of an estimator performing well simply because the consequences of the error processes are not allowed to show themselves in too small samples. To investigate this argument further, we therefore conduct another simulation, using 500 observations. Although our data-generating process have been used in several other studies, cited above, the larger-sample properties of this particular data-generating process have not been investigated before, to our knowledge.

Table 1: Percentiles of t-ratios using 50 observations. IV1 is using $\lambda = 6.25$; IV2 is using $\lambda = 1600$; IV3 is using $\lambda = 129600$;

	2.5%	5%	50%	95%	97.5%	2.5%	5%	50%	95%	97.5%
					$\sigma_{21} = -0.85$					
			$\theta_{21} = .8$					$\theta = .4$		
OLS	-3.60	-3.16	-1.21	0.54	0.93	-3.63	-3.17	-1.22	0.53	0.93
$IV1$	-2.42	-2.02	-0.30	1.34	1.77	-2.45	-2.05	-0.3	1.34	1.76
$IV2$	-2.19	-1.83	-0.08	1.61	1.96	-2.26	-1.84	-0.08	1.61	1.97
$IV3$	-2.0	-1.69	-0.02	1.57	1.89	-2.02	-1.71	-0.02	1.58	1.89
			$\theta_{21} = 0$					$\theta_{21} = -.8$		
OLS	-3.52	-3.1	-1.08	0.74	1.12	-3.24	-2.76	-0.70	1.18	1.58
$IV1$	-2,52	-2.09	-0.26	1.48	1.9	-2.62	-2.14	-0.16	1.73	2.11
$IV2$	-2.29	-1.87	-0.06	1.66	2.04	-2.39	-1.99	-0.02	1.81	2.21
$IV3$	-2.05	-1.74	0.0	1.60	1.91	-2.10	-1.76	0.0	1.65	1.99
					$\sigma_{21} = -0.5$					
			$\theta_{21} = .8$					$\theta_{21} = .4$		
OLS	-2.78	-2.33	-0.33	1.57	2.0	-2.88	-2.44	-0.42	1.48	1.88
IV1	-2.48	-2.03	-0.08	1.86	2.28	-2.5	-2.06	-0.1	1.83	2.28
IV2	-2.32	-1.91	-0.01	1.89	2.27	-2.32	-1.91	-0.01	1.84	2.27
IV3	-2.03	-1.71	-0.02	1.69	2.04	-2.32	-1.91	-0.01	1.84	2.27
			$\theta_{21} = 0$					$\theta_{21} = -.8$		
OLS	-2.92	-2.46	-0.48	1.43	1.84	-2.81	-2.38	-0.42	1.47	1.93
IV1	-2.53	-2.06	-0.12	1.82	2.22	-2.52	-2.05	-0.1	1.84	2.28
IV2	-2.35	-1.92	-0.03	1.87	2.25	-2.36	-1.96	-0.03	1.85	2.24
IV3	-2.08	-1.74	-0.01	1.70	2.04	-2.06	-1.72	-0.01	1.68	2.02
					$\sigma_{21} = 0.5$					
			$\theta_{21} = .8$					$\theta_{21} = .4$		
OLS	-1.02	-0.66	1.12	3.08	3.5	-1.01	-0.64	1.11	3.1	3.56
IV1	-1.8	-1.38	0.28	2.07	2.54	-1.77	-1.4	0.28	2.11	2.53
IV2	-2.0	-1.62	0.08	1.91	2.3	-1.99	-1.62	0.07	1.93	2.33
IV3	-1.88	-1.6	0.02	1.76	2.13	-1.90	-1.58	0.02	1.76	2.14
			$\theta_{21} = 0$					$\theta_{21} = -.8$		
OLS	-1.11	-0.72	1.03	3.01	3.49	-1.79	-1.38	0.5	2.55	2.96
IV1	-1.84	-1.45	0.26	2.09	2.52	-2.17	-1.8	0.14	2.14	2.58
IV2	-2.03	-1.65	0.09	1.91	2.34	-2.17	-1.82	0.03	1.98	2.37
IV3	-1.87	-1.59	0.03	1.75	2.1	-1.99	-1.66	0.04	1.75	2.09

5.2. Large sample properties: 500 observations. To control for the effects of the autocovariance structure, and the implied Newey-West corrections, we first focus on the effect of endogeneity alone. We report the results of the estimators when the data-generating process contains no autocorrelation in the errors induced by moving average processes.

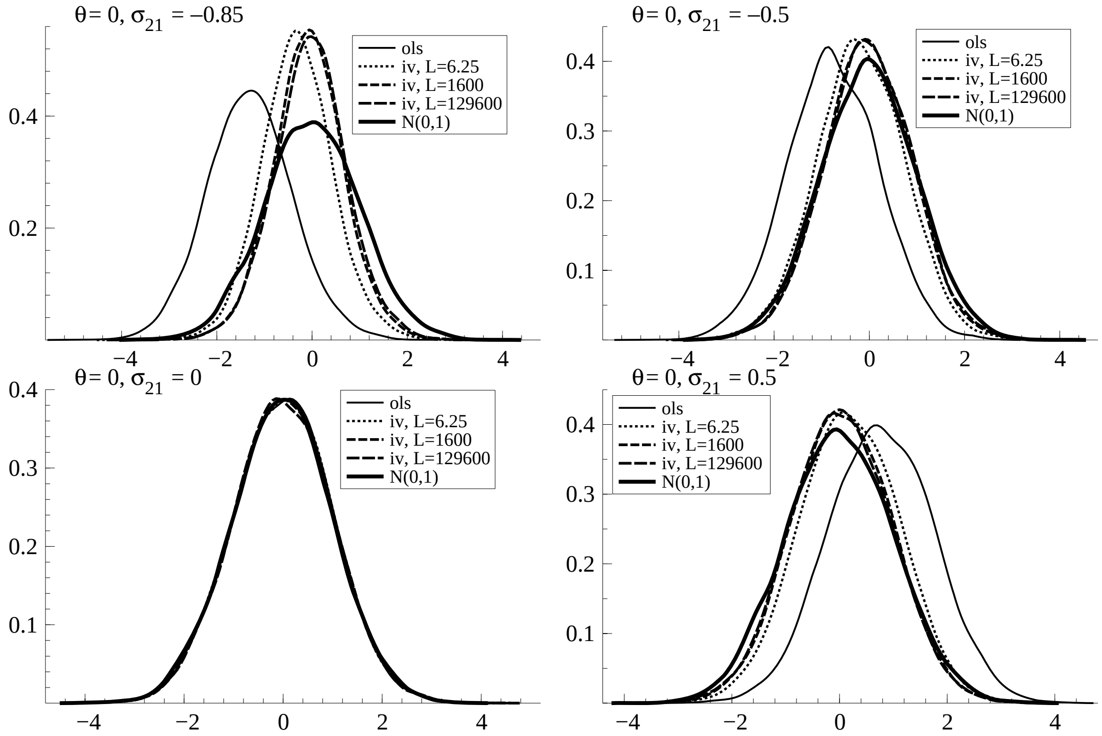

Figure 4: Distribution of t-ratios with $\theta = \mathbf{0}$, using 500 observations.

Without moving average errors: $\theta = \mathbf{0}$. As shown in figure 4, the results reflect the asymptotic theory very clearly. Without any endogeneity, $\sigma_{21} = 0$, normality is reproduced with all estimators. In particular the case of normality of OLS, as well as IV, is clearly shown in the lower left panel of figure 4. In most cases the IV-estimators provided a very good correction approximation to gaussian inference, with the correction in the right direction in all cases. Except for the case of negative covariance $\sigma_{21} = -0.85$, the simulated distributions are practically indistinguishable from the standard normal. The percentiles are reported in table 3 and summarize the impressions of figure 4. The case of strict exogeneity, $\sigma_{21} = 0$, gives normality for all estimators—as they should. With endogeneity, the IV-estimators gives good corrections with reasonable approximations to normality, the approximations generally improving with the smoothness of the instrument.

Table 2: Percentiles of t-ratios using 500 observations and no moving average processes, so $\theta = 0$. IV1 is using $\lambda = 6.25$; IV2 is using $\lambda = 1600$; IV3 is using $\lambda = 129600$;

	2.5%	5%	50%	95%	97.5%	2.5%	5%	50%	95%	97.5%
			$\sigma_{21} = -.85$					$\sigma_{21} = -.5$		
OLS	-3.02	-2.76	-1.31	0.19	0.51	-2.66	-2.34	-0.78	0.82	1.12
IV1	-1.76	-1.53	-0.32	0.95	1.24	-2.00	-1.70	-0.19	1.32	1.62
IV2	-1.55	-1.29	-0.08	1.16	1.42	-1.87	-1.56	-0.05	1.46	1.77
IV3	-1.53	-1.29	-0.03	1.23	1.51	-1.85	-1.54	-0.02	1.48	1.82
			$\sigma_{21} = 0$					$\sigma_{21} = .5$		
OLS	-1.96	-1.66	0.01	1.62	1.93	-1.11	-0.80	0.77	2.36	2.64
IV1	-1.96	-1.66	0.01	1.62	1.95	-1.64	-1.34	0.17	1.68	1.96
IV2	-1.97	-1.67	0.00	1.62	1.94	-1.77	-1.46	0.04	1.55	1.85
IV3	-1.96	-1.66	-0.01	1.64	1.98	-1.79	-1.50	0.02	1.55	1.83

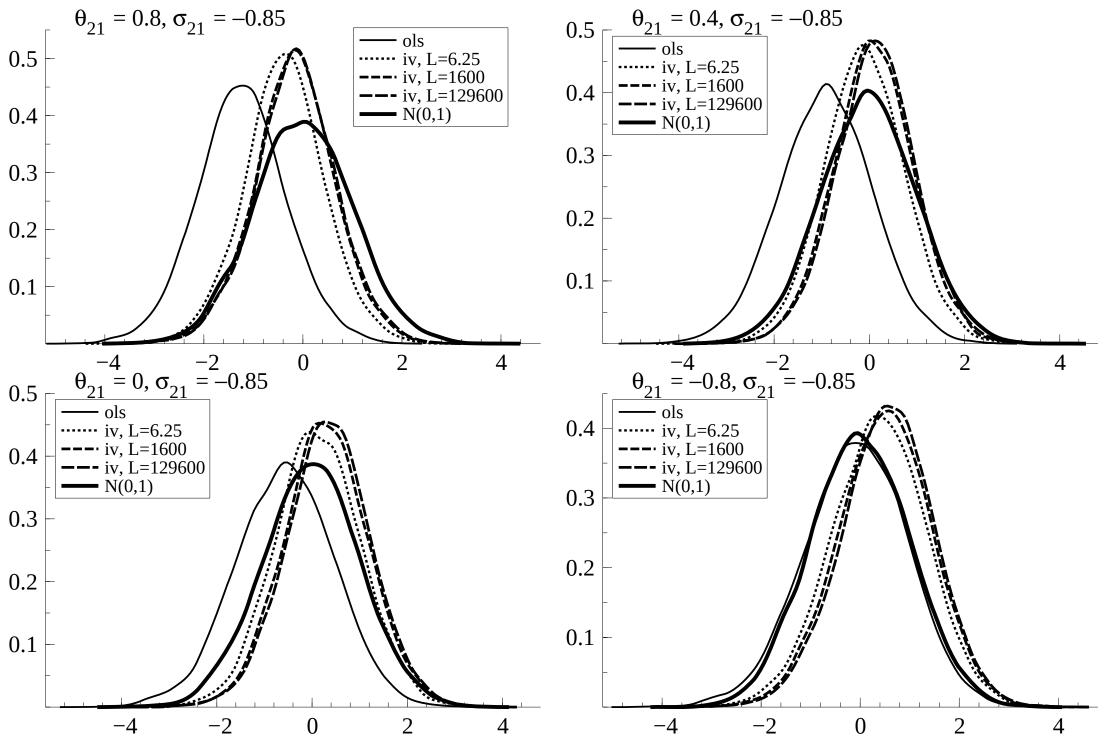

Figure 5: Distribution of t-ratios with $\sigma_{21} = -0.85$, using 500 observations.

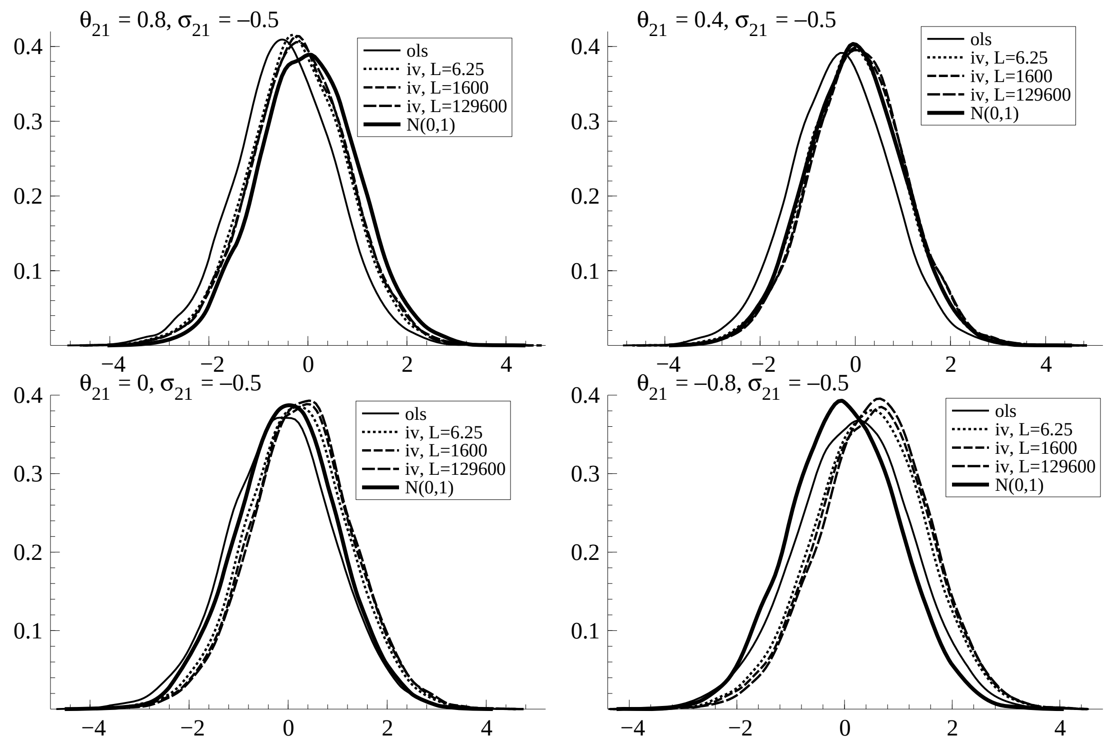

Figure 6: Distribution of t-ratios with $\sigma_{21} = -0.5$, using 500 observations.

5.3. With moving average errors: $\theta \neq 0$. In the presence of moving average errors, the IV-estimators are still always correcting in the right direction, except in the case of negative covariance and negative moving average processes $\sigma_{21} = \{-0.85, -0.5\}$ and $\theta_{21} = -0.8$. In these cases the OLS-estimator produces little bias in the t-ratios,while the IV-estimators produce negatively biased inference—as shown in the lower right panels of figures 5–6 and in table 3. The general impression, however is that with a bigger sample, the effects of moving average processes are more pronounced—as conjectured earlier.[2]

[2] Preliminary experiments, however, shows that other estimators like fully modified Phillips Hansen, Phillips & Loretan and Johansen produced similar biases.

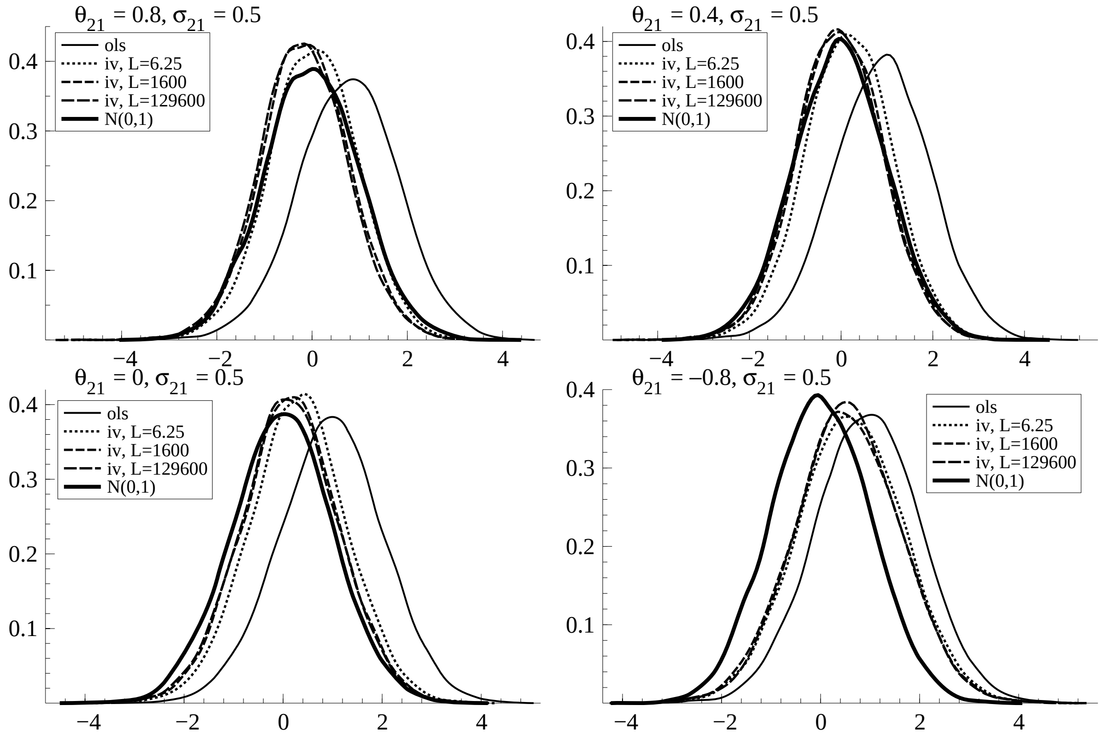

Figure 7: Distribution of t-ratios with $\sigma_{21} = 0.5$, using 500 observations.

Table 3: Percentiles of t-ratios using 500 observations. IV1 is using $\lambda = 6.25$; IV2 is using $\lambda = 1600$; IV3 is using $\lambda = 129600$;

	2.5%	5%	50%	95%	97.5%	2.5%	5%	50%	95%	97.5%
$\sigma_{21} = -0.85$										
	$\theta_{21} = .8$					$\theta = .4$				
OLS	-2.99	-2.69	-1.24	0.24	0.56	-2.87	-2.53	-0.91	0.70	1.04
*IV*1	-2.02	-1.73	-0.37	0.96	1.26	-1.80	-1.51	-0.11	1.30	1.62
*IV*2	-1.82	-1.54	0.17	1.17	1.47	-1.59	-1.31	0.08	1.48	1.79
*IV*3	-1.80	-1.55	-0.14	1.23	1.53	-1.57	-1.27	0.14	1.55	1.82
	$\theta_{21} = 0$					$\theta_{21} = -.8$				
OLS	-2.66	-2.26	-0.56	1.08	1.39	-2.18	-1.82	-0.07	1.55	1.87
*IV*1	-1.65	-1.34	0.11	1.61	1.89	-1.55	-1.22	0.39	1.92	2.20
*IV*2	-1.43	-1.15	0.29	1.73	2.05	-1.37	-1.09	0.51	2.00	2.27
*IV*3	-1.41	-1.10	0.35	1.80	2.09	-1.30	-1.02	0.57	2.04	2.32
$\sigma_{21} = -0.5$										
	$\theta_{21} = .8$					$\theta_{21} = .4$				
OLS	-2.08	2.44	–0.46	1.17	1.50	-2.36	-2.00	-0.30	1.39	1.72
IV1	-2.20	-1.84	-0.23	1.40	1.72	-2.03	-1.67	-0.02	1.66	1.96
IV2	-2.13	-1.80	-0.16	1.46	1.78	-1.95	-1.60	0.05	1.74	2.04
IV3	-2.10	-1.82	-0.16	1.45	1.80	-1.91	-1.60	0.07	1.76	2.05
	$\theta_{21} = 0$					$\theta_{21} = -.8$				
OLS	-2.28	-1.90	-0.13	1.61	1.95	-1.98	-1.62	0.19	1.90	2.22
IV1	-1.91	-1.53	0.17	1.87	2.20	-1.66	-1.28	0.47	2.12	2.44
IV2	-1.80	-1.44	0.26	1.95	2.28	-1.57	-1.18	0.55	2.18	2.48
IV3	-1.75	-1.39	0.29	1.96	2.29	-1.44	-1.10	0.59	2.19	2.50
$\sigma_{21} = 0.5$										
	$\theta_{21} = .8$					$\theta_{21} = .4$				
OLS	-1.32	-0.97	0.78	2.46	2.80	-1.21	-0.86	0.90	2.60	2.91
IV1	-1.83	-1.50	0.04	1.55	1.86	-1.71	-1.41	0.17	1.71	2.00
IV2	-1.95	-1.64	-0.14	1.39	1.66	-1.85	-1.55	-0.02	1.56	1.83
IV3	-2.02	-1.69	-0.19	1.34	1.66	-1.90	-1.59	-0.05	1.52	1.82
	$\theta_{21} = 0$					$\theta_{21} = -.8$				
OLS	-1.12	-0.77	0.98	2.68	2.99	-1.12	-0.81	0.96	2.67	3.02
IV1	-1.62	-1.31	0.32	1.92	2.21	-1.51	-1.16	0.60	2.37	2.69
IV2	-1.77	-1.44	0.16	1.78	2.08	-1.57	-1.22	0.52	2.29	2.62
IV3	-1.82	-1.49	0.13	1.77	2.05	-1.49	-1.17	0.55	2.28	2.61

6. Conclusions

In static single equation cointegration regression models the OLS estimator will have a non-standard distribution unless regressors are strictly exogenous. In the literature a number of estimators have been suggested to deal with this problem especially by the use of semi-nonparametric estimators. In the present paper we suggest an IV estimator where the Hodrick-Prescott filtered trends are used as instruments for the regressors. As a consequence a limiting gaussian distribution characterizes the estimators. Simulations are used to examine the properties of the estimator in finite samples. The results so far are very encouraging, both in small and larger samples. Suggestions for future research are to investigate instruments based on other trend cycle decompositions as well as to investigate dynamic representations of the estimator.

REFERENCES

Banerjee, A., J. J. Dolado, J. W. Galbraith and D. F. Hendry (1993). *Co-integration, Error Correction and the Econometric Analysis of Non-stationary Data*. Oxford University Press, Oxford.

Banerjee, A., J. J. Dolado, D. F. Hendry and G. W. Smith (1986). Exploring Equilibrium Relationships in Econometrics through Static Models: some Monte Carlo Evidence. *Oxford Bulletin of Economics and Statistics*, *48*, 253–277.

Davidson, J. (2000). *Econometric Theory*. Blackwell, Oxford.

Engle, R. F. and C. W. J. Granger (1987). Co-integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, *55*, 251–276.

Gonzalo, J. (1994). Five Alternative Methods of Estimating Long Run Equilibrium Relationships. *Journal of Econometrics*, *60*, 1–31.

Hansen, B. E. and P. C. B. Phillips (1990). Estimation and Inference in Models of Cointegration: A Simulation Study. In *Advances in Econometrics*, vol. 8, 225–48. JAI Press Inc.

Harvey, A. (1989). *Forecasting. Structural Time Series and the Kalman Filter*. Cambridge University Press, Cambridge.

Haug, A. (1996). Tests for Cointegration: A Monte Carlo Comparison. *Journal of Econometrics*, *71*, 81–115.

Johansen, S. (1988). Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control*, *12*, 231–254.

Newey, W. K. and K. D. West (1987). A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, *55*, 703–8.

Park, J. Y. (1992). Canonical Cointegrating Regressions. *Econometrica*, *60*, 119–43.

Phillips, P. C. B. (2006). Optimal estimation of cointegrated systems with irrelevant instruments. Discussion Paper 1547, Cowles Foundation for Research in Economics, Yale University.

Phillips, P. C. B. and S. Durlauf (1986). Multiple Time Series Regression with Integrated Processes. *Review of Economic Studies*, *53*, 473–95.

Phillips, P. C. B. and B. E. Hansen (1990). Statistical Inference in Instrumental Variables Regression with I(1) Processes. *The Review of Economic Studies*, *57*, 99–125.

Phillips, P. C. B. and M. Loretan (1991). Estimating Long Run Economic Equilibria. *Review of Economic Studies*, *58*, 407–436.

Saikkonen, P. (1991). Asymptotically Efficient Estimation of Cointegrating Regressions. *Econometric Theory*, *7*, 1–21.

Stock, J. and M. Watson (1988). Testing for Common Trends. *Journal of the American Statistical Association*, *83*, 1097–1107.

Stock, J. and M. W. Watson (1993). A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems. *Econometrica*, *61*, 783–820.

Stock, J. H. (1987). Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors. *Econometrica*, *55*, 1035–1056.