Tests for Overidentifying Restrictions in Factor-Augmented VAR Models

Xu Han*

Job Market Paper

This Version: January 2012

Abstract

This paper develops tests for overidentifying restrictions in Factor-Augmented Vector Autoregressive (FAVAR) models. The FAVAR combines a high-dimensional factor model and a conventional VAR for the latent factors. The identification of structural shocks in FAVAR can lead to restrictions on the factor loadings of many variables, so it can involve infinitely many identifying restrictions as the number of cross sections goes to infinity. Our focus is to test the joint null hypothesis that all the restrictions are satisfied. Conventional tests cannot be used due to the large dimension. We transform the infinite-dimensional problem into a finite-dimensional one by combining the individual statistics across the cross section dimension. We find the limit distribution of our joint test statistic under the null hypothesis and prove that it is consistent against the alternative that a fraction of or all identifying restrictions are violated. The Monte Carlo results show that the joint test statistic has good finite-sample size and power. We implement our tests to an updated version of Stock and Watson's (2005) data set. The proposed test rejects the null hypotheses that the number of fast shocks is two or more, but does not reject the null that there is only one fast shock, which is the monetary policy shock. This result is further confirmed by the impulse responses of major macroeconomic variables to the monetary policy shock: the impulse responses based on one fast shock are generally more economically plausible than those based on two or more fast shocks; and the price puzzle is either considerably reduced or entirely solved for all price indexes when only one fast shock is used.

Key words: Factor Model, Identification, Structural Impulse Responses, Monetary Policy Shock

JEL Classification: C33, C52, E37, E52

*Department of Economics, Box 8110, North Carolina State University, Raleigh, NC 27695-8110. E-mail: xhan@ncsu.edu.

1 Introduction

Dynamic Factor Models (DFM) have received more and more attention from empirical macroeconomists because a few number of factors can explain a substantial amount of variations of major economic variables (Sargent and Sims, 1977; Stock and Watson, 2002; Giannone, Reichlin, and Sala, 2004). One important application of factor models is the Factor-Augmented Vector Autoregressive (FAVAR) analysis introduced by Bernanke, Boivin, and Eliasz (2005; BBE hereafter). They introduced factors estimated from large panel data sets into conventional VAR models. The main advantage of FAVAR relative to conventional VAR is that it utilizes the information in high-dimensional data sets to identify the space spanned by the structural shocks without the loss of parsimony. Due to such advantages, FAVAR has become increasingly popular in the empirical macroeconomic literature during recent years (for example, Stock and Watson (2005); Boivin, Giannoni, and Mihov (2009); Gilchrist, Yankov and Zakraisek (2009); Bianchi, Mumtaz and Surico (2009); Forni and Gambetti (2010); Eickmeier, Lemke, and Marcellino (2011), among others).

In spite of the various applications of FAVAR, the existing literature has not addressed the overidentification problem in FAVAR models except Stock and Watson (2005). Unlike the conventional structural VAR analysis where the number of restrictions for identifying structural shocks is usually small, the FAVAR tends to involve a large number of identifying restrictions. This is because factors are estimated from many macroeconomic time series and imposing an identification condition on a factor results in imposing an identification condition on each of the macroeconomic time series used to estimate the factor. In such a setup, the number of identifying restrictions is much larger than the number of structural shocks and the system is highly overidentified. If a substantial amount of these restrictions are violated, then it is likely that the structural shocks cannot be consistently estimated, and policy analysis such as impulse response functions based on a misspecified identification scheme may produce very misleading results.

In this paper, we consider testing the joint null hypothesis that all the identifying restrictions are satisfied against the alternative that a non-negligible fraction of or all restrictions are violated. We follow the setup of Stock and Watson (2010) and the identification scheme of Stock and Watson (2005; SW hereafter). SW conduct a structural analysis of monetary policy by imposing contemporaneous timing restrictions in the FAVAR model. They classify all the variables into three groups: slow variables (GDP, wages, etc), the federal funds rate, and fast variables (asset returns, etc). They also partition the structural shocks into three groups: slow shocks, the monetary policy shock, and fast shocks. The identification scheme of SW assumes that the slow variables are not affected by the monetary policy shock and fast shocks within the same month, that the federal funds rate is not affected by the fast shocks within the same month, and that the fast variables can respond contemporaneously to all shocks. Since the number of slow variables is large, SW's scheme imposes many zero restrictions on the coefficients of the structural MA representation of FAVAR. Our focus is to develop a joint test to check whether these zero restrictions hold simultaneously.

Conventional overidentification tests are usually designed in a finite-dimensional framework, so they cannot be applied to test the null hypothesis that involves parameters whose number goes to infinity as the sample size grows. As argued by Han and Inoue (2011), directly accommodating the conventional test to the infinite dimension case is technically challenging for several reasons. First, to construct such statistics, one needs to estimate an infinite-dimensional covariance matrix, but the norm of the difference between the estimated and true covariance matrices can be very large even if each entry of the estimated matrix converges in probability. Second, taking the inverse of such a high dimensional matrix will amplify the estimation error dramatically and lead to very inaccurate results (Ledoit and Wolf, 2004). Also, if the time dimension is smaller than the dimension of the covariance matrix, then the sample covariance computed in a conventional way is singular. Finally, unlike the finite dimension test statistics that have Chi-square limit distributions, the limit distribution of overidentification test statistics with infinite degrees of freedom, even if it is well-defined, is likely to be nonstandard.

To the best of our knowledge, SW is the only one that considers testing the overidentifying restrictions in FAVAR models. They regress each slow variable on the estimated structural shocks and test whether the coefficients on the monetary policy shock and fast shocks are zeros by conventional Wald statistics. The main problem of this method is that it is an equation-by-equation test and it cannot control the overall type I error for the joint null hypothesis that all identifying restrictions hold simultaneously. Hence, if some of the equation-by-equation statistics rejects the null hypothesis but some do not, then it will be difficult to make a decision whether we should reject the joint null hypothesis or not. In their empirical analysis based on a post-war monthly US data, SW establish an FAVAR model with four slow shocks, a monetary policy shock and two fast shocks. They implement their methods to test the coefficients on the monetary policy shock and fast shocks in 67 equations and they find that 49 test statistics reject the null hypothesis at the 5% level. The rejection rate, $73.1\% = 49/67$, seems to be too high compared to 5%, so it somehow indicates that the setup is possibly misspecified. However, since these equation-by-equation statistics are not independent, it is still hard to say whether the joint null hypothesis should be rejected or not based on the rejection rate. Thus, it is necessary to develop a joint test for all identification restrictions.

The contribution of this paper is threefold. First, we propose a new equation-by-equation statistic and establish its limit theory. The intuition of this statistic is similar to that of SW, but the estimation procedure is different from theirs. SW does not provide a formal proof for the limit distribution of their statistic, and it may be difficult to formally prove their result because their statistic is based on a high-dimensional reduced rank regression which involves inverting a large singular matrix when the time dimension is less than the number of slow variables. This paper circumvents this difficulty and develops a new statistic with theoretically justified limit theory both under the null and alternative hypotheses. Second, based on the proposed equation-by-equation statistic, this paper develops a statistic that can test the joint null hypothesis that all identifying restrictions hold simultaneously. To the best of our knowledge, this is the first test that deals with such kind of null hypothesis in the literature. The intuition is to combine the equation-by-equation statistics across the cross section dimension, so that the infinite-dimensional problem reduces to a finite-dimensional one. Under some regularity conditions, we establish the limit theory under the null hypothesis and prove that the joint test statistic is consistent against the alternative that a substantial amount of identifying restrictions are violated. The Monte Carlo results show that the joint test statistic has good finite-sample size and power. Finally, this paper extends the theoretical results found by Bai (2003). He shows that the factors estimated from observed data can be treated as if they were observed as long as $\sqrt{T}/N \to 0$ and N, $T \to \infty$, where N and T denote the cross section and time dimensions, respectively. In SW's FAVAR setup, not only the factors but also data are estimated, i.e. the factors are estimated from estimated data. We show that the main results of Bai still hold under some regularity conditions. This result is a by-product when we try to establish the limit theory for our statistics, but we believe that it would be useful for future research in the literature of dynamic factor models.

Although the main focus of this paper is on SW's contemporaneous timing restrictions, our test can be applied to test other similar identifying restrictions in the literature as well. For example, Gilchrist, Yankov and Zakraisek (2009) use a setup very close to that of SW to investigate the impulse responses of credit spreads and macroeconomic variables to the credit shock. They assume that the credit shock does not affect the macroeconomic variables contemporaneously, which can be tested using our statistics. Furthermore, we can test BBE's identifying restrictions, even though they are slightly different from those of SW. BBE assume that the monetary policy shock does not contemporaneously affect other factors that are estimated from slow variables, and then they estimate a structural VAR model that consists of the federal funds rate and other factors. Although this VAR itself is just identified, the factors other than the monetary policy instrument are estimated under the assumption that slow variables are not contemporaneously affected by the monetary policy shock, which implies a larger number of restrictions. In this sense, it is similar to the identification scheme of SW and these restrictions are testable by our statistics[1].

We implement our tests to an updated version of SW's data set. The tests reject the null hypotheses that the number of fast shocks is two or three, but they do not reject the null that there is only one fast shock, which is the monetary policy shock by definition. Interestingly, this result provides some evidence to support the BBE's identifying restrictions in which factors other than the monetary policy shock are assumed to be slow. We also check the number of slow and fast shocks by information criteria, which should provide consistent estimates as N and $T \to \infty$. However, the information criteria find a contradictory result that the number of slow shocks is greater than the total number of structural shocks. The reason is that the information criteria could lead to biased estimates in a finite sample. In such a scenario, our testing procedure is the only way to evaluate the identifying restrictions in FAVAR.

Furthermore, we compute the impulse responses of major macroeconomic variables to the monetary policy shock based on $\kappa_F = 1$ and 2, where κ_F denotes the potential number of fast shocks. It turns out that the impulse responses based on $\kappa_F = 1$ are generally more economically plausible than those based on $\kappa_F = 2$. Moreover, we investigate the impulse responses of different price indexes. Compared to $\kappa_F = 2$ which leads to persistent positive responses to a contractionary monetary policy shock, $\kappa_F = 1$ either substantially reduces or completely solves the price puzzle (Sims, 1992) in all price indexes. Hence, these results confirm that our tests are useful to select correct identifying restrictions in FAVAR models.

The rest of the paper is organized as follows. Section 2 briefly describes the setup of FAVAR and the contemporaneous timing restrictions considered by SW. Section 3 proposes the equation-by-equation and joint test statistics for the overidentifying restrictions in the FAVAR, and the asymptotic properties are established under the null and alternative hypotheses. Section 4 investigates the finite-sample size and power of our statistics using Monte Carlo experiments. Section 5 provides an empirical application of our test statistics

[1] Besides the above examples, our test statistics can be also applied to test the identifying restrictions in other papers that use similar setup of BBE (for example, Mumtaz and Surico (2009); Eickmeier, Lemke and Marcellino (2011)).

and computes the impulse response functions for major macroeconomic variables based on a monthly US data set. Section 6 concludes.

2 The FAVAR Models and Contemporaneous Timing Restrictions

2.1 The FAVAR Models

In this subsection, we briefly review that setup of FAVAR models. Let $X_t = [X_{1t}, ..., X_{Nt}]'$ be an $N-$dimentional vector of stationary time series variables observed for $t = 1, ..., T$. Suppose that the number of common factors is q. The DFM can be expressed as:

$$X_{it} = \tilde{\lambda}_i(L) f_t + e_{it} \tag{2.1}$$

$$f_t = \phi(L) f_{t-1} + \eta_t \tag{2.2}$$

where f_t is the $q \times 1$ vector of common dynamic factors at period t, $\tilde{\lambda}_i(L)$ are the dynamic factor loadings for series i, consisting of a $1 \times q$ vector lag polynomial, η_t is a $q-$dimensional innovation for dynamic factors at time t, and e_{it} is the idiosyncratic shock for series i at period t.

Suppose that $\tilde{\lambda}_i(L)$ has a finite degree p_0. Define the static factor $F_t = [f_t' \ f_{t-1}' ... f_{t-p_0}']'$, where the dimension of F_t is $r \times 1$. The DFM has the following static representation:

$$X_t = \Lambda F_t + e_t \tag{2.3}$$

$$F_t = \Phi(L) F_{t-1} + G\eta_t \tag{2.4}$$

where Λ is the static factor loading matrix of F_t, $e_t = [e_{1t}, ..., e_{Nt}]'$, $\Phi(L)$ is $r \times r$ matrix lag polynomial, and G is $r \times q$ matrix of zeros and ones. Suppose e_{it} is modeled as an autoregressive process:

$$e_{it} = \delta_i(L) e_{it-1} + v_{it} \tag{2.5}$$

Let $D(L) = \begin{bmatrix} \delta_1(L) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \delta_N(L) \end{bmatrix}$, and $v_t = [v_{1t}, ..., v_{Nt}]'$, so we have

$$e_t = D(L)e_{t-1} + v_t \tag{2.6}$$

Combining Equations (2.3), (2.4) and (2.6), we can re-write the static representation in the following VAR form:

$$\begin{bmatrix} F_t \\ X_t \end{bmatrix} = \begin{bmatrix} \Phi(L) & 0 \\ \Lambda\Phi(L) - D(L)\Lambda & D(L) \end{bmatrix} \begin{bmatrix} F_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{F_t} \\ \varepsilon_{X_t} \end{bmatrix} \tag{2.7}$$

where

$$\begin{bmatrix} \varepsilon_{F_t} \\ \varepsilon_{X_t} \end{bmatrix} = \begin{bmatrix} I \\ \Lambda \end{bmatrix} G\eta_t + \begin{bmatrix} 0 \\ v_t \end{bmatrix} \tag{2.8}$$

If the VAR model (2.7) is invertible, then we have the MA form of the DFM:

$$X_t = B(L)\eta_t + e_t$$

where $B(L) = \Lambda[I - \Phi(L)L]^{-1}G$, which can be derived by substituting (2.4) into (2.3). Let ζ_t denote the q dimensional structural shocks to the dynamic factors, and $E\zeta_t\zeta_t' = I_q$. Assume that the reduced form innovations η_t are linear combinations of ζ_t in the following way:

$$A\eta_t = \zeta_t \tag{2.9}$$

Let $\tilde{B}(L) = B(L)A^{-1}$, then we have the structural MA representation for X_t:

$$X_t = \tilde{B}(L)\zeta_t + e_t \tag{2.10}$$

Note that if the structural shocks ζ_t is a vector white noise process, then the link between

Equations (2.1) and (2.10) is that the filters $\tilde{\lambda}_i(L)$'s are replaced with $\tilde{B}(L)$ such that the dynamic factors are transformed into orthonormal white noise processes. Thus, if we think of ζ_t as the transformed dynamic factors, then Model (2.10) is exactly the DFM considered by Forni et al. (2000).

2.2 Contemporaneous Timing Restrictions in FAVAR

Contemporaneous timing restrictions are commonly applied to identify the monetary policy shock in structural VAR models. It is always assumed that the monetary policy shock does not affect variables, such as output, consumptions, etc, within the same period. Such identification scheme implies that some of the coefficients in the structural MA must be zeros. In the conventional structural VAR analysis, variables are properly ordered such that the monetary policy shock can be identified using Cholesky decomposition.

Now, we consider SW's identification scheme for the monetary policy shock in the FAVAR model. Note that

$$\varepsilon_{X_t} = \tilde{B}_0 \zeta_t + v_t \tag{2.11}$$

where $\tilde{B}_0$ is an $N \times q$ coefficient matrix which is the zero-lag term in $\tilde{B}(L)$. By Equations (2.8) and (2.9), we have $\tilde{B}_0 = \Lambda G A^{-1}$. Thus, the contemporaneous timing restriction is the same as restricting some elements of $\tilde{B}_0$ to be zeros. SW categorize variables into three groups: slow variables, the federal funds rate and the fast variables; they also categorize shocks into three groups: slow shocks, the monetary policy shock and the fast shocks. They assume that the slow variables are only affected by the slow shocks contemporaneously, that the federal funds rate can be affected by slow and monetary policy shocks within the same period, and that the fast variables can be affected by all shocks contemporaneously. Let us use superscripts 'S', 'R', and 'F' to denote the three groups, respectively. Then we have $X_t = [X_t^{S'}, X_t^R, X_t^{F'}]'$, $\varepsilon_{X_t} = [\varepsilon_{X_t}^{S'}, \varepsilon_{X_t}^R, \varepsilon_{X_t}^{F'}]'$, and $\zeta_t = [\zeta_t^{S'}, \zeta_t^R, \zeta_t^{F'}]'$, where X_t^S and $\varepsilon_{X_t}^S$ are $N_S \times 1$ vectors, X_t^R, $\varepsilon_{X_t}^R$, and ζ_t^R are scalers, X_t^F and $\varepsilon_{X_t}^F$ are $N_F \times 1$ vectors, ζ_t^S is a $q_S \times 1$ vector, and ζ_t^F is a $q_F \times 1$ vector. Note that $N_S + N_F + 1 = N$ and $q_S + q_F + 1 = q$. Given the contemporaneous timing restrictions on $\tilde{B}_0$, Equation (2.11) can be rewritten as:

$$\begin{bmatrix} \varepsilon_{X_t}^S \\ \varepsilon_{X_t}^R \\ \varepsilon_{X_t}^F \end{bmatrix} = \begin{bmatrix} \tilde{B}_{0,SS} & 0_{N_S\times 1} & 0_{N_S\times q_F} \\ \tilde{B}_{0,RS} & \tilde{B}_{0,RR} & 0_{1\times q_F} \\ \tilde{B}_{0,FS} & \tilde{B}_{0,FR} & \tilde{B}_{0,FF} \end{bmatrix} \begin{bmatrix} \zeta_t^S \\ \zeta_t^R \\ \zeta_t^F \end{bmatrix} + v_t \tag{2.12}$$

Note that the number of zeros in $\tilde{B}_0$ is equal to $(1+q_F)\times N_S+q_F$, but identifying A only requires $q(q+1)/2$ restrictions. If $N_S \gg q$, then the system is overidentified. We are interested in whether all the zeros restrictions on $\tilde{B}_0$ in Equation (2.12) are satisfied or not. Let us partition $\tilde{B}_0$ into the following block matrix:

$$\tilde{B}_0 = \begin{bmatrix} \tilde{B}_{0,SS} & \tilde{B}_{0,SR} & \tilde{B}_{0,SF} \\ \tilde{B}_{0,RS} & \tilde{B}_{0,RR} & \tilde{B}_{0,RF} \\ \tilde{B}_{0,FS} & \tilde{B}_{0,FR} & \tilde{B}_{0,FF} \end{bmatrix}$$

Consider the following null hypothesis:

$$H_0^*: \ \tilde{B}_{0,SR} = 0_{N_s\times 1}, \ \tilde{B}_{0,SF} = 0_{N_s\times q_F}, \text{ and } \tilde{B}_{0,RF} = 0_{1\times q_F}$$

This null hypothesis involves testing an infinite number of restrictions at the limit. SW argue that the conventional statistics that have asymptotic Chi-square distribution are expected to have poor performance due to the large number of restrictions. Thus, they only implement equation-by-equation hypothesis testing for Equation (2.12). Their null hypothesis can be written as

$$\begin{aligned} H_0^{SW}: \quad & \text{for given } i \\ & \begin{cases} \tilde{B}_{0,SR}^i = 0 \text{ and } \tilde{B}_{0,SF}^i = 0_{1\times q_F} & \text{if } i \in \text{slow group} \\ \tilde{B}_{0,RF} = 0_{1\times q_F} & \text{if } i = R \end{cases} \end{aligned}$$

where $\tilde{B}_{0,SR}^i$ and $\tilde{B}_{0,SF}^i$ denote the i^{th} row of $\tilde{B}_{0,SR}$ and $\tilde{B}_{0,SF}$, respectively, and $i=R$ means that the i^{th} variable is the federal funds rate.

It is worth noting that as $N_S \to \infty$, there will always exist some equation specific test statistics that reject the null H_0^{SW}, even if H_0^* is true. For example, if all the equation specific test statistics are independent, then about 5% of the statistics will reject the null hypothesis based on the critical value at the 5% level. Thus, even if the equation specific test rejects the null hypothesis H_0^{SW} for some i, it does not necessarily mean that we should also reject the null hypothesis H_0^*.

3 Testing the Overidentifying Restrictions

In the rest of the paper, we consider the following model:

$$X_t = \Lambda F_t + e_t \tag{3.1}$$

$$F_t = \sum_{j=1}^{p} \Phi_j F_{t-j} + G\eta_t \tag{3.2}$$

where X_t is an $N-$dimensional vector, F_t is the $r-$dimensional static factor at time t, $\Lambda = [\lambda_1, ..., \lambda_N]'$ is the static factor loading matrix ($N \times r$), Φ_j's are autoregressive coefficients of F_t, G is a $r \times q$ matrix, and η_t is a $q-$dimensional innovation of F_t with $q \leq r$. Unlike the derivation of the FAVAR model (2.7), quasi-demeaning Equation (3.1) is not necessary to test the overidentifying restrictions, so we do not make assumptions (such as Equation (2.6)) on the dynamics of the idiosyncratic components e_t.

Let $\mathcal{F}_t = [F'_{t-1}, ..., F'_{t-p}]'$, $\Phi = [\Phi_1, ..., \Phi_p]$, and η_t is linked to the structural shocks ζ_t by Equation (2.9). Substituting Equations (3.2) and (2.9) into Equation (3.1), we have:

$$X_t = \Pi\mathcal{F}_t + \Gamma\zeta_t + e_t \qquad \text{for t = p + 1, ..., T} \tag{3.3}$$

where $\Pi = \Lambda\Phi$ and $\Gamma = \Lambda G A^{-1}$. The matrix form representation of Equation (3.3) is:

$$X = \mathcal{F}\Pi' + \zeta\Gamma' + e \tag{3.4}$$

where $X = [X_{p+1}, ..., X_T]'$, $\mathcal{F} = [\mathcal{F}_{p+1}, ..., \mathcal{F}_T]'$, $\zeta = [\zeta_{p+1}, ..., \zeta_T]'$, and $e = [e_{p+1}, ..., e_T]'$.

In this paper, variables are classified as slow and fast variables, and structural shocks are classified as slow and fast shocks. The fast shocks do not have contemporaneous impacts on slow variables, whereas slow shocks are allowed to have contemporaneous impacts on both slow and fast variables. By this definition, the monetary policy shock is treated as a fast shock and the federal fund rate is a member of the fast variables. This is different from Equation (2.12), which uses a three-group setup and imposes the restriction that $\tilde{B}_{0,RF} = 0$ to identify the monetary policy shock. We use a two-group setup for three reasons: first, classifying the monetary policy shock as a fast shock does not affect the zero restrictions on the factor loadings of slow variables. Second, classifying the federal funds rate as a fast variable only eliminates the restriction that the fast shocks should not have contemporaneous effects on the federal funds rate, but this involves only a fixed number of restrictions, which can be handled by a conventional testing procedure with finite dimension. Finally, the following section shows that jointly testing the overidentifying restrictions requires estimates of the fast shocks. Since the estimates of factors are extracted from a large number of variables, adding one more variable in the fast group will not affect the estimates asymptotically. Hence, we focus on the following model:

$$\begin{bmatrix} X_t^S \\ X_t^F \end{bmatrix} = \begin{bmatrix} \Pi^S \\ \Pi^F \end{bmatrix} \mathcal{F}_t + \begin{bmatrix} \Gamma^{SS} & \Gamma^{SF} \\ \Gamma^{FS} & \Gamma^{FF} \end{bmatrix} \begin{bmatrix} \zeta_t^S \\ \zeta_t^F \end{bmatrix} + \begin{bmatrix} e_t^S \\ e_t^F \end{bmatrix} \tag{3.5}$$

where X_t^S is an $N_S \times 1$ vector of slow variables, X_t^F is an $N_F \times 1$ vector of fast variables, e_t^S and e_t^F are idiosyncratic shocks of X_t^S and X_t^F, respectively, ζ_t^S is a $q_S \times 1$ vector of slow structural shocks, and ζ_t^F is a $q_F \times 1$ vector of fast structural shocks. In this setup, we have $N_S + N_F = N$, $q_S + q_F = q$, $X_t = [X_t^{S'}, X_t^{F'}]'$, $\zeta_t = [\zeta_t^{S'}, \zeta_t^{F'}]'$, and $e_t = [e_t^{S'}, e_t^{F'}]'$. Equation (3.5) can be expressed in the following matrix form:

$$X^S = \mathcal{F}\Pi^{S'} + \zeta^S \Gamma^{SS'} + \zeta^F \Gamma^{SF'} + e^S \tag{3.6}$$

$$X^F = \mathcal{F}\Pi^{F\prime} + \zeta^S\Gamma^{FS\prime} + \zeta^F\Gamma^{FF\prime} + e^F \tag{3.7}$$

where $X^S = [X^S_{p+1}, ..., X^S_T]'$ is a $(T-p) \times N^S$ matrix of slow variables, $X^F = [X^F_{p+1}, ..., X^F_T]'$ is a $(T-p) \times N^F$ matrix of fast variables, $\zeta^S = [\zeta^S_{p+1}, ..., \zeta^S_T]'$ is a $(T-p) \times q^S$ matrix of slow structural shocks, $\zeta^F = [\zeta^F_{p+1}, ..., \zeta^F_T]'$ is a $(T-p) \times q^F$ matrix of fast structural shocks, $e^S = [e^S_{p+1}, ..., e^S_T]'$ is the idiosyncratic shocks of slow variables, and $e^F = [e^F_{p+1}, ..., e^F_T]'$ is the idiosyncratic shocks of fast variables. Note that $X = [X^S \vdots X^F]$, $\zeta = [\zeta^S \vdots \zeta^F]$ and $e = [e^S \vdots e^F]$.

Note that $X - \mathcal{F}\Pi = \zeta\Gamma' + e$ follows a factor structure, where the structural shocks ζ is a $q-$dimensional factor for $X - \mathcal{F}\Pi$. Amengual and Watson (2007) show that the number of factors in $X - \mathcal{F}\Pi$ can be consistently estimated by implementing Bai and Ng's (2002) information criteria on $X - \hat{\mathcal{F}}\hat{\Pi}$, where $\hat{\mathcal{F}}$ is constructed by stacking the principal component estimator $\hat{F}$ and $\hat{\Pi}$ is the OLS estimator from regression of X_t on $\hat{\mathcal{F}}_t$. Thus, we treat q as known in the rest of the this section[2].

Now, we impose the contemporaneous timing restriction that the fast shock ζ^F_t does not affect X^S_t i.e. $\Gamma^{SF} = 0$. This implies that $X^S - \mathcal{F}\Pi^{S\prime} = \zeta^S\Gamma^{SS\prime} + e^S$ follows a q_S-factor structure, in contrast with the full sample $X - \mathcal{F}\Pi$ that follows a $q-$factor structure. Hence, testing zeros coefficient restrictions in Equation (3.6) is same as comparing the numbers of factors in $X^S - \mathcal{F}\Pi^{S\prime}$ and $X - \mathcal{F}\Pi$.

One noteworthy thing is that Equations (3.5), (3.6) and (3.7) are the true model and q_F is an unknown parameter $\in \{0, 1, ..., q\}$. If $q_F = 0$, then all the structural shocks are slow in the sense that they can affect X^S_t contemporaneously, so Γ^{SF} and Γ^{FF} are $N_S \times 0$ and $N_F \times 0$ matrices, respectively. If $q_F = q$, then all the structural shocks are fast as they do not affect X^S_t contemporaneously, so Γ^{SF} is an $N_S \times q$ zero matrix and Γ^{SS} and Γ^{FS} are $N_S \times 0$ and $N_F \times 0$ matrices, respectively. If $1 \leq q_F \leq q-1$, then slow factors and fast factors co-exist, and Γ^{SF} is an $N_S \times q_F$ zero matrix, which is the true identifying restriction that one should

[2] q can be also consistently determined by methods proposed by Bai and Ng (2007) and Hallin and Liska (2007)

impose if q_F were known. To test the value of q_F, we consider the following hypotheses:

$$\begin{aligned} H_0 : \; q_F &= \kappa_F \\ H_1 : \; q_F &< \kappa_F \end{aligned} \tag{3.8}$$

κ_F is the candidate value of q_F, and it means that we impose $N_S \times \kappa_F$ zero restrictions in the upper right corner of Γ to identify the structural shocks, so (3.3) becomes:

$$X_t = \Pi \mathcal{F}_t + \begin{bmatrix} * & 0_{N_S \times \kappa_F} \\ * & * \end{bmatrix} \zeta_t + e_t \tag{3.9}$$

where the asterisks denote unrestricted entries in Γ. We set $\kappa_F \in \{1, ..., q\}$, and $\kappa_F = 0$ is ruled out because the structural shocks cannot be identified if no restriction is imposed. Hence, q_F is always greater than zero under the null hypothesis H_0. Under the alternative hypothesis H_1, however, $q_F = 0$ is allowed, indicating that no fast shock exists for the current classification of X^S and X^F. If the test rejects H_0: $q_F = 1$ in favor of H_1: $q_F < 1$, then one may need to consider re-classifying the slow and fast variables.

Remarks:

(1) The hypotheses in (3.8) transform the infinite-dimensional problem to a finite-dimensional one. The original null hypothesis that $\Gamma^{SF} = 0_{N_S \times \kappa_F}$ is slightly stronger the null considered in (3.8). To see this, suppose that Γ^{SF} is $N_S \times \kappa_F$ and that a fixed number of entries in Γ^{SF} are non-zero, so $\Gamma^{SF} \neq 0_{N_S \times \kappa_F}$ and the original null hypothesis should be rejected. On the contrary, the null hypothesis that $q_F = \kappa_F$ still holds, because a fixed number of non-zero entries in Γ^{SF} will not change the number of factors in $X^S - \mathcal{F}\Pi^{S'}$. In fact, only a fixed number of non-zeros in Γ^{SF} will not affect the principal component estimates in large samples, so we do not want a test that that is powerful against a very small number of violations in Γ^{SF}. In this sense, the transformed null and alternative hypotheses are more favorable than the original ones. Also, the transformed the hypotheses do cover the cases where we want

the statistics to have power. For instance, if a fraction of entries in Γ^{SF} are non-zeros so that the number of factors in $X^S - \mathcal{F}\Pi^{S'}$ is greater than $q - \kappa_F$, then H_0 in (3.8) will be rejected in favor of H_1.

(2) Onatski (2009) proposes a test for the number of factors in large factor models. However, this method is not applicable to test our hypotheses (3.8) for two technical reasons. First, testing (3.8) implies that we need to compare the dimensions of ζ_t^S and ζ_t, which cannot be estimated from $X^S - \mathcal{F}\Pi^{S'}$ and $X - \mathcal{F}\Pi$ because $\mathcal{F}$ and Π are not observed. Onatski's test would be applicable if $X^S - \mathcal{F}\Pi^{S'}$ and $X - \mathcal{F}\Pi$ were observed. Since we can only use the feasible analogs, $X^S - \hat{\mathcal{F}}\hat{\Pi}^{S'}$ and $X - \hat{\mathcal{F}}\hat{\Pi}$, to estimate ζ_t^S and ζ_t, one must take into account the estimation errors in $\hat{\mathcal{F}}$ and $\hat{\Pi}$, which may change the limit distribution of Onatski's test statistic. Second, Onatski (2009) proposes both dynamic and static versions of statistics for the number of factors. On the one hand, the static version requires the the idiosyncratic component to be Gaussian and serially uncorrelated, which are very strong assumptions compared to the literature on factor models. On the other hand, the dynamic version requires $N = o(T^{1/2-1/d} \log^{-1} T)^{6/13}$ for some $d > 2$ if no Gaussianity is imposed. Bai and Ng (2006) show that when $\sqrt{T}/N \to 0$ as N and $T \to \infty$, the estimated factors can be treated as if they were observed, and this nice property is widely used in the FAVAR literature. It is clear that Onatski's condition on the relative rate between N and T contradicts Bai and Ng's condition. Thus, to maintain the Bai-Ng property, one cannot apply Onatski's statistic to test (3.8).

3.1 The Statistics

In this subsection, we will propose statistics to test the the null hypothesis that $H_0: \; q_F = \kappa_F$ against the alternative hypothesis that $H_1: \; q_F < \kappa_F$. Let us first define some notations: for any matrix Z, the projection matrix, denoted as P_Z, is set equal to $Z(Z'Z)^{-1}Z'$, and the residual maker, denoted as M_Z, is set equal to $I - P_Z$. The statistics are computed using the following steps:

(1) The estimated static factors, denoted as $\hat{F}$, is $\sqrt{T}$ times the eigenvectors corresponding to the r largest eigenvalues of the $T \times T$ matrix XX'. Let $\hat{F}_t$ denote the transpose of the t^{th} row of $\hat{F}$. Define $\hat{\mathcal{F}}_t = [\hat{F}'_{t-1}, ..., \hat{F}'_{t-p}]'$ and $\hat{\mathcal{F}} = [\hat{\mathcal{F}}_{p+1}, ..., \hat{\mathcal{F}}_T]'$.

(2) Define $\hat{X} = M_{\hat{\mathcal{F}}} X$, $\hat{X}^S = M_{\hat{\mathcal{F}}} X^S$, and $\hat{X}^F = M_{\hat{\mathcal{F}}} X^F$. Set the estimated slow structural shocks, denoted as $\hat{\zeta}^S$, equal to $\sqrt{T-p}$ times the eigenvectors corresponding to the $\kappa_S = q - \kappa_F$ largest eigenvalues of the $(T-p) \times (T-p)$ matrix $\hat{X}^S \hat{X}^{S\prime}$. Let $\hat{\zeta}_t^S$ denote the transpose of the t^{th} row of $\hat{\zeta}^S$.

(3) Define $\tilde{X} = M_{\hat{\zeta}^S} \hat{X}$, $\tilde{X}^S = M_{\hat{\zeta}^S} \hat{X}^S$, and $\tilde{X}^F = M_{\hat{\zeta}^S} \hat{X}^F$. Set the estimated fast structural shocks, denoted as $\hat{\zeta}^F$, equal to $\sqrt{T-p}$ times the eigenvectors corresponding to the κ_F largest eigenvalues of the $(T-p) \times (T-p)$ matrix $\tilde{X}\tilde{X}'$. Let $\hat{\zeta}_t^F$ denote the transpose of the t^{th} row of $\hat{\zeta}^F$.

(4) Let $X_i = [X_{i(P+1)}, ..., X_{iT}]'$ denote the observations of the i^{th} variable. Accordingly, $\hat{X}_i = M_{\hat{\mathcal{F}}} X_i$ and $\tilde{X}_i = M_{\hat{\zeta}^S} \hat{X}_i$. Define $\tilde{e}_i = M_{\hat{\zeta}^F} \tilde{X}_i$. Let $i \in S$ abbreviate that the i^{th} variable belongs to the slow group. Define the individual statistic for the i^{th} variable:

$$w_i = \tilde{X}_i' \hat{\zeta}^F \hat{\Omega}_i^{-1} \hat{\zeta}^{F\prime} \tilde{X}_i / (T-p), \quad i \in S$$

where $\hat{\Omega}_i = (T-p)^{-1} \sum_{t=p+1}^{T} \hat{\zeta}_t^F \hat{\zeta}_t^{F\prime} \tilde{e}_{it}^2$.

(5) Define the joint statistic for all slow variables:

$$W = \left(\sum_{i \in S} \tilde{X}_i' \hat{\zeta}^F \right) \hat{\Omega}^{-1} \left(\hat{\zeta}^{F\prime} \sum_{i \in S} \tilde{X}_i \right) / (T-p) N_S$$

where $\hat{\Omega} = (T-p)^{-1} N_S^{-1} \sum_{t=p+1}^{T} \sum_{i \in S} \hat{\zeta}_t^F \hat{\zeta}_t^{F\prime} \tilde{e}_{it}^2$.

3.2 Asymptotics under the Null Hypothesis

Let $M < \infty$ and $m \in (0, 1)$ be constants that do not depend on N or T.

Assumption 1:

(a) $E\|F_t\|^4 < M$, $T^{-1}\sum_{t=1}^{T} F_t F_t' \to_p \Sigma_F$, and $T^{-1}\sum_{t=p+1}^{T} \mathcal{F}_t \mathcal{F}_t' \to_p \Sigma_{\mathcal{F}}$, as $T \to \infty$ for some positive definite matrices Σ_F and $\Sigma_{\mathcal{F}}$.

(b) $E(\zeta_t \zeta_t') = I_q$, $E\|\zeta_t\|^4 < M$, $E(\zeta_s \zeta_t') = 0$ for any $s \neq t$, and $T^{-1}\sum_{p+1}^{T} \zeta_t \zeta_t' \to_p \; I_q$.

(c) $E\| \sum_{t=p+1}^{T} \zeta_t^S \zeta_t^{F'} / \sqrt{T-p} \|^2 < M$ for $q_F > 0$ and $E\| \sum_{t=p+1}^{T} \zeta_t \mathcal{F}_t' / \sqrt{T-p} \|^2 < M$.

Assumption 2:

(a) $E\|\lambda_i\|^4 \leq M$, and $\|\Lambda'\Lambda/N - \Sigma_\Lambda\| \to_p 0$ for some $r \times r$ positive definite matrix Σ_Λ.

(b) G has rank q. $\|G\| \leq M$ and $\|\Phi\| \leq M$. A is non-singular and $\|A\| \leq M$.

(c) $\|\Gamma^{SS'}\Gamma^{SS'}/N_S - \Sigma_{\Gamma^{SS}}\| \to_p 0$ and $\|\Gamma^{FF'}\Gamma^{FF}/N_F - \Sigma_{\Gamma^{FF}}\| \to_p 0$ for some positive definite matrices $\Sigma_{\Gamma^{SS}}(q_S \times q_S)$ and $\Sigma_{\Gamma^{FF}}(q_F \times q_F)$ for $q_F > 0$, and $N_S/N \to m$.

Assumption 3:

(a) $E(e_{it}) = 0$, $E\,|e_{it}|^8 \leq M$ for all i and t.

(b) $E(e_s' e_t / N) = E(N^{-1}\sum_{i=1}^{N} e_{is} e_{it}) = \gamma_N(s,\, t)$, $|\gamma_N(s,\, s)| \leq M$ for all s, and $T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}$ $|\gamma_N(s,\, t)| \leq M$. $E(e_s^{S'} e_t^S / N_S) = E(N_S^{-1}\sum_{i\in S} e_{is} e_{it}) = \gamma^S(s,\, t)$, $|\gamma^S(s,\, s)| \leq M$ for all s, and $T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T} |\gamma^S(s,\, t)| \leq M$.

(c) $E(e_{it} e_{k,t-j}) = \tau_{ik,t,j}$ with $|\tau_{ik,t,j}| \leq |\tau_{ik}|$ for some τ_{ik} and for all t and $j = 0, ..., p$. In addition, $N^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N} |\tau_{ik}| \leq M$.

(d) $E(e_{it} e_{ks}) = \tau_{ik,ts}$, and $(NT)^{-1}\sum_{i=1}^{N}\sum_{k=1}^{N}\sum_{s=1}^{T}\sum_{t=1}^{T} |\tau_{ik,ts}| \leq M$.

(e) for every (t, s), $E\left| N^{-1/2}\sum_{i=1}^{N}[e_{is} e_{it} - E(e_{is} e_{it})] \right|^4 \leq M$ and $E\left| N_S^{-1/2}\sum_{i\in S}[e_{is} e_{it} - E(e_{is} e_{it})] \right|^4$ $\leq M$.

(f) for each t, $E(N_S^{-1/2}\sum_{i\in S} e_{it})^2 \leq M$.

Assumption 4:

λ_i, ζ_t, and e_{it} are mutually independent groups.

Assumption 5: For every $t \leq T$ and for every $i \leq N$:

(a) $\sum_{s=1}^{T} |\gamma_N(s,\, t)| \leq M$, and $\sum_{s=1}^{T} |\gamma^S(s,\, t)| \leq M$.

(b)$\sum_{k=1}^{N} |\tau_{ki}| \leq M$.

Assumption 6:

(a) for each t, $E\left\|\frac{1}{\sqrt{NT}}\sum_{s=1}^{T}\sum_{k=1}^{N} F_s[e_{ks}e_{kt} - E(e_{ks}e_{kt})]\right\|^2 \leq M$, $E\left\|\frac{1}{\sqrt{NT}}\sum_{s=p+1}^{T}\sum_{k=1}^{N} \zeta_s[e_{ks}e_{kt} - E(e_{ks}e_{kt})]\right\|^2 \leq M$, and $E\left\|\frac{1}{\sqrt{N_S T}}\sum_{s=p+1}^{T}\sum_{k\in S} \zeta_s[e_{ks}e_{kt} - E(e_{ks}e_{kt})]\right\|^2 \leq M$.

(b) for each i and for $j = 0, 1, ..., p$, $E\left\|\frac{1}{\sqrt{NT}}\sum_{t=p+1}^{T}\sum_{k=1}^{N} \lambda_k\left[e_{k,t-j}e_{it} - E(e_{k,t-j}e_{it})\right]\right\|^2 \leq M$. For each i, $E\left\|\frac{1}{\sqrt{N_S T}}\sum_{t=p+1}^{T}\sum_{k\in S} \lambda_k\left[e_{kt}e_{it} - E(e_{kt}e_{it})\right]\right\|^2 \leq M$.

(c)$E\left\|\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}\sum_{k=1}^{N} F_t\lambda_k' e_{kt}\right\| \leq M$. For $j = 0, 1, ..., p$, $E\left\|\frac{1}{\sqrt{NT}}\sum_{t=p+1}^{T}\sum_{k=1}^{N} \lambda_k e_{k,t-j}\mathcal{F}_t'\right\|^2 \leq M$, $E\left\|\frac{1}{\sqrt{NT}}\sum_{t=p+1}^{T}\sum_{k=1}^{N} \lambda_k e_{k,t-j}\zeta_t'\right\|^2 \leq M$, and $E\left\|\frac{1}{\sqrt{N_S T}}\sum_{t=p+1}^{T}\sum_{k\in S} \lambda_k e_{kt}\mathcal{F}_t'\right\|^2 \leq M$, and $E\left\|\frac{1}{\sqrt{N_S T}}\sum_{t=p+1}^{T}\sum_{k\in S} \lambda_k e_{kt}\zeta_t'\right\|^2 \leq M$.

(d) for each t, $E\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} \lambda_i e_{it}\right\|^2 \leq M$, and $E\left\|\frac{1}{\sqrt{N_S}}\sum_{i\in S} \lambda_i e_{it}\right\|^2 \leq M$.

(e) for each i, $E\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T} F_t e_{it}\right\|^4 \leq M$ and $E\left\|\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} \zeta_t e_{it}\right\|^4 \leq M$. For each i and for $j = 1, ..., p$, $E\left\|\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} F_{t-j}' e_{it}\right\|^4 \leq M$.

(f) $E\left\|\frac{1}{\sqrt{N_S T}}\sum_{t=p+1}^{T}\sum_{i\in S} \mathcal{F}_t e_{it}\right\|^2 \leq M$ and $E\left\|\frac{1}{\sqrt{N_S T}}\sum_{t=p+1}^{T}\sum_{i\in S} \zeta_t e_{it}\right\|^2 \leq M$.

Assumption 7:

(a) $N_S^{-1/2}\sum_{i\in S}\lambda_i = O_p(1)$.

(b) For each t and for $j = 0, 1, ..., p$, $E\left\|\frac{1}{\sqrt{N_S}}\sum_{i\in S} \lambda_i e_{i,t-j}e_{it}\right\|^2 \leq M$.

(c) For $j = 0, 1, ..., p$, $E\left\|\frac{1}{\sqrt{N_S N T}}\sum_{t=p+1}^{T}\sum_{k\neq i}\sum_{i\in S} \lambda_k e_{k,t-j}e_{it}\right\|^2 \leq M$ and $E\left\|\frac{1}{N_S\sqrt{T}}\sum_{t=p+1}^{T}\sum_{k\in S,k\neq i}\sum_{i\in S} \lambda_k e_{kt}e_{it}\right\|^2 \leq M$.

Assumption 8: under the null hypothesis,

(a) $\frac{1}{\sqrt{T-p}}\sum_{t=p+1}^{T}\zeta_t^F e_{it} \to_d N(0, \Omega_i)$ and $(T-p)^{-1}\sum_{t=p+1}^{T}\zeta_t^F\zeta_t^{F'}e_{it}^2 \to_p \Omega_i$ for $i \in S$ and some positive definite matrix Ω_i.

(b) $\frac{1}{\sqrt{(T-p)N_S}}\sum_{t=p+1}^{T}\sum_{i\in S}\zeta_t^F e_{it} \to_d N(0, \Omega)$ and $[(T-p)N_S]^{-1}\sum_{t=p+1}^{T}\sum_{i\in S}\zeta_t^F\zeta_t^{F'}e_{it}^2 \to_p \Omega$, where Ω is positive definite.

Assumptions 1 - 6 are either from or slight modifications of those in the factor model literature. Assumption 1(a) is almost the same as Assumption A of Bai (2003), and the only difference is that we also requires the positive definiteness of $\Sigma_{\mathcal{F}}$, which follows from Assumption A10 of Amengual and Watson (2007). Assumptions 1(b) and 2(c) provide regularity conditions on the structural shocks. Assumption 1(b) requires that the structural shocks are serially uncorrelated and have unit variance, which is standard in the dynamic factor model literature (see Forni et al., 2000). The first moment inequality in Assumption 1(c) holds if structural shocks are orthogonal, which follows from Forni et al. (2000), and the second moment inequality in Assumption 1(c) is not stringent because $E\zeta_t\mathcal{F}_t' = 0$ by Assumption 1(b). Assumption 2(a) is similar to Assumption B in Bai and Ng (2006), and assumption 2(c) is just an analog of 2(a). The condition that $N_S/N \to$ a constant $\in (0,1)$ ensures that both slow and fast variables are a non-negligible fraction of the full sample. Assumptions 3 and 5 follow from Assumption C and E of Bai (2003), while Assumption 4 is the same as Assumption D in Bai and Ng (2006). Assumption 6 is not stringent because all the sums in this assumption involve zero mean random variables. This assumption is similar to Assumption F of Bai (2003), and the main difference is that we introduce lags of F_t, e_{it} and ζ_t into the sums.

Assumption 7 imposes some further restrictions on the factor loadings. Assumption 7(a) requires that the sum of factor loadings in the slow variables is $O_p(\sqrt{N})$. This will hold if λ_i's are centered around zero, i.e. some variables have positive loadings and some have negative, so that the sum of λ_i's for $i \in S$ diverge at rate $\sqrt{N}$ by the central limit theorem. Assumptions 7(b) and 7(c) imply similar restriction to that of Assumption 7(a). A simple sufficient condition for Assumptions 7(b) and 7(c) is that e_{it}'s and λ_i's are independent with $E\lambda_i = 0$, but 7(b) and 7(c) can also hold for weakly correlated e_{it}'s and λ_i's. The role of Assumption 7 is to ensure that the difference between the $\left(\hat{\zeta}^{F\prime} \sum_{i\in S} \tilde{X}_i\right) / \sqrt{(T-p)N_S}$ and its infeasible analog $H_{\zeta^F}'\zeta^F \sum_{i\in S} e_i / \sqrt{(T-p)N_S}$ is $o_p(1)$, where H_{ζ^F} is a non-singular rotation matrix (see Appendix A3), so that the limit distribution of W is the same as that of W's infeasible analog. This assumption imposes more restrictions on λ_i than the conventional factor model literature, however, it seems to hold for the data set used in Section 5 and other commonly used data sets such as SW.[3]

[3]The largest elements of $|\sum_{i\in S} \hat{\lambda}_i / \sqrt{N_S}|$ are less than 3 for both the data set used in Section 5 and data set of SW, where

Assumptions 8(a) and 8(b) are simply central limit theorems. This assumption is applied to establish the asymptotic null distributions of w_i and W, which are summarized in the following theorem.

THEOREM 1: If $\sqrt{T}/N \to 0$ and the null hypothesis H_0: $q_F = \kappa_F$ holds, then
(i) under Assumptions 1 - 6 and 8, $w_i \to_d \ \chi^2_{q_F}$ for $i \in S$.
(ii) under Assumptions 1 - 8, $W \to_d \ \chi^2_{q_F}$.

Remarks:

(1) The detailed proof of Theorem 1 is provided in Appendix B. The basic idea of the proof is to show that $w_i - w_i^* = o_p(1)$ and $W - W^* = o_p(1)$, where $w_i^* \equiv e_i'\zeta^F(\sum_{t=p+1}^T \zeta_t^F \zeta_t^{F\prime} e_{it}^2)^{-1}\zeta^{F\prime}e_i$ and $W^* \equiv \sum_{i\in S} e_i'\zeta^F(\sum_{t=p+1}^T \sum_{i\in S} \zeta_t^F \zeta_t^{F\prime} e_{it}^2)^{-1}\zeta^{F\prime} \sum_{i\in S} e_i$, so w_i^* and W^* are simply the infeasible analogs of w_i and W. Note that w_i is just the Wald statistic that tests the coefficients in the regression of $\tilde{X}_i$ on $\hat{\zeta}^F$. Theorem 1 shows that one can apply the conventional Wald statistic as if $\tilde{X}_i$ and $\hat{\zeta}^F$ are observed. This is a new result in the literature. Bai (2003) shows that factors estimated from *observed* data can be treated as if they were observed as long as $\sqrt{T}/N \to 0$ and N, $T \to \infty$. We extends Bai's result in the sense that $\hat{\zeta}^F$ can be still treated as observed even if $\hat{\zeta}^F$ is estimated from $\tilde{X}$ which is also estimated. This property plays a central role in establishing the limit distributions of w_i and W, and it may be potentially useful in other inferential problems where both data and factors are not observed but estimated.

(2) Note that the variance matrices in w_i and W are estimated without imposing the null hypothesis. We can also construct LM-like test by imposing the null hypothesis. Define

$$lm_i = \tilde{X}_i'\hat{\zeta}^F\tilde{\Omega}_i^{-1}\hat{\zeta}^{F\prime}\tilde{X}_i/(T-p), \quad i \in S$$

$\hat{\lambda}_i$ is estimated from the regression of X_i on the estimated factors by principal components. Thus, Assumption 7(a) is likely to hold for these data sets. If one encounters a data set that has too large numbers in $|\sum_{i\in S}\hat{\lambda}_i/\sqrt{N_S}|$ to satisfy Assumption 7(a), one could adjust the signs of some X_i's such that Assumption 7 holds, because changing the signs of X_i's will not affect the estimated factor space asymptotically.

$$LM = \left(\sum_{i\in S} \tilde{X}_i' \hat{\zeta}^F\right) \tilde{\Omega}^{-1} \left(\hat{\zeta}^{F\prime} \sum_{i\in S} \tilde{X}_i\right) / (T-p) N_S$$

where $\tilde{\Omega}_i = (T-p)^{-1} \sum_{t=p+1}^{T} \hat{\zeta}_t^F \hat{\zeta}_t^{F\prime} \tilde{X}_{it}^2$ and $\tilde{\Omega} = (T-p)^{-1} N_S^{-1} \sum_{t=p+1}^{T} \sum_{i\in S} \hat{\zeta}_t^F \hat{\zeta}_t^{F\prime} \tilde{X}_{it}^2$. The following corollary shows the asymptotic distributions of lm_i and LM.

Corollary 1: If $\sqrt{T}/N \to 0$ and the null hypothesis H_0: $q_F = \kappa_F$ holds, then
(i) under Assumptions 1 - 6 and 8, $lm_i \to_d \chi^2_{q_F}$ for $i \in S$.
(ii) under Assumptions 1 - 8, $LM \to_d \chi^2_{q_F}$.

3.3 Asymptotics under the Alternative Hypothesis

Recall that the estimator $\hat{\zeta}^S$ is equal to $\sqrt{T-p}$ times the eigenvectors corresponding to the $\kappa_S = q - \kappa_F$ largest eigenvalues of the $(T-p) \times (T-p)$ matrix $\hat{X}^S \hat{X}^{S\prime}$. Under the alternative hypothesis $\kappa_S < q_S$, $\hat{\zeta}^S$ is $(T-p) \times \kappa_S$, so it is impossible for $\hat{\zeta}^S$ to consistently estimate the space spanned by ζ^S, which is $(T-p) \times q_S$. Let $\bar{\zeta}^S$ equal $\sqrt{T-p}$ times the eigenvectors corresponding to the q_S largest eigenvalues of $\hat{X}^S \hat{X}^{S\prime}$ and $\bar{V}_S$ be the diagonal matrix consisting of the first q^S largest eigenvalues of $(1/N^S T)\hat{X}^S \hat{X}^{S\prime}$ in decreasing order. Hence, we have

$$(1/N^S T)\hat{X}^S \hat{X}^{S\prime} \bar{\zeta}^S = \bar{\zeta}^S \bar{V}_S \tag{3.10}$$

Define $\bar{H}_{\zeta^S} = (\Gamma^{SS\prime}\Gamma^{SS}/N_S)(\zeta^{S\prime}\bar{\zeta}^S/T)\bar{V}_S^{-1}$, $\Gamma^S \equiv [\Gamma^{SS\prime} \vdots \Gamma^{FS\prime}]'$ and $\Gamma^F = [\Gamma^{SF\prime} \vdots \Gamma^{FF\prime}]'$, so we have $\Gamma = [\Gamma^S \vdots \Gamma^F]$ and

$$X = \mathcal{F}\Pi' + \zeta^S \Gamma^{S\prime} + \zeta^F \Gamma^{F\prime} + e \tag{3.11}$$

Lemma C1 in the Appendix C shows that $\bar{H}_{\zeta^S}$ is asymptotically non-singular, so Equation (3.11) can be rewritten as

$$\begin{aligned} X &= \mathcal{F}\Pi' + \zeta^S \bar{H}_{\zeta^S} \bar{H}_{\zeta^S}^{-1} \Gamma^{S\prime} + \zeta^F \Gamma^{F\prime} + e \\ &= \mathcal{F}\Pi' + \zeta^S \bar{H}_{\zeta^S} \Xi^{S\prime} + \zeta^F \Gamma^{F\prime} + e \end{aligned} \tag{3.12}$$

where $\Xi^{S'} \equiv \bar{H}_{\zeta^S}^{-1}\Gamma^{S'}$. Partition Ξ^S as $[\Xi^S_{1:\kappa_S}\vdots\Xi^S_{\kappa_S+1:q_S}]$, where $\Xi^S_{1:\kappa_S}$ is the first κ_S columns of Ξ^S and $\Xi^S_{\kappa_S+1:q_S}$ is the last $q_S-\kappa_S$ columns of Ξ^S. Let $\xi^S_{i,\kappa_S+1:q_S}$ denote the transpose of the i^{th} row of $\Xi^S_{\kappa_S+1:q_S}$. To show that our test statistics are consistent under the alternative hypothesis, we make the following assumption.

Assumption 9:

There exist constants $0 < \alpha \leq 0.5$ and $C > 0$, such that

$$\text{Prob}\left(\left\|\frac{T^{\alpha/2}}{\sqrt{N_S}}\sum_{i\in S}\xi^S_{i,\kappa_S+1:q_S}\right\| > C\right) \to 1$$

as N and $T \to \infty$.

Assumption 9 ensures that the joint statistic W diverges to infinity under the alternative hypothesis. One simple sufficient condition for Assumption 9 is that $\sum_{i\in S}\xi^S_{i,\kappa_S+1:q_S}/N_S$ converges to a non-zero constant. For the case where $\sum_{i\in S}\xi^S_{i,\kappa_S+1:q_S}/N_S \to_p 0$, this assumption can still hold under other sufficient conditions. For example, if the central limit theorem holds such that $\sum_{i\in S}\xi^S_{i,\kappa_S+1:q_S}/\sqrt{N_S}$ is normally distributed at the limit, then $T^{\alpha/2}\sum_{i\in S}\xi^S_{i,\kappa_S+1:q_S}/\sqrt{N_S}$ will diverge for any $\alpha > 0$ as long as N and $T \to \infty$ at the same rate.

Theorem 2: Suppose that $\sqrt{T}/N \to 0$ and the alternative hypothesis H_1: $q_F < \kappa_F$ holds.
(i) under Assumptions 1 - 6, if $\xi^S_{i,\kappa_S+1:q_S} \neq 0$ for $i \in S$, then $w_i \to \infty$ as N and $T \to \infty$.
(ii) under Assumptions 1 - 6 and 9, $W \to \infty$ if $\sqrt{N}/T^{1-\alpha/2} \to 0$ as N and $T \to \infty$.

Remarks:

(1) The divergence rate of w_i is T, whereas the divergence rate of W depends on the asymptotic behavior of $\sum_{i\in S}\xi^S_{i,\kappa_S+1:q_S}/\sqrt{N_S}$. Assumption 9 implies that the divergence rate of W is no slower than $T^{1-\alpha}$. For the case where $\sum_{i\in S}\xi^S_{i,\kappa_S+1:q_S}/N_S$ converges to non-zero constant, the divergence rate of W is NT, which is very fast compared to conventional test statistic.

When $\sum_{i\in S}\xi^S_{i,\kappa_S+1:q_S}/N_S$ converges to zero, the divergence rate can be slower than T. For example, if $\sum_{i\in S}\xi^S_{i,\kappa_S+1:q_S}/\sqrt{N_S}$ converges in distribution to a normal random variable and N and T diverge at the same rate, then $T^{\alpha/2}\sum_{i\in S}\xi^S_{i,\kappa_S+1:q_S}/\sqrt{N_S}$ will diverge for any positive α. In such a case, the divergence rate of W is less than but arbitrarily close to T.

(2) The condition that $\sqrt{N}/T^{1-\alpha/2}\to 0$ is slightly stronger than $\sqrt{N}/T\to 0$ used by Bai (2003), but it still allows a wide range of relative rate between N and T. For example, both $\sqrt{N}/T^{1-\alpha/2}\to 0$ and $\sqrt{T}/N\to 0$ will be satisfied for $\alpha<0.5$ if N and T are proportional to each other, which seems to be a reasonable assumption for typical macroeconomic data sets in the DFM literature.

(3) The following corollary shows that lm_i and LM are also consistent tests against the alternative.

Corollary 2: Suppose that $\sqrt{T}/N\to 0$ and the alternative hypothesis H_0: $q_F<\kappa_F$ holds.
(i) under Assumptions 1 - 6, if $\xi^S_{i,\kappa_S+1:q_S}\neq 0$ for $i\in S$, then $lm_i\to\infty$ as N and $T\to\infty$.
(ii) under Assumptions 1 - 6 and 9, $LM\to\infty$ if $\sqrt{N}/T^{1-\alpha/2}\to 0$ as N and $T\to\infty$.

4 Monte Carlo Simulations

In the Monte Carlo experiments we investigate the finite sample properties of our statistics and some other statistics. It is noteworthy that there are no theoretically verified alternatives to both our individual and joint statistics to test the contemporaneous timing restrictions in the FAVAR framework of SW. However, it would be interesting to see how these unverified alternatives perform relative to our statistics. Hence, we also explore the size-power properties of the following statistics: the individual statistic of SW, Bonferroni and pooled statistics based on our individual statistics, Onatski's (2009) statistics for number of dynamic and static factors. The individual statistic of SW, denoted as sw_i, is computed as follows: let $\hat{\Sigma}(r\times r)$ denote the sample covariance of the residual from the VAR of $\hat{F}_t$, estimate the $q-$dimensional innovations $\hat{\eta}_t$ for the dynamic factors by using the spectral decomposition of $\hat{\Sigma}$, estimate the $q-\kappa_F$ dimensional slow shocks $\hat{\eta}_t^S$ by conducting a reduced rank regression of $\hat{X}_t^S = M_{\hat{\mathcal{F}}} X_t^S$ on $\hat{\eta}_t$, estimate the κ_F-dimensional fast shocks $\hat{\eta}_t^F$ that are orthogonal to $\hat{\eta}_t^S$, regress $\hat{X}_{it}^S$ on $\hat{\eta}_t^S$ and $\hat{\eta}_t^F$, and test whether the coefficients on $\hat{\eta}_t^F$ are zeros or not by the conventional Wald statistic. Furthermore, we also compute the Bonferroni and pooled statistics based on our individual test statistic. The Bonferroni statistic, denoted as B_w, is simply the maximum of w_i for $i \in S$, and the 5% critical value is equal to $\mathcal{X}^{-1}(1-5\%/N_S)$, where $\mathcal{X}$ is the chi-square CDF with the degree of freedom κ_F. The pooled statistic, denoted as P_w, is set equal to $\left(\sum_{i\in S} w_i - \kappa_F N_S\right)/\sqrt{2N_S\kappa_F}$ and the 5% critical values are ± 1.96. The critical values of the Bonferroni and pooled statistics are based on the sequential limit argument that first $w_i \to_d \ \chi^2_{\kappa_F}$ and then $N_S \to \infty$. The problem of the sequential limit is that the convergence of w_i relies on N, N_S and $T \to \infty$ simultaneously, so $w_i \to_d \ \chi^2_{\kappa_F}$ and $N_S \to \infty$ cannot be separated into two sequential steps. The sequential limit and simultaneous limit are not always equivalent (see Phillips and Moon, 1999). Hence, we would like to include these two statistics because they allow us to see what would be the size distortion based on the incorrect limit distributions derived from the sequential limit. For Onatski's (2009) statistics, we use R_{dyn} and R_{stat} to denote the statistics for the numbers of dynamic and static factors, respectively. Note that the number of dynamic factors in $X^S - \mathcal{F}\Pi^{S'}$ is the same as the number of static factors, so both R_{dyn} and R_{stat} test the same null hypothesis that $q_S = q - \kappa_F$ against the alternative that $q-\kappa_F \leq q_S \leq k_{max}$. To implement Onatski's tests, $X^S - \mathcal{F}\Pi^{S'}$ is replaced by its feasible analog $M_{\hat{\mathcal{F}}} X^S$ and k_{max} is set equal to 8.

In all Monte Carlo experiments, r and q are selected by IC_{p1} of Bai and Ng (2002), and the number of lags of F_t is assumed to be known[4]. The number of replications is 5000 in each data generating process (DGP).

The data is produced by the model: $X_{it} = F_t\lambda_i' + \omega e_{it}$. Under the null hypothesis, we use

[4]We also use BIC to select the number of lags of F_t, and the results are robust and not reported

the following DGPs:

N1: $f_{\ell t} = \rho_f f_{\ell,t-1} + \eta_{\ell t}$ $\eta_{\ell t} \sim i.i.d.\ N(0,\ 1-\rho_f^2)$ for $\ell = 1, ..., q$. $F_t = [f_{1t}, ..., f_{qt}, f_{1,t-1}, ..., f_{r-q,t-1}]'$.
$e_{it} = \sigma_i \mu_{it}$, $\sigma_i \sim i.i.d.\ U(0.5,\ 1.5)$, $\mu_{it} = \rho_\mu \mu_{it-1} + \epsilon_{it}$, $\epsilon_{it} \sim i.i.d.\ N(0,\ 1-\rho_\mu)$.
$\lambda_i = [\lambda_{i1}, ..., \lambda_{ir}]'$. For $i = 1, ..., N_S$, set $[\lambda_{i1}, ..., \lambda_{iq_F}] = 0_{1\times q_F}$ and draw $\lambda_{ij} \sim i.i.d.\ N(0,1)$ for $j = q_F + 1, ..., r$ with $q_F \leq q$. For $i = N_S + 1, ..., N$, $\lambda_{ij} \sim i.i.d.\ N(0,1)$ for $j = 1, ..., r$.
Set $\omega = \sqrt{12(r - q_F)/13}$ for $i = 1, ..., N_S$, and set $\omega = \sqrt{12r/13}$ for $i = N_S + 1, ..., N$.

N2: F_t and λ_i are generated in the same way as in DGP N1. Set $\omega = \sqrt{12(r - q_F)}$ for $i = 1, ..., N_S$, and set $\omega = \sqrt{12r}$ for $i = N_S + 1, ..., N$. $e_{it} \sim i.i.d.\ U(-0.5,\ 0.5)$.

In DGPs N1 and N2, the factor loadings are generated in a way such that factors $f_{1t}, ..., f_{q_F t}$ do not affect X_{it} for $i = 1, ..., N_S$. Hence, the subscripts $i = 1, ..., N_S$ stand for slow variables, while the subscripts $i = N_S + 1, ..., N$ stand for the fast variables. ω is chosen such that the factors explain 50% variation in the data. We set $N_S = 0.5N$, $\rho_f = \rho_\mu = 0.5$, and $(r,\ q,\ q_F) \in \{(5,\ 3,\ 1),\ (5,\ 4,\ 3)\}$.

The results under DGP N1 are summarized in Table 1. The upper and lower panels report the results for $(r,\ q,\ q_F) = (5, 3, 1)$ and $(r,\ q,\ q_F) = (5, 4, 3)$, respectively. The first two columns of Table 1 are simply the numbers of observations in cross section and time dimensions. The numbers in columns 3 – 5 are computed in the following way: for each simulated sample, we compute the ratio between the number of rejections by the individual tests at the 5% level and the number of slow variables N_S, and then we take the mean of these ratios from 5000 simulated samples. Hence, columns 3 – 5 show the finite-sample size properties of the individual statistics. It is noteworthy that the size of sw_i is about 60% when $N_S = 250 > T = 200$. The reason is that the reduced rank regression will invert a $N_S \times N_S$ sample covariance matrix, which is singular when $T < N_S$. Compared to sw_i, w_i and lm_i do not have the large size distortion problem when $N_S > T$, and their size is approaching the nominal level as N and T become larger.

Columns 6 – 11 of Table 1 report the effective size of different joint tests. First, Bonferroni and Pooled statistics always reject the null much more often than the nominal level. This confirms that the critical values based on the sequential limit argument are invalid and that the asymptotic distributions of B_w and P_w are likely to be non-standard as N and $T \to \infty$ simultaneously. Second, Onatski's (2009) tests seem to work well under DGP1 except that R_{stat} rejects slightly too often when $N = 200$, $T = 500$, $r = 5$, $q = 4$, and $q_F = 3$. Finally, the size of our statistics is slightly higher than the normal level in small samples, but the size distortion become smaller as the sample size increases.

Table 2 summarizes the results under DGP N2, in which the idiosyncratic shocks are drawn from the uniform distribution $U(-0.5,\ 0.5)$. The size properties of B_w, P_w, W, and LM are almost the same as those under DGP N1. However, both R_{dyn} and R_{stat} tend to over-reject especially when $r = 5$, $q = 4$, and $q_F = 3$. The reasons are twofold: (1) Onatski (2009) derives the asymptotic distribution of R_{stat} based on the Gaussianity of the idiosyncratic shocks. Under DGP N2, the Gaussianity assumption is violated, and the size distortion of R_{stat} is very large when $r = 5$, $q = 4$, and $q_F = 3$. This implies that the asymptotic distribution of R_{stat} crucially relies on the distribution of the idiosyncratic shocks. Hence, the inference using R_{stat} can be misleading if Gaussianity does not hold. (2) The asymptotic distributions of R_{dyn} and R_{stat} are developed when the data are observed. Recall that we use the feasible $M_{\hat{\mathcal{F}}}X^S$ instead of its infeasible counterpart $X^S - \mathcal{F}\Pi^{S'}$ to compute Onatski's statistics. The estimation errors in $M_{\hat{\mathcal{F}}}X^S$ may change the asymptotic distributions of R_{dyn} and R_{stat}. This explains why R_{dyn}, which does not relies on the Gaussianity assumption, also has a non-trivial size distortion. Hence, implementing Onatski's statistics to test the overidentifying restrictions in FAVAR may lead to substantial over rejection of the null hypothesis.

Under the alternative hypothesis, we use the following DGPs.

A1: F_t and e_{it} are generated in the same way as in DGP N1.
$\lambda_i = [\lambda_{i1}, ..., \lambda_{ir}]'$. For $i = 1, ..., (1-a)N_S$, set $[\lambda_{i1}, ..., \lambda_{i\kappa_F}] = 0_{1\times\kappa_F}$ and draw $\lambda_{ij} \sim i.i.d.\ N(0,1)$ for $j = \kappa_F+1, ..., r$ with $\kappa_F \leq q$. For $i = (1-a)N_S+1, ..., N$, $\lambda_{ij} \sim i.i.d.\ N(0,1)$ for $j = 1, ..., r$. $a \in \{0.2, 0.4, 0.6, 0.8, 1\}$. $\omega = \sqrt{12r/13}$.

A2: F_t and e_{it} are generated in the same way as in DGP N1.

$\lambda_i = [\lambda_{i1}, ..., \lambda_{ir}]'$. For $i = 1, ..., N_S$, draw $\lambda_{ij} \sim i.i.d.\ N(0, b^2)$ for $j = 1, ..., \kappa_F$ and draw $\lambda_{ij} \sim i.i.d.\ N(0, 1)$ for $j = \kappa_F + 1, ..., r$ with $\kappa_F \leq q$. For $i = N_S + 1, ..., N$, $\lambda_{ij} \sim i.i.d.\ N(0, 1)$ for $j = 1, ..., r$. $b \in \{0.2, 0.4, 0.6, 0.8, 1\}$. $\omega = \sqrt{12r/13}$.

In both DGPs A1 and A2, we set $N_S = 0.5N$, $\rho_f = \rho_\mu = 0.5$, and $(r,\ q,\ \kappa_F) \in \{(5,\ 3,\ 1), (5,\ 4,\ 3)\}$. Note that the true number of fast shock, q_F, is zero under both DGPs A1 and A2.

In DGP A1, factors $f_{1t}, ..., f_{\kappa_F t}$ do not affect X_{it} for $i = 1, ..., (1 - a)N_S$, so α controls the fraction of slow variables that have non-zero factor loadings on $f_{1t}, ..., f_{\kappa_F t}$. Table 3 shows how the power[5] changes as α increases. Columns $4 - 9$ show the results when $r = 5$, $q = 3$, and $\kappa_F = 1$, and columns $10 - 15$ show the results when $r = 5$, $q = 4$, and $\kappa_F = 3$. First, it is clear that the averaged rejection rates of w_i and lm_i increase as N and T increase. Second, the Onatski's statistics do not have much power for small a, N and T. For example, when $r = 5$, $q = 3$, $\kappa_F = 1$, $a = 0.4$ and $N = T = 200$, the power of R_{dyn} is only 6% and the power of R_{stat} is 16%, whereas the power of W and LM is 59% and 55.5%, respectively. It is remarkable that the power against small a is a desired property. In practice, it is unlikely that all the slow variables are misclassified, but it is likely that economists are not sure whether some variables should be classified as slow or fast. In the latter case, we want a test that is powerful to detect the violation of identifying restrictions by only aN_S many variables, especially when a is small. Furthermore, R_{stat} seems to be more powerful than our joint statistics when a is close to one and N and T are large, but this power is suspicious due to potential size distortion of R_{stat} shown in Table 2. Finally, the power of W and LM increase as κ_F increases, which is expected because the tests should be more powerful when there are more wrong restrictions.

In DGP A2, we investigate the relationship between the power and the variance of the factor loadings in slow groups. b controls the extent to which the null hypothesis is violated. Larger b allows the loadings in slow variables to deviate further away from zero. When $b = 1$, DGP A2 is the same as DGP A1 with $\alpha = 1$. Table 4 reports how the power changes as b

[5]Since the Bonferroni and Pooled statistics always have large size distortion and sw_i has large size distortion when $N_S > T$, we do not report their power.

increases. The pattern in Table 4 is similar to that in Table 3. Both the average rejection rates of w_i and lm_i and the power of W and LM increase as N, T and κ_F increase. R_{dyn} is always less powerful than our joint tests, except when $r = 5$, $q = 3$, $\kappa_F = 1$, $b = 1$, and $N = T = 500$. R_{stat} is less powerful than W and LM for small b but more powerful as b, N, and T become larger.

In sum, SW's individual statistic sw_i has large size distortion when $N_S > T$, but our w_i and lm_i have good size and power in finite samples. The Bonferroni and pooled statistics have large size distortion because the critical values based on the sequential limit are invalid in the factor models where N and T go to infinity simultaneously. Onatski's R_{dyn} and R_{Stat} could have substantial size distortions due to violating the Gaussianity assumption and neglecting the estimation errors. They are also less powerful when only a fraction of slow variables violate the identifying restrictions in their factor loadings or when the factor loadings of all slow variables violate the identifying restrictions but their deviations of from zero are not large. Our joint test statistics have reasonably good size and power in finite samples: their size distortion decreases and power increases as N and $T \to \infty$.

5 Empirical Results

In the section, we implement our statistics to a data set, and investigate the impulse responses of major macroeconomic variables to the monetary policy shock, which is identified based on different numbers of fast shocks. The data set is an updated version[6] of the one used by SW. It consists of monthly observations of 125 U.S. macroeconomic time series from 1960:1 through 2007:12, and the number of slow variables is 64. The series are transformed by taking logarithms and/or differencing so that the transformed series are approximately stationary. The transformation mainly follows from SW. For example, all the interest rates variables are transformed by taking the first differences. However, we transform all the prices by taking the first rather than the second differences of logarithms, which follows from BBE and Forni

[6]Some of the variables used by SW are dropped due to the missing values.

and Gambetti (2010). The full list of variables along with the corresponding transformations is given in Appendix D.

5.1 Testing the Overidentifying Restrictions

We first implement Bai and Ng's (2002) IC_{p1} and IC_{p2} to determine the number of static factors in the transformed series[7]. The upper bound of r is set equal to 10, which is also selected as the estimate for r by both IC_{p1} and IC_{p2}. This is different from the view in the forecasting literature (for example, Stock and Watson (2002)), which finds that 1 or 2 factors are needed to to explain the variation of US macroeconomic variables. However, it is common to use large estimates for the number of static factors in the FAVAR literature. For example, SW set $\hat{r} = 9$, whereas Forni and Gambetti (2010) set $\hat{r} = 10$. The reason is that using large $\hat{r}$ tend to better estimate the space spanned by the structural shocks, which is crucial in the structural VAR analysis. Thus, we adopt 10 static factors in our empirical application.

Before implementing our hypothesis testing procedure, it is of interest to see the performance of information criteria. We apply the estimators of Amengual and Watson (2007) and Bai and Ng (2007) to determine the values of q and q_S. Amengual and Watson's estimator is computed in the following way: regress the transformed data on the lags of $\hat{F}_t$, store the residuals, and then use Bai and Ng's (2002) IC_{p1} and IC_{p2} to determine the number static factors in the residuals. Bai and Ng (2007) propose two estimators, q_3 and q_4, for the number of dynamic factors. These two estimators are constructed as follows: fit a VAR for $\hat{F}_t$, compute the sample variance of the residuals, denoted as $\hat{\Sigma}$, truncate small eigenvalues of $\hat{\Sigma}$ to zeros by some thresholds, and then choose the number of non-zero eigenvalues of $\hat{\Sigma}$ as the estimate for q. For both Amengual and Watson's and Bai and Ng's estimators, the number of lags of $\hat{F}_t$ is set equal to 2, which is selected by BIC. We also use 4 lags for a robustness check, and the results do not change. The estimate for q is equal to 6 by IC_{p1}, 5 by IC_{p2}, 6

[7]As usual in the literature, all the transformed series are normalized to have zero mean and unit variance.

by q_3, and 5 by q_4. The estimate for q_S is equal to 7 by IC_{p1}.[8] It is remarkable that these estimates cause a problem if one wants to use them in the FAVAR setup. The estimate of q_S is greater than that of q, which contradicts the fact that $q_S \leq q$. This problem might be caused by the limited number of observations in the cross section dimension (recall that N_S is only 64), so that IC_{p1}, IC_{p2}, q_3 and q_4 lead to finite-sample biases in the estimates, although they are consistent as N and $T \to \infty$. In a word, these results show that information criteria can give biased estimates for q and q_S due to the limited sample size, so that one cannot setup the restrictions to identify the monetary policy shock in the FAVAR model. Hence, our hypothesis testing procedure is the only way to evaluate the specification in such scenarios.

We next implement our statistics to test the null hypothesis H_0: $q_F = \kappa_F$ against the alternative hypothesis $H_1 : q_F < \kappa_F$. Since $q = q_S + q_F$, these hypotheses are equivalent to the null hypothesis that $q_S = q - \kappa_F$ against the alternative hypothesis that $q_S > q - \kappa_F$, which can be tested by Onatski's (2009) statistics. Table 5 reports the p-values of our and Onatski's test statistics for different values of q and κ_F. The numbers outside the parentheses are the p-values when the number of lags of $\hat{F}_t$ is equal to 2 and the numbers inside the parentheses are the p-values when the number of lags of $\hat{F}_t$ is equal to 4. In general, the p-values of W and LM are small for $\kappa_F = 2$ and 3, but relatively large for $\kappa_F = 1$, so our statistics suggest that the number of fast shock is equal to 1. Since the monetary policy shock is assumed not to affect slow variables contemporaneously, the only fast shock in this data set is identified as the monetary policy shock.

Compared to our statistics, Onatski's (2009) statistics have large p-values for almost all values of q, κ_F and the number of lags. The null hypotheses H_0: $q_F = \kappa_F$ are not rejected at the 5% level for $\kappa_F = 1,\ 2,\ 3$, except that R_{dyn} rejects the setup where $q = 4$, $\kappa_F = 3$, and the number of lags of $\hat{F}_t$ is 2. These results indicate that one can impose $\kappa_F = 1,\ 2$ or 3 in the FAVAR to identify the fast shocks. Recall that the simulation results show that Onatski's

[8] Note that q_S is not necessarily equal to the number of dynamic factors in the slow group, because slow shocks are defined as the dynamic factors that can affect slow variables contemporaneously, but the dynamic factors in the slow variables could be either fast or slow. Hence, q_3 and q_4 are not valid to estimate q_S.

tests are not powerful especially when the identifying restrictions are mildly violated and the sample size is relatively small. Thus, the identifying restriction that $\kappa_F = 2$ or 3 could be wrong, but the R_{dyn} and R_{stat} are not able to detect the misspecification.

5.2 Impulses Responses to the Monetary Policy Shock

In this subsection, we compare the impulse response functions based on different values of κ_F to check which value of κ_F generates more plausible results. We consider the following FAVAR model:

$$\begin{bmatrix} F_t \\ X_t \end{bmatrix} = \begin{bmatrix} \Phi(L) & 0 \\ \Upsilon(L) & D(L) \end{bmatrix} \begin{bmatrix} F_{t-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{F_t} \\ \varepsilon_{X_t} \end{bmatrix} \tag{5.1}$$

where

$$\begin{bmatrix} \varepsilon_{F_t} \\ \varepsilon_{X_t} \end{bmatrix} = \begin{bmatrix} I \\ \Lambda \end{bmatrix} G\eta_t + \begin{bmatrix} 0 \\ v_t \end{bmatrix}$$

The setup in (5.1) is almost the same as that of (2.7), except that $\Lambda\Phi(L) - D(L)\Lambda$ is replaced by an unrestricted lag polynomial. We use 2 lags of F_t and 6 lags of X_t in the regression. The results are robust to using 4 lags of F_t or 8 lags of X_t. Based on the results by information criteria, we set $q = 6$ as our benchmark, and the results are robust to $q = 5,\ 7,\ 8$.

The Impulse Responses are computed as follows:

(1) Given $\hat{F}_t$ estimated from X using principal components, get the residuals $\hat{\varepsilon}_{F_t}$ and $\hat{\varepsilon}_{X_t}$ from the regression (5.1).

(2) Compute the sample variance of $\hat{\varepsilon}_{F_t}$, denoted as $\hat{\Sigma}$. Define $\hat{\eta}_t = \Delta' \hat{\varepsilon}_{F_t}$, where Δ is a $r \times q$ matrix consisting of eigenvectors that correspond to the q largest eigenvalues of $\hat{\Sigma}$. Impose the identifying restriction that $q_F = \kappa_F$ and run a reduced rank regression of $\hat{\varepsilon}^S_{X_t}$ on $\hat{\eta}_t$ to get an estimate of the $q - \kappa_F$ dimensional slow shock $\hat{\eta}^S_t$, where $\hat{\varepsilon}^S_{X_t}$ is the residuals of slow variables from (5.1).

(3) Estimate the monetary policy shock by $\hat{\eta}^R_t = \text{Proj}(\hat{\varepsilon}^R_{X_t} | \hat{\eta}_t) - \text{Proj}(\hat{\varepsilon}^R_{X_t} | \hat{\eta}^S_t)$, where the

projections are implemented by OLS. If $\kappa_F \geq 2$, then regress $\hat{\eta}$ on $[\hat{\eta}^S \vdots \hat{\eta}^R]$ and get the residuals $\hat{z}$. Estimate the additional fast factors $\hat{\eta}^F$ as the eigenvectors corresponding to $\kappa_F - 1$ largest eigenvalues of zz'. Normalize $[\hat{\eta}^S \vdots \hat{\eta}^R \vdots \hat{\eta}^F]$ so that they have unity variances. Denote the normalized estimates as $\hat{\zeta}^S$, $\hat{\zeta}^R$, and $\hat{\zeta}^F$.

(4) Regress $[\hat{\varepsilon}'_{F_t},\ \hat{\varepsilon}'_{X_t}]'$ on $\hat{\zeta}_t^S$ and $\hat{\zeta}_t^R$ for $\kappa_F = 1$ and regress $[\hat{\varepsilon}'_{F_t},\ \hat{\varepsilon}'_{X_t}]'$ on $\hat{\zeta}_t^S$, $\hat{\zeta}_t^R$, and $\hat{\zeta}_t^F$ for $\kappa_F \geq 2$ and store the estimates for the coefficients. Compute impulse responses based on these estimated coefficients and the companion form of (5.1).

Confidence bands of the impulse responses are obtained by a bootstrap technique. To create bootstrapped samples of the $T \times N$ data X^b, we use a slight modification of the procedure proposed by Yamamoto (2011). Given the principal component estimator $\hat{F}$, estimate $\hat{\Lambda} = X'\hat{F}/T$, $\hat{e} = X - \hat{F}\hat{\Lambda}'$, $\hat{\Phi}(L)$ and $\hat{\varepsilon}_{F_t}$, where $\hat{\Phi}(L)$ and $\hat{\varepsilon}_{F_t}$ are the estimates of $\Phi(L)$ and ε_{F_t} in (5.1). Re-sample the demeaned residuals $\hat{\varepsilon}_{F_t} - \bar{\hat{\varepsilon}}_{F_t}$ with replacement and label it as $\varepsilon^b_{F_t}$. Generate the bootstrapped factor F_t^b by $F_t^b = \hat{\Phi}(L)F^b_{t-1} + \varepsilon^b_{F_t}$. Since Yamamoto's (2011) bootstrap algorithm does not allow $\hat{e}_t$ to be serially correlated, we modify his way of generating e^b using the re-sampling procedure proposed by Ludvigson and Ng (2009). By equation (2.5), for each i, we estimate $\hat{\delta}_i(L)$ and residuals $\hat{v}_{it}$, which are then re-centered. Re-sample the $N \times 1$ vector $\hat{v}_t$ to get v_t^b so that the cross-section correlation structure is preserved, and then generate e_t^b using v_t^b and $\hat{\delta}_i(L)$. Next, the bootstrapped data X^b is constructed as $X^b = F^b\hat{\Lambda}' + e^b$ for $b = 1, ..., 2000$. Given the re-sampled data X^b, the impulse responses are computed following steps (1) – (4).

Figure 1 demonstrates the impulse responses of major macroeconomic variables to a unity variance contractionary monetary policy shock. The solid curves are computed using the identifying restriction that $\kappa_F = 1$, whereas the dashed curves are computed using the identifying restriction that $\kappa_F = 2$.[9] We report the bootstrap 68% and 95% confidence bands for the impulse responses for $\kappa_F = 1$. Although the impulse responses based on $\kappa_F = 2$

[9]The results based on $\kappa_F = 3$ are similar to or even worse than those based on $\kappa_F = 2$, so they are not reported in the paper.

are within the 95% confidence bands of the impulse responses based on $\kappa_F = 1$ for some variables, the improvement from using $\kappa_F = 1$ is still considerable. In general, the impulse responses based on $\kappa_F = 1$ are more economically plausible than those based on $\kappa_F = 2$. For example, when $\kappa_F = 2$, the output has a peak equal to 0.17 percent, and response becomes negative 10 months after the contractionary monetary policy shock. This result contradicts the conventional view that a monetary tightening would be expected to cause a decline in real output over time rather than an increase. When $\kappa_F = 1$, i.e., the monetary policy shock is the only fast shock in the data set, both the magnitude and duration of the positive response decrease: the output has a peak equal to 0.05 percent, and response becomes negative 4 months after the contractionary monetary policy shock. Hence, $\kappa_F = 1$ generates a result much closer to the prediction by economic theory than $\kappa_F = 2$.

It is remarkable that the positive impulse response of real output after a monetary tightening has already been noticed in the empirical macroeconomics literature. For example, Uhlig (2005) finds a positive response of real output after a contractionary monetary policy shock by imposing sign restrictions on the impulse responses of prices, non-borrowed reserves and the federal funds rate but no restrictions on the impulse response of real output. Similar results are derived by Inoue and Kilian (2011) based on a new inferential technique on the impulse response functions. The reason that Uhlig does not impose restrictions on the impulse response of real output is that he wished to be agnostic about it. In this sense, our estimation procedure has some similarity to Uhlig's because all the impulse response functions are estimated by OLS without imposing restrictions that are used in the identification of structural shocks. Although the problem found by Uhlig is not completed solved in our FAVAR framework, the results are more consistent with the traditional economic theory when we impose identifying restrictions that are not rejected by our joint tests.

Another improvement from imposing $\kappa_F = 1$ instead of $\kappa_F = 2$ is that the price puzzle (Sims, 1992) is considerably reduced in the former setup. When we impose $\kappa_F = 2$ in the FAVAR, the CPI has a persistent positive response lasting for about 2 years. When we impose

$\kappa_F = 1$, the price puzzle in CPI almost disappears. Figure 2 further investigates the impulse responses of different price indexes. Compared to $\kappa_F = 2$ which leads to persistent positive responses to a contractionary monetary policy shock, $\kappa_F = 1$ either substantially reduces or completely solves the price puzzle in all price indexes. This further confirms that the monetary policy shock is likely to be the only fast shock in this data set and that Onatski's tests fail to reject $\kappa_F = 2$ or 3 due to their lack of power.

Moreover, the responses of other variables based on $\kappa_F = 1$ are generally more consistent with economic theory in terms of signs and magnitudes. For example, the monetary tightening leads to an immediate reduction in the real consumption and employment when $\kappa_F = 1$, but it generates positive responses when $\kappa_F = 2$. Also, the response of consumer expectation has the expected sign for $\kappa_F = 1$ but entirely wrong sign for $\kappa_F = 2$. For some variables, such as capacity utilization, unemployment, orders, inventories and commodity price, the impulse responses after the monetary policy shock have the "unexpected" signs for both $\kappa_F = 1$ and $\kappa_F = 2$. However, the results based on $\kappa_F = 1$ are still much better than those based on $\kappa_F = 2$, because the unexpected parts of the responses are much smaller in magnitude and much shorter in duration for $\kappa_F = 1$.

6 Conclusions

In this paper, we develop test statistics for the overidentifying restrictions in FAVAR models. Unlike the conventional structural VAR analysis, the FAVAR can involve a large number of identifying restrictions but a few structural shocks, so the system is highly overidentified. We focus on testing the joint null hypothesis that all the identifying restrictions are satisfied. Since the number of restrictions goes to infinity as the sample size grows, conventional tests are not applicable. Our new joint statistics solve this problem by combining the individual statistics across the cross section dimension, so that the infinite-dimensional problem reduces to a finite-dimensional one. Under some regularity conditions, we find the asymptotic distribution of our statistic under the null hypothesis and prove that it is consistent against the alternative that a substantial amount of identifying restrictions are violated. In the Monte Carlo experiments, we find that our statistics have relatively good size and power in finite samples. Also, the simulation results confirm that other alternative test statistics do not perform well, so our statistics are the only valid candidates to test the overidentifying restrictions in FAVAR models.

In the empirical application, we estimate an FAVAR model using an updated version of Stock and Watson's (2005) data set. We follow the setup of Stock and Watson (2010) and the identification scheme of Stock and Watson (2005). The tests reject the null hypotheses that the number of fast shocks is two or three, but they do not reject the null that there is only one fast shock, which is the monetary policy shock by definition. This result provides some evidence to support the Bernanke, Boivin, and Eliasz's (2005) identification assumption that factors other than the monetary policy shock are slow. Also, Our results show that information criteria could give contradictory estimates for number of shocks in small sample, so the new testing procedure is the only way to evaluate the specification when information criteria fail. To evaluate the performance of our tests in this data set, we compare the impulse responses based on different values of κ_F, where κ_F denotes the potential number of fast shocks. In general, the impulse responses of major macroeconomic variables are more economically plausible based on $\kappa_F = 1$ than those based on $\kappa_F = 2$. Moreover, compared to $\kappa_F = 2$ which leads to persistent positive responses to a contractionary monetary policy shock, $\kappa_F = 1$ either substantially reduces or completely solves the price puzzle in all price indexes.

Finally, the test statistics proposed by this paper mainly focus the contemporaneous timing restrictions, which are widely used in the FAVAR models. Additionally, other identification schemes, such as long run restrictions and sign restrictions, can be applied in FAVAR. We leave for future research to generalize our statistics to cover these identification schemes.

Appendix

A Preliminary Lemmas under the Null Hypothesis

To conserve space, we only state the lemmas in the file, and the detailed proofs of lemmas in Appendix A are available in the supplement appendix.

A.1 Lemmas on $\hat{F}$ and $\hat{\mathcal{F}}$

Let that V_X is the $r \times r$ diagonal matrix of the first r largest eigenvalues of $(1/TN)XX'$ in decreasing order. Recall that the estimated factor matrix $\hat{F}$ is $\sqrt{T}$ times eigenvectors corresponding to the r largest eigenvalues of XX'. Therefore, we have $(1/NT)XX'\hat{F}V_X^{-1} = \hat{F}$, which can be expanded as:

$$\frac{1}{NT}(F\Lambda'\Lambda F' + F\Lambda' e' + e\Lambda F' + ee')\hat{F}V_X^{-1} = \hat{F} \tag{A.1}$$

Define $H_F = (\Lambda'\Lambda/N)(F'\hat{F}/T)V_X^{-1}$, so we have:

$$\hat{F} - FH = \frac{1}{NT}\left(F\Lambda' e'\hat{F} + e\Lambda F'\hat{F} + ee'\hat{F}\right)V_X^{-1} \tag{A.2}$$

$$\hat{F}_t - H_F' F_t = V_X^{-1}\left(\frac{1}{T}\sum_{s=1}^{T}\hat{F}_s\gamma_N(s,t) + \frac{1}{T}\sum_{s=1}^{T}\hat{F}_s\left[\frac{e_s' e_t}{N} - \gamma_N(s,t)\right] + \frac{1}{NT}\sum_{s=1}^{T}\hat{F}_s F_s'\Lambda' e_t + \frac{1}{NT}\sum_{s=1}^{T}\hat{F}_s e_s'\Lambda F_t\right) \tag{A.3}$$

Note that $T^{-1}\sum_{t=1}^{T}\|\hat{F}_t\|^2 = O_p(1)$ due to identification restriction $\hat{F}'\hat{F}/T = I_r$. Also, $\|H_F\| = O_p(1)$ because $\|H_F\| \leq \|\hat{F}'\hat{F}/T\|^{\frac{1}{2}}\|F'F/T\|^{\frac{1}{2}}\|\Lambda'\Lambda/N\|\|V_X^{-1}\| = O_p(1)$ by Assumptions 1, 2 and the following lemma.

Lemma A1: (Lemmas, A.1, A.3, B.1, and B.2 of Bai (2003)) Let $\delta_{NT} = \min\{\sqrt{N},\ \sqrt{T}\}$.

(i) Under Assumptions 1 - 4, $T^{-1}\sum_{t=1}^{T}\|\hat{F}_t - H_F' F_t\|^2 = O_p\left(\delta_{NT}^{-2}\right)$.

(ii) Under Assumptions 1 - 4, $V_X \to_p V$, where V is the diagonal matrix consisting of the eigenvalues of $\Sigma_F\Sigma_\Lambda$.

(iiI) Under Assumptions 1 - 6, $T^{-1}(\hat{F} - FH_F)'e_i = O_p\left(\delta_{NT}^{-2}\right)$.

(iv) Under Assumptions 1 - 6, $T^{-1}(\hat{F} - FH_F)'F = O_p\left(\delta_{NT}^{-2}\right)$.

Lemma A1(i) means that $\hat{F}_t$ is an consistent estimate for rotated factor $H_F' F_t$. Analogously, we can define $H_{\mathcal{F}}$ such that $\hat{\mathcal{F}}_t$ is a consistent estimate of $H_{\mathcal{F}}'\mathcal{F}_t$. Recall that $\hat{\mathcal{F}}_t = [\hat{F}_{t-1}', ..., \hat{F}_{t-p}']'$ and $\mathcal{F}_t = [F_{t-1}', ..., F_{t-p}']'$. Let $H_{\mathcal{F}} \equiv I_p \otimes H_F$, so Lemma A1(i) implies that $T^{-1}\sum_{p+1}^{T}\|\hat{\mathcal{F}}_t - H_{\mathcal{F}}'\mathcal{F}_t\|^2 = O_p\left(\delta_{NT}^{-2}\right)$, which is summarized in the following lemma.

Lemma A2: Under Assumptions 1 - 4,

(i) $T^{-1}\sum_{p+1}^{T}\|\hat{\mathcal{F}}_t - H_{\mathcal{F}}'\mathcal{F}_t\|^2 = O_p\left(\delta_{NT}^{-2}\right)$.

(ii) H_F and $H_{\mathcal{F}}$ are non-singular as N and $T \to \infty$.

Lemma A3:

(i) Under Assumptions 1 - 6, $T^{-1}(\hat{\mathcal{F}} - \mathcal{F}H_{\mathcal{F}})'e_i = O_p\left(\delta_{NT}^{-2}\right)$.

(ii) Under Assumptions 1 - 7, $T^{-1}N_S^{-1/2}(\hat{\mathcal{F}} - \mathcal{F}H_{\mathcal{F}})'\sum_{i\in S}e_i = O_p\left(\delta_{NT}^{-2}\right)$.

Lemma A4: Under Assumptions 1 - 6,

(i) $T^{-1}(\hat{\mathcal{F}} - \mathcal{F}H_{\mathcal{F}})'[\mathcal{F}\vdots\zeta] = O_p\left(\delta_{NT}^{-2}\right)$.

(ii) $\|(T^{-1}\hat{\mathcal{F}}'\hat{\mathcal{F}})^{-1}\| = O_p(1)$.

A.2 Lemmas on $\hat{\zeta}^S$ under the Null Hypothesis

Note that Γ^{SF} in Equation (3.6) is always zero, so X^S can be expressed as

$$X^S = \mathcal{F}\Pi^{S'} + \zeta^S\Gamma^{SS'} + e^S \tag{A.4}$$

By the definition of $\hat{X}^S$, we have

$$\hat{X}^S = M_{\hat{\mathcal{F}}}X^S = \zeta^S\Gamma^{SS'} + e^S - P_{\hat{\mathcal{F}}}(\zeta^S\Gamma^{SS'} + e^S) + M_{\hat{\mathcal{F}}}\mathcal{F}\Pi^{S'} \tag{A.5}$$

Recall that the estimator $\hat{\zeta}^S$ is equal to $\sqrt{T-p}$ times the eigenvectors corresponding to the $\kappa_S = q - \kappa_F$ largest eigenvalues of the $(T-p)\times(T-p)$ matrix $\hat{X}^S\hat{X}^{S'}$. Under the null hypothesis, $q_S = \kappa_S$ and the dimension of $\hat{\zeta}^S$ is the same as that of ζ^S. Hence, we have

$$(1/N^S T)\hat{X}^S\hat{X}^{S'}\hat{\zeta}^S = \hat{\zeta}^S\hat{V}_S \tag{A.6}$$

where $\hat{V}_S$ is the diagonal matrix consisting of the first q^S largest eigenvalues of $(1/N^S T)\hat{X}^S\hat{X}^{S'}$ in decreasing order. Let V_S be the diagonal matrix consisting of the first q^S largest eigenvalues of $(\zeta^S\Gamma^{SS'} + e^S)(\zeta^S\Gamma^{SS'} + e^S)'/N^S T$ in decreasing order. Amengual and Watson (2007) show that $\hat{V}_S - V_S \to_p 0$. Also, Lemma A1(ii) implies that V_S converges to a non-singular diagonal matrix consisting of the eigenvalues of $\Sigma_{\zeta^S}\Sigma_{\Gamma^{SS}} = \Sigma_{\Gamma^{SS}}$. Hence, $\hat{V}_S$ is non-singular asymptotically and $\|\hat{V}_S^{-1}\| = O_p(1)$. Also, note that $T^{-1}\sum_{t=p+1}^{T}\|\hat{\zeta}_t^S\|$ is $O_p(1)$ because $\hat{\zeta}^{S'}\hat{\zeta}^S/(T-p) = I_{q_S}$.

Now, substituting Equation (A.5) into Equation (A.6) gives

$$\frac{1}{N_S T}\left[\zeta^S\Gamma^{SS'} + e^S - P_{\hat{\mathcal{F}}}(\zeta^S\Gamma^{SS'} + e^S) + M_{\hat{\mathcal{F}}}\mathcal{F}\Pi^{S'}\right]\left[\zeta^S\Gamma^{SS'} + e^S - P_{\hat{\mathcal{F}}}(\zeta^S\Gamma^{SS'} + e^S) + M_{\hat{\mathcal{F}}}\mathcal{F}\Pi^{S'}\right]'\hat{\zeta}^S\hat{V}_S^{-1} = \hat{\zeta}^S$$

Define $H_{\zeta^S} = (\Gamma^{SS'}\Gamma^{SS}/N_S)(\zeta^{S'}\hat{\zeta}^S/T)\hat{V}_S^{-1}$. We have $H_{\zeta^S} = O_p(1)$ because $\|H_{\zeta^S}\| \leq \|\hat{\zeta}^{S'}\hat{\zeta}^S/T\|^{\frac{1}{2}}\|\zeta^{S'}\zeta^S/T\|^{\frac{1}{2}}$ $\|\Gamma^{SS'}\Gamma^{SS}/N_S\|\|\hat{V}_S^{-1}\| = O_p(1)$ by Assumptions 1 and 2. Analogous to Lemma A1(i), the following lemma shows that $\hat{\zeta}_t^S$ is a consistent estimate for $H'_{\zeta^S}\zeta_t^S$ and that the convergence rate is the same as in Lemma A1(i).

Lemma A5: Under Assumptions 1 - 6, $T^{-1}\sum_{t=p+1}^{T}\|\hat{\zeta}_t^{\,S} - H'_{\zeta^S}\zeta_t^S\|^2 = O_p\left(\delta_{NT}^{-2}\right)$.

Lemma A6: Under Assumptions 1 - 6, $T^{-1}\sum_{t=p+1}^{T}\|b_{\ell t}\|^2 = O_p(\delta_{NT}^{-4})$ for $\ell = 1, 2, 4, 5, 6, 8, 10, ..., 15$, where $b_{\ell t}$'s are defined in the proof of Lemma A5.

Lemma A7:

(i) Under Assumptions 1 - 6, $T^{-1}(\hat{\zeta}^S - \zeta^S H_{\zeta^S})'e_i = O_p(\delta_{NT}^{-2})$.

(ii) Under Assumptions 1 - 7, $T^{-1}N_S^{-1/2}(\hat{\zeta}^S - \zeta^S H_{\zeta^S})'\sum_{i\in S}e_i = O_p(\delta_{NT}^{-2})$.

Lemma A8: Under Assumptions 1 - 6,

(i) $T^{-1}(\hat{\zeta}^S - \zeta^S H_{\zeta^S})'\zeta = O_p(\delta_{NT}^{-2})$.

(ii) H_{ζ^S} is non-singular as N and $T \to \infty$.

A.3 Lemmas on $\hat{\zeta}^F$ under the Null Hypothesis

By the definitions of $\tilde{X}$, we have

$$\begin{aligned}
\tilde{X} = M_{\hat{\zeta}^S}\hat{X} = M_{\hat{\zeta}^S}M_{\hat{\mathcal{F}}}X &= M_{\hat{\zeta}^S}M_{\hat{\mathcal{F}}}(\mathcal{F}\Pi' + \zeta\Gamma' + e) \\
&= M_{\hat{\zeta}^S}M_{\hat{\mathcal{F}}}(\mathcal{F}\Pi' + [\zeta^S\vdots\zeta^F]\begin{bmatrix} \Gamma^{SS'} & \Gamma^{FS'} \\ \Gamma^{SF'} & \Gamma^{FF'} \end{bmatrix} + e) \\
&= M_{\hat{\zeta}^S}M_{\hat{\mathcal{F}}}(\mathcal{F}\Pi' + \zeta^S[\Gamma^{SS'}\vdots\Gamma^{FS'}] + \zeta^F[\Gamma^{SF'}\vdots\Gamma^{FF'}] + e) \\
&= M_{\hat{\zeta}^S}M_{\hat{\mathcal{F}}}(\mathcal{F}\Pi' + \zeta^S\Gamma^{S'} + \zeta^F\Gamma^{F'} + e) \qquad \text{(A.7)}
\end{aligned}$$

where we use the definitions that $\Gamma^S = [\Gamma^{SS'}\vdots\Gamma^{FS'}]'$, $\Gamma^F = [\Gamma^{SF'}\vdots\Gamma^{FF'}]'$ and $\Gamma = [\Gamma^S\vdots\Gamma^F]$. Define $G = [\mathcal{F}\vdots\zeta^S]$, $\Theta = [\Pi\vdots\Gamma^S]$, and $\hat{G} = [\hat{\mathcal{F}}\vdots\hat{\zeta}^S]$. Note that $\hat{\mathcal{F}}'\hat{\zeta}^S = 0$, so $M_{\hat{\zeta}^S}M_{\hat{\mathcal{F}}} = M_{\hat{G}}$. Hence, Equation (A.7) can be rewritten as

$$\tilde{X} = M_{\hat{G}}X = \zeta^F\Gamma^{F'} + e - P_{\hat{G}}(\zeta^F\Gamma^{F'} + e) + M_{\hat{G}}G\Theta' \qquad \text{(A.8)}$$

Recall that the estimate of fast structural shocks $\hat{\zeta}^F$ is equal to $\sqrt{T-p}$ times the eigenvectors corresponding to the κ_F largest eigenvalues of the $(T-p)\times(T-p)$ matrix $\tilde{X}\tilde{X}'$. Under the null hypothesis, $q_F = \kappa_F > 0$ and the dimension of $\hat{\zeta}^F$ is the same as that of ζ_F. Hence, we have

$$(1/NT)\tilde{X}\tilde{X}'\hat{\zeta}^F = \hat{\zeta}^F\hat{V}_F \qquad \text{(A.9)}$$

where $\hat{V}_F$ is the diagonal matrix consisting of the first q_F largest eigenvalues of $(1/NT)\tilde{X}\tilde{X}'$ in decreasing order. Note that $T^{-1}\sum_{t=p+1}^{T}\|\hat{\zeta}_t^F\|$ is $O_p(1)$ because $\hat{\zeta}^{F'}\hat{\zeta}^F/(T-p) = I_{q_F}$.

Let V_F be the diagonal matrix consisting of the first q_F largest eigenvalues of $(\zeta^F\Gamma^{F'}+e)(\zeta^F\Gamma^{F'}+e)'/NT$ in decreasing order. Note that the difference between $\tilde{X}$ and $\zeta^F\Gamma^{F'}+e$ is that $\tilde{X}$ contains some additional estimation error terms defined in Equation (A.8). Amengual and Watson (2007) show that such estimation errors will not affect the limits of the first q_F largest eigenvalues of $\tilde{X}\tilde{X}'/NT$, i.e. $\hat{V}_F - V_F \to_p 0$. Also, Lemma A1(ii) implies that V_F converges to a non-singular diagonal matrix consisting of the eigenvalues of $\lim_{N,T\to\infty}(\zeta^{F'}\zeta^F/T)(\Gamma^{F'}\Gamma^F/N) = (1-m)\Sigma_F\Sigma_{\Gamma^{FF}}$, where we use the facts that $\Gamma^{SF} = 0$ and that $N_F/N = 1 - N_S/N \to 1-m > 0$. Hence, $\hat{V}_F$ is non-singular asymptotically and $\|\hat{V}_F^{-1}\| = O_p(1)$.

Now, substituting Equation (A.8) into Equation (A.9) gives

$$\frac{1}{NT}\left[\zeta^F\Gamma^{F'} + e - P_{\hat{G}}(\zeta^F\Gamma^{F'}+e) + M_{\hat{G}}G\Theta'\right]\left[\zeta^F\Gamma^{F'} + e - P_{\hat{G}}(\zeta^F\Gamma^{F'}+e) + M_{\hat{G}}G\Theta'\right]'\hat{\zeta}^F\hat{V}_F^{-1} = \hat{\zeta}^F$$

Define $H_{\zeta^F} = (\Gamma^{F'}\Gamma^F/N)(\zeta^{F'}\hat{\zeta}^F/T)\hat{V}_F^{-1}$. We have $H_{\zeta^F} = O_p(1)$ because $\|H_{\zeta^F}\| \le \|\hat{\zeta}^{F'}\hat{\zeta}^F/T\|^{\frac{1}{2}}\|\zeta^{F'}\zeta^F/T\|^{\frac{1}{2}}$ $\|\Gamma^{F'}\Gamma^F/N\|\|\hat{V}_F^{-1}\| = O_p(1)$ by Assumptions 1 and 2. Analogous to Lemmas A1(i) and A5, the following lemma shows that $\hat{\zeta}_t^F$ is a consistent estimate for $H'_{\zeta^F}\zeta_t^F$ and that the convergence rate is the same as in Lemmas A1(i) and A5.

Lemma A9: Under Assumptions 1 - 6, $T^{-1}\sum_{t=p+1}^{T}\|\hat{\zeta}_t^F - H'_{\zeta^F}\zeta_t^F\|^2 = O_p\left(\delta_{NT}^{-2}\right)$.

A sketch of the proof: (see the supplement appendix for details)

First, we show that $\hat{G}'\hat{\zeta}^F = 0$. Note that $\hat{\zeta}^{S'}\hat{\zeta}^F = \hat{\zeta}^{S'}\tilde{X}\tilde{X}'\hat{\zeta}^F\hat{V}_F^{-1}/NT = 0$ because $\hat{\zeta}^{S'}\tilde{X} = 0$. Also, $\hat{\mathcal{F}}'\hat{\zeta}^F = \hat{\mathcal{F}}'M_{\hat{\zeta}^S}\hat{X}\tilde{X}'\hat{\zeta}^F\hat{V}_F^{-1}/NT = \hat{\mathcal{F}}'\hat{X}\tilde{X}'\hat{\zeta}^F\hat{V}_F^{-1}/NT = 0$ because $\hat{\mathcal{F}}'P_{\hat{\zeta}^S} = 0$ and $\hat{\mathcal{F}}'\hat{X} = 0$. Hence,

$\hat{G}'\hat{\zeta}^F = [\hat{\mathcal{F}} \vdots \hat{\zeta}^S]'\hat{\zeta}^F = 0$.

Now, define

$$H_G = \begin{bmatrix} H_{\mathcal{F}} & 0 \\ 0 & H_{\zeta^S} \end{bmatrix}$$

Since $G = [\mathcal{F} \vdots \zeta^S]$ and $\hat{G} = [\hat{\mathcal{F}} \vdots \hat{\zeta}^S]$, we have $T^{-1}\sum_{t=p+1}^{T} \|\hat{G}_t - H_G' G_t\|^2 = O_p(\delta_{NT}^{-2})$ by Lemmas A2(i) and A5. Since both $H_{\mathcal{F}}$ and H_{ζ^S} are non-singular asymptotically by Lemmas A2(ii) and A8(ii), so H_G is also non-singular asymptotically and $\|H_G^{-1}\| = O_p(1)$. Also,

$$\|(\hat{G}'\hat{G}/T)^{-1}\| = O_p(1) \tag{A.10}$$

To see this, note that $\hat{\mathcal{F}}'\hat{\zeta}^S = 0$, so

$$\frac{\hat{G}'\hat{G}}{T} = \begin{bmatrix} \hat{\mathcal{F}}'\hat{\mathcal{F}}/T & 0 \\ 0 & \hat{\zeta}^{S'}\hat{\zeta}^S/T \end{bmatrix}$$

Since $\|(\hat{\mathcal{F}}'\hat{\mathcal{F}}/T)^{-1}\| = O_p(1)$ by Lemma A4(ii) and $\hat{\zeta}^{S'}\hat{\zeta}^S/T = I_{q_S}(T-p)/T \to I_{q_S}$, we have $\|(\hat{G}'\hat{G}/T)^{-1}\| = O_p(1)$. Furthermore, for an $N \times k_0$ matrix Z, since $\Theta = [\Pi \vdots \Gamma^S]$, $\|\Theta' Z\|$ and $\|\Gamma^{F'} Z\|$ can be always bounded by $\|[\Pi \vdots \Gamma]' Z\|$ and $\|\Gamma' Z\|$, respectively. This trick can be used for any term containing Θ or Γ^F.

Based on the above results, it is straightforward to prove that $T^{-1}\sum_{t=p+1}^{T}\|c_{\ell t}\|^2 = O_p\left(\delta_{NT}^{-2}\right)$ for $\ell = 1, ..., 15$ using very similar arguments in the proof of $T^{-1}\sum_{t=p+1}^{T}\|b_{\ell t}\|^2 = O_p\left(\delta_{NT}^{-2}\right)$ for $\ell = 1, ..., 15$, because $c_{\ell t}$ has a very similar structure to that of $b_{\ell t}$. To save space, we skip the detailed proof. Q.E.D.

Lemma A10: Under Assumptions 1 - 6, $T^{-1}\sum_{t=p+1}^{T}\|c_{\ell t}\|^2 = O_p(\delta_{NT}^{-4})$ for $\ell = 1, 2, 4, 5, 6, 8, 10, ..., 15$, where $c_{\ell t}$'s are defined in the proof of Lemma A9.

Lemma A11:

(i) Under Assumptions 1 - 6, $T^{-1}(\hat{\zeta}^F - \zeta^F H_{\zeta^F})' e_i = O_p(\delta_{NT}^{-2})$.

(ii) Under Assumptions 1 - 7, $T^{-1}N_S^{-1/2}(\hat{\zeta}^F - \zeta^F H_{\zeta^F})' \sum_{i \in S} e_i = O_p(\delta_{NT}^{-2})$.

B Proof of Theorem 1

Lemma B1:

(i) Under Assumptions 1 - 6, $T^{-1}\left(\hat{\zeta}^{F'}\tilde{X}_i - H_{\zeta^F}'\zeta^{F'} e_i\right) = O_p(\delta_{NT}^{-2})$ for $i \in S$ if $q_S = \kappa_S$.

(ii) Under Assumptions 1 - 7, $T^{-1}N_S^{-1/2}\left(\sum_{i \in S}\hat{\zeta}^{F'}\tilde{X}_i - \sum_{i \in S} H_{\zeta^F}'\zeta^{F'} e_i\right) = O_p(\delta_{NT}^{-2})$ if $q_S = \kappa_S$.

Proof:

(i) Recall Equation (3.6). Since $q_S = \kappa_S$ under the null hypothesis, we have

$$X^S = \mathcal{F}\Pi^{S'} + \zeta^S \Gamma^{SS'} + e^S = G\Theta^{S'} + e^S$$

where $\Theta^S = [\Pi^S \vdots \Gamma^{SS}](N_S \times (rp + q_S))$. Hence, we have $X_i = G\theta_i + e_i$ and

$$\tilde{X}_i = M_{\hat{\zeta}^S} M_{\hat{\mathcal{F}}} X_i = M_{\hat{G}} X_i = e_i - P_{\hat{G}} e_i + M_{\hat{G}} G\theta_i \tag{B.1}$$

where θ_i is the transpose of the i^{th} row of Θ^S. Since $\hat{G}'\hat{\zeta}^F = 0$ (see the proof of Lemma A9), we have $\hat{\zeta}^{F'} P_{\hat{G}} = 0$ and

$$\begin{aligned}\frac{1}{T}(\hat{\zeta}^{F'}\tilde{X}_i - H'_{\zeta^F}\zeta^{F'}e_i) &= \frac{1}{T}(\hat{\zeta}^F - \zeta^F H_{\zeta^F})'e_i + \frac{1}{T}\hat{\zeta}^{F'} M_{\hat{G}} G\theta_i \\ &= \frac{1}{T}(\hat{\zeta}^F - \zeta^F H_{\zeta^F})'e_i + \frac{1}{T}\hat{\zeta}^{F'} G\theta_i \qquad \text{(B.2)}\end{aligned}$$

where $T^{-1}(\hat{\zeta}^F - \zeta^F H_{\zeta^F})'e_i = O_p(\delta_{NT}^{-2})$ by Lemma A11(i). Hence, it is sufficient to show that $T^{-1}\hat{\zeta}^{F'}G\theta_i = O_p(\delta_{NT}^{-2})$. To see this, note that $\|H_G^{-1}\| = O_p(1)$ (see the proof of Lemma A9) and $E\|\theta_i\| \leq (\|\Phi\| + \|GA^{-1}\|)E\|\lambda_i\| < \infty$ by Assumptions 2(a) and 2(b). Now, we have $T^{-1}\hat{\zeta}^{F'}G\theta_i = O_p(\delta_{NT}^{-2})$, because

$$\begin{aligned}\frac{1}{T}\hat{\zeta}^{F'}G\theta_i &= \frac{1}{T}\hat{\zeta}^{F'}(GH_G - \hat{G})H_G^{-1}\theta_i \\ &= \frac{1}{T}(\hat{\zeta}^F - \zeta^F H_{\zeta^F})'(GH_G - \hat{G})H_G^{-1}\theta_i + \frac{1}{T}H'_{\zeta^F}\zeta^{F'}(GH_G - \hat{G})H_G^{-1}\theta_i \qquad \text{(B.3)}\end{aligned}$$

where the first term is $O_p(\delta_{NT}^{-2})$ by Cauchy-Schwarz inequality and Lemmas A2, A5 and A9, and the second term is $O_p(\delta_{NT}^{-2})$ by Lemmas A4(i) and A8(i).

(ii) Similarly to Equation (B.2), we have

$$\begin{aligned}\frac{1}{T\sqrt{N_S}}\left(\sum_{i\in S}\hat{\zeta}^{F'}\tilde{X}_i - \sum_{i\in S}H'_{\zeta^F}\zeta^{F'}e_i\right) &= \frac{1}{T}(\hat{\zeta}^F - \zeta^F H_{\zeta^F})'\frac{1}{\sqrt{N_S}}\sum_{i\in S}e_i + \frac{1}{T}\hat{\zeta}^{F'}M_{\hat{G}}G\frac{1}{\sqrt{N_S}}\sum_{i\in S}\theta_i \\ &= \frac{1}{T}(\hat{\zeta}^F - \zeta^F H_{\zeta^F})'\frac{1}{\sqrt{N_S}}\sum_{i\in S}e_i + \frac{1}{T}\hat{\zeta}^{F'}G\frac{1}{\sqrt{N_S}}\sum_{i\in S}\theta_i\end{aligned}$$

where the first term is $O_p(\delta_{NT}^{-2})$ by Lemma A11(ii) and the second term is $O_p(\delta_{NT}^{-2})$ because $T^{-1}\hat{\zeta}^{F'}G = O_p(\delta_{NT}^{-2})$ (see the proof in part (i) of this lemma) and $N_S^{-1/2}\sum_{i\in S}\theta_i = O_p(1)$ by Assumption 7(a). Q.E.D.

Lemma B2: Under Assumptions 1 - 6, if $q_S = \kappa_S$, then

(i) $T^{-1}\sum_{t=p+1}^T \|\hat{\zeta}_t^F - H'_{\zeta^F}\zeta_t^F\|^4 = O_p(T/\delta_{NT}^4)$, $T^{-1}\sum_{t=p+1}^T \|\hat{G}_t - H'_G G_t\|^4 = O_p(T/\delta_{NT}^4)$, $T^{-1}\sum_{t=p+1}^T \|\hat{\zeta}_t^F\|^4 = O_p(T/\delta_{NT}^4) + O_p(1)$, and $T^{-1}\sum_{t=p+1}^T \|\hat{G}_t\|^4 = O_p(T/\delta_{NT}^4) + O_p(1)$.

(ii) If $\sqrt{T}/N \to 0$, then $T^{-1}\sum_{t=p+1}^T |\tilde{e}_{it} - e_{it}|^4 = o_p(1)$ for $i \in S$ and $(N_S T)^{-1}\sum_{t=p+1}^T \sum_{i\in S}|\tilde{e}_{it} - e_{it}|^4 = o_p(1)$

Proof:

(i) Since Lemma A9 shows that $T^{-1}\sum_{t=p+1}^T \|\hat{\zeta}_t^F - H'_{\zeta^F}\zeta_t^F\|^2 = O_p(\delta_{NT}^{-2})$, we have

$$T^{-1}\sum_{t=p+1}^T \|\hat{\zeta}_t^F - H'_{\zeta^F}\zeta_t^F\|^4 \leq T^{-1}\sum_{t=p+1}^T \|\hat{\zeta}_t^F - H'_{\zeta^F}\zeta_t^F\|^2 (\sum_{t=p+1}^T \|\hat{\zeta}_t^F - H'_{\zeta^F}\zeta_t^F\|^2) = O_p\left(\frac{T}{\delta_{NT}^4}\right).$$

Similarly, Lemmas A2 and A5 imply that $T^{-1}\sum_{t=p+1}^T \|\hat{G}_t - H'_G G_t\|^4 = O_p(T/\delta_{NT}^4)$. Also, $T^{-1}\sum_{t=p+1}^T \|\hat{\zeta}_t^F\|^4$

can be bounded by

$$
\begin{aligned}
T^{-1}\sum_{t=p+1}^{T}\|\hat{\zeta}_t^{\,F}\|^4 &\leq \frac{8}{T}\sum_{t=p+1}^{T}\|\hat{\zeta}_t^{\,F}-H'_{\zeta^F}\zeta_t^F\|^4+\frac{8}{T}\sum_{t=p+1}^{T}\|H'_{\zeta^F}\zeta_t^F\|^4 \\
&= O_p\left(\frac{T}{\delta_{NT}^4}\right)+O_p(1)
\end{aligned}
$$

where the $O_p(1)$ term follows from Assumption 1(b) and the fact that $\|H_{\zeta^F}\|^4 = O_p(1)$. The proof of $T^{-1}\sum_{t=p+1}^{T}\|\hat{G}_t\|^4 = O_p(T/\delta_{NT}^4)+O_p(1)$ is similar.

(ii) By the definition of $\tilde{e}_i$ and the fact that $\hat{G}'\hat{\zeta}^F=0$, we have $\tilde{e}_i = M_{\hat{\zeta}^F}\tilde{X}_i = e_i - P_{\hat{G}}e_i - P_{\hat{\zeta}^F}e_i + G\theta_i - P_{\hat{G}}G\theta_i - P_{\hat{\zeta}^F}G\theta_i$. Thus,

$$
\tilde{e}_{it}-e_{it} = d_{1,it}+d_{2,it}+d_{3,it}+d_{4,it}
$$

where $d_{1,it} = -\hat{G}'_t\left(\hat{G}'\hat{G}/T\right)^{-1}\hat{G}'e_i/T$, $d_{2,it} = -\hat{\zeta}_t^{F'}\left(\hat{\zeta}^{F'}\hat{\zeta}^F/T\right)^{-1}\hat{\zeta}^{F'}e_i/T$, $d_{3,it} = G'_t\theta_i - \hat{G}'_t\left(\hat{G}'\hat{G}/T\right)^{-1}\hat{G}'G\theta_i/T$, and $d_{4,it} = -\hat{\zeta}_t^{F'}\left(\hat{\zeta}^{F'}\hat{\zeta}^F/T\right)^{-1}\hat{\zeta}^{F'}G\theta_i/T$. Note that $|\tilde{e}_{it}-e_{it}|^4 \leq 64(d_{1,it}^4+d_{2,it}^4+d_{3,it}^4+d_{4,it}^4)$, so it is enough to consider the bounds of $d_{\ell,it}^4$ for $\ell = 1,2,3,4$.

First, $T^{-1}\sum_{t=p+1}^{T}d_{1,it}^4$ can be bounded by

$$
\begin{aligned}
T^{-1}\sum_{t=p+1}^{T}d_{1,it}^4 &\leq \left(\frac{1}{T}\sum_{t=p+1}^{T}\|\hat{G}_t\|^4\right)\left\|\left(\frac{\hat{G}'\hat{G}}{T}\right)^{-1}\right\|^4\left\|\frac{\hat{G}'e_i}{T}\right\|^4 \\
&\leq 64\left[O_p\left(\frac{T}{\delta_{NT}^4}\right)+O_p(1)\right]\left\|\left(\frac{\hat{G}'\hat{G}}{T}\right)^{-1}\right\|^4\left(\left\|\frac{(\hat{G}-GH_G)'e_i}{T}\right\|^4+\left\|\frac{H'_GG'e_i}{T}\right\|^4\right) \\
&= \left[O_p\left(\frac{T}{\delta_{NT}^4}\right)+O_p(1)\right]O_p(1)O_p\left(\frac{1}{\delta_{NT}^4}\right)
\end{aligned}
$$

where $T^{-1}\sum_{t=p+1}^{T}\|\hat{G}_t\|^4 = O_p(T/\delta_{NT}^4)+O_p(1)$ by part (i) of this lemma, $\|(\hat{G}'\hat{G}/T)^{-1}\| = O_p(1)$ by Equation (A.10), $\|T^{-1}(\hat{G}-GH_G)'e_i\|^4 \leq \left(T^{-1}\sum_{t=p+1}^{T}\|\hat{G}_t-H'_GG_t\|^2T^{-1}\sum_{t=p+1}^{T}e_{it}^2\right)^2 = O_p(\delta_{NT}^{-4})$ by Lemmas A2 and A5 and Assumption 3(a), and $\|T^{-1}G'e_i\|^4 = O_p(T^{-2})$ by Assumption 6(e).

Similarly, $(N_ST)^{-1}\sum_{t=p+1}^{T}\sum_{i\in S}d_{1,it}^4$ can be bounded by

$$
\frac{1}{TN_S}\sum_{t=p+1}^{T}\sum_{i\in S}d_{1,it}^4 \leq \left[O_p\left(\frac{T}{\delta_{NT}^4}\right)+O_p(1)\right]O_p(1)\frac{1}{N_S}\sum_{i\in S}\left(\left\|\frac{(\hat{G}-GH_G)'e_i}{T}\right\|^4+\left\|\frac{H'_GG'e_i}{T}\right\|^4\right)
$$

where $N_S^{-1}\sum_{i\in S}\|T^{-1}H'_GG'e_i\|^4 = O_p(T^{-2})$ by Assumption 6(e), and $N_S^{-1}\sum_{i\in S}\|T^{-1}(\hat{G}-GH_G)'e_i\|^4 = O_p(\delta_{NT}^{-4})$ because

$$
\begin{aligned}
\frac{1}{N_S}\sum_{i\in S}\left\|\frac{(\hat{G}-GH_G)'e_i}{T}\right\|^4 &\leq \frac{1}{N_S}\sum_{i\in S}\left(\frac{1}{T}\sum_{t=p+1}^{T}\|\hat{G}_t-H'_GG_t\|^2\frac{1}{T}\sum_{t=p+1}^{T}e_{it}^2\right)^2 \\
&\leq \left(\frac{1}{T}\sum_{t=p+1}^{T}\|\hat{G}_t-H'_GG_t\|^2\right)^2\frac{1}{N_ST}\sum_{i\in S}\sum_{t=p+1}^{T}e_{it}^4 = O_p\left(\frac{1}{\delta_{NT}^4}\right)O_p(1)
\end{aligned}
$$

by Lemmas A2 and A5 and Assumption 3(a). Since $\sqrt{T}/N \to 0$, we have $T/\delta_{NT}^4 \to 0$. Thus, $T^{-1}\sum_{t=p+1}^{T} d_{1,it}^4 = o_p(1)$ and $(N_S T)^{-1}\sum_{t=p+1}^{T}\sum_{i\in S} d_{1,it}^4 = o_p(1)$. The structure of $d_{2,it}$ is similar to that of $d_{1,it}$, so it is straightforward to prove that $T^{-1}\sum_{t=p+1}^{T} d_{2,it}^4 = o_p(1)$ and $(N_S T)^{-1}\sum_{t=p+1}^{T}\sum_{i\in S} d_{2,it}^4 = o_p(1)$ using a similar argument.

For term $d_{3,it}$, we have

$$\begin{aligned}
T^{-1}\sum_{t=p+1}^{T} d_{3,it}^4 &\leq T^{-1}\sum_{t=p+1}^{T}\left\| G_t' - \hat{G}_t'\left(\frac{\hat{G}'\hat{G}}{T}\right)^{-1}\frac{\hat{G}'G}{T}\right\|^4 \|\theta_i\|^4 \\
&\leq T^{-1}\sum_{t=p+1}^{T}\left[\|(G_t'H_G - \hat{G}_t')H_G^{-1}\|^4 + \left\|\hat{G}_t'\left(\frac{\hat{G}'\hat{G}}{T}\right)^{-1}\frac{\hat{G}'(GH_G-\hat{G})H_G^{-1}}{T}\right\|^4\right]\|\theta_i\|^4 \\
&\leq \left[O_p\left(\frac{T}{\delta_{NT}^4}\right) + T^{-1}\sum_{t=p+1}^{T}\|\hat{G}_t\|^4\left\|\left(\frac{\hat{G}'\hat{G}}{T}\right)^{-1}\right\|^4\left\|\frac{\hat{G}'(GH_G-\hat{G})H_G^{-1}}{T}\right\|^4\right]\|\theta_i\|^4 \\
&= \left\{O_p\left(\frac{T}{\delta_{NT}^4}\right) + \left[O_p\left(\frac{T}{\delta_{NT}^4}\right) + O_p(1)\right]O_p(1)O_p\left(\frac{1}{\delta_{NT}^4}\right)\right\}O_p(1) = O_p\left(\frac{T}{\delta_{NT}^4}\right)
\end{aligned}$$

where $\|T^{-1}\hat{G}'(GH_G - \hat{G})\|^4 \leq \left(T^{-1}\sum_{t=p+1}^{T}\|\hat{G}_t - H_G'G_t\|^2 T^{-1}\sum_{t=p+1}^{T}\|\hat{G}_t\|^2\right)^2 = O_p(\delta_{NT}^{-4})$ by Lemmas A2 and A5, and $\|\theta_i\|^4 \leq (\|\Phi\| + \|GA^{-1}\|)^4\|\lambda_i\|^4 = O_p(1)$ by Assumptions 2(a) and 2(b). As $\sqrt{T}/N \to 0$, $T^{-1}\sum_{t=p+1}^{T} d_{3,it}^4 = o_p(1)$. Also, since $N_S^{-1}\sum_{i\in S}\|\theta_i\|^4 = O_p(1)$ by Assumptions 2(a) and 2(b), we have $(N_S T)^{-1}\sum_{t=p+1}^{T}\sum_{i\in S} d_{3,it}^4 = o_p(1)$.

Lastly, since $\hat{\zeta}^{F'}\hat{\zeta}^F/(T-p) = I_{q_F}$, we can bound $T^{-1}\sum_{t=p+1}^{T} d_{4,it}^4$ by

$$\begin{aligned}
T^{-1}\sum_{t=p+1}^{T} d_{4,it}^4 &\leq T^{-1}\sum_{t=p+1}^{T}\|\hat{\zeta}_t^{F'}\|^4\left\|\frac{\hat{\zeta}^{F'}G}{T-p}\right\|^4\|\theta_i\|^4 \\
&= \left[O_p\left(\frac{T}{\delta_{NT}^4}\right) + O_p(1)\right]O_p\left(\frac{1}{\delta_{NT}^8}\right)O_p(1)
\end{aligned}$$

where $\|\hat{\zeta}^{F'}G/(T-p)\|^4 = O_p(\delta_{NT}^{-8})$ is due to the proof following Equation (B.3). As $\sqrt{T}/N \to 0$, $T^{-1}\sum_{t=p+1}^{T} d_{4,it}^4 = o_p(1)$. Also, since $N_S^{-1}\sum_{i\in S}\|\theta_i\|^4 = O_p(1)$ by Assumptions 2(a) and 2(b), we have $(N_S T)^{-1}\sum_{t=p+1}^{T}\sum_{i\in S} d_{4,it}^4 = o_p(1)$. Q.E.D.

Lemma B3: Under Assumptions 1 - 6, if $\sqrt{T}/N \to 0$ and $q_S = \kappa_S$, then

(i)

$$(T-p)^{-1}\left[\left(\sum_{t=p+1}^{T}\hat{\zeta}_t^F\hat{\zeta}_t^{F'}\tilde{e}_{it}^2\right) - H_{\zeta^F}'\left(\sum_{t=p+1}^{T}\zeta_t^F\zeta_t^{F'}e_{it}^2\right)H_{\zeta^F}\right] \to_p 0$$

(ii)

$$(T-p)^{-1}N_S^{-1}\left[\left(\sum_{t=p+1}^{T}\sum_{i\in S}\hat{\zeta}_t^F\hat{\zeta}_t^{F'}\tilde{e}_{it}^2\right) - H_{\zeta^F}'\left(\sum_{t=p+1}^{T}\sum_{i\in S}\zeta_t^F\zeta_t^{F'}e_{it}^2\right)H_{\zeta^F}\right] \to_p 0$$

Proof:

By Lemma B2(ii) and Assumption 3(a), we have

$$\frac{1}{T-p}\sum_{t=p+1}^{T}\tilde{e}_{it}^4 \le \frac{8}{T-p}\sum_{t=p+1}^{T}(\tilde{e}_{it}-e_{it})^4 + \frac{8}{T-p}\sum_{t=p+1}^{T}e_{it}^4 = O_p(1) \tag{B.4}$$

$$\frac{1}{(T-p)N_S}\sum_{i\in S}\sum_{t=p+1}^{T}\tilde{e}_{it}^4 \le \frac{8}{(T-p)N_S}\sum_{i\in S}\sum_{t=p+1}^{T}(\tilde{e}_{it}-e_{it})^4 + \frac{8}{(T-p)N_S}\sum_{i\in S}\sum_{t=p+1}^{T}e_{it}^4 = O_p(1) \tag{B.5}$$

(i)

$$\begin{aligned}
&\frac{1}{T-p}\sum_{t=p+1}^{T}\hat{\zeta}_t^F\hat{\zeta}_t^{F'}\tilde{e}_{it}^2 - H'_{\zeta^F}\left(\frac{1}{T-p}\sum_{t=p+1}^{T}\zeta_t^F\zeta_t^{F'}e_{it}^2\right)H_{\zeta^F}\\
\le\;& \frac{1}{T-p}\sum_{t=p+1}^{T}(\hat{\zeta}_t^F\hat{\zeta}_t^{F'} - H'_{\zeta^F}\zeta_t^F\zeta_t^{F'}H_{\zeta^F})\tilde{e}_{it}^2 + \frac{1}{T-p}\sum_{t=p+1}^{T}H'_{\zeta^F}\zeta_t^F\zeta_t^{F'}H_{\zeta^F}(\tilde{e}_{it}^2-e_{it}^2)
\end{aligned}$$

The first term is bounded by

$$\begin{aligned}
&\left(\frac{1}{T-p}\sum_{t=p+1}^{T}\|\hat{\zeta}_t^F\hat{\zeta}_t^{F'} - H'_{\zeta^F}\zeta_t^F\zeta_t^{F'}H_{\zeta^F}\|^2\frac{1}{T-p}\sum_{t=p+1}^{T}\tilde{e}_{it}^4\right)^{\frac{1}{2}}\\
\le\;& \left(\frac{2}{T-p}\sum_{t=p+1}^{T}\|\hat{\zeta}_t^F(\hat{\zeta}_t^{F'} - \zeta_t^{F'}H_{\zeta^F})\|^2 + \|(\hat{\zeta}_t^{\,F} - H'_{\zeta^F}\zeta_t^F)\zeta_t^{F'}H_{\zeta^F}\|^2\right)^{\frac{1}{2}}\left(\frac{1}{T-p}\sum_{t=p+1}^{T}\tilde{e}_{it}^4\right)^{\frac{1}{2}}\\
\le\;& \Bigg[2\left(\frac{1}{T-p}\sum_{t=p+1}^{T}\|\hat{\zeta}_t^F\|^4\frac{1}{T-p}\sum_{t=p+1}^{T}\|\hat{\zeta}_t^{F'} - \zeta_t^{F'}H_{\zeta^F}\|^4\right)^{\frac{1}{2}} + \\
&2\left(\frac{1}{T-p}\sum_{t=p+1}^{T}\|\zeta_t^{F'}H_{\zeta^F}\|^4\frac{1}{T-p}\sum_{t=p+1}^{T}\|\hat{\zeta}_t^{\,F} - H'_{\zeta^F}\zeta_t^F\|^4\right)^{\frac{1}{2}}\Bigg]^{\frac{1}{2}} O_p(1) = o_p(1)
\end{aligned}$$

because $(T-p)^{-1}\sum_{t=p+1}^{T}\|\hat{\zeta}_t^{F'} - \zeta_t^{F'}H_{\zeta^F}\|^4 = o_p(1)$ and $(T-p)^{-1}\sum_{t=p+1}^{T}\|\hat{\zeta}_t^F\|^4 = O_p(1)$ by Lemma B(i) and the assumption that $\sqrt{T}/N \to 0$, $(T-p)^{-1}\sum_{t=p+1}^{T}\|\zeta_t^{F'}H_{\zeta^F}\|^4 = O_p(1)$ by Assumption 1(b), and $(T-p)^{-1}\sum_{t=p+1}^{T}\tilde{e}_{it}^4 = O_p(1)$ by Equation (B.4). The second term can be bounded by

$$\begin{aligned}
&\left(\frac{1}{T-p}\sum_{t=p+1}^{T}\|H'_{\zeta^F}\zeta_t^F\zeta_t^{F'}H_{\zeta^F}\|^2\frac{1}{T-p}\sum_{t=p+1}^{T}(\tilde{e}_{it}^2-e_{it}^2)^2\right)^{\frac{1}{2}}\\
\le\;& \|H_{\zeta^F}\|^2\left(\frac{1}{T-p}\sum_{t=p+1}^{T}\|\zeta_t^F\|^4\right)^{\frac{1}{2}}\left(\frac{2}{T-p}\sum_{t=p+1}^{T}\tilde{e}_{it}^2(\tilde{e}_{it}-e_{it})^2 + e_{it}^2(\tilde{e}_{it}-e_{it})^2\right)^{\frac{1}{2}}\\
=\;& O_p(1)\left\{\left[\frac{2}{T-p}\sum_{t=p+1}^{T}\tilde{e}_{it}^4\frac{2}{T-p}\sum_{t=p+1}^{T}(\tilde{e}_{it}-e_{it})^4\right]^{\frac{1}{2}} + \left[\frac{2}{T-p}\sum_{t=p+1}^{T}e_{it}^4\frac{2}{T-p}\sum_{t=p+1}^{T}(\tilde{e}_{it}-e_{it})^4\right]^{\frac{1}{2}}\right\}^{\frac{1}{2}} = o_p(1)
\end{aligned}$$

because $(T-p)^{-1}\sum_{t=p+1}^{T}\|\zeta_t^F\|^4 = O_p(1)$ by Assumption 1(b), $(T-p)^{-1}\sum_{t=p+1}^{T}\tilde{e}_{it}^4 = O_p(1)$ by Equation

(B.4), $(T-p)^{-1}\sum_{t=p+1}^{T} e_{it}^4 = O_p(1)$ by Assumption 3(a), and $(T-p)^{-1}\sum_{t=p+1}^{T}(\tilde{e}_{it}-e_{it})^4 = o_p(1)$ by Lemma B2(ii).

(ii) The proof of this part is similar to that of part (i).

$$\begin{aligned}
& \frac{1}{N_S T}\left[\left(\sum_{t=p+1}^{T}\sum_{i\in S}\hat{\zeta}_t^F\hat{\zeta}_t^{F'}\tilde{e}_{it}^2\right) - H'_{\zeta^F}\left(\sum_{t=p+1}^{T}\sum_{i\in S}\zeta_t^F\zeta_t^{F'}e_{it}^2\right)H_{\zeta^F}\right] \\
= & \frac{1}{N_S T}\sum_{t=p+1}^{T}\sum_{i\in S}(\hat{\zeta}_t^F\hat{\zeta}_t^{F'} - H'_{\zeta^F}\zeta_t^F\zeta_t^{F'}H_{\zeta^F})\tilde{e}_{it}^2 + \frac{1}{N_S T}\sum_{t=p+1}^{T}\sum_{i\in S}H'_{\zeta^F}\zeta_t^F\zeta_t^{F'}H_{\zeta^F}(\tilde{e}_{it}^2 - e_{it}^2)
\end{aligned}$$

The first term is bounded by

$$\begin{aligned}
& \left(\frac{1}{T-p}\sum_{t=p+1}^{T}\|\hat{\zeta}_t^F\hat{\zeta}_t^{F'} - H'_{\zeta^F}\zeta_t^F\zeta_t^{F'}H_{\zeta^F}\|^2\frac{1}{T-p}\sum_{t=p+1}^{T}\left(\frac{1}{N_S}\sum_{i\in S}\tilde{e}_{it}^2\right)^2\right)^{\frac{1}{2}} \\
\leq & \; o_p(1)\left[\frac{1}{(T-p)N_S}\sum_{t=p+1}^{T}\sum_{i\in S}\tilde{e}_{it}^4\right]^{\frac{1}{2}} = o_p(1)O_p(1)
\end{aligned}$$

where the $o_p(1)$ terms follows from the proof in part (i) of this lemma and $(T-p)^{-1}N_S^{-1}\sum_{t=p+1}^{T}\sum_{i\in S}\tilde{e}_{it}^4 = O_p(1)$ follows from Equation (B.5). The second term is bounded by

$$\begin{aligned}
& \left\{\frac{1}{T}\sum_{t=p+1}^{T}\|H'_{\zeta^F}\zeta_t^F\zeta_t^{F'}H_{\zeta^F}\|^2\frac{1}{T}\sum_{t=p+1}^{T}\left[\frac{1}{N_S}\sum_{i\in S}(\tilde{e}_{it}^2 - e_{it}^2)\right]^2\right\}^{\frac{1}{2}} \\
\leq & \; O_p(1)\left\{\frac{2}{T}\sum_{t=p+1}^{T}\left[\frac{1}{N_S}\sum_{i\in S}\tilde{e}_{it}(\tilde{e}_{it}-e_{it})\right]^2 + \frac{2}{T}\sum_{t=p+1}^{T}\left[\frac{1}{N_S}\sum_{i\in S}e_{it}(\tilde{e}_{it}-e_{it})\right]^2\right\}^{\frac{1}{2}} \\
\leq & \; O_p(1)\left[\frac{2}{TN_S}\sum_{t=p+1}^{T}\sum_{i\in S}\tilde{e}_{it}^2(\tilde{e}_{it}-e_{it})^2 + \frac{2}{TN_S}\sum_{t=p+1}^{T}\sum_{i\in S}e_{it}^2(\tilde{e}_{it}-e_{it})^2\right]^{\frac{1}{2}} \\
\leq & \; O_p(1)\left\{\left[\frac{2}{TN_S}\sum_{t=p+1}^{T}\sum_{i\in S}\tilde{e}_{it}^4\frac{2}{TN_S}\sum_{t=p+1}^{T}\sum_{i\in S}(\tilde{e}_{it}-e_{it})^4\right]^{\frac{1}{2}} + \left[\frac{2}{TN_S}\sum_{t=p+1}^{T}\sum_{i\in S}e_{it}^4\frac{2}{TN_S}\sum_{t=p+1}^{T}\sum_{i\in S}(\tilde{e}_{it}-e_{it})^4\right]^{\frac{1}{2}}\right\}^{\frac{1}{2}} \\
= & \; o_p(1)
\end{aligned}$$

where $(N_S T)^{-1}\sum_{t=p+1}^{T}\sum_{i\in S}\tilde{e}_{it}^4 = O_p(1)$ follows from Equation (B.5), $(N_S T)^{-1}\sum_{t=p+1}^{T}\sum_{i\in S}e_{it}^4 = O_p(1)$ follows from Assumption 3(a), and $(N_S T)^{-1}\sum_{t=p+1}^{T}\sum_{i\in S}(\tilde{e}_{it}-e_{it})^4 = o_p(1)$ follows from Lemma B2(ii). Q.E.D.

Proof of Theorem 1:

We will first show that H_{ζ^F} is non-singular as N and $T \to \infty$. To see this, consider the bound of $T^{-1}\|\hat{\zeta}^{F'}\hat{\zeta}^F - H'_{\zeta^F}\zeta^{F'}\zeta^F H_{\zeta^F}\|$

$$\frac{1}{T}\|\hat{\zeta}^{F'}\hat{\zeta}^F - H'_{\zeta^F}\zeta^{F'}\zeta^F H_{\zeta^F}\| \leq \frac{1}{T}\|\hat{\zeta}^{F'}(\hat{\zeta}^F - \zeta^F H_{\zeta^F})\| + \frac{1}{T}\|(\hat{\zeta}^F - \zeta^F H_{\zeta^F})'\zeta^F H_{\zeta^F}\| = O_p(\delta_{NT}^{-1})$$

by Lemma A9 and Cauchy-Schwarz inequality. Since $\hat{\zeta}^{F'}\hat{\zeta}^F/(T-p) = I_{q_F}$ and $\zeta^{F'}\zeta^F/T \to_p I_{q_F}$ by Assumption 1(b), it must follow that H_{ζ^F} is non-singular as N and $T \to \infty$, otherwise $(\hat{\zeta}^{F'}\hat{\zeta}^F - H'_{\zeta^F}\zeta^{F'}\zeta^F H_{\zeta^F})/T$ cannot converge to zero.

(i) Define $w_i^* \equiv e_i'\zeta^F(\sum_{t=p+1}^T \zeta_t^F\zeta_t^{F'}e_{it}^2)^{-1}\zeta^{F'}e_i$, and it can be rewritten as

$$w_i^* \quad = \quad \frac{e_i'\zeta^F H_{\zeta^F}}{\sqrt{T-p}}\left(\frac{1}{T-p}\sum_{t=p+1}^T H'_{\zeta^F}\zeta_t^F\zeta_t^{F'}H_{\zeta^F}e_{it}^2\right)^{-1}\frac{H'_{\zeta^F}\zeta^{F'}e_i}{\sqrt{T-p}} \to_d \chi^2_{q_F}$$

by Assumption 8(a). Note that Lemma B1(i) and the condition that $\sqrt{T}/N \to 0$ imply that

$$\frac{1}{\sqrt{T-p}}(H'_{\zeta^F}\zeta^{F'}e_i - \hat{\zeta}^{F'}\tilde{X}_i) = O_p\left(\frac{\sqrt{T}}{\delta_{NT}^2}\right) = o_p(1) \tag{B.6}$$

Since $(T-p)^{-1}\sum_{t=p+1}^T \zeta_t^F\zeta_t^{F'}e_{it}^2 \to_p$ a positive definite matrix Ω_i and H_{ζ^F} is non-singular asymptotically, $\|[\sum_{t=p+1}^T H'_{\zeta^F}\zeta_t^F\zeta_t^{F'}H_{\zeta^F}e_{it}^2/(T-p)]^{-1}\| = O_p(1)$. Also, we have $(T-p)^{-1}\left(\sum_{t=p+1}^T \hat{\zeta}_t^F\hat{\zeta}_t^{F'}\tilde{e}_{it}^2 - \sum_{t=p+1}^T H'_{\zeta^F}\zeta_t^F\zeta_t^{F'}H_{\zeta^F}e_{it}^2\right) = o_p(1)$ by Lemma B3(ii), so $(T-p)^{-1}\sum_{t=p+1}^T \hat{\zeta}_t^F\hat{\zeta}_t^{F'}\tilde{e}_{it}^2$ is positive definite asymptotically and $\|[\sum_{t=p+1}^T \hat{\zeta}_t^F\hat{\zeta}_t^{F'}\tilde{e}_{it}^2/(T-p)]^{-1}\| = O_p(1)$. Thus,

$$\begin{aligned}
&\left\|\left(\frac{1}{T-p}\sum_{t=p+1}^T \hat{\zeta}_t^F\hat{\zeta}_t^{F'}\tilde{e}_{it}^2\right)^{-1} - \left(\frac{1}{T-p}\sum_{t=p+1}^T H'_{\zeta^F}\zeta_t^F\zeta_t^{F'}H_{\zeta^F}e_{it}^2\right)^{-1}\right\| \\
\leq\ &\left\|\left(\frac{1}{T-p}\sum_{t=p+1}^T H'_{\zeta^F}\zeta_t^F\zeta_t^{F'}H_{\zeta^F}e_{it}^2\right)^{-1}\right\|\left\|\frac{1}{T-p}\sum_{t=p+1}^T\left(H'_{\zeta^F}\zeta_t^F\zeta_t^{F'}H_{\zeta^F}e_{it}^2 - \hat{\zeta}_t^F\hat{\zeta}_t^{F'}\tilde{e}_{it}^2\right)\right\|\left\|\left(\frac{1}{T-p}\sum_{t=p+1}^T \hat{\zeta}_t^F\hat{\zeta}_t^{F'}\tilde{e}_{it}^2\right)^{-1}\right\| \\
=\ &O_p(1)o_p(1)O_p(1)
\end{aligned} \tag{B.7}$$

Combining Equations (B.6) and (B.7), we have $w_i - w_i^* = o_p(1)$ and $w_i \to_d \chi^2_{q_F}$ for $i \in S$.

(ii) Define $W^* \equiv \sum_{i\in S}e_i'\zeta^F(\sum_{t=p+1}^T\sum_{i\in S}\zeta_t^F\zeta_t^{F'}e_{it}^2)^{-1}\zeta^{F'}\sum_{i\in S}e_i$, and it can be rewritten as

$$W^* \equiv \frac{\sum_{i\in S}e_i'\zeta^F H_{\zeta^F}}{\sqrt{(T-p)N_S}}\left(\frac{\sum_{t=p+1}^T\sum_{i\in S}H'_{\zeta^F}\zeta_t^F\zeta_t^{F'}H_{\zeta^F}e_{it}^2}{(T-p)N_S}\right)^{-1}\frac{H'_{\zeta^F}\zeta^{F'}\sum_{i\in S}e_i}{\sqrt{(T-p)N_S}} \to_d \chi^2_{q_F}$$

by Assumption 8(b). Note that Lemma B1(ii) and the condition that $\sqrt{T}/N \to 0$ imply that

$$\frac{1}{\sqrt{(T-p)N_S}}(\sum_{i\in S}H'_{\zeta^F}\zeta^{F'}e_i - \sum_{i\in S}\hat{\zeta}^{F'}\tilde{X}_i) = O_p\left(\frac{\sqrt{T}}{\delta_{NT}^2}\right) = o_p(1) \tag{B.8}$$

Also, since $(T-p)^{-1}N_S^{-1}\sum_{t=p+1}^T\sum_{i\in S}\zeta_t^F\zeta_t^{F'}e_{it}^2 \to_p \Omega$, H_{ζ^F} is non-singular asymptotically, and $(T-p)^{-1}N_S^{-1}\sum_{t=p+1}^T\sum_{i\in S}\left(\hat{\zeta}_t^F\hat{\zeta}_t^{F'}\tilde{e}_{it}^2 - H'_{\zeta^F}\zeta_t^F\zeta_t^{F'}H_{\zeta^F}e_{it}^2\right) \to_p 0$ by Lemma B3(ii), we have

$$\left(\frac{1}{(T-p)N_S}\sum_{t=p+1}^T\sum_{i\in S}\hat{\zeta}_t^F\hat{\zeta}_t^{F'}\tilde{e}_{it}^2\right)^{-1} - \left(\frac{1}{(T-p)N_S}\sum_{t=p+1}^T\sum_{i\in S}H'_{\zeta^F}\zeta_t^F\zeta_t^{F'}H_{\zeta^F}e_{it}^2\right)^{-1} = o_p(1) \tag{B.9}$$

by the same argument for proving Equation (B.7). Combining Equations (B.8) and (B.9), we have $W - W^* = o_p(1)$ and $W \to_d \ \chi^2_{q_F}$. Q.E.D.

C Proof of Theorem 2

Let us re-consider $\hat{X}^S$:

$$\hat{X}^S = M_{\hat{\mathcal{F}}} X^S = \zeta^S \Gamma^{SS'} + e^S - P_{\hat{\mathcal{F}}}(\zeta^S \Gamma^{SS'} + e^S) + M_{\hat{\mathcal{F}}} \mathcal{F} \Pi^{S'} \tag{C.1}$$

Recall that the estimator $\hat{\zeta}^S$ is equal to $\sqrt{T-p}$ times the eigenvectors corresponding to the $\kappa_S = q - \kappa_F$ largest eigenvalues of the $(T-p) \times (T-p)$ matrix $\hat{X}^S \hat{X}^{S'}$. Under the alternative hypothesis, $\kappa_S < q_S$, $\hat{\zeta}^S$is $(T-p) \times \kappa_S$, so it can only span the subspace of ζ^S, which is $(T-p) \times q_S$. Let $\bar{\zeta}^S$ equal $\sqrt{T-p}$ times the eigenvectors corresponding to the q_S largest eigenvalues of $\hat{X}^S \hat{X}^{S'}$ and $\bar{V}_S$ be the diagonal matrix consisting of the first q^S largest eigenvalues of $(1/N^S T)\hat{X}^S \hat{X}^{S'}$ in decreasing order. Hence, we have

$$(1/N^S T)\hat{X}^S \hat{X}^{S'} \bar{\zeta}^S = \bar{\zeta}^S \bar{V}_S \tag{C.2}$$

Recall that $\bar{H}_{\zeta^S} = (\Gamma^{SS'}\Gamma^{SS}/N_S)(\zeta^{S'}\bar{\zeta}^S/T)\bar{V}_S^{-1}$. Similar to H_{ζ^S}, $\bar{H}_{\zeta^S}$ is also $O_p(1)$. Since $\bar{\zeta}^S$ is $(T-p) \times q_S$, Lemmas A5 and A8 in Appendix A.2 are directly applicable to $\bar{\zeta}^S$, which is summarized by the following Lemma.

Lemma C1: Under Assumptions 1 - 6,

(i) $T^{-1} \sum_{t=p+1}^{T} \|\bar{\zeta}^S_t - \bar{H}'_{\zeta^S}\zeta^S_t\|^2 = O_p(\delta^{-2}_{NT})$.

(ii) $T^{-1}(\bar{\zeta}^S - \zeta^S \bar{H}_{\zeta^S})'\zeta = O_p(\delta^{-2}_{NT})$.

(iii) $\bar{H}_{\zeta^S}$ is non-singular as N and $T \to \infty$.

Proof:

The same as the proof of Lemmas A5 and A8. Q.E.D.

By the definition of $\hat{\zeta}^S$, $\hat{\zeta}^S$ is equal to the first κ_S columns of $\bar{\zeta}^S$. Let $\bar{\zeta}^S$ be partitioned as $\bar{\zeta}^S = [\bar{\zeta}^S_{1:\kappa_S} \vdots \bar{\zeta}^S_{\kappa_S+1:q_S}]$, where $\bar{\zeta}^S_{1:\kappa_S}$ is a $(T-p) \times \kappa_S$ matrix consisting of the first κ_S columns of $\bar{\zeta}^S$ and $\bar{\zeta}^S_{\kappa_S+1:q_S}$ is a $(T-p) \times \kappa_F$ matrix consisting of the last $\kappa_F = q_S - \kappa_S$ columns of $\bar{\zeta}^S$. Thus, we have $\hat{\zeta}^S = \bar{\zeta}^S_{1:\kappa_S}$. Define $\bar{H}^S_{1:\kappa_S} = (\Gamma^{SS'}\Gamma^{SS}/N_S)(\zeta^{S'}\bar{\zeta}^S_{1:\kappa_S}/T)\bar{V}^{-1}_{S,1:\kappa_S}$ and $\bar{H}^S_{\kappa_S+1:q_S} = (\Gamma^{SS'}\Gamma^{SS}/N_S)(\zeta^{S'}\bar{\zeta}^S_{\kappa_S+1:q_S}/T)\bar{V}^{-1}_{S,\kappa_S+1:q_S}$, where $\bar{V}_{S,1:\kappa_S}$ is a diagonal matrix consisting of the first κ^S largest eigenvalues of $(1/N^S T)\hat{X}^S \hat{X}^{S'}$ in descending order, and $\bar{V}_{S,\kappa_S+1:q_S}$ is a is a diagonal matrix consisting of the $(\kappa^S+1)^{th}$ through the q_S largest eigenvalues of $(1/N^S T)\hat{X}^S \hat{X}^{S'}$ in descending order. Hence, $\bar{H}_{\zeta^S} = [\bar{H}^S_{1:\kappa_S} \vdots \bar{H}^S_{\kappa_S+1:q_S}]$ and $\bar{V}_S = \begin{bmatrix} \bar{V}_{S,1:\kappa_S} & 0 \\ 0 & \bar{V}_{S,\kappa_S+1:q_S} \end{bmatrix}$.

Now, re-write Equation (A.7) as

$$\begin{aligned}
\tilde{X} = M_{\hat{\zeta}^S} M_{\hat{\mathcal{F}}} X &= M_{\hat{\zeta}^S} M_{\hat{\mathcal{F}}} (\mathcal{F}\Pi' + \zeta^S \Gamma^{S'} + \zeta^F \Gamma^{F'} + e) \\
&= M_{\hat{\zeta}^S} M_{\hat{\mathcal{F}}} (\mathcal{F}\Pi' + \zeta^S \bar{H}_{\zeta^S} \bar{H}^{-1}_{\zeta^S} \Gamma^{S'} + \zeta^F \Gamma^{F'} + e) \\
&= M_{\hat{\zeta}^S} M_{\hat{\mathcal{F}}} (\mathcal{F}\Pi' + [\zeta^S \bar{H}^S_{1:\kappa_S} \vdots \zeta^S \bar{H}^S_{\kappa_S+1:q_S}] \Xi^{S'} + \zeta^F \Gamma^{F'} + e)
\end{aligned} \tag{C.3}$$

where we use the definition that $\Xi^{S'} \equiv \bar{H}_{\zeta^S}^{-1}\Gamma^{S'}$. Recall that $\Xi^S = [\Xi_{1:\kappa_S}^S \vdots \Xi_{\kappa_S+1:q_S}^S]$, where $\Xi_{1:\kappa_S}^S$ is the first κ_S columns of Ξ^S and $\Xi_{\kappa_S+1:q_S}^S$ is the last $q_S - \kappa_S$ columns of Ξ^S. Let $\upsilon^S \equiv \zeta^S \bar{H}_{1:\kappa_S}^S$, $\upsilon^F = [\zeta^S \bar{H}_{\kappa_S+1:q_S}^S \vdots \zeta^F]$, $\Psi^S = \Xi_{1:\kappa_S}^S$, $\Psi^F = [\Xi_{\kappa_S+1:q_S}^S \vdots \Gamma^F]$, and $\Psi = [\Psi^S \vdots \Psi^F]$, so Equation (C.4) reduces to

$$\tilde{X} = M_{\hat{\zeta}^S} M_{\hat{\mathcal{F}}}(\mathcal{F}\Pi' + \upsilon^S \Psi^{S'} + \upsilon^F \Psi^{F'} + e) \tag{C.4}$$

Note that $\hat{\zeta}^S$ is an estimate for $\upsilon^S = \zeta^S \bar{H}_{1:\kappa_S}^S$ and that

$$T^{-1}\sum_{t=p+1}^{T} \|\hat{\zeta}_t^S - \upsilon_t^S\|^2 = T^{-1}\sum_{t=p+1}^{T} \|\hat{\zeta}_t^S - \bar{H}_{1:\kappa_S}^{S'}\zeta_t^S\|^2 \leq T^{-1}\sum_{t=p+1}^{T} \|\bar{\zeta}_t^S - \bar{H}_{\zeta^S}'\zeta_t^S\|^2 = O_p(\delta_{NT}^{-2}) \tag{C.5}$$

by Lemma C1(i) and the fact that $\bar{\zeta}_{1:\kappa_S}^S \equiv \hat{\zeta}^S$. Let $\hat{\mathcal{G}} = [\hat{\mathcal{F}} \vdots \hat{\zeta}^S]$, $\mathcal{G} = [\mathcal{F} \vdots \upsilon^S]$, and $\mathcal{L} = [\Pi \vdots \Psi^S]$. Note that $\hat{\mathcal{F}}'\hat{\zeta}^S = 0$, so $M_{\hat{\zeta}^S} M_{\hat{\mathcal{F}}} = M_{\hat{\mathcal{G}}}$. Hence, $\tilde{X}$ can be rewritten as

$$\begin{aligned}\tilde{X} &= M_{\hat{\zeta}^S} M_{\hat{\mathcal{F}}}(\mathcal{F}\Pi' + \upsilon^S \Psi^{S'} + \upsilon^F \Psi^{F'} + e) \\ &= \upsilon^F \Psi^{F'} + e - P_{\hat{\mathcal{G}}}(\upsilon^F \Psi^{F'} + e) + M_{\hat{\mathcal{G}}}\mathcal{G}\mathcal{L}'\end{aligned} \tag{C.6}$$

Lemma C2: Under Assumption 2, $\Psi^{F'}\Psi^F/N \to_p$ a $\kappa_F \times \kappa_F$ positive definite matrix.

Proof:

Note that $\Psi = [\Psi^S \vdots \Psi^F] = [\Gamma^S \bar{H}_{\zeta^S}^{-1'} \vdots \Gamma^F] = \Gamma D = \Lambda G A^{-1} D$, where $D = \begin{bmatrix} \bar{H}_{\zeta^S}^{-1'} & 0 \\ 0 & I_{q_F} \end{bmatrix}$ is a non-singular matrix asymptotically. Hence,

$$N^{-1}\Psi'\Psi = N^{-1}D'A^{-1'}G'\Lambda'\Lambda G A^{-1} D$$

Since $\Lambda'\Lambda/N \to_p$ a $r \times r$ positive definite matrix Σ_Λ, rank$(G) = q$, and A is non-singular, $N^{-1}\Psi'\Psi$ is positive definite asymptotically. Since $\Psi^{F'}\Psi^F/N$ is a principal submatrix of $N^{-1}\Psi'\Psi$, it must be positive definite asymptotically. Q.E.D.

Lemma C3: Under Assumptions 1 - 6,

(i) Define $H_{\mathcal{G}} = \begin{bmatrix} H_{\mathcal{F}} & 0 \\ 0 & I_{\kappa_S} \end{bmatrix}$. $H_{\mathcal{G}}$ is non-singular and $T^{-1}\sum_{t=p+1}^{T} \|\hat{\mathcal{G}}_t - H_{\mathcal{G}}'\mathcal{G}_t\|^2 = O_p(\delta_{NT}^{-2})$, where $\hat{\mathcal{G}}_t$ and $\mathcal{G}_t$ are transposes of the t^{th} row of $\hat{\mathcal{G}}$ and $\mathcal{G}$, respectively.

(ii) $T^{-1}\upsilon^{S'}\upsilon^F = O_p(\delta_{NT}^{-2}) + O_p(T^{-1/2})$.

Proof:

(i) $H_{\mathcal{G}}$ is non-singular because $H_{\mathcal{F}}$ is non-singular by Lemma A2(ii). Also, note that $T^{-1}\sum_{t=p+1}^{T} \|\hat{\mathcal{F}}_t - H_{\mathcal{F}}'\mathcal{F}_t\|^2 = O_p(\delta_{NT}^{-2})$ by Lemma A2(i) and that $T^{-1}\sum_{t=p+1}^{T} \|\hat{\zeta}_t^S - \upsilon_t^S\|^2 = O_p(\delta_{NT}^{-2})$ by Lemma C1 (see Equation (C.5)). Since $\hat{\mathcal{G}} = [\hat{\mathcal{F}} \vdots \hat{\zeta}^S]$ and $\mathcal{G} = [\mathcal{F} \vdots \upsilon^S]$, it follows that $T^{-1}\sum_{t=p+1}^{T} \|\hat{\mathcal{G}}_t - H_{\mathcal{G}}'\mathcal{G}_t\|^2 = O_p(\delta_{NT}^{-2})$.

(ii) Since $\upsilon^S \equiv \zeta^S \bar{H}_{1:\kappa_S}^S$ and $\upsilon^F = [\zeta^S \bar{H}_{\kappa_S+1:q_S}^S \vdots \zeta^F]$.

$$\|T^{-1}\upsilon^{S'}\upsilon^F\| \leq \|T^{-1}\bar{H}_{1:\kappa_S}^{S'}\zeta^{S'}\zeta^S \bar{H}_{\kappa_S+1:q_S}^S\| + \|T^{-1}\bar{H}_{1:\kappa_S}^{S'}\zeta^{S'}\zeta^F\|$$

where the second term is $O_p(T^{-1/2})$ by Assumption 1(c). For the first term, note that $T^{-1}\bar{\zeta}_{1:\kappa_S}^{S'}\bar{\zeta}_{\kappa_S+1:q_S}^S = 0$

by the orthogonality of eigenvectors, so we have

$$\begin{aligned}
& \|T^{-1}\bar{H}^{S'}_{1:\kappa_S}\zeta^{S'}\zeta^{S}\bar{H}^{S}_{\kappa_S+1:q_S}\| \\
\leq\ & \|T^{-1}(\zeta^{S}\bar{H}^{S}_{1:\kappa_S}-\bar{\zeta}^{S}_{1:\kappa_S})'\zeta^{S}\bar{H}^{S}_{\kappa_S+1:q_S}\| + \|T^{-1}\bar{\zeta}^{S'}_{1:\kappa_S}(\zeta^{S}\bar{H}^{S}_{\kappa_S+1:q_S}-\bar{\zeta}^{S}_{\kappa_S+1:q_S})\| + \|T^{-1}\bar{\zeta}^{S'}_{1:\kappa_S}\bar{\zeta}^{S}_{\kappa_S+1:q_S}\| \\
\leq\ & \|T^{-1}(\zeta^{S}\bar{H}^{S}-\bar{\zeta}^{S})'\zeta^{S}\bar{H}_{\zeta^S}\| + \|T^{-1}\bar{\zeta}^{S'}(\zeta^{S}\bar{H}_{\zeta^S}-\bar{\zeta}^{S})\| \\
\leq\ & 2\|T^{-1}(\zeta^{S}\bar{H}_{\zeta^S}-\bar{\zeta}^{S})'\zeta^{S}\bar{H}_{\zeta^S}\| + \|T^{-1}(\bar{\zeta}^{S}-\zeta^{S}\bar{H}_{\zeta^S})'(\zeta^{S}\bar{H}_{\zeta^S}-\bar{\zeta}^{S})\| \\
=\ & O_p(\delta_{NT}^{-2})
\end{aligned}$$

by Lemma C1. Q.E.D.

Now, let us consider the estimator $\hat{\zeta}^F$ under the alternative hypothesis. Recall that $\hat{\zeta}^F$ is equal to $\sqrt{T-p}$ times the eigenvectors corresponding to the κ_F largest eigenvalues of the $(T-p)\times(T-p)$ matrix $\tilde{X}\tilde{X}'$, so we have

$$(1/NT)\tilde{X}\tilde{X}'\hat{\zeta}^F = \hat{\zeta}^F\hat{U}_F \tag{C.7}$$

where $\hat{U}_F$ is the diagonal matrix consisting of the first κ_F largest eigenvalues of $(1/NT)\tilde{X}\tilde{X}'$ in decreasing order. Let U_F be the diagonal matrix consisting of the first κ_F largest eigenvalues of $(\upsilon^F\Psi^{F'}+e)(\upsilon^F\Psi^{F'}+e)'/NT$ in decreasing order. We have

$$\|\hat{U}_F^{-1}\| = O_p(1) \tag{C.8}$$

To see this, first note that $\upsilon^{F'}\upsilon^F/T$ has full rank asymptotically because $\upsilon^F = [\zeta^S\bar{H}^S_{\kappa_S+1:q_S}\vdots\zeta^F]$ is of full column rank (note that $\bar{H}_{\zeta^S}$ is non-singular asymptotically, so its submatrix$\bar{H}^S_{\kappa_S+1:q_S}$ has rank equal to $q_S-\kappa_S$). By Lemma C2, $\Psi^{F'}\Psi^F/N$ also has full rank asymptotically, so Lemma A1(ii) is applicable to U_F, i.e. U_F converges to a non-singular diagonal matrix consisting of the eigenvalues of $\lim_{N,T\to\infty}(\upsilon^{F'}\upsilon^F/T)(\Psi^{F'}\Psi^F/N)$. Additionally, the difference between $\tilde{X}$ and $\upsilon^F\Psi^{F'}+e$ is that $\tilde{X}$ contains some additional estimation error terms defined in Equation (C.6). Again, by Amengual and Watson's (2007) results, we know that such estimation errors will not affect the limits of the first κ_F largest eigenvalues of $\tilde{X}\tilde{X}'/NT$, i.e. $\hat{U}_F - U_F \to_p 0$. Hence, $\hat{U}_F$ is non-singular asymptotically and $\|\hat{U}_F^{-1}\| = O_p(1)$.

Now, substituting Equation (C.6) into Equation (C.7) gives

$$\frac{1}{NT}\left[\upsilon^F\Psi^{F'}+e-P_{\hat{\mathcal{G}}}(\upsilon^F\Psi^{F'}+e)+M_{\hat{\mathcal{G}}}\mathcal{G}\mathcal{L}'\right]\left[\upsilon^F\Psi^{F'}+e-P_{\hat{\mathcal{G}}}(\upsilon^F\Psi^{F'}+e)+M_{\hat{\mathcal{G}}}\mathcal{G}\mathcal{L}'\right]'\hat{\zeta}^F\hat{U}_F^{-1} = \hat{\zeta}^F$$

Note that $T^{-1}\sum_{t=p+1}^T\|\hat{\zeta}_t^F\|$ is $O_p(1)$ because $\hat{\zeta}^{F'}\hat{\zeta}^F/(T-p) = I_{q_F}$. Define $H_{\upsilon^F} = (\Psi^{F'}\Psi^F/N)(\upsilon^{F'}\hat{\zeta}^F/T)\hat{U}_F^{-1}$. We have $H_{\upsilon^F} = O_p(1)$ because $\|H_{\upsilon^F}\| \leq \|\hat{\zeta}^{F'}\hat{\zeta}^F/T\|^{\frac{1}{2}}\|\upsilon^{F'}\upsilon^F/T\|^{\frac{1}{2}}\|\Psi^{F'}\Psi^F/N\|\|\hat{U}_F^{-1}\| = O_p(1)$ by Lemma C2 and (C.8). Analogous to Lemmas A1(i), A5 and A9, the following lemma shows that $\hat{\zeta}_t^F$ is a consistent estimate for $H'_{\upsilon^F}\upsilon_t^F$ and that the convergence rate is the same as in Lemmas A1(i), A5, and A9.

Lemma C4: Under Assumptions 1 - 6, $T^{-1}\sum_{t=p+1}^T\|\hat{\zeta}_t^F - H'_{\upsilon^F}\upsilon_t^F\|^2 = O_p\left(\delta_{NT}^{-2}\right)$.

Proof:

Note that Equation (C.4) has almost the same structure as (A.7). The only difference is that Equation(A.7) deals with structural shock ζ, whereas Equation (C.4) deals with υ, which is a rotation of ζ. Note that Assumptions 1 - 6 are stated in terms of ζ and Γ. To prove Lemma C4, it is sufficient to show that these assumptions hold if ζ and Γ are replaced by υ and Ψ. Assumptions 1(a), 2(a), 2(b), 3, 4, and 5 are the same whether the model is expressed in terms of ζ or υ. Also, since $\upsilon = [\zeta^S \bar{H}_{\zeta^S} \vdots \zeta^F] = \zeta D^{-1'}$ and $\Psi = [\Gamma^S \bar{H}_{\zeta^S}^{-1'} \vdots \Gamma^F] = \Gamma D$, where $D = \begin{bmatrix} \bar{H}_{\zeta^S}^{-1'} & 0 \\ 0 & I_{q_F} \end{bmatrix}$ is a asymptotically non-singular by Lemma C1 and $D = O_p(1)$, the boundedness conditions in Assumption 6 still hold if we replace ζ_t with υ_t. For the same reason, Assumption 2(b) can be rewritten as "$E\|\upsilon_t\|^4 < M$, $E(\upsilon_s \upsilon_t') = 0$ for any $s \neq t$, and $T^{-1}\sum_{p+1}^T \upsilon_t \upsilon_t'$ is a $q \times q$ positive definite matrix as $T \to \infty$". Furthermore, the role of Assumption 1(c) for the proof of Lemma 9 is that $T^{-1}\zeta^{S'}\zeta^F$ can be bounded by $O(T^{-1/2}) \leq O_p(\delta_{NT}^{-1})$, but we have already proved that $T^{-1}\upsilon^{S'}\upsilon^F = O_p(\delta_{NT}^{-2}) + O_p(T^{-1/2}) \leq O_p(\delta_{NT}^{-1})$, which can serve to play the same role as Assumption 1(c). For Assumption 2(c), the role of $\Gamma^{FF'}\Gamma^{FF}/N_F \to_p \Sigma_{\Gamma^{FF}}$ is to ensure that $\Gamma^{F'}\Gamma^F/N$ is positive definite asymptotically, and Lemma C2 shows that $\Psi^{F'}\Psi^F/N \to_p$ a $\kappa_F \times \kappa_F$ positive definite matrix. Lastly, the analog of Γ^{SS} in (C.4) is the first N_S rows of $\Psi^S = \Xi^S_{1:\kappa_S}$ corresponding to the slow variables. Let Ψ^{SS} and Ξ^{SS} denote the first N_S rows of Ψ^S and Ξ^S, respectively. Since $\Xi^{S'} \equiv \bar{H}_{\zeta^S}^{-1}\Gamma^{S'}$ and $\Xi^{SS'} \equiv \bar{H}_{\zeta^S}^{-1}\Gamma^{SS'}$, $N_S^{-1}\Xi^{SS'}\Xi^{SS}$ is positive definite asymptotically by the assumption that $\Gamma^{SS'}\Gamma^{SS'}/N_S \to_p \Sigma_{\Gamma^{SS}}$ and Lemma C1(iii) that $\bar{H}_{\zeta^S}$ is non-singular. $N_S^{-1}\Psi^{SS'}\Psi^{SS}$ is a principal submatrix of $N_S^{-1}\Xi^{SS'}\Xi^{SS}$, so it is also positive definite asymptotically. Since all analogous assumptions in terms of υ and Ψ are satisfied, Lemma C4 holds by Lemma A9. Q.E.D.

Proof of Theorem 2:

We will first show that H_{υ^F} is non-singular as N and $T \to \infty$. To see this, consider the bound of $T^{-1}\|\hat{\zeta}^{F'}\hat{\zeta}^F - H_{\upsilon^F}'\upsilon^{F'}\upsilon^F H_{\upsilon^F}\|$

$$\frac{1}{T}\|\hat{\zeta}^{F'}\hat{\zeta}^F - H_{\upsilon^F}'\upsilon^{F'}\upsilon^F H_{\upsilon^F}\| \leq \frac{1}{T}\|\hat{\zeta}^{F'}(\hat{\zeta}^F - \upsilon^F H_{\upsilon^F})\| + \frac{1}{T}\|(\hat{\zeta}^F - \upsilon^F H_{\upsilon^F})'\upsilon^F H_{\upsilon^F}\| = O_p(\delta_{NT}^{-1})$$

by Lemma C4 and Cauchy-Schwarz inequality. Note that $\upsilon^{F'}\upsilon^F/T$ is positive definite asymptotically because it is a principal submatrix of $\upsilon'\upsilon/T = D^{-1}\zeta'\zeta D^{-1'}/T$, which is asymptotically positive definite by Assumption 1(b) and Lemma C1(iii). Also, since $\hat{\zeta}^{F'}\hat{\zeta}^F/(T-p) = I_{\kappa_F}$, it must follow that H_{υ^F} is non-singular as N and $T \to \infty$, otherwise $(\hat{\zeta}^{F'}\hat{\zeta}^F - H_{\upsilon^F}'\upsilon^{F'}\upsilon^F H_{\upsilon^F})/T$ cannot converge to zero.

Furthermore, the proof of Lemma C4 has shown that the model in Equation (C.4) satisfies Assumptions 1 - 6 with ζ and Γ replaced by υ and Ψ. Thus, we have the following result analogous to Lemma B3.

$$(T-p)^{-1}\left[\left(\sum_{t=p+1}^T \hat{\zeta}_t^F \hat{\zeta}_t^{F'} \tilde{e}_{it}^2\right) - H_{\upsilon^F}'\left(\sum_{t=p+1}^T \upsilon_t^F \upsilon_t^{F'} e_{it}^2\right) H_{\upsilon^F}\right] \to_p 0$$

$$(T-p)^{-1}N_S^{-1}\left[\left(\sum_{t=p+1}^T \sum_{i\in S} \hat{\zeta}_t^F \hat{\zeta}_t^{F'} \tilde{e}_{it}^2\right) - H_{\upsilon^F}'\left(\sum_{t=p+1}^T \sum_{i\in S} \upsilon_t^F \upsilon_t^{F'} e_{it}^2\right) H_{\upsilon^F}\right] \to_p 0$$

Note that $(T-p)^{-1}\sum_{t=p+1}^T \upsilon_t^F \upsilon_t^{F'} e_{it}^2$ and $[N_S(T-p)]^{-1}\sum_{t=p+1}^T \sum_{i\in S} \upsilon_t^F \upsilon_t^{F'} e_{it}^2$ are $O_p(1)$ by Assumptions 1 and 3, so $(T-p)^{-1}\sum_{t=p+1}^T \hat{\zeta}_t^F \hat{\zeta}_t^{F'} \tilde{e}_{it}^2$ and $[N_S(T-p)]^{-1}\sum_{t=p+1}^T \sum_{i\in S} \hat{\zeta}_t^F \hat{\zeta}_t^{F'} \tilde{e}_{it}^2$ are also $O_p(1)$.

(i) We will show that $w_i = \tilde{X}_i'\hat{\zeta}^F(\sum_{t=p+1}^T \hat{\zeta}_t^F \hat{\zeta}_t^{F'} \tilde{e}_{it}^2)^{-1}\hat{\zeta}^{F'}\tilde{X}_i \to \infty$ as N and $T \to \infty$ for $i \in S$. First, consider $\hat{\zeta}^{F\prime}\tilde{X}_i/(T-p)$, which can be rewritten as

$$\begin{aligned}\frac{1}{T-p}\hat{\zeta}^{F\prime}\tilde{X}_i &= \frac{1}{T-p}\hat{\zeta}^{F\prime}[\upsilon^F\psi_i^F + e_i - P_{\hat{\mathcal{G}}}(\upsilon^F\psi_i^F + e_i) + M_{\hat{\mathcal{G}}}\mathcal{G}\mathcal{L}_i] \\ &= \frac{1}{T-p}\hat{\zeta}^{F\prime}(\upsilon^F\psi_i^F + e_i + \mathcal{G}\mathcal{L}_i)\end{aligned}$$

where ψ_i^F and $\mathcal{L}_i$ are the transposes of the i^{th} rows of Ψ^F and $\mathcal{L}$, respectively, and we use the fact that $\hat{\zeta}^{F\prime}\hat{\mathcal{G}} = \hat{\zeta}^{F\prime}[\hat{\mathcal{F}}\vdots\hat{\zeta}^S] = 0$. Note that $\|\hat{\zeta}^{F\prime}e_i/(T-p)\| \leq \|(\hat{\zeta}^F - \upsilon^F H_{\upsilon^F})'e_i/(T-p)\| + \|H'_{\upsilon^F}\upsilon^{F\prime}e_i/(T-p)\| = O_p(\delta_{NT}^{-1}) + O_p(T^{-1/2})$ by Lemma C4, Cauchy-Schwarz inequality and Assumptions 3(a) and 6(e). Additionally, by the fact that $\hat{\zeta}^{F\prime}\hat{\mathcal{G}} = 0$, we have

$$\begin{aligned}\left\|\frac{1}{T-p}\hat{\zeta}^{F\prime}\mathcal{G}\mathcal{L}_i\right\| &= \left\|\frac{1}{T-p}\hat{\zeta}^{F\prime}(\mathcal{G}H_{\mathcal{G}} - \hat{\mathcal{G}})H_{\mathcal{G}}^{-1}\mathcal{L}_i\right\| \\ &\leq \left\|\frac{1}{T-p}(\hat{\zeta}^F - \upsilon^F H_{\upsilon^F})'(\mathcal{G}H_{\mathcal{G}} - \hat{\mathcal{G}})H_{\mathcal{G}}^{-1}\mathcal{L}_i\right\| + \|H_{\upsilon^F}\|\left\|\frac{1}{T-p}\upsilon^{F\prime}(\mathcal{G}H_{\mathcal{G}} - \hat{\mathcal{G}})H_{\mathcal{G}}^{-1}\mathcal{L}_i\right\| \\ &\leq \left\|\frac{1}{T-p}(\hat{\zeta}^F - \upsilon^F H_{\upsilon^F})'(\mathcal{G}H_{\mathcal{G}} - \hat{\mathcal{G}})H_{\mathcal{G}}^{-1}\mathcal{L}_i\right\| + \|H_{\upsilon^F}\|\left\|\frac{1}{T-p}D^{-1}\zeta'(\mathcal{G}H_{\mathcal{G}} - \hat{\mathcal{G}})H_{\mathcal{G}}^{-1}\mathcal{L}_i\right\| \\ &= O_p\left(\frac{1}{\delta_{NT}^2}\right)\end{aligned}$$

by Lemmas C3(i), C4, A4(i) and C1(ii). Lastly, we consider $\hat{\zeta}^{F\prime}\upsilon^F\psi_i^F/(T-p)$. Recall that $\Xi^{S\prime} \equiv \bar{H}_{\zeta^S}^{-1}\Gamma^{S\prime}$, $\Xi^S = [\Xi_{1:\kappa_S}^S \vdots \Xi_{\kappa_S+1:q_S}^S]$, $\Psi^F = [\Xi_{\kappa_S+1:q_S}^S \vdots \Gamma^F]$, and $\upsilon^F = [\zeta^S\bar{H}_{\kappa_S+1:q_S}^S \vdots \zeta^F]$. Since $\Gamma^F = [\Gamma^{SF\prime} \vdots \Gamma^{FF\prime}]' = [0_{q_F\times N_S} \vdots \Gamma^{FF\prime}]'$, the last q_F entries in ψ_i^F are zeros for $i \in S$ and $\upsilon^F\psi_i^F = \zeta^S\bar{H}_{\kappa_S+1:q_S}^S\xi_{i,\kappa_S+1:q_S}^S$, where $\xi_{i,\kappa_S+1:q_S}^S$ is the transpose of the i^{th} row of $\Xi_{\kappa_S+1:q_S}^S$. Hence, we have

$$\begin{aligned}&\frac{1}{T-p}\hat{\zeta}^{F\prime}\upsilon^F\psi_i^F \\ =\;& \frac{1}{T-p}\hat{\zeta}^{F\prime}\zeta^S\bar{H}_{\kappa_S+1:q_S}^S\xi_{i,\kappa_S+1:q_S}^S \\ =\;& \left[\frac{1}{T-p}H'_{\upsilon^F}\upsilon^{F\prime}\zeta^S + \frac{1}{T-p}(\hat{\zeta}^F - \upsilon^F H_{\upsilon^F})'\zeta^S\right]\bar{H}_{\kappa_S+1:q_S}^S\xi_{i,\kappa_S+1:q_S}^S \\ =\;& \left[\frac{1}{T-p}H'_{\upsilon^F}[\zeta^S\bar{H}_{\kappa_S+1:q_S}^S \vdots \zeta^F]'\zeta^S + \frac{1}{T-p}(\hat{\zeta}^F - \upsilon^F H_{\upsilon^F})'\zeta^S\right]\bar{H}_{\kappa_S+1:q_S}^S\xi_{i,\kappa_S+1:q_S}^S \\ =\;& \left\{H'_{\upsilon^F}\begin{bmatrix}\left(\bar{H}_{\kappa_S+1:q_S}^{S\prime}\frac{\zeta^{S\prime}\zeta^S}{T-p}\right) \\ \left(\frac{1}{T-p}\zeta^{F\prime}\zeta^S\right)\end{bmatrix} + \frac{1}{T-p}(\hat{\zeta}^F - \upsilon^F H_{\upsilon^F})'\zeta^S\right\}\bar{H}_{\kappa_S+1:q_S}^S\xi_{i,\kappa_S+1:q_S}^S \\ =\;& H'_{\upsilon^F}\begin{bmatrix}\bar{H}_{\kappa_S+1:q_S}^{S\prime}\frac{\zeta^{S\prime}\zeta^S}{T-p}\bar{H}_{\kappa_S+1:q_S}^S \\ O_p(1/\sqrt{T})\end{bmatrix}\xi_{i,\kappa_S+1:q_S}^S + O_p\left(\frac{1}{\delta_{NT}}\right) \qquad \text{(C.9)}\end{aligned}$$

where the $O_p(1/\sqrt{T})$ term follows from Assumption 1(c) and the $o_p(\delta_{NT}^{-1})$ follows from Lemma C4 and Cauchy-Schwarz inequality. Note that $\bar{H}_{\kappa_S+1:q_S}^S$ is off full column rank by the fact that $\bar{H}_{\zeta^S}$ is non-singular, so $\bar{H}_{\kappa_S+1:q_S}^{S\prime}\frac{\zeta^{S\prime}\zeta^S}{T-p}\bar{H}_{\kappa_S+1:q_S}^S$ is of full rank asymptotically. Since $\xi_{i,\kappa_S+1:q_S}^S$ is assumed to be non-zero and H_{υ^F} is of full rank, Equation (C.9) bounded away from zero asymptotically. Also, we have proved that

$\hat{\zeta}^{F'} e_i/(T-p) = o_p(1)$ and $\hat{\zeta}^{F'} \mathcal{G}\mathcal{L}_i/(T-p) = o_p(1)$, so $\hat{\zeta}^{F'} \tilde{X}_i/(T-p)$ is bounded away from zero asymptotically. Thus, we have the desired result:

$$\text{plim}_{N,T\to\infty} \inf \left(\frac{w_i}{T-p} \right) = \text{plim}_{N,T\to\infty} \inf \left[\frac{\tilde{X}_i' \hat{\zeta}^F}{T-p} \left(\frac{\sum_{t=p+1}^{T} \hat{\zeta}_t^F \hat{\zeta}_t^{F'} \tilde{e}_{it}^2}{T-p} \right)^{-1} \frac{\hat{\zeta}^{F'} \tilde{X}_i}{T-p} \right] > 0$$

(ii) We will show that

$$W = \left(\sum_{i\in S} \tilde{X}_i' \hat{\zeta}^F \right) \left(\sum_{t=p+1}^{T} \sum_{i\in S} \hat{\zeta}_t^F \hat{\zeta}_t^{F'} \tilde{e}_{it}^2 \right)^{-1} \left(\hat{\zeta}^{F'} \sum_{i\in S} \tilde{X}_i \right) \to \infty$$

as N and $T \to \infty$. First, using Equation (C.6), $\hat{\zeta}^{F'} \sum_{i\in S} \tilde{X}_i/[\sqrt{N_S}(T-p)]$ can be rewritten as

$$\begin{aligned}
\frac{1}{\sqrt{N_S}(T-p)} \hat{\zeta}^{F'} \sum_{i\in S} \tilde{X}_i &= \frac{1}{\sqrt{N_S}(T-p)} \hat{\zeta}^{F'} \sum_{i\in S} [\upsilon^F \psi_i^F + e_i - P_{\hat{\mathcal{G}}}(\upsilon^F \psi_i^F + e_i) + M_{\hat{\mathcal{G}}} \mathcal{G}\mathcal{L}_i] \\
&= \frac{1}{\sqrt{N_S}(T-p)} \hat{\zeta}^{F'} \sum_{i\in S} (\upsilon^F \psi_i^F + e_i + \mathcal{G}\mathcal{L}_i)
\end{aligned}$$

where ψ_i^F and $\mathcal{L}_i$ are the transposes of the i^{th} rows of Ψ^F and $\mathcal{L}$, respectively, and we use the fact that $\hat{\zeta}^{F'} \hat{\mathcal{G}} = \hat{\zeta}^{F'} [\hat{\mathcal{F}} \vdots \hat{\zeta}^S] = 0$. Note that

$$\begin{aligned}
& \left\| \frac{1}{\sqrt{N_S}(T-p)} \hat{\zeta}^{F'} \sum_{i\in S} e_i \right\| \\
\leq & \left\| \frac{1}{T-p} (\hat{\zeta}^F - \upsilon^F H_{\upsilon^F})' \frac{\sum_{i\in S} e_i}{\sqrt{N_S}} \right\| + \left\| \frac{1}{T-p} H_{\upsilon^F}' \upsilon^{F'} \frac{\sum_{i\in S} e_i}{\sqrt{N_S}} \right\| \\
\leq & \left[\frac{1}{T-p} \sum_{t=p+1}^{T} \|\hat{\zeta}_t^F - H_{\upsilon^F}' \upsilon_t^F\|^2 \frac{1}{T-p} \sum_{t=p+1}^{T} \left\| \frac{\sum_{i\in S} e_{it}}{\sqrt{N_S}} \right\|^2 \right]^{\frac{1}{2}} + \|H_{\upsilon^F}\| \left\| \frac{1}{T-p} D^{-1} \zeta' \frac{\sum_{i\in S} e_i}{\sqrt{N_S}} \right\| \\
= & O_p\left(\frac{1}{\delta_{NT}} \right) + O_p\left(\frac{1}{\sqrt{T}} \right) \qquad \text{(C.10)}
\end{aligned}$$

by Lemma C4 and Assumptions 3(f) and 6(f). Additionally, by the fact that $\hat{\zeta}^{F\prime}\hat{\mathcal{G}} = 0$, we have

$$
\begin{aligned}
& \left\| \frac{1}{\sqrt{N_S}(T-p)} \hat{\zeta}^{F\prime} \mathcal{G} \sum_{i \in S} \mathcal{L}_i \right\| \\
= & \left\| \frac{1}{T-p} \hat{\zeta}^{F\prime} (\mathcal{G} H_{\mathcal{G}} - \hat{\mathcal{G}}) H_{\mathcal{G}}^{-1} \frac{1}{\sqrt{N_S}} \sum_{i \in S} \mathcal{L}_i \right\| \\
\leq & \left\| \frac{1}{T-p} (\hat{\zeta}^F - \upsilon^F H_{\upsilon^F})' (\mathcal{G} H_{\mathcal{G}} - \hat{\mathcal{G}}) H_{\mathcal{G}}^{-1} \frac{1}{\sqrt{N_S}} \sum_{i \in S} \mathcal{L}_i \right\| + \|H_{\upsilon^F}\| \left\| \frac{1}{T-p} \upsilon^{F\prime} (\mathcal{G} H_{\mathcal{G}} - \hat{\mathcal{G}}) H_{\mathcal{G}}^{-1} \frac{1}{\sqrt{N_S}} \sum_{i \in S} \mathcal{L}_i \right\| \\
\leq & \left\| \frac{1}{T-p} (\hat{\zeta}^F - \upsilon^F H_{\upsilon^F})' (\mathcal{G} H_{\mathcal{G}} - \hat{\mathcal{G}}) H_{\mathcal{G}}^{-1} \frac{1}{\sqrt{N_S}} \sum_{i \in S} \mathcal{L}_i \right\| + \|H_{\upsilon^F}\| \left\| \frac{1}{T-p} D^{-1} \zeta' (\mathcal{G} H_{\mathcal{G}} - \hat{\mathcal{G}}) H_{\mathcal{G}}^{-1} \frac{1}{\sqrt{N_S}} \sum_{i \in S} \mathcal{L}_i \right\| \\
= & O_p\left(\frac{\sqrt{N}}{\delta_{NT}^2} \right)
\end{aligned}
\tag{C.11}
$$

by Lemmas C3(i), C4, A4(i) and C1(ii). Lastly, we consider $\hat{\zeta}^{F\prime} \upsilon^F \sum_{i \in S} \psi_i^F / [\sqrt{N_S}(T-p)]$. Recall that $\Xi^{S\prime} \equiv \bar{H}_{\zeta^S}^{-1} \Gamma^{S\prime}$, $\Xi^S = [\Xi_{1:\kappa_S}^S \vdots \Xi_{\kappa_S+1:q_S}^S]$, $\Psi^F = [\Xi_{\kappa_S+1:q_S}^S \vdots \Gamma^F]$, and $\upsilon^F = [\zeta^S \bar{H}_{\kappa_S+1:q_S}^S \vdots \zeta^F]$. Since $\Gamma^F = [\Gamma^{SF\prime} \vdots \Gamma^{FF\prime}]' = [0_{q_F \times N_S} \vdots \Gamma^{FF\prime}]'$, the last q_F entries in ψ_i^F are zeros for $i \in S$ and $\upsilon^F \psi_i^F = \zeta^S \bar{H}_{\kappa_S+1:q_S}^S \xi_{i,\kappa_S+1:q_S}^S$, where $\xi_{i,\kappa_S+1:q_S}^S$ is the transpose of the i^{th} row of $\Xi_{\kappa_S+1:q_S}^S$. Hence, we have

$$
\begin{aligned}
& \frac{1}{\sqrt{N_S}(T-p)} \hat{\zeta}^{F\prime} \upsilon^F \sum_{i \in S} \psi_i^F \\
= & \frac{1}{T-p} \hat{\zeta}^{F\prime} \zeta^S \bar{H}_{\kappa_S+1:q_S}^S \frac{1}{\sqrt{N_S}} \sum_{i \in S} \xi_{i,\kappa_S+1:q_S}^S \\
= & \left[\frac{1}{T-p} H_{\upsilon^F}' \upsilon^{F\prime} \zeta^S + \frac{1}{T-p} (\hat{\zeta}^F - \upsilon^F H_{\upsilon^F})' \zeta^S \right] \bar{H}_{\kappa_S+1:q_S}^S \frac{1}{\sqrt{N_S}} \sum_{i \in S} \xi_{i,\kappa_S+1:q_S}^S \\
= & \left[\frac{1}{T-p} H_{\upsilon^F}' [\zeta^S \bar{H}_{\kappa_S+1:q_S}^S \vdots \zeta^F]' \zeta^S + \frac{1}{T-p} (\hat{\zeta}^F - \upsilon^F H_{\upsilon^F})' \zeta^S \right] \bar{H}_{\kappa_S+1:q_S}^S \frac{1}{\sqrt{N_S}} \sum_{i \in S} \xi_{i,\kappa_S+1:q_S}^S \\
= & \left\{ H_{\upsilon^F}' \left[\begin{array}{c} \left(\bar{H}_{\kappa_S+1:q_S}^{S\prime} \frac{\zeta^{S\prime} \zeta^S}{T-p} \right) \\ \left(\frac{1}{T-p} \zeta^{F\prime} \zeta^S \right) \end{array} \right] + \frac{1}{T-p} (\hat{\zeta}^F - \upsilon^F H_{\upsilon^F})' \zeta^S \right\} \bar{H}_{\kappa_S+1:q_S}^S \frac{1}{\sqrt{N_S}} \sum_{i \in S} \xi_{i,\kappa_S+1:q_S}^S \\
= & H_{\upsilon^F}' \left\{ \left[\begin{array}{c} \bar{H}_{\kappa_S+1:q_S}^{S\prime} \frac{\zeta^{S\prime} \zeta^S}{T-p} \bar{H}_{\kappa_S+1:q_S}^S \\ O_p(1/\sqrt{T}) \end{array} \right] + O_p\left(\frac{1}{\delta_{NT}} \right) \right\} \frac{1}{\sqrt{N_S}} \sum_{i \in S} \xi_{i,\kappa_S+1:q_S}^S
\end{aligned}
\tag{C.12}
$$

where the $O_p(1/\sqrt{T})$ term follows from Assumption 1(c) and the $o_p(\delta_{NT}^{-1})$ follows from Lemma C4 and Cauchy-Schwarz inequality. Using the results in (C.10), (C.11) and (C.12), we have

$$
\begin{aligned}
& \frac{T^{\alpha/2}}{\sqrt{N_S}(T-p)} \hat{\zeta}^{F\prime} \sum_{i \in S} \tilde{X}_i \\
= & H_{\upsilon^F}' \left\{ \left[\begin{array}{c} \bar{H}_{\kappa_S+1:q_S}^{S\prime} \frac{\zeta^{S\prime} \zeta^S}{T-p} \bar{H}_{\kappa_S+1:q_S}^S \\ O_p(1/\sqrt{T}) \end{array} \right] + O_p\left(\frac{1}{\delta_{NT}} \right) \right\} \frac{T^{\alpha/2}}{\sqrt{N_S}} \sum_{i \in S} \xi_{i,\kappa_S+1:q_S}^S + O_p\left(\frac{T^{\alpha/2}}{\delta_{NT}} \right) + O_p\left(\frac{T^{\alpha/2}\sqrt{N}}{\delta_{NT}^2} \right)
\end{aligned}
\tag{C.13}
$$

Note that the $O_p\left(\frac{T^{\alpha/2}}{\delta_{NT}}\right)$ term is $o_p(1)$ because $0 < \alpha \leq 0.5$ and $\sqrt{T}/N \to 0$ as N and $T \to \infty$, that the $O_p\left(\frac{T^{\alpha/2}\sqrt{N}}{\delta_{NT}^2}\right)$ term is $o_p(1)$ because $\sqrt{N}/T^{1-\alpha/2} \to 0$ as N and $T \to \infty$, and that H_{v^F} and $\bar{H}^{S'}_{\kappa_S+1:q_S}\frac{\zeta^{S'}\zeta^S}{T-p}\bar{H}^S_{\kappa_S+1:q_S}$ are asymptotically non-singular, so (C.13) is bounded away from zero asymptotically. Since $\hat{\Omega}$ is positive definite and $O_p(1)$, we have the desired result:

$$\text{plim}_{N,T\to\infty}\inf\left(\frac{T^\alpha}{T-p}W\right) = \text{plim}_{N,T\to\infty}\inf\left[\frac{T^{\alpha/2}\sum_{i\in S}\tilde{X}_i'\hat{\zeta}^F}{\sqrt{N_S}(T-p)}\left(\frac{\sum_{t=p+1}^T\sum_{i\in S}\hat{\zeta}_t^F\hat{\zeta}_t^{F'}\tilde{e}_{it}^2}{(T-p)N_S}\right)^{-1}\frac{T^{\alpha/2}\hat{\zeta}^{F'}\sum_{i\in S}\tilde{X}_i}{\sqrt{N_S}(T-p)}\right] > 0$$

which implies that $W \to \infty$. Q.E.D.

D Data Description

Transformations: 1 = levels, 2 = first differences, 4 = logs, 5= first differences of logs of the original series.

S denote slow variables and F denotes fast variables.

Short Name	Mnemonic	Long Label	F/S	Trans.
PI	A0M052	Personal income (AR, bil. Chain 2005$)	S	5
PI less transfers	A0M051	Personal income less transfer payments (AR, bil. Chain 2005$)	S	5
Consumption	A0M224_R	Real Consumption (AC) a0m224/gmdc (a0m224 is from TCB)	S	5
M&T sales	A0M057	Manufacturing and trade sales (mil. Chain 2005 $)	S	5
Retail sales	A0M059	Sales of retail stores (mil. Chain 2000 $)	S	5
IP: total	IPS10	Industrial Production Index - Total Index	S	5
IP: products	IPS11	Industrial Production Index - Products, Total	S	5
IP: final prod	IPS299	Industrial Production Index - Final Products	S	5
IP: cons gds	IPS12	Industrial Production Index - Consumer Goods	S	5
IP: cons dble	IPS13	Industrial Production Index - Durable Consumer Goods	S	5
IP:cons nondble	IPS18	Industrial Production Index - Nondurable Consumer Goods	S	5
IP:bus eqpt	IPS25	Industrial Production Index - Business Equipment	S	5
IP: matls	IPS32	Industrial Production Index - Materials	S	5
IP: dble mats	IPS34	Industrial Production Index - Durable Goods Materials	S	5
IP:nondble mats	IPS38	Industrial Production Index - Nondurable Goods Materials	S	5
IP: mfg	IPS43	Industrial Production Index - Manufacturing (Sic)	S	5
IP: res util	IPS307	Industrial Production Index - Residential Utilities	S	5
IP: fuels	IPS306	Industrial Production Index - Fuels	S	5
NAPM prodn	PMP	Napm Production Index (Percent)	S	1
Cap util	B00004	Capacity Utilization	S	2
Emp CPS total	LHEM	Civilian Labor Force: Employed, Total (Thous.,Sa)	S	5
Emp CPS nonag	LHNAG	Civilian Labor Force: Employed, Nonagric.Industries (Thous.,Sa)	S	5
U: all	LHUR	Unemployment Rate: All Workers, 16 Years & Over (%,Sa)	S	2
U: mean duration	LHU680	Unemploy.By Duration: Average(Mean)Duration In Weeks (Sa)	S	2
U < 5 wks	LHU5	Unemploy.By Duration: Persons Unempl.Less Than 5 Wks (Thous.,Sa)	S	5
U 5-14 wks	LHU14	Unemploy.By Duration: Persons Unempl.5 To 14 Wks (Thous.,Sa)	S	5
U 15+ wks	LHU15	Unemploy.By Duration: Persons Unempl.15 Wks + (Thous.,Sa)	S	5
U 15-26 wks	LHU26	Unemploy.By Duration: Persons Unempl.15 To 26 Wks (Thous.,Sa)	S	5
U 27+ wks	LHU27	Unemploy.By Duration: Persons Unempl.27 Wks + (Thous,Sa)	S	5
UI claims	A0M005	Average Weekly Initial Claims, Unemploy. Insurance (Thous.) (TCB)	S	5
Emp: total	CES002	Employees On Nonfarm Payrolls: Total Private	S	5
Emp: gds prod	CES003	Employees On Nonfarm Payrolls - Goods-Producing	S	5
Emp: mining	CES006	Employees On Nonfarm Payrolls - Mining	S	5
Emp: const	CES011	Employees On Nonfarm Payrolls - Construction	S	5
Emp: mfg	CES015	Employees On Nonfarm Payrolls - Manufacturing	S	5
Emp: dble gds	CES017	Employees On Nonfarm Payrolls - Durable Goods	S	5
Emp: nondbles	CES033	Employees On Nonfarm Payrolls - Nondurable Goods	S	5
Emp: services	CES046	Employees On Nonfarm Payrolls - Service-Providing	S	5
Emp: TTU	CES048	Employees On Nonfarm Payrolls - Trade, Transportation, And Utilities	S	5
Emp: wholesale	CES049	Employees On Nonfarm Payrolls - Wholesale Trade	S	5
Emp: retail	CES053	Employees On Nonfarm Payrolls - Retail Trade	S	5
Emp: FIRE	CES088	Employees On Nonfarm Payrolls - Financial Activities	S	5
Emp: Govt	CES140	Employees On Nonfarm Payrolls - Government	S	5

Avg hrs	CES151	Avg Weekly Hrs: Private Nonfarm - Goods-Producing	S	1
Overtime: mfg	CES155	Avg Weekly Hrs: Private Nonfarm Payrolls - Mfg Overtime Hours	S	2
Avg hrs: mfg	A0M001	Average Weekly Hours, Mfg. (Hours) (TCB)	S	1
NAPM empl	PMEMP	Napm Employment Index (Percent)	S	1
CPI-U: all	PUNEW	CPI-U: All Items (82-84=100,Sa)	S	5
CPI-U: apparel	PU83	Cpi-U: Apparel & Upkeep (82-84=100,Sa)	S	5
CPI-U: transp	PU84	Cpi-U: Transportation (82-84=100,Sa)	S	5
CPI-U: medical	PU85	Cpi-U: Medical Care (82-84=100,Sa)	S	5
CPI-U: comm.	PUC	Cpi-U: Commodities (82-84=100,Sa)	S	5
CPI-U: dbles	PUCD	Cpi-U: Durables (82-84=100,Sa)	S	5
CPI-U: services	PUS	Cpi-U: Services (82-84=100,Sa)	S	5
CPI-U: ex food	PUXF	Cpi-U: All Items Less Food (82-84=100,Sa)	S	5
CPI-U: ex shelter	PUXHS	Cpi-U: All Items Less Shelter (82-84=100,Sa)	S	5
CPI-U: ex med	PUXM	Cpi-U: All Items Less Medical Care (82-84=100,Sa)	S	5
PCE defl	JPC	Pce, Impl Pr Defl:Pce (1987=100)	S	5
PCE defl: dlbes	JPCD	Pce, Impl Pr Defl:Pce; Durables (1987=100)	S	5
PCE defl: nondble	JPCN	Pce, Impl Pr Defl:Pce; Nondurables (1996=100)	S	5
PCE defl: services	JPCSV	Pce, Impl Pr Defl:Pce; Services (1987=100)	S	5
AHE: goods	CES275	Avg Hourly Earnings: Private Nonfarm Payrolls - Goods-Producing	S	5
AHE: const	CES277	Avg Hourly Earnings: Private Nonfarm Payrolls - Construction	S	5
AHE: mfg	CES278	Avg Hourly Earnings: Private Nonfarm Payrolls - Manufacturing	S	5
HStarts: Total	HSFR	Housing Starts:Nonfarm(1947-58);Total Farm&Nonfarm(1959-)(Thous.,Saar)	F	4
HStarts: NE	HSNE	Housing Starts:Northeast (Thous.U.)S.A.	F	4
HStarts: MW	HSMW	Housing Starts:Midwest(Thous.U.)S.A.	F	4
HStarts: South	HSSOU	Housing Starts:South (Thous.U.)S.A.	F	4
HStarts: West	HSWST	Housing Starts:West (Thous.U.)S.A.	F	4
BP: total	HSBR	Housing Authorized: Total New Priv Housing Units (Thous.,Saar)	F	4
BP: NE	HSBNE	Houses Authorized By Build. Permits:Northeast(Thou.U.)S.A	F	4
BP: MW	HSBMW	Houses Authorized By Build. Permits:Midwest(Thou.U.)S.A.	F	4
BP: South	HSBSOU	Houses Authorized By Build. Permits:South(Thou.U.)S.A.	F	4
BP: West	HSBWST	Houses Authorized By Build. Permits:West(Thou.U.)S.A.	F	4
PMI	PMI	Purchasing Managers' Index (Sa)	F	1
PMNO	NAPM new ordrs	Napm New Orders Index (Percent)	F	1
PMDEL	NAPM vendor del	Napm Vendor Deliveries Index (Percent)	F	1
PMNV	NAPM Invent	Napm Inventories Index (Percent)	F	1
A0M008	Orders: cons gds	Mfrs' New Orders, Consumer Goods And Materials (Bil. Chain 1982 $) (TCB)	F	5
A0M007	Orders: dble gds	Mfrs' New Orders, Durable Goods Industries (Bil. Chain 2000 $) (TCB)	F	5
A0M027	Orders: cap gds	Mfrs' New Orders, Nondefense Capital Goods (Mil. Chain 1982 $) (TCB)	F	5
A1M092	Unf orders: dble	Mfrs' Unfilled Orders, Durable Goods Indus. (Bil. Chain 2000 $) (TCB)	F	5
A0M070	M&T invent	Manufacturing And Trade Inventories (Bil. Chain 2000 $) (TCB)	F	5
A0M077	M&T invent/sales	Ratio, Mfg. And Trade Inventories To Sales (Based On Chain 2000 $) (TCB)	F	2
FM1	M1	Money Stock: M1 (Bil$,Sa)	F	5
FM2	M2	Money Stock: M2 (Bil$,Sa)	F	5
A0M106	M2 (real)	Money Supply - M2 In 1996 Dollars (Bci)	F	5
FMFBA	MB	Monetary Base, Adj For Reserve Requirement Changes(Mil$,Sa)	F	5
FMRRA	Reserves tot	Depository Inst Reserves:Total, Adj For Reserve Req Chgs(Mil$,Sa)	F	5
FMRNBA	Reserves nonbor	Depository Inst Reserves:Nonborrowed,Adj Res Req Chgs(Mil$,Sa)	F	5
A0M101	C&I loans	Commercial and Industrial Loans Outstanding (mil. Chain 2005$)	F	5
CCINRV	Cons credit	Consumer Credit Outstanding - Nonrevolving(G19)	F	5
A0M095	Inst cred/PI	Ratio, Consumer Installment Credit To Personal Income (Pct.) (TCB)	F	2
FSPCOM	S&P 500	S&P's Common Stock Price Index: Composite (1941-43=10)	F	5
FSPIN	S&P: indust	S&P's Common Stock Price Index: Industrials (1941-43=10)	F	5

FSDXP	S&P div yield	S&P's Composite Common Stock: Dividend Yield (% Per Annum)	F	2
FSPXE	S&P PE ratio	S&P's Composite Common Stock: Price-Earnings Ratio (%,Nsa)	F	5
FYFF	FedFunds	Interest Rate: Federal Funds (Effective) (% Per Annum,Nsa)	F	2
CP90/RMCP3	Commpaper	Cmmercial Paper Rate (AC)	F	2
FYGM3	3 mo T-bill	Interest Rate: U.S.Treasury Bills,Sec Mkt,3-Mo.(% Per Ann,Nsa)	F	2
FYGM6	6 mo T-bill	Interest Rate: U.S.Treasury Bills,Sec Mkt,6-Mo.(% Per Ann,Nsa)	F	2
FYGT1	1 yr T-bond	Interest Rate: U.S.Treasury Const Maturities,1-Yr.(% Per Ann,Nsa)	F	2
FYGT5	5 yr T-bond	Interest Rate: U.S.Treasury Const Maturities,5-Yr.(% Per Ann,Nsa)	F	2
FYGT10	10 yr T-bond	Interest Rate: U.S.Treasury Const Maturities,10-Yr.(% Per Ann,Nsa)	F	2
FYAAAC	Aaabond	Bond Yield: Moody's Aaa Corporate (% Per Annum)	F	2
FYBAAC	Baabond	Bond Yield: Moody's Baa Corporate (% Per Annum)	F	2
SCP90(AC)	CP-FF spread	cp90-fyff (AC)	F	1
sFYGM3	3 mo-FF spread	fygm3-fyff (AC)	F	1
sFYGM6	6 mo-FF spread	fygm6-fyff (AC)	F	1
sFYGT1	1 yr-FF spread	fygt1-fyff (AC)	F	1
sFYGT5	5 yr-FFspread	fygt5-fyff (AC)	F	1
sFYGT10	10yr-FF spread	fygt10-fyff (AC)	F	1
sFYAAAC	Aaa-FF spread	fyaaac-fyff (AC)	F	1
sFYBAAC	Baa-FF spread	fybaac-fyff (AC)	F	1
EXRSW	Ex rate: Switz	Foreign Exchange Rate: Switzerland (Swiss Franc Per U.S.$)	F	5
EXRJAN	Ex rate: Japan	Foreign Exchange Rate: Japan (Yen Per U.S.$)	F	5
EXRUK	Ex rate: UK	Foreign Exchange Rate: United Kingdom (Cents Per Pound)	F	5
EXRCAN	EX rate: Canada	Foreign Exchange Rate: Canada (Canadian $ Per U.S.$)	F	5
PWFSA	PPI: fin gds	Producer Price Index: Finished Goods (82=100,Sa)	F	5
PWFCSA	PPI: cons gds	Producer Price Index: Finished Consumer Goods (82=100,Sa)	F	5
PWIMSA	PPI: int mat'ls	Producer Price Index:I ntermed Mat.Supplies & Components(82=100,Sa)	F	5
PWCMSA	PPI: crude mat'ls	Producer Price Index: Crude Materials (82=100,Sa)	F	5
PSCCOM	Commod: spot price	Spot market price index: bls & crb: all commodities(1967=100)	F	5
PMCP	NAPM com price	Napm Commodity Prices Index (Percent)	F	1
U0M083	Consumer expect	U. Of Mich. Index Of Consumer Expectations(Bcd-83)	F	2

References

[1] Amengual, D. and Watson, M. (2007), "Consistent Estimation of the Number of Dynamic Factors in a Large N And T Panel, *Journal of Business and Economic Statistics*, 25(1): 91-96.

[2] Bai, J. (2003), "Inferential Theory for Factor Models of Large Dimensions," *Econometrica* 71, 135-172.

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

[4] Bai, J., and S. Ng (2006), "Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-augmented Regressions," *Econometrica* 74, 1133-1150.

[5] Bai, J., and S. Ng (2007), "Determining the Number of Primitive Shocks in Factor Models," *Journal of Business and Economic Statistics*, 25(1): 52-60.

[6] Bernanke, B.S., J. Boivin and P. Eliasz (2005), "Measuring the Effects of Monetary Policy: a Factor-augmented Vector Autoregressive (FAVAR) Approach", *Quarterly Journal of Economics* 120, 387–422.

[7] Bianchi, F., H. Mumtaz, and P. Surico (2009), "The Great Moderation of the Term Structure of U.K. Interest Rates," *Journal of Monetary Economics*, 56, 856-871.

[8] Boivin, J. and M. Giannoni, and I. Mihov (2009), " Sticky Prices and Monetary Policy: Evidence from Disaggregated US Data," *American Economic Review*, 99:1, 350-384.

[9] Eickmeier, S., W. Lemke, and M. Marcellino (2011), "Classical Time-varying FAVAR Models – Estimation, Forecasting and Structural Analysis," CEPR Discussion Paper No. DP8321.

[10] Forni, M. and L. Gambetti (2010), "The Dynamic Effects of Monetary Policy: A Structural Factor Model Approach," *Journal of Monetary Economics*, 57(2), 203-216.

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

[12] Giannone, D., L. Reichlin, and L. Sala (2004), "Monetary Policy in Real Time", NBER Macroeconomics Annual, 2004.

[13] Gilchrist, S., V. Yankov, and E. Zakraisek (2009), "Credit Market Shocks and Economic Fluctuations: Evidence from Corporate Bond and Stock Markets," *Journal of Monetary Economics*, 56, 471-493.

[14] Hallin, M. and Liska R. (2007), "The Generalized Dynamic Factor Model: Determining the Number of Factors", *Journal of the American Statistical Association*, 102, 603-617.

[15] Han, X and A. Inoue (2011), "Tests for Parameter Stability in Dynamic Factor Models," Manuscript, NC State University.

[16] Inoue, A. and L. Kilian (2011), "Inference on Impulse Response Functions inStructural VAR Models," CEPR Discussion Paper No. DP8419.

[17] Ledoit, O. and M. Wolf (2004), "A Well-conditioned Estimator for Large-dimensional Covariance Matrices," *Journal of Multivariate Analysis*, 88, 365-411.

[18] Ludvigson, S and S. Ng (2009), "A Factor Analysis of Bond Risk Premia," *Review of Financial Studies*, 22(12), 5027-5067.

[19] Mumtaz, H. and P. Surico (2009), "The Transmission of International Shocks: A Factor-Augmented VAR Approach," *Journal of Money, Credit and Banking*, 41, Issue Supplement s1. 71-100.

[20] Onatski, A. (2009), "Testing Hypotheses about the Number of Factors in Large Factor Models," *Econometrica*, 77(5), 1447-1479.

[21] Phillips, P.C.B. and H.R. Moon (1999), "Linear Regression Limit Theory for Nonstationary Panel Data," *Econometrica*, 67, 1057-1111.

[22] Sargent, T.J. and C.A. Sims (1977), "Business Cycle Modelling without Pretending to Have Too Much a-priori Economic Theory.", in: Sims et al., eds., New Methods in Business Cycle Research (Federal Reserve Bank of Minneapolis, Minneapolis).

[23] Sims, C.A. (1992), "Interpreting the Macroeconomic Time Series Facts, The Effects of Monetary Policy," *European Economic Review* 36, 975-1011.

[24] Stock, J.H., and M.W. Watson (2002), "Macroeconomic forecasting using diffusion indexes", *Journal of Business and Economic Statistics* 20, 147-162.

[25] Stock, J.H., and M.W. Watson (2005), "Implications of Dynamic Factor Models for VAR Analysis," NBER Working Paper No. 11467.

[26] Stock, J.H., and M.W. Watson (2010), "Dynamic Factor Models," Chapter 2 in M. Clements and D. Hendry (eds), Oxford Handbook of Economic Forecasting.

[27] Uhlig, H. (2005), “What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure,” *Journal of Monetary Economics*, 52, 381-419.

[28] Yamamoto, Y. (2011), “Bootstrap Inference of Impulse Response Functions in Factor-Augmented Vector Autoregressions,” Working Paper, University of Alberta.

Table 1: Size of Tests for the Overidentifying Restrictions under DGP N1 with 5000 Replications

$N_S = 0.5N$

$r = 5,\ q = 3,\ q_F = 1$

		Individual Tests			Bonferroni&Pooled		Onatski		New Joint Tests	
N	T	sw_i	w_i	lm_i	B_w	P_w	R_{dyn}	R_{stat}	W	LM
200	200	0.131	0.086	0.077	0.241	0.478	0.067	0.061	0.080	0.077
200	300	0.124	0.078	0.072	0.178	0.365	0.053	0.058	0.079	0.078
200	500	0.127	0.073	0.070	0.157	0.308	0.057	0.063	0.072	0.072
300	200	0.109	0.078	0.070	0.209	0.476	0.052	0.051	0.074	0.072
300	300	0.096	0.070	0.065	0.154	0.331	0.056	0.051	0.068	0.067
300	500	0.096	0.065	0.062	0.127	0.239	0.051	0.053	0.067	0.066
500	200	0.590	0.074	0.066	0.220	0.552	0.054	0.051	0.073	0.070
500	300	0.084	0.067	0.062	0.149	0.358	0.055	0.053	0.064	0.063
500	500	0.077	0.061	0.058	0.106	0.230	0.054	0.047	0.063	0.062

$r = 5,\ q = 4,\ q_F = 3$

		Individual Tests			Bonferroni&Pooled		Onatski		New Joint Tests	
N	T	sw_i	w_i	lm_i	B_w	P_w	R_{dyn}	R_{stat}	W	LM
200	200	0.126	0.101	0.073	0.312	0.713	0.066	0.079	0.087	0.080
200	300	0.117	0.088	0.071	0.231	0.540	0.067	0.089	0.075	0.069
200	500	0.115	0.079	0.069	0.162	0.403	0.065	0.134	0.073	0.070
300	200	0.113	0.092	0.065	0.284	0.752	0.059	0.067	0.074	0.068
300	300	0.099	0.080	0.063	0.209	0.543	0.064	0.072	0.066	0.061
300	500	0.091	0.070	0.060	0.144	0.344	0.066	0.098	0.059	0.057
500	200	0.648	0.088	0.062	0.302	0.863	0.059	0.056	0.072	0.067
500	300	0.090	0.076	0.059	0.198	0.623	0.061	0.060	0.070	0.065
500	500	0.077	0.066	0.057	0.131	0.350	0.059	0.066	0.062	0.060

Note: The nominal size is 5%. The first two columns are the numbers of observations in cross section and time dimensions. The numbers in columns $3-5$ are the means of individual tests' rejection rates from 5000 simulated samples. Columns $6-11$ are the rejection rates of Bonferroni, Pooled, Onatski, and our new joint test statistics. The upper panel of Table 1 summarizes the results based on the setup where $r = 5$, $q = 3$, and $q_F = 1$, whereas the lower panel of Table 1 summarizes the results based on the setup where $r = 5$, $q = 4$, and $q_F = 3$.

Table 2: Size of Tests for the Overidentifying Restrictions under DGP N2 with 5000 Replications

$N_S = 0.5N$

$r = 5$, $q = 3$, $q_F = 1$

		Individual Tests			Bonferroni&Pooled		Onatski		New Joint Tests	
N	T	sw_i	w_i	lm_i	B_w	P_w	R_{dyn}	R_{stat}	W	LM
200	200	0.077	0.083	0.075	0.248	0.470	0.054	0.074	0.080	0.076
200	300	0.073	0.083	0.077	0.217	0.460	0.059	0.107	0.081	0.080
200	500	0.070	0.086	0.082	0.229	0.500	0.067	0.175	0.087	0.086
300	200	0.075	0.074	0.066	0.212	0.400	0.054	0.058	0.064	0.063
300	300	0.068	0.071	0.065	0.173	0.353	0.055	0.069	0.073	0.070
300	500	0.064	0.070	0.067	0.156	0.329	0.050	0.129	0.076	0.074
500	200	0.577	0.067	0.060	0.190	0.367	0.053	0.054	0.069	0.067
500	300	0.066	0.064	0.059	0.149	0.272	0.050	0.064	0.058	0.057
500	500	0.060	0.062	0.059	0.123	0.228	0.051	0.076	0.060	0.059

$r = 5$, $q = 4$, $q_F = 3$

		Individual Tests			Bonferroni&Pooled		Onatski		New Joint Tests	
N	T	sw_i	w_i	lm_i	B_w	P_w	R_{dyn}	R_{stat}	W	LM
200	200	0.089	0.101	0.073	0.359	0.719	0.101	0.320	0.081	0.073
200	300	0.077	0.094	0.075	0.273	0.615	0.135	0.551	0.084	0.078
200	500	0.069	0.093	0.082	0.256	0.596	0.174	0.846	0.089	0.087
300	200	0.088	0.090	0.064	0.325	0.714	0.080	0.204	0.078	0.069
300	300	0.074	0.081	0.064	0.234	0.552	0.088	0.406	0.078	0.075
300	500	0.066	0.076	0.066	0.174	0.461	0.121	0.754	0.070	0.068
500	200	0.625	0.083	0.058	0.324	0.764	0.056	0.122	0.061	0.056
500	300	0.075	0.074	0.057	0.215	0.536	0.065	0.228	0.059	0.055
500	500	0.064	0.067	0.058	0.154	0.358	0.073	0.580	0.062	0.060

Note: The nominal size is 5%. The first two columns are the numbers of observations in cross section and time dimensions. The numbers in columns $3-5$ are the means of individual tests' rejection rates from 5000 simulated samples. Columns $6-11$ are the rejection rates of Bonferroni, Pooled, Onatski, and our new joint test statistics. The upper panel of Table 2 summarizes the results based on the setup where $r = 5$, $q = 3$, and $q_F = 1$, whereas the lower panel of Table 2 summarizes the results based on the setup where $r = 5$, $q = 4$, and $q_F = 3$.

Table 3: Power of Tests for the Overidentifying Restrictions under DGP A1 with 5000 Replications

$N_S = 0.5N$

			$r=5$, $q=3$, $\kappa_F=1$						$r=5$, $q=4$, $\kappa_F=3$					
			Individual		Onatski		New		Individual		Onatski		New	
a	N	T	w_i	lm_i	R_{dyn}	R_{stat}	W	LM	w_i	lm_i	R_{dyn}	R_{stat}	W	LM
0.2	200	200	0.242	0.236	0.056	0.059	0.470	0.444	0.313	0.284	0.056	0.080	0.798	0.751
0.2	300	300	0.239	0.233	0.058	0.107	0.549	0.526	0.293	0.276	0.060	0.121	0.859	0.832
0.2	500	500	0.243	0.239	0.055	0.594	0.624	0.612	0.281	0.271	0.060	0.218	0.927	0.911
0.4	200	200	0.409	0.399	0.060	0.160	0.590	0.555	0.530	0.504	0.059	0.103	0.899	0.848
0.4	300	300	0.420	0.413	0.065	0.607	0.655	0.626	0.516	0.500	0.074	0.264	0.948	0.917
0.4	500	500	0.434	0.431	0.137	0.999	0.729	0.704	0.509	0.500	0.097	0.972	0.972	0.955
0.6	200	200	0.566	0.556	0.078	0.406	0.642	0.597	0.738	0.719	0.077	0.169	0.937	0.898
0.6	300	300	0.596	0.590	0.145	0.936	0.711	0.677	0.737	0.726	0.086	0.753	0.961	0.934
0.6	500	500	0.624	0.621	0.499	1.000	0.770	0.741	0.737	0.730	0.240	1.000	0.980	0.967
0.8	200	200	0.674	0.665	0.116	0.641	0.685	0.634	0.888	0.877	0.083	0.400	0.959	0.916
0.8	300	300	0.723	0.718	0.292	0.994	0.736	0.701	0.904	0.898	0.151	0.967	0.976	0.954
0.8	500	500	0.770	0.768	0.816	1.000	0.798	0.764	0.915	0.912	0.620	1.000	0.988	0.977
1	200	200	0.710	0.702	0.146	0.753	0.698	0.651	0.965	0.958	0.114	0.624	0.968	0.934
1	300	300	0.762	0.757	0.392	0.999	0.753	0.706	0.980	0.978	0.283	0.995	0.977	0.956
1	500	500	0.817	0.815	0.922	1.000	0.813	0.775	0.990	0.990	0.863	1.000	0.991	0.977

Note: The parameter a controls the fraction of slow variables that violate the zero restrictions in their loadings on the fast shocks. The numbers in columns 3 – 4 and 9 –10 are the means of rejection rates of our individual tests from 5000 simulated samples. Columns 5 – 8 and 11 – 14 are the rejection rates of Onatski and our new joint test statistics. The left panel of Table 3 summarizes the results based on the setup where $r = 5$, $q = 3$, and $\kappa_F = 1$, whereas the right panel of Table 3 summarizes the results based on the setup where $r = 5$, $q = 4$, and $\kappa_F = 3$. The true number of fast shocks $q_F = 0$.

Table 4: Power of Tests for the Overidentifying Restrictions under DGP A2 with 5000 Replications

$N_S = 0.5N$

			$r=5$, $q=3$, $\kappa_F=1$						$r=5$, $q=4$, $\kappa_F=3$					
			Individual		Onatski		New		Individual		Onatski		New	
b	N	T	w_i	lm_i	R_{dyn}	R_{stat}	W	LM	w_i	lm_i	R_{dyn}	R_{stat}	W	LM
0.2	200	200	0.246	0.232	0.057	0.059	0.212	0.202	0.441	0.389	0.060	0.064	0.372	0.344
0.2	300	300	0.292	0.283	0.060	0.052	0.253	0.245	0.515	0.482	0.056	0.055	0.456	0.434
0.2	500	500	0.368	0.363	0.055	0.049	0.336	0.328	0.634	0.617	0.053	0.054	0.592	0.576
0.4	200	200	0.455	0.443	0.057	0.054	0.426	0.406	0.762	0.728	0.057	0.058	0.737	0.693
0.4	300	300	0.525	0.517	0.056	0.053	0.494	0.477	0.830	0.814	0.057	0.070	0.809	0.780
0.4	500	500	0.609	0.605	0.053	0.299	0.604	0.588	0.900	0.893	0.051	0.124	0.894	0.878
0.6	200	200	0.592	0.581	0.052	0.112	0.570	0.540	0.892	0.874	0.061	0.079	0.892	0.844
0.6	300	300	0.655	0.649	0.058	0.493	0.629	0.601	0.931	0.923	0.059	0.227	0.926	0.895
0.6	500	500	0.724	0.721	0.091	0.999	0.718	0.697	0.963	0.961	0.066	0.975	0.960	0.944
0.8	200	200	0.674	0.664	0.081	0.462	0.661	0.620	0.944	0.934	0.074	0.236	0.945	0.905
0.8	300	300	0.729	0.724	0.166	0.968	0.710	0.673	0.966	0.962	0.090	0.864	0.967	0.940
0.8	500	500	0.788	0.785	0.581	1.000	0.783	0.751	0.983	0.982	0.343	1.000	0.982	0.969
1	200	200	0.710	0.702	0.146	0.753	0.698	0.651	0.965	0.958	0.114	0.624	0.968	0.934
1	300	300	0.762	0.757	0.392	0.999	0.753	0.706	0.980	0.978	0.283	0.995	0.977	0.956
1	500	500	0.817	0.815	0.922	1.000	0.813	0.775	0.990	0.990	0.863	1.000	0.991	0.977

Note: The parameter b controls the standard deviation of the loadings that are assumed to be zeros by the identification restrictions. The numbers in columns 3 – 4 and 9 –10 are the means of rejection rates of our individual tests from 5000 simulated samples. Columns 5 – 8 and 11 – 14 are the rejection rates of Onatski and our new joint test statistics. The left panel of Table 3 summarizes the results based on the setup where $r = 5$, $q = 3$, and $\kappa_F = 1$, whereas the right panel of Table 3 summarizes the results based on the setup where $r = 5$, $q = 4$, and $\kappa_F = 3$. The true number of fast shocks $q_F = 0$.

Table 5: p-values for testing H_0: $q_F = \kappa_F$ against H_1: $q_F < \kappa_F$

	W					LM				
	$q=4$	$q=5$	$q=6$	$q=7$	$q=8$		$q=4$	$q=5$	$q=6$	$q=7$	$q=8$
$\kappa_F = 1$	0.933	0.755	0.385	0.053	0.363	$\kappa_F = 1$	0.934	0.759	0.388	0.055	0.368
	(0.810)	(0.858)	(0.293)	(0.263)	(0.438)		(0.814)	(0.861)	(0.300)	(0.265)	(0.440)
$\kappa_F = 2$	0.000	0.000	0.020	0.017	0.063	$\kappa_F = 2$	0.000	0.000	0.025	0.022	0.069
	(0.029)	(0.160)	(0.025)	(0.027)	(0.020)		(0.035)	(0.172)	(0.030)	(0.032)	(0.023)
$\kappa_F = 3$	0.000	0.000	0.000	0.000	0.000	$\kappa_F = 3$	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.059)	(0.000)	(0.000)	(0.000)		(0.000)	(0.082)	(0.000)	(0.000)	(0.000)
	R_{dyn}					R_{stat}				
	$q=4$	$q=5$	$q=6$	$q=7$	$q=8$		$q=4$	$q=5$	$q=6$	$q=7$	$q=8$
$\kappa_F = 1$	0.346	0.749	0.819	0.668	0.384	$\kappa_F = 1$	0.791	0.870	0.854	0.706	0.499
	(0.912)	(0.971)	(0.992)	(0.926)	(0.705)		(0.222)	(0.189)	(0.151)	(0.110)	(0.843)
$\kappa_F = 2$	0.392	0.346	0.749	0.819	0.668	$\kappa_F = 2$	0.841	0.791	0.870	0.854	0.706
	(0.459)	(0.912)	(0.971)	(0.992)	(0.926)		(0.255)	(0.222)	(0.189)	(0.151)	(0.110)
$\kappa_F = 3$	0.043	0.392	0.346	0.749	0.819	$\kappa_F = 3$	0.288	0.841	0.791	0.870	0.854
	(0.505)	(0.459)	(0.912)	(0.971)	(0.992)		(0.184)	(0.255)	(0.222)	(0.189)	(0.151)

Note: W and LM are our new statistics, and R_{dyn} and R_{stat} are the dynamic and static versions Onatski's (2009) statistics, respectively. The numbers outside the parentheses are the p-values when the number of lags of $\hat{F}_t$ is equal to 2, and the numbers inside the parentheses are the p-values when the number of lags is equal to 4.

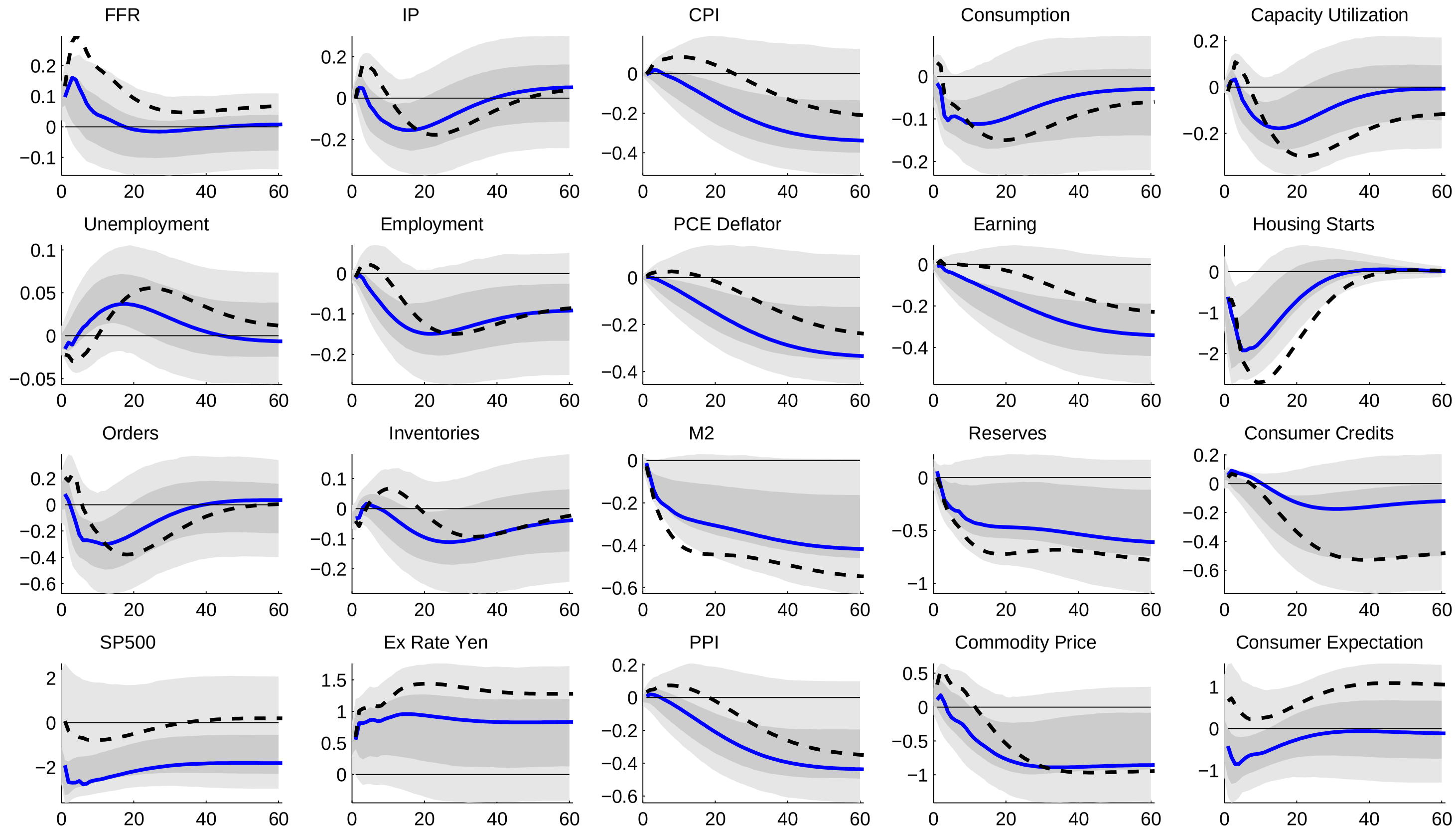

Figure 1: The Impulse Responses (in Percentage) of Macroeconomic Variables to a Unity Variance Contractionary Monetary Policy Shock

The solid lines are based on the identifying restriction $\kappa_F = 1$ with 68% and 95% bootstrap confidence bands. The dashed lines are based on the identifying restriction $\kappa_F = 2$.

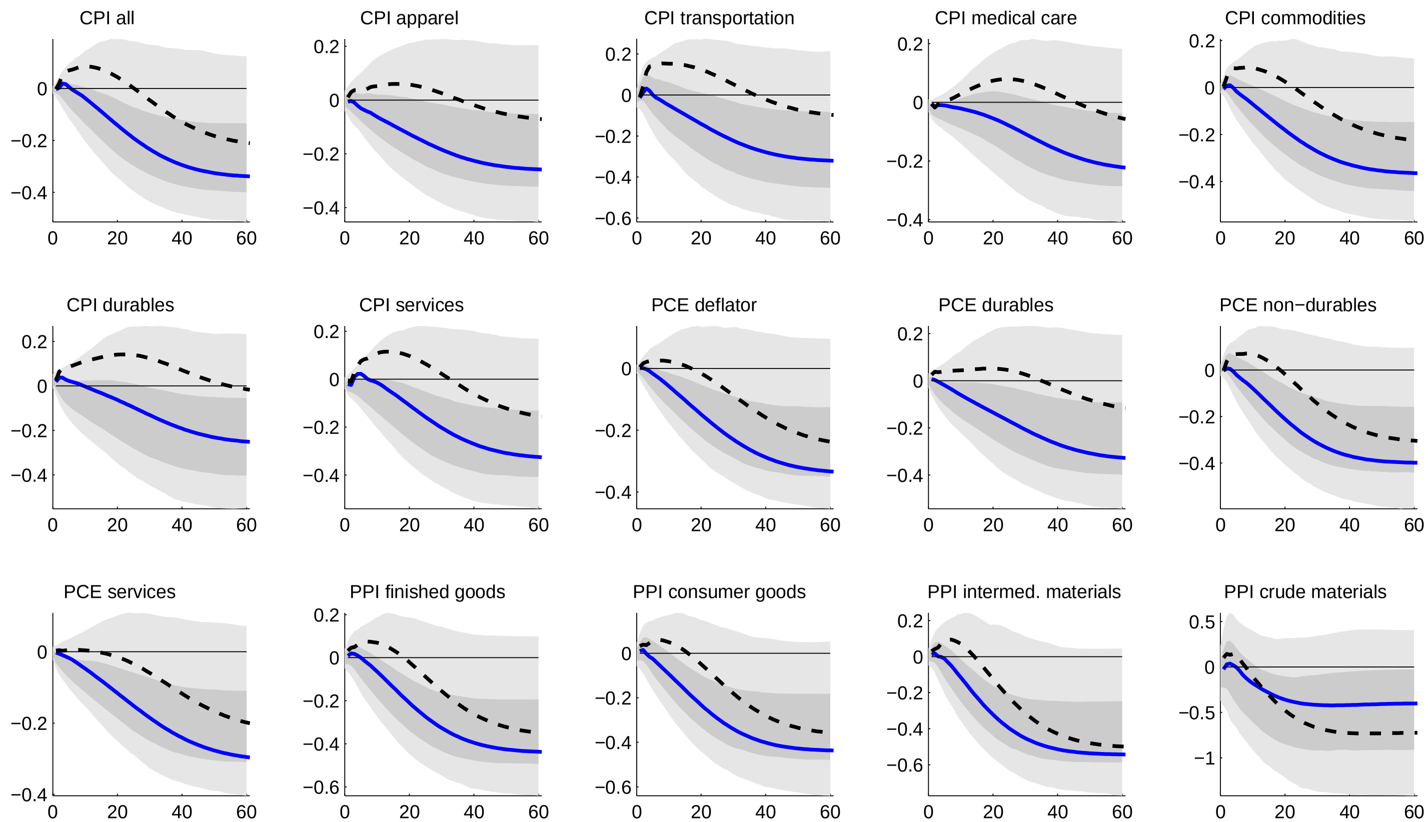

Figure 2: The Impulse Responses (in Percentage) of Price Indexes to a Unity Variance Contractionary Monetary Policy Shock

The solid lines are based on the identifying restriction $\kappa_F = 1$ with 68% and 95% bootstrap confidence bands. The dashed lines are based on the identifying restriction $\kappa_F = 2$.