Supplement to "Semiparametric Estimates of Monetary Policy Effects: String Theory Revisited" - More on Data Construction and Inference*

Joshua D. Angrist
MIT and NBER

Òscar Jordà
Federal Reserve Bank of San Francisco
and U.C. Davis

Guido M. Kuersteiner
University of Maryland

June 16, 2016

The supplemental appendix provides additional information on how we construct our main policy predictor. It also develops the asymptotic theory that justifies our estimators, explains how we compute the standard errors and introduces our specification tests.

I Market Based Expectations in Meeting Months

This section discusses the construction of s_t^1, the indicator of the risk-adjusted expected change in the federal funds rate target based on futures prices for months in which there was an FOMC meeting scheduled. Notice that this variable is a key ingredient of the vector z_t when we construct the ordered propensity score model $p^j(z_t, \psi)$. The discussion in this section borrows many elements from Hamilton (2008). Using the same notation as in the previous section, the effective federal funds rate observed on day k of month t can be described as

$$r_{t,k} = \overline{r}_{t,k} + u_{t,k}$$

where $u_{t,k}$ is a deviation from target. Where possible, we omit referencing t to any particular month and we use κ to denote the total number of days in the month generically to simplify the exposition. Historically, the Fed has been able to keep the effective federal funds rate trading within a few basis points of the target. However, on occasion there can be considerable fluctuations that are related to the *seasonality* of the maintenance period.

Denote by $f_{t,k}^0$ the spot federal funds rate contract. The contract settlement r_t^a is the average of $r_{t,k}$

*Thanks go to Jordi Galì, Ivan Werning, and seminar participants at the Bank of England, the European Central Bank, the Federal Reserve Bank of San Francisco, the NBER Summer Institute 2013, and Yale University for helpful suggestions and comments. The views expressed herein are solely the responsibility of the authors and should not be interpreted as representing the views of the Federal Reserve Bank of San Francisco or the Board of Governors of the Federal Reserve System. Early Elias, Helen Irvin, and Daniel Molitor provided expert research assistance.

Contact Information: angrist@mit.edu; oscar.jorda@sf.frb.org and ojorda@ucdavis.edu; kuersteiner@econ.umd.edu

over the month. Assuming the month has κ days

$$r_t^a = \frac{1}{\kappa}\sum_{k=1}^{\kappa} r_{t,k}$$

and using the risk neutral measure to price the spot federal funds rate contract

$$f_{t,k}^0 = E_{t,k}^Q(r_t^a).$$

Suppose there is a change Δ in the target, effective on day k^* of that month. The target rate is $\overline{r}_{t,k} = \overline{r}$ for $k = 1, ..., (k^* - 1)$ and $\overline{r}_{t,k} = \overline{r} + \Delta$ for $k = k^*, ..., \kappa$ (assuming only one target change to simplify the exposition). Accordingly, the futures rate observed on day k^* within the month is

$$\begin{aligned} f_{t,k^*}^0 &= E_{t,k^*}^Q\left[\frac{1}{\kappa}\sum_{k=1}^{\kappa} r_{t,k}\right] = E_{t,k^*}^Q\left[\frac{1}{\kappa}\sum_{k=1}^{\kappa}(\overline{r}_{t,k} + u_{t,k})\right] = \\ & \frac{k^*-1}{\kappa}\overline{r} + \frac{\kappa - (k^*-1)}{\kappa}E_{t,k^*}^Q\left[\overline{r} + \Delta\right] + \frac{1}{\kappa}\sum_{k=1}^{k^*} u_{t,k} + \frac{1}{\kappa}\sum_{k=k^*+1}^{\kappa} E_{t,k^*}^Q\left[u_{t,k}\right]. \end{aligned}$$

Notice that $u_{t,k} = r_{t,k} - \overline{r}$ is observed in the data on a daily basis.

Let $\mu = (k^* - 1)/\kappa$, that is, the proportion of the month before the target changes. Let

$$v_{t,k^*} = \frac{1}{\kappa}\sum_{k=1}^{k^*} u_{t,k} + \frac{1}{\kappa}\sum_{k=k^*+1}^{\kappa} E_{t,k^*}^Q\left[u_{t,k}\right]$$

then it is easy to see that by solving for the optimal predictor[1] of the target change, $E_{t,k^*}^Q\left[\Delta\right]$, we obtain

$$s_t^1 \equiv E_{t,k^*}^Q\left[\Delta\right] = \frac{1}{1-\mu}(f_{t,k^*}^0 - \overline{r}) - \frac{1}{1-\mu}v_{t,k^*}.$$

The first term is simply the difference between the federal funds rate futures price minus the target before the change, scaled to reflect where within the month the target change took place. The second term is usually close to zero. For any day of the month before the target change, $u_{t,k}$ is directly observable from $u_{t,k} = r_{t,k} - \overline{r}$. But for the remaining days in the month, we need to construct a model for $E_{t,k^*}^Q\left[u_{t,k}\right]$. Following Hamilton (2008), we assume that $u_{t,k}$ follows an AR(1) process (we implicitly assume that professional investors are risk-neutral as is customary and therefore do not add a correction for unknown risk aversion) and include a rich set of dummies, one for each day of the maintenance period, so as to capture weekend effects (reserves held on Fridays count for Saturdays and Sundays therefore inducing more volatility on the federal funds rate on Fridays), as well as end-of-the-maintenance-period effects (when banks may be willing to pay extra to make up for any reserve shortfalls before the maintenance

[1]Optimality is in terms of a mean square criterion under the risk neutral measure.

period is over). Following Hamilton (2008), we also include a dummy for the last day of the month and for the last day of the calendar year.

II Inference

This part of the appendix contains a detailed description of our estimators, confidence intervals and test statistics. The limiting distribution of the estimators and specification tests as well as a detailed discussion of the regularity conditions is given. We demonstrate that our regularity conditions imply the conditions in Newey and West (1994), justifying the use of their robust standard error estimator.

1 Setup

For ease of reference we repeat a number of definitions from the main paper. The identification restriction is:

Condition 1 *Selection on observables:*

$$y_{t,l}^{\psi}(d_j) \perp D_t | z_t \textit{ for all } l \geq 0 \textit{ and for all } d_j, \textit{ with } \psi \textit{ fixed; } \psi \in \Psi.$$

Let

$$\delta_{t,j}(\psi) = \delta_{t,j}(z_t, \psi) = \frac{1\{D_t = d_j\}}{p^j(z_t, \psi)} - \frac{1\{D_t = d_0\}}{p^0(z_t, \psi)}$$

and define the residual weights as $\ddot{\delta}_{t,j} = \delta_{t,j}(\widehat{\psi}) - \widehat{\delta}_{t,j}$ where $\widehat{\delta}_{t,j}$ is the predicted value formed from a regression of $\delta_{t,j}(\widehat{\psi})$ on z_t, the variables included in the propensity score model. Define $\widehat{h}_{j,t} = Y_{t,L}\ddot{\delta}_{t,j}$ and hence $\hat{h}_t = (\widehat{h}_{1,t}', ..., \widehat{h}_{J,t})'$. Therefore,

$$\hat{\theta} = T^{-1} \sum\nolimits_{t=1}^{T} \hat{h}_t. \tag{1}$$

The estimator $\hat{\theta}$ can also be obtained as the solution to the following minimum distance problem:

$$\hat{\theta} = \arg\min_{\theta} \left(T^{-1} \sum\nolimits_{t=1}^{T} \hat{h}_t - \theta\right)' \Omega^{-1} \left(T^{-1} \sum\nolimits_{t=1}^{T} \hat{h}_t - \theta\right), \tag{2}$$

Below we discuss estimates of the spectral density of $\hat{h}_t$ that take into account first step estimation of ψ. First note that our estimates of the optimal Ω are equivalent to estimates of the optimal weight matrix given in Hansen (2008, Section 4.2).

Assume $\hat{\psi}$ is the maximum likelihood estimator with representation

$$T^{1/2}\left(\hat{\psi} - \psi\right) = \Omega_{\psi}^{-1} T^{-1/2} \sum_{t=1}^{T} l(D_t, z_t, \psi_0) + o_p(1) \tag{3}$$

where $\Omega_\psi = E\left[l(D_t, z_t, \psi_0) l(D_t, z_t, \psi_0)'\right]$ and the function

$$l(D_t, z_t, \psi) = \sum_{j=0}^{J} \frac{1\left\{D_t = d_j\right\}}{p_t^j\left(z_t, \psi\right)} \frac{\partial p_t^{d_j}\left(z_t, \psi\right)}{\partial \psi}$$

is the score of the maximum likelihood estimator. Define the population projection π_y as

$$\pi_y = \arg\min_b E\left[\left\|Y_{t,L} - b z_t\right\|^2\right],$$

define $\vartheta = \left(\psi', (\operatorname{vec} \pi_y)'\right)'$ and let $h_t\left(\vartheta_0\right) = \left(Y_{t,L} - \pi_y z_t\right) \delta_{t,j}\left(\psi_0\right)$. The representation in (3) is used to expand $\hat{h}_t$ around ψ_0 leading to $\hat{\theta} - \theta_0 = T^{-1} \sum_{t=1}^{T} v_t\left(\vartheta_0\right) + o_p\left(T^{-1/2}\right)$ where $v_t\left(\vartheta_0\right) = h_t\left(\vartheta_0\right) - \theta_0 + \dot{h}(\vartheta_0)\Omega_\psi^{-1} l(D_t, z_t, \psi_0)$ and $\dot{h}(\vartheta_0) = E\left[\partial h_t\left(\vartheta_0\right) / \partial \psi'\right]$. The covariance matrix Ω_θ is the typical spectrum at frequency zero matrix of $v_t\left(\vartheta_0\right)$ found in the HAC-standard error literature (see Newey and West (1994)) and is given by

$$\Omega_\theta = \sum_{i=-\infty}^{\infty} E\left[v_t\left(\vartheta_0\right) v_{t-i}\left(\vartheta_0\right)'\right] \tag{4}$$

The formula for Ω_θ takes into account that the 'observations' $\hat{h}_t$ used to compute the sample averages are based on estimated, rather than observed data. Confidence intervals for θ can be constructed from Ω_θ. We use the procedure in Newey and West (1994) to estimate Ω_θ. Below, we provide further details regarding regularity conditions needed for the Newey West procedure.

Using $\hat{\vartheta} = \left(\hat{\psi}', (\operatorname{vec} \hat{\pi}_y)'\right)'$ where $\hat{\pi}_y$ is the OLS estimator in a regression of $Y_{t,L}$ on z_t we estimate Ω_θ from the sample averages

$$\widehat{\dot{h}}(\hat{\vartheta}) = T^{-1} \sum_{t=1}^{T} \partial h_t\left(\hat{\vartheta}\right) / \partial \psi', \;\; \hat{\Omega}_\psi = -T^{-1} \sum_{t=1}^{T} \frac{\partial l(D_t, z_t, \hat{\psi})}{\partial \psi'}$$

and by letting $v_t\left(\hat{\vartheta}\right) = h_t\left(\hat{\vartheta}\right) - \hat{\theta} + \widehat{\dot{h}}(\hat{\vartheta})\hat{\Omega}_\psi^{-1} l(D_t, z_t, \hat{\psi})$. As in Newey and West (1994), we use the Bartlett kernel with prewhitening and a data-dependent plug in estimator to obtain the necessary bandwidth parameter.

The Newey and West procedure is implemented as follows. Prewhitening is achieved by fitting a AR(1) model to each element $v_{t,j}\left(\hat{\vartheta}\right)$ of $v_t\left(\hat{\vartheta}\right)$. For this purpose define the autoregressive parameter estimate

$$\hat{A}_{jj} = \sum_{t=2}^{T} v_{t,j}\left(\hat{\vartheta}\right) v_{t-1,j}\left(\hat{\vartheta}\right)' \left(\sum_{t=2}^{T} v_{t-1,j}\left(\hat{\vartheta}\right) v_{t-1,j}\left(\hat{\vartheta}\right)'\right)'$$

and let $\hat{r}_t\left(\hat{\vartheta}\right) = v_t\left(\hat{\vartheta}\right) - \hat{A} v_{t-1}\left(\hat{\vartheta}\right)$ where $\hat{A}$ is a diagonal matrix with diagonal elements $\hat{A}_{jj}$. Then define $\hat{\Omega}_{\theta,j} = (T-1)^{-1} \sum_{t=j+1}^{T} \hat{r}_t\left(\hat{\vartheta}\right) \hat{r}_{t-j}\left(\hat{\vartheta}\right)'$ for $j \geqq 0$ and $\hat{\Omega}_{\theta,j} = \hat{\Omega}_{\theta,-j}'$ for $j < 0$. Let $\mathbf{1} = [1, ..., 1]'$ be an r-dimensional vector where r is the dimension of θ. Define $\hat{\sigma}_j = \mathbf{1}' \hat{\Omega}_{\theta,j} \mathbf{1}$, $\hat{s}^{(q)} = \sum_{j=-n}^{n} |j|^q \hat{\sigma}_j$ and

$\hat{\gamma} = c_\gamma \left(\hat{s}^{(1)}/\hat{s}^{(0)}\right)^{2/3}$ where[2] $c_\gamma = 1.1447$ and $n = \left\lfloor 4\left(T/100\right)^{2/9} \right\rfloor$ where $\lfloor . \rfloor$ denotes the integer part of a real number. Set the bandwidth parameter to $\hat{B} = \left\lfloor \hat{\gamma} T^{1/3} \right\rfloor$.

The estimator for Ω_θ is now defined as

$$\hat{\Omega}_\theta = \left(I_r - \hat{A}\right)^{-1} \left(\hat{\Omega}_{\theta,0} + \sum_{j=1}^{\hat{B}} \left(1 - \frac{j}{\hat{B}+1}\right) \left(\hat{\Omega}_{\theta,j} + \hat{\Omega}'_{\theta,j}\right)\right) \left(I_r - \hat{A}\right)^{-1}.$$

An important diagnostic for our purposes looks at whether lagged macro aggregates are independent of policy changes conditional on the policy propensity score. In other words we would like to show that the policy shocks implicitly defined by our score model look to be "as good as randomly assigned." Angrist and Kuersteiner (2011) develop semiparametric tests that can be used for this purpose.

The specification tests are based on the following fact. If w_t is a vector of k_w elements of z_t or χ_{t-1}, then correct specification of the propensity score implies that

$$E\left[\delta_{t,j}\left(\psi_0\right)|w_t\right] = 0 \text{ for all } j = 1, ..., J.$$

All J conditional moment restrictions, or a subset of them, can be summarized into a vector. Let $\mathcal{D}_t\left(z_t, \psi\right) = \left(\delta_{t,j_1}\left(\psi\right), ..., \delta_{t,j_k}\left(\psi\right)\right)$. Set $k \leq J$ and $1 \leq j_1 < ... < j_k \leq J$. In our case, we use this setup to focus on $d_j = \{-.25, 0, .25\}$. Then, $E\left[\mathcal{D}_t\left(z_t, \psi_0\right)|w_t\right] = 0$ must hold. To test this condition, consider the unconditional moment restriction $E\left[\mathcal{D}_t(z_t, \psi_0) \otimes w_t\right] = 0$. Since our estimators are based on $\ddot{\delta}_{t,j}$we similarly define our test based on $\ddot{\delta}_{t,j}$. For this purpose, let $\ddot{\mathcal{D}}_t\left(z_t, \psi\right) = \left(\ddot{\delta}_{t,j_1}\left(\psi\right), ..., \ddot{\delta}_{t,j_k}\left(\psi\right)\right)$ and consider the test statistic $T^{-1/2}\sum_{t=1}^T \ddot{\mathcal{D}}_t(z_t, \hat{\psi}) \otimes w_t$. Let π_w be the population projection parameter of a projection of w_t onto z_t, and $\hat{\pi}_w$ the corresponding sample OLS estimator. Define $\xi = \left(\psi', \operatorname{vec}\left(\pi_w\right)'\right)'$, let $m_t\left(\xi\right) = \left(\mathcal{D}_t(z_t, \psi)\right) \otimes \left(w_t - \pi_w z_t\right)$ and define $\bar{m}\left(\xi\right) = T^{-1}\sum_{t=1}^T m_t\left(\xi\right)$. It then follows that $T^{-1}\sum_{t=1}^T \ddot{\mathcal{D}}_t(z_t, \hat{\psi}) \otimes w_t = \bar{m}\left(\hat{\xi}\right)$ where $\hat{\xi} = \left(\hat{\psi}', \operatorname{vec}\left(\hat{\pi}_w\right)'\right)'$ and we base our statistic on $\bar{m}\left(\hat{\xi}\right)$. The limiting distribution of $\bar{m}\left(\hat{\xi}\right)$ is affected by the fact that ψ_0 is estimated. Define $\dot{m}\left(\xi\right) = E\left[\partial m_t\left(\xi\right)/\partial \psi'\right]$, $\hat{m}_t = m_t\left(\hat{\xi}\right)$ and consider the expansion

$$\hat{m}_t = m_t\left(\xi_0\right) + \dot{m}\left(\xi_0\right)\Omega_\psi^{-1} l(D_t, z_t, \psi_0) + o_p\left(T^{-1/2}\right).$$

A key insight is that under the null-hypothesis, $\hat{m}_t$ is approximately a martingale difference sequence and thus is mean zero. This feature significantly simplifies estimation of the asymptotic variance normalizing the test. Then, letting $\bar{m} = \bar{m}\left(\hat{\xi}\right)$, $\nu_t\left(\xi_0\right) = m_t\left(\xi_0\right) + \dot{m}\left(\xi_0\right)\Omega_\psi^{-1} l(D_t, z_t, \psi_0)$ and $\hat{V} = T^{-1}\sum_{t=1}^T \nu_t\left(\hat{\xi}\right)\nu_t\left(\hat{\xi}\right)'$ leads to the test statistic

$$T\bar{m}'\hat{V}^{-1}\bar{m} \rightarrow_d \chi^2_{(k \cdot k_w)} \tag{5}$$

[2] See Newey and West (1994, Tables I and II).

under the null hypothesis that $E\left[\mathbf{1}\left\{D_t = j\right\}|z_t\right] = p^j(z_t, \psi_0)$. The limiting distribution in (5) is established below.

2 Regularity Conditions

Assume that $\{\chi_t\}_{t=-\infty}^{\infty}$ is strictly stationary with values in the measurable space $(\mathbb{R}^r, \mathcal{B}^r)$ where $\mathcal{B}^r$ is the Borel σ-field on $\mathbb{R}^r$ and r is fixed with $2 \leq r < \infty$. Let $\mathcal{A}_k^l = \sigma\left(\chi_k, ..., \chi_l\right)$ be the sigma field generated by $\chi_k, ..., \chi_l$. The sequence χ_t is φ-mixing if

$$\varphi_m = \sup_l \left[\sup_{A \in \mathcal{A}_{l+m}^{\infty}, B \in \mathcal{A}_{-\infty}^{l}, P(B)>0} \left| \Pr\left(A|B\right) - P\left(A\right) \right| \right] \to 0 \text{ as } m \to \infty.$$

Condition 2 *Let χ_t be a stationary, φ-mixing sequence such that for some $2 < p < \infty$ the φ-mixing coefficient of χ_t satisfies $\varphi_m \leq cm^{-\frac{1+p}{p-4/p}}$ for some bounded constant $c > 0$. For each element $\chi_{t,j}$ of χ_t it follows that $E\left[\left|\chi_{t,j}\right|^p\right] < \infty$.*

Condition 2 implies that $\sum_{m=1}^{\infty} \varphi_m^{1-1/p} < \infty$ as required for Corollary 3.9 of McLeish (1975a). In addition, φ_m satisfies (2.6) of McLeish (1975b) required for a strong law of large numbers. This follows because for any $p > 2$ the inequality $p/(p-2) < (1+p)/(p-4/p)$ holds and, since $p > 2$, the moment restrictions imposed below are stronger than required by McLeish. Using Corollary A.2 of Hall and Heyde (1980), and assuming that, for each element $v_{t,j}\left(\vartheta_0\right)$ of $v_t\left(\vartheta_0\right)$, $E\left[\left|v_{t,j}\left(\vartheta_0\right)\right|^p\right] < \infty$ it also follows that $\sum_{m=1}^{\infty} |m|^q \left\|E\left[v_t\left(\vartheta_0\right) v_{t-m}\left(\vartheta_0\right)'\right]\right\| < \infty$ for some $q > 7/4$ as required by Assumption 2 of Newey and West (1994) when the Bartlett kernel is used. If the size of the mixing coefficients is weakened to $-\left(1+p\right)/\left(p-2/p\right)$ then Assumption 2 of Newey and West holds for all $p > 2 + \sqrt{6}$ and some $q > 7/4$. Also note that $p > 2$ is sufficient to satisfy Assumption 3 of Newey and West (1994) when the Bartlett kernel is used as suggested here.

The next condition states that the propensity score $p(z_t, \theta)$ is the correct parametric model for the conditional expectation of D_t and lists a number of additional regularity conditions.

Condition 3 *Let Θ be a compact subset of $\mathbb{R}^{k_\vartheta}$ where k_ϑ is the dimension of ϑ. Let $\psi_0 \in \Psi \subset \Theta$ where $\Psi \subset \mathbb{R}^{k_\psi}$ is a compact set and $k_\psi < \infty$. Assume that $E\left[1\left\{D_t = d_j\right\}|z_t\right] = p_t^j\left(z_t, \psi_0\right)$ and for all $\psi \neq \psi_0$ it follows $E\left[1\left\{D_t = d_j\right\}|z_t\right] \neq p^j(z_t|\psi)$. Assume that $p^j(z_t|\psi)$ is differentiable a.s. for $\vartheta \in \left\{\vartheta \in \Theta | \left\|\vartheta - \vartheta_0\right\| \leq \delta\right\} := N_\delta(\vartheta_0)$ for some $\delta > 0$. Let $N(\vartheta_0)$ be a compact subset of the union of all neighborhoods $N_\delta\left(\vartheta_0\right)$ where $\partial p^j(z_t|\psi)/\partial\psi$, $\partial^2 p^j(z_t|\psi)/\partial\psi_i\partial\psi_j$ exists and assume that $N(\vartheta_0)$ is not empty. Assume that for all $j \in \{0, ..., J\}$ and some $\delta_0 > 0$ and any $\delta > 0$, ϑ, ϑ^* with $\left\|\vartheta - \vartheta^*\right\| < \delta \leq \delta_0$ there exists a random variable B_t which is a measurable function of D_t, z_t and $Y_{t,L}$ and a constant $\alpha > 0$ such that for all i*

$$\left\|h_{t,j}\left(\vartheta\right) - h_{t,j}\left(\vartheta^*\right)\right\| \leq B_t \left\|\vartheta - \vartheta^*\right\|^{\alpha},$$

and

$$\begin{aligned}
\left\| \partial h_{t,j}\left(\vartheta\right)/\partial\vartheta - \partial h_{t,j}\left(\vartheta^{*}\right)/\partial\vartheta \right\| &\leq B_{t}\left\|\vartheta-\vartheta^{*}\right\|^{\alpha} &(6)\\
\left\| \partial^{2} h_{t,j}\left(\vartheta\right)/\partial\vartheta\partial\vartheta' - \partial^{2} h_{t,j}\left(\vartheta^{*}\right)/\partial\vartheta\partial\vartheta' \right\| &\leq B_{t}\left\|\vartheta-\vartheta^{*}\right\|^{\alpha} &(7)\\
\left\| z_{t}\left(\delta_{t,j}\left(\psi\right)-\delta_{t,j}\left(\psi^{*}\right)\right) \right\| &\leq B_{t}\left\|\psi-\psi^{*}\right\|^{\alpha} &(8)
\end{aligned}$$

and $\vartheta, \vartheta^{*} \in \operatorname{int} N\left(\vartheta_{0}\right)$*. Let* $h_{t,j,i}\left(\vartheta\right)$ *be the i-th element of* $h_{t,j}\left(\vartheta\right)$ *and* ϑ_{k} *the k-th element of* ϑ*. Assume* $E\left[\left|B_{t}\right|^{p}\right]<\infty$*, and for all* i,j,k *that* $E\left[\left|h_{t,j,i}\left(\vartheta_{0}\right)\right|^{p}\right]<\infty$*,* $E\left[\left|\partial h_{t,j,i}\left(\vartheta_{0}\right)/\partial\vartheta_{k}\right|^{p}\right]<\infty$*, and*

$$E\left[\left|\partial^{2} h_{t,j,i}\left(\vartheta_{0}\right)/\left(\partial\vartheta_{k}\partial\vartheta_{k'}\right)\right|^{p}\right]<\infty.$$

Condition 4 *Assume that* $\hat{\vartheta}-\vartheta_{0}=o_{p}\left(1\right), T^{1/2}\left(\hat{\psi}-\psi_{0}\right)=\Omega_{\psi}^{-1}T^{-1/2}\sum_{t=1}^{T} l(D_{t},z_{t},\psi_{0})+o_{p}\left(1\right).$ *Assume that* $E\left[z_{t}z_{t}'\right]$ *is positive definite. Let* $l_{i}(D_{t},z_{t},\psi_{0})$ *be the i-th element of* $l(D_{t},z_{t},\psi)$*. Let* p *be given as in Condition 2 and assume that* $E\left[\left\|l(D_{t},z_{t},\psi_{0})\right\|^{p}\right]<\infty$*,* $\sup_{\psi\in N(\vartheta_{0})}\left\|l(D_{t},z_{t},\psi)\right\|\leq B_{t}$*,*

$$\sup_{\psi\in N(\vartheta_{0})}\left\|\partial l(D_{t},z_{t},\psi)/\partial\psi\right\|\leq B_{t}$$

and $\sup_{\psi\in N(\vartheta_{0})}\left\|\partial^{2} l_{i}(D_{t},z_{t},\psi)/\partial\psi\partial\psi'\right\|\leq B_{t}$*.*

Condition 5 *Assume that* Ω_{ψ} *is positive definite for all* ψ *in some neighborhood* $N\subset\Psi$ *such that* $\psi_{0}\in\operatorname{int} N$ *and* $0<\left\|\Omega_{\psi}\right\|<\infty$ *for all* $\psi\in N$*. Assume that* Ω_{θ} *defined in (4) is positive definite.*

Conditions 2, 3 and 4 imply that Assumption 2 of Newey and West is satisfied. The results of their paper thus apply to the estimates of Ω_{θ} proposed here.

Regularity conditions for the specification tests are given below.

Condition 6 *Let* $N(\xi_{0})$ *a neighborhood of* ξ_{0} *defined similarly to the one in Condition 3. Let* p *be given as in Condition 2. For some random variable* B_{t} *which is a measurable function of* D_{t}, z_{t} *and* w_{t} *and for which* $E\left[B_{t}^{p}\right]<\infty$*, it holds that for some* $\varepsilon>0$ *and* ξ,ξ^{*} *with* $\left\|\xi-\xi^{*}\right\|<\delta\leq\delta_{0}$ *and* $\xi,\xi^{*}\in\operatorname{int} N\left(\xi_{0}\right)$ *that*

i) $E\left[\left\|m_{t}\left(\xi_{0}\right)\right\|^{p+\varepsilon}\right]<\infty$*,* $E\left[\left\|\partial m_{t}\left(\xi_{0}\right)/\partial\xi'\right\|^{p+\varepsilon}\right]<\infty$*,* $E\left[\left\|l(D_{t},z_{t},\psi_{0})\right\|^{p+\varepsilon}\right]<\infty$

ii) $\left\|l(D_{t},z_{t},\psi)-l(D_{t},z_{t},\psi^{*})\right\|\leq B_{t}\left\|\psi-\psi^{*}\right\|^{\alpha},$

iii) $\left\|\partial m_{t}\left(\xi\right)/\partial\xi'-\partial m_{t}\left(\xi^{*}\right)/\partial\xi'\right\|\leq B_{t}\left\|\xi-\xi^{*}\right\|^{\alpha}.$

3 Proofs

The following theorem establishes consistency and the limiting distribution of our estimator.

Theorem 1 *Let $\hat{\theta}$ be defined in (1) and assume that Conditions 1, 2, 3, 4, and 5 hold. Then, $\hat{\theta} \to_p \theta$ and*

$$T^{1/2}\left(\hat{\theta}-\theta\right) \xrightarrow{d} N\left(0, \Omega_\theta\right)$$

where Ω_θ is defined in (4).

Proof. Let $Z = (z_1, ..., z_T)'$, $Y_L = (Y_{1,L}, ..., Y_{T,L})'$ and $\delta_j\left(\hat{\psi}\right) = \left(\delta_{1,j}\left(\hat{\psi}\right), ..., \delta_{T,j}\left(\hat{\psi}\right)\right)'$. Define the population projection π_y as $\pi_y = \arg\min_b E\left[\|Y_{t,L} - bz_t\|^2\right]$ and sample analog $\hat{\pi}_y = Y_L' Z\left(Z'Z\right)^{-1}$. Recall that $\hat{h}_{t,j} = Y_{t,L}\left(\delta_{t,j}\left(\hat{\psi}\right) - \hat{\delta}_{t,j}\right)$ where $\hat{\delta}_{t,j} = z_t'(Z'Z)^{-1}Z'\delta_j\left(\hat{\psi}\right)$ and let $h_{t,j}\left(\vartheta_0\right) = \left(Y_{t,L} - \pi_y z_t\right)\delta_{t,j}\left(\psi_0\right)$. First observe that

$$\begin{aligned}
\sum_{t=1}^{T} \hat{h}_{t,j} &= \sum_{t=1}^{T} Y_{t,L}\left(\delta_{t,j}\left(\hat{\psi}\right) - \hat{\delta}_{t,j}\right) \\
&= \sum_{t=1}^{T} Y_{t,L}\delta_{t,j}\left(\hat{\psi}\right) - \sum_{t=1}^{T} Y_{t,L}z_t'(Z'Z)^{-1}\sum_{s=1}^{T} z_s'\delta_{s,j}\left(\hat{\psi}\right) \\
&= \sum_{t=1}^{T} Y_{t,L}\delta_{t,j}\left(\hat{\psi}\right) - \hat{\pi}_y\sum_{s=1}^{T} z_s'\delta_{s,j}\left(\hat{\psi}\right) \\
&= \sum_{t=1}^{T}\left(Y_{t,L} - \hat{\pi}_y z_t'\right)\delta_{t,j}\left(\hat{\psi}\right).
\end{aligned}$$

By the Mean Value Theorem we then obtain

$$\begin{aligned}
T^{1/2}\left(\hat{\theta}_j - \theta_{0,j}\right) &= T^{-1/2}\sum_{t=1}^{T}\hat{h}_{t,j} - \theta_{0,j} \qquad (9)\\
&= T^{-1/2}\sum_{t=1}^{T}\left(Y_{t,L} - \pi_y z_t\right)\delta_{t,j}\left(\hat{\psi}\right) - \theta_0 + \left(\pi_y - \hat{\pi}_y\right)T^{-1/2}\sum_{t=1}^{T} z_t\delta_{t,j}\left(\hat{\psi}\right) \\
&= T^{-1/2}\sum_{t=1}^{T} h_{t,j}\left(\vartheta_0\right) - \theta_0 + T^{-1}\sum_{t=1}^{T}\partial h_{t,j}\left(\vartheta_0\right)/\partial\psi' T^{1/2}\left(\hat{\psi} - \psi_0\right) \\
&\quad + T^{-1}\sum_{t=1}^{T}\left(\partial h_{t,j}\left(\check{\vartheta}\right)/\partial\psi' - \partial h_{t,j}\left(\vartheta_0\right)/\partial\psi'\right)T^{1/2}\left(\hat{\psi} - \psi_0\right) \\
&\quad + \left(\pi_y - \hat{\pi}_y\right)T^{-1/2}\sum_{t=1}^{T} z_t\delta_{t,j}\left(\hat{\psi}\right)
\end{aligned}$$

where $\left\|\check{\vartheta} - \vartheta_0\right\| \leq \left\|\hat{\vartheta} - \vartheta_0\right\|$ and $\partial h_t\left(\vartheta\right)/\partial\psi' = \left[\partial h_{t,1}\left(\vartheta\right)/\partial\psi', ..., \partial h_{t,J}\left(\vartheta\right)/\partial\psi'\right]$ with

$$\partial h_{t,j}\left(\vartheta\right)/\partial\psi = \left(Y_{t,L} - \pi_y z_t\right)\left(-\frac{D_{t,j}}{p^j\left(z_t,\psi\right)^2}\frac{\partial p^j\left(z_t,\psi\right)}{\partial\psi} + \frac{D_{t,0}}{p^0\left(z_t,\psi\right)^2}\frac{\partial p^0\left(z_t,\psi\right)}{\partial\psi}\right). \qquad (10)$$

By (6) it follows that for δ_0 given in Condition 3 and any δ such that $\delta_0 > \delta > 0$,

$$\begin{aligned}
& P\left(\left\|T^{-1}\sum_{t=1}^{T}\left(\partial h_{t,j}\left(\check{\vartheta}\right)/\partial\psi' - \partial h_{t,j}\left(\vartheta_0\right)/\partial\psi'\right)\right\| > \eta\right) \qquad (11)\\
\leq\; & P\left(\sup_{\|\vartheta-\vartheta_0\|\leq\delta}\left\|T^{-1}\sum_{t=1}^{T}\left(\partial h_{t,j}\left(\vartheta\right)/\partial\psi' - \partial h_{t,j}\left(\vartheta_0\right)/\partial\psi'\right)\right\| > \eta, \left\|\check{\vartheta} - \vartheta_0\right\| < \delta\right) + P\left(\left\|\check{\vartheta} - \vartheta_0\right\| \geq \delta\right) \\
=\; & \frac{E\left[\left|B_t\right|^p\right]\delta^{p\alpha}}{\eta^p} + P\left(\left\|\check{\vartheta} - \vartheta_0\right\| \geq \delta\right)
\end{aligned}$$

where both terms can be made arbitrarily small by choosing $\eta = \sqrt{\delta}$ and $\delta > 0$ for T large enough by using Conditions 4 and 3. By McLeish (1975b, Theorem 2.10) $T^{-1}\sum_{t=1}^{T} \partial h_{t,j}(\vartheta_0)/\partial\psi' \xrightarrow{p} \dot{h}_j(\vartheta_0)$ where we defined $E\left[\partial h_{t,j}(\vartheta_0)/\partial\psi'\right] = \dot{h}_j(\vartheta_0)$. This implies that the third term in (9) is $o_p(1)$.

For the last term in (9) note that $(\pi_y - \hat{\pi}_y) = O_p\left(T^{-1/2}\right)$ by McLeish (1975b, Theorem 2.10), Corollary 3.9 of McLeish (1975a) and standard arguments for linear regressions. Now consider

$$\begin{aligned} & (\pi_y - \hat{\pi}_y) T^{-1/2} \sum\nolimits_{t=1}^{T} z_t \delta_{t,j}\left(\hat{\psi}\right) \\ = \; & T^{1/2}(\pi_y - \hat{\pi}_y) T^{-1}\sum\nolimits_{t=1}^{T} z_t\delta_{t,j}(\psi_0) \\ & + T^{1/2}(\pi_y - \hat{\pi}_y) T^{-1}\sum\nolimits_{t=1}^{T} z_t\left(\delta_{t,j}\left(\hat{\psi}\right) - \delta_{t,j}(\psi_0)\right). \end{aligned} \tag{12}$$

The first term in (12) is $o_p(1)$ because from $E\left[z_t\delta_{t,j}(\psi_0)\right] = 0$ it follows that

$$T^{-1}\sum_{t=1}^{T} z_t\delta_{t,j}(\psi_0) = o_p(1). \tag{13}$$

For the second term in (12) use Condition 3 to show that

$$T^{-1}\sum\nolimits_{t=1}^{T} z_t\left(\delta_{t,j}(\psi_0) - \delta_{t,j}\left(\hat{\psi}\right)\right) = o_p(1) \tag{14}$$

by arguments similar to those in (11). Then, (13) and (14) establish that (12) is $o_p(1)$. It then follows from (12) and (14) that (9) is

$$\begin{aligned} & T^{-1/2}\sum\nolimits_{t=1}^{T} h_{t,j}(\vartheta_0) - \theta_0 \\ & + T^{-1}\sum\nolimits_{t=1}^{T} \partial h_{t,j}(\vartheta_0)/\partial\psi' T^{1/2}\left(\hat{\psi} - \psi_0\right) + o_p(1) \\ = \; & T^{-1/2}\sum\nolimits_{t=1}^{T}\left[h_{t,j}(\vartheta_0) - \theta_0 + \dot{h}_j(\vartheta_0)\Omega_\psi^{-1} l(D_t, z_t, \psi_0)\right] + o_p(1). \end{aligned}$$

Stack $h_t(\vartheta) = \left[h_{t,1}(\vartheta)', ..., h_{t,J}(\vartheta)'\right]'$ and $\dot{h}(\vartheta) = \left[\dot{h}_1(\vartheta)', ..., \dot{h}_J(\vartheta)'\right]'$, let

$$v_t(\vartheta_0) = h_t(\vartheta_0) - \theta + \dot{h}(\vartheta_0)\Omega_\psi^{-1} l(D_t, z_t, \psi_0)$$

and $v_{t,j}(\vartheta_0)$ is the j-th element of $v_t(\vartheta_0)$. Note that $v_{t,j}(\vartheta_0)$ is β-mixing with $E\left[v_{t,j}(\vartheta_0)\right] = 0$. Then it follows that

$$T^{-1}E\left[\sum\nolimits_{t=1}^{\tau}\sum\nolimits_{t=s}^{\tau} v_t(\vartheta_0)\, v_s(\vartheta_0)'\right] \tag{15}$$

$$= \sum_{j=-T+1}^{T-1}\left(1 - \frac{|j|}{T}\right) E\left[v_1(\vartheta_0)\, v_{1-j}(\vartheta_0)'\right] \to \Omega_\theta \tag{16}$$

by stationarity of $v_t = v_t(\vartheta_0)$ and the Toeplitz lemma. Fix $\lambda \in \mathbb{R}^k$ with $\|\lambda\| = 1$ and let $S_T = T^{-1/2}\sum_{t=1}^{T}\lambda' v_t$. Then, $E\left[S_T^2\right] \to \lambda'\Omega_\theta\lambda > 0$ by (15) and Condition 5. In addition

$$E\left[\left|\lambda' v_t\right|^p\right] \leq E\left[\left(\sum_{l=1}^{k}|\lambda_l|\,|\tilde{v}_{t,l}|\right)^p\right] \leq \left(\sum_{l=1}^{k}|\lambda_l|^{\frac{p}{p-1}}\right)^{p-1} E\left[\sum_{l=1}^{k}|\tilde{v}_{t,l}|^p\right]$$

by Hölder's inequality (Magnus and Neudecker, 1988, p.220) and where $\tilde{v}_{t,l}$ is the l-th element of v_t. Since $p/(p-1) \leq 2$ and $\|\lambda\| = 1$ it follows that $\sum_{l=1}^{k}|\lambda_l|^{\frac{p}{p-1}} < k$. Denote by $h_{t,j}(\vartheta_0)$ and $\theta_{(j)}$ the j-th element of $h_t(\vartheta_0)$ and θ respectively and by $\dot{h}_j(\vartheta_0)$ the j-th row of $\dot{h}(\vartheta_0)$. Then,

$$\begin{aligned} E\left[|\tilde{v}_{t,j}|^p\right] &\leq E\left[\left(|h_{t,j}(\vartheta_0)| + \left|\theta_{(j)}\right| + \left\|\dot{h}_j(\vartheta_0)\right\|\left\|\Omega_\psi^{-1}\right\|\|l(D_t, z_t, \psi_0)\|\right)^p\right] \\ &\leq 3^{p-1}\left(E\left[|h_{t,j}(\vartheta_0)|^p\right] + |\theta_j|^p + \left|\dot{h}_j(\vartheta_0)\right|^p\left\|\Omega_\psi^{-1}\right\|^p\|l(D_t, z_t, \psi_0)\|^p\right) \end{aligned}$$

again by Hölder's inequality. It follows that $\left|\theta_{(j)}\right|^p \leq E\left[|h_{t,j}(\vartheta_0)|^p\right]$ by Jensen's inequality and $\left\|\Omega_\psi^{-1}\right\|^p < \infty$ by Condition 5. Similarly, $E\left[\|l(D_t, z_t, \psi_0)\|^p\right] < \infty$ by Condition 4 and

$$\left|\dot{h}_j(\vartheta_0)\right|^p \leq E\left[|\partial h_{t,j}(\vartheta_0)/\partial\psi|^p\right] < \infty$$

by Condition 3. By Condition 3 $E\left[|h_{t,j}(\vartheta_0)|^p\right] < \infty$ such that $E\left[|\tilde{v}_{t,j}|^p\right] < \infty$. These arguments together with Condition 2 show that all the conditions of Corollary 3.9 of McLeish (1975a) are satisfied. Thus, $S_T \to_d N\left(0, \lambda'\Omega_\theta\lambda\right)$. The result now follows from the Cramer-Wold theorem.

Consistency of $\hat{\theta}$ follows directly from the asymptotic distribution which implies that $T^{1/2}\left(\hat{\theta} - \theta\right) = O_p(1)$ such that $\hat{\theta} = \theta + o_p(1)$. ■

The following theorem establishes the limiting distribution of the test statistic in (5).

Theorem 2 *Assume that Conditions 2, 3, 4, 5 and 6 hold. For $\nu_t = \nu_t(\xi_0)$ let $V_t = \nu_t\nu_t' - V$ where V is a fixed, positive definite matrix. Assume that for any element $\nu_{t,i}$ of ν_t, $E\left[|\nu_{t,i}|^{p+\varepsilon}\right] < \infty$ where ε is the same as in Condition 6. Then,*

$$T\bar{m}'\hat{V}^{-1}\bar{m} \to_d \chi^2_{(k\cdot k_w)}$$

Proof. First consider $\sum_{t=1}^{T} \ddot{\mathcal{D}}_t(z_t, \hat{\psi}) \otimes w_t$ with representative element

$$\begin{aligned}
\sum_{t=1}^{T} \ddot{\delta}_{t,j}(\psi) w_t &= \sum_{t=1}^{T} \left(\delta_{t,j}\left(\hat{\psi}\right) - z_t'(Z'Z)^{-1} Z' \delta_j \left(\hat{\psi}\right)\right) w_t \\
&= \sum_{t=1}^{T} \left(\delta_{t,j}\left(\hat{\psi}\right) - \sum_{s=1}^{T} \delta_{js}\left(\hat{\psi}\right) z_s'(Z'Z)^{-1} z_t \right) w_t \\
&= \sum_{t=1}^{T} \delta_{t,j}\left(\hat{\psi}\right) w_t - \sum_{t=1}^{T} \delta_{js}\left(\hat{\psi}\right) z_s' \hat{\pi}_w' \\
&= \sum_{t=1}^{T} \delta_{t,j}\left(\hat{\psi}\right) (w_t - \hat{\pi}_w z_s).
\end{aligned}$$

Thus, the test we consider is based on $\delta_{t,j}\left(\hat{\psi}\right)(w_t - \hat{\pi}_w z_s)$. Recall $\hat{m}_t = \left(\mathcal{D}_t(z_t, \hat{\psi})\right) \otimes (w_t - \hat{\pi}_w z_t)$ such that for $m_t(\xi) = (\mathcal{D}_t(z_t, \psi)) \otimes (w_t - \pi_w z_t)$ and $m_{t,0} = m_t(\xi_0)$ and the mean value theorem it follows that

$$\hat{m}_t = m_t(\xi_0) + \partial m_t\left(\check{\xi}\right)/\partial\psi' \left(\hat{\psi} - \psi_0\right)$$

with $\left\|\check{\xi} - \xi_0\right\| \leq \left\|\hat{\xi} - \xi_0\right\|$. Using (3) as well as Condition 4 and setting $\widehat{\dot{m}}(\xi) = T^{-1}\sum_{t=1}^{T} \partial m_t(\xi)/\partial\psi'$ we obtain

$$\begin{aligned}
T^{-1/2}\sum_{t=1}^{T} \hat{m}_t &= T^{-1/2}\sum_{t=1}^{T} m_{t,0} + \widehat{\dot{m}}\left(\check{\xi}\right) \Omega_\psi^{-1} T^{-1/2} \sum_{t=1}^{T} l(D_t, z_t, \psi_0) + (\pi_w - \hat{\pi}_w) T^{-1/2} \sum\nolimits_{t=1}^{T} z_t \delta_{t,j}\left(\hat{\psi}\right) + o_p(1). \\
&= T^{-1/2}\sum_{t=1}^{T} \left(m_{t,0} + \dot{m}(\xi_0) \Omega_\psi^{-1} l(D_t, z_t, \psi_0)\right) + o_p(1)
\end{aligned}$$

where the last line follows by the same arguments as in the proof of Theorem 1. With $\nu_t(\xi_0) = m_t(\xi) + \dot{m}(\xi_0)\Omega_\psi^{-1} l(D_t, z_t, \psi_0)$ it follows from Corollary 3.9 of McLeish (1975a) that

$$T^{-1/2}\sum_{t=1}^{T} \hat{m}_t = T^{-1/2}\sum_{t=1}^{T} \nu_t(\xi_0) + o_p(1) \to^d N(0, V) \tag{17}$$

where $V = E\left[\nu_t(\xi_0)\nu_t(\xi_0)'\right]$ is a $(k \cdot k_w) \times (k \cdot k_w)$ non-singular matrix. A detailed verification of the conditions is omitted but follows the same line of argument as given in the proof of Theorem 1 above. To estimate V, define

$$\hat{\nu}_t = \hat{m}_t + \widehat{\dot{m}}\left(\hat{\xi}\right) \hat{\Omega}_\psi^{-1} l(D_t, z_t, \hat{\psi})$$

with

$$\hat{\Omega}_\psi = -T^{-1}\sum_{t=1}^{T} \frac{\partial l(D_t, z_t, \hat{\psi})}{\partial \psi'}.$$

Let

$$\hat{V} = T^{-1} \sum_{t=1}^{T} \hat{\nu}_t \hat{\nu}_t'.$$

By arguments similar to the proof of Theorem 1 it follows that

$$\hat{\Omega}_\psi \to_p \Omega_\psi \tag{18}$$

and

$$\widehat{\dot{m}}\left(\hat{\xi}\right) \to_p \dot{m}\left(\xi_0\right). \tag{19}$$

Next, expand

$$\begin{aligned}
\hat{\nu}_t &= m_{t,0} + \partial m_t\left(\check{\xi}\right)/\partial\psi'\left(\hat{\psi} - \psi_0\right) \\
&\quad + \left(\widehat{\dot{m}}\left(\hat{\xi}\right)\hat{\Omega}_\psi^{-1} - \dot{m}\left(\xi_0\right)\Omega_\psi^{-1}\right) l(D_t, z_t, \hat{\psi}) \\
&\quad + \dot{m}\left(\xi_0\right)\Omega_\psi^{-1}\left(l(D_t, z_t, \hat{\psi}) - l(D_t, z_t, \psi_0)\right) \\
&\quad + \dot{m}\left(\xi_0\right)\Omega_\psi^{-1} l(D_t, z_t, \psi_0)
\end{aligned}$$

and recalling $\nu_t = m_{t,0} + \dot{m}\left(\xi_0\right)\Omega_\psi^{-1} l(D_t, z_t, \psi_0)$. Then,

$$\left\|T^{-1}\sum_{t=1}^{T} \hat{\nu}_t\hat{\nu}_t' - V\right\| \leq \left\|T^{-1}\sum_{t=1}^{T} \left(\hat{\nu}_t\hat{\nu}_t - \nu_t\nu_t'\right)\right\| + \left\|T^{-1}\sum_{t=1}^{T} \nu_t\nu_t' - V\right\| \tag{20}$$

where the second term on the RHS of (20) is $o_p(1)$ by Theorem 2.10 of McLeish (1995b). Next, consider

$$T^{-1}\sum_{t=1}^{T} \left(\hat{\nu}_t\hat{\nu}_t' - \nu_t\nu_t'\right) = T^{-1}\sum_{t=1}^{T} \left(\hat{\nu}_t - \nu_t\right)\left(\hat{\nu}_t - \nu_t\right)' + \nu_t\left(\hat{\nu}_t - \nu_t\right)' - \left(\hat{\nu}_t - \nu_t\right)\nu_t' \tag{21}$$

where

$$\begin{aligned}
\hat{\nu}_t - \nu_t &= \partial m_t\left(\check{\xi}\right)/\partial\psi'\left(\hat{\psi} - \psi_0\right) + \left(\widehat{\dot{m}}\left(\hat{\xi}\right)\hat{\Omega}_\psi^{-1} - \dot{m}\left(\xi_0\right)\Omega_\psi^{-1}\right) l(D_t, z_t, \hat{\psi}) \\
&\quad + \dot{m}\left(\xi_0\right)\Omega_\psi^{-1}\left(l(D_t, z_t, \hat{\psi}) - l(D_t, z_t, \psi_0)\right).
\end{aligned} \tag{22}$$

Thus,

$$\begin{aligned}
T^{-1}\sum_{t=1}^{T} \nu_t\left(\hat{\nu}_t - \nu_t\right)' &= T^{-1}\sum_{t=1}^{T} \nu_t\left(\partial m_t\left(\check{\xi}\right)/\partial\psi'\left(\hat{\psi} - \psi_0\right)\right)' \\
&\quad + T^{-1}\sum_{t=1}^{T} \nu_t\left(\left(\widehat{\dot{m}}\left(\hat{\xi}\right)\hat{\Omega}_\psi^{-1} - \dot{m}\left(\xi_0\right)\Omega_\psi^{-1}\right) l(D_t, z_t, \hat{\psi})\right)' \\
&\quad + T^{-1}\sum_{t=1}^{T} \nu_t\left(\dot{m}\left(\xi_0\right)\Omega_\psi^{-1}\left(l(D_t, z_t, \hat{\psi}) - l(D_t, z_t, \psi_0)\right)\right)' \\
&\equiv R_1 + R_2 + R_3.
\end{aligned} \tag{23}$$

For R_1 note that

$$\begin{aligned} \|R_1\| &\leq \left\|T^{-1}\sum_{t=1}^{T}\nu_t\partial m_t(\xi_0)/\partial\psi'\right\|\left\|\hat{\psi}-\psi_0\right\| \\ &\quad +T^{-1}\sum_{t=1}^{T}\|\nu_t\|\left\|\partial m_t(\xi_0)/\partial\psi'-\partial m_t\left(\check{\xi}\right)/\partial\psi'\right\|\left\|\hat{\psi}-\psi_0\right\| \end{aligned} \tag{24}$$

where $\left\|\hat{\psi}-\psi_0\right\| = O_p\left(T^{-1/2}\right)$ and

$$T^{-1}\sum_{t=1}^{T}\nu_t\partial m_t(\xi_0)/\partial\psi' = O_p(1) \tag{25}$$

because

$$E\left[\left\|\nu_t\partial m_t(\xi_0)/\partial\psi'\right\|^{(p+\epsilon)/2}\right] \leq \left(E\left[\|\nu_t\|^{p+\epsilon}\right]E\left[\left\|\partial m_t(\xi_0)/\partial\psi'\right\|^{p+\epsilon}\right]\right)^{1/2} < \infty$$

by Condition 6 and by Theorem 2.10 of McLeish (1975b).[3] The second term in (24) can be bounded with probability approaching 1 as $T\to\infty$, using Condition 6(iii), and noting that

$$\left\|\partial m_t(\xi_0)/\partial\psi'-\partial m_t\left(\check{\xi}\right)/\partial\psi'\right\| \leq B_t\left\|\check{\xi}-\xi_0\right\|^{\alpha},$$

by

$$\begin{aligned} & T^{-1}\sum_{t=1}^{T}\|\nu_t\|\left\|\partial m_t(\xi_0)/\partial\psi'-\partial m_t\left(\check{\xi}\right)/\partial\psi'\right\|\left\|\hat{\psi}-\psi_0\right\| \\ \leq\ & \left\|\hat{\xi}-\xi_0\right\|^{1+\alpha}T^{-1}\sum_{t=1}^{T}\|\nu_t\||B_t| \end{aligned} \tag{26}$$

where $E\left[\|\nu_t\|^{(p+\epsilon)/2}|B_t|^{(p+\epsilon)/2}\right] \leq \left(E\left[\|\nu_t\|^{p+\epsilon}\right]E\left[|B_t|^{p+\epsilon}\right]\right)^{1/2} < \infty$ by Condition 6. This again implies that

$$T^{-1}\sum_{t=1}^{T}\|\nu_t\||B_t| = O_p(1) \tag{27}$$

by McLeish (1975b). Now (25) and (26) imply that $R_1 = o_p(1)$.

For R_2 note that using Condition 6(ii), w.p.a.1 as $T\to\infty$,

$$\begin{aligned} \|R_2\| &\leq \left\|\hat{\dot{m}}\left(\hat{\xi}\right)\hat{\Omega}_{\psi}^{-1}-\dot{m}(\xi_0)\Omega_{\psi}^{-1}\right\|T^{-1}\sum_{t=1}^{T}\|\nu_t\|\|l(D_t,z_t,\psi_0)\| \\ &\quad +\left\|\hat{\dot{m}}\left(\hat{\xi}\right)\hat{\Omega}_{\psi}^{-1}-\dot{m}(\xi_0)\Omega_{\psi}^{-1}\right\|T^{-1}\sum_{t=1}^{T}\|\nu_t\|\left\|l(D_t,z_t,\psi_0)-l\left(D_t,z_t,\hat{\psi}\right)\right\| \\ &\leq \left\|\hat{\dot{m}}\left(\hat{\xi}\right)\hat{\Omega}_{\psi}^{-1}-\dot{m}(\xi_0)\Omega_{\psi}^{-1}\right\|T^{-1}\sum_{t=1}^{T}\|\nu_t\|\|l(D_t,z_t,\psi_0\| \\ &\quad +\left\|\hat{\dot{m}}\left(\hat{\xi}\right)\hat{\Omega}_{\psi}^{-1}-\dot{m}(\xi_0)\Omega_{\psi}^{-1}\right\|T^{-1}\sum_{t=1}^{T}\|\nu_t\||B_t|\left\|\hat{\psi}-\psi_0\right\|^{\alpha} \end{aligned}$$

where $E\left[\left(\|\nu_t\|\|l(D_t,z_t,\psi_0)\|\right)^{(p+\epsilon)/2}\right] < \infty$ as before. Then, $T^{-1}\sum_{t=1}^{T}\|\nu_t\|\|l(D_t,z_t,\psi_0)\| = O_p(1)$ and (18), (19) and (27) imply that $R_2 = o_p(1)$.

[3] We use McLeish (1975), Equation (2.12) and stationarity to establish Condition (2.11) of Theorem (2.10).

For R_3 note that

$$
\begin{aligned}
& \left\| T^{-1} \sum_{t=1}^{T} \nu_t \left(\dot{m}\left(\xi_0\right) \Omega_{\psi}^{-1} \left(l(D_t, z_t, \hat{\psi}) - l(D_t, z_t, \psi_0) \right) \right)' \right\| \\
\leq & \left\| \dot{m}\left(\xi_0\right) \Omega_{\psi}^{-1} \right\| T^{-1} \sum_{t=1}^{T} \left\| \nu_t \right\| \left\| l(D_t, z_t, \hat{\psi}) - l(D_t, z_t, \psi_0) \right\| \\
& \left\| \dot{m}\left(\xi_0\right) \Omega_{\psi}^{-1} \right\| T^{-1} \sum_{t=1}^{T} \left\| \nu_t \right\| \left| B_t \right| \left\| \hat{\psi} - \psi_0 \right\|
\end{aligned}
$$

where $\left\| \hat{\psi} - \psi_0 \right\| = o_p\left(1\right)$ by Condition 4. Then, $R_3 = o_p\left(1\right)$ follows from (27). The term $T^{-1} \sum_{t=1}^{T} \left(\hat{\nu}_t - \nu_t\right) \times \left(\hat{\nu}_t - \nu_t\right)'$ in (21) can be analyzed in the same way as $T^{-1} \sum_{t=1}^{T} \nu_t \left(\hat{\nu}_t - \nu_t\right)'$ but the details are omitted. It follows that $T^{-1} \sum_{t=1}^{T} \left(\hat{\nu}_t \hat{\nu}_t' - \nu_t \nu_t'\right) = o_p\left(1\right)$ which in turn implies that

$$
\hat{V} - V = o_p\left(1\right). \tag{28}
$$

Then, for $\bar{m} = T^{-1} \sum_{t=1}^{T} \hat{m}_t$, the statistic $T\bar{m}'\hat{V}^{-1}\bar{m}$ is asymptotically $\chi^2_{(k \cdot k_w)}$ because of (17), (28) and the continuous mapping theorem. ■

References

HALL, P., AND C. HEYDE (1980): *Martingale Limit Theory and its Application.* Academic Press.

HAMILTON, J. D. (2008): "Assessing Monetary Policy Effects Using Daily Federal Funds Futures Contracts," *Federal Reserve Bank of St. Louis Review*, 90(4), 377–393.

MAGNUS, J. R., AND H. NEUDECKER (1988): *Matrix Differential Calculus with Applications in Statistics and Econometrics.* John Wiley and Sons.

MCLEISH, D. (1975a): "Invariance Principles for Dependent Variables," *Zeitschrift f. Wahrscheinlichkeitsrechnung und verw. Gebiete*, 32, 165–178.

MCLEISH, D. L. (1975b): "A Maximal Inequality and Dependent Strong Laws," *The Annals of Probability*, 3(5), 829–839.