Department of Economics

A Shrinkage Instrumental Variable Estimator for Large Datasets

Andrea Carriero, George Kapetanios and Massimiliano Marcellino

Working Paper No. 626 March 2008 ISSN 1473-0278

A Shrinkage Instrumental Variable Estimator for Large Datasets

A. Carriero
Queen Mary, University of London

G. Kapetanios
Queen Mary, University of London

M. Marcellino
Bocconi University and EUI

March 14, 2008

Abstract

This paper proposes and discusses an instrumental variable estimator that can be of particular relevance when many instruments are available. Intuition and recent work (see, e.g., Hahn (2002)) suggest that parsimonious devices used in the construction of the final instruments, may provide effective estimation strategies. Shrinkage is a well known approach that promotes parsimony. We consider a new shrinkage 2SLS estimator. We derive a consistency result for this estimator under general conditions, and via Monte Carlo simulation show that this estimator has good potential for inference in small samples.

Keywords: Instrumental Variable Estimation, 2SLS, Shrinkage, Bayesian Regression.

JEL classification: C13, C23, C51.

1 Introduction

Recent work in instrumental variable estimation has considered two distinct routes. The first is one where instrumental variables are only weakly correlated with the endogenous explanatory variables of an instrumental variables (IV) regression. Work by, e.g., Phillips (1983), Rothenberg (1984), Stock and Yogo (2003b) and Chao and Swanson (2005) consider a natural measure of instrument weakness (or strength) in a linear IV framework to be the so-called concentration parameter. In standard analysis the concentration parameter is taken to grow at the rate of the sample size whereas in the case of weak instruments this parameter grows more slowly or in the extreme case introduced and considered by Staiger and Stock (1997) it remains finite asymptotically. In the case of weak instruments, the properties of IV estimators such two stage least squares (2SLS) and limited information maximum likelihood

(LIML) are affected relative to the case of strong instruments and the estimators may, in fact, be inconsistent.

Another direction in IV research involves the case where the number of available instruments is large. This approach was first taken by Morimune (1983) and later generalized by Bekker (1994). Other relevant papers include Donald and Newey (2001), Hahn, Hausman, and Kuersteiner (2001), Hahn (2002), and Chao and Swanson (2004). More recently, the two different stands have been combined to provide a comprehensive framework for the analysis of the properties of IV estimators in the case of many weak instruments. Work on this includes Hansen, Hausman, and Newey (2006), Stock and Yogo (2003a), Newey (2004) and Chao and Swanson (2005). The last paper is closest to the spirit of the analysis of the current paper. A clear conclusion from this work suggests that inconsistency of IV estimators is a probable outcome when many weak instruments are used.

With this in mind, a further recent development focuses on considering parsimonious modeling assumptions for the set of instruments to avoid IV estimator inconsistency. In particular, Kapetanios and Marcellino (2006) and Bai and Ng (2006) suggest that imposing a factor structure on the set of instruments, extracting estimates of these factors and using them as instruments can be very useful. Of course, an issue with this approach is the need to assume a factor structure, albeit a possibly weak one, as discussed in detail in Kapetanios and Marcellino (2006). Simulation evidence suggests that if no factor structure exists then assuming one is problematic for IV estimation as one would expect. Another approach similar to Kapetanios and Marcellino (2006) but designed to parsimoniously summarise large sets of instruments in the complete absence of a factor structure is suggested by Kapetanios and Marcellino (2007). The basic idea is that a finite number of cross-sectional weighted averages of the available instruments can, under certain conditions, be valid instruments themselves.

The current paper provides a new approach that is valid for IV estimation more generally than those discussed in the previous paragraph, but shares a common characteristic of parsimony. In particular, we suggest that a shrinkage estimator be considered in the first stage of IV regression to construct appropriate instruments which can then be used in a standard way to estimate the parameters of the structural equation. Shrinkage promotes parsimony in the first stage of estimation. There is a reasonably strong case for parsimony to be made for IV estimation. Hahn (2002) provides grounds for parsimony in terms of optimal inference when many instruments are available. We provide a theoretical justification for the validity of the new shrinkage IV estimator. Further, we carry out a Monte Carlo study which provides clear evidence in favor of the new estimator compared with existing estimators.

The paper is structured as follows: Section 2 presents the theoretical results. Section 3

reports results of the Monte Carlo study. Finally, Section 4 concludes. Proofs are relegated to an Appendix.

2 Theoretical Results

The model is given by

$$y_{1n} = Y_{2n}\beta + u_n \tag{1}$$

$$Y_{2n} = Z_n\Pi_n + V_n \tag{2}$$

where y_{1n} and Y_{2n} are respectively an $n \times 1$ vector and an $n \times G$ matrix of observations on the $G+1$ endogenous variables of the system, Z_n is an $n \times K_n$ matrix of observations on the K_n instrumental variables, and $u_n = (u_1, ..., u_i, ..., u_n)'$ and $V_n = (v_1, ..., v_i, ..., v_n)'$ are, respectively, an $n \times 1$ vector and an $n \times G$ matrix of random disturbances.

We propose a two stage shrinkage estimator for β obtained as follows: In the first stage, we obtain instruments by using a standard shrinkage estimator to estimate Π_n in (2). Then, we use these instruments in a standard fashion to obtain a second stage estimator for β. For simplicity we use the following shrinkage estimator:

$$\hat{\Pi}_n = (Z_n'Z_n + s_nI)^{-1}Z_n'Y_{2n}$$

Then, straightforwardly, the two stage estimator is given by

$$\hat{\beta}_{SKG} = \left(Y_{2n}'Z_n(Z_n'Z_n + s_nI)^{-1}Z_n'Y_{2n}\right)^{-1} Y_{2n}'Z_n(Z_n'Z_n + s_nI)^{-1}Z_n'y_{1n} \tag{3}$$

We refer to this estimator as the 2SLS Shrinkage ($2SLSS$) estimator. This estimator becomes of interest if the shrinkage parameter s_n becomes large enough to promote parsimony asymptotically. As we will see, for this it is required that $n/s_n = o(1)$. We make the following assumptions.

Assumption 1 *(i) $K_n \to \infty$ as $n \to \infty$ such that $K_n/n \to \tau$, $0 \leq \tau \leq C < \infty$. (ii) $\forall n$, $Z_n'Z_n + s_nI$ has full rank. (iii) There exist two nondecreasing sequences of real numbers, r_n and s_n, such that as $n \to \infty$ $r_n/n \to \mu$ for some nonnegative constant μ, $n/s_n = o(1)$ and $s_n/nK_n = o(1)$, and such that*

$$\frac{q_n\Pi_n'Z_n'Z_n(Z_n'Z_n + s_nI)^{-1}Z_n'Z_n\Pi_n}{r_n} \to \Psi \tag{4}$$

where $q_n = s_n/n$, almost surely for some positive definite matrix Ψ and

$$\frac{q_n\Pi_n'Z_n'Z_n(Z_n'Z_n + s_nI)^{-1}Z_n'Z_n(Z_n'Z_n + s_nI)^{-1}Z_nZ_n\Pi_n}{r_n} \to 0$$

almost surely. (iv) The eigenvalues of $Z_n'Z_n/n$ are bounded away from zero and infinity for all n.

Assumption 2 *(i) Z_n and $\eta_i = (u_i, v_i')'$ are independent for all i, n, (ii) $\eta_i \sim i.i.d.(0, \Sigma)$, where $\Sigma = \begin{pmatrix} \sigma_{uu} & \sigma_{Vu}' \\ \sigma_{Vu} & \Sigma_{VV} \end{pmatrix}$, (iii) η_i has finite fourth moments.*

Given the above, we have the following Theorem

Theorem 1 *Let $P_{Z_n}^{s_n} = Z_n(Z_n'Z_n + s_nI)^{-1}Z_n'$. Let $q_n = s_n/n$ such that $q_n \to \infty$ and $\frac{K_n}{q_n} \to \infty$ and $\frac{r_n}{q_n} \to \infty$. Let the shrinkage estimator be given by*

$$\hat{\beta}_{SKG} = \left(Y_{2n}'P_{Z_n}^{s_n}Y_{2n}\right)^{-1}\left(Y_{2n}'P_{Z_n}^{s_n}y_{1n}\right)$$

Assume that $\frac{K_n}{r_n} \to 0$. Then, $\hat{\beta}_{SKG}$ is consistent for β_0.

Remark 1 *Some comments on the assumptions are in order. In particular Assumption 1(iii) is worthy of comment. The first part of assumption 1(iii) is the counterpart of the assumption relating to the concentration parameter made usually in the literature concerning the 2SLS and other IV estimators. Note that there is no need for the sequence r_n satisfying Assumption 1(iii) for the 2SLSS estimator to be the same or have the same order of magnitude as that required for the 2SLS estimator.*

Remark 2 *The importance of parsimony for IV estimation has been pointed out by Hahn (2002) who conjectured that a 2SLS estimator using a small subset of available instruments, when the number of available instruments is large, may be optimal. We view our shrinkage estimator in the same spirit as the estimator suggested by Hahn (2002). It is important to note Condition 1 of Hahn (2002) which requires that the fit of a parsimonious estimator be comparable to that of the 2SLS estimator using all instruments. In this sense it is reasonable to expect that the fit of the shrinkage estimator may, under certain conditions relating to the structure of Π_n, be close to that of the 2SLS estimator using all instruments, thereby implying that the r_n sequence for the 2SLSS estimator be of a larger order of magnitude than the analogous sequence for the 2SLS estimator. However, it is difficult to envisage specific conditions for Π_n that ensure this is the case.*

Remark 3 *We have chosen to focus on the simplest shrinkage estimator on the grounds of simplifying the asymptotic analysis. This estimator shrinks, in a uniform way, the parameter estimates towards zero. It may in fact be more appropriate to shrink towards a nonzero constant or vary the degree of shrinkage depending on the instrument. For such shrinkage*

estimators the condition (4) would have a different form and therefore it is entirely possible that such estimator will have different and possibly superior consistency properties, depending of course on the data generating process for z_t. We leave theoretical investigation of this possibility to future work mainly because there are many possibilities for the shrinkage setup. However, in the Monte Carlo section we consider uniform shrinkage to a non-zero constant and obtain interesting results.

3 Monte Carlo Evidence

In this section we provide a Monte Carlo study of the 2SLS Shrinkage ($2SLSS$) estimator and its relative performance compared to the traditional $2SLS$ estimator, given by $(Y_{2n}'Z_n(Z_n'Z_n)^{-1}Z_n'Y_{2n})^{-1}Y_{2n}'Z_n(Z_n'Z_n)^{-1}Z_n'y_{1n}$, and the bias corrected Nagar's B2SLS estimator given by $\left(Y_{2n}'Z_n(Z_n'Z_n)^{-1}Z_n'Y_{2n}+\frac{K_n-2}{n}Y_{2n}'Y_{2n}\right)^{-1}\left(Y_{2n}'Z_n(Z_n'Z_n)^{-1}Z_n'y_{1n}+\frac{K_n-2}{n}Y_{2n}'y_{1n}\right)$. The basic setup of the Monte Carlo experiments is:

$$y_i = x_i + \epsilon_i,\ i = 1, ..., n \tag{5}$$

$$z_{ij} = e_{ij}, \quad j = 1, ..., K_n,\ i = 1, ..., n \tag{6}$$

$$x_i = \sum_{j=1}^{K_n} K_n^{-1/2}(1+\alpha_j)z_{ij} + u_i, \tag{7}$$

where $e_{ij} \sim i.i.d.N(0,1)$, $cov(e_{il}, e_{sj}) = 0$ for $i \neq s$ or $l \neq j$, d is either 1 or 2 with probability 0.5, $\alpha_j \sim N(0, c^2)$ with $c = 0.1, 0.5, 1$. Let $\kappa_i = (\epsilon_i, u_i)'$. Then, $\kappa_i = P\eta_i$, where $\eta_i = (\eta_{1,i}, \eta_{2,i})'$, $\eta_{j,i} \sim i.i.d.N(0,1)$ and $P = [p_{ij}]$, $p_{ij} \sim i.i.d.N(0,1)$. The errors e_{ij} and u_s are independent for each i and s.

The $2SLSS$ estimator is computed for a grid of values of the tightness parameter s_n. In particular we use the grid $s_n = 0, 10, 10^3, 10^5$. For $s_n = 0$ the $2SLSS$ and $2SLS$ are equivalent, therefore we do not report results for this case. Higher values of s_n correspond to stronger shrinkage. We consider two different shrinkage setups: one where we shrink towards zero and one where we shrink towards $1/\sqrt{K_n}$. The latter corresponds to the actual setup of the Monte Carlo data generation process. However, we have also considered shrinking towards $1/K_n$ with very similar results[1] to those for $1/\sqrt{K_n}$ giving us some comfort that the actual choice of the non-zero constant may not be crucial. In all cases the $2SLS$, $B2SLS$ and $2SLSS$ estimators have negligible biases and we therefore concentrate on their variances.

Results are reported in Tables 1-2. Both tables are organized so that on the rows are reported results for different numbers of observations n while on the columns are displayed

[1]These results are not reported but are available upon request.

results for different proportions of the number of instruments to the number of observations, i.e. K_n/n. Both tables are divided in subpanels providing results for the three cases $c = 0.1, 0.5, 1$.

Table 1 shows results on the $2SLS$ and the $B2SLS$ estimator. In particular, the panels on the left hand side of the table display the variances of the $2SLS$ estimator, while the panels on the right hand side display the ratio $VAR(B2SLS)/VAR(2SLS)$. Therefore, in the subpanels on the right, a figure smaller than 1 signals that the $B2SLS$ is more efficient than $2SLS$. The $B2SLS$ substantially improves on the traditional $2SLS$ in all the cases in which $K_n < n$ (with large n) while in the case $n = K_n$ the two estimators are by construction equivalent[2].

Table 2 shows the ratios $VAR(2SLSS)/VAR(2SLS)$, therefore a figure below 1 signals that the variance of the $2SLSS$ estimator is smaller than that of the traditional $2SLS$. The panels on the left report the results for the case in which the shrinking is towards $q = 0$ while the panels on the right report results for the case in which the shrinking is towards $q = 1/\sqrt{K_n}$. Two main results are apparent. First, the $2SLSS$ features a systematically smaller variance than both $2SLS$ and, to a smaller extent, B2SLS. Second, when both n and K_n are large the variances of $2SLSS$ with prior mean $q = 1/\sqrt{K_n}$ become remarkably small.

Finally we focus on the case $K_n > n$. As for this case the competitor estimators are not implementable, we provide results only for the $2SLSS$. Table 3 displays the variances of the $2SLSS$ estimator in the cases $K_n/n = 1$ and $K_n/n = 1.1$, as well as their ratio. The variance ratios are systematically around 1, showing that the $2SLSS$ estimator can handle the $K_n > n$ case as efficiently as the case $n = K_n$.

These results confirm our theoretical findings and, further, show that using shrinkage in the first stage may significantly improve the small sample efficiency of the estimator. Our results for the case $q = 1/\sqrt{K_n}$ suggest that shrinking the coefficients towards an appropriate direction might improve the results even further, possibly indicating that the consistency properties of this shrinkage estimator are different to those of the simple one analysed theoretically in the previous section. As we noted in Remark 3 this is entirely possible since the relevant consistency condition (4) will be different for the two shrinkage estimators. This is a topic of interest for future work.

[2]The two estimators are also equivalent to OLS when $n = K_n$. The two equivalences are obvious once one notes that for $n = K_n$, $Y_{2n}'Z_n(Z_n'Z_n)^{-1}Z_n'Y_{2n} = Y_{2n}'Y_{2n}$ and $Y_{2n}'Z_n(Z_n'Z_n)^{-1}Z_n'y_{1n} = Y_{2n}'y_{1n}$.

4 Conclusion

Estimation of structural equations using instrumental variable techniques, in the presence of a large number of, possibly weak, instruments, is a topic that has received a lot of attention in the literature. Most work has focused on the properties of existing estimators in the case of many, possibly weak, instruments. These estimators include the 2SLS estimator and the LIML estimator.

This paper is part of a small literature that discusses estimators that can be of particular relevance when many instruments are available. Intuition and recent work (see, e.g., Hahn (2002)) suggests that parsimonious devices used in the construction of the final instruments, may provide effective estimation strategies. Shrinkage is a well known approach that promotes parsimony. We consider a new shrinkage 2SLS estimator. We derive a consistency result for this estimator under general conditions, and via Monte Carlo simulations show that it has also good potential for inference in small samples.

An open and interesting question for future research relates to the choice of the shrinkage parameter, s_n. It is of interest to develop a data-dependent way of determining this. An interesting possibility is to derive approximations of the MSE of the 2SLS shrinkage estimator and optimise the choice of s_n with respect to this measure, in the spirit of Donald and Newey (2001). We consider such an investigation to be the next step in our research agenda on this topic.

Appendix

Lemmas

Lemma 1 *Let assumptions 1-2 hold. Let* $P_{Z_n}^{s_n} = Z_n(Z_n'Z_n + s_nI)^{-1}Z_n'$. *Define* $q_n = s_n/n$ *such that* $q_n \to \infty$, $\frac{K_n}{q_n} \to \infty$ *and* $\frac{r_n}{q_n} \to \infty$. *Then, (i) for some constant* C,

$$V_n'P_{Z_n}^{s_n}V_n \Big/ \left(\frac{CK_n}{q_n}\right) = \Sigma_{VV} + O_p\left(\sqrt{\frac{q_n}{K_n}}\right),$$

(ii) for some contant C,

$$V_n'P_{Z_n}^{s_n}u_n \Big/ \left(\frac{CK_n}{q_n}\right) = \sigma_{Vu} + O_p\left(\sqrt{\frac{q_n}{K_n}}\right),$$

(iii)

$$V_n'P_{Z_n}^{s_n}Z_n\Pi_n \Big/ \left(\frac{r_n}{q_n}\right) = O_p\left(\sqrt{\frac{q_n}{r_n}}\right)$$

and (iv)

$$u_n'P_{Z_n}^{s_n}Z_n\Pi_n \Big/ \left(\frac{r_n}{q_n}\right) = O_p\left(\sqrt{\frac{q_n}{r_n}}\right).$$

Proof. C denotes constants which may be differ across contexts. To prove (i) it is sufficient to prove the statement for the g,h-th element of $V_n'P_{Z_n}^{s_n}V_n$ denoted by $V_{gn}'P_{Z_n}^{s_n}V_{hn}$ where V_{gn} denotes the g-th column of V_n. It is sufficient to show that

$$E\left(V_{gn}'P_{Z_n}^{s_n}V_{hn} \Big/ \left(\frac{CK_n}{q_n}\right) - \Sigma_{VV}^{gh}\right)^2 = O_p\left(\frac{q_n}{K_n}\right)$$

where Σ_{VV}^{gh} denotes the g,h-th element of Σ_{VV}. So, denoting the (i,j)-th element of $P_{Z_n}^{s_n}$ by $p_{ij,n}^{s_n}$, we have

$$E\left(V_{gn}'P_{Z_n}^{s_n}V_{hn} \Big/ \left(\frac{CK_n}{q_n}\right) - \Sigma_{VV}^{gh}\right)^2 =$$

$$\left(\frac{q_n}{CK_n}\right)^2 \sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} E\left(p_{ij,n}^{s_n}p_{kl,n}^{s_n}\right) E\left(v_{ig}v_{jh}v_{kg}v_{lh}\right) -$$

$$\frac{2q_n\Sigma_{VV}^{gh}}{CK_n} \sum_{i=1}^{n}\sum_{j=1}^{n} E\left(p_{ij,n}^{s_n}\right) E\left(v_{ig}v_{jh}\right) + \left(\Sigma_{VV}^{gh}\right)^2 =$$

$$\left(\frac{q_n}{CK_n}\right)^2 E\left(v_{ig}^2v_{jh}^2\right)\left[\sum_{i=1}^{n} E\left(\left(p_{ii,n}^{s_n}\right)^2\right)\right] + \left(\frac{\sqrt{2}q_n}{CK_n}\right)^2 \Sigma_{VV}^{gg}\Sigma_{VV}^{hh}\left[\sum_{i=2}^{n}\sum_{j=1}^{i-1} E\left(\left(p_{ij,n}^{s_n}\right)^2\right)\right] +$$

$$\left\{\left(\frac{\sqrt{2}\Sigma_{VV}^{gh}q_n}{CK_n}\right)^2 \left[\sum_{i=2}^{n}\sum_{j=1}^{i-1} E\left(p_{ii,n}^{s_n}p_{jj,n}^{s_n} + \left(p_{ii,n}^{s_n}\right)^2\right)\right] -\right.$$

$$\left.\frac{2\left(\Sigma_{VV}^{gh}\right)^2 q_n}{CK_n}\sum_{i=1}^{n}E\left(p_{ii,n}^{s_n}\right)+\left(\Sigma_{VV}^{gh}\right)^2\right\}=A_{1n}+A_{2n}+A_{3n}$$

We examine each of A_{1n}, A_{2n} and A_{3n} in turn. Starting with A_{1n} we have that

$$\left(\frac{q_n}{CK_n}\right)^2 E\left(v_{ig}^2 v_{jh}^2\right)\left[\sum_{i=1}^{n}E\left(\left(p_{ii,n}^{s_n}\right)^2\right)\right]\leq$$

$$\left(\frac{q_n}{CK_n}\right)^2\sqrt{E\left(v_{ig}^4\right)}\sqrt{E\left(v_{jh}^4\right)}\left[\sum_{i=1}^{n}E\left(\left(p_{ii,n}^{s_n}\right)^2\right)\right]\leq$$

$$\left(\frac{q_n}{CK_n}\right)\sqrt{E\left(v_{ig}^4\right)}\sqrt{E\left(v_{jh}^4\right)}=O\left(\frac{q_n}{K_n}\right).$$

The second inequality follows from the fact that

$$\sum_{i=1}^{n}E\left(\left(p_{ii,n}^{s_n}\right)^2\right)\leq\sum_{i=1}^{n}E\left(p_{ii,n}^{s_n}\right)$$

which follows from the fact that $0\leq p_{ii,n}^{s_n}\leq 1$. This in turn follows from the fact that $0\leq p_{ii,n}\leq 1$ where $p_{ii,n}$ is the (i,i)-th element of $Z_n(Z_n'Z_n)^{-1}Z_n'$. The result then follows from Lemma 2. Next, focusing on A_{2n} we have that

$$\left(\frac{\sqrt{2}q_n}{CK_n}\right)^2\Sigma_{VV}^{gg}\Sigma_{VV}^{hh}\left[\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left(\left(p_{ij,n}^{s_n}\right)^2\right)\right]\leq$$

$$\left(\frac{q_n}{CK_n}\right)^2\Sigma_{VV}^{gg}\Sigma_{VV}^{hh}\left[\sum_{i=1}^{n}E\left(\left(p_{ii,n}^{s_n}\right)^2\right)+2\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left(\left(p_{ij,n}^{s_n}\right)^2\right)\right]$$

But

$$\left[\sum_{i=1}^{n}E\left(\left(p_{ii,n}^{s_n}\right)^2\right)+2\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left(\left(p_{ij,n}^{s_n}\right)^2\right)\right]=tr\left(E\left(P_{Z_n}^{s_n}P_{Z_n}^{s_n}\right)\right)\leq\frac{CK_n}{q_n^2}$$

by Lemma 2. So

$$\left(\frac{\sqrt{2}q_n}{CK_n}\right)^2\Sigma_{VV}^{gg}\Sigma_{VV}^{hh}\left[\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left(\left(p_{ij,n}^{s_n}\right)^2\right)\right]\leq\frac{Cq_n\Sigma_{VV}^{gg}\Sigma_{VV}^{hh}}{K_n}=O\left(\frac{q_n}{K_n}\right)$$

Fnally, we consider A_{3n}. We have that

$$|A_{3n}|=\left|\left(\frac{\sqrt{2}\Sigma_{VV}^{gh}q_n}{CK_n}\right)^2\left[\sum_{i=2}^{n}\sum_{j=1}^{i-1}E\left(p_{ii,n}^{s_n}p_{jj,n}^{s_n}+\left(p_{ii,n}^{s_n}\right)^2\right)\right]-\right.$$

$$\left.\frac{2\left(\Sigma_{VV}^{gh}\right)^2 q_n}{CK_n}\sum_{i=1}^{n}E\left(p_{ii,n}^{s_n}\right)+\left(\Sigma_{VV}^{gh}\right)^2\right|=$$

$$\left|\left(\frac{\Sigma_{VV}^{gh}q_n}{CK_n}\right)^2\left\{\left(tr\left(E\left(P_{Z_n}^{s_n}\right)\right)\right)^2+tr\left(E\left(P_{Z_n}^{s_n}P_{Z_n}^{s_n}\right)\right)-2\sum_{i=1}^{n}E\left(\left(p_{ii,n}^{s_n}\right)^2\right)\right\}-\right.$$
$$\left.\frac{2\left(\Sigma_{VV}^{gh}\right)^2q_n}{CK_n}\sum_{i=1}^{n}E\left(p_{ii,n}^{s_n}\right)+\left(\Sigma_{VV}^{gh}\right)^2\right|\leq$$
$$\left|\left(\frac{\Sigma_{VV}^{gh}q_n}{CK_n}\right)^2\left\{\left(tr\left(E\left(P_{Z_n}^{s_n}\right)\right)\right)^2-2\sum_{i=1}^{n}E\left(\left(p_{ii,n}^{s_n}\right)^2\right)\right\}-\right. \tag{8}$$
$$\left.\frac{2\left(\Sigma_{VV}^{gh}\right)^2q_n}{CK_n}\sum_{i=1}^{n}E\left(p_{ii,n}^{s_n}\right)+\left(\Sigma_{VV}^{gh}\right)^2\right|+\left|\left(\frac{\Sigma_{VV}^{gh}q_n}{CK_n}\right)^2tr\left(E\left(P_{Z_n}^{s_n}P_{Z_n}^{s_n}\right)\right)\right|$$

The second term of (8) is $O\left(\frac{1}{K_n}\right)=o\left(\frac{q_n}{K_n}\right)$ by Lemma 2. Focusing on the first term we note that by Lemma 2 there exists a constant such that $tr\left(E\left(P_{Z_n}^{s_n}\right)\right)=\frac{CK_n}{q_n}$. Thus, we have

$$\left|\left(\frac{\Sigma_{VV}^{gh}q_n}{CK_n}\right)^2\left\{\left(tr\left(E\left(P_{Z_n}^{s_n}\right)\right)\right)^2-2\sum_{i=1}^{n}E\left(\left(p_{ii,n}^{s_n}\right)^2\right)\right\}-\right.$$
$$\left.\frac{2\left(\Sigma_{VV}^{gh}\right)^2q_n}{CK_n}\sum_{i=1}^{n}E\left(p_{ii,n}^{s_n}\right)+\left(\Sigma_{VV}^{gh}\right)^2\right|\leq$$
$$\frac{\left(\Sigma_{VV}^{gh}\right)^2}{K_n}+\frac{2\left(\Sigma_{VV}^{gh}\right)^2q_n^2\sum_{i=1}^{n}E\left(\left(p_{ii,n}^{s_n}\right)\right)}{K_n^2}\leq\frac{C\left(\Sigma_{VV}^{gh}\right)^2q_n}{K_n}=O\left(\frac{q_n}{K_n}\right)$$

where the first inequality follows from Lemma 2. This concludes the proof of part (i) of Lemma 1. Part (ii) is proven similarly. Next, we move on to part (iii). We have

$$E\left(\left\|\frac{q_nV_n'P_{Z_n}^{s_n}\Pi_n}{r_n}\right\|\right)^2=$$
$$E\left(\left\|\frac{q_n^2\Pi_n'Z_n'Z_n(Z_n'Z_n+s_nI)^{-1}Z_n'V_nV_n'Z_n(Z_n'Z_n+s_nI)^{-1}Z_n'Z_n\Pi_n}{r_n^2}\right\|\right)^2=$$
$$tr\left(\Sigma_{VV}\right)E\left(tr\left[\frac{q_n^2\Pi_n'Z_n'Z_n(Z_n'Z_n+s_nI)^{-1}Z_n'Z_n(Z_n'Z_n+s_nI)^{-1}Z_n'Z_n\Pi_n}{r_n^2}\right]\right)\leq$$
$$\frac{q_n}{r_n}tr\left(\Sigma_{VV}\right)E\left(tr\left[\frac{q_n\Pi_n'Z_n'Z_n(Z_n'Z_n+s_nI)^{-1}Z_n'Z_n(Z_n'Z_n+s_nI)^{-1}Z_n'Z_n\Pi_n}{r_n}\right]\right)\leq$$
$$\frac{q_n}{r_n}tr\left(\Sigma_{VV}\right)E\left(tr\left[\frac{q_n\Pi_n'Z_n'Z_n(Z_n'Z_n+s_nI)^{-1}Z_n'Z_n\Pi_n}{r_n}\right]\right)\leq$$
$$\frac{q_nC}{r_n}=O\left(\frac{q_n}{r_n}\right)$$

where the last but one inequality follows by the second part of assumption 1(iii) and the last by the first part of assumption 1(iii). Part (iv) can be proved similarly. ■

Lemma 2 *Let* $P_{Z_n}^{s_n} = Z_n(Z_n'Z_n + s_nI)^{-1}Z_n'$ *and let assumptions 1-2 hold. Then, for all* n, $tr\left(P_{Z_n}^{s_n}\right)$, *and therefore* $E\left(tr\left(P_{Z_n}^{s_n}\right)\right)$ *are bounded from above by* $\frac{CnK_n}{s_n}$. *Further, for all* n, $tr\left(\left(P_{Z_n}^{s_n}\right)^2\right)$,*and therefore,* $E\left(tr\left(\left(P_{Z_n}^{s_n}\right)^2\right)\right)$, *are bounded from above by* $\frac{Cn^2K_n}{s_n^2}$.

Proof. We have that

$$tr\left(Z_n(Z_n'Z_n + s_nI)^{-1}Z_n'\right) = tr\left(Z_n'Z_n(Z_n'Z_n + s_nI)^{-1}\right) =$$

$$tr\left(\frac{Z_n'Z_n}{n}\left(\frac{Z_n'Z_n}{n} + \frac{s_n}{n}I\right)^{-1}\right) = tr\left(\left(I + \frac{s_n}{n}\left(\frac{Z_n'Z_n}{n}\right)^{-1}\right)^{-1}\right) \tag{9}$$

We next use standard results (see, e.g., Bai and Golub (1997)) on upper bounds of the trace of the inverse of a matrix. We first note the following: By assumption 1(iv), for all n, all eigenvalues of $\frac{Z_n'Z_n}{n}$, and therefore those of $\left(\frac{Z_n'Z_n}{n}\right)^{-1}$, are bounded and bounded away from zero. As a result all eigenvalues of $\frac{s_n}{n}\left(\frac{Z_n'Z_n}{n}\right)^{-1}$, and therefore of $I + \frac{s_n}{n}\left(\frac{Z_n'Z_n}{n}\right)^{-1}$, are $O\left(\frac{s_n}{n}\right)$. Then, we use Kantorovich's inequality for a square $m \times m$ matrix A, which states that for the i,i-th element of A^{-1} the following holds:

$$(A^{-1})_{ii} \leq \frac{1}{4a_{ii}}\left(\frac{\alpha}{\beta} + \frac{\beta}{\alpha} + 2\right) \tag{10}$$

where a_{ii} is the i,i-th element of A, $\alpha = \min_i \lambda_i(A)$, $\beta = \max_i \lambda_i(A)$ and λ_i denotes the i-th eigenvalue of A. Applying this result to our case gives

$$tr\left(\left(I + \frac{s_n}{n}\left(\frac{Z_n'Z_n}{n}\right)^{-1}\right)^{-1}\right) \leq C\frac{nK_n}{s_n}$$

noting firstly that both $\frac{\alpha}{\beta}$ and $\frac{\beta}{\alpha}$, in (10), in our case are $O(1)$ and secondly that $a_{ii} = O\left(\frac{s_n}{n}\right)$. Note that since we assume that $n = o(s_n)$, it follows that

$$\lim_{n\to\infty} E\left(tr\left(Z_n(Z_n'Z_n + s_nI)^{-1}Z_n'\right)\right) = o(K_n).$$

This proves the first part of the Lemma. Given the preceding analysis, in order to prove the second part of the Lemma, it is simply sufficient to analyse the behaviour of $tr\left(\left(I + \frac{s_n}{n}\left(\frac{Z_n'Z_n}{n}\right)^{-1}\right)^{-2}\right)$. Then, we have

$$tr\left(\left(I + \frac{s_n}{n}\left(\frac{Z_n'Z_n}{n}\right)^{-1}\right)^{-2}\right) =$$

$$tr\left(\left(\left(I + \frac{s_n}{n}\left(\frac{Z_n'Z_n}{n}\right)^{-1}\right)\left(I + \frac{s_n}{n}\left(\frac{Z_n'Z_n}{n}\right)^{-1}\right)\right)^{-1}\right) =$$

$$tr\left(\left(I+2\frac{s_n}{n}\left(\frac{Z_n'Z_n}{n}\right)^{-1}+\left(\frac{s_n}{n}\right)^2\left(\frac{Z_n'Z_n}{n}\right)^{-2}\right)^{-1}\right)=tr\left(\left(I+\tilde{\Sigma}_n\right)^{-1}\right)$$

where

$$\tilde{\Sigma}_n=2\frac{s_n}{n}\left(\frac{Z_n'Z_n}{n}\right)^{-1}+\left(\frac{s_n}{n}\right)^2\left(\frac{Z_n'Z_n}{n}\right)^{-2}$$

Since all eigenvalues of $\left(\frac{Z_n'Z_n}{n}\right)^{-1}$ are bounded, and bounded away from zero, the same follows for the eigenvalues of $\left(\frac{Z_n'Z_n}{n}\right)^{-2}$. Then, it is easy to see that all eigenvalues of $I+\tilde{\Sigma}_n$ are $O\left(\left(\frac{s_n}{n}\right)^2\right)$. Then, by a similar analysis as that used for $tr\left(\left(I+\frac{s_n}{n}\left(\frac{Z_n'Z_n}{n}\right)^{-1}\right)^{-1}\right)$ we have that

$$tr\left(\left(I+\frac{s_n}{n}\Sigma^{-1}\right)^{-2}\right)\leq C\frac{n^2K_n}{s_n^2}$$

■

Proof of Theorem 1

We have that

$$\hat{\beta}_{SKG}-\beta_0=\left(Y_{2n}'P_{Z_n}^{s_n}Y_{2n}\right)^{-1}\left(Y_{2n}'P_{Z_n}^{s_n}u_n\right)=\left(\frac{q_nY_{2n}'P_{Z_n}^{s_n}Y_{2n}}{r_n}\right)^{-1}\left(\frac{q_nY_{2n}'P_{Z_n}^{s_n}u_n}{r_n}\right) \tag{11}$$

We analyse each term of the product of the RHS of (11) in turn. We have,

$$\frac{q_nY_{2n}'P_{Z_n}^{s_n}Y_{2n}}{r_n}=\frac{q_n\Pi_n'Z_n'P_{Z_n}^{s_n}Z_n\Pi_n}{r_n}+\frac{q_nV_n'P_{Z_n}^{s_n}Z_n\Pi_n}{r_n}+$$
$$\frac{q_n\Pi_n'Z_n'P_{Z_n}^{s_n}V_n}{r_n}+\left(\frac{K_n}{r_n}\right)\frac{q_nV_n'P_{Z_n}^{s_n}V_n}{K_n}\xrightarrow{p}\Psi$$

by the fact that $\frac{K_n}{r_n}\to 0$, Assumption 1(iii) and Lemma 1(i),(iii). Next, we have that

$$\frac{q_nY_{2n}'P_{Z_n}^{s_n}u_n}{r_n}=\frac{q_n\Pi_n'Z_n'P_{Z_n}^{s_n}u_n}{r_n}+\left(\frac{K_n}{r_n}\right)\frac{q_nV_n'P_{Z_n}^{s_n}u_n}{K_n}$$

By Lemma 1 (ii), $\frac{q_nV_n'P_{Z_n}^{s_n}u_n}{K_n}=O_p(1)$ and so

$$\left(\frac{K_n}{r_n}\right)\frac{q_nV_n'P_{Z_n}^{s_n}u_n}{K_n}=O_p\left(\frac{K_n}{r_n}\right)$$

Further, by lemma 1 (iv)

$$\frac{q_n\Pi_n'Z_n'\tilde{P}_{Z_n}^{s_n}u_n}{r_n}=O_p\left(\sqrt{\frac{q_n}{r_n}}\right)$$

Overall,

$$\frac{q_nY_{2n}'\tilde{P}_{Z_n}^{s_n}u_n}{r_n}=O_p\left(\frac{K_n}{r_n}\right)+O_p\left(\sqrt{\frac{q_n}{r_n}}\right)=o_p(1)$$

under the assumptions of Theorem 1.

References

Bai, J., and S. Ng (2006): "Instrumental Variable Estimation in a Data Rich Environment," *Mimeo*.

Bai, Z., and G. H. Golub (1997): "Bounds for the Trace of the Inverse and the Determinant of Symmetric Postivive Definite Matrices," *Annals of Numerical Mathematics*, 4, 29–38.

Bekker, P. A. (1994): "Alternative Approximations to the Distributions of Instrumental Variable Estimators," *Econometrica*, 62, 657–681.

Chao, J. C., and N. R. Swanson (2004): "Asymptotic Normality of Single-Equation Estimators for the Case with a Large Number of Weak Instruments," in *Frontiers of Analysis and Applied Research: Essays in Honor of Peter C. B. Phillips*, ed. by D. Corbae, S. Durlauf, and B. Hansen. Cambridge, U.K.: Cambridge University Press.

——— (2005): "Consistent Estimation with a Large Number of Weak Instruments," *Econometrica*, 73, 1673–1692.

Donald, S. G., and W. K. Newey (2001): "Choosing the Number of Instruments," *Econometrica*, 69, 1161–1191.

Hahn, J. (2002): "Optimal Inference with Many Instruments," *Econometric Theory*, 18, 140–168.

Hahn, J., J. Hausman, and G. Kuersteiner (2001): "Accuracy of Higher Order Bias and MSE Approximations," *Econometrics Journal*, 7, 272–306.

Hansen, C., J. Hausman, and W. K. Newey (2006): "Many Instruments, Weak Instruments and Microeconometric Practice," *Working Paper, MIT*.

Kapetanios, G., and M. Marcellino (2006): "Factor-GMM Estimation with Large Sets of Possibly Weak Instruments," *Queen Mary Working Paper No. 577*.

——— (2007): "Cross-sectional Averaging and Instrumental Variable Estimation with Many Weak Instruments," *mimeo*.

Morimune, K. (1983): "Approximate Distributions of k-class Estimators when the Degree of Overidentification is Large Compared with the Sample Size," *Econometrica*, 51, 821–842.

Newey, W. K. (2004): "Many Weak Moment Asymptotics for the Continuously Updated GMM Estimator," *Working Paper, MIT*.

Phillips, P. C. B. (1983): "Exact Small Sample Theory in the Simultaneous Equations Model," in *Handbook of Econometrics, Vol. I*, ed. by Z. Griliches, and M. D. Intriligator, p. 449516. Amsterdam: North-Holland.

Rothenberg, T. (1984): "Approximating the Distributions of Econometric Estimators and Test Statistics," in *Handbook of Econometrics, Vol. II*, ed. by Z. Griliches, and M. D. Intriligator, pp. 881–935. Amsterdam: North-Holland.

Staiger, D., and J. H. Stock (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65, 557–586.

Stock, J. H., and M. Yogo (2003a): "Asymptotic Distributions of Instrumental Variables Statistics with Many Weak Instruments," in *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg*, ed. by D. W. K. Andrews, and J. H. Stock. Cambridge, U.K.: Cambridge University Press.

——— (2003b): "Testing for Weak Instruments in Linear IV Regression," in *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg*, ed. by D. W. K. Andrews, and J. H. Stock, pp. 80–108. Cambridge, U.K.: Cambridge University Press.

Table 1. Variance of the 2SLS and B2SLS estimates.

		$2SLS$					$B2SLS/2SLS$ **(RATIO)**				
	K_n/n	0.6	0.7	0.8	0.9	1	0.6	0.7	0.8	0.9	1
n				$c = 0.1$					$c = 0.1$		
50		0.214	0.229	0.246	0.263	0.281	1.446	1.636	1.924	2.180	1.000
100		0.203	0.220	0.242	0.258	0.271	0.723	0.867	1.257	1.874	1.000
200		0.200	0.220	0.239	0.253	0.270	0.498	0.555	0.704	1.385	1.000
400		0.198	0.217	0.239	0.251	0.272	0.346	0.368	0.434	0.684	1.000
				$c = 0.5$					$c = 0.5$		
50		0.193	0.206	0.223	0.239	0.256	1.187	1.280	1.565	2.138	1.000
100		0.181	0.198	0.218	0.234	0.246	0.664	0.765	0.888	1.734	1.000
200		0.177	0.196	0.214	0.228	0.244	0.474	0.505	0.594	0.902	1.000
400		0.175	0.193	0.213	0.226	0.245	0.331	0.349	0.392	0.567	1.000
				$c = 1$					$c = 1$		
50		0.147	0.162	0.175	0.188	0.199	0.862	0.918	1.142	1.589	1.000
100		0.138	0.153	0.166	0.181	0.191	0.628	0.628	0.744	0.965	1.000
200		0.132	0.148	0.162	0.176	0.191	0.449	0.453	0.495	0.623	1.000
400		0.131	0.145	0.162	0.175	0.188	0.319	0.317	0.348	0.442	1.000

Table 2. Variance Ratio of the 2SLSS vs 2SLS estimates.

		$q=0$					$q=K_n^{-0.5}$				
c	K_n/n	0.6	0.7	0.8	0.9	1	0.6	0.7	0.8	0.9	1
	n	$s_n=10$					$s_n=10$				
0.1	50	0.924	0.905	0.886	0.870	0.854	0.844	0.828	0.816	0.804	0.791
	100	0.943	0.928	0.913	0.896	0.877	0.895	0.883	0.871	0.858	0.842
	200	0.964	0.953	0.940	0.923	0.904	0.938	0.929	0.917	0.902	0.885
	400	0.979	0.971	0.961	0.946	0.928	0.965	0.959	0.949	0.935	0.917
		$s_n=10^3$					$s_n=10^3$				
	50	0.818	0.786	0.757	0.741	0.722	0.550	0.503	0.460	0.433	0.411
	100	0.784	0.751	0.731	0.716	0.699	0.393	0.362	0.331	0.314	0.306
	200	0.772	0.749	0.729	0.714	0.700	0.336	0.322	0.313	0.310	0.308
	400	0.782	0.761	0.743	0.728	0.714	0.392	0.388	0.389	0.392	0.395
		$s_n=10^5$					$s_n=10^5$				
	50	0.815	0.782	0.752	0.736	0.717	0.574	0.529	0.484	0.459	0.437
	100	0.774	0.739	0.718	0.702	0.685	0.405	0.374	0.336	0.317	0.303
	200	0.746	0.721	0.700	0.684	0.670	0.282	0.256	0.239	0.225	0.214
	400	0.733	0.710	0.690	0.674	0.661	0.201	0.182	0.167	0.157	0.143
		$s_n=10$					$s_n=10$				
0.5	50	0.922	0.901	0.881	0.863	0.845	0.853	0.835	0.819	0.805	0.790
	100	0.942	0.925	0.907	0.889	0.868	0.900	0.885	0.870	0.855	0.837
	200	0.962	0.950	0.935	0.917	0.896	0.939	0.928	0.915	0.899	0.879
	400	0.978	0.969	0.959	0.942	0.921	0.966	0.958	0.948	0.932	0.912
		$s_n=10^3$					$s_n=10^3$				
	50	0.818	0.782	0.750	0.731	0.711	0.611	0.559	0.508	0.475	0.448
	100	0.781	0.745	0.720	0.703	0.684	0.434	0.397	0.361	0.341	0.330
	200	0.763	0.736	0.715	0.697	0.682	0.366	0.348	0.336	0.331	0.326
	400	0.771	0.749	0.729	0.711	0.695	0.418	0.410	0.408	0.408	0.409
		$s_n=10^5$					$s_n=10^5$				
	50	0.815	0.779	0.745	0.726	0.705	0.651	0.601	0.546	0.513	0.488
	100	0.771	0.732	0.707	0.689	0.670	0.460	0.422	0.377	0.353	0.337
	200	0.736	0.708	0.685	0.668	0.651	0.320	0.288	0.269	0.250	0.237
	400	0.721	0.696	0.674	0.656	0.640	0.227	0.205	0.186	0.175	0.159
		$s_n=10$					$s_n=10$				
1	50	0.921	0.898	0.873	0.848	0.827	0.875	0.853	0.831	0.808	0.790
	100	0.936	0.920	0.896	0.875	0.850	0.908	0.891	0.870	0.851	0.828
	200	0.958	0.944	0.927	0.906	0.881	0.942	0.929	0.913	0.892	0.868
	400	0.976	0.966	0.952	0.934	0.908	0.967	0.958	0.945	0.926	0.901
		$s_n=10^3$					$s_n=10^3$				
	50	0.832	0.789	0.752	0.720	0.695	0.787	0.727	0.665	0.590	0.567
	100	0.770	0.744	0.705	0.683	0.656	0.530	0.482	0.445	0.411	0.393
	200	0.750	0.721	0.692	0.670	0.650	0.442	0.417	0.396	0.380	0.368
	400	0.755	0.727	0.703	0.682	0.660	0.477	0.464	0.455	0.449	0.441
		$s_n=10^5$					$s_n=10^5$				
	50	0.830	0.787	0.749	0.717	0.690	0.919	1.026	0.823	0.700	0.678
	100	0.760	0.733	0.693	0.671	0.642	0.611	0.550	0.513	0.464	0.442
	200	0.724	0.694	0.663	0.641	0.619	0.442	0.392	0.350	0.323	0.297
	400	0.703	0.672	0.647	0.625	0.603	0.301	0.271	0.247	0.228	0.211

Table 3. Variances of the 2SLSS. Case $K_n > n$

	$c = 0.1000$			$c = 0.5000$			$c = 1.0000$		
K_n/N	1	1.1	ratio	1	1.1	ratio	1	1.1	ratio
N	$q = K_n^{-0.5}; s_n = \mathbf{10}$								
50	0.220	0.225	1.022	0.201	0.205	1.021	0.159	0.162	1.018
100	0.231	0.238	1.032	0.208	0.215	1.033	0.160	0.166	1.037
200	0.242	0.248	1.026	0.217	0.223	1.027	0.167	0.172	1.029
400	0.249	0.250	1.004	0.224	0.226	1.009	0.171	0.174	1.017
	$q = K_n^{-0.5}; s_n = \mathbf{10}^3$								
50	0.115	0.112	0.976	0.114	0.111	0.975	0.113	0.109	0.972
100	0.082	0.084	1.018	0.080	0.082	1.020	0.075	0.076	1.017
200	0.084	0.087	1.037	0.080	0.083	1.034	0.071	0.073	1.031
400	0.107	0.111	1.036	0.100	0.104	1.039	0.084	0.087	1.041
	$q = K_n^{-0.5}; s_n = \mathbf{10}^5$								
50	0.122	0.119	0.977	0.124	0.121	0.970	0.137	0.129	0.942
100	0.081	0.082	1.009	0.082	0.083	1.010	0.084	0.085	1.007
200	0.059	0.057	0.971	0.059	0.057	0.971	0.059	0.057	0.963
400	0.039	0.040	1.020	0.039	0.040	1.018	0.039	0.040	1.023
	$q = 0; s_n = \mathbf{10}$								
50	0.237	0.241	1.019	0.214	0.218	1.018	0.166	0.169	1.017
100	0.241	0.248	1.029	0.216	0.222	1.030	0.165	0.170	1.035
200	0.247	0.253	1.025	0.222	0.227	1.026	0.169	0.174	1.028
400	0.252	0.253	1.004	0.226	0.228	1.008	0.172	0.175	1.017
	$q = 0; s_n = \mathbf{10}^3$								
50	0.200	0.203	1.015	0.181	0.182	1.009	0.141	0.141	1.002
100	0.192	0.197	1.023	0.170	0.175	1.025	0.128	0.131	1.027
200	0.191	0.196	1.030	0.168	0.173	1.030	0.124	0.128	1.033
400	0.194	0.196	1.011	0.171	0.173	1.016	0.125	0.128	1.026
	$q = 0; s_n = \mathbf{10}^5$								
50	0.199	0.201	1.014	0.179	0.181	1.008	0.140	0.140	1.001
100	0.188	0.192	1.021	0.167	0.170	1.022	0.125	0.128	1.024
200	0.182	0.188	1.029	0.160	0.165	1.028	0.118	0.122	1.030
400	0.180	0.181	1.010	0.157	0.159	1.015	0.114	0.117	1.024

This working paper has been produced by
the Department of Economics at
Queen Mary, University of London

Department of Economics
Queen Mary, University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 5096
Fax: +44 (0)20 8983 3580
Web: www.econ.qmul.ac.uk/papers/wp.htm