# Department of Economics

## Factor-GMM Estimation with Large Sets of Possibly Weak Instruments

George Kapetanios and Massimiliano Marcellino

Working Paper No. 577 October 2006 ISSN 1473-0278

Queen Mary
University of London

# Factor-GMM Estimation with Large Sets of Possibly Weak Instruments

George Kapetanios*
Queen Mary University of London

Massimiliano Marcellino†
IEP-Bocconi University, IGIER and CEPR

Preliminary Version: October 2006

**Abstract**

This paper analyses the use of factor analysis for instrumental variable estimation when the number of instruments tends to infinity. We consider cases where the unobserved factors are the optimal instruments but also cases where the factors are not necessarily the optimal instruments but can provide a summary of a large set of instruments. Further, the situation where many weak instruments exist is also considered in the context of factor models. Theoretical results, simulation experiments and empirical applications highlight the relevance and simplicity of Factor-GMM estimation.

*J.E.L. Classification:* C32, C51, E52

*Keywords:* Factor models, Principal components, Instrumental variables, GMM, weak instruments, DSGE models

*Department of Economics, Queen Mary, University of London, Mile End Rd., London E1 4NS. Email: G.Kapetanios@qmul.ac.uk

†IGIER - Università Bocconi, Via Salasco 5, 20136, Milano, Italy. Phone: +39-02-5836-3327. Fax: +39-02-5836-3302. E-mail: massimiliano.marcellino@uni-bocconi.it

# 1 Introduction

The paradigm of a factor model is very appealing and has been used extensively in economic analyses. Underlying the factor model is the idea that a large number of economic variables can be adequately modeled by a small number of indicator variables or shocks. Factor analysis has been used fruitfully to model, among other cases, asset returns, macroeconomic aggregates and Engel curves (see, e.g., Stock and Watson (1989), Lewbel (1991) and others).

Most analyses have traditionally been focused on small datasets meaning that the number of variables, $N$, to be modeled via a factor model is finite. Recently, Stock and Watson (2002) have put forward the case for analysing large datasets via factor analysis, where $N$ is allowed to tend to infinity. Stock and Watson (2002) suggest the use of principal components for estimating factors in this context. Similar work in a more general setting has been carried out by, e.g., Forni, Hallin, Lippi, and Reichlin (2000) and Forni, Hallin, Lippi, and Reichlin (2004) in which use of dynamic principal components has been made.

Most of this work has focused on exploiting the parsimony of factor models to provide a more adequate reduced form modelling tool in various contexts such as, e.g, forecasting. However, a facet where factor analysis has been less widely considered is the estimation of structural relationships. Estimation of regression models in the presence of endogeneity of the regressors is a well established area of research where the use of instruments is made to provide valid estimation and inference methods. In the past, the relevant literature has analysed cases where the number of instruments tends to infinity as the sample size grows. Eminent examples of this work are Morimune (1983) and Bekker (1994). Clearly once allowance is made for an increasing number of instruments, the relevance of tools that parsimoniously represent an increasing number of variables, such as factor models, becomes of relevance.

From an empirical point of view, Favero, Marcellino, and Neglia (2005) show that using factors extracted from a large set of macroeconomic variables as additional instruments in GMM estimation of forward looking Taylor rules for the US and Europe, substantially improves the efficiency of the parameter estimators. Beyer, Farmer, Henry, and Marcellino (2005) extend the analysis to a system context, where a Taylor rule is jointly estimated with a forward looking output equation and a hybrid Phillips curve, along the lines of Galí and Gertler (1999), finding again substantial gains in the GMM estimator's efficiency when adding factors to the instrument set. The present paper provides a theoretical explanation for such empirical findings, and more generally a theory for Factor-GMM estimation in the presence of a large set of instruments.

Another paper that analyzes the interface of factor models and instrumental variable estimation is Bai and Ng (2006b). In a similar vein to our paper they consider the case where the regressors are linear functions of a set of unobserved factors which are also underlying an expanding set of observed instruments. Clearly, under these circumstances it is intuitive to expect that knowledge of the factors would provide a superior estimator to the one using the observed set of instruments. Factor analysis can provide an estimate of the factors and thereby enable feasible factor-IV estimation. Bai and Ng (2006b) and our paper independently analyse this possibility.

However, assuming that the endogenous regressors are only functions of the factors is clearly a restrictive assumption. We therefore, generalise our and the analysis of Bai and Ng (2006b) to the case where regressors are either only functions of the observed instruments or both the observed instruments and the unobserved factors. In these two more general cases the superiority of a Factor-GMM estimator is not obvious. In fact we show that, when regressors are only functions of the observed instruments, standard IV estimation is preferable to using the factors as instruments, even when the factors are known. In the more general case where the endogenous variables depend both on a small set of key instruments and on the factors, proving the superiority of either Factor-GMM or standard GMM estimation is not possible, since the ranking depends on the parameter values.

Our analysis of Factor-GMM also evaluates another important aspect of IV estimation that has been recently explored in the literature. This is the possibility that instruments are weak, in the sense that their relation to the endogenous regressors is local-to-zero. A key reference in the context of a finite set of instruments is Staiger and Stock (1997). In that paper the strength of the correlation of the regressors and the instruments is measured in terms of what is refereed to as a concentration parameter. In standard IV estimation this parameters diverges at a rate equal to the number of observations. Staiger and Stock (1997) consider the case of a constant concentration parameter, which implies that the IV estimator is no longer consistent. The work of Staiger and Stock (1997) has been extended in a variety of ways. However, in our view, the most interesting generalization relates to combining the framework of many instruments with the framework of weak instruments. This was first done in the literature by Chao and Swanson (2005), and subsequently generalised extensively by Han and Phillips (2006). We consider a number of elements of their analysis in the context of Factor-GMM estimation, which provides a richer framework for analysing weak instruments than those previously adopted.

The paper is structured as follows: Section 2 develops the theoretical properties of factor-based instrumental variable estimators. Section 3 provides an extensive Monte Carlo study of the Factor-IV estimator for a wide variety of settings, including ones where the instruments are weak or many, and cases where the regressors depend on either the unobserved factors or the observed instruments or both. Section 4 provides empirical illustrations of the new method. Finally, Section 5 concludes.

# 2 Theory

In this Section we study the properties of factor-based Instrumental Variable (IV) estimators. In the first subsection we derive results for the standard IV case where the errors are uncorrelated and homoskedastic, which is useful to provide insights on the working of factors as instruments. In the second subsection we extend the analysis to general error structures.

## 2.1 Factor-IV estimation

Let the equation of interest be

$$y_t = x_t'\beta + \epsilon_t, \quad t = 1, \dots, T, \tag{1}$$

where the $k$ regressors in $x_t'$ are possibly correlated with the error term $\epsilon_t$.[1] A standard source of correlation in the IV literature is measurement error, which could be widespread in macroeconomic applications, where the variables are typically expressed as deviations from an unobservable equilibrium value. Another source of endogeneity is, of course, simultaneity, which is again widespread in applied macroeconomic applications based on single equation estimation. A more specific source of endogeneity in forward looking models, such as the new generation of DSGE models, is the use of expectations of future variables as regressors, which are then typically replaced by their true values for estimation, see for example the literature on Taylor rules or hybrid Phillips curves (e.g., Clarida, Galí, and Gertler (1998) or Galí and Gertler (1999)).

[1] In practice, only a subset of the $k$ regressors could be correlated with the error term, but for the sake of simplicity we will assume that all of them are endogenous.

Let us assume that there exist $N$ instrumental variables, $z_t$, generated by a factor model with $r$ factors:

$$z_t = \Lambda_N^{0\prime} f_t + e_t, \tag{2}$$

where $r$ is much smaller than $N$. Therefore, each instrumental variable can be decomposed into a common component (an element of $\Lambda_N^{0\prime} f_t$) that is driven by a few common forces, the factors, and an idiosyncratic component (an element of $e_t$). When the latter is small compared to the former, the information in the large set of $N$ instrumental variables $z_t$ can be efficiently summarized by the $r$ factors $f_t$.

We consider three different data generation mechanisms for $x_t$ that allow for non-zero correlation between $x_t$ and $\epsilon_t$, and a different degree of efficiency of the instruments $z_t$ and of the factors $f_t$. They are given by

$$x_t = A_{NZ}^{0\prime} z_t + u_t, \tag{3}$$

$$x_t = A_N^{0\prime} f_t + u_t, \tag{4}$$

and

$$x_t = A_{NZ}^{0\prime} z_t + A_N^{0\prime} f_t + u_t, \tag{5}$$

with $E(u_t \epsilon_t) \neq 0$ to introduce simultaneity in (1). In (3) the endogenous variables depend directly on the instruments. Therefore, the optimal instuments in this case are $z_t$ rather than $f_t$ and, conditional on $z_t$, the factors are irrelevant. However, when $N$ is very large, possibly larger than $T$, the standard IV estimator becomes inefficient or unfeasible. In this context, the factors could become useful again, since they provide a coincise summary of the information in $z_t$.

In (4), the endogenous variables depend directly on the factors. This is the case also considered by Bai and Ng (2006b), and it represents the most favourable situation for Factor-IV estimation, since the original instruments $z_t$ become irrelevant, conditional on the factors.[2]

In (5), the endogenous variables depend both on (possibly a subset of) the instrumental variables and on (possibly a subset of) the factors. This appears to be the most interesting case from an economic point of view. For example, future inflation can be expected to depend on a set of key macroeconomic indicators, such as monetary policy, oil prices and unit labor

[2] More specifically, Bai and Ng (2006b) assume that only a subset of the regressors in (1) are endogenous, say $x_{2t}$. The $x_{2t}$ variables depend on a few factors, say $f_t$, which are a subset of the set of factors that drives the $N$ exogenous variables $z_t$, say $F_t$. They also discuss procedures for selecting $f_t$ from $F_t$.

costs, on the past values of inflation itself due to persistence, but also on the behaviour of a large set of other variables, such as developments at the sectoral or regional level, that can be well summarized by a few factors (see Beck, Hubrich and Marcellino (2006)). A similar reasoning holds for unobservable variables, such as the output gap. Unfortunately, under (5) it is not possible to provide a unique ranking of the standard IV and of the Factor-(only)-IV estimators, since the ranking depends on the loading matrices in (5). However, it is clear that in this context the optimal estimator requires a combination of standard instruments and factors.

Stacking observations across time for all models presented above gives:

$$y = X\beta + \epsilon \tag{6}$$

$$Z = F\Lambda_N^0 + v \tag{7}$$

$$X = ZA_{NZ}^0 + u \tag{8}$$

$$X = FA_N^0 + u \tag{9}$$

$$X = ZA_{NZ}^0 + FA_N^0 + u \tag{10}$$

where $y = (y_1, ..., y_T)'$, $X = (x_1, ..., x_T)'$, $Z = (z_1, ..., z_T)'$, $F = (f_1, ..., f_T)'$, $u = (u_1, ..., u_T)'$, $v = (v_1, ..., v_T)'$ and $\epsilon = (\epsilon_1, ..., \epsilon_T)'$. Let $\hat{F}$ denote the Stock and Watson (2002) principal component estimator of $F$. We consider two alternative two stages least squares estimators for the parameters of interest, $\beta$:

$$\hat{\beta} = \left(X'\hat{F}(\hat{F}'\hat{F})^{-1}\hat{F}'X\right)^{-1} X'\hat{F}(\hat{F}'\hat{F})^{-1}\hat{F}'y \tag{11}$$

and

$$\tilde{\beta} = \left(X'Z(Z'Z)^{-1}Z'X\right)^{-1} X'Z(Z'Z)^{-1}Z'y \tag{12}$$

We also define the infeasible estimator given by

$$\bar{\beta} = \left(X'F(F'F)^{-1}F'X\right)^{-1} X'F(F'F)^{-1}F'y \tag{13}$$

We make the following assumptions:

**Assumption 1** *1. $E||f_t||^4 \leq M < \infty$, $T^{-1}\sum_{t=1}^{T} f_t f_t' \xrightarrow{p} \Sigma$ for some $k \times k$ positive definite matrix $\Sigma$.*

*2. $E(e_{i,t}) = 0$, $E|e_{i,t}|^8 \leq M$*

*3. For $\tau_{i,j,t,s} \equiv E(e_{i,t}e_{j,s})$ the following hold*

- $(NT)^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} |\sum_{i=1}^{N} \tau_{i,i,t,s}| \leq M$
- $|1/N \sum_{i=1}^{N} \tau_{i,i,s,s}| \leq M$ *for all* $s$
- $N^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N} |\tau_{i,j,s,s}| \leq M$
- $(NT)^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \sum_{i=1}^{N} \sum_{j=1}^{N} |\tau_{i,j,t,s}| \leq M$
- *For every* $(t, s)$, $E|(N)^{-1/2} \sum_{i=1}^{N}(e_{i,s}e_{i,t} - \tau_{i,i,s,t})|^4 \leq M$

**Assumption 2** $\epsilon_t$ *is a martingale difference sequence with finite fourth moment and* $E(\epsilon_t|\mathcal{F}_t) = \sigma^2 < \infty$ *where* $\mathcal{F}_t$ *is the* $\sigma$*-field generated by* $(f_s, z_s)$, $s \leq t$.

**Assumption 3** $(x_t', z_t')$ *are jointly stationary.* $z_t$ *is predetermined, so that* $E(z_{it}\epsilon_t = 0)$, $i = 1, ..., N$. *The probability limit of* $\frac{z_t z_t'}{T}$ *is finite and nonsingular.* $E(z_t x_t')$ *has full column rank* $k$. $x_t$ *and* $z_t$ *have finite fourth moments.*

Assumption 1 is standard in the factor literature. In particular, it is used in Stock and Watson (2002), Bai and Ng (2002) and Bai (2003) to prove consistency and asymptotic normality (at certain rates) of the principal component based estimator of the factors, and by Bai and Ng (2006a) to show consistency of the parameter estimators in factor augmented regressions. Assumption 3 guarantees that standard IV estimation using $z_t$ as instruments is feasible, and Assumption 2 that it is efficient. Assumption 2 will be relaxed in the following subsection, while Assumptions 1 and 3 are assumed to hold throughout the paper.

We now present a number of results related to the above setup. Note that we have allowed for $\Lambda_N^0$, $A_{NZ}^{0\prime}$ and $A_N^{0\prime}$ to depend on $N$.

The first result we present analyses conditions under which a local-to-zero factor loading matrix in (2) leads to a model that loses its defining factor characteristic which are commonly taken to imply that the largest $r$ eigenvalues of the variance covariance of $z_t$ tend to infinity. This setup implies that if (3) holds then the true (or, by extension, estimated) factors are weak instruments. Further, if (4) holds then the observed instruments, $z_t$, are weak instruments. Formally we note that instruments are referred to as weak when $A_{NZ}^{0\prime} Z' Z A_{NZ}^0$ or $A_N^{0\prime} F' F A_N^0$ is less than $O_p(T)$. This is the definition usually adopted in the relevant literature (see, e.g., Chao and Swanson (2005)).

**Theorem 1** *Let* $\Lambda_N^0 = \Lambda^0/N^\alpha$, $A_{NZ}^0 = A_Z^0$ *and* $A_N^0 = A^0$. *The eigenvalues of the population variance covariance matrix of* $Z$ *are bounded for* $\alpha \geq 1/2$ *for all* $N$.

**Proof.** The covariance matrix of $Z$, $\Sigma_Z$, is given by $\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0+\Sigma_v$ where $\Sigma_f$ and $\Sigma_v$ are the covariance matrices of $F$ and $v$ respectively. By Weyl's theorem (see 5.3.2(9) of Lutkepohl (1996)) the eigenvalues of $\Sigma_Z$ are bounded if the eigenvalues of $\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0$ and $\Sigma_v$ are bounded. By assumption the eigenvalues of $\Sigma_v$ are bounded. Hence we examine $\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0$. By Schwarz, Rutishauser, and Stiefel (1973), the eigenvalues of $\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0$, will be bounded if the column sum norm of $\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0$ is bounded. But every element of $\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0$ is $O(N^{-2\alpha})$. Hence the column sum norm of $\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0$ is O(1) for all $\alpha\geq 1/2$. Hence the result follows. ■

In the case considered in the above Theorem, the factor model is no longer identifiable and common and idiosyncratic components cannot be distinguished. This possibility is studied in some detail in Onatski (2006). The next result concerns estimation of the factors for a local-to-zero factor loading matrix.

**Theorem 2** *Let $\Lambda_N^0=\Lambda^0/N^\alpha$ where $0<\alpha<1/4$. Then, $\hat{F}-FH'=o_p(1)$ for some nonsingular matrix $H$, as long as $N^{2\alpha}=o(T^{1/2})$. Further,*

$$\frac{1}{T}\sum_{t=1}^{T}\left\|\hat{f}_t-Hf_t\right\|^2=O_p\left(\min\left(N^{4\alpha}T^{-1},N^{-1+4\alpha}\right)\right)=O_p\left(N^{4\alpha-1}\right)$$

*since $N=o(T)$.*

**Proof.** We follow the proof of Theorem 1 of Bai and Ng (2002). A crucial difference is that because of the local nature of $\Lambda_N^0$ we use a different normalisation for $\Lambda$. Therefore, rather than the normalisation $\Lambda'\Lambda/N=I$, we use $\Lambda'\Lambda/N^{1-2\alpha}=I$. This leads to the mathematical identities $\hat{F}=N^{-1+2\alpha}X\tilde{\Lambda}$ and $\tilde{\Lambda}=T^{-1}X'\tilde{F}$ where $\tilde{F}$ is the solution to the optimisation problem of maximising $tr(F'(X'X)F)$ subject to $F'F/T=I$. Let $H=(\tilde{F}'F/T)(\Lambda_N^{0\prime}\Lambda_N^0/N^{1-2\alpha})$. Then,

$$\hat{f}_t-Hf_t=N^{2\alpha}T^{-1}\sum_{s=1}^{T}\tilde{f}_s\gamma_N(s,t)+N^{2\alpha}T^{-1}\sum_{s=1}^{T}\tilde{f}_s\zeta_{st}+N^{2\alpha}T^{-1}\sum_{s=1}^{T}\tilde{f}_s\eta_{st}+N^{2\alpha}T^{-1}\sum_{s=1}^{T}\tilde{f}_s\xi_{st} \tag{14}$$

where $\gamma_N(s,t)=E(e_s'e_t/N)$

$$\zeta_{st}=e_s'e_t/N-\gamma_N(s,t) \tag{15}$$

$$\eta_{st}=f_s'^0\Lambda_N^0e_t/N \tag{16}$$

$$\xi_{st}=f_t'^0\Lambda_N^0e_s/N=\eta_{ts} \tag{17}$$

It is easy to see that

$$||\hat{f}_t-Hf_t||^2\leq 4(a_t+b_t+c_t+d_t) \tag{18}$$

where

$$a_t = N^{4\alpha} T^{-2} \left\| \sum_{s=1}^{T} \tilde{f}_s \gamma_N(s,t) \right\|^2 \tag{19}$$

$$b_t = N^{4\alpha} T^{-2} \left\| \sum_{s=1}^{T} \tilde{f}_s \zeta_{st} \right\|^2 \tag{20}$$

$$c_t = N^{4\alpha} T^{-2} \left\| \sum_{s=1}^{T} \tilde{f}_s \eta_{st} \right\|^2 \tag{21}$$

$$d_t = N^{4\alpha} T^{-2} \left\| \sum_{s=1}^{T} \tilde{f}_s \xi_{st} \right\|^2$$

It follows that

$$1/T \sum_{t=1}^{T} ||\hat{f}_t - H f_t||^2 \leq c/T \sum_{t=1}^{T} (a_t + b_t + c_t + d_t) \tag{22}$$

for some constant $c$. Now

$$\left\| \sum_{s=1}^{T} \tilde{f}_t \gamma_N(s,t) \right\|^2 \leq \left( \sum_{s=1}^{T} \left\| \tilde{f}_s \right\|^2 \right) \left( \sum_{s=1}^{T} \gamma_N^2(s,t) \right) \tag{23}$$

which implies that

$$1/T \sum_{t=1}^{T} a_t = O_p(N^{4\alpha} T^{-1}) \tag{24}$$

Following the analysis for $N^{-4\alpha} b_t$ given in the proof of Bai and Ng (2002) we see that

$$1/N^{4\alpha} T \sum_{t=1}^{T} b_t = O_p(N^{-1}) \tag{25}$$

and hence

$$1/T \sum_{t=1}^{T} b_t = O_p(N^{-1+4\alpha}) = o_p(1) \tag{26}$$

as long as $a < 1/4$. Finally we look at $c_t$. $d_t$ can be treated similarly.

$$c_t = N^{4\alpha} T^{-2} \left\| \sum_{s=1}^{T} \tilde{f}_s \eta_{st} \right\|^2 = N^{4\alpha} T^{-2} \left\| \sum_{s=1}^{T} \tilde{f}_s f_s'^0 \Lambda_N^0 e_t / N \right\|^2 = \tag{27}$$

$$N^{2\alpha} T^{-2} \left\| \sum_{s=1}^{T} \tilde{f}_s f_s'^0 \Lambda^0 e_t / N \right\|^2 \leq$$

$$N^{-2+4\alpha} \left\| \Lambda^0 e_t \right\|^2 \left( T^{-1} \sum_{s=1}^{T} \left\| \tilde{f}_s \right\|^2 \right) \left( T^{-1} \sum_{s=1}^{T} \| f_s \|^2 \right) = N^{-2+4\alpha} \left\| \Lambda^0 e_t \right\|^2 O_p(1)$$

So

$$1/T \sum_{t=1}^{T} c_t = O_p(1) N^{-1+4\alpha} T^{-1} \sum_{t=1}^{T} \left\| \frac{\Lambda^0 e_t}{N^{1/2}} \right\|^2 = O_p(N^{-1+4\alpha}) \tag{28}$$

■

Therefore, a sufficient condition for estimation in the local to zero case is the presence of a relatively strong local-to-zero factor model $(\alpha < 1/4)$ and a slow rate of increase for $N$. Note also the tradeoff between $\alpha$ and the allowable rate of increase for $N$.

The next theorem provides the asymptotic distribution of the alternative IV estimators for the case where $N$ is finite, and shows that, as expected, their relative efficiency depends on the generating mechanism of the data.

**Theorem 3** *Assuming that $f_t$ is observed, then for finite $N$ the asymptotic variance covariance matrix of $\sqrt{T}(\bar{\beta}-\beta)$ under (3)-(5) are given, up to the same scalar constant of proportionality, by*

$$var\left(\sqrt{T}(\bar{\beta}-\beta)\right)=\left(A_{NZ}^{0\prime}\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0A_{NZ}^0\right)^{-1} \tag{29}$$

$$var\left(\sqrt{T}(\bar{\beta}-\beta)\right)=\left(A_N^{0\prime}\Sigma_fA_N^0\right)^{-1} \tag{30}$$

*and*

$$var\left(\sqrt{T}(\bar{\beta}-\beta)\right)=\left(\left(A_{NZ}^{0\prime}\Lambda_N^{0\prime}+A_N^{0\prime}\right)\Sigma_f\left(\Lambda_N^0A_{NZ}^0+A_N^0\right)\right)^{-1} \tag{31}$$

*respectively. The asymptotic variance covariance matrix of $\sqrt{T}(\tilde{\beta}-\beta)$ under (3)-(5) are given, up to the same scalar constant of proportionality, by*

$$var\left(\sqrt{T}(\tilde{\beta}-\beta)\right)=\left(A_{NZ}^{0\prime}\left(\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0+\Sigma_v\right)A_{NZ}^0\right)^{-1} \tag{32}$$

$$var\left(\sqrt{T}(\tilde{\beta}-\beta)\right)=\left(A_N^{0\prime}\Sigma_f\Lambda_N^{0\prime}\left(\Lambda_N^0\Sigma_f\Lambda_N^{0\prime}+\Sigma_v\right)^{-1}\Lambda_N^0\Sigma_fA_N^0\right)^{-1} \tag{33}$$

*and*

$$var\left(\sqrt{T}(\tilde{\beta}-\beta)\right)=\left(\left(A_N^{0\prime}\left(\Sigma_f\Lambda_N^0+\Sigma_v\right)+A_N^{0\prime}\Sigma_f\Lambda_N^{0\prime}\right)\left(\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0+\Sigma_v\right)^{-1}\right. \tag{34}$$
$$\left.\left(A_N^{0\prime}\left(\Sigma_f\Lambda_N^0+\Sigma_v\right)+A_N^{0\prime}\Sigma_f\Lambda_N^{0\prime}\right)'\right)^{-1}$$

*respectively. The difference between the RHS of (29) and (32) is a positive semidefinite matrix. The difference between the RHS of (33) and (30) is a positive semidefinite matrix.*

**Proof.** Asymptotic normality for the estimators follows straightforwardly from the martingale difference central limit theorem given Assumption 2. We then examine the asymptotic variances. The general expressions for the covariance matrices of $\sqrt{T}(\bar{\beta}-\beta)$ and $\sqrt{T}(\tilde{\beta}-\beta)$ are given by the probability limits as $T\to\infty$ of $\left(\frac{X'F}{T}\left(\frac{F'F}{T}\right)^{-1}\frac{F'X}{T}\right)^{-1}$ and $\left(\frac{X'Z}{T}\left(\frac{Z'Z}{T}\right)^{-1}\frac{Z'X}{T}\right)^{-1}$. We begin by deriving results under (4) as it is more straightforward.

The following probability limits, using standard laws of large numbers and the uncorrelatedness of $u_t$ and $v_t$, give the required ingredients for the results

$$\frac{X'F}{T} = \frac{\left(A_N^{0'}F' + u'\right)F}{T} \xrightarrow{p} A_N^{0'}\Sigma_f \tag{35}$$

$$\frac{F'F}{T} \xrightarrow{p} \Sigma_f \tag{36}$$

$$\frac{X'Z}{T} = \frac{\left(A_N^{0'}F' + u'\right)\left(F\Lambda_N^0 + v\right)}{T} \xrightarrow{p} A_N^{0'}\Sigma_f\Lambda_N^0 \tag{37}$$

$$\frac{Z'Z}{T} = \frac{\left(F\Lambda_N^0 + v\right)'\left(F\Lambda_N^0 + v\right)}{T} \xrightarrow{p} \Lambda_N^{0'}\Sigma_f\Lambda_N^0 + \Sigma_v \tag{38}$$

Then, (30) and (33) easily follow. Similarly, under (3), (35) and (38) become

$$\frac{X'F}{T} = \frac{\left(A_{NZ}^{0'}\left(\Lambda_N^{0'}F' + v'\right) + u'\right)F}{T} \xrightarrow{p} A_{NZ}^{0'}\Lambda_N^{0'}\Sigma_f \tag{39}$$

$$\frac{X'Z}{T} = \frac{\left(A_{NZ}^{0'}\left(\Lambda_N^{0'}F' + v'\right) + u'\right)\left(F\Lambda_N^0 + v\right)}{T} \xrightarrow{p} A_{NZ}^{0'}\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0 + A_{NZ}^{0'}\Sigma_v = A_{NZ}^{0'}\left(\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0 + \Sigma_v\right) \tag{40}$$

Hence, (29) and (32) easily follow. Finally, under (5), (35) and (38) become

$$\frac{X'F}{T} = \frac{\left(A_{NZ}^{0'}\left(\Lambda_N^{0'}F' + v'\right) + A_N^{0'}F' + u'\right)F}{T} \xrightarrow{p} \left(A_{NZ}^{0'}\Lambda_N^{0'} + A_N^{0'}\right)\Sigma_f \tag{41}$$

$$\frac{X'Z}{T} = \frac{\left(A_{NZ}^{0'}\left(F'\Lambda_N^{0'} + v'\right) + A_N^{0'}F' + u'\right)\left(F\Lambda_N^0 + v\right)}{T} \xrightarrow{p} A_{NZ}^{0'}\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0 + A_N^{0'}\left(\Sigma_v + \Sigma_f\Lambda_N^0\right) \tag{42}$$

We next examine

$$\left(A_{NZ}^{0'}\Lambda_N^{0'}\Sigma_f\Lambda_N^0 A_{NZ}^0\right)^{-1} - \left(A_{NZ}^{0'}\left(\Lambda_N^{0'}\Sigma_f\Lambda_N^0 + \Sigma_v\right)A_{NZ}^0\right)^{-1}$$

which is positive semidefinite (psd) if

$$A_{NZ}^{0'}\left(\Lambda_N^{0'}\Sigma_f\Lambda_N^0 + \Sigma_v\right)A_{NZ}^0 - A_{NZ}^{0'}\Lambda_N^{0'}\Sigma_f\Lambda_N^0 A_{NZ}^0 \tag{43}$$

is psd. But (43) is psd since

$$\Lambda_N^{0'}\Sigma_f\Lambda_N^0 + \Sigma_v - \Lambda_N^{0'}\Sigma_f\Lambda_N^0 = \Sigma_v$$

Hence the result follows. Finally we examine

$$\left(A_N^{0'}\Sigma_f\Lambda_N^{0'}\left(\Lambda_N^0\Sigma_f\Lambda_N^{0'} + \Sigma_v\right)^{-1}\Lambda_N^0\Sigma_f A_N^0\right)^{-1} - \left(A_N^{0'}\Sigma_f A_N^0\right)^{-1}$$

which is psd if

$$A_N^{0'}\Sigma_f A_N^0 - A_N^{0'}\Sigma_f\Lambda_N^{0'}\left(\Lambda_N^0\Sigma_f\Lambda_N^{0'} + \Sigma_v\right)^{-1}\Lambda_N^0\Sigma_f A_N^0 \tag{44}$$

is. But (44) is psd if

$$A_N^{0\prime}\Sigma_f A_N^0 - A_N^{0\prime}\Sigma_f \Lambda_N^{0\prime}\left(\Lambda_N^0\Sigma_f\Lambda_N^{0\prime}\right)^{-1}\Lambda_N^0\Sigma_f A_N^0 \tag{45}$$

is. Define $\tilde{A}_N^0 = \Sigma_f^{1/2}A_N^0$ and $\tilde{\Lambda}_N^0 = \Lambda_N^0\Sigma_f^{1/2}$. Then, (45) becomes

$$\tilde{A}_N^{0\prime}\tilde{A}_N^0 - \tilde{A}_N^{0\prime}\tilde{\Lambda}_N^{0\prime}\left(\tilde{\Lambda}_N^0\tilde{\Lambda}_N^{0\prime}\right)^{-1}\tilde{\Lambda}_N^0\tilde{A}_N^0 = \tilde{A}_N^{0\prime}\left(I - \tilde{\Lambda}_N^{0\prime}\left(\tilde{\Lambda}_N^0\tilde{\Lambda}_N^{0\prime}\right)^{-1}\tilde{\Lambda}_N^0\right)\tilde{A}_N^0 \tag{46}$$

which is psd since $I - \tilde{\Lambda}_N^{0\prime}\left(\tilde{\Lambda}_N^0\tilde{\Lambda}_N^{0\prime}\right)^{-1}\tilde{\Lambda}_N^0$ is. Hence the result follows. ■

The next result provides asymptotic equivalence between the feasible and infeasible Factor-IV estimators in the case of strong instruments and diverging $N$.

**Theorem 4** *Let $\Lambda_N^0 = \Lambda^0$, $A_{NZ}^0 = A_Z^0$ and $A_N^0 = A^0$. If $\sqrt{T}/N = o(1)$ then*

$$\sqrt{T}(\bar{\beta} - \beta) - \sqrt{T}(\hat{\beta} - \beta) = o_p(1) \tag{47}$$

**Proof.** We need to prove that

$$\sqrt{T}\left(\left(X'F(F'F)^{-1}F'X\right)^{-1}X'F(F'F)^{-1}F'\epsilon - \left(X'\hat{F}(\hat{F}'\hat{F})^{-1}\hat{F}'X\right)^{-1}X'\hat{F}(\hat{F}'\hat{F})^{-1}\hat{F}'\epsilon\right) = o_p(1) \tag{48}$$

or

$$\left(\frac{X'F}{T}\left(\frac{F'F}{T}\right)^{-1}\frac{F'X}{T}\right)^{-1}\frac{X'F}{T}\left(\frac{F'F}{T}\right)^{-1}\frac{F'\epsilon}{T^{1/2}} - \tag{49}$$

$$\left(\frac{X'\hat{F}}{T}\left(\frac{\hat{F}'\hat{F}}{T}\right)^{-1}\frac{\hat{F}'X}{T}\right)^{-1}\frac{X'\hat{F}}{T}\left(\frac{\hat{F}'\hat{F}}{T}\right)^{-1}\frac{\hat{F}'\epsilon}{T^{1/2}} = o_p(1)$$

(49) follows if

$$\frac{X'F}{T}\left(\frac{F'F}{T}\right)^{-1}\frac{F'X}{T} - \frac{X'\hat{F}}{T}\left(\frac{\hat{F}'\hat{F}}{T}\right)^{-1}\frac{\hat{F}'X}{T} = o_p(1) \tag{50}$$

and

$$\frac{X'F}{T}\left(\frac{F'F}{T}\right)^{-1}\frac{F'\epsilon}{T^{1/2}} - \frac{X'\hat{F}}{T}\left(\frac{\hat{F}'\hat{F}}{T}\right)^{-1}\frac{\hat{F}'\epsilon}{T^{1/2}} = o_p(1) \tag{51}$$

(50) and (51) follow if

$$\frac{X'F}{T} - \frac{X'\hat{F}}{T} = o_p(1) \tag{52}$$

$$\frac{F'F}{T} - \frac{\hat{F}'F}{T} = o_p(1) \tag{53}$$

and

$$\sqrt{T}\left(\frac{F'\epsilon}{T} - \frac{\hat{F}'\epsilon}{T}\right) = o_p(1) \tag{54}$$

hold. We examine (52)-(54). They can all be written as

$$A_T \frac{1}{T} \sum_{t=1}^{T} (\hat{f}_t - H f_t) q_t' = o_p(1) \tag{55}$$

where $A_T$ is $1, 1$ and $\sqrt{T}$ and $q_t$ is $x_t$, $f_t$ and $\epsilon_t$ respectively for (52)-(54). By Lemma A.1 of Bai and Ng (2006a) we have that

$$\frac{1}{T} \sum_{t=1}^{T} (\hat{f}_t - H f_t) q_t' = O_p\left(\min(N,T)^{-1}\right) \tag{56}$$

as long as $q_t$ has finite fourth moments, nonsingular covariance matrix and satisfies a central limit theorem. These conditions are satisfied for $x_t$, $f_t$ and $\epsilon_t$ via assumptions 2 and 3. Hence, (52)-(53) follow, while (54) follows if $\sqrt{T}/N = o(1)$. ■

The following theorem provides results for the case of many weak instruments. Note that in this case the instruments are weak by virtue of the specification of the relationships (3)-(5) rather than the presence of a local-to-zero factor loading matrix in (2). The latter possibility is not allowed since the rate at which the estimated factors converge to the space spanned by the true factors is affected by the presence of a local-to-zero factor loading matrix as detailed in Theorem 2. However, a distributional result for $\hat{\beta}$ in the presence of a local-to-zero factor loading matrix is given in Theorem 6.

**Theorem 5** *Let one of (3), (4) or (5) hold. Let $N = O(T^{\gamma})$, $\gamma > 1/2$. Let $\Lambda_N^0 = \Lambda^0$. Further, let every element of $\Lambda^0 A_{NZ}^0$ be $O(N^{-\beta}) = O(T^{-\alpha})$, where $\alpha = \gamma\beta$, $0 \leq \alpha < 1/2$. Also let every element of $A_N^0$ be $O(N^{-\delta}) = O(T^{-\vartheta})$, where $\vartheta = \gamma\delta$, $0 \leq \vartheta < 1/2$. Then, under (3),*

$$T^{1/2-\alpha}\left(\hat{\beta} - \beta\right) \xrightarrow{d} N\left(0, \sigma_\epsilon^2 \left(\Upsilon' \Sigma_f \Upsilon\right)^{-1}\right) \tag{57}$$

*Under (4)*

$$T^{1/2-\alpha}\left(\hat{\beta} - \beta\right) \xrightarrow{d} N\left(0, \sigma_\epsilon^2 \left(\Psi' \Sigma_f \Psi\right)^{-1}\right) \tag{58}$$

*and under (5), if $\vartheta > \alpha$*

$$T^{1/2-\alpha}\left(\hat{\beta} - \beta\right) \xrightarrow{d} N\left(0, \sigma_\epsilon^2 \left(\Psi' \Sigma_f \Psi\right)^{-1}\right) \tag{59}$$

*if $\vartheta < \alpha$,*

$$T^{1/2-\vartheta}\left(\hat{\beta} - \beta\right) \xrightarrow{d} N\left(0, \sigma_\epsilon^2 \left(\Upsilon' \Sigma_f \Upsilon\right)^{-1}\right) \tag{60}$$

*and if $\vartheta = \alpha$,*

$$T^{1/2-\alpha}\left(\hat{\beta} - \beta\right) \xrightarrow{d} N\left(0, \sigma_\epsilon^2 \left(\left(\Upsilon + \Psi\right)' \Sigma_f \left(\Upsilon + \Psi\right)\right)^{-1}\right) \tag{61}$$

*where*

$$\lim_{T\to\infty} \frac{\Lambda^0 A^0_{N(T)Z}}{T^{-\alpha}} = \Upsilon \tag{62}$$

$$\lim_{T\to\infty} \frac{A^0_{N(T)}}{T^{-\vartheta}} = \Psi \tag{63}$$

*and $\Upsilon$ and $\Psi$ are nonsingular matrices.*

**Proof.** We establish (57). (58)-(61) follow similarly. We examine the asymptotic distribution of

$$T^{1/2-\alpha} \left( X'\hat{F}(\hat{F}'\hat{F})^{-1}\hat{F}'X \right)^{-1} X'\hat{F}(\hat{F}'\hat{F})^{-1}\hat{F}'\epsilon \tag{64}$$

By theorem 4 and (55) for $A_T = o_p(T^{1/2})$ it is sufficient to examine the asymptotic distribution of

$$\left( \frac{X'F}{T^{1-\alpha}} \left(\frac{F'F}{T}\right)^{-1} \frac{F'X}{T^{1-\alpha}} \right)^{-1} \frac{X'F}{T^{1-\alpha}} \left(\frac{F'F}{T}\right)^{-1} \frac{F'\epsilon}{T^{1/2}} \tag{65}$$

A standard central limit theorem suffices to show that under assumptions 1-3

$$T^{-1/2}F'\epsilon \xrightarrow{d} N\left(0, \sigma^2_\epsilon \Sigma_f\right) \tag{66}$$

We examine the limits of $\frac{F'F}{T}$ and $\frac{X'F}{T^{1-\alpha}}$. The first is given in (36). We examine the second. We have

$$\frac{X'F}{T^{1-\alpha}} = \frac{\left(A^{0\prime}_{NZ}\left(\Lambda^{0\prime}_N F' + v'\right) + u'\right) F}{T^{1-\alpha}} = \frac{A^{0\prime}_{NZ}\Lambda^{0\prime}_N}{T^{-\alpha}}\frac{F'F}{T} + \frac{A^{0\prime}_{NZ}v'F}{T^{1-\alpha}} + \frac{u'F}{T^{1-\alpha}} \tag{67}$$

The second and third terms of the RHS of (67) tend to zero since $1-\alpha > 1/2$. The first term tends to $\Upsilon'\Sigma_f$. Hence, the result follows. ■

**Remark 1** *The assumption that the elements of $\Lambda^0_N$, $A^0_{NZ}$ and $A^0_N$ are deterministic can be relaxed to allow for the possibility of random elements that are independent of $F$, $\epsilon$, $u$ and $v$. Then, the conditions (62) and (63) would be modified to ones involving stochastic convergence.*

For the next result we need the following lemma.

**Lemma 1** *Let $\Lambda^0_N = \Lambda^0/N^\alpha$. Let $a < 1/4$ and $N^{2\alpha} = o(T^{1/2})$ . Then,*

$$\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t - Hf_t)q_t' = O_p\left(N^{2\alpha-1}\right) \tag{68}$$

*as long as $q_t$ has finite fourth moments, nonsingular covariance matrix and satisfies a central limit theorem.*

**Proof.** We follow the proof of Lemma A.1 of Bai and Ng (2006a). Using (15)-(17) we get

$$\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t - Hf_t)q_t' = N^{2\alpha}T^{-2}\sum_{t=1}^{T}\left(\sum_{s=1}^{T}\tilde{f}_s\gamma_N(s,t)\right)q_t' + N^{2\alpha}T^{-2}\sum_{t=1}^{T}\left(\sum_{s=1}^{T}\tilde{f}_s\zeta_{st}\right) + \tag{69}$$

$$N^{2\alpha}T^{-2}\sum_{t=1}^{T}\left(\sum_{s=1}^{T}\tilde{f}_s\eta_{st}\right)q_t' + N^{2\alpha}T^{-2}\sum_{t=1}^{T}\left(\sum_{s=1}^{T}\tilde{f}_s\xi_{st}\right)q_t'$$

The first two terms of (69) apart from the normalisation $N^{2\alpha}$ are the same as those analysed in Lemma A.1 of Bai and Ng (2006a). Thus, under the assumption of the Lemma for $q_t$, we immediately get that they are $O_p\left(N^{2\alpha}T^{-1/2}\min(N,T)^{-1/2}\right)$ and $O_p\left(N^{2\alpha-1/2}\min(N,T)^{-1/2}\right)$ respectively. The third and the fourth term of (69) are analysed similarly. We focus on the third term. We have

$$N^{2\alpha}T^{-2}\frac{1}{T}\sum_{t=1}^{T}\left(\sum_{s=1}^{T}\tilde{f}_s\eta_{st}\right)q_t' = N^{2\alpha}T^{-2}\sum_{t=1}^{T}\left(\sum_{s=1}^{T}Hf_s\eta_{st}\right)q_t' + N^{2\alpha}T^{-2}\sum_{t=1}^{T}\left(\sum_{s=1}^{T}\left(\tilde{f}_s - Hf_s\right)\eta_{st}\right)q_t' \tag{70}$$

The first term on the RHS of (70) can be written as

$$N^{2\alpha}T^{-2}\sum_{t=1}^{T}\left(\sum_{s=1}^{T}Hf_s\eta_{st}\right)q_t' = N^{2\alpha}\left(H\frac{1}{T}\sum_{t=1}^{T}f_sf_s'\right)\frac{1}{NT}\sum_{t=1}^{T}\Lambda_N^0e_tq_t' =$$

$$N^{2\alpha}\left(H\frac{1}{T}\sum_{t=1}^{T}f_sf_s'\right)\frac{N^{-2\alpha}}{NT}\sum_{t=1}^{T}\Lambda^0e_tq_t' = \left(H\frac{1}{T}\sum_{t=1}^{T}f_sf_s'\right)\frac{1}{NT}\sum_{t=1}^{T}\Lambda^0e_tq_t' = O_p\left((NT)^{-1/2}\right)$$

For the second term of (70) we have

$$\left\|N^{2\alpha}T^{-2}\sum_{t=1}^{T}\left(\sum_{s=1}^{T}\left(\tilde{f}_s - Hf_s\right)\eta_{st}\right)q_t'\right\| \le \left(\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{f}_s - Hf_s\right\|^2\right)^{1/2}\left(N^{4\alpha}\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\eta_{st}q_t'\right\|^2\right)^{1/2}$$

But

$$\left(\frac{1}{T}\sum_{s=1}^{T}\left\|\tilde{f}_s - Hf_s\right\|^2\right)^{1/2} = O_p\left(N^{2\alpha-1/2}\right)$$

by Theorem 2. Then,

$$N^{4\alpha}\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\eta_{st}q_t'\right\|^2 = N^{4\alpha}\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\frac{f_s'^0\Lambda_N^0e_t}{N}q_t'\right\|^2 = \frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{T}\sum_{t=1}^{T}\frac{f_s'^0\Lambda^0e_t}{N}q_t'\right\|^2 \tag{71}$$

But,

$$\frac{1}{T}\sum_{t=1}^{T}\frac{f_s'^0\Lambda^0e_t}{N}q_t' = O_p(N^{-1/2})$$

and so the RHS of (71) is also $O_p(N^{-1/2})$. As a result, the third term of (69) is $O_p\left(N^{2\alpha-1}\right)$. Thus, overall

$$\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t - Hf_t)q_t' = O_p\left(N^{2\alpha}\min(N,T)^{-1}\right) = O_p\left(N^{2\alpha-1}\right)$$

since $N = o(T)$. ■

Then, we have the following theorem

**Theorem 6** *Let $\Lambda_N^0 = \Lambda^0/N^\alpha$. Let $a < 1/4$, $N^{2\alpha} = o(T^{1/2})$ and $N^{2\alpha-1}T^{1/2} = o(1)$. Then, Theorem 4 and (57) of Theorem 5 under (3) follow where*

$$\lim_{T\to\infty}\frac{\Lambda_{N(T)}^0 A_{N(T)Z}^0}{T^{-\alpha}} = \Upsilon \tag{72}$$

*and $A_{N(T)Z}^0 = A_Z^0$.*

**Proof.** The results follow from Lemma 1, (55) and the proofs of Theorems 4 and 5. ■

In summary, for some data generating processes, using $z_t$ rather than $f_t$ is preferable in the case of finite $N$, as detailed in Theorem 3. However, this result is reversed when $N$ tends to infinity. First, as $N \to \infty$, it becomes feasible to estimate the unobserved factors $f_t$ consistently even for local-to-zero factor models, as discussed in Theorem 1, and estimation of the factors does not to matter for the asymptotic properties of the Factor-IV estimators, as discussed in Theorems 4 and 6. Moreover, whereas estimation using the estimated factors remains consistent and asymptotically normal even in the case where $z_t$ are weak instruments, as discussed in Theorems 5 and 6, standard IV estimation can be inconsistent if the number of instrument increases fast enough, as discussed by Bekker (1994) and Chao and Swanson (2005).

## 2.2 Factor-GMM estimation

We now relax assumption 2 and allow for correlation and heteroskedasticity in the errors $\epsilon$ of equation (1). We formalise this with the following assumption, which substitutes assumption 2:

**Assumption 4** *$\epsilon_t$ is a zero mean process with finite variance. The process $z_t\epsilon_t$ and, by implication, $f_t\epsilon_t$, satisfies the conditions for the application of some central limit theorem for weakly dependent processes, with a zero mean asymptotic normal limit. The probability limits of $\frac{F'\epsilon\epsilon' F}{T}$ and $\frac{Z'\epsilon\epsilon' Z}{T}$, denoted by $S_{f\epsilon}$ and $S_{z\epsilon}$ exist and are nonsingular.*

We further add the following regularity condition.

**Assumption 5** *$E[(z_{ti}x_{tj})^2]$ exists and is finite for i=1,...,N and j=1,...,k,*

**Remark 2** *Assumption 4 is a high level assumption. It is given in this form for generality. More primitive conditions on $\epsilon_t$ such as, e.g., mixing with polynomially declining mixing coefficients or near epoque dependence (see, e.g, Davidson (1994)) are sufficient for Assumption 4 to hold.*

As long as the instruments remain uncorrelated with the errors at all leads and lags, the estimators $\hat{\beta}$ and $\widetilde{\beta}$ in (11) and (12) remain consistent and asymptotically normal. Furthermore, we have:

**Theorem 7** *For finite $N$, the asymptotic variance covariance matrix of $\sqrt{T}(\bar{\beta}-\beta)$ under (3)-(4) are given by*

$$var\left(\sqrt{T}(\bar{\beta}-\beta)\right) = \left(A_{NZ}^{0\prime}\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0 A_{NZ}^0\right)^{-1} A_{NZ}^{0\prime}\Lambda_N^{0\prime}\Sigma_f^{-1}S_{f\epsilon}\Sigma_f^{-1}\Lambda_N^0 A_{NZ}^0 \left(A_{NZ}^{0\prime}\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0 A_{NZ}^0\right)^{-1} \quad (73)$$

$$var\left(\sqrt{T}(\bar{\beta}-\beta)\right) = \left(A_N^{0\prime}\Sigma_f A_N^0\right)^{-1} A_N^{0\prime}\Sigma_f^{-1}S_{f\epsilon}\Sigma_f^{-1}A_N^0 \left(A_N^{0\prime}\Sigma_f A_N^0\right)^{-1} \quad (74)$$

*where $S_{f\epsilon} = E(F'\epsilon\epsilon' F)$. The asymptotic variance covariance matrix of $\sqrt{T}(\tilde{\beta}-\beta)$ under (3)-(4) are given by*

$$var\left(\sqrt{T}(\tilde{\beta}-\beta)\right) = \left(A_{NZ}^{0\prime}\left(\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0+\Sigma_v\right)A_{NZ}^0\right)^{-1} \quad (75)$$

$$A_{NZ}^{0\prime}S_{z\epsilon}A_{NZ}^0\left(A_{NZ}^{0\prime}\left(\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0+\Sigma_v\right)A_{NZ}^0\right)^{-1}$$

$$var\left(\sqrt{T}(\tilde{\beta}-\beta)\right) = \left(A_N^{0\prime}\Sigma_f\Lambda_N^{0\prime}\left(\Lambda_N^0\Sigma_f\Lambda_N^{0\prime}+\Sigma_v\right)^{-1}\Lambda_N^0\Sigma_f A_N^0\right)^{-1} \quad (76)$$

$$\left(A_N^{0\prime}\Sigma_f\Lambda_N^{0\prime-1}S_{z\epsilon}^{-1}\Lambda_N^0\Sigma_f A_N^0\right)\left(A_N^{0\prime}\Sigma_f\Lambda_N^{0\prime}\left(\Lambda_N^0\Sigma_f\Lambda_N^{0\prime}+\Sigma_v\right)^{-1}\Lambda_N^0\Sigma_f A_N^0\right)^{-1}$$

**Proof.** The general expressions for the covariance matrices of $\sqrt{T}(\bar{\beta}-\beta)$ and $\sqrt{T}(\tilde{\beta}-\beta)$ are given by the probability limits as $T\to\infty$ of

$$\left(\frac{X'F}{T}\left(\frac{F'F}{T}\right)^{-1}\frac{F'X}{T}\right)^{-1}\frac{X'F}{T}\left(\frac{F'F}{T}\right)^{-1}\frac{F'\epsilon\epsilon'F}{T}\left(\frac{F'F}{T}\right)^{-1}\frac{F'X}{T}\left(\frac{X'F}{T}\left(\frac{F'F}{T}\right)^{-1}\frac{F'X}{T}\right)^{-1}$$

and

$$\left(\frac{X'Z}{T}\left(\frac{Z'Z}{T}\right)^{-1}\frac{Z'X}{T}\right)^{-1}\frac{X'Z}{T}\left(\frac{Z'Z}{T}\right)^{-1}\frac{Z'\epsilon\epsilon'Z}{T}\left(\frac{Z'Z}{T}\right)^{-1}\frac{Z'X}{T}\left(\frac{X'Z}{T}\left(\frac{Z'Z}{T}\right)^{-1}\frac{Z'X}{T}\right)^{-1}.$$

The probability limits of $\frac{X'F}{T}$, $\frac{F'F}{T}$, $\frac{X'Z}{T}$ and $\frac{Z'Z}{T}$ are as in Theorem 3, while

$$\frac{F'\epsilon\epsilon'F}{T}\xrightarrow{p}S_{f\epsilon}, \quad (77)$$

$$\frac{Z'\epsilon\epsilon' Z}{T} \xrightarrow{p} S_{z\epsilon}. \tag{78}$$

■

Notice that when the errors $\epsilon$ are uncorrelated and homoskedastic, it is (up to a scalar constant) $S_{f\epsilon} = E(F'F) = \Sigma_f$ and $S_{z\epsilon} = E(Z'Z) = \Lambda_N^{0\prime}\Sigma_f\Lambda_N^0 + \Sigma_v$. Therefore, the variance covariance matrices of $\hat{\beta}$ and $\widetilde{\beta}$ reduce to those derived in Theorem 3. In practice, $S_{f\epsilon}$ and $S_{z\epsilon}$ can be estimated by a HAC procedure, such as that developed in Newey and West (1987). For example, using a Bartlett kernel, we have

$$\widehat{S}_{z\epsilon,h} = \widehat{\Phi}_0 + \sum_{j=1}^{h}(1-\frac{j}{h+1})(\widehat{\Phi}_j + \widehat{\Phi}_j')$$
$$\widehat{\Phi}_j = T^{-1}\sum_{T=j+1}^{T}\widehat{\epsilon}_t^2 z_t z_t',$$

where $h$ is the length of the window, $\widehat{\epsilon}_t = y_t - x_t' b$, and $b$ is a consistent estimator for $\beta$. We focus on the Newey and West (1987) HAC procedure using the Bartlett kernel in the rest of the section.

A remaining problem with the two stage least square estimators $\hat{\beta}$ and $\widetilde{\beta}$ is that they are not efficient in the presence of a general error structure. In fact, the efficient estimators in this context are obtained by GMM estimation with either $S_{f\epsilon}^{-1}$ or $S_{z\epsilon}^{-1}$ as the weighting matrix. Using standard methods, the resulting estimators are

$$\widehat{b} = \left(X'\hat{F}\widehat{S}_{\widehat{f}\epsilon}^{-1}\hat{F}'X\right)^{-1} X'\hat{F}\widehat{S}_{\widehat{f}\epsilon}^{-1}\hat{F}'y \tag{79}$$

$$\widetilde{b} = \left(X'Z\widehat{S}_{z\epsilon}^{-1}Z'X\right)^{-1} X'Z\widehat{S}_{z\epsilon}^{-1}Z'y \tag{80}$$

and

$$\bar{b} = \left(X'FS_{f\epsilon}^{-1}F'X\right)^{-1} X'FS_{f\epsilon}^{-1}F'y, \tag{81}$$

When the errors are uncorrelated and homoskedastic, these expressions simplify to those in (11)-(13).

**Theorem 8** *Assuming that $f_t$ is observed, then for finite $N$ the asymptotic variance covariance matrix of $\sqrt{T}(\bar{b}-\beta)$ under (3)-(5) are given by*

$$var\left(\sqrt{T}(\bar{b}-\beta)\right) = \left(A_{NZ}^{0\prime}\Lambda_N^{0\prime}\Sigma_f S_{f\epsilon}^{-1}\Sigma_f\Lambda_N^0 A_{NZ}^0\right)^{-1} \tag{82}$$

$$var\left(\sqrt{T}(\bar{b}-\beta)\right) = \left(A_N^{0\prime}\Sigma_f S_{f\epsilon}^{-1}\Sigma_f A_N^0\right)^{-1} \tag{83}$$

*and*

$$var\left(\sqrt{T}(\bar{b}-\beta)\right)=\left(\left(A_{NZ}^{0\prime}\Lambda_N^{0\prime}+A_N^{0\prime}\right)\Sigma_f S_{f\epsilon}^{-1}\Sigma_f\left(\Lambda_N^0 A_{NZ}^0+A_N^0\right)\right)^{-1} \tag{84}$$

*respectively. The asymptotic variance covariance matrix of* $\sqrt{T}(\widetilde{b}-\beta)$ *under (3)-(5) are given by*

$$var\left(\sqrt{T}(\widetilde{b}-\beta)\right)=\left(A_{NZ}^{0\prime}\left(\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0+\Sigma_v\right)S_{z\epsilon}^{-1}\left(\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0+\Sigma_v\right)A_{NZ}^0\right)^{-1} \tag{85}$$

$$var\left(\sqrt{T}(\widetilde{b}-\beta)\right)=\left(A_N^{0\prime}\Sigma_f\Lambda_N^{0\prime}S_{z\epsilon}^{-1}\Lambda_N^0\Sigma_f A_N^0\right)^{-1} \tag{86}$$

*and*

$$var\left(\sqrt{T}(\widetilde{b}-\beta)\right)=\left(\left(A_{NZ}^{0\prime}\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0+A_N^{0\prime}\left(\Sigma_f\Lambda_N^0+\Sigma_v\right)\right)S_{z\epsilon}^{-1}\right. \tag{87}$$
$$\left.\left(A_{NZ}^{0\prime}\Lambda_N^{0\prime}\Sigma_f\Lambda_N^0+A_N^{0\prime}\left(\Sigma_f\Lambda_N^0+\Sigma_v\right)\right)'\right)^{-1}$$

*respectively.*

**Proof.** The general expressions for the covariance matrices of $\sqrt{T}(\bar{b}-\beta)$ and $\sqrt{T}(\widetilde{b}-\beta)$ are given by the probability limits as $T\to\infty$ of $\left(\frac{X'F}{T}S_{f\epsilon}^{-1}\frac{F'X}{T}\right)^{-1}$ and $\left(\frac{X'Z}{T}\widehat{S}_{z\epsilon}^{-1}\frac{Z'X}{T}\right)^{-1}$. The results follow from those in the Proof of Theorem 3 and consistency of the HAC estimator of $S$. ■

Notice that the higher precision of $\bar{b}$ with respect to $\widetilde{b}$ when the model is (4) (and of $\widetilde{b}$ for 3) follows from the choice of the weighting matrix.

**Theorem 9** *Let* $\Lambda_N^0=\Lambda^0$, $A_{NZ}^0=A_Z^0$ *and* $A_N^0=A^0$. *If* $\sqrt{T}/N=o(1)$ *then*

$$\sqrt{T}(\bar{b}-\beta)-\sqrt{T}(\widehat{b}-\beta)=o_p(1) \tag{88}$$

**Proof.** We need to prove that

$$\sqrt{T}\left(\left(X'FS_{f\epsilon}^{-1}F'X\right)^{-1}X'FS_{f\epsilon}^{-1}F'\epsilon-\left(X'\hat{F}\widehat{S}_{\hat{f}\epsilon}^{-1}\hat{F}'X\right)^{-1}X'\hat{F}\widehat{S}_{\hat{f}\epsilon}^{-1}\hat{F}'\epsilon\right)=o_p(1) \tag{89}$$

or

$$\left(\frac{X'F}{T}S_{f\epsilon}^{-1}\frac{F'X}{T}\right)^{-1}\frac{X'F}{T}S_{f\epsilon}^{-1}\frac{F'\epsilon}{T^{1/2}}- \tag{90}$$
$$\left(\frac{X'\hat{F}}{T}\widehat{S}_{\hat{f}\epsilon}^{-1}\frac{\hat{F}'X}{T}\right)^{-1}\frac{X'\hat{F}}{T}\widehat{S}_{\hat{f}\epsilon}^{-1}\frac{\hat{F}'\epsilon}{T^{1/2}}=o_p(1)$$

From the proof of Theorem 4 we already know that

$$\frac{X'F}{T}-\frac{X'\hat{F}}{T}=o_p(1) \tag{91}$$

and

$$\sqrt{T}\left(\frac{F'\epsilon}{T}-\frac{\hat{F}'\epsilon}{T}\right)=o_p(1). \tag{92}$$

Then we have,

$$\widehat{S}_{\widehat{f}\epsilon,h}=\widehat{\Phi}_0+\sum_{j=1}^{h}(1-\frac{j}{h+1})(\widehat{\Phi}_j+\widehat{\Phi}_j')$$
$$\widehat{\Phi}_j=T^{-1}\sum_{T=j+1}^{T}\widehat{\epsilon}_t^2\widehat{f}_t\widehat{f}_t',$$

so that

$$\Phi_j-\widehat{\Phi}_j=T^{-1}\sum_{T=j+1}^{T}\widehat{\epsilon}_t^2\left(f_tf_t'-\widehat{f}_t\widehat{f}_t'\right).$$

The theorem is complete if we show formally that $\Phi_j-\widehat{\Phi}_j=o_p(h^{-1})$. We have

$$\left\|T^{-1}\sum_{T=j+1}^{T}\widehat{\epsilon}_t^2\left(f_tf_t'-\widehat{f}_t\widehat{f}_t'\right)\right\|\leq C_1\left\|T^{-1}\sum_{T=j+1}^{T}\widehat{\epsilon}_t^2f_t'\left(Hf_t-\widehat{f}_t'\right)\right\|\leq$$

$$C_2\left\|T^{-1}\sum_{T=j+1}^{T}\epsilon_t^2f_t'\left(Hf_t-\widehat{f}_t'\right)\right\|+C_3\left\|T^{-1}\sum_{T=j+1}^{T}(\widehat{\epsilon}_t-\epsilon_t)f_t'\left(Hf_t-\widehat{f}_t'\right)\right\|$$

for some constants $C_1$,$C_2$ and $C_3$. But, by (55) and $\sqrt{T}/N=o(1)$, $\left\|T^{-1}\sum_{T=j+1}^{T}\epsilon_t^2f_t'\left(Hf_t-\widehat{f}_t'\right)\right\|=o_p\left(T^{-1/2}\right)$ as long as $\epsilon_t$ has finite eighth moments. Then,

$$\left\|T^{-1}\sum_{T=j+1}^{T}(\widehat{\epsilon}_t-\epsilon_t)f_t'\left(Hf_t-\widehat{f}_t'\right)\right\|\leq C_4\left\|T^{-1}(b-\beta)\sum_{T=j+1}^{T}x_tf_t'\left(Hf_t-\widehat{f}_t'\right)\right\|$$

$$\leq C_5\left\|b-\beta\right\|\left\|T^{-1}\sum_{T=j+1}^{T}x_tf_t'\left(Hf_t-\widehat{f}_t'\right)\right\|$$

for some constants $C_4$ and $C_5$. Again by (55) and $\sqrt{T}/N=o(1)$, $\left\|T^{-1}\sum_{T=j+1}^{T}x_tf_t'\left(Hf_t-\widehat{f}_t'\right)\right\|=o_p\left(T^{-1/2}\right)$. By consistency of $b$, $\|b-\beta\|=o_p(1)$. Hence $\Phi_j-\widehat{\Phi}_j=o_p(h^{-1})$ as long as $h=o(T^{1/2})$. ■

Next, we have

**Theorem 10** *Under the same assumptions of Theorem 5, under (3),*

$$T^{1/2-\alpha}\left(\widehat{b}-\beta\right)\xrightarrow{d}N\left(0,\left(\Upsilon'\Sigma_fS_{f\epsilon}^{-1}\Sigma_f\Upsilon'\right)^{-1}\right) \tag{93}$$

*Under (4)*

$$T^{1/2-\alpha}\left(\widehat{b}-\beta\right)\xrightarrow{d}N\left(0,\left(\Psi'\Sigma_f S_{f\epsilon}^{-1}\Sigma_f\Psi\right)^{-1}\right) \tag{94}$$

*and under (5), if $\vartheta > \alpha$*

$$T^{1/2-\alpha}\left(\widehat{b}-\beta\right)\xrightarrow{d}N\left(0,\left(\Psi'\Sigma_f S_{f\epsilon}^{-1}\Sigma_f\Psi\right)^{-1}\right) \tag{95}$$

*if $\vartheta < \alpha$,*

$$T^{1/2-\alpha}\left(\widehat{b}-\beta\right)\xrightarrow{d}N\left(0,\left(\Upsilon'\Sigma_f S_{f\epsilon}^{-1}\Sigma_f\Upsilon'\right)^{-1}\right) \tag{96}$$

*and if $\vartheta = \alpha$,*

$$T^{1/2-\alpha}\left(\hat{\beta}-\beta\right)\xrightarrow{d}N\left(0,\left((\Psi+\Upsilon)'\Sigma_f S_{f\epsilon}^{-1}\Sigma_f(\Psi+\Upsilon)\right)^{-1}\right) \tag{97}$$

*where*

$$\lim_{T\to\infty}\frac{\Lambda^0_{N(T)}A^0_{N(T)Z}}{T^{\alpha}}=\Upsilon \tag{98}$$

$$\lim_{T\to\infty}\frac{A^0_{N(T)}}{T^{\vartheta}}=\Psi \tag{99}$$

*and $\Upsilon$ and $\Psi$ are nonsingular matrices.*

**Proof.** It follows from the proof of theorem 5, given Theorem 8 and consistency of $\widehat{S}_{\hat{f}\epsilon}$. ■

Finally, note that that the results of Theorem 6 follow straightforwardly for the GMM case.

# 3 Monte Carlo Study

This section presents a detailed Monte Carlo study of the relative performance of the standard and factor IV estimator. We consider all setups and estimators we have proposed in Section 2. The setup is as follows:

$$y_t=\sum_{i=1}^{k}x_{it}+\epsilon_t \tag{100}$$

$$z_{it}=\sum_{j=1}^{r}N^{-p}f_{jt}+c_2e_{it},\quad i=1,...,N \tag{101}$$

$$x_{it}=\sum_{j=1}^{N}N^{-1/2}z_{jt}+u_{it},\quad i=1,...,n_x \tag{102}$$

$$x_{it}=\sum_{j=1}^{r}c_1^{-1}r^{-1/2}f_{jt}+u_{it},\quad i=1,...,n_x \tag{103}$$

and

$$x_{it} = \sum_{j=1}^{N} N^{-1/2} z_{jt} + \sum_{j=1}^{r} c_1^{-1} r^{-1/2} f_{jt} + u_{it}, \quad i = 1, ..., n_x \tag{104}$$

where $k = 1$, $r = 1$, $e_{it} \sim i.i.d.N(0,1)$, $f_{it} \sim i.i.d.N(0,1)$ and $cov(e_{it}, e_{sj}) = 0$ for $i \neq s$. Let $\kappa_t = (\epsilon_t, u_{1t}, ..., u_{kt})'$. Then, $\kappa_t = P\eta_t$, where $\eta_t = (\eta_{1,t}, ..., \eta_{k+1,t})'$, $\eta_{i,t} \sim i.i.d.N(0,1)$ and $P = [p_{ij}]$, $p_{ij} \sim i.i.d.N(0,1)$. We do not consider heteroskedastic and/or correlated errors because we want to compare the standard and Factor-IV estimators without the possible complications arising from estimation of the HAC variance covariance matrix of the errors. Note that $\sum_{j=1}^{N} N^{-1/2} z_{jt}$ and $\sum_{j=1}^{r} c_1^{-1} r^{-1/2} f_{jt}$ are both $O_p(1)$ as $N, T \to \infty$. Hence, there is no difference in the weakness of the instruments between (102) and (103). This is separate from the issue of a local-to-zero factor model which in our Monte Carlo is explored via the choice of $p$ in (101)

For the setup corresponding to (3) we use (100), (101) and (102) where $c_2 = 0.5, 1, 4$. This is the framework where the standard IV estimator should perform well, at least for limited values of $N$. However, the Factor-IV estimator should also produce reasonable results, since the factors are a proxy for the $z$ variables, and be better than the standard IV estimator for large values of $N$. A larger value of $c_2$ corresponds to a larger idiosyncratic component for the $z$ variables, so that the factors provide a worse approximation for the $z$ variables.

For the setup corresponding to (4) we use (100), (101) and (103) where $c_2 = 1$ and $c_1 = 0.5, 1, 4$. In this case the Factor-IV estimator should be systematically better than the standard IV-estimator, at least as long as $p$ and $c_1$ are not too large. The parameter $c_1$ measures the "strength" of the factors as instruments, which decreases for higher values of $c_1$. Instead, the parameter $p$ controls the "strength" of the factors in the factor model for $z_t$, which decreases for higher values of $p$.

Finally, for the setup corresponding to (5) we use (100), (101) and (104) where $c_2 = 0.5, 1, 4$ and $c_1 = 0.5, 1, 4$. This is perhaps the most interesting case for empirical applications, since the endogenous regressors depend both on the $z_t$ variables and on the factors.

In all cases $p = 0, 0.1, 0.25, 0.33, 0.45, 0.5$, and we consider the following combinations of $N$ and $T$: $N = (30, 50, 100, 200)$, $T = (30, 50, 100, 200)$.[3] Results on the variance of the standard and Factor-IV estimators are presented in Tables 1-6. As expected, the biases

[3]Notice that when $N > T$ we use generalized inverses in the computation of the first step of the standard (two-stages) IV-estimator.

of the alternative estimators are small for all cases, detailed tables are available upon request.

Results make interesting reading. Looking first at the results for the experiment using (4), where the endogenous variables depend on the factors, we note that the performance of the standard IV estimator improves with $T$, but the extent of that improvement diminishes with $N$, as expected since the estimator is adversely affected by large $N$ (see Table 1). When $c_1$ increases, the dependence between the endogenous variables and the factors decreases, as well as that between the variables and the $z_t$ that are a proxi for the factors. Therefore, the variance of the IV estimator increases. Similarly, when $p$ increases the link between the factors and $z_t$ is weaker, so that $z_t$ is a worse proxy for the factors and the variance of the standard IV estimator increases.

From Table 2, when $p = 0$, the Factor-IV estimator improves with the sample size $T$, and the degree of the improvement is not affected by $N$. Instead, as expected. the Factor-IV estimator is adversely affected by a increase in $p$ and $c_1$. In particular, notice that when $p$ is larger than 0.25, increasing $N$ increases the variance of the estimator, which remains large even for $T = 200$. Actually, from Theorem 2 we know that in this case the Factor-estimator may no longer be consistent. Paradoxically, in this case using $z_t$, which contains information on the true factors $f_t$, is better than estimating $f_t$ with the principal component based estimator. When $p = 0$, in line with Theorem 5, the Factor-IV estimator is better in terms of variance for all values of $c_1$ and combinations of $N$ and $T$ (except, $T = 30$ and $c_1 = 4$, namely, in a very small sample). However, when $p$ increases, the sample size $T$ must be larger and larger for the Factor-IV to have a lower variance than standard IV estimator for large values of $c_1$. The worst results in terms of variance are for $c_1 = 4$ and $p = 0.5$, namely, a weak link between the endogenous variables and the factors, combined with a weak link between the $z_t$ variables and the factors. In this case the standard IV estimator has a smaller variance also for $T = 200$ when $N$ is large.

Overall, the Factor-IV estimator is better than the standard IV estimator as long as the factor model remains identified. When the parameter $p$ increases, the performance of the estimator of the factor $(\widehat{f_t})$ deteriorates, causing a larger variance for the factor-IV estimator of the parameters of the structural equation. A weaker link between the endogenous variables and the factors (large $c_1$) also increases the variance of the factor-IV estimator, but more so for the standard IV estimator.

Moving on to the experiment using (3), we note from Table 3 that the overall performance of the standard IV estimator improves, as expected since now the endogenous variables depend on $z_t$. Three additional features that emerge are the following. First, in general, when $p$ increases the variance of the standard IV estimator increases. This could seem surprising, since $p$ is a parameter of the factor model. However, when $p$ increases, the variance of each of the $z$ variables decreases, and therefore the explanatory power of equation (102) for $x_t$ decreases. Second, when $c_2$ increases and $p$ is large, the variance of the standard IV estimator decreases. This is because a larger $c_2$ increases the variance of $z$, but the effect is visible only when the common part $N^{-p} f_{jt}$ is small.

The performance of the Factor-IV estimator is evaluated in Table 4. The main finding is that, as long as the parameter $p$ is zero or small (namely there is a proper factor structure), the Factor-GMM estimator is even better than the standard IV-estimator, with the difference shrinking with the sample size $T$, as expected from the theory. For intermediate values of $p$, the parameter $c_2$ becomes relevant, and the Factor-IV is preferable only for small values of $c_2$ (small idiosyncratic component in factor model, so that the factors are a good proxy for all the $z$ variables). Finally, when $p$ is large, the standard-IV estimator becomes the best.

The results obtained for the experiment using (5), reported in Tables 5 and 6, are in line with those observed in the simpler cases given by (3) and (4). In particular, as long as $p$ is small, the Factor-IV estimator has a lower variance that the standard IV estimator for virtually any value of $c_1$, $c_2$, $N$ and $T$. When $p$ is larger than 0.25, the $c_1$ parameter becomes relevant and the Factor-IV estimator is better only for low values of $c_1$ (relatively large role of factors compared to the $z$ variables in explaining the endogenous variable). When $p$ is large, $0.45 - 0.50$, the N dimension becomes also relevant, in particular for limited sample size T, and the Factor-IV estimator is worse for large values of $N$ (this is the case where the additional variables in the $z$ set have only limited information on the factors).

The next issue we evaluate is variable preselection (i.e. selecting the variables that enter the factor analysis), since it may be conducive to better results in various modelling situations such as forecasting macroeconomic variables (see Boivin and Ng (2006)). To assess whether such a procedure may have some relevance to our work, we consider the setup of equation (4) but, prior to using the instruments $z_t$ either for standard or factor IV estimation, we preselect the 50% of the instruments with the highest correlation with $x_t$

(since we consider experiments with one $x_t$ variable only). Then, we carry out standard and factor IV estimation as usual. Results for a subset of the experiments ($c_1 = 0.5$ only) are reported in Tables 7-8.

Standard IV estimation is significantly improved when instrument preselection occurs. This is, of course, intuitive as the best instrument are retained. Factor IV estimation improves as well, and to a larger extent than standard IV estimation in many cases. The improvement is most apparent for high values of $p$. When $p$ is high, variable preselection plays a double role: it selects instruments correlated with the target, but because of this the selected instruments are also more correlated among themselves and therefore will likely present a stronger factor structure. Overall, factor IV estimation remains superior to standard IV, as long as the factor model remains identified.

Finally, we consider a different setup where 50% of the instruments are generated by (101) with $p$ varying in the range $0 - 0.5$, whereas the remaing instruments are generated by (101) with $p = 0$. Results for both the case where $z_t$ are preselected and where preselection is not undertaken, are reported in Tables 9-12 for $c_1 = 0.5$.

For the standard IV estimation, preselection once more helps improving the efficiency of the estimator. For the factor IV with preselection, it is interesting to note that the variances reported in Table 10 are much smaller than those in Table 8 when $p > 0.25$. However, the gains are smaller, though still systematic, when compared with the case without preselection in Table 12. This pattern arises because the factors can still be well estimated by the 50% of instruments with a well defined factor structure, so that variable preselection plays only a minor role in terms of strengthening the factor stucture. If a smaller fraction of the instruments presents a well defined factor structure, variable preselection plays a larger role in reducing the variance of the factor-IV estimator, as we will also see in the empirical applications.

In summary, the simulation results indicate that, as long as there is a well-defined factor structure, the Factor-IV estimator is preferable to the standard IV-estimator, even when the endogenous variable depends on the instrumental variables rather than on the factors. When the factor structure is loose, other parameters, such as the size of the idiosyncratic component in the factor model or the parameters in the equation for the endogenous variable, become important. Finally, when the factor structure is very weak, the standard IV estimator has a lower variance than the Factor-IV estimator. A comparable ranking can be expected for the GMM case, since HAC estimation of the variance covariance matrix of the residuals should have a similar impact on the efficiency of the standard and Factor-GMM estimators.

# 4 Empirical Applications

In this Section we discuss two empirical applications of the factor GMM estimation. The former concerns estimation of a forward looking Taylor rule, along the lines of Clarida, Galí, and Gertler (1998) (CGG), Clarida, Galí, and Gertler (2000) (CGG2)) and Favero, Marcellino, and Neglia (2005). The latter focuses on estimation of a New-Keynesian Phillips curve, along the lines of Galí and Gertler (1999) (GG 1999) and Beyer, Farmer, Henry, and Marcellino (2005).

For the Taylor rule, we adopt the following specification :

$$r_t = \alpha + (1-\rho)\beta(\pi_{t+12} - \pi_t^*) + (1-\rho)\gamma(y_t - y_t^*) + \rho r_{t-1} + \epsilon_t, \tag{105}$$

where $\epsilon_t = (1-\rho)\beta(\pi_{t+12}^e - \pi_{t+12}) + v_t$, and $v_t$ is an i.i.d. error. We use the federal funds rate for $r_t$, annual cpi inflation for $\pi_t$, 2% as a measure of the inflation target $\pi_t^*$, and the potential output $y_t^*$ is the Hodrick Prescott filtered version of the IP series.

Estimation of equation (105) presents several problems. First $\pi_{t+12}$ is correlated with the error term $\epsilon_t$. Second, the error term is correlated over time. In particular, under correct specification of the model, $\epsilon_t$ should be an MA(11) since it contains the forecast error $\pi_{t+12}^e - \pi_{t+12}$. Finally, the output gap is likely measured with error, so that $y_t - y_t^*$ can be also correlated with the error term. These problems can be handled by GMM estimation, with a correction for the MA component in the error $\epsilon_t$ and a proper choice of instruments.

In particular, we use a HAC estimator for the weighting matrix, based on a Bartlett kernel with Newey and West (1994) automatic bandwith selection. For the set of instruments, in the base case the choice is similar to that in CGG and CGG2. We use one lag of the output gap, inflation, commodity price index, unemployment and interest rate. Then we also include factors extracted from a large dataset of macroeconomic and financial variables for the US, the same used in Stock and Watson (2005) that contains 132 time series for the period 1959-2003. If the factors contain useful information, more precise estimates of the parameters should be obtained, as we have seen from a theoretical point of view and in the Monte Carlo simulations.

We focus on the period 1986-2003, since Beyer, Farmer, Henry, and Marcellino (2005) have detected instability in Phillips curves and Taylor rules estimated on a longer sample with an earlier start date. We consider factors estimated in three ways. First, from the whole dataset (All data). Second. from subsets of nominal, real and financial variables (Split data). Third, from variables pre-selected with the Boivin and Ng (2006) criterion (Select data). Pre-selecting the variables has a double effect in this context. First, by only retaining series related to the endogenous variable it attenuates the weak instrument problem. Second, the similarity among the retained series can be expected to be higher than that among all the series, so that the common component can be expected to be dominant with respect to the idiosyncratic component, which decreases problems of a weak factor structure.

The number of factors is determined by the Bai and Ng (2002) criteria, that suggest 8 factors from all data, 2 for nominal and financial variables, and 8 for real variables. Hence, the largest variability in the data seems to be in the real series. For the Select data we use just one factor. Actually, in this case only 13 variables remain in the dataset, those with a correlation with yearly inflation higher than 0.4 in absolute value, so that the Bai and Ng (2002) criteria that have a large $N$ justification cannot be used. However, one factor already explains more than 60% of the variability of these 13 variables. We use one lag of each factor for All data and Split data, 12 lags for the single factor from Select data. Adding additional lags of the variables or factors is not helpful in this application.

The results are reported in Table 13. For the base case, the estimated values for $\beta$ and $\gamma$ are, respectively, about 2.3 and 1, and the fact that the output gap matters less than inflation is not surprising. The persistence parameter, $\rho$, is about 0.88, in line with other studies. An LM test for the null hypothesis of no correlation in the residuals of an MA(11) model for $\widehat{\epsilon}_t$ does not reject the null hypothesis, which provides evidence in favor of the correct dynamic specification of the Taylor rule in (105). The p-value of the J-statistic for instrument validity is 0.11, so that the null hypothesis is not rejected at the conventional level of 10%.

Adding the factors to the instrument set does not improve the precision of the estimators of $\rho$, $\gamma$ and $\beta$ when using All data or Split Data. However, when the Select data factor is used, there is a reduction in the variance of the estimator of $\rho$ of about 20%, and of about 4% and 30% for $\gamma$ and $\beta$. With this set of factors there are also no significant changes in the parameter estimates, while there is an improvement in the p-value of the J-statistic. It is also worth mentioning that when only factors are used as instruments, the precision of the GMM estimators decreases substantially, which suggests that a combined use of key macro variables and factors is the optimal solution. Finally, a regression of future (12 months ahead) inflation on the instruments indicates that each set of factors is significant at the 10% level when added to the macro variables.

For the second example, the New-Keynesian Phillips curve is specified as,

$$\pi_t = c + \gamma\pi_{t+1} + \alpha x_t + \rho\pi_{t-1} + \epsilon_t, \tag{106}$$

where $\epsilon_t = \gamma(\pi^e_{t+1} - \pi_{t+1}) + v_t$, and $v_t$ is an i.i.d. error. Moreover, $\pi_t$ is annual CPI inflation, $\pi^e_{t+1}$ is the forecast of $\pi_{t+1}$ made in period $t$, and $x_t$ is a real forcing variable (unemployment, with reference to Okun's law, as in e.g. Beyer and Farmer (2003). [4]

As for the Taylor rule, $\pi_{t+1}$ is correlated with the error term $\epsilon_t$, which in turn is correlated over time. Hence, we estimate the parameters of (106) by GMM, with a correction for the MA component in the error $\epsilon_t$, and the same four sets of instruments as for the Taylor rule (but using the second lag of inflation).

The results are reported in Table 14. For the base case, the coefficient of the forcing variable is not statistically significant (though it has the correct sign), while the coefficients of the backward and forward looking components of inflation, $\rho$ and $\gamma$, are similar and close to 0.5. Adding the factors to the instrument set improves the precision of the estimators of all parameters, with the best results again from the Select data factors. For the latter, the gains are about 10% for $\alpha$ and 120% for $\gamma$ and $\rho$. Moreover, a regression of future (1 month ahead) inflation on the instruments indicates that only the Select data factors are strongly significant when added to the set of macroeconomic regressors.

Since the number of variables in the "Select data" set is relatively small, they could be directly used as additional instrumental variables instead of the factor that drives all of them. However, it turns out that the resulting parameter estimators are substantially less efficient than the factor-GMM estimators for both the Taylor rule and the hybrid Phillips curve.

In summary, these two examples confirm the relevance of factors as additional instruments for GMM estimation. Moreover, and in line with the results for forecasting, variable pre-selection appears to be relevant for the extraction of the factors to be used as (additional) instruments in GMM estimation.

[4] The results are qualitatively similar using the output gap.

# 5 Conclusions

The use of factor models has become very popular in the last few years, following the seminal work of Stock and Watson (2002) and Forni, Hallin, Lippi, and Reichlin (2000). Paralleling the developments in the VAR literature in the '80s and '90s, so far factor models have been mainly used for reduced form modelling and forecasting. However, recently there has a been an interest in more structural applications of factor analysis. Stock and Watson (2005), Giannone, Reichlin, and Sala (2002) and Kapetanios and Marcellino (2006) have shown that it is possible to obtain more realistic impulse response functions in a structural factor model. Favero, Marcellino, and Neglia (2005) and Beyer, Farmer, Henry, and Marcellino (2005) have estimated structural forward looking equations, such as those typically encountered in DSGE models, by means of factor augmented GMM estimation.

In this paper, and in a related independent article by Bai and Ng (2006b), we develop the theoretical underpinnings of Factor-GMM estimation. We show that when the endogenous variables in a structural equation are explained by a set of unobservable factors, which are also the driving forces of a larger set of instrumental variables, using the estimated factors as instuments rather than the large set of instrumental variables yields sizeable efficiency gains. Bai and Ng (2006b) show that a similar finding remains valid in a system framework, and the same would be true for our methodology.

We then extend the basic results in two directions. First, we evaluate what happens when the endogenous variables depend on a large set of instrumental variables rather than on the factors, or on a combination of them. We show theoretically that in this case the ranking of the standard and Factor-IV estimators is no longer clear-cut, since it depends on the specific parameter values. However, in an extensive set of simulation experiments, we have found that Factor-IV estimation seems to be more efficient also in this context.

Second, we evaluate what happens when either the factor structure is weak, or the instruments are weak, or both. When the factor structure is weak, the by now standard principal component based estimators of the factors are no longer consistent, basically because the factor model is no longer identified. However, we show that these factor estimators remain consistent even if the factor loadings in the factor model converge to zero, but at a sufficiently slow rate as a function of $N$. In this case, it is still possible to use Factor-IV estimators with well defined asymptotic properties.

When the instruments are weak, it is also possible to derive standard and Factor-IV estimators with well defined asymptotic properties, when the parameters in the equation that relates the instruments (or the factors) to the endogenous variables converge to zero at a sufficiently slow rate.

Both types of "weaknesses", in the factor structure and/or in the instruments, imply a slower convergence rate of the instrumental variable estimators. The simulation experiments indicate that, at least in our designs, a weak factor structure is more relevant than a weak instrument situation. Moreover, in the presence of a well defined factor structure but with weak instruments, Factor-IV estimation is in general more efficient than standard IV estimation, intuitively because the information in a large set of weak instruments in condensed in just a few variables.

Finally, we have applied Factor-GMM for the estimation of a Taylor rule and of a hybrid Phillips curve for the US, using factors extracted from a large set of macroeconomic variables. The findings confirm the empirical relevance of the theoretical results in this paper, in particular when the instrumental variables are pre-selected in a first stage, based on their correlation with the endogenous variable(s). Variable pre-selection can in fact alleviate both the weak instrument problem, since only instruments correlated with the target variable(s) are retained, and the weak factor structure problem, since more homogeneous variables are retained. In such a context, the gains from Factor-IV estimation with respect to standard-IV estimation can be fully exploited.

# References

Bai, J., and S. Ng (2002): "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70, 191–221.

——— (2006a): "Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions," *Econometrica*, 74, 1133–1150.

——— (2006b): "Instrumental Variable Estimation in a Data Rich Environment," *Mimeo.*

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

Beyer, A., and R. Farmer (2003): "On the indeterminacy of New-Keynesian economics," *ECB Working Paper No. 323.*

Beyer, A., R. Farmer, J. Henry, and M. Marcellino (2005): "Factor analysis in a new-Keynesian model," *ECB Working Paper No. 510.*

Boivin, J., and S. Ng (2006): "Are More Data Always Better for Factor Analysis?," *Journal of Econometrics*, 127, 169–194.

Chao, J. C., and N. R. Swanson (2005): "Consistent Estimation with a Large Number of Weak Instruments," *Econometrica*, 73, 1673–1692.

Clarida, R., J. Galí, and M. Gertler (1998): "Monetary policy rules in practice: Some international evidence," *European Economic Review*, 42, 1033–1067.

——— (2000): "Monetary policy rules and macroeconomic stability: evidence and some theory," *Quarterly Journal of Economics*, 115, 147–180.

Davidson, J. (1994): *Stochastic Limit Theory.* Oxford University Press.

Favero, C., M. Marcellino, and F. Neglia (2005): "Principal components at work: the empirical analysis of monetary policy with large datasets," *Journal of Applied Econometrics*, Forthcoming.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000): "The Generalised Factor Model: Identification and Estimation," *Review of Economics and Statistics*, 82, 540–554.

——— (2004): "The Generalised Factor Model: Consistency and Rates," *Journal of Econometrics*, Forthcoming.

Galí, J., and M. Gertler (1999): "Inflation Dynamics: A structural econometric approach," *Journal of Monetary Economics*, 44, 195–222.

Giannone, D., L. Reichlin, and L. Sala (2002): "Tracking Greenspan: systematic and unsystematic monetary policy revisited," *CEPR Working Paper No. 3550.*

Han, C., and P. C. B. Phillips (2006): "GMM with Many Moment Conditions," *Econometrica*, 74, 147–182.

Kapetanios, G., and M. Marcellino (2006): "Impulse response functions from structural dynamic factor models: A Monte Carlo evaluation," *CEPR Working Paper No. 5621.*

Lewbel, A. (1991): "The Rank of Demand Systems: Theory and Nonparametric Estimation," *Econometrica*, 59, 711–730.

Lutkepohl, H. (1996): *Handbook of Matrices.* Wiley.

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., and K. West (1994): "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631–653.

Newey, W., and K. D. West (1987): "A Simple, Positive Semi-Definite Heteroscedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Onatski, A. (2006): "Asymptotic distribution of the principal components estimator of large factor models when factors are relatively weak," *mimeo.*

Schwarz, H., H. R. Rutishauser, and E. Stiefel (1973): *Numerical Analysis of Symmetric Matrices.* Prentice Hall.

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

Stock, J. H., and M. W. Watson (1989): "New Indices of Coincident and Leading Indicators," in *NBER Macroeconomics Annual 1989*, ed. by O. J. Blanchard, and S. Fischer. Cambridge, M.I.T. Press.

——— (2002): "Macroeconomic Forecasting Using Diffusion Indices," *Journal of Business and Economic Statistics*, 20, 147–162.

——— (2005): "Implications of dynamic factor models for VAR analysis," *mimeo*.

# 6 Table Appendix

## 6.1 Setup of equation (4)

Table 1. Variance of standard IV estimator

| | | $c_1$ | 0.5 | | | | 1 | | | | 4 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $r$ | $p$ | T/N | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 |
| | 0 | 30 | 0.251 | 0.252 | 0.255 | 0.256 | 0.372 | 0.371 | 0.376 | 0.371 | 0.576 | 0.579 | 0.601 | 0.586 |
| | 0 | 50 | 0.179 | 0.246 | 0.251 | 0.248 | 0.284 | 0.370 | 0.371 | 0.368 | 0.522 | 0.577 | 0.574 | 0.580 |
| | 0 | 100 | 0.112 | 0.154 | 0.239 | 0.243 | 0.182 | 0.248 | 0.365 | 0.370 | 0.405 | 0.484 | 0.584 | 0.570 |
| | 0 | 200 | 0.066 | 0.089 | 0.147 | 0.243 | 0.108 | 0.152 | 0.241 | 0.356 | 0.285 | 0.356 | 0.476 | 0.575 |
| | 0.1 | 30 | 0.251 | 0.250 | 0.249 | 0.250 | 0.366 | 0.381 | 0.371 | 0.374 | 0.580 | 0.575 | 0.579 | 0.579 |
| | 0.1 | 50 | 0.181 | 0.242 | 0.241 | 0.252 | 0.277 | 0.366 | 0.371 | 0.367 | 0.528 | 0.573 | 0.576 | 0.573 |
| | 0.1 | 100 | 0.109 | 0.150 | 0.234 | 0.239 | 0.181 | 0.240 | 0.358 | 0.356 | 0.412 | 0.491 | 0.568 | 0.574 |
| | 0.1 | 200 | 0.069 | 0.091 | 0.149 | 0.236 | 0.111 | 0.161 | 0.240 | 0.348 | 0.291 | 0.370 | 0.472 | 0.581 |
| | 0.25 | 30 | 0.256 | 0.250 | 0.262 | 0.253 | 0.371 | 0.372 | 0.368 | 0.381 | 0.579 | 0.584 | 0.591 | 0.574 |
| | 0.25 | 50 | 0.187 | 0.251 | 0.245 | 0.248 | 0.289 | 0.362 | 0.366 | 0.366 | 0.533 | 0.575 | 0.575 | 0.566 |
| | 0.25 | 100 | 0.120 | 0.161 | 0.242 | 0.237 | 0.193 | 0.259 | 0.364 | 0.357 | 0.415 | 0.485 | 0.567 | 0.563 |
| 1 | 0.25 | 200 | 0.073 | 0.097 | 0.154 | 0.247 | 0.122 | 0.159 | 0.245 | 0.356 | 0.303 | 0.379 | 0.487 | 0.567 |
| | 0.33 | 30 | 0.257 | 0.253 | 0.261 | 0.255 | 0.378 | 0.374 | 0.373 | 0.364 | 0.572 | 0.585 | 0.589 | 0.575 |
| | 0.33 | 50 | 0.192 | 0.251 | 0.246 | 0.247 | 0.295 | 0.365 | 0.362 | 0.361 | 0.538 | 0.561 | 0.568 | 0.574 |
| | 0.33 | 100 | 0.123 | 0.165 | 0.240 | 0.238 | 0.201 | 0.266 | 0.359 | 0.363 | 0.423 | 0.490 | 0.564 | 0.567 |
| | 0.33 | 200 | 0.077 | 0.103 | 0.158 | 0.241 | 0.127 | 0.174 | 0.248 | 0.359 | 0.314 | 0.389 | 0.497 | 0.572 |
| | 0.45 | 30 | 0.259 | 0.258 | 0.253 | 0.254 | 0.369 | 0.374 | 0.371 | 0.374 | 0.582 | 0.576 | 0.589 | 0.596 |
| | 0.45 | 50 | 0.201 | 0.248 | 0.254 | 0.251 | 0.318 | 0.362 | 0.366 | 0.363 | 0.540 | 0.580 | 0.588 | 0.572 |
| | 0.45 | 100 | 0.140 | 0.182 | 0.240 | 0.243 | 0.227 | 0.279 | 0.356 | 0.355 | 0.463 | 0.508 | 0.568 | 0.567 |
| | 0.45 | 200 | 0.089 | 0.125 | 0.169 | 0.237 | 0.148 | 0.195 | 0.273 | 0.359 | 0.364 | 0.425 | 0.508 | 0.564 |
| | 0.5 | 30 | 0.251 | 0.252 | 0.257 | 0.251 | 0.368 | 0.371 | 0.385 | 0.373 | 0.585 | 0.579 | 0.568 | 0.580 |
| | 0.5 | 50 | 0.205 | 0.240 | 0.252 | 0.249 | 0.326 | 0.364 | 0.350 | 0.358 | 0.550 | 0.582 | 0.568 | 0.579 |
| | 0.5 | 100 | 0.155 | 0.189 | 0.240 | 0.244 | 0.238 | 0.292 | 0.359 | 0.361 | 0.476 | 0.526 | 0.567 | 0.567 |
| | 0.5 | 200 | 0.097 | 0.129 | 0.183 | 0.234 | 0.162 | 0.209 | 0.290 | 0.357 | 0.366 | 0.446 | 0.511 | 0.564 |

| Table 2. Variance of factor IV estimator | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $c_1$ | | 0.5 | | | | 1 | | | | 4 | | |
| $r$ | $p$ | T/N | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 |
| | 0 | 30 | 0.146 | 0.147 | 0.148 | 0.147 | 0.217 | 0.241 | 0.215 | 0.210 | 0.764 | 0.868 | 0.720 | 0.857 |
| | 0 | 50 | 0.105 | 0.107 | 0.101 | 0.106 | 0.156 | 0.143 | 0.162 | 0.161 | 0.438 | 0.528 | 0.474 | 0.433 |
| | 0 | 100 | 0.070 | 0.073 | 0.075 | 0.070 | 0.106 | 0.107 | 0.104 | 0.108 | 0.247 | 0.242 | 0.227 | 0.242 |
| | 0 | 200 | 0.052 | 0.050 | 0.052 | 0.050 | 0.072 | 0.074 | 0.074 | 0.073 | 0.154 | 0.160 | 0.153 | 0.149 |
| | 0.1 | 30 | 0.153 | 0.148 | 0.145 | 0.150 | 0.233 | 0.250 | 0.225 | 0.223 | 0.829 | 0.795 | 0.772 | 0.941 |
| | 0.1 | 50 | 0.109 | 0.110 | 0.113 | 0.108 | 0.163 | 0.162 | 0.163 | 0.154 | 0.463 | 0.468 | 0.452 | 0.411 |
| | 0.1 | 100 | 0.074 | 0.074 | 0.071 | 0.073 | 0.104 | 0.102 | 0.103 | 0.106 | 0.262 | 0.248 | 0.232 | 0.238 |
| | 0.1 | 200 | 0.051 | 0.053 | 0.054 | 0.050 | 0.075 | 0.074 | 0.072 | 0.075 | 0.162 | 0.151 | 0.151 | 0.147 |
| | 0.25 | 30 | 0.205 | 0.158 | 0.175 | 0.157 | 0.262 | 0.293 | 0.258 | 0.272 | 1.129 | 0.810 | 0.917 | 0.858 |
| | 0.25 | 50 | 0.116 | 0.116 | 0.122 | 0.119 | 0.168 | 0.162 | 0.173 | 0.174 | 0.548 | 0.468 | 0.436 | 0.497 |
| | 0.25 | 100 | 0.079 | 0.083 | 0.078 | 0.076 | 0.117 | 0.117 | 0.114 | 0.110 | 0.253 | 0.253 | 0.251 | 0.248 |
| 1 | 0.25 | 200 | 0.054 | 0.053 | 0.054 | 0.054 | 0.079 | 0.077 | 0.073 | 0.073 | 0.175 | 0.166 | 0.158 | 0.160 |
| | 0.33 | 30 | 0.303 | 0.238 | 0.296 | 0.365 | 0.409 | 0.391 | 0.487 | 0.480 | 1.105 | 0.975 | 1.132 | 1.228 |
| | 0.33 | 50 | 0.126 | 0.129 | 0.134 | 0.134 | 0.202 | 0.209 | 0.186 | 0.216 | 0.547 | 0.607 | 0.695 | 0.656 |
| | 0.33 | 100 | 0.089 | 0.085 | 0.082 | 0.085 | 0.125 | 0.119 | 0.117 | 0.122 | 0.345 | 0.283 | 0.311 | 0.281 |
| | 0.33 | 200 | 0.056 | 0.055 | 0.057 | 0.057 | 0.083 | 0.083 | 0.080 | 0.080 | 0.177 | 0.172 | 0.171 | 0.171 |
| | 0.45 | 30 | 0.800 | 1.021 | 1.169 | 1.348 | 1.053 | 1.184 | 1.436 | 1.565 | 1.753 | 1.711 | 1.872 | 1.775 |
| | 0.45 | 50 | 0.433 | 0.678 | 0.878 | 1.125 | 0.787 | 0.872 | 1.195 | 1.355 | 1.347 | 1.459 | 1.787 | 1.866 |
| | 0.45 | 100 | 0.108 | 0.211 | 0.388 | 0.698 | 0.160 | 0.249 | 0.579 | 0.944 | 0.489 | 0.672 | 0.936 | 1.375 |
| | 0.45 | 200 | 0.071 | 0.069 | 0.073 | 0.157 | 0.100 | 0.105 | 0.110 | 0.148 | 0.242 | 0.329 | 0.245 | 0.376 |
| | 0.5 | 30 | 1.134 | 1.237 | 1.353 | 1.451 | 1.217 | 1.468 | 1.570 | 1.796 | 1.758 | 1.975 | 1.994 | 1.919 |
| | 0.5 | 50 | 0.749 | 1.004 | 1.475 | 1.562 | 0.833 | 1.103 | 1.429 | 1.780 | 1.816 | 1.592 | 1.764 | 1.938 |
| | 0.5 | 100 | 0.361 | 0.611 | 0.846 | 1.183 | 0.484 | 0.703 | 1.103 | 1.355 | 0.964 | 1.210 | 1.430 | 1.872 |
| | 0.5 | 200 | 0.080 | 0.106 | 0.346 | 0.738 | 0.123 | 0.182 | 0.508 | 0.840 | 0.274 | 0.418 | 0.869 | 1.203 |

## 6.2 Setup of equation (3)

Table 3. Variance of standard IV estimator

| | | $c_2$ | 0.5 | | | | 1 | | | | 4 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $r$ | $p$ | T/N | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 |
| | 0 | 30 | 0.041 | 0.030 | 0.020 | 0.013 | 0.039 | 0.029 | 0.020 | 0.014 | 0.037 | 0.030 | 0.021 | 0.013 |
| | 0 | 50 | 0.029 | 0.023 | 0.016 | 0.011 | 0.029 | 0.024 | 0.016 | 0.011 | 0.027 | 0.023 | 0.016 | 0.011 |
| | 0 | 100 | 0.019 | 0.016 | 0.012 | 0.008 | 0.020 | 0.015 | 0.012 | 0.008 | 0.018 | 0.015 | 0.012 | 0.009 |
| | 0 | 200 | 0.013 | 0.010 | 0.008 | 0.006 | 0.013 | 0.011 | 0.008 | 0.006 | 0.013 | 0.010 | 0.008 | 0.006 |
| | 0.1 | 30 | 0.061 | 0.047 | 0.034 | 0.024 | 0.060 | 0.045 | 0.034 | 0.024 | 0.053 | 0.044 | 0.033 | 0.024 |
| | 0.1 | 50 | 0.042 | 0.041 | 0.027 | 0.020 | 0.040 | 0.040 | 0.027 | 0.020 | 0.038 | 0.036 | 0.027 | 0.019 |
| | 0.1 | 100 | 0.029 | 0.025 | 0.022 | 0.016 | 0.026 | 0.025 | 0.022 | 0.016 | 0.025 | 0.022 | 0.022 | 0.015 |
| | 0.1 | 200 | 0.019 | 0.016 | 0.014 | 0.013 | 0.019 | 0.016 | 0.014 | 0.013 | 0.017 | 0.015 | 0.013 | 0.012 |
| | 0.25 | 30 | 0.120 | 0.105 | 0.082 | 0.065 | 0.113 | 0.098 | 0.079 | 0.061 | 0.087 | 0.078 | 0.068 | 0.056 |
| | 0.25 | 50 | 0.083 | 0.094 | 0.074 | 0.057 | 0.078 | 0.090 | 0.073 | 0.055 | 0.060 | 0.070 | 0.058 | 0.049 |
| | 0.25 | 100 | 0.050 | 0.055 | 0.068 | 0.052 | 0.050 | 0.052 | 0.066 | 0.051 | 0.040 | 0.043 | 0.052 | 0.043 |
| 1 | 0.25 | 200 | 0.034 | 0.034 | 0.038 | 0.049 | 0.030 | 0.032 | 0.036 | 0.047 | 0.025 | 0.026 | 0.031 | 0.040 |
| | 0.33 | 30 | 0.173 | 0.154 | 0.136 | 0.108 | 0.157 | 0.142 | 0.123 | 0.104 | 0.105 | 0.099 | 0.092 | 0.082 |
| | 0.33 | 50 | 0.117 | 0.145 | 0.124 | 0.107 | 0.106 | 0.138 | 0.114 | 0.097 | 0.073 | 0.089 | 0.085 | 0.076 |
| | 0.33 | 100 | 0.069 | 0.085 | 0.120 | 0.103 | 0.066 | 0.079 | 0.111 | 0.092 | 0.045 | 0.054 | 0.078 | 0.071 |
| | 0.33 | 200 | 0.045 | 0.050 | 0.067 | 0.097 | 0.040 | 0.047 | 0.062 | 0.090 | 0.029 | 0.032 | 0.044 | 0.067 |
| | 0.45 | 30 | 0.266 | 0.253 | 0.251 | 0.234 | 0.223 | 0.225 | 0.213 | 0.212 | 0.133 | 0.126 | 0.130 | 0.126 |
| | 0.45 | 50 | 0.186 | 0.248 | 0.241 | 0.233 | 0.159 | 0.215 | 0.208 | 0.202 | 0.091 | 0.120 | 0.119 | 0.117 |
| | 0.45 | 100 | 0.109 | 0.151 | 0.236 | 0.228 | 0.091 | 0.128 | 0.201 | 0.195 | 0.053 | 0.069 | 0.113 | 0.110 |
| | 0.45 | 200 | 0.068 | 0.089 | 0.140 | 0.223 | 0.060 | 0.074 | 0.122 | 0.196 | 0.035 | 0.042 | 0.063 | 0.108 |
| | 0.5 | 30 | 0.303 | 0.298 | 0.301 | 0.296 | 0.257 | 0.258 | 0.253 | 0.257 | 0.137 | 0.132 | 0.137 | 0.134 |
| | 0.5 | 50 | 0.212 | 0.293 | 0.297 | 0.293 | 0.179 | 0.246 | 0.252 | 0.245 | 0.092 | 0.130 | 0.127 | 0.132 |
| | 0.5 | 100 | 0.128 | 0.182 | 0.285 | 0.294 | 0.109 | 0.149 | 0.237 | 0.240 | 0.059 | 0.074 | 0.123 | 0.125 |
| | 0.5 | 200 | 0.081 | 0.113 | 0.182 | 0.286 | 0.065 | 0.090 | 0.148 | 0.239 | 0.036 | 0.044 | 0.071 | 0.117 |

Table 4. Variance of factor IV estimator

| | | $c_2$ | 0.5 | | | | 1 | | | | 4 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $r$ | $p$ | T/N | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 |
| | 0 | 30 | 0.035 | 0.027 | 0.019 | 0.013 | 0.034 | 0.027 | 0.019 | 0.014 | 0.033 | 0.028 | 0.020 | 0.013 |
| | 0 | 50 | 0.027 | 0.020 | 0.014 | 0.010 | 0.026 | 0.021 | 0.015 | 0.010 | 0.025 | 0.020 | 0.014 | 0.010 |
| | 0 | 100 | 0.019 | 0.014 | 0.010 | 0.007 | 0.019 | 0.014 | 0.010 | 0.007 | 0.018 | 0.014 | 0.010 | 0.007 |
| | 0 | 200 | 0.013 | 0.010 | 0.007 | 0.005 | 0.013 | 0.010 | 0.007 | 0.005 | 0.012 | 0.010 | 0.007 | 0.005 |
| | 0.1 | 30 | 0.050 | 0.039 | 0.031 | 0.023 | 0.046 | 0.040 | 0.031 | 0.023 | 0.052 | 0.041 | 0.030 | 0.023 |
| | 0.1 | 50 | 0.037 | 0.030 | 0.022 | 0.017 | 0.035 | 0.030 | 0.023 | 0.017 | 0.035 | 0.028 | 0.022 | 0.017 |
| | 0.1 | 100 | 0.027 | 0.021 | 0.016 | 0.012 | 0.024 | 0.021 | 0.016 | 0.012 | 0.024 | 0.019 | 0.016 | 0.012 |
| | 0.1 | 200 | 0.018 | 0.015 | 0.011 | 0.009 | 0.017 | 0.015 | 0.011 | 0.009 | 0.016 | 0.014 | 0.011 | 0.008 |
| | 0.25 | 30 | 0.079 | 0.074 | 0.061 | 0.051 | 0.080 | 0.070 | 0.060 | 0.050 | 0.359 | 0.630 | 0.401 | 0.525 |
| | 0.25 | 50 | 0.062 | 0.053 | 0.045 | 0.039 | 0.059 | 0.052 | 0.046 | 0.038 | 0.173 | 0.220 | 0.131 | 0.095 |
| | 0.25 | 100 | 0.043 | 0.038 | 0.031 | 0.027 | 0.042 | 0.036 | 0.031 | 0.027 | 0.051 | 0.047 | 0.031 | 0.035 |
| 1 | 0.25 | 200 | 0.030 | 0.025 | 0.022 | 0.018 | 0.027 | 0.025 | 0.022 | 0.019 | 0.025 | 0.023 | 0.021 | 0.017 |
| | 0.33 | 30 | 0.106 | 0.100 | 0.088 | 0.081 | 0.302 | 0.144 | 0.139 | 0.150 | 0.843 | 0.786 | 0.743 | 0.833 |
| | 0.33 | 50 | 0.083 | 0.074 | 0.065 | 0.059 | 0.078 | 0.074 | 0.085 | 0.125 | 0.592 | 0.567 | 0.705 | 0.744 |
| | 0.33 | 100 | 0.054 | 0.049 | 0.045 | 0.040 | 0.051 | 0.045 | 0.042 | 0.040 | 0.249 | 0.314 | 0.401 | 0.589 |
| | 0.33 | 200 | 0.038 | 0.035 | 0.031 | 0.028 | 0.035 | 0.033 | 0.030 | 0.028 | 0.092 | 0.054 | 0.073 | 0.106 |
| | 0.45 | 30 | 0.300 | 0.457 | 0.505 | 0.731 | 0.757 | 0.715 | 1.172 | 1.175 | 1.043 | 1.318 | 1.174 | 1.242 |
| | 0.45 | 50 | 0.118 | 0.266 | 0.198 | 0.330 | 0.378 | 0.470 | 0.684 | 0.874 | 0.764 | 1.100 | 1.044 | 1.168 |
| | 0.45 | 100 | 0.076 | 0.078 | 0.074 | 0.297 | 0.091 | 0.165 | 0.238 | 0.613 | 0.807 | 0.709 | 0.929 | 1.169 |
| | 0.45 | 200 | 0.052 | 0.051 | 0.049 | 0.050 | 0.048 | 0.048 | 0.054 | 0.101 | 0.503 | 0.631 | 0.750 | 0.919 |
| | 0.5 | 30 | 0.455 | 0.824 | 0.967 | 1.288 | 0.896 | 1.206 | 1.343 | 1.462 | 1.121 | 1.144 | 1.107 | 1.367 |
| | 0.5 | 50 | 0.189 | 0.305 | 0.514 | 1.133 | 0.520 | 0.853 | 0.984 | 1.219 | 0.989 | 0.981 | 1.305 | 1.205 |
| | 0.5 | 100 | 0.089 | 0.090 | 0.188 | 0.363 | 0.331 | 0.313 | 0.583 | 1.023 | 0.774 | 0.785 | 1.215 | 1.107 |
| | 0.5 | 200 | 0.058 | 0.061 | 0.065 | 0.082 | 0.056 | 0.111 | 0.208 | 0.551 | 0.625 | 0.684 | 0.850 | 1.092 |

## 6.3 Setup of equation (5)

Table 5. Variance of standard IV estimator

| | | | $c_2$ | 0.5 | | | | 1 | | | | 4 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $r$ | $p$ | $c_1$ | T/N | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 |
| | 0 | 0.5 | 30 | 0.031 | 0.025 | 0.017 | 0.013 | 0.029 | 0.024 | 0.017 | 0.012 | 0.030 | 0.024 | 0.017 | 0.012 |
| | 0 | 0.5 | 50 | 0.022 | 0.020 | 0.014 | 0.010 | 0.022 | 0.020 | 0.014 | 0.010 | 0.022 | 0.020 | 0.014 | 0.010 |
| | 0 | 0.5 | 100 | 0.015 | 0.013 | 0.010 | 0.007 | 0.015 | 0.013 | 0.010 | 0.007 | 0.015 | 0.013 | 0.010 | 0.007 |
| | 0 | 0.5 | 200 | 0.010 | 0.009 | 0.007 | 0.006 | 0.011 | 0.009 | 0.006 | 0.005 | 0.010 | 0.009 | 0.007 | 0.005 |
| | 0 | 1 | 30 | 0.032 | 0.025 | 0.018 | 0.013 | 0.032 | 0.024 | 0.018 | 0.013 | 0.032 | 0.025 | 0.018 | 0.013 |
| | 0 | 1 | 50 | 0.023 | 0.021 | 0.014 | 0.010 | 0.024 | 0.021 | 0.014 | 0.010 | 0.024 | 0.020 | 0.014 | 0.010 |
| | 0 | 1 | 100 | 0.016 | 0.013 | 0.011 | 0.007 | 0.016 | 0.014 | 0.011 | 0.007 | 0.016 | 0.013 | 0.011 | 0.008 |
| | 0 | 1 | 200 | 0.011 | 0.009 | 0.007 | 0.006 | 0.011 | 0.009 | 0.007 | 0.006 | 0.011 | 0.009 | 0.007 | 0.006 |
| | 0 | 4 | 30 | 0.035 | 0.026 | 0.018 | 0.013 | 0.034 | 0.026 | 0.019 | 0.013 | 0.034 | 0.026 | 0.018 | 0.013 |
| | 0 | 4 | 50 | 0.024 | 0.021 | 0.015 | 0.011 | 0.025 | 0.022 | 0.015 | 0.011 | 0.024 | 0.021 | 0.015 | 0.011 |
| | 0 | 4 | 100 | 0.017 | 0.014 | 0.011 | 0.008 | 0.016 | 0.014 | 0.011 | 0.008 | 0.016 | 0.014 | 0.012 | 0.008 |
| 1 | 0 | 4 | 200 | 0.011 | 0.009 | 0.007 | 0.006 | 0.012 | 0.009 | 0.008 | 0.006 | 0.011 | 0.009 | 0.007 | 0.006 |
| | 0.1 | 0.5 | 30 | 0.043 | 0.033 | 0.026 | 0.020 | 0.041 | 0.034 | 0.026 | 0.021 | 0.042 | 0.033 | 0.026 | 0.020 |
| | 0.1 | 0.5 | 50 | 0.031 | 0.028 | 0.022 | 0.016 | 0.031 | 0.028 | 0.022 | 0.016 | 0.030 | 0.028 | 0.021 | 0.016 |
| | 0.1 | 0.5 | 100 | 0.019 | 0.018 | 0.018 | 0.013 | 0.020 | 0.018 | 0.017 | 0.013 | 0.020 | 0.018 | 0.017 | 0.013 |
| | 0.1 | 0.5 | 200 | 0.013 | 0.013 | 0.011 | 0.010 | 0.014 | 0.012 | 0.011 | 0.010 | 0.014 | 0.012 | 0.011 | 0.010 |
| | 0.1 | 1 | 30 | 0.044 | 0.037 | 0.027 | 0.022 | 0.046 | 0.037 | 0.028 | 0.021 | 0.045 | 0.037 | 0.028 | 0.021 |
| | 0.1 | 1 | 50 | 0.032 | 0.031 | 0.023 | 0.018 | 0.032 | 0.031 | 0.023 | 0.017 | 0.032 | 0.030 | 0.023 | 0.017 |
| | 0.1 | 1 | 100 | 0.021 | 0.019 | 0.018 | 0.013 | 0.021 | 0.020 | 0.018 | 0.014 | 0.021 | 0.020 | 0.019 | 0.014 |
| | 0.1 | 1 | 200 | 0.014 | 0.013 | 0.011 | 0.011 | 0.015 | 0.013 | 0.012 | 0.011 | 0.014 | 0.013 | 0.012 | 0.011 |
| | 0.1 | 4 | 30 | 0.047 | 0.038 | 0.029 | 0.022 | 0.047 | 0.040 | 0.030 | 0.021 | 0.048 | 0.039 | 0.030 | 0.022 |
| | 0.1 | 4 | 50 | 0.035 | 0.033 | 0.024 | 0.018 | 0.036 | 0.033 | 0.024 | 0.018 | 0.033 | 0.032 | 0.024 | 0.018 |
| | 0.1 | 4 | 100 | 0.023 | 0.020 | 0.020 | 0.014 | 0.022 | 0.020 | 0.020 | 0.015 | 0.022 | 0.021 | 0.019 | 0.014 |
| | 0.1 | 4 | 200 | 0.015 | 0.014 | 0.013 | 0.011 | 0.015 | 0.014 | 0.012 | 0.012 | 0.015 | 0.014 | 0.012 | 0.012 |
| | 0.25 | 0.5 | 30 | 0.065 | 0.059 | 0.049 | 0.044 | 0.063 | 0.057 | 0.050 | 0.040 | 0.066 | 0.060 | 0.048 | 0.044 |
| | 0.25 | 0.5 | 50 | 0.044 | 0.053 | 0.044 | 0.038 | 0.045 | 0.049 | 0.043 | 0.034 | 0.045 | 0.051 | 0.042 | 0.036 |
| | 0.25 | 0.5 | 100 | 0.029 | 0.030 | 0.038 | 0.031 | 0.029 | 0.032 | 0.039 | 0.031 | 0.030 | 0.031 | 0.037 | 0.031 |
| | 0.25 | 0.5 | 200 | 0.020 | 0.020 | 0.022 | 0.028 | 0.019 | 0.020 | 0.022 | 0.027 | 0.020 | 0.020 | 0.021 | 0.028 |
| | 0.25 | 1 | 30 | 0.073 | 0.063 | 0.056 | 0.048 | 0.072 | 0.065 | 0.053 | 0.045 | 0.075 | 0.066 | 0.056 | 0.048 |
| | 0.25 | 1 | 50 | 0.051 | 0.059 | 0.048 | 0.040 | 0.051 | 0.057 | 0.048 | 0.041 | 0.050 | 0.060 | 0.048 | 0.040 |
| | 0.25 | 1 | 100 | 0.033 | 0.034 | 0.042 | 0.034 | 0.032 | 0.034 | 0.044 | 0.034 | 0.032 | 0.034 | 0.042 | 0.036 |
| | 0.25 | 1 | 200 | 0.021 | 0.022 | 0.025 | 0.031 | 0.021 | 0.022 | 0.026 | 0.033 | 0.021 | 0.022 | 0.024 | 0.032 |
| | 0.25 | 4 | 30 | 0.073 | 0.067 | 0.055 | 0.048 | 0.072 | 0.069 | 0.056 | 0.050 | 0.073 | 0.066 | 0.055 | 0.048 |
| | 0.25 | 4 | 50 | 0.050 | 0.058 | 0.050 | 0.041 | 0.052 | 0.062 | 0.051 | 0.041 | 0.051 | 0.061 | 0.051 | 0.041 |
| | 0.25 | 4 | 100 | 0.032 | 0.037 | 0.044 | 0.037 | 0.035 | 0.035 | 0.044 | 0.037 | 0.033 | 0.036 | 0.044 | 0.038 |
| 1 | 0.25 | 4 | 200 | 0.023 | 0.022 | 0.027 | 0.034 | 0.021 | 0.022 | 0.026 | 0.034 | 0.021 | 0.022 | 0.025 | 0.034 |
| | 0.33 | 0.5 | 30 | 0.081 | 0.073 | 0.067 | 0.063 | 0.077 | 0.074 | 0.069 | 0.061 | 0.081 | 0.074 | 0.067 | 0.061 |
| | 0.33 | 0.5 | 50 | 0.056 | 0.070 | 0.059 | 0.056 | 0.055 | 0.069 | 0.063 | 0.055 | 0.055 | 0.066 | 0.061 | 0.055 |
| | 0.33 | 0.5 | 100 | 0.036 | 0.041 | 0.056 | 0.049 | 0.036 | 0.040 | 0.056 | 0.051 | 0.036 | 0.040 | 0.056 | 0.049 |
| | 0.33 | 0.5 | 200 | 0.023 | 0.025 | 0.031 | 0.046 | 0.023 | 0.025 | 0.031 | 0.047 | 0.023 | 0.025 | 0.032 | 0.046 |
| | 0.33 | 1 | 30 | 0.092 | 0.085 | 0.078 | 0.071 | 0.093 | 0.086 | 0.076 | 0.070 | 0.093 | 0.087 | 0.078 | 0.069 |
| | 0.33 | 1 | 50 | 0.062 | 0.080 | 0.073 | 0.063 | 0.063 | 0.079 | 0.069 | 0.064 | 0.065 | 0.079 | 0.070 | 0.062 |
| | 0.33 | 1 | 100 | 0.040 | 0.047 | 0.063 | 0.056 | 0.040 | 0.045 | 0.064 | 0.057 | 0.039 | 0.046 | 0.066 | 0.055 |
| | 0.33 | 1 | 200 | 0.027 | 0.028 | 0.036 | 0.055 | 0.026 | 0.028 | 0.036 | 0.053 | 0.027 | 0.029 | 0.036 | 0.054 |

Continued on next page

| continued from previous page | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $c_2$ | | 0.5 | | | | 1 | | | | 4 | | |
| $r$ | $p$ | $c_1$ | T/N | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 |
| | 0.33 | 4 | 30 | 0.088 | 0.085 | 0.079 | 0.071 | 0.089 | 0.084 | 0.077 | 0.070 | 0.092 | 0.085 | 0.075 | 0.071 |
| | 0.33 | 4 | 50 | 0.062 | 0.078 | 0.070 | 0.065 | 0.063 | 0.077 | 0.070 | 0.063 | 0.060 | 0.078 | 0.070 | 0.063 |
| | 0.33 | 4 | 100 | 0.037 | 0.046 | 0.064 | 0.058 | 0.038 | 0.044 | 0.065 | 0.058 | 0.038 | 0.046 | 0.065 | 0.057 |
| | 0.33 | 4 | 200 | 0.025 | 0.028 | 0.036 | 0.055 | 0.027 | 0.028 | 0.035 | 0.055 | 0.025 | 0.028 | 0.037 | 0.054 |
| | 0.45 | 0.5 | 30 | 0.105 | 0.103 | 0.100 | 0.102 | 0.107 | 0.104 | 0.101 | 0.098 | 0.103 | 0.105 | 0.100 | 0.098 |
| | 0.45 | 0.5 | 50 | 0.074 | 0.096 | 0.092 | 0.088 | 0.075 | 0.098 | 0.090 | 0.090 | 0.073 | 0.095 | 0.092 | 0.091 |
| | 0.45 | 0.5 | 100 | 0.046 | 0.060 | 0.087 | 0.086 | 0.047 | 0.058 | 0.089 | 0.086 | 0.046 | 0.055 | 0.091 | 0.087 |
| | 0.45 | 0.5 | 200 | 0.030 | 0.035 | 0.050 | 0.083 | 0.031 | 0.034 | 0.050 | 0.084 | 0.029 | 0.035 | 0.050 | 0.083 |
| | 0.45 | 1 | 30 | 0.122 | 0.121 | 0.123 | 0.120 | 0.124 | 0.119 | 0.118 | 0.117 | 0.124 | 0.124 | 0.121 | 0.119 |
| | 0.45 | 1 | 50 | 0.084 | 0.115 | 0.113 | 0.108 | 0.087 | 0.114 | 0.112 | 0.107 | 0.085 | 0.115 | 0.116 | 0.110 |
| | 0.45 | 1 | 100 | 0.054 | 0.068 | 0.106 | 0.102 | 0.053 | 0.069 | 0.106 | 0.102 | 0.052 | 0.069 | 0.104 | 0.103 |
| | 0.45 | 1 | 200 | 0.036 | 0.041 | 0.061 | 0.098 | 0.034 | 0.040 | 0.060 | 0.101 | 0.034 | 0.042 | 0.060 | 0.101 |
| | 0.45 | 4 | 30 | 0.112 | 0.108 | 0.108 | 0.104 | 0.109 | 0.108 | 0.105 | 0.104 | 0.111 | 0.109 | 0.108 | 0.105 |
| | 0.45 | 4 | 50 | 0.077 | 0.101 | 0.101 | 0.098 | 0.076 | 0.100 | 0.101 | 0.095 | 0.074 | 0.102 | 0.098 | 0.099 |
| | 0.45 | 4 | 100 | 0.047 | 0.060 | 0.092 | 0.093 | 0.046 | 0.058 | 0.095 | 0.089 | 0.047 | 0.060 | 0.093 | 0.091 |
| 1 | 0.45 | 4 | 200 | 0.030 | 0.036 | 0.052 | 0.088 | 0.031 | 0.035 | 0.053 | 0.088 | 0.029 | 0.036 | 0.053 | 0.089 |
| | 0.5 | 0.5 | 30 | 0.116 | 0.118 | 0.117 | 0.119 | 0.117 | 0.119 | 0.117 | 0.113 | 0.113 | 0.120 | 0.117 | 0.115 |
| | 0.5 | 0.5 | 50 | 0.080 | 0.107 | 0.110 | 0.107 | 0.082 | 0.107 | 0.110 | 0.108 | 0.079 | 0.112 | 0.107 | 0.108 |
| | 0.5 | 0.5 | 100 | 0.050 | 0.067 | 0.105 | 0.101 | 0.051 | 0.066 | 0.104 | 0.103 | 0.052 | 0.067 | 0.103 | 0.102 |
| | 0.5 | 0.5 | 200 | 0.034 | 0.041 | 0.060 | 0.097 | 0.034 | 0.039 | 0.060 | 0.101 | 0.034 | 0.039 | 0.061 | 0.099 |
| | 0.5 | 1 | 30 | 0.138 | 0.137 | 0.137 | 0.137 | 0.137 | 0.139 | 0.139 | 0.133 | 0.136 | 0.138 | 0.141 | 0.137 |
| | 0.5 | 1 | 50 | 0.096 | 0.130 | 0.129 | 0.129 | 0.096 | 0.128 | 0.126 | 0.131 | 0.095 | 0.127 | 0.125 | 0.128 |
| | 0.5 | 1 | 100 | 0.059 | 0.077 | 0.123 | 0.125 | 0.060 | 0.079 | 0.123 | 0.126 | 0.059 | 0.076 | 0.122 | 0.124 |
| | 0.5 | 1 | 200 | 0.040 | 0.047 | 0.073 | 0.119 | 0.039 | 0.048 | 0.074 | 0.122 | 0.038 | 0.047 | 0.072 | 0.120 |
| | 0.5 | 4 | 30 | 0.118 | 0.117 | 0.118 | 0.118 | 0.117 | 0.120 | 0.117 | 0.116 | 0.117 | 0.118 | 0.118 | 0.117 |
| | 0.5 | 4 | 50 | 0.081 | 0.108 | 0.110 | 0.110 | 0.077 | 0.109 | 0.108 | 0.106 | 0.082 | 0.110 | 0.109 | 0.109 |
| | 0.5 | 4 | 100 | 0.049 | 0.067 | 0.103 | 0.102 | 0.049 | 0.063 | 0.102 | 0.103 | 0.050 | 0.064 | 0.104 | 0.103 |
| | 0.5 | 4 | 200 | 0.033 | 0.039 | 0.059 | 0.099 | 0.033 | 0.040 | 0.059 | 0.098 | 0.032 | 0.038 | 0.058 | 0.100 |

| Table 6. Variance of factor IV estimator | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | $c_2$ | 0.5 | | | | 1 | | | | 4 | | | |
| $r$ | $p$ | $c_1$ | T/N | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 |
| | 0 | 0.5 | 30 | 0.028 | 0.023 | 0.016 | 0.012 | 0.028 | 0.023 | 0.017 | 0.012 | 0.028 | 0.023 | 0.017 | 0.012 |
| | 0 | 0.5 | 50 | 0.021 | 0.017 | 0.013 | 0.009 | 0.021 | 0.017 | 0.013 | 0.009 | 0.021 | 0.018 | 0.013 | 0.009 |
| | 0 | 0.5 | 100 | 0.015 | 0.012 | 0.009 | 0.006 | 0.015 | 0.012 | 0.009 | 0.006 | 0.014 | 0.012 | 0.009 | 0.007 |
| | 0 | 0.5 | 200 | 0.010 | 0.008 | 0.006 | 0.005 | 0.010 | 0.008 | 0.006 | 0.004 | 0.010 | 0.009 | 0.006 | 0.004 |
| | 0 | 1 | 30 | 0.030 | 0.024 | 0.017 | 0.013 | 0.029 | 0.024 | 0.017 | 0.013 | 0.029 | 0.023 | 0.017 | 0.013 |
| | 0 | 1 | 50 | 0.023 | 0.018 | 0.013 | 0.009 | 0.022 | 0.018 | 0.013 | 0.010 | 0.023 | 0.018 | 0.013 | 0.010 |
| | 0 | 1 | 100 | 0.016 | 0.013 | 0.009 | 0.007 | 0.015 | 0.013 | 0.009 | 0.006 | 0.015 | 0.012 | 0.010 | 0.007 |
| | 0 | 1 | 200 | 0.011 | 0.009 | 0.006 | 0.005 | 0.011 | 0.009 | 0.006 | 0.005 | 0.011 | 0.009 | 0.006 | 0.005 |
| | 0 | 4 | 30 | 0.031 | 0.024 | 0.017 | 0.013 | 0.031 | 0.025 | 0.018 | 0.013 | 0.030 | 0.025 | 0.017 | 0.013 |
| | 0 | 4 | 50 | 0.023 | 0.018 | 0.014 | 0.010 | 0.023 | 0.019 | 0.014 | 0.010 | 0.023 | 0.019 | 0.014 | 0.010 |
| | 0 | 4 | 100 | 0.017 | 0.013 | 0.010 | 0.007 | 0.016 | 0.013 | 0.009 | 0.007 | 0.016 | 0.013 | 0.010 | 0.007 |
| 1 | 0 | 4 | 200 | 0.011 | 0.009 | 0.007 | 0.005 | 0.012 | 0.009 | 0.007 | 0.005 | 0.011 | 0.009 | 0.007 | 0.005 |
| | 0.1 | 0.5 | 30 | 0.037 | 0.030 | 0.025 | 0.020 | 0.038 | 0.030 | 0.025 | 0.020 | 0.036 | 0.030 | 0.025 | 0.019 |
| | 0.1 | 0.5 | 50 | 0.028 | 0.023 | 0.019 | 0.015 | 0.027 | 0.024 | 0.019 | 0.015 | 0.028 | 0.024 | 0.018 | 0.015 |
| | 0.1 | 0.5 | 100 | 0.018 | 0.016 | 0.014 | 0.011 | 0.019 | 0.016 | 0.013 | 0.010 | 0.019 | 0.017 | 0.013 | 0.010 |
| | 0.1 | 0.5 | 200 | 0.013 | 0.012 | 0.009 | 0.007 | 0.013 | 0.011 | 0.009 | 0.007 | 0.013 | 0.011 | 0.009 | 0.007 |
| | 0.1 | 1 | 30 | 0.037 | 0.032 | 0.025 | 0.021 | 0.039 | 0.033 | 0.025 | 0.020 | 0.039 | 0.034 | 0.027 | 0.021 |
| | 0.1 | 1 | 50 | 0.028 | 0.024 | 0.020 | 0.016 | 0.030 | 0.024 | 0.020 | 0.016 | 0.030 | 0.025 | 0.020 | 0.016 |
| | 0.1 | 1 | 100 | 0.020 | 0.017 | 0.014 | 0.011 | 0.020 | 0.017 | 0.014 | 0.011 | 0.020 | 0.018 | 0.014 | 0.011 |
| | 0.1 | 1 | 200 | 0.014 | 0.013 | 0.010 | 0.008 | 0.014 | 0.012 | 0.010 | 0.008 | 0.014 | 0.012 | 0.010 | 0.008 |
| | 0.1 | 4 | 30 | 0.047 | 0.036 | 0.028 | 0.021 | 0.044 | 0.035 | 0.029 | 0.021 | 0.047 | 0.043 | 0.028 | 0.021 |
| | 0.1 | 4 | 50 | 0.032 | 0.027 | 0.021 | 0.016 | 0.032 | 0.026 | 0.021 | 0.017 | 0.030 | 0.027 | 0.021 | 0.016 |
| | 0.1 | 4 | 100 | 0.022 | 0.018 | 0.014 | 0.011 | 0.021 | 0.018 | 0.014 | 0.012 | 0.021 | 0.018 | 0.015 | 0.011 |
| | 0.1 | 4 | 200 | 0.015 | 0.012 | 0.010 | 0.008 | 0.015 | 0.013 | 0.010 | 0.008 | 0.014 | 0.013 | 0.010 | 0.008 |
| | 0.25 | 0.5 | 30 | 0.051 | 0.047 | 0.042 | 0.040 | 0.053 | 0.048 | 0.042 | 0.035 | 0.052 | 0.048 | 0.042 | 0.037 |
| | 0.25 | 0.5 | 50 | 0.039 | 0.037 | 0.032 | 0.028 | 0.039 | 0.036 | 0.032 | 0.028 | 0.040 | 0.037 | 0.032 | 0.028 |
| | 0.25 | 0.5 | 100 | 0.027 | 0.025 | 0.022 | 0.019 | 0.026 | 0.025 | 0.022 | 0.019 | 0.027 | 0.024 | 0.022 | 0.020 |
| | 0.25 | 0.5 | 200 | 0.019 | 0.017 | 0.016 | 0.014 | 0.018 | 0.017 | 0.016 | 0.014 | 0.019 | 0.017 | 0.015 | 0.014 |
| | 0.25 | 1 | 30 | 0.057 | 0.053 | 0.047 | 0.236 | 0.058 | 0.053 | 0.326 | 0.041 | 0.060 | 0.055 | 0.047 | 0.041 |
| | 0.25 | 1 | 50 | 0.044 | 0.038 | 0.034 | 0.031 | 0.044 | 0.039 | 0.035 | 0.031 | 0.043 | 0.040 | 0.034 | 0.030 |
| | 0.25 | 1 | 100 | 0.029 | 0.027 | 0.025 | 0.021 | 0.029 | 0.026 | 0.024 | 0.021 | 0.028 | 0.027 | 0.025 | 0.022 |
| | 0.25 | 1 | 200 | 0.020 | 0.019 | 0.017 | 0.014 | 0.021 | 0.019 | 0.017 | 0.015 | 0.020 | 0.018 | 0.016 | 0.014 |
| | 0.25 | 4 | 30 | 0.670 | 0.465 | 0.474 | 0.356 | 0.634 | 0.458 | 0.553 | 0.337 | 0.382 | 0.643 | 0.373 | 0.465 |
| | 0.25 | 4 | 50 | 0.268 | 0.224 | 0.229 | 0.150 | 0.187 | 0.238 | 0.217 | 0.127 | 0.162 | 0.239 | 0.325 | 0.087 |
| | 0.25 | 4 | 100 | 0.082 | 0.041 | 0.027 | 0.024 | 0.052 | 0.033 | 0.027 | 0.028 | 0.039 | 0.030 | 0.028 | 0.039 |
| 1 | 0.25 | 4 | 200 | 0.022 | 0.019 | 0.018 | 0.016 | 0.021 | 0.019 | 0.017 | 0.016 | 0.021 | 0.020 | 0.018 | 0.016 |
| | 0.33 | 0.5 | 30 | 0.061 | 0.059 | 0.055 | 0.054 | 0.061 | 0.059 | 0.056 | 0.054 | 0.061 | 0.058 | 0.053 | 0.051 |
| | 0.33 | 0.5 | 50 | 0.046 | 0.043 | 0.041 | 0.037 | 0.045 | 0.042 | 0.041 | 0.036 | 0.046 | 0.044 | 0.040 | 0.039 |
| | 0.33 | 0.5 | 100 | 0.032 | 0.031 | 0.028 | 0.026 | 0.032 | 0.029 | 0.028 | 0.027 | 0.032 | 0.031 | 0.027 | 0.026 |
| | 0.33 | 0.5 | 200 | 0.022 | 0.021 | 0.020 | 0.018 | 0.022 | 0.020 | 0.019 | 0.019 | 0.021 | 0.021 | 0.020 | 0.018 |
| | 0.33 | 1 | 30 | 0.275 | 0.163 | 0.109 | 0.091 | 0.135 | 0.135 | 0.352 | 0.166 | 0.116 | 0.154 | 0.142 | 0.100 |
| | 0.33 | 1 | 50 | 0.051 | 0.050 | 0.047 | 0.063 | 0.052 | 0.050 | 0.049 | 0.048 | 0.052 | 0.052 | 0.045 | 0.045 |
| | 0.33 | 1 | 100 | 0.035 | 0.033 | 0.032 | 0.028 | 0.035 | 0.033 | 0.032 | 0.029 | 0.036 | 0.033 | 0.031 | 0.028 |
| | 0.33 | 1 | 200 | 0.025 | 0.022 | 0.022 | 0.021 | 0.024 | 0.022 | 0.021 | 0.020 | 0.024 | 0.024 | 0.022 | 0.021 |
| | 0.33 | 4 | 30 | 0.687 | 0.763 | 0.681 | 0.961 | 0.714 | 0.686 | 0.924 | 1.009 | 0.680 | 0.821 | 0.719 | 0.681 |
| | 0.33 | 4 | 50 | 0.453 | 0.556 | 0.427 | 0.508 | 0.324 | 0.452 | 0.546 | 0.642 | 0.513 | 0.548 | 0.747 | 0.622 |
| | 0.33 | 4 | 100 | 0.258 | 0.406 | 0.362 | 0.360 | 0.194 | 0.301 | 0.538 | 0.437 | 0.361 | 0.465 | 0.452 | 0.478 |
| | 0.33 | 4 | 200 | 0.198 | 0.060 | 0.143 | 0.109 | 0.055 | 0.032 | 0.052 | 0.080 | 0.053 | 0.034 | 0.061 | 0.075 |

Continued on next page

continued from previous page

| | | | $c_2$ | 0.5 | | | | 1 | | | | 4 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $r$ | $p$ | $c_1$ | T/N | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 | 30 | 50 | 100 | 200 |
| | 0.45 | 0.5 | 30 | 0.137 | 0.191 | 0.453 | 0.476 | 0.154 | 0.160 | 0.389 | 0.529 | 0.095 | 0.190 | 0.341 | 0.524 |
| | 0.45 | 0.5 | 50 | 0.063 | 0.075 | 0.198 | 0.352 | 0.062 | 0.063 | 0.227 | 0.294 | 0.062 | 0.062 | 0.242 | 0.216 |
| | 0.45 | 0.5 | 100 | 0.040 | 0.040 | 0.041 | 0.042 | 0.040 | 0.042 | 0.039 | 0.042 | 0.040 | 0.039 | 0.042 | 0.043 |
| | 0.45 | 0.5 | 200 | 0.028 | 0.028 | 0.027 | 0.028 | 0.028 | 0.027 | 0.028 | 0.028 | 0.027 | 0.027 | 0.027 | 0.027 |
| | 0.45 | 1 | 30 | 0.490 | 0.700 | 0.970 | 1.010 | 0.504 | 0.581 | 0.704 | 0.932 | 0.485 | 0.652 | 0.884 | 0.960 |
| | 0.45 | 1 | 50 | 0.417 | 0.457 | 0.621 | 0.767 | 0.365 | 0.382 | 0.587 | 0.718 | 0.261 | 0.328 | 0.561 | 0.800 |
| | 0.45 | 1 | 100 | 0.153 | 0.114 | 0.257 | 0.427 | 0.049 | 0.070 | 0.200 | 0.524 | 0.054 | 0.071 | 0.224 | 0.370 |
| | 0.45 | 1 | 200 | 0.035 | 0.031 | 0.039 | 0.048 | 0.032 | 0.032 | 0.057 | 0.046 | 0.032 | 0.033 | 0.035 | 0.062 |
| | 0.45 | 4 | 30 | 1.015 | 1.127 | 0.994 | 1.262 | 0.895 | 1.082 | 1.265 | 1.170 | 0.902 | 1.003 | 1.061 | 1.369 |
| | 0.45 | 4 | 50 | 0.811 | 0.870 | 1.242 | 0.987 | 0.778 | 0.948 | 0.905 | 1.199 | 0.943 | 1.100 | 0.888 | 1.170 |
| | 0.45 | 4 | 100 | 0.666 | 0.839 | 0.865 | 0.885 | 0.682 | 0.572 | 0.952 | 1.029 | 0.559 | 0.764 | 0.691 | 1.061 |
| 1 | 0.45 | 4 | 200 | 0.343 | 0.434 | 0.688 | 0.758 | 0.456 | 0.608 | 0.639 | 0.876 | 0.267 | 0.592 | 0.730 | 0.926 |
| | 0.5 | 0.5 | 30 | 0.169 | 0.416 | 0.766 | 1.093 | 0.447 | 0.393 | 0.873 | 0.966 | 0.295 | 0.537 | 0.682 | 1.001 |
| | 0.5 | 0.5 | 50 | 0.086 | 0.208 | 0.439 | 0.673 | 0.108 | 0.208 | 0.423 | 0.561 | 0.110 | 0.429 | 0.241 | 0.701 |
| | 0.5 | 0.5 | 100 | 0.044 | 0.047 | 0.076 | 0.255 | 0.045 | 0.046 | 0.070 | 0.302 | 0.046 | 0.054 | 0.069 | 0.218 |
| | 0.5 | 0.5 | 200 | 0.031 | 0.032 | 0.033 | 0.041 | 0.031 | 0.031 | 0.033 | 0.039 | 0.032 | 0.031 | 0.032 | 0.037 |
| | 0.5 | 1 | 30 | 0.625 | 0.743 | 1.031 | 1.195 | 0.839 | 0.934 | 1.024 | 1.186 | 0.703 | 0.851 | 1.102 | 1.157 |
| | 0.5 | 1 | 50 | 0.497 | 0.769 | 0.980 | 1.162 | 0.646 | 0.621 | 0.942 | 1.045 | 0.522 | 0.652 | 0.830 | 1.074 |
| | 0.5 | 1 | 100 | 0.228 | 0.222 | 0.618 | 0.950 | 0.200 | 0.468 | 0.671 | 0.734 | 0.278 | 0.323 | 0.575 | 0.946 |
| | 0.5 | 1 | 200 | 0.039 | 0.065 | 0.324 | 0.609 | 0.036 | 0.047 | 0.096 | 0.477 | 0.037 | 0.110 | 0.124 | 0.443 |
| | 0.5 | 4 | 30 | 1.076 | 0.938 | 1.197 | 1.143 | 1.083 | 1.268 | 1.123 | 1.083 | 1.142 | 1.081 | 1.301 | 1.428 |
| | 0.5 | 4 | 50 | 0.828 | 0.983 | 1.062 | 1.247 | 0.868 | 1.013 | 1.256 | 1.247 | 0.997 | 1.200 | 1.263 | 1.217 |
| | 0.5 | 4 | 100 | 0.783 | 0.928 | 1.028 | 1.120 | 0.957 | 0.934 | 1.027 | 1.112 | 0.700 | 0.875 | 1.106 | 1.160 |
| | 0.5 | 4 | 200 | 0.474 | 0.604 | 0.861 | 0.960 | 0.655 | 0.541 | 0.761 | 1.062 | 0.430 | 0.857 | 0.951 | 1.017 |

## 6.4 Setup of equation (4) with homogeneous factor loadings and with variable preselection

| Table 7. Variance of standard IV estimator with preselection | | | | | | |
|---|---|---|---|---|---|---|
| $r$ | $p$ | T/N | 30 | 50 | 100 | 200 |
| | 0 | 30 | 0.170 | 0.230 | 0.263 | 0.260 |
| | 0 | 50 | 0.125 | 0.148 | 0.251 | 0.246 |
| | 0 | 100 | 0.085 | 0.104 | 0.141 | 0.241 |
| | 0 | 200 | 0.060 | 0.066 | 0.083 | 0.131 |
| | 0.1 | 30 | 0.182 | 0.230 | 0.256 | 0.254 |
| | 0.1 | 50 | 0.132 | 0.161 | 0.242 | 0.246 |
| | 0.1 | 100 | 0.092 | 0.105 | 0.148 | 0.236 |
| | 0.1 | 200 | 0.061 | 0.068 | 0.089 | 0.141 |
| | 0.25 | 30 | 0.192 | 0.239 | 0.253 | 0.248 |
| | 0.25 | 50 | 0.150 | 0.171 | 0.250 | 0.247 |
| | 0.25 | 100 | 0.099 | 0.120 | 0.160 | 0.241 |
| 1 | 0.25 | 200 | 0.071 | 0.080 | 0.108 | 0.155 |
| | 0.33 | 30 | 0.194 | 0.236 | 0.251 | 0.253 |
| | 0.33 | 50 | 0.152 | 0.188 | 0.244 | 0.249 |
| | 0.33 | 100 | 0.110 | 0.130 | 0.169 | 0.243 |
| | 0.33 | 200 | 0.080 | 0.090 | 0.116 | 0.164 |
| | 0.45 | 30 | 0.205 | 0.242 | 0.253 | 0.257 |
| | 0.45 | 50 | 0.171 | 0.199 | 0.242 | 0.250 |
| | 0.45 | 100 | 0.128 | 0.150 | 0.185 | 0.239 |
| | 0.45 | 200 | 0.095 | 0.111 | 0.136 | 0.178 |
| | 0.5 | 30 | 0.219 | 0.250 | 0.259 | 0.254 |
| | 0.5 | 50 | 0.177 | 0.203 | 0.252 | 0.245 |
| | 0.5 | 100 | 0.139 | 0.160 | 0.191 | 0.239 |
| | 0.5 | 200 | 0.098 | 0.118 | 0.147 | 0.190 |

| Table 8. Variance of Factor IV estimator with preselection | | | | | | |
|---|---|---|---|---|---|---|
| $r$ | $p$ | T/N | 30 | 50 | 100 | 200 |
| | 0 | 30 | 0.130 | 0.133 | 0.135 | 0.139 |
| | 0 | 50 | 0.105 | 0.103 | 0.109 | 0.099 |
| | 0 | 100 | 0.074 | 0.077 | 0.072 | 0.073 |
| | 0 | 200 | 0.055 | 0.054 | 0.053 | 0.053 |
| | 0.1 | 30 | 0.140 | 0.135 | 0.128 | 0.136 |
| | 0.1 | 50 | 0.107 | 0.107 | 0.102 | 0.104 |
| | 0.1 | 100 | 0.078 | 0.076 | 0.076 | 0.074 |
| | 0.1 | 200 | 0.054 | 0.054 | 0.052 | 0.055 |
| | 0.25 | 30 | 0.146 | 0.147 | 0.143 | 0.132 |
| | 0.25 | 50 | 0.116 | 0.112 | 0.117 | 0.114 |
| | 0.25 | 100 | 0.080 | 0.086 | 0.086 | 0.084 |
| 1 | 0.25 | 200 | 0.061 | 0.061 | 0.065 | 0.066 |
| | 0.33 | 30 | 0.158 | 0.156 | 0.153 | 0.149 |
| | 0.33 | 50 | 0.126 | 0.127 | 0.121 | 0.125 |
| | 0.33 | 100 | 0.092 | 0.094 | 0.096 | 0.097 |
| | 0.33 | 200 | 0.070 | 0.067 | 0.070 | 0.075 |
| | 0.45 | 30 | 0.246 | 0.490 | 0.281 | 0.397 |
| | 0.45 | 50 | 0.218 | 0.195 | 0.400 | 0.274 |
| | 0.45 | 100 | 0.112 | 0.112 | 0.116 | 0.121 |
| | 0.45 | 200 | 0.082 | 0.086 | 0.087 | 0.094 |
| | 0.5 | 30 | 0.417 | 0.644 | 0.549 | 0.817 |
| | 0.5 | 50 | 0.309 | 0.344 | 0.510 | 0.726 |
| | 0.5 | 100 | 0.161 | 0.266 | 0.374 | 0.244 |
| | 0.5 | 200 | 0.088 | 0.097 | 0.105 | 0.157 |

## 6.5 Setup of equation (4) with heterogeneous factor loadings and with or without variable preselection

| Table 9. Variance of standard IV estimator with preselection | | | | | | |
|---|---|---|---|---|---|---|
| $r$ | $p$ | T/N | 30 | 50 | 100 | 200 |
| | 0 | 30 | 0.172 | 0.226 | 0.257 | 0.257 |
| | 0 | 50 | 0.129 | 0.150 | 0.243 | 0.249 |
| | 0 | 100 | 0.086 | 0.097 | 0.138 | 0.243 |
| | 0 | 200 | 0.059 | 0.064 | 0.083 | 0.131 |
| | 0.1 | 30 | 0.174 | 0.229 | 0.252 | 0.253 |
| | 0.1 | 50 | 0.132 | 0.163 | 0.249 | 0.245 |
| | 0.1 | 100 | 0.085 | 0.103 | 0.150 | 0.239 |
| | 0.1 | 200 | 0.058 | 0.062 | 0.091 | 0.150 |
| | 0.25 | 30 | 0.179 | 0.231 | 0.255 | 0.255 |
| | 0.25 | 50 | 0.127 | 0.162 | 0.248 | 0.249 |
| | 0.25 | 100 | 0.082 | 0.098 | 0.148 | 0.243 |
| 1 | 0.25 | 200 | 0.056 | 0.064 | 0.090 | 0.151 |
| | 0.33 | 30 | 0.176 | 0.226 | 0.256 | 0.259 |
| | 0.33 | 50 | 0.126 | 0.166 | 0.251 | 0.248 |
| | 0.33 | 100 | 0.082 | 0.100 | 0.148 | 0.243 |
| | 0.33 | 200 | 0.053 | 0.061 | 0.090 | 0.147 |
| | 0.45 | 30 | 0.183 | 0.227 | 0.258 | 0.256 |
| | 0.45 | 50 | 0.131 | 0.163 | 0.242 | 0.246 |
| | 0.45 | 100 | 0.084 | 0.100 | 0.151 | 0.244 |
| | 0.45 | 200 | 0.055 | 0.063 | 0.093 | 0.146 |
| | 0.5 | 30 | 0.178 | 0.229 | 0.250 | 0.260 |
| | 0.5 | 50 | 0.131 | 0.162 | 0.249 | 0.241 |
| | 0.5 | 100 | 0.083 | 0.098 | 0.151 | 0.245 |
| | 0.5 | 200 | 0.055 | 0.061 | 0.089 | 0.146 |

| Table 10. Variance of Factor IV estimator with preselection | | | | | | |
|---|---|---|---|---|---|---|
| $r$ | $p$ | T/N | 30 | 50 | 100 | 200 |
| | 0 | 30 | 0.133 | 0.132 | 0.137 | 0.137 |
| | 0 | 50 | 0.106 | 0.100 | 0.098 | 0.102 |
| | 0 | 100 | 0.075 | 0.072 | 0.072 | 0.073 |
| | 0 | 200 | 0.053 | 0.053 | 0.052 | 0.050 |
| | 0.1 | 30 | 0.135 | 0.135 | 0.131 | 0.130 |
| | 0.1 | 50 | 0.108 | 0.106 | 0.100 | 0.101 |
| | 0.1 | 100 | 0.073 | 0.072 | 0.072 | 0.068 |
| | 0.1 | 200 | 0.051 | 0.049 | 0.050 | 0.049 |
| | 0.25 | 30 | 0.145 | 0.141 | 0.135 | 0.140 |
| | 0.25 | 50 | 0.102 | 0.103 | 0.105 | 0.102 |
| | 0.25 | 100 | 0.075 | 0.073 | 0.074 | 0.073 |
| 1 | 0.25 | 200 | 0.054 | 0.053 | 0.051 | 0.052 |
| | 0.33 | 30 | 0.144 | 0.135 | 0.136 | 0.137 |
| | 0.33 | 50 | 0.107 | 0.107 | 0.104 | 0.102 |
| | 0.33 | 100 | 0.074 | 0.072 | 0.072 | 0.072 |
| | 0.33 | 200 | 0.051 | 0.051 | 0.054 | 0.050 |
| | 0.45 | 30 | 0.142 | 0.138 | 0.136 | 0.143 |
| | 0.45 | 50 | 0.107 | 0.105 | 0.110 | 0.106 |
| | 0.45 | 100 | 0.075 | 0.076 | 0.074 | 0.075 |
| | 0.45 | 200 | 0.052 | 0.052 | 0.051 | 0.053 |
| | 0.5 | 30 | 0.143 | 0.140 | 0.137 | 0.132 |
| | 0.5 | 50 | 0.110 | 0.109 | 0.106 | 0.105 |
| | 0.5 | 100 | 0.078 | 0.076 | 0.076 | 0.070 |
| | 0.5 | 200 | 0.052 | 0.052 | 0.051 | 0.052 |

| Table 11. Variance of standard IV estimator without preselection | | | | | | |
|---|---|---|---|---|---|---|
| $r$ | $p$ | T/N | 30 | 50 | 100 | 200 |
| | 0 | 30 | 0.256 | 0.252 | 0.254 | 0.251 |
| | 0 | 50 | 0.184 | 0.246 | 0.246 | 0.248 |
| | 0 | 100 | 0.113 | 0.151 | 0.237 | 0.247 |
| | 0 | 200 | 0.069 | 0.090 | 0.145 | 0.236 |
| | 0.1 | 30 | 0.259 | 0.259 | 0.249 | 0.255 |
| | 0.1 | 50 | 0.177 | 0.251 | 0.245 | 0.252 |
| | 0.1 | 100 | 0.114 | 0.150 | 0.241 | 0.244 |
| | 0.1 | 200 | 0.067 | 0.090 | 0.143 | 0.235 |
| | 0.25 | 30 | 0.255 | 0.259 | 0.255 | 0.259 |
| | 0.25 | 50 | 0.176 | 0.245 | 0.248 | 0.245 |
| | 0.25 | 100 | 0.114 | 0.152 | 0.240 | 0.239 |
| 1 | 0.25 | 200 | 0.069 | 0.091 | 0.148 | 0.242 |
| | 0.33 | 30 | 0.251 | 0.258 | 0.257 | 0.252 |
| | 0.33 | 50 | 0.184 | 0.251 | 0.249 | 0.245 |
| | 0.33 | 100 | 0.109 | 0.149 | 0.245 | 0.246 |
| | 0.33 | 200 | 0.068 | 0.092 | 0.148 | 0.232 |
| | 0.45 | 30 | 0.256 | 0.258 | 0.251 | 0.253 |
| | 0.45 | 50 | 0.180 | 0.244 | 0.245 | 0.245 |
| | 0.45 | 100 | 0.115 | 0.153 | 0.243 | 0.246 |
| | 0.45 | 200 | 0.068 | 0.090 | 0.150 | 0.239 |
| | 0.5 | 30 | 0.256 | 0.258 | 0.256 | 0.260 |
| | 0.5 | 50 | 0.186 | 0.249 | 0.248 | 0.248 |
| | 0.5 | 100 | 0.108 | 0.153 | 0.238 | 0.242 |
| | 0.5 | 200 | 0.069 | 0.092 | 0.147 | 0.241 |

| Table 12. Variance of Factor IV estimator without preselection | | | | | | |
|---|---|---|---|---|---|---|
| $r$ | $p$ | T/N | 30 | 50 | 100 | 200 |
| | 0 | 30 | 0.147 | 0.153 | 0.151 | 0.147 |
| | 0 | 50 | 0.108 | 0.106 | 0.104 | 0.100 |
| | 0 | 100 | 0.073 | 0.073 | 0.071 | 0.074 |
| | 0 | 200 | 0.054 | 0.053 | 0.050 | 0.049 |
| | 0.1 | 30 | 0.149 | 0.142 | 0.154 | 0.146 |
| | 0.1 | 50 | 0.104 | 0.112 | 0.107 | 0.109 |
| | 0.1 | 100 | 0.075 | 0.075 | 0.073 | 0.074 |
| | 0.1 | 200 | 0.053 | 0.052 | 0.052 | 0.049 |
| | 0.25 | 30 | 0.145 | 0.149 | 0.148 | 0.147 |
| | 0.25 | 50 | 0.111 | 0.109 | 0.111 | 0.110 |
| | 0.25 | 100 | 0.074 | 0.074 | 0.073 | 0.075 |
| 1 | 0.25 | 200 | 0.052 | 0.053 | 0.051 | 0.052 |
| | 0.33 | 30 | 0.156 | 0.150 | 0.146 | 0.150 |
| | 0.33 | 50 | 0.111 | 0.106 | 0.106 | 0.111 |
| | 0.33 | 100 | 0.076 | 0.075 | 0.071 | 0.074 |
| | 0.33 | 200 | 0.054 | 0.052 | 0.051 | 0.052 |
| | 0.45 | 30 | 0.160 | 0.148 | 0.149 | 0.149 |
| | 0.45 | 50 | 0.109 | 0.108 | 0.108 | 0.108 |
| | 0.45 | 100 | 0.079 | 0.073 | 0.073 | 0.071 |
| | 0.45 | 200 | 0.052 | 0.049 | 0.052 | 0.049 |
| | 0.5 | 30 | 0.155 | 0.144 | 0.146 | 0.153 |
| | 0.5 | 50 | 0.110 | 0.106 | 0.107 | 0.103 |
| | 0.5 | 100 | 0.070 | 0.076 | 0.074 | 0.075 |
| | 0.5 | 200 | 0.055 | 0.052 | 0.052 | 0.052 |

## 6.6 Empirical Results

| Table 13. Results for Hybrid Phillips curve | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | **First stage regression (infl+1)** | | |
| | | $\alpha$ | $\gamma$ | $\rho$ | **R2-adj** | **S.E. regr** | **Pval J-stat** | **R2-adj** | **S.E. regr** | **Pval F-stat** |
| | | | | | | | | | | |
| **Base** | | -0.002 | 0.538 | 0.462 | 0.98 | 0.16 | 0.62 | 0.12 | 0.002 | |
| | st. err | 0.007 | 0.048 | 0.047 | | | | | | |
| | | | | | | | | | | |
| **Factors** | | 0.000 | 0.513 | 0.492 | 0.98 | 0.16 | 0.30 | 0.11 | 0.002 | 0.48 |
| **All data** | st. err | 0.006 | 0.038 | 0.038 | | | | | | |
| | | | | | | | | | | |
| **Factors** | | 0.003 | 0.527 | 0.478 | 0.98 | 0.16 | 0.28 | 0.14 | 0.002 | 0.25 |
| **Split data** | st. err | 0.006 | 0.038 | 0.037 | | | | | | |
| | | | | | | | | | | |
| **Factors** | | -0.002 | 0.500 | 0.509 | 0.98 | 0.15 | 0.12 | 0.23 | 0.002 | 0.00 |
| **Select data** | st. err | 0.006 | 0.021 | 0.020 | | | | | | |
| | | | | | | | | | | |

Notes: The estimated equation is $r_t = \alpha + (1-\rho)\beta(\pi_{t+12} - \pi_t^*) + (1-\rho)\gamma(y_t - y_t^*) + \rho r_{t-1} + \epsilon_t$ (see text for details). The parameters are estimated by GMM over 1986.01-2003.12. In the base case (no factors) the set of instruments used includes lags of the output gap, unemployment, inflation, interest rate and commodity price index. In the Factors cases, the SW factors are added to the instruments. In particular, in "All data" the (8) factors are extracted from the whole dataset; in "Split data" the factors are extracted from separate datasets for nominal (2), real (8) and financial variables (2); in "Select data" the (1) factor extracted from.a subset of the variables selected with the Boivin and Ng (2006) criterion. The number of factors is based on the Bai and Ng (2002) criteria for each dataset, except "Select data" where it is set to one. We use one lag of each factor, but 12 lags for the Select data factor. The last three columns contain statistics related to the first-stage regression of the one-year ahead expected inflation on the set of instruments used. In particular, we report the adjusted $R^2$, the standard error of the regression and the F-test for the joint significance of the coefficients on factors, when factors are added to the baseline model.

| Table 14. Results for Taylor rule | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | |
| | | | | | | | | **First stage regression (infl+12)** | | |
| | | $\rho$ | $\gamma$ | $\beta$ | **R2-adj** | **S.E. regr** | **Pval J-stat** | **R2-adj** | **S.E. regr** | **Pval F-stat** |
| | | | | | | | | | | |
| **Base** | | 0.883 | 0.993 | 2.310 | 0.98 | 0.27 | 0.11 | 0.12 | 0.002 | |
| | st. err | 0.037 | 0.241 | 0.278 | | | | | | |
| | | | | | | | | | | |
| **Factors** | | 0.908 | 1.261 | 2.905 | 0.98 | 0.27 | 0.13 | 0.15 | 0.002 | 0.05 |
| **All data** | st. err | 0.024 | 0.291 | 0.394 | | | | | | |
| | | | | | | | | | | |
| **Factors** | | 0.929 | 1.291 | 3.346 | 0.98 | 0.26 | 0.14 | 0.18 | 0.002 | 0.01 |
| **Split data** | st. err | 0.021 | 0.335 | 0.567 | | | | | | |
| | | | | | | | | | | |
| **Factors** | | 0.884 | 1.122 | 2.251 | 0.98 | 0.27 | 0.52 | 0.15 | 0.002 | 0.08 |
| **Select data** | st. err | 0.028 | 0.233 | 0.204 | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |

Notes: The estimated equation is $\pi_t = c + \alpha(ur_t) + \gamma(\pi_{t+1}) + \rho\pi_{t-1} + \epsilon_t$ (see text for details). The parameters are estimated by GMM over 1986.01-2003.12. In the base case (no factors) the set of instruments used includes lags of the output gap, unemployment, inflation, interest rate and commodity price index. In the Factors cases, the (first lag of the) SW factors are added to the instruments. In particular, in "All data" the (8) factors are extracted from the whole dataset; in "Split data" the factors are extracted from separate datasets for nominal (2), real (8) and financial variables (2); in "Select data" the (1) factor is extracted from.a subset of the variables selected with the Boivin and Ng (2006) criterion. The number of factors is based on the Bai and Ng (2002) criteria for each dataset, except "Select data" where it is set to one. We use one lag of each factor, but 12 lags for the Select data factor. The last three columns contain statistics related to the first-stage regression of the one-month ahead expected inflation on the set of instruments used. In particular, we report the adjusted $R^2$, the standard error of the regression and the F-test for the joint significance of the coefficients on factors, when factors are added to the baseline model.

Queen Mary
University of London

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