Factor based identification-robust inference in IV regressions*

G. Kapetanios
Queen Mary, University of London

L. Khalaf
Carleton University

M. Marcellino
European University Institute, Bocconi University and CEPR

March 18, 2011

Abstract

Robust methods for IV inference have received considerable attention recently. Their analysis has raised a variety of problematic issues such as size/power trade-offs resulting from omitted, weak or many instruments. The popular Anderson-Rubin test has the right size when the underlying first-stage model (that is, the model linking the structural equation's right-hand-side endogenous variables to available instruments) is closed or is incomplete. Alternative methods are available that may outperform this statistic assuming a closed first-stage specification (that is, assuming that all instruments are accounted for). We show that information-reduction methods provide a useful and practical solution to this and related problems. Formally, we propose factor-based modifications to three popular weak-instruments-robust statistics, and illustrate their validity asymptotically and in finite samples. Results are derived using asymptotic settings that are commonly used in both the factor-analysis and weak-instruments literatures. For the Anderson-Rubin statistic, we also provide analytical finite sample results under usual assumptions. An illustrative Monte Carlo study reveals the following. Firstly, our factor-based corrections circumvent the size problems resulting from omitted or many instruments and improve the power of the Anderson-Rubin statistic. Secondly, once corrected through factor reduction, all considered statistics perform equally well. Results suggest that factor-reduction holds promise as a unifying solution to the missing instruments problem as well as a generic attractive approach to robust inference in the presence of many or weak instruments. An empirical study on New Keynesian Phillips Curves suggests that our factor-based methods can bridge the gap between structural and statistical macroeconomic models.

Keywords: Instrumental Regression; Weak Instruments; Identification-Robust Inference; Factor Model; Principle Components; New Keynesian Phillips Curve.

*We would like to thank Jean-Marie Dufour and Abderrahim Taamouti for useful comments and for sharing their code and data. The usual disclaimers apply.

1 Introduction

Identification concerns have revolutionized (see Stock (2010)) the practice of instrumental-variable (IV) based econometrics. Identification problems, which refer to the possibility of inferring characteristics of interest from observable data, are particularly enduring in IV regressions. In such contexts, identifying structural parameters depends on the quality of available instruments, and it is often quite difficult in economics to find informative instruments from observed data. Such difficulties have been at the heart of IV-based econometrics since the early 1990s and various inference methods are now available that are considered identification-robust [IdR], that is, that correct for the possibility of weak instruments.[1] Formally, identification-robustness in this context implies that test sizes or confidence set coverages are valid whether instruments are weak or strong.

The field has been profoundly affected by the introduction of such methods. Nevertheless, important questions are still standing and concern, in particular, problems arising from over-identification or from reliance on "too many" instruments. Such problems stem from the fact that commonly used economic models rarely provide theory-based guidance for instrument choice. Consequently, the number of instruments used in practice is often much larger than the number of instrumented variables. Despite its intuitive appeal, such a practice uses up degrees-of-freedom which causes power losses. In this paper, we propose an information reduction technique aiming to address this problem.

Successful application of most IV-based methods hinges on selecting useful instruments. But the effects of over-identification on IdR methods are particularly consequential, for the following reason. Assume, to set focus (a complete framework is specified below), that interest centers on testing, in the context of the IV regression $Y = \mathbf{Y}_1\beta + U$ where Y is a T-dimensional vector and $\mathbf{Y}_1$ is a $T \times G$ matrix of endogenous regressors, that $\beta = \beta^0$ where β^0 is known. In this case, one of the first proposed IdR methods that may be traced back to Anderson and Rubin (1949) [thereafter AR] requires testing, in the context of an artificial regression of $Y - \mathbf{Y}_1\beta^0$ on k instruments, the exclusion of these k instruments. Mapping the test into a regression on instruments allows one to use standard techniques whose size is not affected by the quality of instruments. However, in over-identified applications, testing G restrictions [on β] would require assessing k constraints [the number of instruments tested

[1] See for example Dufour (1997), Staiger and Stock (1997), Wang and Zivot (1998), Stock and Wright (2000), Stock, Wright, and Yogo (2002), Kleibergen (2002), Dufour (2003), Moreira (2003), Dufour and Taamouti (2005), Kleibergen (2005), Andrews, Moreira, and Stock (2006), Dufour and Taamouti (2006), Andrews and Stock (2007), and Chaudhuri, Richardson, Robins, and Zivot (2010).

out] where k may be much larger than G. A substantial part of the above cited literature has thus focused on assessing the resulting over-identification costs, and on proposing power-saving alternatives.

The statistics proposed by Kleibergen (2002) and Moreira (2003) in the linear IV setting provide such an improvement asymptotically. However, as argued in Dufour (2003) and Dufour and Taamouti (2007), the validity [in terms of size control] of both methods requires, in contrast with the AR method, a complete model; this covers the structural equation as well as the underlying first stage regression (that is, the regression linking the equation's right-hand-side endogenous variables to available instruments). Assuming full conditioning information thus implies accounting for all relevant instruments, which is rather difficult if not often practically impossible, in the absence of economic-theory-based justification for instrument choice. Formally, Dufour and Taamouti (2007) show that departures from a full information setting, and more specifically, omitted instruments, can cause serious size distortions for both Kleibergen's and Moreira's methods. The AR test is immune to this problem. The message one may thus draw from Dufour and Taamouti (2007) is that Kleibergen's and Moreira's statistics outperform the AR method in much too stringent settings. In this paper, we revisit this conclusion.

Unless economic theory suggests the full set of conditioning variables lending credibility to a complete conditioning assumption (which is rare in both macro- and micro-economic applications), missing instruments are a concrete problem. Indeed, in view of Dufour and Taamouti's results, one may wonder [redirecting the question from Bound, Jaeger, and Baker (1995)] whether the "cure" promised by some of the recent IdR methods can still, in some cases, be "worse than the disease". Of course, the latter question, in its original setting, was motivated by asymptotic failures. In contrast, as may be understood from the above motivating example, the missing instruments bias raises dimensionality rather than flawed asymptotics issues. The present paper reflects such a perspective, via our focus on information-reduction.

Information reduction methods including principle components and factor analysis are popular nowadays to analyze a wide spectrum of economic models. Their usefulness for improving standard inference IV methods including the 2SLS estimator has recently been demonstrated; see Bai and Ng (2010), Kapetanios and Marcellino (2010) and the references therein. In this paper, we assess the worth of principle-components based methods in the context of IdR IV-techniques.

When factor approaches are combined with IV-regressions, identification requires both strong instruments as well as strong factor structures. These effects are hard to disentangle. While research on weak instruments has come a long way in the past decade, work on joint weak-instruments and weak-factor structure problems are rather scarce, as may be checked from Bai and Ng (2010) and Kapetanios and Marcellino (2010). In this regard, this paper has four main contributions.

First, we extend the AR method as well as the K-test from Kleibergen (2002) and Moreira (2003) to cover the case where instruments are selected via principle components based methods; we refer to resulting statistics as their factor-based counterparts. Formally, we consider the case where endogenous regressors ($\mathbf{Y}_1$ in the above example) depend, weakly or strongly, on a - possibly large - number of - possibly - unobservable factors, and where the set of available instruments depends, weakly or strongly, on these factors. In this context, we show that the factor-based criteria achieve size control asymptotically, given commonly used regularity assumptions.

Secondly, we show that the factor-based AR statistic remains finite sample exact under usual assumptions. In particular, we prove that this statistic remains pivotal irrespective of whether instruments are weak or strong, or whether the underlying factor structure is weak or strong. To do this, we track the key nuisance parameters that control identification in the statistic's null distribution and show how they cancel out, under the null hypothesis. We provide a simple proof of the latter invariance result, that does not require Gaussian errors. This result also does not call for any assumptions on the relative size of T, the available sample size, and k, the overall number of available instruments, as long as the number of retained factors is much smaller than T. Our factor-AR test is thus useful even if k is larger than T, in which case the original AR test would be infeasible.

Thirdly, we show, via a Monte Carlo study modeled after the designs from Dufour and Taamouti (2007) that our factor-reduction based modification circumvents the size problems associated with Kleibergen's and Moreira's statistics, and improves the power of the AR statistic. Once size corrected, Kleibergen's and Moreira's factor-based tests do not outperform the modified AR test; overall, no one test uniformly dominates the other following instrument reduction. Information-reduction thus provides a promising win/win solution to the size/power trade-offs arising from over-identification.

Finally, we apply our proposed inference methods to structural inflation models. We consider estimating New Keynesian Phillips Curve (NKPC) equations with US data, and three different specifications based on Gali and Gertler (1999) and Gali, Gertler, and Lopez-Salido (2001), Benati (2008) and Kapetanios and Marcellino (2010). IdR inference on the

NKPC has gained popularity in macroeconometrics.[2] Our focus on the NKPC is motivated by the recent survey of Schorfheide (2010). To address some of the empirical challenges that plague Dynamic Stochastic General Equilibrium (DSGE) models, Schorfheide (2010) proposes, among other strategies: (i) to use identification-robust methods, and (ii) to *"connect"* the structure - in some way - to "*richer data-sets*".[3] Our exercise takes both suggestions in consideration. Specifically, we combine model-based instruments with factors based on Stock and Watson (2005). Results suggest that factor based estimation can bridge the gap between data and theory.

The paper is structured as follows. In Section 2 we review the IdR inferential procedures. In Section 3 we introduce and analyze the factor based AR statistic, while in Section 4 we study the factor version of the other IdR procedures. In Section 5 we consider several extensions of the basic framework, including non i.i.d. errors, parameter non-linearity, instrument and factor weakness, and selection of the number of factors. In Section 6 we assess the finite sample size and power properties of the alternative factor based IdR methods. Our empirical analysis is reported in section 7. In Section 8 we summarize our main findings and conclude.

2 Robust inferential procedures in IV regressions

This section provides a brief overview of the robust IV inference methods we consider in this paper within their initially proposed setting. The Data Generation Process (DGP) used in this section is thus motivational; our extended statistical framework is formally presented in section 3.1.

Adopting the notation [maintained throughout the paper] that vectors are identified by capital letters, e.g. $Z = (z_1, ..., z_T)'$ is $T \times 1$, and matrices by bold capital letters, e.g. $\mathbf{Z} = (Z_1, ..., Z_k)$ is $T \times k$, where T denotes the sample size, we consider the system

$$Y = \mathbf{Y}_1\beta + U, \tag{1}$$
$$\mathbf{Y}_1 = \mathbf{X\Pi} + \mathbf{V},$$

[2] See Dufour, Khalaf, and Kichian (2010a), Kleibergen and Mavroeidis (2009) and companion discussions and references therein, Magnusson and Mavroeidis (2010), Nason and Smith (2008) and Dufour, Khalaf, and Kichian (2006)

[3] On the identification of DSGE models, see also Canova and Sala (2009), Iskrev (2010), Komunjer and Ng (2009), Guerron-Quintana, Inoue, and Kilian (2009) and the references therein. On exploiting data-rich approaches in DSGEs including factor analysis, see also Consolo, Favero, and Paccagnini (2009) and the references therein.

where $\mathbf{Y}_1$ is a $T \times G$ matrix of endogenous regressors, $\mathbf{X}$ is a $T \times k$ matrix of valid instruments (exogenous variables), with $k \geq G$, and β and $\mathbf{\Pi}$ are, respectively, a $G \times 1$ vector and $k \times G$ matrix of unknown parameters, and the $T \times (G+1)$ matrix $[U : \mathbf{V}]$ of errors terms have mean zero and are *i.i.d.* across rows.

In this context, the AR test statistic associated with testing

$$H_0 : \beta = \beta^0 \tag{2}$$

using $\mathbf{X}$ as instruments, takes the form

$$AR(\beta^0|\mathbf{X}) = \frac{T-k}{k} \frac{(Y - \mathbf{Y}_1\beta^0)' \left[I - M(\mathbf{X})\right] (Y - \mathbf{Y}_1\beta^0)}{(Y - \mathbf{Y}_1\beta^0)' \left[M(\mathbf{X})\right] (Y - Y_1\beta^0)}$$

where

$$M(\mathbf{A}) = I - N(\mathbf{A})$$
$$N(\mathbf{A}) = \mathbf{A} \left(\mathbf{A}'\mathbf{A}\right)^{-1} \mathbf{A}'$$

for any full-column rank matrix $\mathbf{A}$. This statistic may be viewed as the usual F-criterion associated with testing for the exclusion of $\mathbf{X}$ in the artificial regression of $Y - \mathbf{Y}_1\beta^0$ on $\mathbf{X}$. If in addition we assume that

$$(u_t, v_{1t}, ..., v_{Gt})' \overset{i.i.d.}{\sim} N(0, \Sigma), \quad t = 1, ..., T \tag{3}$$

and that U and $\mathbf{X}$ are independent, then the AR statistic follows in finite samples an $F(k, T-k)$ null distribution. Note that, since this null distribution does not depend on $\mathbf{\Pi}$, it is robust to instrument weakness.

Following the usual classical regression analysis, the Gaussian hypothesis on the U error terms and strong exogeneity of instruments can be relaxed so that, under standard regularity conditions [see for example Dufour and Jasiak (2001)],

$$\left(k \times AR(\beta^0|\mathbf{X})\right) \overset{asy}{\sim} \chi^2(k). \tag{4}$$

We will refer to this asymptotic version of AR as ARS.

Dufour and Taamouti (2006) [see also Dufour (2003) and Dufour, Khalaf, and Kichian (2006)] have shown that a point-optimal although infeasible instrument exists and corresponds to

$$\bar{Z} = \mathbf{X\Pi}.$$

$\bar{Z}$ achieves information reduction, for the associated test amounts to testing the exclusion of the $T \times G$ variables in $\bar{Z}$, even if $k > G$. Dufour (2003) shows that if the OLS estimator

$$\hat{\mathbf{\Pi}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}_1 \tag{5}$$

resulting from the unrestricted reduced form multivariate regression of $\mathbf{Y}_1$ on $\mathbf{X}$ is used in $\bar{Z}$, then the associated AR-test coincides with the LM procedure defined by Wang and Zivot (1998). In addition, see also Dufour, Khalaf, and Kichian (2006), the K-test from Kleibergen (2002) may be obtained using $\bar{Z}$ where $\mathbf{\Pi}$ is replaced by its constrained reduced form maximum likelihood estimates imposing the structure:

$$\hat{\mathbf{\Pi}}^0 = \hat{\mathbf{\Pi}} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\left[Y - \mathbf{Y}_1\beta^0\right] \frac{\left[Y - \mathbf{Y}_1\beta^0\right]' M(\mathbf{X})\mathbf{Y}_1}{\left[Y - \mathbf{Y}_1\beta^0\right]' M(\mathbf{X}) \left[Y - \mathbf{Y}_1\beta^0\right]}. \tag{6}$$

To correct for plug-in effects, Dufour (2003) recommends split sample estimation techniques, where the first sub-sample is used to estimate $\mathbf{\Pi}$ and the second to the associated test based on the latter estimate. The asymptotic distribution of the K-test

$$\left(k \times AR(\beta^0|\hat{\mathbf{\Pi}}^0)\right) \overset{asy}{\sim} \chi^2\,(G)$$

follows under the weak-IV regularity conditions in Kleibergen (2002), which require a closed model.

To introduce the statistic of Moreira (2003), we recall that in a Gaussian closed model with fixed regressors, the LR statistic associated with H_0 and a known reduced form error covariance matrix Ω is given by

$$LR(\beta^0|\mathbf{X}) = \frac{b_0'\mathbf{Y}'N(\mathbf{X})\mathbf{Y}b_0}{b_0'\mathbf{\Omega}b_0} - \min_b \frac{b'\mathbf{Y}'N(\mathbf{X})\mathbf{Y}b}{b'\mathbf{\Omega}b}$$

where $\mathbf{Y}{=}(Y, \mathbf{Y}_1)$, $b = (1, -\beta')'$ and $b_0 = (1, -\beta^{0\prime})'$. The null distribution of LR conditional on the statistic

$$\mathbf{T}_0 = \mathbf{X}'\mathbf{Y}\mathbf{\Omega}^{-1}\left(\beta^0, I\right)',$$

that is, given that $\mathbf{T}_0$ takes the value $\mathbf{t}_0$, does not depend on $\mathbf{\Pi}$; the latter distribution can be computed (as a function of $\mathbf{t}_0$) and used to construct a test robust to $\mathbf{\Pi}$, i.e., to weak instruments. Furthermore, Moreira suggests the plug-in statistic

$$\begin{aligned} LR_1(\beta^0|\mathbf{X}) &= \frac{b_0'\mathbf{Y}'N(\mathbf{X})\mathbf{Y}b_0}{b_0'\hat{\mathbf{\Omega}}b_0} - \hat{\lambda} \\ \hat{\lambda} &= \min_\beta \frac{b'\mathbf{Y}'N(\mathbf{X})\mathbf{Y}b}{b'\hat{\mathbf{\Omega}}b} \\ \hat{\Omega} &= \mathbf{Y}'M(\mathbf{X})\mathbf{Y}/(T-k) \end{aligned}$$

where $\hat{\lambda}$ corresponds to the smallest eigenvalue of $\hat{\mathbf{\Omega}}^{-1/2}\mathbf{Y}'N(\mathbf{X})\mathbf{Y}\hat{\mathbf{\Omega}}^{-1/2}$, and another variant based on the LR statistic with an unknown $\mathbf{\Omega}$

$$LR_2(\beta^0|\mathbf{X}) = \frac{T}{2}\ln\left(1+\frac{b_0'\mathbf{Y}'N(\mathbf{X})\mathbf{Y}b_0}{b_0'\mathbf{Y}'M(\mathbf{X})\mathbf{Y}_0}\right) - \frac{T}{2}\ln\left(1+\frac{\hat{\lambda}}{T-k}\right).$$

These statistics can be used with the same critical values as with $LR(\beta^0|\mathbf{X})$, using for $\mathbf{t}_0$ its plug-in counterpart, obtained by replacing $\mathbf{\Omega}$ by $\hat{\mathbf{\Omega}}$ in $\mathbf{t}_0$.

As we will see in more details in the next section, the AR statistic is robust not only to instrument weakness but also to instrument omission, while several of the other procedures that rely on a specific parametric model for $\mathbf{Y}_1$ are robust to weak instruments but not to instrument omission. In addition, all the inferential procedures that we have described in this section, with the exception of AR, experience severe size distortions when the number of instruments k increases. On the other hand, in this case the AR test can have very low power. In the next two sections we will introduce factor based versions of the robust inferential procedures, and discuss why they can be expected to have better size and power characteristics than the standard procedures we have just presented.

3 Factor based AR-test

3.1 The factor set-up

Using a similar notation as before, let us now consider the DGP

$$Y = \mathbf{Y}_1\beta + U, \tag{7}$$

$$\mathbf{Y}_1 = \mathbf{F\Pi} + X_3\delta + \mathbf{V}, \tag{8}$$

$$\mathbf{X} = \mathbf{F\Lambda} + \mathbf{E}. \tag{9}$$

The main differences with respect to the DGP (1) in Section 2 are that now the endogenous variables $\mathbf{Y}_1$ depend on r unobservable and independent factors, with $r < k$, grouped in the $T \times r$ matrix $\mathbf{F}$, and possibly on another valid instrument, X_3, orthogonal to the factors and with associated $1 \times G$ vector of coefficients δ. In turn, each element of $\mathbf{X}$ depends on the common factors $\mathbf{F}$ via the $r \times k$ loadings $\mathbf{\Lambda}$, and on an idiosyncratic component, i.e., an element of the $T \times k$ matrix $\mathbf{E}$.[4] Conditions on the errors $\mathbf{E} = (e_1, ..., e_T)'$, which are uncorrelated with U, are specified below.

[4] Our analysis formally covers the case where instruments may not have a factor structure.

Assumption 1 *1.* $E||f_t||^4 \leq M < \infty$, $T^{-1}\sum_{t=1}^{T} f_t f_t' \xrightarrow{p} \mathbf{\Sigma}_f$ *for some* $r \times r$ *positive definite matrix* $\mathbf{\Sigma}_f$. $\mathbf{\Lambda}$ *has bounded elements. Further* $||\mathbf{\Lambda}'\mathbf{\Lambda}/k - \mathbf{D}|| \rightarrow 0$, *as* $k \rightarrow \infty$, *where* $\mathbf{D}$ *is a positive definite matrix.*

2. $E(e_{i,t}) = 0$, $E|e_{i,t}|^8 \leq M$ *where* $e_t = (e_{1,t}, ..., e_{k,t})'$. *The variance of* e_t *is denoted by* $\mathbf{\Sigma}_e$. f_s *and* e_t *are independent for all* s, t.

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

- $(kT)^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T} |\sum_{i=1}^{N} \tau_{i,i,t,s}| \leq M$
- $|1/k \sum_{i=1}^{N} \tau_{i,i,s,s}| \leq M$ *for all* s
- $k^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N} |\tau_{i,j,s,s}| \leq M$
- $(kT)^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N} |\tau_{i,j,t,s}| \leq M$
- *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 $(u_t, v_{1t}, ..., v_{Gt})' \overset{i.i.d.}{\sim} N(0, \Sigma_{uv}), t = 1, ..., T.$

Assumption 3 $E(x_{it}u_t) = 0$, $i = 1, ..., k$. $E(x_t x_t')$ *is nonsingular for all* k *and* t. $E(x_t x_t')$ *has full column rank* k. x_t *have finite fourth moments.* u_t *have finite eighth moments.*

Assumption 1 is standard in the factor literature. In particular, it is used in Stock and Watson (2002b), Stock and Watson (2002a), 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 (2006) to show consistency of the parameter estimators in factor augmented regressions. Assumption 3 guarantees that standard IV estimation using (possibly a subset of) x_t as instruments is feasible, and Assumption 2 that it is efficient. Assumption 2 can be relaxed to allow for a more general error structure, while Assumptions 1 and 3 are assumed to hold throughout the paper, unless otherwise stated or modified.

Kapetanios and Marcellino (2010) and Bai and Ng (2010) suggest the use of factors as instruments to construct IV (and GMM) estimators of the parameters β in the model (7). They show that the resulting Factor-IV estimators are more efficient than the standard estimators based on $\mathbf{X}$ as instruments, and can even handle the case where the number of instruments k exceeds the number of observations T, as long as r is small. We are instead interested in deriving factor based robust inferential procedures to test hypotheses on β. In this section we focus on the Factor-AR statistic, while in the next section we consider the Factor version of the other statistics discussed in Section 2.

3.2 The factor AR test

In the factor context defined so far, and labeling $\mathbf{Z}$ as any $T \times r$ matrix of valid instruments, the AR test statistic associated with testing $H_0 : \beta = \beta^0$ takes the form

$$AR(\beta^0|\mathbf{Z}) = \frac{T-r}{r} \frac{(Y - \mathbf{Y}_1\beta^0)' [I - M(\mathbf{Z})] (Y - \mathbf{Y}_1\beta^0)}{(Y - \mathbf{Y}_1\beta^0)' [M(\mathbf{Z})] (Y - \mathbf{Y}_1\beta^0)}.$$

As in the previous section, this statistic may be viewed as the usual F-criterion associated with testing for the exclusion of $\mathbf{Z}$ in the artificial regression of $Y - \mathbf{Y}_1\beta^0$ on $\mathbf{Z}$.

The finite sample properties of this statistic can be motivated via the following decomposition:

$$\begin{aligned} Y - \mathbf{Y}_1\beta^0 &= \mathbf{Y}_1 \left(\beta - \beta^0\right) + U, \\ &= \mathbf{F\Pi} \left(\beta - \beta^0\right) + X_3\delta \left(\beta - \beta^0\right) + U + \mathbf{V} \left(\beta - \beta^0\right), \end{aligned} \tag{10}$$

which, under the null hypothesis, yields

$$Y - \mathbf{Y}_1\beta^0 = U.$$

Plugging the latter into the formula for the statistic, we see that $AR(\beta^0|\mathbf{Z})$ is distributed, under the null hypothesis for finite T, like the pivotal criterion

$$\overline{AR}(\beta^0|\mathbf{Z}) = \frac{T-r}{T} \frac{U' [I - M(\mathbf{Z})] U}{U' [M(\mathbf{Z})] U} \tag{11}$$

irrespective of the values of both $\mathbf{\Pi}$, δ and $\mathbf{\Lambda}$. Robustness to the quality of instruments, to omitted instruments as well as to presence of a factor structure obtains from the latter result. In addition, the error terms $\mathbf{E}$ (and $\mathbf{V}$) are also evacuated-out; consequently, the statistic's null distribution does not depend on the specification of the idiosyncratic errors in the factor model for $\mathbf{X}$. The results on the robustness to the specification of the factor model are new, while the other ones were derived, e.g., in Dufour and Taamouti (2006) (DT) using $\mathbf{X}$ as instruments. Hence, we have also shown that the invariance results from DT do not require using $\mathbf{X}$ for $\mathbf{Z}$ in the formula for $AR(\beta^0|\mathbf{Z})$, and are unaffected by our introduction of the factor specification for $\mathbf{X}$, and by the use of the factors as instruments.

Equation (10) can be also used to assess the key determinants of the power performance of the Factor-AR statistic. In particular, for a given difference $(\beta - \beta^0)$, the power will decrease in the case of weak instruments ("small" $\mathbf{\Pi}$) and omitted relevant instruments ($X_3\delta$). When the $\mathbf{X}$ are used as instruments, the strength of the factor structure ($\mathbf{\Lambda}$) also matters.

The discussion so far did not consider that the factors are unobservable and they have to be estimated. Stock and Watson (2002b) and Stock and Watson (2002a) suggested the use of the principal components of $\mathbf{X}$ as estimators for $\mathbf{F}$, $\widehat{\mathbf{F}}$. Kapetanios and Marcellino (2010) and Bai and Ng (2010) show that the use of $\widehat{\mathbf{F}}$ instead of $\mathbf{F}$ in the construction of Factor-IV estimators does not affect their asymptotic properties under fairly general conditions (we refer to them for details). The main requirement is that $T^{0.5}/k$ is $o_p(1)$, namely, the number of instruments grows faster than the sample size.

In terms of finite sample distribution of the Factor-AR statistic, we do not need any assumptions on the relative size of T and k. In fact, (10) shows that, when the other hypotheses on the model are satisfied and r valid instruments are used (and the principal components are valid since they are just linear combinations of the elements of $\mathbf{X}$), the distribution of the statistic under the null does not change when $\widehat{\mathbf{F}}$ is used for $\mathbf{F}$. This result is noteworthy, since if the number of available instruments is too large relative to the sample size, the original AR test is infeasible. Our factor-based reduction overcomes this problem.

In addition, we will now show that under the same assumptions of Kapetanios and Marcellino (2010) and when $T^{0.5}/k$ is $o_p(1)$ the Factor-AR statistics based on $\mathbf{F}$ and $\widehat{\mathbf{F}}$ are asymptotically numerically equivalent. The proofs for this and subsequent theorems are provided in the appendix.

Theorem 1 *Under assumptions 1-3, under the null hypothesis (2) and when* $T^{0.5}/k = o_p(1)$*,* $(\overline{AR}(\beta^0|\mathbf{F}) - \overline{AR}(\beta^0|\widehat{\mathbf{F}}))$ *is* $o_p(1)$*.*

To conclude, a similar result applies for the Factor-ARS statistic, whose distribution is therefore the same as that of the ARS test with r instruments.

4 Other factor based robust inferential procedures

In this section we introduce the Factor-based version of the other robust inferential procedures described in Section 2, assuming that the DGP is (7), and show that factor estimation does not affect their asymptotic distribution.

Starting with the Factor-LM statistic that extends the procedure suggested by Wang and Zivot (1998), and given the results in Dufour and Taamouti (2006), assuming that $\delta = 0$, we need to show that $(\overline{AR}(\beta^0|\mathbf{F}\hat{\mathbf{\Pi}}) - \overline{AR}(\beta^0|\widehat{\mathbf{F}}\hat{\mathbf{\Pi}}))$ is $o_p(1)$.

Theorem 2 *Under the assumptions in Kapetanios and Marcellino (2010) and when* $T^{0.5}/k$ *is* $o_p(1)$*,* $(\overline{AR}(\beta^0|\mathbf{F}\hat{\mathbf{\Pi}}) - \overline{AR}(\beta^0|\widehat{\mathbf{F}}\hat{\mathbf{\Pi}}))$ *is* $o_p(1)$*.*

The proof is similar to that for the Factor-AR test and the result follows from the fact that $\hat{\mathbf{F}} - \mathbf{FH}$ is $o_p(1)$, where $\mathbf{H}$ is a non singular rotation matrix, since $\mathbf{H}$ does not enter in the computation of the statistic.

Proceeding as in Dufour (2003) and Dufour, Khalaf, and Kichian (2006) it can be shown that the Factor version of Kleibergen's (2002)'s K-test, Factor-K, may be obtained using the instruments $\overline{\mathbf{Z}} = \mathbf{F}\hat{\mathbf{\Pi}}_F^0$, where $\hat{\mathbf{\Pi}}_F^0$ replaces $\mathbf{\Pi}$ by its constrained reduced form OLS estimates imposing the structure [here, again, under $\delta = 0$, that is when the first stage specification is fully specified as a closed model]:

$$\hat{\mathbf{\Pi}}_F^0 = \hat{\mathbf{\Pi}} - (\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}' \left[Y - \mathbf{Y}_1\beta^0\right] \frac{[Y - \mathbf{Y}_1\beta^0]' \, M(\mathbf{F})\mathbf{Y}_1}{[Y - \mathbf{Y}_1\beta^0]' \, M(\mathbf{F}) \, [Y - \mathbf{Y}_1\beta^0]}. \tag{12}$$

The asymptotic distribution for the Factor-K test remains

$$\left(r \times AR(\beta^0|\hat{\mathbf{\Pi}}_F^0)\right) \overset{asy}{\sim} \chi^2\left(r\right)$$

under the weak-IV regularity conditions in Kleibergen (2002), since once again the statistic is invariant to the use of $\mathbf{F}$ or $\mathbf{FH}$ and the principal components are consistent for the space spanned by the true factors.

Finally, we introduce the Factor version of Moreira's statistics. The Factor-LR statistic can be written as

$$Factor - LR(\beta^0|\widehat{\mathbf{F}}) = \frac{b_0'\mathbf{Y}'N(\widehat{\mathbf{F}})\mathbf{Y}b_0}{b_0'\mathbf{\Omega}b_0} - \hat{\lambda}$$

$$\hat{\lambda} = \min_b \frac{b'\mathbf{Y}'N(\widehat{\mathbf{F}})\mathbf{Y}b}{b'\mathbf{\Omega}b}$$

where $\mathbf{Y}= (Y, \mathbf{Y}_1)$, $b = (1, -\beta')'$, $b_0 = (1, -\beta^{0\prime})'$. If the factors are known, the Factor-LR statistic has the same limiting distribution as the standard LR statistic. Hence, we only need to show that $Factor - LR(\beta^0|\mathbf{F}) - Factor - LR(\beta^0|\widehat{\mathbf{F}})$ is $o_p(1)$.

Theorem 3 *Under the assumptions in Moreira (2003) and our assumptions 1-3, and when $T^{0.5}/k = o_p(1)$, $Factor - LR(\beta^0|\mathbf{F}) - Factor - LR(\beta^0|\widehat{\mathbf{F}})$ is $o_p(1)$.*

The Factor versions of Moreira's LR1 and LR2 statistics are constructed along the same lines, and a similar result applies for their asymptotic distribution.

To conclude we note that, as we mentioned, the statistics considered in this section are not robust to instrument omission ($\delta \neq 0$). However, empirically, the factor versions should be much more robust than the standard versions of the statistics, since in practical situations it can be expected that X_3 rather than being orthogonal to $\mathbf{X}$ will be also driven by the factors.

5 Extensions

5.1 Heteroscedasticity and Serial Correlation

We now relax assumption 2 and allow for correlation and heteroskedasticity in the errors U of equation (7) when considering the Factor AR test. We formalize this with the following assumption, which substitutes assumption 2:

Assumption 4 *u_t is a zero mean process with finite variance. The process $f_t u_t$ satisfies the conditions for the application of some central limit theorem for weakly dependent processes, with a zero mean asymptotic normal limit.*

$$S_{fu} = \lim_{T\to\infty}\left[E\left(\left[T^{-1}\sum_{t=1}^{T} u_t^2 f_t f_t^{'}\right]\left[T^{-1}\sum_{t=1}^{T} u_t^2 f_t f_t^{'}\right]^{'}\right)\right]$$

exists and is nonsingular.

Assumption 4 is a high level assumption. It is given in this form for generality. More primitive conditions on ϵ_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.

In principle, the AR method [see for example Stock and Wright (2000), Dufour, Khalaf, and Kichian (2010a), Dufour, Khalaf, and Kichian (2010b)] can be adapted to accommodate such deviations from the *i.i.d.* case. To do so, revisit the artificial regression of $Y - \mathbf{Y}_1\beta^0$ on $\mathbf{Z}$ from section 3.1. In this context, the usual Newey-West Wald-HAC statistic may be used instead of the F-statistic:

$$AR\text{-HAC}(\beta^0|\hat{\mathbf{F}}) = T^{-1}\left(Y - \mathbf{Y}_1\beta^0\right)'\hat{\mathbf{F}}\left(\hat{S}_{\hat{f}u}\right)\hat{\mathbf{F}}'\left(Y - \mathbf{Y}_1\beta^0\right) \tag{13}$$

where $\hat{S}_{\hat{f}u}$ is an estimate of S_{fu} that can be obtained by a HAC procedure, such as that developed in Newey and West (1987). For example, using a Bartlett kernel, we have

$$\widehat{S}_{\hat{f}u,h} = \widehat{\Phi}_0 + \sum_{j=1}^{h}\left(1 - \frac{j}{h+1}\right)(\widehat{\Phi}_j + \widehat{\Phi}_j^{'}), \widehat{\Phi}_j = T^{-1}\sum_{T=j+1}^{T}\widehat{u}_t^{\hat{f}}\widehat{u}_{t-j}^{\hat{f}}\hat{f}_t\hat{f}_{t-j}^{'},$$

where h is the length of the window (bandwidth), $\hat{U}^{\hat{f}} = (\widehat{u}_1^{\hat{f}}, ..., \widehat{u}_T^{\hat{f}})'$, $\hat{U}^{\hat{f}} = (Y - \mathbf{Y}_1\beta^0) - \hat{\mathbf{F}}\hat{\vartheta}^{\hat{f}}$, $\hat{\vartheta}^{\hat{f}} = (\hat{\mathbf{F}}'\hat{\mathbf{F}})^{-1}\hat{\mathbf{F}}'(Y - \mathbf{Y}_1\beta^0)$. Then, assuming the use of the Bartlett kernel, we have the following Theorem that mirrors Theorem 1.

Theorem 4 *Under assumptions 1,3 and 4, under the null hypothesis (2) assuming that $\widehat{S}_{\hat{f}u,h}$ is used as an estimate of S_{fu} with $h = o\left(T^{1/2}\right)$, and when $T^{0.5}/k = o_p(1)$, (AR-HAC($\beta^0|\hat{\mathbf{F}}$) $-$ AR-HAC($\beta^0|\mathbf{F}$)) is $o_p(1)$.*

If $\mathbf{F}$ and $\mathbf{U}$ satisfy the usual exogeneity, stationarity and ergodicity assumptions that validate the usual Newey-West statistic, then the AR-HAC($\beta^0|\mathbf{F}$) statistic will follow a limiting $\chi^2(k)$ under the null hypothesis, and so under Theorem 4 will AR-HAC($\beta^0|\hat{\mathbf{F}}$). Because the residuals $\tilde{u}_t$ are a function of β^0, inverting this statistic cannot be done analytically since the method of Dufour and Taamouti (2005) will no longer work. Numerical methods will be required, as applied for example by Dufour, Khalaf, and Kichian (2010a), Dufour, Khalaf, and Kichian (2010b).

5.2 Nonlinearity

We next consider relaxing the assumption of linearity in parameters. Then, the model becomes

$$\begin{aligned} Y &= \mathbf{Y}_1 \varphi\left(\beta\right) + U, \qquad (14) \\ \mathbf{Y}_1 &= \mathbf{F\Pi} + X_3\delta + \mathbf{V}, \\ \mathbf{X} &= \mathbf{F\Lambda} + \mathbf{E}, \end{aligned}$$

where $\varphi\left(.\right)$ is a possibly non-linear function of β and the latter remains the structural parameter of interest. The above defined AR and factor AR statistics remain valid in finite samples to test $\beta = \beta^0$, regardless of whether the function $\varphi\left(.\right)$ is well behaved or presents discontinuities. Combining non-linearity with possibly non-i.i.d. errors leads to the framework of Stock and Wright (2000), in which case the Wald-HAC criterion would correspond (numerically) or will at least be asymptotically equivalent to a GMM objective function.

5.3 Instrument and factor weakness

In this subsection we briefly discuss the implications of weakness in either instruments or factors for the performance (mostly in terms of power) of the proposed tests. We start by briefly examining the effect of instrument weakness. Instrument weakness relates to (8). In particular, to discuss the issue of weak instruments we replace (8) by

$$\mathbf{Y}_1 = \mathbf{F\Pi}_T + \mathbf{V}$$

where now $\mathbf{\Pi}_T$ is some non-stochastic sequence of matrices over T, and for simplicity we have set $\delta = 0$. Weakness in this context is equivalent to the idea that $\Pi_T'\mathbf{F}'\mathbf{F}\Pi_T = o_p(T)$. In particular, it is reasonable to model weakness by specifying that $\|\Pi_T\| = O(T^{-\vartheta})$, $0 \leq \vartheta < 1/2$. The effect of instrument weakness is that the tests have lower power. In particular, focusing on the AR test although similar results hold for the other tests, it is easy to see that

$$\overline{AR}(\beta^0|\widehat{\mathbf{F}}) = O_p\left(T^{1-2\vartheta}\right) \tag{15}$$

when the true value of the coefficient is $\beta^1 \neq \beta^0$. To see this we have that

$$AR(\beta^0|\mathbf{F}) = \frac{T-r}{r}\frac{(U + (\mathbf{F\Pi}_T + \mathbf{V})(\beta^1 - \beta^0))'\,\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\,(U + \mathbf{F\Pi}_T(\beta^1 - \beta^0))}{(U + \mathbf{F\Pi}_T(\beta^1 - \beta^0))'\,[I - \mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}']\,(U + \mathbf{F\Pi}_T(\beta^1 - \beta^0))}$$

But since

$$\|U'U\| = O_p(T), \quad \left\|U'\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'U\right\| = O_p(1), \quad \left\|(\beta^1 - \beta^0)'\mathbf{\Pi}_T'\mathbf{F}'\mathbf{F\Pi}_T(\beta^1 - \beta^0)\right\| = O_p\left(T^{1-2\vartheta}\right)$$

it follows that

$$\frac{T-r}{r}\frac{1}{(U + \mathbf{F\Pi}_T(\beta^1 - \beta^0))'\,[I - \mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}']\,(U + \mathbf{F\Pi}_T(\beta^1 - \beta^0))} = O_p(1)$$

and

$$\left(U + \mathbf{F\Pi}_T(\beta^1 - \beta^0)\right)'\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\left(U + \mathbf{F\Pi}_T(\beta^1 - \beta^0)\right) = O_p\left(T^{1-2\vartheta}\right)$$

thus proving (15).

Next, we examine the question of weak factor models. In this case (9) is replaced by

$$\mathbf{X} = \mathbf{F}\mathbf{\Lambda}_k + \mathbf{E}.$$

where $\Lambda_k = \mathbf{\Lambda}/k^{\alpha}$, $0 < \alpha < 1$. While such a model of factor loadings may appear restrictive, discussions in Kapetanios and Marcellino (2010) suggest that it can accommodate a variety of different weak factor loading structures. The idea for this model is that the factor structure is weak in the sense that, as the cross-sectional dimension of the dataset increases, the factors explain a diminishing proportion of the variance of the data whose limit as $k \to \infty$, is zero. Kapetanios and Marcellino (2010) discuss this model in detail. For our purposes it is important to note what the implications of this weakness is for the inferential tools we propose. The main effect of the weakness of the factor structure is that factor estimates approximate the true factors sufficiently well to enable Theorems 1-4 to hold, only under stricter conditions. In particular while Theorems 1-4 hold if $T^{0.5}/k = o_p(1)$, the same results hold if $\alpha < 1/4$, $k = o(T^{1/4\alpha})$ and $C_{kT}^{-1}T^{1/2} = o(1)$ where

$$C_{kT} \equiv \min\left(k^{1/2-3\alpha}T^{1/2}, k^{1-3\alpha}, k^{-2\alpha}\min(k, T)\right). \tag{16}$$

The above follows immediately from Lemma 1 and Theorem 5 of Kapetanios and Marcellino (2010).

Another possibility that we wish to explore relates to the case that the instrument dataset does not in fact support any factor structure. In this case, the factors $\mathbf{F}$ appear in (8) but not in (9). This is of course an eccentric setup since it implies that the set of instruments are completely irrelevant for $\mathbf{Y}_1$. Clearly, in this setup any tests, using $\mathbf{X}$ as instruments or as a basis to extract factors and use them as instruments, cannot be expected to exhibit any power. However, it would be reassuring to know that even in this case the factor tests behave well under the null hypothesis. The next Theorem provides this result.

Theorem 5 *Let assumptions 1 and 3, and the null hypothesis (2) hold. Let $\Lambda = 0$ in (9). Let the following hold: u_t is a martingale difference sequence with finite variance,*

$$\sup_{w,\ w'w=1} E\left[(w'x_t)^2\right] \leq \infty, \tag{17}$$

$$\frac{1}{T}\sum_{t=1}^{T}(w'x_t)^2 \xrightarrow{p} w'\Sigma_e w, \quad \textit{uniformly over } w, \textit{ such that } w'w = 1, \tag{18}$$

$$\sigma_f^2 = p\lim_{T,k_2} \hat{w}'\Sigma_e\hat{w} \tag{19}$$

exists and is finite, where $\hat{w}$ denotes the normalized eigenvector associated with the largest eigenvalue of $\frac{1}{T}E'E$. Then

$$k \times \overline{AR}(\beta^0|\widehat{\mathbf{F}}) \xrightarrow{d} \chi^2(k)$$

5.4 Selecting the number of factors

In a conventional factor analysis, the number of factors is estimated (possibly allowing for zero as an outcome) using information criteria (IC) as proposed *e.g.* by Bai and Ng (2002). Penalties with such IC are however too strict when factor structures are weak. The problem (specifically that of estimating the number of factors) within Factor-IV, calls for a different although related treatment. Broadly, three issues deserve notice.

First, results of the previous section suggest reliance on IC that account for weakness, as proposed *e.g.* by Kapetanios and Marcellino (2010). Second, the exact distributional results in section 3.2 are unaffected, as long as the IC are applied to the instruments set only (in a way that does not use information on the model's endogenous regressors).

Finally, and perhaps more importantly, recall that our framework - unlike standard factor IV - does not require $r \geq G$. In particular, size control is achieved, at least asymptotically, even when Λ is zero (in which case $r = 0$). From the power perspective: (i) consistency does not require using *all* relevant factors, and (ii) adding more factors even when relevant does not necessarily translate into effective power gains. On balance, our Monte Carlo results reported below support the following strategy. If more than zero factors are selected via IC, we recommend disregarding the number of factors suggested by IC, and using G factors instead.

6 Monte Carlo evaluation

In order to assess the finite sample behavior of different factor based testing procedures that should be robust in the presence of weak and possibly missing instruments, we now conduct a set of simulation experiments. The statistics under analysis are the factor based versions of the AR test, of its asymptotic version relying on the $\chi^2(k_2)/k_2$ distribution (ARS), of the LR and LM tests proposed by Wang and Zivot (1998), of the K-test by Kleibergen (2002), and of the two versions of the conditional LR test (LR1 and LR2) of Moreira (2003). The same set of statistics was analyzed by DT, but using the instrument set rather than factors as instruments, so that we have a natural benchmark for our results. Moreover, while DT only consider the size of the statistics, we are also interested in their finite sample power.

The Data Generation Process (DGP) follows that by DT but imposes a factor structure. Specifically, the DGP is

$$\begin{aligned} Y &= Y_1\beta_1 + Y_2\beta_2 + U, \qquad (20)\\ (Y_1, Y_2) &= (F_1, F_2)\mathbf{\Pi}_2 + X_3\delta + (V_1, V_2),\\ \mathbf{X}_2 &= (F_1, F_2)\mathbf{\Lambda} + \mathbf{E}, \end{aligned}$$

$$(u_t, v_{1t}, v_{2t})' \overset{i.i.d.}{\sim} N(0, \Sigma),\ \Sigma = \begin{pmatrix} 1 & .8 & .8 \\ .8 & 1 & .3 \\ .8 & .3 & 1 \end{pmatrix},\ t = 1, ..., T.$$

Here $\mathbf{X}_2$ is a $T \times k_2$ matrix of valid instruments and X_3 is a $T \times 1$ omitted instrument vector which is not taken into account when computing the different statistics. DT generate each of the elements of $\mathbf{X}_2$ as *i.i.d.* $N(0, 1)$ variables. Instead, we assume that $\mathbf{X}_2$ is generated by a factor model, where F_1 and F_2 are $T \times 1$ unobservable independent factors whose elements are generated respectively as *i.i.d.* $N(0, 0.4)$ and *i.i.d.* $N(0, 0.2)$, $\mathbf{\Lambda}$ is a $2 \times k_2$

matrix of loadings whose elements are all i.i.d. draws from $N(1,1)$, while the error terms collected in the $T \times k_2$ matrix $\mathbf{E}$ are $i.i.d.$ $N(0,0.4)$ and independent of the factors and all the other errors in the model. Hence, each element of $\mathbf{X}_2$ is $i.i.d.$ $N(0,1)$ as in DT, but it is correlated with each other element of $\mathbf{X}_2$ because of the common factors F_1 and F_2 that drive $\mathbf{X}_2$. The elements of X_3 are instead generated as $i.i.d.$ $N(0,1)$, independent of $\mathbf{X}_2$.

About the parameter values, we follow DT and set $\beta_1 = 1/2$, $\beta_2 = 1$, $\delta = \lambda(1,1)'$, where λ can take the values 0 or 1. The larger the value of λ, the more relevant the omitted instrument X_3. The correlation of each element of U with the corresponding elements of V_1 and V_2 is equal to 0.8, so that the variables Y_1 and Y_2 are endogenous and the instrumental variables $\mathbf{X}_2$ are necessary.

The matrix $\mathbf{\Pi}_2$ controls the strength of the instruments and it is defined as $\mathbf{\Pi}_2 = \rho\mathbf{\Pi}$, where ρ is either 0.01 (weak instruments) or 1 (strong instruments), and as in DT $\mathbf{\Pi}$ is obtained from the identity matrix by keeping the first k_2 lines and the first 2 columns. The number of instruments k_2 varies from 2 to 50, the sample size is $T = 100$, the number of replications is $N = 1000$, and the conditional LR critical values are computed using the same number of replications.

For each statistic, we consider first a version based on the $\mathbf{X}_2$ instruments, and then a factor based version where the unobservable factors F_1 and F_2 are substituted by the first two principal components of $\mathbf{X}_2$. We assess the size of the tests by computing the empirical rejection probability of the null hypothesis $H_0 : \beta = (1/2, 1)'$, when the nominal level of the tests is 5%. To evaluate the (size-adjusted) power, we test the same null hypothesis $H_0 : \beta = (1/2, 1)'$ when $\beta = (1/2 + x, 1 + x)'$ and x is either 0.1 or 0.5.

Size and power are evaluated for four parameters combinations. It can be either $\delta = 0$, so that there is no omitted instrument and as in DT this is a benchmark for comparison with other cases, or $\delta = 1$, so that there is an omitted instrument. For each value of δ, we then consider designs with either weak identification ($\rho = 0.01$) or strong identification ($\rho = 1$). The size and power results for instrument based statistics are presented in Tables 1 and 2 for, respectively, $\delta = 0$ and $\delta = 1$. Tables 3 and 4 present the corresponding results for the (estimated) factor based statistics. In each Table we report results for different values of k_2, ranging from 2 to 50.

Starting with the size results in Tables 1-2, which can be directly compared with those in DT, we find qualitatively similar results to DT even though our instruments are generated by the factor model. In particular, the sizes of the tests K, LR1 and LR2 can be seriously affected by the omission of a relevant instrument, with empirical rejection frequencies even higher than those of DT, and as high as 99% (rather than 5%) when $\delta = 10$.[5] Again as in DT, the conditional LR (LR1 and LR2) tests are better behaved than the K test, but the size distortions remain sizeable, in the range 17%-60%, even in the presence of strong identification. On the positive side, the AR test is very robust to instrument exclusion and weak identification. The ARS is only slightly worse than AR, with largest empirical rejection frequency of about 16%. As noted by DT, the mild distortion in ARS size is due to the use of the chi-square critical value rather than the Fisher critical value.

For all the tests, except AR, the size distortions increase progressively with the number of instruments. This comment suggests that if it is possible to summarize the information in the large instrument set, the size distortions should be limited. The factor structure allows to summarize the information with the estimated factors, and the size panels of Tables 3-4 confirm that when the two estimated factors are used as instruments the size distortions of ALL the factor based test statistics basically disappear. All the empirical rejection frequencies are in the range 4%-6% for any value of the δ parameter, which controls the relevance of the omitted instrument, of the ρ parameter, which controls the instrument strength, and of the k_2 parameter, which controls the number of instruments. This is a very noticeable result, which emphasizes the relevance of factor based testing in IV regressions.

It turns out that the use of factors improves not only the size but also the power properties of all the test statistics. A comparison of Tables 1 and 3 suggests that when the instruments are weak all statistics have no power. However, with strong instruments and $x = 0.1$, the factor based tests are systematically more powerful than the standard statistics. In addition, while the power of the standard statistics decreases with the number of instruments, k_2, that of the factor based statistics is stable. When $x = 0.5$, all statistics have power close to one. When $\delta = 1$, and the omitted instrument becomes relevant, Tables 2 and 4 indicate that the power gains from the factor based tests augment. In particular, in this case the power of the standard statistics becomes close to zero when k_2 increases, while that of the factor based tests remains equal to one.

A further experiment considers the size of the factor and standard tests when $\Lambda = 0$ in (9), a case considered theoretically in subsection 5.3. We report rejection probabilities under the null in Table 5. It turns out that all factor tests behave very well even though the dataset used to extract factors does not in fact contain any factor structure. On the other hand, standard tests have problematic rejection probabilities as expected.

[5] Results for the $\delta = 10$ case are not reported but are available upon request.

A final set of experiments relates to the question of how many factors to use, addressed in subsection 5.4. Here we consider the case of strong instruments ($\rho = 1$) with an extra missing instrument ($\delta = 1$). For power results we set $x = 0.1$. We allow the true (r) and assumed ($\hat{r}$) number of factors to vary. In particular, we consider three experiments: Experiment FN_A: $r = 3$, $\hat{r} = 3$, Experiment FN_B: $r = 3$, $\hat{r} = 2$ and Experiment FN_C: $r = 1$, $\hat{r} = 2$. $G = 2$ throughout. As discussed in subsection 5.4 our preferred strategy is to set $\hat{r} = 2$, since $G = 2$. For experiments FN_A and FN_B we only present results for the factor tests while for the experiment FN_C we also present comparative results for the standard tests. The figures are in Table 6. While for FN_A we get slight size distortions in the form of underrejections for a few tests, these problems are in general reduced in experiments FN_B. Power results remain good throughout, irrespective of the true number of factors. From experiments FN_C, it turns out that the generally better size and power performance of the factor statistics with respect to the standard statistics remains unchanged when the number of factors is overestimated ($r = 1$, $\hat{r} = 2$). Overall, it is clear from the results in Table 6 that our preferred strategy of setting $\hat{r} = G$ makes sense, once one takes into account the overall performance of the test both under the null and alternative hypothesis.

In summary, the simulation experiments indicate very clearly that using factors as instruments can solve the size problems of all the identification robust statistics, and increase their power, including that of the AR-statistic.

7 Empirical Analysis

For our empirical analysis, we consider the New Keynesian Phillips Curve (NKPC) which is perhaps one of the first structural macroeconomic models that have been estimated using identification-robust methods. The NKPC is a dynamic relationship resulting from a limited or full-information equilibrium model, between inflation and a driving variable such as the output gap, unemployment or real marginal costs. Its identification remains a concern, despite major advances in this literature.

Instruments commonly used for empirical works on the NKPC include lags of each model's dependent and forcing variables, which are natural choices when model-consistency is desired. Based on a larger information set, researchers have singled-out additional useful instruments. Examples include: (i) the long-short interest rate spread [Gali and Gertler (1999) and Gali, Gertler, and Lopez-Salido (2001)], (ii) lags of model dependent and forcing variables from various competing specifications [Dufour, Khalaf, and Kichian (2010b)], and (iii) factors extracted via principle components from the 132 variables in Stock and Watson (2005) [Kapetanios and Marcellino (2010)]. We adopt the latter data-rich perspective. Importantly, and differently from Kapetanios and Marcellino (2010), we propose factors that combine model-consistent with statistically-supported variables. We thus assume that relevant instruments are driven by a few common forces, on which factors can provide an exhaustive summary. In this sense, factors may parsimoniously capture all relevant (model-based and statistical) information excluded from the theoretical NKPC.

A prototypical NKPC equation takes the form

$$\pi_t = \lambda s_t + \gamma_f E_t \pi_{t+1} + \gamma_b \pi_{t-1} + \eta_t, \quad t = 1, ..., T \tag{21}$$

where π_t is inflation, s_t is a driving variable measured via the output gap, unemployment or real marginal costs. For estimation purposes, rational expectations are imposed

$$\pi_{t+1} = E_t \pi_{t+1} + \upsilon_{t+1} \tag{22}$$

leading to

$$\pi_t = \lambda s_t + \gamma_f \pi_{t+1} + \gamma_b \pi_{t-1} + \epsilon_t, \quad t = 1, ..., T \tag{23}$$

where ϵ_t includes expectation errors. If υ_t is correlated with η_t, serial dependence cannot be ruled out for ϵ_t (see e.g. Mavroeidis (2005) and Mavroeidis (2004)). Time dependence may also be necessary to fit the data, so HAC estimation of the NKPC is usual.

We consider three different specifications representative of the large literature on the NKPC. First we estimate (23) as in Kapetanios and Marcellino (2010), using US unemployment as a proxy for s_t. Secondly, we estimate, with the same data and instrument sets, a more stylized version of (23) based on Benati (2008) in which

$$\gamma_f = \frac{\phi}{1 + \phi\vartheta}, \quad \gamma_b = \frac{\vartheta}{1 + \phi\vartheta}, \tag{24}$$

where ϑ captures the extent of indexation to past inflation and ϕ is the subjective discount rate. In this case, we use the output gap for s_t as in Benati (2008).[6] Finally, we revisit the initial specification of Gali and Gertler (1999) and Gali, Gertler, and Lopez-Salido (2001)

[6] Note that for strategic reasons Benati imposes i.i.d. errors for this equation and estimates it within a DSGE model with three equations modeling inflation, the output gap and interest rate. Within this system, serial dependence feeds through via AR(1) shocks to the output gap and interest rate equations.

in which

$$\lambda = \frac{(1-\omega)(1-\theta)(1-\beta\theta)}{\theta + \omega[1-\theta(1-\beta)]}, \tag{25}$$

$$\gamma_f = \frac{\beta\theta}{\theta + \omega[1-\theta(1-\beta)]}, \quad \gamma_b = \frac{\omega}{\theta + \omega[1-\theta(1-\beta)]}, \tag{26}$$

where θ measures the degree of price stickiness, ω the degree of "backwardness" in price setting, and β the discount factor. In this specification, following Gali and Gertler (1999), the forcing variable is represented by marginal costs.[7]

Because of our limited-information approach, we allow for serial dependence in all three specifications. We thus rely on the HAC statistic AR-HAC [in (13)]. We invert this statistic, that is we collect the set of parameter values that are not rejected at the 5% level, over a model-relevant search set (reported below in each case). If this set is empty, then the model is rejected at the 5% level. The largest test p-value over the search set provides a measure of model fit, in the spirit of a J-type test; the associated parameter values (that is, the parameter values that maximize this p-value) can be considered as point estimates. We report these point estimates yet we warn that the p-value function was rather flat in parameters, which is expected given documented identification difficulties in this literature. We thus prefer to rely on the projection-based confidence sets that adequately reflect estimation uncertainty.

We use monthly data, from January 1986 to December 2003, taken from Stock and Watson (2005) as in Kapetanios and Marcellino (2010). The marginal cost (defined as in Gali and Gertler (1999)) quarterly time series is interpolated using cubic splines in order to obtain a monthly series. We first produce results using the set of instruments commonly used in this literature (see e.g. Clarida, Galí, and Gertler (1998)): first lag of the output gap, second lag of inflation, first and second lag of the US federal funds rate, first lag of unemployment, and the first lag of the commodity price index. In Table 7, these six instruments are denoted as our 'baseline' set. We next combine the baseline set with 8 factors extracted via principle components from the 132 macroeconomic variables in the Stock and Watson (2005) data set.[8] These factors combine available statistical information (from the reference Stock and Watson (2005) data set) on gap and inflation,

[7]We also used marginal costs with (24) and the output gap with (25), and survey expectations as a proxy for $E_t\pi_{t+1}$ as in Dufour, Khalaf, and Kichian (2006). Results - available upon request - are qualitatively similar to those reported in Table 7.

[8]We refer to Stock and Watson (2005) for a detailed description of the variables and data transformations. We use 8 factors as in Kapetanios and Marcellino (2010) since this is the number selected by the Bai and Ng (2002) selection criterion.

with model-based information. From the latter combined set, we extract two factors (again by principle components) and the first lags of these are used as our instrument set. We retain two factors in line with Section 5.4, since the considered equation includes two endogenous regressors.

For the model with unemployment our estimation search set is -.99 to 0 and 0.0 to .99 for γ_f and γ_b. In the gap-based model imposing (24), our search set for ϕ, ϑ and λ is 0.00 to .99. In the marginal cost case imposing (25), the search set for θ and β is 0.01 to .99 and 0.01 to .90 for ω. Four lags are used for the HAC procedure; projection-based confidence sets are obtained numerically via a grid search with a step of .01. Results are presented in Table 7. We report the projection-based 5% confidence sets and the least-rejected parameters [as point estimates] and associated maximum p-value. Yet as noted above, we prefer to interpret set estimates in the presence of a rather flat objective function.

For the unemployment based linear equation, baseline instruments provide some information on λ. Aside from confirming that the coefficients are significantly non-zero (a point nevertheless worth noting), the confidence sets on γ_f and γ_b are practically non-informative. With two factors as instruments, the confidence set for λ is much tighter, narrowing down from $[-0.54, -0.01]$ to $[-0.08, -0.01]$. The estimated sets for the other two coefficients remain basically non-informative, although zero is still ruled out. With reference to the considered base case, our data-rich approach provides a more precise estimate of λ.

When the output gap is used as a driving variable and (24) is imposed for the backward and forward looking terms, results with baseline instruments reveal serious identification problems. Indeed, we recover the whole search set for all coefficients. When factors are used instead, the confidence sets for the gap coefficient λ tightens up importantly to the [0.00, 0.23] range; the remaining parameters remain under-identified. In contrast to the linear unemployment-based case, zero cannot be ruled-out for all parameters. Yet here again, with reference to the considered base case, our data-rich approach allows to identify λ.

In the case of the marginal-cost driven specification imposing (25), our factor based confidence sets provide the same results as the baseline case. The confidence set for the Calvo parameter θ is the rather wide $[0.53, .99]$ range; the other two parameters are under-identified.

Taken collectively, our results suggest that a factor-based reduced form decreases estimation uncertainty for the slope NKPC, in two out of the three specifications considered, and relative to a conventional benchmark, leaves unaltered the support of the data for the third specification. These results must be interpreted recognizing that instrument selection is unlikely to solve all challenges that plague the NKPC, as a single-equation or within a DSGE model. Indeed, with or without factors, two out of the three parameters in each of the three models considered remained under-identified. While instrument weakness may explain such findings, the survey in Schorfheide (2010) suggests that identifying the NKPC raises much deeper problems. Our results confirm nevertheless that linking data and theory via factor-based first stage regressions is a worthy objective. On balance, and although important identification concerns remain, our data-rich instrument set dominates (and at worst replicates) the considered benchmark selective-information set.

8 Conclusions

In this paper we focus on identification robust inferential procedures and make four main contributions to the literature. First, we introduce the factor-based counterparts of the AR method as well as of the K-test from Kleibergen (2002) and Moreira (2003). In our framework the endogenous regressors depend, weakly or strongly, on a number of unobservable factors, and the possibly large set of available instruments depends, weakly or strongly, on these factors. In this context, we show that our proposed factor-based procedures achieve size control asymptotically, given commonly used regularity assumptions.

Second, we demonstrate that the factor-based AR statistic remains finite sample exact under usual assumptions. Specifically, we show that the statistic remains pivotal whether instruments are weak or strong, and whether the underlying factor structure is weak or strong. In addition, the result does not require any assumptions on the relative size of T, the available sample size, and k, the overall number of available instruments, as long as the number of retained factors is reasonably smaller than T. Our factor-AR test is thus useful even if k is larger than T, a case where the original AR test would be infeasible.

Third, using a Monte Carlo study with a design similar to that in Dufour and Taamouti (2007), we show that our factor-based approach circumvents the size problems associated with Kleibergen's and Moreira's statistics, and improves the power of the AR statistic. Once size corrected, Kleibergen's and Moreira's factor-based tests do not outperform the modified AR test. Overall, no one test uniformly dominates the others following instrument reduction. Information-reduction thus provides a promising win/win solution to the size/power trade-offs arising from over-identification.

Finally, with an empirical study on inflation models we provide evidence that our factor-based methods are easily implementable and can bridge the gap between structural and statistical macroeconomic models.

References

Anderson, T. W., and H. Rubin (1949): "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics*, 20, 46–63.

Andrews, D. W. K., M. J. Moreira, and J. H. Stock (2006): "Optimal Two-Sided Invariant Similar Tests for Instrumental Variables Regression," *Econometrica*, 74, 715–752.

Andrews, D. W. K., and J. H. Stock (2007): "Testing with Many Weak Instruments," *Journal of Econometrics*, 138, 24–46.

Bai, J. (2003): "Inferential Theory for Factor Models of Large Dinensions," *Econometrica*, 71, 135–173.

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

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

——— (2010): "Instrumental Variable Estimation in a Data Rich Environment," *Econometric Theory*.

Benati, L. (2008): "Investigating Inflation Persistence across Monetary Regimes," *Quarterly Journal of Economics*, 123, 1005–60.

Bound, J., D. A. Jaeger, and R. M. Baker (1995): "Problems With Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable Is Weak," *Journal of the American Statistical Association*, 90, 443–450.

Canova, F., and L. Sala (2009): "Back to Square One: Identification Issues in DSGE Models," *Journal of Monetary Economics*, 54, 431–49.

Chaudhuri, S., T. Richardson, J. Robins, and E. Zivot (2010): "A New Split-Sample Score Test in Linear Instrumental Variables Regression," *Econometric Theory*, 26, 1820–1837.

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

Consolo, A., C. Favero, and A. Paccagnini (2009): "On the Statistical Identification of DSGE Models," *Journal of Econometrics*, 150, 99–115.

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

Dufour, J.-M. (1997): "Some Impossibility Theorems in Econometrics, with Applications to Structural and Dynamic Models," *Econometrica*, 65, 1365–1389.

——— (2003): "Identification, Weak Instruments and Statistical Inference in Econometrics," *Canadian Journal of Economics*, 36(4), 767–808.

Dufour, J.-M., and J. Jasiak (2001): "Finite Sample Limited Information Inference Methods for Structural Equations and Models with Generated Regressors," *International Economic Review*, 42, 815–843.

Dufour, J.-M., L. Khalaf, and M. Kichian (2006): "Inflation dynamics and the New Keynesian Phillips Curve: an identification robust econometric analysis," *Journal of Economic Dynamics and Control*, 30, 1707–1728.

——— (2010a): "Estimation Uncertainty in Structural Inflation Models with Real Wage Rigidities," *Computational Statistics and Data Analysis*, 54, 2554–2561.

——— (2010b): "On the Precision of Calvo Parameter Estimates in Structural NKPC Models," *Journal of Economic Dynamics and Control*, 34, 1582–1595.

Dufour, J.-M., and M. Taamouti (2005): "Projection-Based Statistical Inference in Linear Structural Models with Possibly Weak Instruments," *Econometrica*, 73, 1351–1365.

——— (2006): "Point-Optimal Instruments and Generalized Anderson-Rubin Procedures for Nonlinear Models," Discussion paper, C.R.D.E., Université de Montréal.

——— (2007): "Further Results on Projection-Based Inference in IV Regressions with Weak, Collinear or Missing Instruments," *Journal of Econometrics*, 139, 133–153.

Gali, J., and M. Gertler (1999): "Inflation Dynamics: A Structural Econometric Analysis," *Journal of Monetary Economics*, 44, 195–222.

Gali, J., M. Gertler, and J. D. Lopez-Salido (2001): "European Inflation Dynamics," *European Economic Review*, 45, 1237–1270.

Guerron-Quintana, P., A. Inoue, and L. Kilian (2009): "Inference in Weakly Identified DSGE Models," Discussion paper, North Carolina State University and University of Michigan.

Iskrev, N. (2010): "Local Identification in DSGE Models," *Journal of Monetary Economics*, 57, 189–202.

Kapetanios, G., and M. Marcellino (2010): "Factor-GMM estimation with large sets of possibly weak instruments," *Computational Statistics and Data Analysis*, 54(11), 2655–2675.

Kleibergen, F. (2002): "Pivotal Statistics for Testing Structural Parameters in Instrumental Variables Regression," *Econometrica*, 70, 1781–1803.

——— (2005): "Testing Parameters in GMM Without Assuming that They Are Identified," *Econometrica*, 73, 1103–1123.

Kleibergen, F., and S. Mavroeidis (2009): "Weak Instrument Robust Tests in GMM and the New Keynesian Phillips Curve," *Journal of Business and Economic Statistics*, 27, 293–311.

Komunjer, I., and S. Ng (2009): "Dynamic Identification of DSGE Models," Discussion paper, University of California San Diego.

Magnusson, L., and S. Mavroeidis (2010): "Identification-robust Minimum Distance estimation of the new Keynesian Phillips curve," *Journal of Money Credit and Banking*, 42, 1582–1595.

Mavroeidis, S. (2004): "Weak identification of forward-looking models in monetary economics," *Oxford Bulletin of Economics and Statistics*, 66, 609–635.

——— (2005): "Identification issues in forward-looking models estimated by GMM with an application to the Phillips curve," *Journal of Money, Credit and Banking*, 37, 421–449.

Moreira, M. J. (2003): "A Conditional Likelihood Ratio Test for Structural Models," *Econometrica*, 71(4), 1027–1048.

Nason, J., and G. Smith (2008): "Identifying the New Keynesian Phillips Curve," *Journal of Applied Econometrics*, 23, 525–51.

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

Schorfheide, F. (2010): "Estimation and Evaluation of DSGE Models: Progress and Challenges," Discussion paper, Invited Lecture at 2010 Econometric Society World Congress in Shanghai.

Staiger, D., and J. H. Stock (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65(3), 557–586.

Stock, J. H. (2010): "The Other Transformation in Econometric Practice: Robust Tools for Inference," *The Journal of Economic Perspectives*, 24, 83–94.

Stock, J. H., and M. W. Watson (2002a): "Forecasting Using Principal Components From a Large Number of Predictors," *Journal of the American Statistical Association*, 97, 1167–1179.

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

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

Stock, J. H., and J. H. Wright (2000): "GMM with Weak Identification," *Econometrica*, 68, 1097–1126.

Stock, J. H., J. H. Wright, and M. Yogo (2002): "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business and Economic Statistics*, 20(4), 518–529.

Wang, J., and E. Zivot (1998): "Inference on Structural Parameters in Instrumental Variable Regression with Weak Instruments," *Econometrica*, 66, 1389–1404.

Appendix

Proof of Theorem 1

We need to show that

$$\frac{(T-r)(Y-\mathbf{Y}_1\beta^0)'\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'(Y-\mathbf{Y}_1\beta^0)}{r(Y-\mathbf{Y}_1\beta^0)'\left(I-\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\right)(Y-\mathbf{Y}_1\beta^0)} - \frac{(T-r)(Y-\mathbf{Y}_1\beta^0)'\hat{\mathbf{F}}(\hat{\mathbf{F}}'\hat{\mathbf{F}})^{-1}\hat{\mathbf{F}}'(Y-\mathbf{Y}_1\beta^0)}{r(Y-\mathbf{Y}_1\beta^0)'\left(I-\hat{\mathbf{F}}(\hat{\mathbf{F}}'\hat{\mathbf{F}})^{-1}\hat{\mathbf{F}}'\right)(Y-\mathbf{Y}_1\beta^0)} = o_p(1). \tag{27}$$

This holds if

$$U'\mathbf{F}\mathbf{H}(\mathbf{H}'\mathbf{F}'\mathbf{F}\mathbf{H})^{-1}\mathbf{H}'\mathbf{F}'U - U\hat{\mathbf{F}}(\hat{\mathbf{F}}'\hat{\mathbf{F}})^{-1}\hat{\mathbf{F}}'U = U\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'U - U\hat{\mathbf{F}}(\hat{\mathbf{F}}'\hat{\mathbf{F}})^{-1}\hat{\mathbf{F}}'U = o_p(1),$$

for any nonsingular rotation matrix $\mathbf{H}$, which is in turn satisfied if

$$\frac{\mathbf{F}'\mathbf{F}}{T} - \frac{\hat{\mathbf{F}}'\mathbf{F}}{T} = o_p(1) \tag{28}$$

and

$$\sqrt{T}\left(\frac{\mathbf{F}'U}{T} - \frac{\hat{\mathbf{F}}'U}{T}\right) = o_p(1) \tag{29}$$

hold. We examine (28)-(29). The left hand side of (28) and that of (29) divided by $\sqrt{T}$ can be re-written as

$$\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t - Hf_t)q_t' \tag{30}$$

where q_t is either f_t or u_t. By Lemma A.1 of Bai and Ng (2006) we have that

$$\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t - Hf_t)q_t' = O_p\left(\min(k,T)^{-1}\right) \tag{31}$$

as long as q_t has finite fourth moments, nonsingular covariance matrix and $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(q_t - E(q_t))$ satisfies a central limit theorem. These conditions are satisfied for f_t and u_t via assumptions 1 and 3. Hence, (28) follows, while (29) follows if $\sqrt{T}/k = o_p(1)$. Note, for later use, that a similar result holds for $q_t = y_t$.

Proof of Theorem 3

Given the consistency of $\widehat{\mathbf{F}}$ for $\mathbf{FH}$ and the invariance of $N(\mathbf{F})$ to the use of $\mathbf{F}$ or $\mathbf{FH}$, we only need to show that

$$\hat{\lambda} - \lambda = o_p(1) \tag{32}$$

where

$$\lambda = \min_b \frac{b'\mathbf{Y}'\left(\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\right)\mathbf{Y}b}{b'\Omega b}$$

and

$$\Omega = \mathbf{Y}'\left(I - \left(\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\right)\right)\mathbf{Y}/(T-k_2).$$

Note that $\hat{\lambda}$ and λ correspond to the smallest eigenvalues of $\hat{\Omega}^{-1/2}\mathbf{Y}'\left(\hat{\mathbf{F}}(\hat{\mathbf{F}}'\hat{\mathbf{F}})^{-1}\hat{\mathbf{F}}'\right)\mathbf{Y}\hat{\Omega}^{-1/2}$ and $\Omega^{-1/2}\mathbf{Y}'\left(\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\right)\mathbf{Y}\Omega^{-1/2}$ respectively Note further, that $\hat{\Omega}^{-1/2}\mathbf{Y}'\left(\hat{\mathbf{F}}(\hat{\mathbf{F}}'\hat{\mathbf{F}})^{-1}\hat{\mathbf{F}}'\right)\mathbf{Y}\hat{\Omega}^{-1/2}$ is a finite dimensional matrix whose every element converges in probability to the respective element of $\Omega^{-1/2}\mathbf{Y}'\left(\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\right)\mathbf{Y}\Omega^{-1/2}$, if we can show that

$$\frac{\mathbf{Y}'\mathbf{F}}{T} - \frac{\mathbf{Y}'\widehat{\mathbf{F}}}{T} = o_p(1), \quad \frac{\mathbf{F}'\mathbf{F}}{T} - \frac{\hat{\mathbf{F}}'\mathbf{F}}{T} = o_p(1) \tag{33}$$

hold. Let us assume for the moment that (33) holds. Then, given that the eigenvalues of a matrix, being polynomial roots, are continuous functions of the elements of the matrix, (32) follows by Slutsky's theorem. Hence, to complete the proof we need to demonstrate (33). But it follows immediately from (30)-(31) in the proof of Theorem 1.

Proof of Theorem 4

We need to show that

$$T^{-1}\left(Y - \mathbf{Y}_1\beta^0\right)'\mathbf{F}\hat{S}_{fu}^{-1}\mathbf{F}'\left(Y - \mathbf{Y}_1\beta^0\right) -$$

$$T^{-1}\left(Y - \mathbf{Y}_1\beta^0\right)'\hat{\mathbf{F}}\hat{S}_{\hat{f}u}^{-1}\hat{\mathbf{F}}'\left(Y - \mathbf{Y}_1\beta^0\right) = o_p(1)$$

or, introducing normalization terms,

$$\frac{U'\hat{\mathbf{F}}}{T^{1/2}}\hat{S}_{\hat{f}u}^{-1}\frac{\hat{\mathbf{F}}'U}{T^{1/2}} - \frac{U'\mathbf{F}}{T^{1/2}}\hat{S}_{fu}^{-1}\frac{\mathbf{F}'U}{T^{1/2}} = o_p(1)$$

where $\widehat{S}_{fu,h} = \Phi_0 + \sum_{j=1}^{h}\left(1 - \frac{j}{h+1}\right)(\Phi_j + \Phi_j'),\ \Phi_j = T^{-1}\sum_{T=j+1}^{T}\hat{u}_t^f\hat{u}_{t-j}^f f_t f_{t-j}'$. In the proofs of previous Theorems we have shown that $\sqrt{T}\left(\frac{\hat{\mathbf{F}}'U}{T} - \frac{\mathbf{F}'U}{T}\right) = o_p(1)$. The proof of the theorem is complete if we show formally that $\Phi_j - \widehat{\Phi}_j = o_p(h^{-1})$. We show below that $\Phi_0 - \widehat{\Phi}_0 = o_p(h^{-1})$. The result for $j > 0$ follows similarly. We have

$$\left\|T^{-1}\sum_{T=j+1}^{T}\left(\widehat{u}_t^f\right)^2 f_t f_t' - T^{-1}\sum_{T=j+1}^{T}\left(\widehat{u}_t^{\hat{f}}\right)^2 \hat{f}_t\hat{f}_t'\right\| \le C_1\left\|T^{-1}\sum_{T=j+1}^{T}\left(\widehat{u}_t^f\right)^2 f_t'\left(Hf_t - \widehat{f}_t\right)\right\| + \tag{34}$$

$$C_2 \left\| T^{-1} \sum_{T=j+1}^{T} \widehat{u}_t^f \left(\widehat{u}_t^f - \widehat{u}_t^{\hat{f}} \right) f_t f_t' \right\|$$

for some constants C_1,C_2. First, we consider the first term of the RHS of (34). We have

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

$$C_3 \left\| T^{-1} \sum_{T=j+1}^{T} u_t^2 f_t' \left(H f_t - \widehat{f}_t \right) \right\| + C_4 \left\| T^{-1} \sum_{T=j+1}^{T} (\widehat{u}_t^f - u_t) f_t' \left(H f_t - \widehat{f}_t \right) \right\|$$

for some constants C_3,C_4. But, by Lemma A.1 of Bai and Ng (2006), if $\sqrt{T}/k = o(1)$,

$$\left\| T^{-1} \sum_{T=j+1}^{T} u_t^2 f_t' \left(H f_t - \widehat{f}_t \right) \right\| = o_p\left(T^{-1/2} \right)$$

as long as u_t has finite eighth moments. Then,

$$\left\| T^{-1} \sum_{T=j+1}^{T} (\widehat{u}_t^f - u_t) f_t' \left(H f_t - \widehat{f}_t \right) \right\| \leq C_5 \left\| \hat{\vartheta}^f - \vartheta \right\| \left\| T^{-1} \sum_{T=j+1}^{T} f_t f_t' \left(H f_t - \widehat{f}_t \right) \right\|$$

for some constant C_5. Again, by Lemma A.1 of Bai and Ng (2006), if $\sqrt{T}/k = o(1)$,

$$\left\| T^{-1} \sum_{T=j+1}^{T} f_t f_t' \left(H f_t - \widehat{f}_t \right) \right\| = o_p\left(T^{-1/2} \right).$$

By consistency of $\hat{\vartheta}^f$ (note that ϑ is equal to zero under the null hypothesis), $\left\| \hat{\vartheta}^f - \vartheta \right\| = o_p(1)$. Next, we consider the second term on the RHS of (34). We have

$$\left\| T^{-1} \sum_{T=j+1}^{T} \widehat{u}_t^f \left(\widehat{u}_t^f - \widehat{u}_t^{\hat{f}} \right) f_t f_t' \right\| \leq C_6 \left\| T^{-1} \sum_{T=j+1}^{T} u_t f_t f_t' \left(H f_t - \widehat{f}_t \right) \right\| \quad (35)$$

for some constant C_6. By similar arguments to those above the RHS of (35) is $o_p\left(T^{-1/2} \right)$ if $\sqrt{T}/k = o(1)$. Hence, $\Phi_0 - \widehat{\Phi}_0 = o_p(h^{-1})$ as long as $h = o(T^{1/2})$.

Proof of Theorem 5

It is sufficient to show that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \hat{f}_t u_t \xrightarrow{d} N(0, \sigma_f^2 \sigma_u^2) \quad (36)$$

and

$$\frac{1}{T} \sum_{t=1}^{T} \hat{f}_t^2 \xrightarrow{p} \sigma_f^2 \quad (37)$$

(37) follows immediately by (19) and (18). (36) can be shown as follows. We first note that x_t and u_t are independent. Then, by the martingale difference assumption for u_t and (17) it follows that $\hat{f}_t u_t$ is a martingale difference with finite variance. Then, (36) follows by the martingale difference central limit theorem.

Table 1. Variables as instruments, no omitted instruments

	AR	ARS	K	LM	LR	LR1	LR2		AR	ARS	K	LM	LR	LR1	LR2
K2								K2							
			Weak instruments								Strong instruments				
			Size								Size				
2	0.039	0.043	0.043	0.038	0.043	0.043	0.043	2	0.057	0.063	0.063	0.057	0.061	0.061	0.063
3	0.054	0.057	0.055	0.051	0.046	0.053	0.054	3	0.048	0.054	0.051	0.018	0.020	0.052	0.054
4	0.047	0.053	0.058	0.044	0.037	0.054	0.057	4	0.049	0.054	0.056	0.011	0.011	0.058	0.060
5	0.046	0.057	0.061	0.041	0.035	0.055	0.056	5	0.050	0.057	0.062	0.006	0.006	0.058	0.063
10	0.048	0.063	0.074	0.040	0.016	0.070	0.074	10	0.043	0.063	0.050	0.000	0.001	0.049	0.052
20	0.058	0.093	0.110	0.034	0.008	0.111	0.106	20	0.043	0.069	0.064	0.000	0.000	0.066	0.069
40	0.055	0.116	0.170	0.015	0.003	0.141	0.137	40	0.057	0.124	0.082	0.000	0.000	0.085	0.082
50	0.053	0.144	0.229	0.003	0.002	0.192	0.181	50	0.040	0.124	0.069	0.000	0.000	0.077	0.081
			Power (size corrected) x=0.1								Power (size corrected) x=0.1				
2	0.064	0.064	0.064	0.064	0.064	0.064	0.064	2	0.147	0.147	0.147	0.147	0.147	0.147	0.147
3	0.037	0.037	0.045	0.037	0.046	0.046	0.044	3	0.179	0.179	0.225	0.185	0.221	0.221	0.224
4	0.061	0.061	0.062	0.060	0.062	0.062	0.062	4	0.142	0.142	0.166	0.117	0.167	0.167	0.165
5	0.041	0.041	0.046	0.041	0.037	0.037	0.034	5	0.199	0.199	0.253	0.156	0.241	0.241	0.241
10	0.047	0.047	0.053	0.047	0.046	0.046	0.047	10	0.138	0.138	0.288	0.034	0.273	0.273	0.272
20	0.044	0.044	0.045	0.042	0.036	0.036	0.036	20	0.127	0.127	0.301	0.001	0.284	0.284	0.294
40	0.054	0.054	0.036	0.052	0.046	0.046	0.060	40	0.078	0.078	0.228	0.002	0.198	0.198	0.187
50	0.035	0.035	0.052	0.029	0.058	0.058	0.057	50	0.109	0.109	0.279	0.000	0.256	0.256	0.247
			Power (size corrected) x=0.5								Power (size corrected) x=0.5				
2	0.073	0.073	0.073	0.073	0.073	0.073	0.073	2	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	0.043	0.043	0.045	0.038	0.051	0.051	0.049	3	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	0.061	0.061	0.061	0.038	0.062	0.062	0.062	4	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.057	0.057	0.053	0.038	0.049	0.049	0.047	5	1.000	1.000	1.000	1.000	1.000	1.000	1.000
10	0.056	0.056	0.058	0.021	0.054	0.054	0.056	10	1.000	1.000	1.000	1.000	1.000	1.000	1.000
20	0.046	0.046	0.045	0.010	0.038	0.038	0.039	20	1.000	1.000	1.000	1.000	1.000	1.000	1.000
40	0.050	0.050	0.047	0.000	0.047	0.047	0.058	40	1.000	1.000	1.000	0.989	1.000	1.000	1.000
50	0.051	0.051	0.041	0.000	0.058	0.058	0.057	50	1.000	1.000	1.000	0.868	1.000	1.000	1.000

Note: The statistics under analysis are the factor based versions of the AR test, of its asymptotic version relying on the $\chi^2(k_2)/k_2$distribution (ARS), of the LR and LM tests proposed by Wang and Zivot (1998), of the K-test by Kleibergen (2002), and of the two versions of the conditional LR test (LR1 and LR2) of Moreira (2003). The Data Generating Process is described in Section 6. K2 indicates the number of instruments.

Table 2. Variables as instruments, omitted instrument, δ=1

	AR	ARS	K	LM	LR	LR1	LR2		AR	ARS	K	LM	LR	LR1	LR2
K2								K2							
	Weak instruments								Strong instruments						
	Size								Size						
2	0.045	0.048	0.048	0.043	0.047	0.047	0.048	2	0.040	0.043	0.043	0.040	0.042	0.042	0.043
3	0.069	0.076	0.090	0.063	0.074	0.076	0.076	3	0.045	0.052	0.039	0.012	0.020	0.042	0.042
4	0.048	0.050	0.104	0.045	0.050	0.052	0.054	4	0.047	0.053	0.051	0.009	0.010	0.048	0.049
5	0.040	0.052	0.127	0.033	0.047	0.049	0.050	5	0.051	0.060	0.054	0.007	0.007	0.059	0.062
10	0.043	0.053	0.305	0.035	0.057	0.059	0.059	10	0.047	0.055	0.069	0.001	0.001	0.065	0.066
20	0.046	0.086	0.514	0.029	0.111	0.120	0.092	20	0.043	0.078	0.070	0.000	0.000	0.075	0.077
40	0.047	0.126	0.826	0.007	0.348	0.379	0.142	40	0.052	0.130	0.117	0.000	0.000	0.156	0.154
50	0.041	0.131	0.856	0.003	0.536	0.571	0.150	50	0.056	0.154	0.177	0.000	0.000	0.183	0.192
	Power (size corrected) x=0.1								Power (size corrected) x=0.1						
2	0.056	0.056	0.056	0.056	0.056	0.056	0.056	2	0.293	0.293	0.293	0.293	0.293	0.293	0.293
3	0.030	0.030	0.039	0.031	0.030	0.030	0.030	3	0.188	0.188	0.226	0.214	0.222	0.222	0.221
4	0.040	0.040	0.062	0.039	0.040	0.040	0.040	4	0.152	0.152	0.177	0.129	0.192	0.192	0.196
5	0.053	0.053	0.082	0.053	0.053	0.053	0.053	5	0.105	0.105	0.141	0.081	0.145	0.145	0.135
10	0.041	0.041	0.143	0.040	0.041	0.041	0.041	10	0.121	0.121	0.131	0.016	0.161	0.161	0.168
20	0.042	0.042	0.189	0.040	0.043	0.043	0.043	20	0.110	0.110	0.198	0.000	0.198	0.198	0.190
40	0.033	0.033	0.078	0.033	0.033	0.033	0.033	40	0.053	0.053	0.046	0.000	0.045	0.045	0.046
50	0.039	0.039	0.086	0.034	0.044	0.044	0.041	50	0.056	0.056	0.033	0.000	0.039	0.039	0.046
	Power (size corrected) x=0.5								Power (size corrected) x=0.5						
2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	2	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	3	1.000	1.000	0.932	1.000	1.000	1.000	1.000
4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	4	1.000	1.000	0.730	1.000	1.000	1.000	1.000
5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	5	1.000	1.000	0.714	1.000	1.000	1.000	1.000
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	10	1.000	1.000	0.483	1.000	1.000	1.000	1.000
20	0.000	0.000	0.000	0.000	0.000	0.000	0.000	20	1.000	1.000	0.307	1.000	0.986	0.986	0.989
40	0.000	0.000	0.000	0.000	0.000	0.000	0.000	40	0.372	0.372	0.118	0.777	0.580	0.580	0.564
50	0.000	0.000	0.000	0.000	0.000	0.000	0.000	50	0.000	0.000	0.007	0.006	0.003	0.003	0.000

Note: The statistics under analysis are the factor based versions of the AR test, of its asymptotic version relying on the $\chi^2(k_2)/k_2$distribution (ARS), of the LR and LM tests proposed by Wang and Zivot (1998), of the K-test by Kleibergen (2002), and of the two versions of the conditional LR test (LR1 and LR2) of Moreira (2003).The Data Generating Process is described in Section 6. K2 indicates the number of instruments.

Table 3. Factors as instruments, no omitted instruments

	AR	ARS	K	LM	LR	LR1	LR2		AR	ARS	K	LM	LR	LR1	LR2
K2								K2							
	Weak instruments								Strong instruments						
	Size								Size						
2	0.051	0.054	0.054	0.051	0.053	0.053	0.054	2	0.049	0.050	0.050	0.046	0.049	0.049	0.050
3	0.033	0.037	0.037	0.033	0.037	0.037	0.037	3	0.036	0.037	0.037	0.036	0.036	0.036	0.037
4	0.052	0.056	0.056	0.049	0.055	0.055	0.056	4	0.047	0.053	0.053	0.046	0.053	0.053	0.053
5	0.048	0.053	0.053	0.044	0.051	0.051	0.053	5	0.035	0.038	0.038	0.034	0.036	0.036	0.038
10	0.051	0.055	0.055	0.050	0.053	0.053	0.055	10	0.052	0.060	0.060	0.050	0.056	0.056	0.060
20	0.052	0.054	0.054	0.051	0.052	0.052	0.054	20	0.041	0.045	0.045	0.041	0.045	0.045	0.045
40	0.042	0.048	0.048	0.041	0.045	0.045	0.048	40	0.042	0.048	0.048	0.039	0.046	0.046	0.048
50	0.045	0.051	0.051	0.045	0.050	0.050	0.051	50	0.053	0.057	0.057	0.052	0.055	0.055	0.057
	Power (size corrected) x=0.1								Power (size corrected) x=0.1						
2	0.047	0.050	0.050	0.045	0.050	0.050	0.050	2	0.206	0.218	0.218	0.203	0.214	0.214	0.218
3	0.046	0.049	0.049	0.044	0.047	0.047	0.049	3	0.273	0.280	0.280	0.268	0.277	0.277	0.280
4	0.056	0.057	0.057	0.054	0.056	0.056	0.057	4	0.222	0.237	0.237	0.216	0.234	0.234	0.237
5	0.040	0.043	0.043	0.039	0.041	0.041	0.043	5	0.273	0.287	0.287	0.271	0.284	0.284	0.287
10	0.057	0.063	0.063	0.055	0.062	0.062	0.063	10	0.308	0.327	0.327	0.302	0.322	0.322	0.327
20	0.052	0.058	0.058	0.052	0.055	0.055	0.058	20	0.281	0.290	0.290	0.277	0.288	0.288	0.290
40	0.053	0.056	0.056	0.053	0.056	0.056	0.056	40	0.304	0.321	0.321	0.298	0.317	0.317	0.321
50	0.048	0.052	0.052	0.048	0.050	0.050	0.052	50	0.291	0.306	0.306	0.287	0.303	0.303	0.306
	Power (size corrected) x=0.5								Power (size corrected) x=0.5						
2	0.047	0.048	0.048	0.044	0.047	0.047	0.048	2	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	0.057	0.061	0.061	0.055	0.059	0.059	0.061	3	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	0.062	0.064	0.064	0.061	0.064	0.064	0.064	4	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.039	0.044	0.044	0.038	0.043	0.043	0.044	5	1.000	1.000	1.000	1.000	1.000	1.000	1.000
10	0.046	0.054	0.054	0.043	0.053	0.053	0.054	10	1.000	1.000	1.000	1.000	1.000	1.000	1.000
20	0.042	0.049	0.049	0.041	0.045	0.045	0.049	20	1.000	1.000	1.000	1.000	1.000	1.000	1.000
40	0.065	0.070	0.070	0.063	0.068	0.068	0.070	40	1.000	1.000	1.000	1.000	1.000	1.000	1.000
50	0.053	0.057	0.057	0.052	0.055	0.055	0.057	50	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Note: The statistics under analysis are the factor based versions of the AR test, of its asymptotic version relying on the $\chi^2(k_2)/k_2$distribution (ARS), of the LR and LM tests proposed by Wang and Zivot (1998), of the K-test by Kleibergen (2002), and of the two versions of the conditional LR test (LR1 and LR2) of Moreira (2003).The Data Generating Process is described in Section 6. K2 indicates the number of instruments.

Table 4. Factors as instruments, omitted instrument, δ=1

	AR	ARS	K	LM	LR	LR1	LR2		AR	ARS	K	LM	LR	LR1	LR2
K2								K2							
	Weak instruments								Strong instruments						
	Size								Size						
2	0.060	0.068	0.068	0.059	0.065	0.065	0.068	2	0.065	0.069	0.069	0.063	0.069	0.069	0.069
3	0.049	0.053	0.053	0.047	0.050	0.050	0.053	3	0.049	0.052	0.052	0.047	0.052	0.052	0.052
4	0.053	0.059	0.059	0.052	0.058	0.058	0.059	4	0.052	0.059	0.059	0.048	0.055	0.055	0.059
5	0.055	0.059	0.059	0.053	0.056	0.056	0.059	5	0.059	0.063	0.063	0.057	0.061	0.061	0.063
10	0.051	0.054	0.054	0.050	0.053	0.053	0.054	10	0.043	0.045	0.045	0.041	0.045	0.045	0.045
20	0.052	0.060	0.060	0.052	0.058	0.058	0.060	20	0.055	0.057	0.057	0.054	0.056	0.056	0.057
40	0.053	0.058	0.058	0.053	0.058	0.058	0.058	40	0.052	0.055	0.055	0.051	0.055	0.055	0.055
50	0.039	0.042	0.042	0.036	0.042	0.042	0.042	50	0.055	0.056	0.056	0.054	0.055	0.055	0.056
	Power (size corrected) x=0.1								Power (size corrected) x=0.1						
2	0.035	0.040	0.040	0.035	0.038	0.038	0.040	2	0.129	0.140	0.140	0.126	0.133	0.133	0.140
3	0.043	0.051	0.051	0.041	0.049	0.049	0.051	3	0.271	0.283	0.283	0.268	0.281	0.281	0.283
4	0.026	0.030	0.030	0.026	0.029	0.029	0.030	4	0.223	0.238	0.238	0.219	0.231	0.231	0.238
5	0.042	0.044	0.044	0.040	0.044	0.044	0.044	5	0.227	0.237	0.237	0.221	0.236	0.236	0.237
10	0.042	0.045	0.045	0.041	0.045	0.045	0.045	10	0.277	0.292	0.292	0.270	0.286	0.286	0.292
20	0.049	0.051	0.051	0.047	0.050	0.050	0.051	20	0.295	0.301	0.301	0.290	0.298	0.298	0.301
40	0.041	0.049	0.049	0.038	0.045	0.045	0.049	40	0.231	0.245	0.245	0.228	0.237	0.237	0.245
50	0.040	0.044	0.044	0.039	0.041	0.041	0.044	50	0.273	0.289	0.289	0.267	0.283	0.283	0.289
	Power (size corrected) x=0.5								Power (size corrected) x=0.5						
2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	2	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	3	1.000	1.000	1.000	1.000	1.000	1.000	1.000
4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	4	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	5	1.000	1.000	1.000	1.000	1.000	1.000	1.000
10	0.000	0.000	0.000	0.000	0.000	0.000	0.000	10	1.000	1.000	1.000	1.000	1.000	1.000	1.000
20	0.000	0.000	0.000	0.000	0.000	0.000	0.000	20	1.000	1.000	1.000	1.000	1.000	1.000	1.000
40	0.000	0.000	0.000	0.000	0.000	0.000	0.000	40	1.000	1.000	1.000	1.000	1.000	1.000	1.000
50	0.000	0.000	0.000	0.000	0.000	0.000	0.000	50	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Note: The statistics under analysis are the factor based versions of the AR test, of its asymptotic version relying on the $\chi^2(k_2)/k_2$distribution (ARS), of the LR and LM tests proposed by Wang and Zivot (1998), of the K-test by Kleibergen (2002), and of the two versions of the conditional LR test (LR1 and LR2) of Moreira (2003).The Data Generating Process is described in Section 6. K2 indicates the number of instruments.

Table 5. Factors/Variables as instruments, no omitted instrument, $\Lambda = 0$ in (9)

K2	AR	ARS	K	LM	LR	LR1	LR2	K2	AR	ARS	K	LM	LR	LR1	LR2
Factors as instruments															
	Weak instruments								Strong instruments						
	Size								Size						
2	0.047	0.053	0.053	0.046	0.052	0.052	0.053	2	0.068	0.073	0.073	0.065	0.071	0.071	0.073
3	0.046	0.051	0.051	0.044	0.049	0.049	0.051	3	0.049	0.054	0.054	0.047	0.051	0.051	0.054
4	0.054	0.057	0.057	0.053	0.057	0.057	0.057	4	0.053	0.055	0.055	0.052	0.054	0.054	0.055
5	0.050	0.055	0.055	0.048	0.054	0.054	0.055	5	0.041	0.044	0.044	0.041	0.043	0.043	0.044
10	0.054	0.057	0.057	0.052	0.057	0.057	0.057	10	0.051	0.059	0.059	0.051	0.056	0.056	0.059
20	0.048	0.053	0.053	0.046	0.052	0.052	0.053	20	0.052	0.056	0.056	0.052	0.053	0.053	0.056
40	0.068	0.074	0.074	0.068	0.073	0.073	0.074	40	0.044	0.047	0.047	0.043	0.047	0.047	0.047
50	0.059	0.064	0.064	0.058	0.062	0.062	0.064	50	0.058	0.062	0.062	0.056	0.059	0.059	0.062
Variables as instruments															
	Weak instruments								Strong instruments						
	Size								Size						
2	0.051	0.058	0.058	0.048	0.057	0.057	0.058	2	0.043	0.050	0.050	0.043	0.050	0.050	.050
3	0.037	0.044	0.039	0.035	0.031	0.047	0.047	3	0.058	0.063	0.064	0.021	0.048	0.069	0.072
4	0.059	0.066	0.053	0.056	0.048	0.064	0.067	4	0.052	0.063	0.045	0.008	0.032	0.057	0.061
5	0.059	0.070	0.060	0.048	0.040	0.065	0.069	5	0.047	0.058	0.057	0.002	0.025	0.057	0.058
10	0.058	0.074	0.070	0.046	0.021	0.077	0.082	10	0.057	0.074	0.077	0.001	0.022	0.069	0.075
20	0.051	0.083	0.075	0.027	0.006	0.086	0.093	20	0.041	0.076	0.079	0.000	0.002	0.090	0.095
40	0.059	0.142	0.204	0.012	0.002	0.168	0.167	40	0.047	0.119	0.094	0.000	0.001	0.134	0.134
50	0.032	0.121	0.239	0.001	0.003	0.172	0.156	50	0.047	0.135	0.123	0.000	0.001	0.159	0.168

Note: The statistics under analysis are the factor based versions of the AR test, of its asymptotic version relying on the $\chi^2(k_2)/k_2$distribution (ARS), of the LR and LM tests proposed by Wang and Zivot (1998), of the K-test by Kleibergen (2002), and of the two versions of the conditional LR test (LR1 and LR2) of Moreira (2003).The Data Generating Process is described in Section 6. K2 indicates the number of instruments.

Table 6. Factors/Variables as instruments, omitted instrument, δ=1, results for different numbers of factors

	AR	ARS	K	LM	LR	LR1	LR2		AR	ARS	K	LM	LR	LR1	LR2
K2								K2							
								$r=3$, $\hat{r}=3$							
			Factor Tests: Size								Factor Tests: Power x=0.1				
3	0.047	0.052	0.046	0.016	0.032	0.046	0.046	3	0.111	0.115	0.131	0.039	0.076	0.130	0.133
4	0.047	0.051	0.052	0.011	0.025	0.051	0.053	4	0.203	0.221	0.253	0.106	0.160	0.238	0.246
5	0.050	0.054	0.055	0.016	0.027	0.055	0.056	5	0.174	0.191	0.212	0.081	0.117	0.211	0.213
10	0.053	0.061	0.049	0.018	0.026	0.047	0.050	10	0.190	0.205	0.229	0.087	0.139	0.237	0.238
20	0.044	0.049	0.066	0.020	0.024	0.063	0.063	20	0.222	0.235	0.260	0.126	0.165	0.259	0.265
40	0.040	0.047	0.051	0.013	0.017	0.053	0.053	40	0.231	0.247	0.297	0.140	0.184	0.296	0.303
50	0.049	0.057	0.062	0.016	0.022	0.063	0.063	50	0.264	0.283	0.322	0.162	0.216	0.325	0.331
								$r=3$, $\hat{r}=2$							
			Factor Tests: Size								Factor Tests: Power x=0.1				
3	0.057	0.060	0.060	0.057	0.059	0.059	0.060	3	0.194	0.201	0.201	0.190	0.199	0.199	0.201
4	0.053	0.056	0.056	0.050	0.055	0.055	0.056	4	0.237	0.254	0.254	0.231	0.246	0.246	0.254
5	0.035	0.038	0.038	0.032	0.037	0.037	0.038	5	0.094	0.100	0.100	0.092	0.096	0.096	0.100
10	0.049	0.052	0.052	0.047	0.052	0.052	0.052	10	0.270	0.285	0.285	0.266	0.282	0.282	0.285
20	0.050	0.053	0.053	0.050	0.052	0.052	0.053	20	0.239	0.257	0.257	0.234	0.249	0.249	0.257
40	0.037	0.043	0.043	0.036	0.041	0.041	0.043	40	0.218	0.230	0.230	0.212	0.229	0.229	0.230
50	0.038	0.051	0.051	0.038	0.045	0.045	0.051	50	0.142	0.152	0.152	0.140	0.146	0.146	0.152
								$r=1$, $\hat{r}=2$							
			Factor Tests: Size								Factor Tests: Power x=0.1				
2	0.049	0.051	0.051	0.047	0.051	0.051	0.051	2	0.128	0.139	0.139	0.127	0.137	0.137	0.139
3	0.042	0.045	0.045	0.042	0.044	0.044	0.045	3	0.121	0.128	0.128	0.116	0.127	0.127	0.128
4	0.046	0.051	0.051	0.046	0.050	0.050	0.051	4	0.134	0.142	0.142	0.132	0.139	0.139	0.142
5	0.048	0.053	0.053	0.047	0.052	0.052	0.053	5	0.146	0.156	0.156	0.143	0.152	0.152	0.156
10	0.054	0.058	0.058	0.052	0.057	0.057	0.058	10	0.149	0.160	0.160	0.146	0.156	0.156	0.160
20	0.041	0.047	0.047	0.041	0.043	0.043	0.047	20	0.140	0.153	0.153	0.138	0.148	0.148	0.153
40	0.046	0.054	0.054	0.046	0.053	0.053	0.054	40	0.149	0.157	0.157	0.144	0.153	0.153	0.157
50	0.055	0.059	0.059	0.055	0.058	0.058	0.059	50	0.130	0.142	0.142	0.126	0.140	0.140	0.142
								$r=1$, $\hat{r}=2$							
			Standard Tests: Size								Standard Tests: Power x=0.1				
2	0.049	0.054	0.054	0.047	0.052	0.052	0.054	2	0.120	0.120	0.120	0.120	0.120	0.120	0.120
3	0.052	0.058	0.061	0.045	0.053	0.061	0.062	3	0.136	0.136	0.160	0.129	0.135	0.135	0.135
4	0.042	0.049	0.051	0.026	0.036	0.049	0.051	4	0.102	0.102	0.138	0.087	0.100	0.100	0.100
5	0.053	0.057	0.068	0.026	0.044	0.059	0.061	5	0.091	0.091	0.129	0.093	0.099	0.099	0.093
10	0.042	0.067	0.074	0.008	0.030	0.071	0.074	10	0.076	0.076	0.169	0.053	0.081	0.081	0.081
20	0.053	0.088	0.141	0.004	0.048	0.114	0.098	20	0.060	0.060	0.097	0.033	0.047	0.047	0.050
40	0.041	0.093	0.226	0.000	0.051	0.194	0.103	40	0.059	0.059	0.184	0.019	0.068	0.068	0.059
50	0.050	0.163	0.389	0.000	0.226	0.425	0.179	50	0.036	0.036	0.113	0.009	0.031	0.031	0.030

Note: The statistics under analysis are the factor based versions of the AR test, of its asymptotic version relying on the $\chi^2(k_2)/k_2$distribution (ARS), of the LR and LM tests proposed by Wang and Zivot (1998), of the K-test by Kleibergen (2002), and of the two versions of the conditional LR test (LR1 and LR2) of Moreira (2003).The Data Generating Process is described in Section 6. K2 indicates the number of instruments. r denotes the true number of factors and $\hat{r}$ denotes the assumed number of factors.

Table 7. Inflation models						
Equation: (23), s_t proxy = unemployment						
Instruments	Baseline Set			Factors		
Parameter	Estimate	Interval	Max-P	Estimate	Interval	Max-P
λ	-0.01	[-0.54, -0.01]	1.00	-0.01	[-0.08, -0.01]	0.800
γ_f	0.07	[0.02, 0.99]		0.87	[0.01, 0.99]	
γ_b	0.95	[0.01, 0.99]		0.15	[0.02, 0.99]	
Equations: (23)-(24), s_t proxy = output gap						
Instruments	Baseline Set			Factors		
Parameter	Estimate	Interval	Max-P	Estimate	Interval	Max-P
λ	0.00	[0.00, 0.99]	1.00	0.01	[0.00, 0.23]	0.993
ϕ	0.90	[0.00, 0.99]		0.94	[0.00, 0.99]	
ϑ	0.91	[0.00, 0.99]		0.99	[0.00, 0.99]	
Equations: (23)-(25), s_t proxy = marginal cost						
Instruments	Baseline Set			Factors		
Parameter	Estimate	Interval	Max-P	Estimate	Interval	Max-P
ω	0.82	[0.01, 0.90]	1.00	0.74	[0.01, 0.90]	0.998
θ	0.95	[0.53, 0.99]		0.90	[0.53, 0.99]	
β	0.99	[0.01, 0.99]		0.99	[0.01, 0.99]	

Note: The models and estimation procedures are described in Section 7.