Yale University
EliScholar – A Digital Platform for Scholarly Publishing at Yale

Cowles Foundation Discussion Papers | Cowles Foundation

12-1-2016

Structural Inference from Reduced Forms with Many Instruments

Peter C.B. Phillips

Wayne Yuan Gao

Follow this and additional works at: https://elischolar.library.yale.edu/cowles-discussion-paper-series

Part of the Economics Commons

Recommended Citation

Phillips, Peter C.B. and Gao, Wayne Yuan, "Structural Inference from Reduced Forms with Many Instruments" (2016). *Cowles Foundation Discussion Papers*. 2524.
https://elischolar.library.yale.edu/cowles-discussion-paper-series/2524

This Discussion Paper is brought to you for free and open access by the Cowles Foundation at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Cowles Foundation Discussion Papers by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

STRUCTURAL INFERENCE FROM REDUCED FORMS WITH MANY INSTRUMENTS

By

Peter C. B. Phillips and Wayne Yuan Gao

October 2016

COWLES FOUNDATION DISCUSSION PAPER NO. 2062

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Structural Inference from Reduced Forms with Many Instruments*

Peter C. B. Phillips

Yale University, University of Auckland,
Singapore Management University & University of Southampton

Wayne Yuan Gao

Yale University

October 24, 2016

Abstract

This paper develops exact finite sample and asymptotic distributions for structural equation tests based on partially restricted reduced form estimates. Particular attention is given to models with large numbers of instruments, wherein the use of partially restricted reduced form estimates is shown to be especially advantageous in statistical testing even in cases of uniformly weak instruments and reduced forms. Comparisons are made with methods based on unrestricted reduced forms, and numerical computations showing finite sample performance of the tests are reported. Some new results are obtained on inequalities between noncentral chi-squared distributions with different degrees of freedom that assist in analytic power comparisons.

Keywords: Endogeneity, Exact distributions, Partial identification, Partially restricted reduced form, Structural inference, Unidentified structure, Weak reduced form.

JEL classifications: C23, C32

*Support of the NSF under Grant SES 12-58258 is gratefully acknowledged.

1 Introduction

Instrumental variable (IV) methods are a commonly used resource in structural equation estimation and testing. While asymptotic theory is still the primary tool of inference in such cases, a substantial body of finite sample theory is now available to guide applied research. Exact finite sample distribution theory for IV estimates began with the work of Basmann (1961), who derived the distribution of the two stage least squares estimator of the coefficient in a special case of a structural equation with two endogenous variables driven by Gaussian errors. The general case was explored in Phillips (1980) who derived the exact distribution of the IV estimator of the coefficients in a structural equation with an arbitrary number (m) of endogenous variables and gave a higher order asymptotic expansion of the exact distribution using a Laplace approximation. Importantly, the exact theory holds for any configuration of strong, weak, or even irrelevant instruments – cases that are determined by the strength of the systematic component of the reduced form as measured by the reduced form parameters or, more specifically, by the matrix noncentrality parameter matrix which appears as a key element in the exact density (Phillips, 1980, equation (12)). This appealing feature has a major bearing on asymptotic theory and the quality of asymptotic approximations.

When the instruments are strong and the noncentrality matrix diverges with the sample size at the usual $O(n)$ rate that applies with stationary data, the exact distribution yields the standard $\sqrt{n}$ asymptotic normal distribution for IV estimates. When the instruments are irrelevant, the exact distribution yields the asymptotic distribution in the unidentified case where order conditions but not rank conditions hold. In this case, the IV estimator converges weakly to a random variable whose distribution is proportional to a t distribution, reflecting the uncertainty in the limit that is implicit in the lack of identification. Importantly, in this case the same asymptotic distribution holds when the Gaussian error assumption is relaxed because a martingale central limit theorem operates with respect to the sample moment components on which the IV estimator is based, as first demonstrated in Phillips (1989). Thus, an invariance principle applies in the unidentified case, just as it does in the strong IV case. An entirely analogous argument based on a noncentral version of the same martingale central limit theorem shows that an invariance principle applies in the weak IV case where the reduced form parameters are local to the origin at a $\sqrt{n}$ rate, so that the limit distribution is simply the exact finite sample distribution (under Gaussian errors) upon simple rescaling of the noncentrality matrix without requiring Gaussianity.

The case of many weak instrument asymptotics may also be obtained quite simply from the exact theory. As the number of instruments K grows, while m remains fixed, then a martingale CLT enables use of the exact results to deliver the appropriate asymptotics under conditions on the rate of expansion of K that ensure consistency of the estimator is retained. These various uses of the exact theory in conjunction with a suitable martingale CLT are explored in Phillips (Forthcoming). The wider literature on weak instrumentation and many instruments is extensive and is not reviewed here. Readers are referred to Andrews and Stock (2005) for an overview of some aspects of that literature, focusing on cases where instrument weakness is induced by localizing coefficients to the origin, as in Staiger and Stock (1997). Other approaches to weak instruments are possible and some alternatives are considered in recent work by Andrews and Guggenberger (Forthcoming) and Phillips (2006; Forthcoming).

Work on structural parameter testing dates back to Anderson and Rubin (1949) and much of the recent literature deals with test statistics that are robust to the strength of instruments. In particular, Kleibergen (2002) constructed a test statistic (the so-called K-statistic) whose limit distribution is chi-squared with degrees of freedom m matching the dimension of the structural parameter, irrespective of the strength and number of the instruments. This reduction in degrees of freedom is a feature of the test statistic that is proposed in the present paper, although the mechanism by which this is accomplished differs. Bekker and Kleibergen (2003) later characterized many instrument asymptotic theory for the K-statistic. Moreira (2003) provided an alternate method of constructing test statistics, particularly a conditional likelihood ratio (CLR) test, that are robust under weak instrumentation, and Andrews, Moreira, and Stock (2006) developed theoretical results on the power envelope within the class of invariant similar tests.

While most exact theory deals with estimation, some recent work has considered restricted reduced form estimation that incorporates information from the structural system. The reduced form is particularly important for forecasting and it seems natural to import structural information into forecasts constructed from the reduced form. Phillips (Forthcoming) derives the exact distribution of forecasts obtained from the partially restricted reduced form (Kakwani and Court, 1972), which carries restrictions from a structural equation that is estimated by IV. The primary effect of importing such structural restrictions into forecasts is to reduce variance in the forecasts. It turns out that even when the structural equation is unidentified, shrinkage still occurs and is generally beneficial in concentrating the forecast distribution and in reducing forecast mean squared error, although this is not universally so, as noted in Kakwani and Court (1972).

Since the reduced form parameters satisfy identifiability relations, these parameters are also useful in testing hypotheses about the structural parameters. The present paper explores this approach to inference. We focus on the identifying relation embedded in the partially restricted reduced form equation, and investigate both the exact finite-sample theory and the asymptotics of the partially restricted reduced-form estimators, which we then use to construct statistics for hypothesis testing about the structural parameters. The most closely related work to the present paper is Chernozhukov and Hansen (2008), who examined the unrestricted reduced-form estimator as a vehicle for structural parameter testing.

The paper's main contribution is to develop both exact finite sample and asymptotic distributions for structural equation tests based on partially restricted reduced form (PRRF) estimates. This approach is shown to be especially advantageous in statistical testing when there are large numbers of instruments and this remains so even in cases of uniformly weak instruments and reduced forms. Our main finding is that the PRRF is useful in raising the power of structural parameter tests when the number of instruments is large especially, but not always, when the instruments are strong and when the focus is on testing whether the structural coefficients are zero[1]. Comparisons are made with tests based on unrestricted reduced forms. Some numerical calculations reporting finite sample performance of these tests

[1]In unstandardized systems, this hypothesis corresponds to absence of endogeneity.

are reported. The paper also contributes by providing new analytic results on inequalities for tail probabilities of noncentral chi-squared distributions with different degrees of freedom. These results assist in making analytic power comparisons between PRRF and unrestricted reduced form (URRF) procedures.

The paper is organized as follows. The model, identifiability relations and reduced form estimates are given in Section 2. Section 3 develops exact and asymptotic distributions of the estimators under standardizing transformations of the model. Section 4 considers hypothesis tests constructed using unrestricted and partially restricted reduced form estimates. Section 5 provides limit theory for large numbers of instruments. Section 6 gives extensions for unstandardized cases and Section 7 concludes. Proofs are provided in the Appendix.

2 Estimation with Instrumental Variables

We consider a single regression equation in the following structural form

$$y_1 = Y_2\beta + u \tag{1}$$

where $[y_1, Y_2]$ is an $n \times (m+1)$ matrix of endogenous variables, u is a vector of structural errors, and β is an $m \times 1$ vector of structural parameters. Extensions to structural equations with included exogenous variables are straightforward and are not considered in what follows for notational simplicity. Let Z be a $n \times K$ matrix of (exogenous) instruments with order condition $K \geq m$ satisfied. The associated reduced form for (1) then has the form

$$Y := [y_1, Y_2] = Z[\pi_1, \Pi_2] + [v_1, V_2] \equiv Z\Pi + V. \tag{2}$$

where $\Pi = [\pi_1, \Pi_2]$ is the reduced form parameter matrix and $V = [v_1, V_2]$ is the reduced form error matrix. The restrictions imposed by the structural equation (1) on the reduced form are

$$\begin{aligned} \pi_1 &= \Pi_2\beta, \\ v_1 &= u + V_2\beta. \end{aligned} \tag{3}$$

Let $\hat{\Pi} = \left(\hat{\pi}_1, \hat{\Pi}_2\right)$ be the unrestricted reduced form (URRF) least squares estimate of Π with $\hat{\pi}_1 = (Z'Z)^{-1} Z'y_1$ and $\hat{\Pi}_2 = (Z'Z)^{-1} Z'Y_2$.

Exploiting the restrictions from the structural form implied by the identifiability relations (3), we may transform the reduced-form equation (2) to the partially restricted form (Kakwani and Court, 1972):

$$\begin{cases} y_1 &= Z\Pi_2\beta + v_1 \\ Y_2 &= Z\Pi_2 + V_2 \end{cases} \tag{4}$$

which leads to the partially restricted reduced-form (PRRF) estimator $\tilde{\Pi}$ of Π (Knight, 1977):

$$\tilde{\Pi} := \left(\tilde{\pi}_1, \hat{\Pi}_2\right), \quad \tilde{\pi}_1 := \hat{\Pi}_2\beta_{IV},$$

where $\beta_{IV} = \left(Y_2' P_Z Y_2\right)^{-1} Y_2' P_Z y_1$ is the IV estimator of β in (1) and $P_Z := Z\left(Z'Z\right)^{+} Z'$. Since $\beta_{IV} = \left(\hat{\Pi}_2'\hat{\Pi}_2\right)' \hat{\Pi}_2'\hat{\pi}_1$, we have $\tilde{\pi}_1 = P_{\hat{\Pi}_2}\hat{\pi}_1$, where $P_{\hat{\Pi}_2} := \hat{\Pi}_2\left(\hat{\Pi}_2'\hat{\Pi}_2\right)^{+}\hat{\Pi}_2'$ is the projection matrix to the range of $\hat{\Pi}_2$.

Interest focuses on the partially restricted reduced-form estimator $\tilde{\pi}_1$, which carries information about the structural equation through the estimate β_{IV}, and the effect of this information on inference. We investigate both exact finite sample and asymptotic distributions. It is already known from Knight (1977) that the partially restricted reduced form estimator has finite moments of all orders under a Gaussian error matrix V. Finite sample density results for the PRRF estimator were first obtained in Phillips (Forthcoming), again for Gaussian errors.

3 Exact and Asymptotic Distributions of the PRRF Estimator

We derive the exact and the asymptotic distributions of the PRRF estimator under three different assumptions concerning instrument strength. For simplicity and without loss of generality (see Phillips, 1983), we employ standardizing transformations so that $Z'Z = nI_K$. For the exact finite sample theory we employ Gaussian error assumptions with $V \overset{d}{=} \mathcal{N}_{n,m+1}\left(\mathbf{0}, I_{n(m+1)}\right)$ where $\mathcal{N}_{n,m+1}\left(A,\Sigma\right)$ signifies an $n\times(m+1)$ matrix normal distribution with mean matrix A and covariance matrix Σ, or in vectorized form $vec\left(V\right) \overset{d}{=} \mathcal{N}\left(vec\left(A\right),\Sigma\right)$. Then

$$\frac{1}{\sqrt{n}}Z'V \overset{d}{=} \mathcal{N}_{K,m}\left(\mathbf{0}, I_{K(m+1)}\right). \tag{A1}$$

To develop asymptotics without Gaussianity, we assume that the rows $\left\{V_{(i)}\right\}_{i=1}^{n}$ of V, coupled with the natural filtration, form an $\mathbb{R}^{m+1}$-valued martingale difference sequence with

$$\mathbb{E}\left[V\right] = \mathbf{0} \text{ and } \mathbb{V}\mathrm{ar}\left[V\right] := \mathbb{V}\mathrm{ar}\left[vec\left(V\right)\right] = I_{n(m+1)}. \tag{A2}$$

Using the martingale central limit theory in Phillips (1989) we have the weak convergence as $n\to\infty$

$$\frac{1}{\sqrt{n}}Z'V = \frac{1}{\sqrt{n}}Z'\left[v_1, V_2\right] \overset{d}{\longrightarrow} \left(\xi,\Xi\right) \overset{d}{=} \mathcal{N}_{K,m+1}\left(\mathbf{0}, I_{K(m+1)}\right). \tag{A3}$$

It is convenient in what follows to expand the probability space as needed so that, by Skorokhod's representation theorem, (A3) may be replaced by strong convergence in that space, giving

$$\frac{1}{\sqrt{n}}Z'V \overset{a.s.}{\longrightarrow} \left(\xi,\Xi\right) \overset{d}{=} \mathcal{N}_{K,m}\left(\mathbf{0}, I_{Km}\right), \text{ as } n\to\infty. \tag{A3$'$}$$

We proceed to characterize the exact finite sample and asymptotic distributions of the PRRF estimator under three different assumptions concerning instrument strengths.

3.1 Strong Instruments

We first consider the case of strong instruments under the following standard condition.

Assumption (S-IV). *(Strong instruments) Π_2 is fixed with respect to n and has full column rank m.*

Since $\Pi = \mathbb{E}[Y|\,Z]$ and $\Pi_2'\Pi_2$ is invertible under strong instruments, the structural parameter β is identified from the reduced-form and satisfies $\beta = \left(\Pi_2'\Pi_2\right)^{-1}\Pi_2'\pi_1$.

Lemma 1. *Under (S-IV) and (A3'), the asymptotic distributions of $\hat{\Pi}$, β_{IV}, and $\tilde{\pi}_1$ are given by*

$$\sqrt{n}\left(\hat{\Pi} - \Pi\right) \xrightarrow{a.s.} (\xi, \Xi) \overset{d}{=} \mathcal{N}_{K,m+1}\left(\mathbf{0}, I_{K(m+1)}\right) \tag{5}$$

$$\sqrt{n}\left(\beta_{IV} - \beta\right) \xrightarrow{a.s.} \left(\Pi_2'\Pi_2\right)^{-1}\Pi_2'\left(\xi - \Xi\beta\right) \overset{d}{=} \mathcal{N}\left(\mathbf{0}, (1+\beta'\beta)\left(\Pi_2'\Pi_2\right)^{-1}\right) \tag{6}$$

$$\sqrt{n}\left(\tilde{\pi}_1 - \pi_1\right) \xrightarrow{a.s.} P_{\Pi_2}\xi + M_{\Pi_2}\Xi\beta \overset{d}{=} \mathcal{N}\left(\mathbf{0}, \beta'\beta \cdot I_K + (1-\beta'\beta)\,P_{\Pi_2}\right), \tag{7}$$

where $M_{\Pi_2} := I_K - P_{\Pi_2}$.

Under (S-IV) and (A1), the finite-sample distributions of $\hat{\Pi}$, β_{IV}, and $\tilde{\pi}_1$ are given by

$$\begin{aligned} \hat{\Pi} &\overset{d}{=} \mathcal{N}_{K,m+1}\left(\Pi, \frac{1}{n}I_{K(m+1)}\right) \\ \beta_{IV} = W_{22}^{-1}w_{21} &\overset{d}{=} \mathcal{MN}\left(\sqrt{n}\left(A_\Xi'A_\Xi\right)^{-1}A_\Xi\Pi_2\beta, \left(A_\Xi'A_\Xi\right)^{-1}\right) \\ \tilde{\pi}_1 &\overset{d}{=} \mathcal{MN}\left(P_{A_\Xi}\Pi_2\beta, \frac{1}{n}P_{A_\Xi}\right) \end{aligned} \tag{8}$$

where $A_\Xi := \sqrt{n}\Pi_2 + \Xi$, and (W_{22}, w_{21}) form blocks of the non-central Wishart matrix

$$W \equiv \begin{array}{c} \begin{array}{cc} \scriptstyle 1 & \scriptstyle m \end{array} \\ \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & W_{22} \end{bmatrix} \end{array} \overset{d}{=} W_{m+1}\left(K, I_{m+1}, MM'\right),$$

with K degrees of freedom, covariance matrix I_{m+1}, and noncentrality matrix

$$MM' = n\begin{bmatrix} \beta' \\ I_m \end{bmatrix}\Pi_2'\Pi_2\left[\beta, I_m\right],$$

using $\mathcal{MN}\left(h(\Xi), H(\Xi)\right)$ to denote the mixed normal distribution with random mean vector $h(\Xi)$ and random covariance matrix $H(\Xi)$, where $\Xi \overset{d}{=} \mathcal{N}_{K,m}\left(\mathbf{0}, I_{Km}\right)$.

Let $\mathbb{AV}\text{ar}\left(\hat{\pi}_1\right)$ and $\mathbb{AV}\text{ar}\left(\tilde{\pi}_1\right)$ denote the asymptotic variances of $\hat{\pi}_1, \tilde{\pi}_1$. Then

$$\mathbb{AV}\text{ar}\left(\hat{\pi}_1\right) - \mathbb{AV}\text{ar}\left(\tilde{\pi}_1\right) = (1-\beta'\beta)\cdot M_{\Pi_2} \geq 0$$

Since M_{Π_2} is positive semidefinite, it follows that $\tilde{\pi}_1$ is asymptotically weakly more efficient than $\hat{\pi}_1$ when $\|\beta\| < 1$. That is, when $\|\beta\| < 1$, $\gamma'\tilde{\pi}_1$ is at least as efficient as $\gamma'\hat{\pi}_1$ in all directions and more efficient in directions $\gamma \notin \mathcal{R}(\Pi_2)$.

The relative efficiency of the PRRF and the URRF estimators depends on the magnitude of two opposite effects. On one hand, the partially restricted estimator $\tilde{\pi}_1 = \hat{\Pi}_2\beta_{IV}$ brings extra information from the structural equation into the estimation of π_1, potentially contributing to more efficient estimation. This is captured mathematically by comparing the two terms $P_{\Pi_2}\xi$ and $\xi = P_{\Pi_2}\xi + (I - P_{\Pi_2})\xi$ in the asymptotic distributions of the PRRF and the URRF estimators, viz., (7) for PRRF, and (5) for URRF: the former achieves a reduction of $M_{\Pi_2} = I - P_{\Pi_2}$ in asymptotic variation relative to the latter. On the other hand, the PRRF estimator introduces estimation error in β_{IV} multiplicatively with $\hat{\Pi}_2$, which can amplify asymptotic variance when that factor is large. This is captured through the addition of the second term $M_{\Pi_2}\Xi\beta$ in the asymptotic distribution of $\tilde{\pi}_1$, which is not present in that of $\hat{\pi}_1$. This extra term contributes to the additional variance component $\beta'\beta \cdot M_{\Pi_2}$ that figures in the asymptotic variance of the PRRF estimator. When the length of β is less than 1, the improvement by incorporating the information from the structural equation dominates the extra variance it brings in, leading to smaller asymptotic variance for the PRRF estimator $\tilde{\pi}_1$.

Since $\mathcal{N}(\Pi_2, n^{-1}I_{Km}) = \Pi_2 + \mathcal{N}(0, n^{-1}I_{Km}) = \Pi_2 + O_p\left(n^{-1/2}\right)$, the representation (8) reveals that the distribution of $\tilde{\pi}_1$ is determined by $\left\{\Pi_2 + n^{-1/2}\mathcal{N}(0, I_{Km})\right\} \times W_{22}^{-1}w_{21}$, showing that a primary contribution to the finite sample distribution is given by $\Pi_2 \times W_{22}^{-1}w_{21}$. The analytic form of this density was obtained recently in Phillips (Forthcoming) and can be derived using the exact density of the matrix quotient $W_{22}^{-1}w_{21}$, which was given in Phillips (1980).

3.2 Totally Irrelevant Instruments

Next consider the case where the instruments are all totally irrelevant for the structural parameter β. This case represents the polar extreme of the strong instrument case. There is no information in the reduced form about the structural coefficients and so β is totally unidentified.

Assumption (I-IV). *(Irrelevant instruments)* $\Pi_2 = \mathbf{0}$.

Lemma 2. *Under (I-IV) and (A3'), the asymptotic distributions of* $\hat{\Pi}$, β_{IV} *and* $\tilde{\pi}_1$ *are given by*

$$\sqrt{n}\hat{\Pi} \xrightarrow{a.s.} (\xi, \Xi) \stackrel{d}{=} \mathcal{N}_{K,m+1}\left(\mathbf{0}, I_{K(m+1)}\right), \tag{9}$$

$$\beta_{IV} \xrightarrow{a.s.} \left(\Xi'\Xi\right)^{-1}\Xi'\xi \stackrel{d}{=} \mathcal{MN}\left(\mathbf{0}, \left(\Xi'\Xi\right)^{-1}\right) \tag{10}$$

$$\sqrt{n}\tilde{\pi}_1 \xrightarrow{a.s.} P_{\Xi}\xi \stackrel{d}{=} \mathcal{MN}(\mathbf{0}, P_{\Xi}) \ \textit{with}\ \Xi \stackrel{d}{=} \mathcal{N}_{K,m}(\mathbf{0}, I_{Km}) \tag{11}$$

Under (I-IV) and (A1), the finite-sample distributions are obtained from the above simply by replacing "$\xrightarrow{a.s.}$" with "$\stackrel{d}{=}$".

Define $\Upsilon := \Xi \left(\Xi'\Xi\right)^{-\frac{1}{2}}$. Since Υ is uniformly distributed on the Stiefel manifold[2] $V_{K,m}$, we may alternatively take the mixture in $\mathcal{MN}\left(\mathbf{0}, P_{\Xi} \equiv \Upsilon\Upsilon'\right)$ with respect to Υ.

Notice that $\mathbb{AV}\mathrm{ar}(\hat{\pi}_1) - \mathbb{AV}\mathrm{ar}(\tilde{\pi}_1) = I_K - \mathbb{E}\left[P_{\Xi}\right] = \mathbb{E}\left[M_{\Xi}\right]$. Since M_{Ξ} is positive semi-definite for each realization of Ξ and M_{Ξ} is well defined almost surely, $\mathbb{E}\left[M_{\Xi}\right]$ is also positive semi-definite. Hence, the PRRF estimator $\tilde{\pi}_1$ is asymptotically more efficient for $\pi_1 = 0$ than the URRF estimator $\hat{\pi}$. Importantly, as the model is totally unidentified, the relative efficiency of $\tilde{\pi}_1$ does not translate in this case into improved inference on the structural parameter β, which will be made clear in Section 4.

3.3 Weak Instruments

Finally we consider the case where the instruments are weak in the conventional sense that the corresponding reduced form coefficients are local to zero.

Assumption (W-IV). *(Weak instruments)* $\Pi_2 = \frac{1}{\sqrt{n}}\Pi_2^*$, *where* Π_2^* *is of full rank* m.

In this case, we say that β is weakly identified. Write $\Pi^* \equiv (\pi_1^*, \Pi_2^*)$ with $\pi_1^* := \Pi_2^*\beta$.

Lemma 3. *Under (W-IV) and (A3'), the asymptotic distributions of* $\hat{\Pi}$, β_{IV} *and* $\tilde{\pi}_1$ *are given by*

$$\begin{aligned}
\sqrt{n}\hat{\Pi} &\xrightarrow{a.s.} \Pi^* + (\xi, \Xi) \overset{d}{=} \mathcal{N}_{K,m+1}\left(\Pi^*, I_{K(m+1)}\right), \\
\beta_{IV} &\xrightarrow{a.s.} \left[\left(\Pi_2^* + \Xi\right)'\left(\Pi_2^* + \Xi\right)\right]^{-1}\left(\Pi_2^* + \Xi\right)'\left(\pi_1^* + \xi\right) \\
\sqrt{n}\tilde{\pi}_1 &\xrightarrow{a.s.} P_{\left(\Pi_2^*+\Xi\right)}\left(\pi_1^* + \xi\right) \overset{d}{=} \mathcal{MN}\left(P_{\left(\Pi_2^*+\Xi\right)}\Pi_2^*\beta, P_{\left(\Pi_2^*+\Xi\right)}\right) \text{ with } \Xi \overset{d}{=} \mathcal{N}_{K,m}\left(\mathbf{0}, I_{Km}\right).
\end{aligned}$$

Under (W-IV) and (A1), the finite-sample distributions can be obtained from above by replacing "$\xrightarrow{a.s.}$*" with "*$\overset{d}{=}$*".*

Again, $\mathbb{AV}\mathrm{ar}\left(\hat{\pi}_1\right) - \mathbb{AV}\mathrm{ar}\left(\tilde{\pi}_1\right) = \mathbb{E}\left[M_{\left(\Pi_2^*+\Xi\right)}\right]$ is positive semi-definite, but notice that the asymptotic distribution of $\tilde{\pi}_1$ is now no longer centered at π_1^* because

$$\mathbb{E}\left[P_{\left(\Pi_2^*+\Xi\right)}\left(\pi_1^* + \xi\right)\right] = \mathbb{E}\left[P_{\left(\Pi_2^*+\Xi\right)}\right]\Pi_2^*\beta \neq \Pi_2^*\beta.$$

Therefore, when instruments are weak, importing structural information into reduced-form estimation introduces bias, as compared with unrestricted reduced-form estimation. This bias is to be expected, because the weak instrument asymptotic theory corresponds to the Gaussian exact distribution and therefore carries all the finite sample parameter dependencies that arise in finite sample theory, including the finite sample bias of the instrumental variable estimator. This heuristic reasoning indicates that there may be some advantage in the use of partially restricted reduced form estimation using the LIML estimator of β because the LIML estimator, while having no finite sample integer moments (e.g., Phillips, 1984), is known to be better centered about β than β_{IV}.

[2] $V_{K,m}$ is the manifold formed by K frames of m dimensional orthonormal vectors. See, for example, Muirhead (2005).

4 Hypothesis Testing on β

The structural parameter β is usually the parameter of interest and we can use the reduced-form estimates $\left(\hat{\Pi}, \tilde{\Pi}\right)$ to construct tests concerning β. Specifically, to test the hypothesis $H_0 : \beta = \beta_0$, we may test the implied relationship between reduced-form parameters $H_0 :\ \pi_1 = \Pi_2\beta_0$. Using the estimates $\left(\hat{\Pi}, \tilde{\Pi}\right)$ respectively, define

$$W_{UR}^{\beta_0} := \frac{n\left(\hat{\pi}_1 - \hat{\Pi}_2\beta_0\right)'\left(\hat{\pi}_1 - \hat{\Pi}_2\beta_0\right)}{1+\beta_0'\beta_0},$$
$$W_{PR}^{\beta_0} := \frac{n\left(\tilde{\pi}_1 - \hat{\Pi}_2\beta_0\right)'\left(\tilde{\pi}_1 - \hat{\Pi}_2\beta_0\right)}{1+\beta_0'\beta_0}.$$

We call $W_{UR}^{\beta_0}$ the unrestricted reduced-form (URRF) test statistic, and $W_{PR}^{\beta_0}$ the partially restricted reduced-form (PRRF) test statistic. The URRF statistic was proposed in Chernozhukov and Hansen (2008), who showed that the asymptotic distribution of the URRF statistic $W_{UR}^{\beta_0}$ is robust to assumptions concerning instrument strength.

Theorem 1. *(Chernozhukov and Hansen, 2008) Under (A3) and the null hypothesis H_0 : $pi_1 = \Pi_2\beta_0$,*

$$W_{UR}^{\beta_0} \xrightarrow{d} \chi_K^2,$$

irrespective of the strength of instruments.

The next result characterizes the asymptotic distribution of the PRRF test statistic $W_{PR}^{\beta_0}$.

Theorem 2. *Under (A3) and the null hypothesis $H_0 :\ \pi_1 = \Pi_2\beta_0$,*

$$W_{PR}^{\beta_0} \xrightarrow{d} \frac{1}{1+\beta_0'\beta_0}\mathcal{M}\chi^2_{m,\lambda_A(\Xi,\beta_0)}, \tag{12}$$

where $\mathcal{M}\chi^2_{m,\lambda_A(\Xi,\beta_0)}$ denotes a mixed noncentral chi-squared distribution with m degrees of freedom and random noncentrality parameter $\lambda_A(\Xi,\beta_0)$ with $\Xi \stackrel{d}{=} \mathcal{N}_{K,m}(\mathbf{0}, I_{Km})$, and the index $A \in \{S, W, I\}$ signifies strong, weak and irrelevant instruments:

(i) $A = S$: under (S-IV) and (A3'),

$$\lambda_S(\Xi,\beta_0) = \beta_0'\Xi' P_{\Pi_2}\Xi\beta_0 \stackrel{d}{=} \beta_0'\beta_0 \cdot \chi_m^2. \tag{13}$$

In this case, equivalently we have

$$W_{PR}^{\beta_0} \xrightarrow{d} \chi_m^2.$$

Under (S-IV) and (A1), the finite-sample distribution is given by

$$W_{PR}^{\beta_0} \stackrel{d}{=} \frac{1}{1+\beta_0'\beta_0}\mathcal{M}\chi^2_{m,\lambda_{S,n}(\Xi,\beta_0)}$$

with

$$\lambda_{S,n}(\Xi,\beta_0) = \beta_0'\Xi' P_{A_\Xi}\Xi\beta_0.$$

(ii) $A = I$*: under (I-IV) and (A3'),*

$$\lambda_I(\Xi,\beta_0) = \beta_0'\Xi'\Xi\beta_0 \stackrel{d}{=} \beta_0'\beta_0 \cdot \chi^2_K. \tag{14}$$

Under (I-IV) and (A1), the finite-sample distribution coincides with the asymptotic distribution.

(iii) $A = W$*: under (W-IV) and (A3'),*

$$\lambda_W(\Xi,\beta_0) = \beta_0'\Xi' P_{\left(\Pi_2^*+\Xi\right)}\Xi\beta_0. \tag{15}$$

Under (W-IV) and (A1), the finite-sample distribution coincides with the asymptotic distribution.

Theorem 2 shows that the differences in the asymptotic distributions under different instrument strengths are embodied in the corresponding noncentrality parameters $\lambda_A(\Xi,\beta_0)$. A smaller noncentrality parameter corresponds to a more concentrated null distribution, and thus a tighter (smaller) critical value in hypothesis testing.

As the family of noncentral chi-squared distributions with the same degree of freedom, say m, are ordered in the sense of first-order stochastic dominance according to their noncentrality parameters, we may compare distributions by comparing the noncentrality parameters for the three instrument strengths. For any β_0 and realization Ξ,

$$\begin{aligned}\lambda_S(\Xi,\beta_0) &\leq \lambda_I(\Xi,\beta_0),\\ \lambda_W(\Xi,\beta_0) &\leq \lambda_I(\Xi,\beta_0),\end{aligned}$$

with the inequality being strict almost surely. So, in the case of irrelevant instruments, the asymptotic distribution of $W_{PR}^{\beta_0}$ first-order stochastically dominates those with strong and weak instruments. This is natural because with both strong and weak instruments the reduced-form estimates contain information about the structural parameter β, while under irrelevant instruments these estimates carry no such information. The comparison between $\lambda_S(\Xi,\beta_0)$ and $\lambda_W(\Xi,\beta_0)$, however, is not immediately clear from (13) and (15).

We may also compare the asymptotic distributions of $W_{PR}^{\beta_0}$ with that of $W_{UR}^{\beta_0}$. Noticing that, regardless of instrument strengths, under (A3')

$$\begin{aligned}W_{UR}^{\beta_0} = \frac{n\left(\hat{\pi}_1 - \hat{\Pi}_2\beta_0\right)'\left(\hat{\pi}_1 - \hat{\Pi}_2\beta_0\right)}{1+\beta_0'\beta_0} &\xrightarrow{a.s.} \frac{(\xi-\Xi\beta_0)'(\xi-\Xi\beta_0)}{1+\beta_0'\beta_0}\\ \geq \frac{(\xi-\Xi\beta_0)' P_\Xi(\xi-\Xi\beta_0)}{1+\beta_0'\beta_0} &\stackrel{d}{=} \frac{1}{1+\beta_0'\beta_0}\mathcal{M}\chi^2_{m,\lambda_I(\Xi,\beta_0)}.\end{aligned}$$

Thus the asymptotic distribution of $W_{UR}^{\beta_0}$ first-order stochastically dominates that of $W_{PR}^{\beta_0}$ under all three instrument strengths.

Moreover, under the particular null hypothesis where $\beta_0 = \mathbf{0}$, the asymptotic distribution of W_{PR}^{0} also becomes invariant to the strength of instruments, as $\lambda_A(\Xi, \mathbf{0}) = 0$ for any Ξ and $A \in \{S, W, I\}$. In this case, we have $W_{PR}^{\mathbf{0}} \xrightarrow{d} \chi_m^2$ in all cases. Importantly, the distribution of the PRRF test statistic has degrees of freedom m, corresponding to the dimension of the vector β being tested, unlike the URRF statistic whose distribution has degrees of freedom K, corresponding to the number of instruments available from the reduced form.

The analysis above focuses on the asymptotic distribution of W_{PR}^{0} under the null hypothesis $H_0 : \beta = \beta_0$. To obtain a complete comparison of hypothesis tests based on the URRF and the PRRF test statistics, we need to compute the power functions.

Consider tests of size α for the null hypothesis $H_0 : \beta = \beta_0$ based on the URRF and the PRRF test statistics. Let $\mathcal{K}_{UR}(\beta_1; \beta_0)$, $\mathcal{K}_{PR}(\beta_1; \beta_0)$ denote the asymptotic power functions of the test of the null hypothesis $H_0 : \beta = \beta_0$ against an alternative involving the localizing parameter β_1 (specified precisely later), and let $\mathcal{K}_{UR}^{(n)}(\beta_1; \beta_0)$, $\mathcal{K}_{PR}^{(n)}(\beta_1; \beta_0)$ denote the corresponding finite-sample power functions. Let $q_k^{1-\alpha}$ denote the $(1-\alpha)$-th quantile of χ_k^2 and $\Psi(x; k, \lambda)$ denote the upper tail (survivor function) probability of $\chi_{k,\lambda}^2$. We state the following lemma before giving the results on power.

Lemma 4.

(i) $\forall x > 0,\ \forall \lambda \geq 0,\ \forall k \in \mathbb{N}_+,\ \Psi(x; k+1, \lambda) > \Psi(x; k, \lambda)$.

(ii) $\forall k \in \mathbb{N}_+,\ \forall x > 0,\ \Psi(x; k, \lambda)$ *is increasing in* λ.

(iii) For any even $k \in \mathbb{N}$ *and for small enough* $\alpha \in (0,1)$, *there exists some* $\overline{\lambda} > 0$ *such that* $\forall \lambda \in \left(0, \overline{\lambda}\right),\ \Psi\left(q_{k+2}^{1-\alpha}; k+2, \lambda\right) < \Psi\left(q_k^{1-\alpha}; k, \lambda\right)$.

The inequality in Lemma 4(iii) appears to be the first result of this kind for noncentral chi-squared distributions. The result gives an inequality for the tail probabilities of noncentral chi-squared variates evaluated at different quantiles and with different degrees of freedom. The result is relevant to power comparisons in many circumstances in which alternative test statistics have finite sample or asymptotic ch-squared distributions with differing degrees of freedom.

A stronger version of Lemma 4(iii) would state: for small enough $\alpha \in (0,1)$, $\forall k \in \mathbb{N}$, $\Psi\left(q_{k+1}^{1-\alpha}; k+1, \lambda\right) < \Psi\left(q_k^{1-\alpha}; k, \lambda\right)$. This inequality seems expected on the following heuristic grounds: a chi-squared distribution with a higher degree of freedom is more dispersed, so a shift in the noncentrality parameter of a given size should have a smaller impact on the tail probability of the resultant chi-squared distribution with a higher degree of freedom when evaluated at corresponding quantiles under the null. We have numerically verified that the inequality holds uniformly for $\alpha \in \{0.01, 0.05, 0.1\}$, $k \leq 200$ and $\lambda \in \{0.01, 0.02, ..., 0.99, 1, 2, ..., 50\}$.[3]

[3]The calculations were performed in MatLab with a machine epsilon of approximately 2×10^{-16}, and

A proof of this more general version of the inequality appears difficult due to the complicated nonlinear dependence of the quantiles $q_k^{1-\alpha}$ on both k and α, as well as the analytic complexity of the chi-squared CDF, which involves an incomplete gamma function. A proof of the power comparison inequality would probably require relatively tight upper and lower bounds on the quantiles of the associated chi-squared distributions, which play a significant role in the comparison. Further analysis of such comparisons is therefore left for future work.

Theorem 3. *(Hypothesis testing against the null $H_0 : \beta = \beta_0$)*

(i) Under (S-IV) and (A3'), and testing H_0 against $H_1 : \beta = \tilde{\beta}_n := \beta_0 + \frac{1}{\sqrt{n}}(\beta_1 - \beta_0)$ with size α, we have

$$\mathcal{K}_{UR}(\beta_1; \beta_0) = \Psi\left(q_K^{1-\alpha};\ K,\ \frac{(\beta_1 - \beta_0)' \Pi_2' \Pi_2 (\beta_1 - \beta_0)}{1 + \beta_0' \beta_0}\right),$$
$$\mathcal{K}_{PR}(\beta_1; \beta_0) = \Psi\left(q_m^{1-\alpha};\ m,\ \frac{(\beta_1 - \beta_0)' \Pi_2' \Pi_2 (\beta_1 - \beta_0)}{1 + \beta_0' \beta_0}\right).$$

Under (S-IV) and (A1),

$$\mathcal{K}_{UR}^{(n)}(\beta_1; \beta_0) = \Psi\left(q_K^{1-\alpha};\ K,\ \frac{(\beta_1 - \beta_0)' \Pi_2' \Pi_2 (\beta_1 - \beta_0)}{1 + \beta_0' \beta_0}\right).,$$
$$\mathcal{K}_{PR}^{(n)}(\beta_1; \beta_0) = \mathbb{E}\left[\Psi\left(\left(1 + \beta_0' \beta_0\right) q_m^{1-\alpha};\ m,\ \kappa_n(\Xi)\right)\right]$$

with

$$\kappa_n(\Xi) = (\Xi\beta_0 - \Pi_2(\beta_1 - \beta_0))' P_{A_\Xi} (\Xi\beta_0 - \Pi_2(\beta_1 - \beta_0)).$$

(ii) Under (I-IV) and (A3'), and testing H_0 against $H_1 : \beta = \beta_1$ with size α, we have

$$\mathcal{K}_{UR}(\beta_1; \beta_0) = \mathcal{K}_{PR}(\beta_1; \beta_0) = \alpha.$$

Under (I-IV) and (A1), the finite-sample power functions coincide with the asymptotic power functions.

(iii) Under (W-IV) and (A3'), and testing H_0 against $H_1 : \beta = \beta_1$ with size α, we have

$$\mathcal{K}_{UR}(\beta_1; \beta_0) = \Psi\left(q_K^{1-\alpha};\ K,\ \frac{(\beta_1 - \beta_0)' \Pi_2^{*\prime} \Pi_2^* (\beta_1 - \beta_0)}{1 + \beta_0' \beta_0}\right)$$
$$\mathcal{K}_{PR}(\beta_1; \beta_0) = \mathbb{E}\left[\Psi\left(\left(1 + \beta_0' \beta_0\right) c_{\beta_0}^{1-\alpha};\ m, (\Xi\beta_0 - \Pi_2^*(\beta_1 - \beta_0))' P_{(\Pi_2^* + \Xi)} (\Xi\beta_0 - \Pi_2^*(\beta_1 - \beta_0))\right)\right]$$

we found that: the maximum difference $\Psi\left(q_{k+1}^{1-\alpha}; k+1, \lambda\right) - \Psi\left(q_k^{1-\alpha}; k, \lambda\right)$ is negative and has a magnitude larger than 10^{-8}, and the maximum log ratio $\log\left(\Psi\left(q_{k+1}^{1-\alpha}; k+1, \lambda\right) / \Psi\left(q_k^{1-\alpha}; k, \lambda\right)\right)$ is also negative and has a magnitude larger than 10^{-5}.

where $\Xi \overset{d}{=} \mathcal{N}_{Km}(\mathbf{0}, I_{Km})$ *and* $c_{\beta_0}^{1-\alpha}$ *is such that*

$$\mathbb{P}\left\{\mathcal{M}\chi^2_{m,\beta_0'\Xi' P_{\left(\Pi_2^*+\Xi\right)}\Xi\beta_0} \geq \left(1+\beta_0'\beta_0\right) c_{\beta_0}^{1-\alpha}\right\} = \alpha.$$

If $\beta_0 = \mathbf{0}$*, then* $c_{\beta_0}^{1-\alpha} = q_m^{1-\alpha}$ *and*

$$\mathcal{K}_{PR}(\beta_1; \mathbf{0}) = \mathbb{E}\left[\Psi\left(q_m^{1-\alpha};\ m,\right)\right]$$

Under (W-IV) and (A1), the finite-sample power functions coincide with the asymptotic power functions.

These results facilitate several power comparisons between the URRF and the PRRF tests. First, with strong instruments, the PRRF test is typically more powerful than the URRF test for any null β_0 and (local) alternative β_1 under the conditions of Lemma 4(iii) or, more extensively, as supported by the stated numerical computations; and we conjecture that the power domination comparison holds more generally. Second, with irrelevant instruments, power equals size and neither the PRRF nor the URRF test is informative about β. Third, with weak instruments, the power comparison is indeterminate. Take for example the case of $\beta_0 = \mathbf{0}$. Recall that $\Psi\left(q_k^{1-\alpha};\ k,\ \lambda\right)$ is increasing in λ and decreasing in k, so $\beta_1'\Pi_2^{*\prime} P_{\left(\Pi_2^*+\Xi\right)}\Pi_2^*\beta_1 < \beta_1'\Pi_2^{*\prime}\Pi_2^*\beta_1$ almost surely. But in overidentified models $m < K$ and so $\Psi\left(q_m^{1-\alpha};\ m,\right)$ may be larger or smaller than $\Psi\left(q_K^{1-\alpha};\ K,\right)$, depending on the values of m, K, Π_2^* and the realization of Ξ. With Ξ integrated out, the power comparison remains dependent on Π_2^* and β_1.

5 Asymptotic Power with $K \to \infty$

To construct a framework that allows for an increasing number of instruments, we assume that

$$\Pi_2 = \left[\Pi_{2(1)}', ..., \Pi_{2(K)}'\right]', \quad \text{with } \Pi_{2(k)} \sim_{iid} \left(\mathbf{0}_{1\times m}, \Omega_{\Pi_2}\right),\ \Omega_{\Pi_2} > 0. \tag{16}$$

Then, as $K \to \infty$, we have $K^{-1}\mathbf{1}_K'\Pi_2 \xrightarrow{a.s.} 0$ and $K^{-1}\Pi_2'\Pi_2 \xrightarrow{a.s.} \Omega_{\Pi_2} > 0$. Next let $K = K(n) \to \infty$ slowly relative to n so that $\frac{K}{n} \to 0$. With this framework, we can derive asymptotic power functions of the URRF test and the PRRF test allowing for different strengths in the increasing number of instruments. First note that under the null $H_0 : \beta = \beta_0$, Assumption A1 or A3, and suitable centering and standardization, the URRF statistic $W_{UR}^{\beta_0}$ is asymptotically normal, viz.,

$$\tilde{W}_{UR}^{\beta_0} = \sqrt{K}\left(\frac{1}{K}W_{UR}^{\beta_0} - 1\right) = \sqrt{K}\left[\frac{\frac{1}{K}n\left(\hat{\pi}_1 - \hat{\Pi}_2\beta_0\right)'\left(\hat{\pi}_1 - \hat{\Pi}_2\beta_0\right)}{1+\beta_0'\beta_0} - 1\right]$$

$$\xrightarrow{d} \lim_{K\to\infty}\sqrt{K}\left[\frac{1}{K}\frac{\left(\xi - \Xi\beta_0\right)'\left(\xi - \Xi\beta_0\right)}{1+\beta_0'\beta_0} - 1\right] \overset{d}{=} \mathcal{N}(0, 2).$$

This limit theory is used to obtain critical values of the URRF statistic when $K \to \infty$.

Theorem 4. *(Hypothesis testing of $H_0 : \beta = \beta_0$ with $K \to \infty$)*

(i) (Strong instruments, $\frac{1}{\sqrt{nK}}$-local alternatives) Suppose that $\Pi_{2(k)} \sim_{iid} (\mathbf{0}_{1\times m}, \Omega_{\Pi_2})$ and consider a size-α test of H_0 against

$$H_1 : \beta = \tilde{\beta}_n := \beta_0 + \frac{1}{\sqrt{nK}} (\beta_1 - \beta_0) \tag{17}$$

based on the (asymptotic distributions of) the test statistics $\tilde{W}_{UR}^{\beta_0} = \sqrt{K}\left(\frac{1}{K} W_U^{\beta_0} - 1\right)$ and $W_{PR}^{\beta_0}$ as $(K, n) \to \infty$ with $\frac{K}{n} \to 0$. Under (S-IV) and (A3'), the PRRF test has nontrivial asymptotic power while the URRF test is asymptotically blind to $O\left(1/\sqrt{nK}\right)$ local alternatives. In particular

$$\mathcal{K}_{PR}(\beta_1; \beta_0) = \Psi\left(q_m^{1-\alpha}; m, \frac{(\beta_1 - \beta_0)' \Omega_{\Pi_2} (\beta_1 - \beta_0)}{1 + \beta_0' \beta_0}\right) > \alpha = \mathcal{K}_{UR}(\beta_1; \beta_0).$$

(ii) (Strong instruments, $\frac{1}{\sqrt{n}K^{1/4}}$-local alternatives) Suppose that $\Pi_{2(k)} \sim_{iid} (\mathbf{0}_{1\times m}, \Omega_{\Pi_2})$ and a size-α test of H_0 is performed against the local alternative

$$H_1 : \beta = \tilde{\beta}_n := \beta_0 + \frac{1}{\sqrt{n}K^{1/4}} (\beta_1 - \beta_0) \tag{18}$$

using $\tilde{W}_{UR}^{\beta_0}$ and $W_{PR}^{\beta_0}$. Then, under (S-IV) and (A3'), the URRF test has nontrivial asymptotic power while the PRRF test has unit power asymptotically:

$$\mathcal{K}_{PR}(\beta_1; \beta_0) = 1 \geq \mathcal{K}_{UR}(\beta_1; \beta_0) = \Phi\left(\Phi^{-1}(\alpha) - \frac{1}{\sqrt{2}} \frac{(\beta_1 - \beta_0)' \Omega_{\Pi_2} (\beta_1 - \beta_0)}{1 + \beta_0' \beta_0}\right).$$

(iii) (Weak instruments, $\frac{1}{\sqrt{nK}}$-local alternatives) Let $\Pi_{2(k)} = \frac{1}{\sqrt{n}} \Pi_{2(k)}^*$ with $\Pi_{2(k)}^* \sim_{iid} \left(\mathbf{0}_{1\times m}, \Omega_{\Pi_2^*}\right)$ and $\Omega_{\Pi_2^*} > 0$. Size-α tests of H_0 are performed against the local alternative

$$H_1 : \beta = \tilde{\beta}_n := \beta_0 + \frac{1}{\sqrt{K}} (\beta_1 - \beta_0)$$

using $\tilde{W}_{UR}^{\beta_0}$ and $\tilde{W}_{PR}^{\beta_0}$, with

$$\tilde{W}_{PR}^{\beta_0} := \begin{cases} W_{PR}^{\mathbf{0}}, & \beta_0 = \mathbf{0}, \\ \frac{1}{\sqrt{K}}\left(W_{PR}^{\beta_0} - \frac{m + K\beta_0' \left(\Omega_{\Pi_2^*} + I_m\right)^{-1} \beta_0}{1 + \beta_0' \beta_0}\right), & \beta_0 \neq \mathbf{0}. \end{cases}$$

Then, under (W-IV) and (A3'),

$$\mathcal{K}_{UR}(\beta_1;\beta_0)=\alpha.$$

$$\mathcal{K}_{PR}(\beta_1;\beta_0)=\begin{cases}\Psi\left(q_m^{1-\alpha};m,\beta_1^{'}\Omega_{\Pi_2^*}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\Omega_{\Pi_2^*}\beta_1\right)\geq\alpha, & \beta_0=\mathbf{0};\\ \Phi\left(\Phi^{-1}(\alpha)+\frac{2\beta_0^{'}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\Omega_{\Pi_2^*}(\beta_1-\beta_0)}{1+\beta_0^{'}\beta_0}\right), & \beta_0\neq\mathbf{0}.\end{cases}$$

(iv) *(Weak instruments, $\frac{1}{\sqrt{n}K^{1/4}}$-local alternatives) Let $\Pi_{2(k)}=\frac{1}{\sqrt{n}}\Pi_{2(k)}^*$ with $\Pi_{2(k)}^*\sim_{iid}\left(\mathbf{0}_{1\times m},\Omega_{\Pi_2^*}\right)$ and $\Omega_{\Pi_2^*}>0$. Size-α tests of H_0 are performed against the local alternative*

$$H_1:\beta=\tilde{\beta}:=\beta_0+\frac{1}{K^{1/4}}(\beta_1-\beta_0)$$

using $\tilde{W}_{UR}^{\beta_0}$ and $\tilde{W}_{PR}^{\beta_0}$. Then, under (W-IV) and (A3'),

$$\mathcal{K}_{UR}(\beta_1;\beta_0)=\Phi\left(\Phi^{-1}(\alpha)-\frac{(\beta_1-\beta_0)^{'}\Omega_{\Pi_2^*}(\beta_1-\beta_0)}{\sqrt{2}\left(1+\beta_0^{'}\beta_0\right)}\right),$$

$$\mathcal{K}_{PR}(\beta_1;\beta_0)=\begin{cases}1, & \beta_0=\mathbf{0}\ \vee\ \beta_0^{'}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\Omega_{\Pi_2^*}(\beta_1-\beta_0)<0,\\ \Phi\left(\Phi^{-1}(\alpha)-\frac{\mu}{\sqrt{\mathbb{AV}ar\left[\tilde{W}_{PR}^{\beta_0}\right]}}\right), & \beta_0\neq\mathbf{0}\ \wedge\ \beta_0^{'}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\Omega_{\Pi_2^*}(\beta_1-\beta_0)=0,\\ 0, & \beta_0^{'}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\Omega_{\Pi_2^*}(\beta_1-\beta_0)>0.\end{cases}$$

where

$$\mu:=\frac{(\beta_1-\beta_0)^{'}\Omega_{\Pi_2^*}^{'}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\Omega_{\Pi_2^*}(\beta_1-\beta_0)}{1+\beta_0^{'}\beta_0}\geq 0,$$

and $\mathbb{AV}ar\left[\tilde{W}_{PR}^{\beta_0}\right]$ is some finite positive constant.

A particularly interesting outcome of Theorem 4 is the robust superiority of the PRRF test over the URRF test for the null hypothesis $H_0:\beta=\mathbf{0}$: irrespective of the instrument strength, so that $\mathcal{K}_{PR}(\beta_1;\mathbf{0})\geq\mathcal{K}_{UR}(\beta_1;\mathbf{0})$. Notably, with weak instruments, the URRF test is blind against local alternatives that converge to the null $\beta=\mathbf{0}$ at rates faster than $\frac{1}{K^{1/4}}$, while the PRRF test is informative against these and other local alternatives that converge at rates up to $\frac{1}{\sqrt{K}}$.

6 Correspondence with the Unstandardized Model

Removing the standardizing transformation (A2) on the variance matrix of V, we now assume that the rows $\left\{V_{(i)}\right\}_{i=1}^n$ of V have common variance $\mathbb{V}ar\left[V_{(i)}\right]=\Omega$, and use the triangular decomposition $\Omega = L'L$ where

$$L = \begin{pmatrix} \omega_{11\cdot 2}^{\frac{1}{2}} & \mathbf{0} \\ \Omega_{22}^{-\frac{1}{2}}\omega_{21} & \Omega_{22}^{\frac{1}{2}} \end{pmatrix}$$

with $\omega_{11\cdot 2} = \omega_{11} - \omega_{21}^{'}\Omega_{22}^{-1}\omega_{21}$. Let $Y^0 := YL^{-1}$, $\Pi^0 := \Pi L^{-1}$ and $V^0 := VL^{-1}$. Then $\mathbb{V}ar\left[V_{(i)}^0\right] = L^{-1'}\Omega L^{-1} = I$ so that this transformation leads to the standardized system (Phillips, 1983). Defining $\beta^0 = L\beta$, we obtain the standardized structural-form $y_1^0 = Y_2^0\beta^0 + u^0$ and corresponding reduced-form $Y^0 = Z\Pi^0 + V^0$, where $y_1^0 = \omega_{11\cdot 2}^{-\frac{1}{2}}\left(y_1 - Y_2\Omega_{22}^{-1}\omega_{21}\right)$, $Y_2^0 = Y_2\Omega_{22}^{-\frac{1}{2}}$, $\beta^0 = \omega_{11\cdot 2}^{-\frac{1}{2}}\Omega_{22}^{\frac{1}{2}}\left(\beta - \Omega_{22}^{-1}\omega_{21}\right)$, and $u^0 = \omega_{11\cdot 2}^{-\frac{1}{2}}u$. All previous results apply to this standardized model with variables superscripted by 0. Importantly, note that $E\left(u_i^0 Y_{2(i)}^0\right) = -\beta^0$, so that β^0 measures the extent of endogeneity in the standardized model.

Hypothesis testing on β in the unstandardized model corresponds to tests on the standardized parameter β^0 in the standardized model. Thus, $H_0 : \beta = \beta_0$ has the following standardized parametric form (with superscripts 0 denoting standardized parameters)

$$H_0 : \beta^0 = \beta_0^0 := \omega_{11\cdot 2}^{-\frac{1}{2}}\Omega_{22}^{\frac{1}{2}}\left(\beta_0 - \Omega_{22}^{-1}\omega_{21}\right).$$

We can then apply the results obtained in the standardized case in previous sections. In particular, we note the following correspondences.

(i) Testing the null hypothesis $H_0 : \beta^0 = 0$ in the standardized case is equivalent to testing

$$H_0 : \beta = \beta^* := \Omega_{22}^{-1}\omega_{21}$$

in the unstandardized system. This particular value β^* corresponds to the null hypothesis that Y_2 is exogenous in the structural equation, viz., $H_0 :\ \mathbb{E}\left[Y_{2(i)}u_{(i)}\right] = \mathbf{0}$. To see this, note that

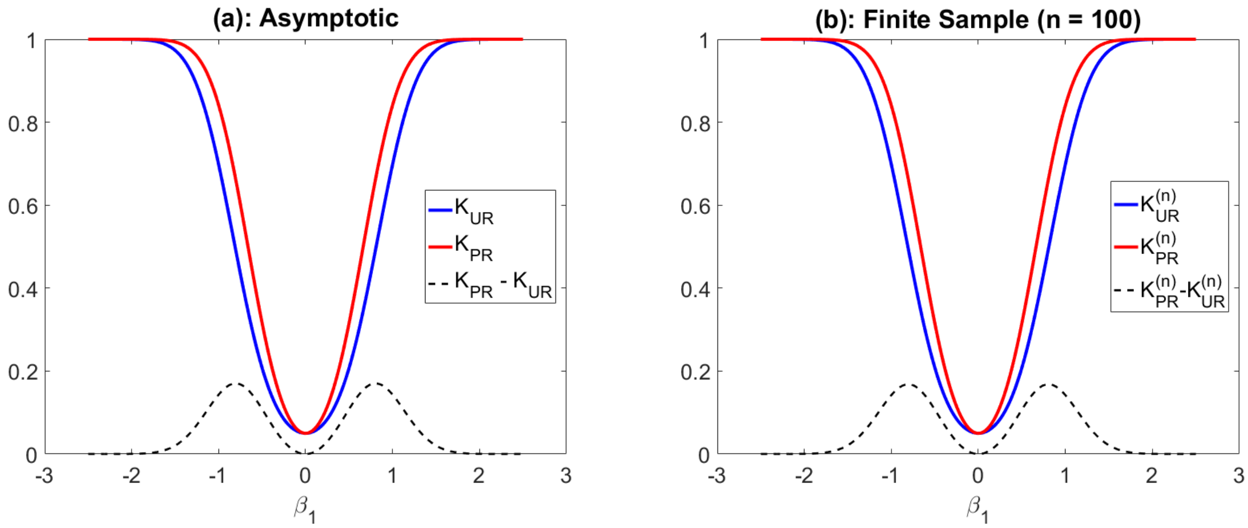
$$\begin{aligned}
\mathbb{E}\left[\underset{m\times 1}{Y_{2(i)}}\underset{1\times 1}{u_{(i)}}\right] &= \mathbb{E}\left[\left(\underset{m\times K}{\Pi_2^{'}}\underset{K\times 1}{Z_{(i)}} + \underset{m\times 1}{V_{2(i)}}\right)\left(\underset{1\times 1}{v_{1(i)}} - \underset{1\times m}{V_{2(i)}^{'}}\underset{m\times 1}{\beta}\right)\right] \\
&= \Pi_2^{'}\mathbb{E}\left[Z_{(i)}v_{1(i)}\right] - \Pi_2^{'}\mathbb{E}\left[Z_{(i)}V_{2(i)}^{'}\right]\beta + \mathbb{E}\left[V_{2(i)}v_{1(i)}\right] - \mathbb{E}\left[V_{2(i)}V_{2(i)}^{'}\right]\beta \\
&= \mathbf{0} + \mathbf{0} + \omega_{21} - \Omega_{22}\beta \\
&= \mathbf{0} \text{ if and only if } \beta = \beta^*.
\end{aligned}$$

Hence, all previous results for testing the null hypothesis $H_0 : \beta = 0$ in the standardized case correspond to tests of exogeneity of Y_2 in the unstandardized structural equation.

(ii) Tests of $H_0 : \beta = 0$ in the unstandardized system similarly correspond to tests in the standardized system of $H_0 : \beta^0 = \omega_{11\cdot 2}^{-\frac{1}{2}}\Omega_{22}^{-\frac{1}{2}}\omega_{21} = \rho/\left(1 - \rho'\rho\right)^{1/2}$, where $\rho = \omega_{11}^{-1/2}\Omega_{22}^{-\frac{1}{2}}\omega_{21}$ is the correlation vector of $y_{1(i)}$ and $Y_{2(i)}$ in the unstandardized model.

(iii) Results for testing a general null hypothesis such as $H_0 : \beta^0 = \beta_0^0$ for arbitrary β_0^0 correspond to similar general hypotheses in the unstandardized system.

Figure 1: Strong IV



7 Numerical Results

7.1 Numerical Evaluation of the Power Functions

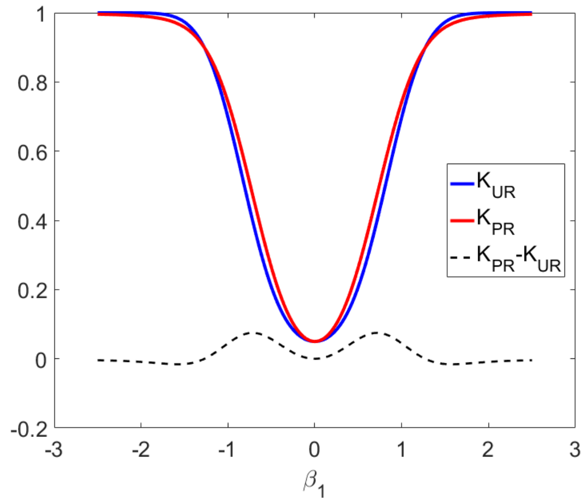
For simplicity in the following calculations we set $m = 1$ and $H_0 : \beta = 0$. The graphics are computed numerically using the analytic forms of the density functions given in Theorems 3 and 4. We consider separately the case of a fixed number K of instruments and the case of many instruments where K increases.

Case (i): Fixed Number of Instruments

We fix $K = 3$ and set $\Pi_2^* \approx (0.5377,\ 1.8339, 2.2588)^{'}$, based on random drawings from a standard normal distribution. For the case with strong instruments, i.e., $\Pi_2 = \Pi_2^*$, we plot in Figure 1a the asymptotic power functions $\mathcal{K}_{UR}(\beta_1; \mathbf{0})$, $\mathcal{K}_{PR}(\beta_1; \mathbf{0})$ of the tests against the local alternative $H_1 : \beta = \frac{1}{\sqrt{n}}\beta_1$ for different values of β_1. We also plot the finite-sample power functions $\mathcal{K}^{(n)}_{UR}(\beta_1; \mathbf{0})$, $\mathcal{K}^{(n)}_{PR}(\beta_1; \mathbf{0})$ of the same tests for $n = 100$ in Figure 1b. For the case with weak instruments, i.e., $\Pi_2 = \frac{1}{\sqrt{n}}\Pi_2^*$, we plot in Figure 2 the asymptotic (and finite-sample) power functions $\mathcal{K}_{UR}(\beta_1; \mathbf{0})$, $\mathcal{K}_{PR}(\beta_1; \mathbf{0})$ of the test against the alternative $H_1 : \beta = \beta_1$ for different values of β_1.

As noted in Section 4, the power comparison in the weak-IV case is ambiguous and depends on the value of Π_2^*. For the case shown in Figure 2, it is clear that power for the PRRF test exceeds that of the URRF test except when both powers are close to unity. To compare the "average" performance of the PRRF and URRF tests across multiple specifications of Π_2^*, we draw 150 different Π_2^* from normal distributions with three configurations of means and variances, and plot the average power functions in the three graphs of Figure 3. These graphs show that PRRF power continues to exceed power of the URRF test for alternatives

Figure 2: Weak IV



close to the null. When the elements of Π_2^* are very small on average as in Figure 3(a), the URRF power tends to exceed PRRF power for alternatives further from the null and when power for both tests is greater than 50%. But when the elements of Π_2^* are centred away from the origin as in Figure 3(c), PRRF power is uniformly greater than URRF power.

Case (ii): Many Instruments $K \to \infty$

For the strong-IV case, we assume $\Pi_{2(k)} \sim_{iid} (0, \Omega_{\Pi_2})$ and set $\Omega_{\Pi_2} = 1$. For the weak-IV case, we assume $\Pi_{2(k)}^* \sim_{iid} \left(0, \Omega_{\Pi_2}^*\right)$ and set $\Omega_{\Pi_2}^* = 1$.

In all of the cases considered, the gains are apparent from using the partially restricted reduced form for testing. The strong instrument and many instrument cases reveal unequivocal gains. The gains are especially evident in cases where there are many instruments (Figures 4 and 5). In the weak instrument case (Figure 2 and 3) the power gains are clear for all alternatives that are close to the null. Loosely speaking, if the magnitude of Π_2^* is large relative to the variance of Ξ, then only for alternatives far from the null where power for the PRRF and URRF tests are both closer to unity does the URRF power function exceed the PRRF power function. So even in the weak instrument case, strength in the remaining signal from the reduced form continues to matter in the performance of structural parameter tests.

8 Conclusion

One advantage of using reduced forms as a vehicle for testing structural hypotheses is that the effect of employing many instruments on testing is immediately apparent in the dimensional linkage between reduced form and structure. The partially restricted reduced form approach takes advantage of this linkage in using the additional information that comes

Figure 3: Weak IV: Average Power

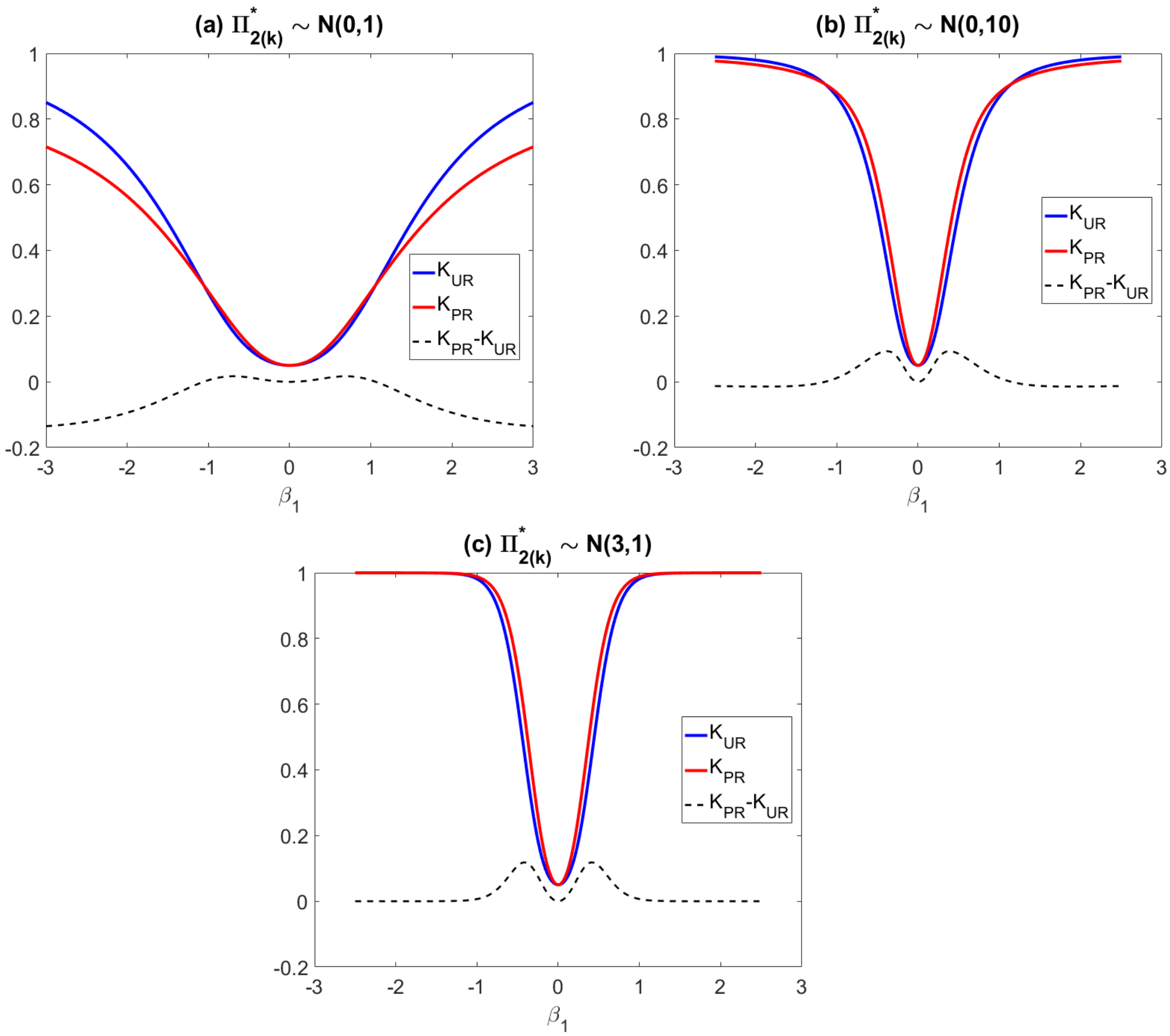


Figure 4: Strong IV, Asymptotic

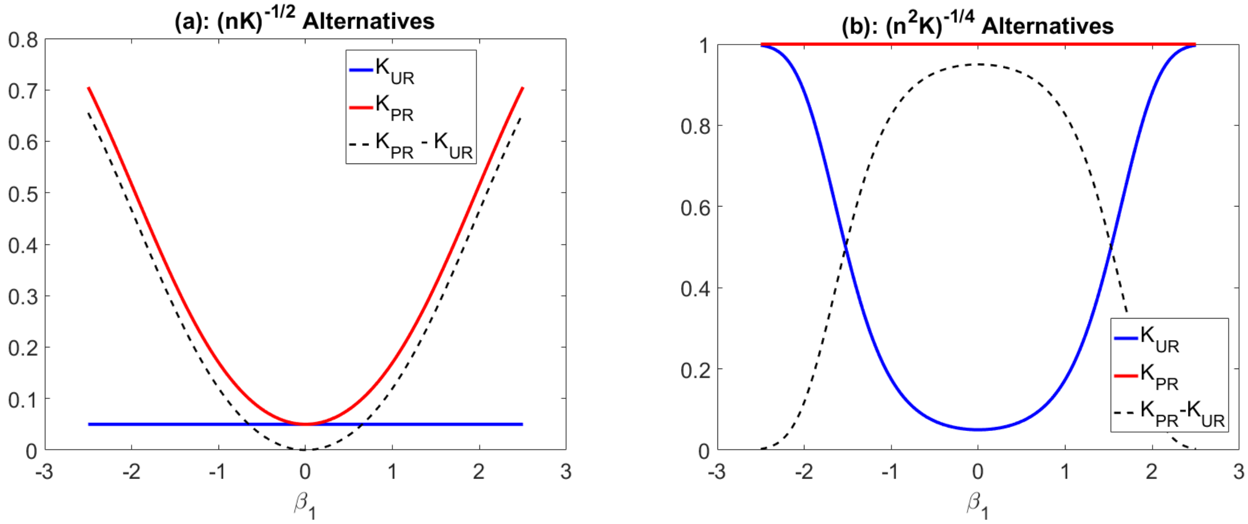
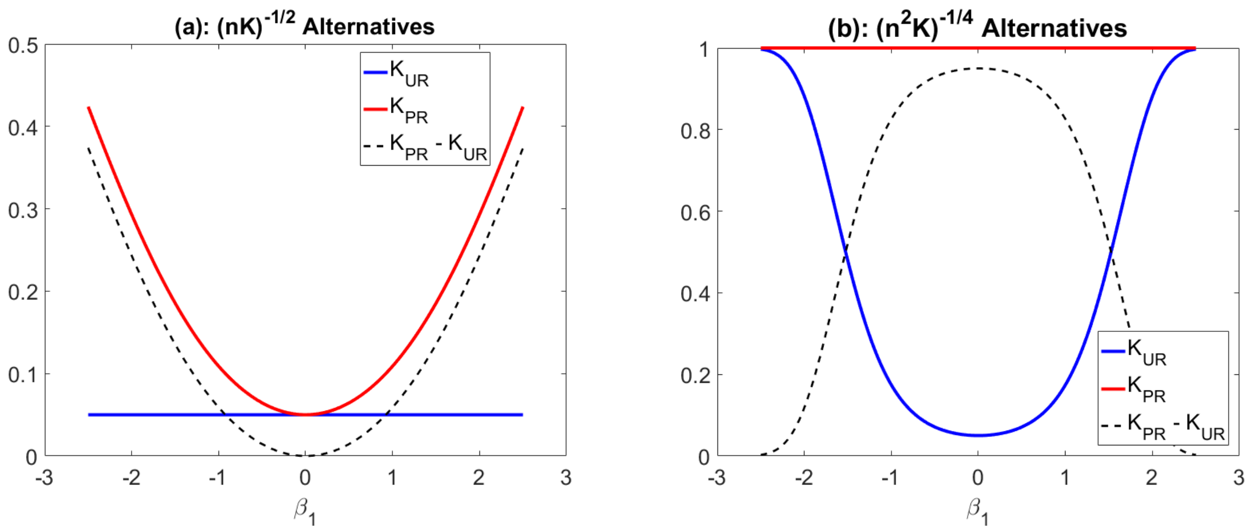


Figure 5: Weak IV



from a higher dimensional reduced form while at the same time exploiting the dimensional reduction of the restrictions that produce the structural parameters. The power gains from lower degrees of freedom in the chi-squared limit theory are especially notable when instruments are strong and the number of instruments is large, as might be expected. Gains are also apparent in weakly identified cases especially for local departures from the null, but they do not hold uniformly in the parameter space in this case. The results of the paper therefore help to reveal how strength and weakness in the reduced form are transmitted to structural coefficient testing. The approach taken in the paper also shows some of the close connections that exist between exact finite sample distributions and asymptotic theory in structural model testing, underlining the value of the trail pioneered by Basmann (1961) and Bergstrom (1962).

References

Anderson, T. W. and H. Rubin (1949): "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *The Annals of Mathematical Statistics*, 46–63. 1

Andrews, D. and J. Stock (2005): *Inference with Weak Instruments*, Cambridge: Cambridge University Press. 1

Andrews, D. W. and P. Guggenberger (Forthcoming): "Asymptotic Size of Kleibergen's LM and Conditional LR Tests for Moment Condition Models," *Econometric Theory.* 1

Andrews, D. W., M. J. Moreira, and J. H. Stock (2006): "Optimal Two-sided Invariant Similar Tests for Instrumental Variables Regression," *Econometrica*, 715–752. 1

Basmann, R. L. (1961): "A Note on the Exact Finite Sample Frequency Functions of Generalized Classical Linear Estimators in Two Leading Over-identified Cases," *Journal of the American Statistical Association*, 56, 619–636. 1, 8

Bekker, P. and F. Kleibergen (2003): "Finite-Sample Instrumental Variables Inference Using an Asymptotically Pivotal Statistic," *Econometric Theory*, 19, 744–753. 1

Bergstrom, A. (1962): "The Exact Sampling Distributions of Least Squares and Maximum Likelihood Estimators of the Marginal Propensity to Consume," *Econometrica: Journal of the Econometric Society*, 480–490. 8

Chen, J. and H. Rubin (1986): "Bounds for the difference between median and mean of Gamma and Poisson distributions," *Statistics & probability letters*, 4, 281–283. 8

Chernozhukov, V. and C. Hansen (2008): "Instrumental Variable Quantile Regression: A Robust Inference Approach," *Journal of Econometrics*, 142, 379–398. 1, 4, 1

Inglot, T. (2010): "Inequalities for quantiles of the chi-square distribution," *Probability and Mathematical Statistics*, 30, 339–351. 8, 8

Kakwani, N. C. and R. Court (1972): "Reduced Form Coefficient Estimation and Forecasting from a Simultaneous Equation Model," *Australian Journal of Statistics*, 14, 143–160. 1, 2

Kleibergen, F. (2002): "Pivotal Statistics for Testing Structural Parameters in Instrumental Variables Regression," *Econometrica*, 70, 1781–1803. 1

Knight, J. L. (1977): "On the Existence of Moments of the Partially Restricted Reduced-Form Estimators from a Simultaneous-Equation Model," *Journal of Econometrics*, 5, 315–321. 2

Laurent, B. and P. Massart (2000): "Adaptive estimation of a quadratic functional by model selection," *Annals of Statistics*, 1302–1338. 8

Moreira, M. J. (2003): "A Conditional Likelihood Ratio Test for Structural Models," *Econometrica*, 71, 1027–1048. 1

Muirhead, R. J. (2005): *Aspects of Multivariate Statistical Theory*, vol. 197, John Wiley & Sons. 2

Phillips, P. C. B. (1980): "The Exact Distribution of Instrumental Variable Estimators in an Equation Containing n+1 Endogenous Variables," *Econometrica: Journal of the Econometric Society*, 861–878. 1, 3.1, 8

——— (1983): "Exact Small Sample Theory in the Simultaneous Equations Model," *Handbook of econometrics*, 1, 449–516. 3

——— (1984): "The Exact Distribution of LIML: I," *International Economic Review*, 249–261. 3.3

——— (1989): "Partially Identified Econometric Models," *Econometric Theory*, 5, 181–240. 1, 3

——— (2006): "A Remark on Bimodality and Weak Instrumentation in Structural Equation Estimation," *Econometric Theory*, 22, 947–960. 1

——— (Forthcoming): "Reduced Forms and Weak Instrumentation," *Econometric Reviews*. 1, 2, 3.1

Staiger, D. and J. Stock (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65, 557–586. 1

Appendix: Proofs

Proof of Lemma 1

Proof. Under (S-IV) and (A3′),

$$\begin{aligned}
\sqrt{n}\left(\hat{\Pi}-\Pi\right) &= \frac{1}{\sqrt{n}}Z'V \xrightarrow{a.s.} (\xi,\Xi)\,,\\
\sqrt{n}\left(\beta_{IV}-\beta\right) &= \left(\hat{\Pi}_2'\hat{\Pi}_2\right)^{-1}\hat{\Pi}_2'\left[\sqrt{n}\left(\hat{\pi}_1-\pi_1\right)-\sqrt{n}\left(\hat{\Pi}_2-\Pi_2\right)\beta\right]\\
&\xrightarrow{a.s.} \left(\Pi_2'\Pi_2\right)^{-1}\Pi_2'\left(\xi-\Xi\beta\right),\\
\sqrt{n}\left(\tilde{\pi}_1-\pi_1\right) &= \sqrt{n}\left(\hat{\Pi}_2-\Pi_2\right)\left(\beta_{IV}-\beta\right)+\Pi_2\sqrt{n}\left(\beta_{IV}-\beta\right)+\sqrt{n}\left(\hat{\Pi}_2-\Pi_2\right)\beta\\
&\xrightarrow{a.s.} P_{\Pi_2}\left(\xi-\Xi\beta\right)+\Xi\beta = P_{\Pi_2}\xi+M_{\Pi_2}\Xi\beta,
\end{aligned}$$

establishing the limit theory (5) - (7).

Under (S-IV) and (A1), $\hat{\Pi} \overset{d}{=} \mathcal{N}_{K,m+1}\left(\Pi, n^{-1}I_{K(m+1)}\right)$. So the matrix quadratic form $n\hat{\Pi}'\hat{\Pi} = Y'P_ZY = Y'CC'Y$, where $C = Z\left(Z'Z\right)^{-1/2}$, is distributed as noncentral Wishart of dimension $m+1$ with covariance matrix I_{m+1} and noncentrality matrix MM' where

$$M' = \mathbb{E}\left(C'Y\right) = \left(Z'Z\right)^{-\frac{1}{2}}Z'Z\Pi = \sqrt{n}\left[\pi_1,\Pi_2\right] = \sqrt{n}\Pi_2\left[\beta, I_m\right].$$

This distribution is written as $W_{m+1}\left(K, I_{m+1}, MM'\right)$. Partitioning the Wishart matrix $W := Y'CC'Y$ conformably with the structural equation (1), we can write the matrix quadratic form

$$W = \begin{array}{c} \\ \\ \end{array}\begin{array}{cc} \scriptstyle 1 & \scriptstyle m \\ \end{array}$$

$$W = \left[\begin{array}{cc} \overset{1}{w_{11}} & \overset{m}{w_{12}} \\ w_{21} & W_{22} \end{array}\right] = \left[\begin{array}{cc} y_1'P_zy_1 & y_1'P_zY_2 \\ Y_2'P_zy_1 & Y_2'P_zY_2 \end{array}\right] \overset{d}{=} W_{m+1}\left(K, I_{m+1}, MM'\right), \tag{19}$$

where the noncentrality matrix

$$MM' = n\left[\begin{array}{c} \beta' \\ I_m \end{array}\right]\Pi_2'\Pi_2\left[\beta, I_m\right], \tag{20}$$

has deficient rank m. Then $\beta_{IV} = \left(\hat{\Pi}_2'\hat{\Pi}_2\right)'\hat{\Pi}_2'\hat{\pi}_1 \overset{d}{=} W_{22}^{-1}w_{21}$, showing the exact finite sample distribution of β_{IV} to be a matrix quotient of the components of the non-central Wishart matrix W. The analytic form of this density is derived in Phillips (1980). We may also write this distribution in mixed normal form as $\beta_{IV} \overset{d}{=} \mathcal{MN}\left(\sqrt{n}\left(A_\Xi'A_\Xi\right)^{-1}A_\Xi\Pi_2\beta, \left(A_\Xi'A_\Xi\right)^{-1}\right)$, where $A_\Xi = \sqrt{n}\Pi_2+\Xi$, by noting that $A_\Xi = C'Y_2 \overset{d}{=} \mathcal{N}_{K,m}\left(\sqrt{n}\Pi_2, I_{K(m+1)}\right)$ and $C'y_1 \overset{d}{=} \mathcal{N}_K\left(\sqrt{n}\Pi_2\beta, I_K\right)$. Continuing under the Gaussian assumption (A1) we have

$$\tilde{\pi}_1 = P_{\hat{\Pi}_2}\hat{\pi}_1 \overset{d}{=} P_{A_\Xi}\left(\Pi_2\beta+\frac{1}{\sqrt{n}}\xi\right) \overset{d}{=} \mathcal{MN}\left(P_{A_\Xi}\Pi_2\beta, \frac{1}{n}P_{A_\Xi}\right)$$

again with $\Xi \overset{d}{=} \mathcal{N}_{K,m}\left(\mathbf{0}, I_{Km}\right)$ and $P_{A_\Xi} = A_\Xi\left(A_\Xi'A_\Xi\right)^{-1}A_\Xi'$. ∎

Proof of Lemma 2

Proof. Under (I-IV) and (A3'),

$$\sqrt{n}\hat{\Pi} = \frac{1}{\sqrt{n}}ZV \xrightarrow{a.s.} (\xi, \Xi) \overset{d}{=} \mathcal{N}_{K,m+1}\left(\mathbf{0}, I_{K(m+1)}\right)$$

$$\beta_{IV} = \left(\sqrt{n}\hat{\Pi}_2'\sqrt{n}\hat{\Pi}_2\right)^{-1}\sqrt{n}\hat{\Pi}_2\sqrt{n}\hat{\pi}_1 \xrightarrow{a.s.} \left(\Xi'\Xi\right)^{-1}\Xi'\xi \overset{d}{=} \mathcal{MN}\left(\mathbf{0}, \left(\Xi'\Xi\right)^{-1}\right)$$

$$\sqrt{n}\tilde{\pi}_1 = \sqrt{n}\hat{\Pi}_2\left(\sqrt{n}\hat{\Pi}_2'\sqrt{n}\hat{\Pi}_2\right)^{-1}\sqrt{n}\hat{\Pi}_2\sqrt{n}\hat{\pi}_1 \xrightarrow{a.s.} P_{\Xi}\xi \overset{d}{=} \mathcal{MN}\left(\mathbf{0}, P_{\Xi}\right)$$

Under (I-IV) and (A1), $\sqrt{n}\hat{\Pi} = \frac{1}{\sqrt{n}}ZV \overset{d}{=} (\xi, \Xi)$, and the stated results follow. ∎

Proof of Lemma 3

Proof. Under (W-IV) and (A3'), notice that

$$\begin{aligned}
\sqrt{n}\hat{\Pi} &= \frac{1}{\sqrt{n}}Z'\left(Z\Pi + V\right) = \Pi^* + \frac{1}{\sqrt{n}}Z'V \xrightarrow{a.s.} \Pi_2^* + (\xi, \Xi),\\
\beta_{IV} &\xrightarrow{a.s.} \left[\left(\Pi_2^* + \Xi\right)'\left(\Pi_2^* + \Xi\right)\right]^{-1}\left(\Pi_2^* + \Xi\right)'\left(\Pi_2^*\beta + \xi\right)\\
\sqrt{n}\tilde{\pi}_1 &= \sqrt{n}\hat{\Pi}_2\left[\sqrt{n}\hat{\Pi}_2'\sqrt{n}\hat{\Pi}_2\right]^{-1}\sqrt{n}\hat{\Pi}_2'\sqrt{n}\hat{\pi}_1\\
&\xrightarrow{a.s.} P_{\left(\Pi_2^*+\Xi\right)}\left(\pi_1^* + \xi\right) = P_{\left(\Pi_2^*+\Xi\right)}\left(\Pi_2^*\beta + \xi\right).
\end{aligned}$$

Under (W-IV) and (A1), $\sqrt{n}\hat{\Pi} \overset{d}{=} \Pi_2^* + (\xi, \Xi)$ and the stated results follow. ∎

Proof of Theorem 2

Proof. We prove the proposition for the three cases with strong, irrelevant and weak instruments separately.

(i). Strong instruments: Under the null $\pi_1 = \Pi_2\beta_0$,

$$\begin{aligned}
\sqrt{n}\left(\tilde{\pi}_1 - \hat{\Pi}_2\beta_0\right) &= \sqrt{n}\hat{\Pi}_2\left(\beta_{IV} - \beta_0\right) = \sqrt{n}\hat{\Pi}_2\left(\left(\hat{\Pi}_2'\hat{\Pi}_2\right)^{-1}\hat{\Pi}_2'\hat{\pi}_1 - \beta_0\right)\\
&= P_{\hat{\Pi}_2}\left[\sqrt{n}\left(\hat{\pi}_1 - \pi_1\right) + \sqrt{n}\left(\Pi_2 - \hat{\Pi}_2\right)\beta_0\right]\\
&\xrightarrow{a.s.} P_{\Pi_2}\left(\xi - \Xi\beta_0\right) \overset{d}{=} \mathcal{N}\left(\mathbf{0}, \left(1 + \beta_0'\beta_0\right)P_{\Pi_2}\right)
\end{aligned}$$

Define $C := \Pi_2\left(\Pi_2'\Pi_2\right)^{-\frac{1}{2}}$. Then $CC' = P_{\Pi_2}$ and $C'C = I_m$, so

$$\begin{aligned}
C'\xi &\overset{d}{=} \mathcal{N}\left(\mathbf{0}, C'C = I_m\right),\\
C'\Xi &\overset{d}{=} \mathcal{N}_{m,m}\left(\mathbf{0}, I_{mm}\right),\\
C'\xi - C'\Xi\beta_0 &\overset{d}{=} \mathcal{N}_{K,m}\left(\mathbf{0}, \left(1 + \beta_0'\beta_0\right)I_m\right).
\end{aligned}$$

Hence,

$$
\begin{aligned}
W_{PR}^{\beta_0} &= \frac{n\left(\tilde{\pi}_1 - \hat{\Pi}_2\beta_0\right)^{'}\left(\tilde{\pi}_1 - \hat{\Pi}_2\beta_0\right)}{1+\beta_0^{'}\beta_0} \\
&\xrightarrow{a.s.} \frac{\left(\xi - \Xi\beta_0\right)^{'} P_{\Pi_2}\left(\xi - \Xi\beta_0\right)}{1+\beta_0^{'}\beta_0} = \frac{\left(C'\xi - C'\Xi\beta_0\right)^{'}\left(C'\xi - C'\Xi\beta_0\right)}{1+\beta_0^{'}\beta_0} \\
&\overset{d}{=} \frac{1}{1+\beta_0^{'}\beta_0}\mathcal{M}\chi^2_{m,\beta_0^{'}\Xi' P_{\Pi_2}\Xi\beta_0} \overset{d}{=} \chi^2_m.
\end{aligned}
$$

As $C'\Xi\beta_0 \overset{d}{=} \mathcal{N}\left(\mathbf{0}, \beta_0^{'}\beta_0 \cdot I_m\right)$,

$$
\frac{1}{\beta_0^{'}\beta_0}\lambda_S\left(\Xi, \beta_0\right) = \frac{\beta_0^{'}\Xi^{'}C}{\sqrt{\beta_0^{'}\beta_0}}\left(\frac{C'\Xi\beta_0}{\sqrt{\beta_0^{'}\beta_0}}\right) \overset{d}{=} \chi^2_m.
$$

Under the Gaussianity assumption (A1), note that

$$
\begin{aligned}
\sqrt{n}\left(\tilde{\pi}_1 - \hat{\Pi}_2\beta_0\right) &= P_{\hat{\Pi}_2}\left[\sqrt{n}\left(\hat{\pi}_1 - \pi_1\right) + \sqrt{n}\left(\Pi_2 - \hat{\Pi}_2\right)\beta_0\right] \\
&\overset{d}{=} P_{A_\Xi}\left(\xi - \Xi\beta_0\right) \overset{d}{=} \mathcal{MN}\left(-P_{A_\Xi}\Xi\beta_0, P_{A_\Xi}\right)
\end{aligned}
$$

Hence,

$$
W_{PR}^{\beta_0} \overset{d}{=} \frac{\left(\xi - \Xi\beta_0\right)^{'} P_{A_\Xi}\left(\xi - \Xi\beta_0\right)}{1+\beta_0^{'}\beta_0} \overset{d}{=} \frac{1}{1+\beta_0^{'}\beta_0}\mathcal{M}\chi^2_{m,\beta_0^{'}\Xi' P_{A_\Xi}\Xi\beta_0}
$$

(ii). Irrelevant instruments: Under the null $\pi_1 = \Pi_2\beta_0 \equiv 0$,

$$
\begin{aligned}
\sqrt{n}\left(\tilde{\pi}_1 - \hat{\Pi}_2\beta_0\right) &= \sqrt{n}\hat{\Pi}_2\left(\beta_{IV} - \beta_0\right) \\
&= \frac{1}{\sqrt{n}}Z'V_2\left[\left(\frac{1}{\sqrt{n}}V_2^{'}Z\frac{1}{\sqrt{n}}Z'V_2\right)^{-1}\frac{1}{\sqrt{n}}V_2^{'}Z\frac{1}{\sqrt{n}}Z'v_1 - \beta_0\right] \\
&\xrightarrow{a.s.} P_\Xi\left(\xi - \Xi\beta_0\right) d = \mathcal{N}\left(-\Xi\beta_0, P_\Xi\right)
\end{aligned}
$$

Define $\Upsilon := \Xi\left(\Xi'\Xi\right)^{-\frac{1}{2}}$. As $\Upsilon'\Upsilon = I$, the conditional distribution of $\Upsilon'\xi$ given any realization of Υ is $\mathcal{N}\left(\mathbf{0}, \Upsilon'\Upsilon = I_m\right)$, which does not depend on Υ, so

$$
\Upsilon'\xi \overset{d}{=} \mathcal{N}\left(\mathbf{0}, I_m\right).
$$

Hence,

$$
\begin{aligned}
W_{PR}^{\beta_0} &= \frac{n\left(\tilde{\pi}_1 - \hat{\Pi}_2\beta_0\right)^{'}\left(\tilde{\pi}_1 - \hat{\Pi}_2\beta_0\right)}{1+\beta_0^{'}\beta_0} \\
&\xrightarrow{a.s.} \frac{\left(\xi - \Xi\beta_0\right)^{'} P_{\Xi}\left(\xi - \Xi\beta_0\right)}{1+\beta_0^{'}\beta_0} = \frac{\left(\Upsilon'\xi - \Upsilon'\Xi\beta_0\right)^{'}\left(\Upsilon'\xi - \Upsilon'\Xi\beta_0\right)}{1+\beta_0^{'}\beta_0} \\
&\stackrel{d}{=} \frac{1}{1+\beta_0^{'}\beta_0} Mixed\_\chi^2_{m,\beta_0'\Xi'\Xi\beta_0}.
\end{aligned}
$$

As $\Xi\beta_0 \stackrel{d}{=} \mathcal{N}\left(\mathbf{0}, \beta_0^{'}\beta_0 \cdot I_K\right)$,

$$
\frac{1}{\beta_0^{'}\beta_0}\lambda_U\left(\Xi, \beta_0\right) = \frac{\beta_0^{'}\Xi'}{\sqrt{\beta_0^{'}\beta_0}}\frac{\Xi\beta_0}{\sqrt{\beta_0^{'}\beta_0}} \stackrel{d}{=} \chi_K^2.
$$

(iii). Weak instruments: Under the null $\pi_1 = \Pi_2\beta_0 \equiv 0$,

$$
\begin{aligned}
\sqrt{n}\left(\tilde{\pi}_1 - \hat{\Pi}_2\beta_0\right) &= \sqrt{n}\hat{\Pi}_2\left(\beta_{IV} - \beta_0\right) \\
&= \left(\Pi_2^* + \frac{1}{\sqrt{n}}ZV_2\right)\left\{\left[\left(\Pi_2^* + \frac{1}{\sqrt{n}}ZV_2\right)^{'}\left(\Pi_2^* + \frac{1}{\sqrt{n}}ZV_2\right)\right]^{-1}\right. \\
&\qquad \left.\left(\Pi_2^* + \frac{1}{\sqrt{n}}ZV_2\right)^{'}\left(\Pi_2^*\beta_0 + \frac{1}{\sqrt{n}}Z'v_1\right) - \beta_0\right\} \\
&= \left(\Pi_2^* + \frac{1}{\sqrt{n}}ZV_2\right)\left[\left(\Pi_2^* + \frac{1}{\sqrt{n}}ZV_2\right)^{'}\left(\Pi_2^* + \frac{1}{\sqrt{n}}ZV_2\right)\right]^{-1} \\
&\qquad \left(\Pi_2^* + \frac{1}{\sqrt{n}}ZV_2\right)^{'}\left(\frac{1}{\sqrt{n}}Z'v_1 - \frac{1}{\sqrt{n}}Z'V_2\beta_0\right) \\
&\xrightarrow{a.s.} P_{\left(\Pi_2^*+\Xi\right)}\left(\xi - \Xi\beta_0\right) \stackrel{d}{=} \mathcal{MN}\left(-P_{\left(\Pi_2^*+\Xi\right)}\Xi\beta_0, P_{\left(\Pi_2^*+\Xi\right)}\right).
\end{aligned}
$$

Define $\Upsilon_{\Pi_2^*} := \left(\Pi_2^* + \Xi\right)\left(\left(\Pi_2^* + \Xi\right)^{'}\left(\Pi_2^* + \Xi\right)\right)^{-\frac{1}{2}}$. As $\Upsilon_{\Pi_2^*}^{'}\Upsilon_{\Pi_2^*} = I$, the condition distribution of $\Upsilon_{\Pi_2^*}^{'}\xi$ given any realization of Ξ is $N\left(\mathbf{0}, \Upsilon_{\Pi_2^*}^{'}\Upsilon_{\Pi_2^*} = I_m\right)$, which does not depend on Ξ, so

$$
\Upsilon_{\Pi_2^*}^{'}\xi \stackrel{d}{=} \mathcal{N}\left(\mathbf{0}, I_m\right).
$$

Hence, as $\varUpsilon_{\Pi_2^*}\varUpsilon'_{\Pi_2^*} = P_{\left(\Pi_2^*+\Xi\right)}$,

$$
\begin{aligned}
W_{PR}^{\beta_0} &= \frac{n\left(\tilde{\pi}_1 - \hat{\Pi}_2\beta_0\right)'\left(\tilde{\pi}_1 - \hat{\Pi}_2\beta_0\right)}{1+\beta_0'\beta_0} \\
&\xrightarrow{a.s.} \frac{\left(\xi - \Xi\beta_0\right)' P_{\left(\Pi_2^*+\Xi\right)}\left(\xi - \Xi\beta_0\right)}{1+\beta_0'\beta_0} = \frac{\left(\varUpsilon'_{\Pi_2^*}\xi - \varUpsilon'_{\Pi_2^*}\Xi\beta_0\right)'\left(\varUpsilon'_{\Pi_2^*}\xi - \varUpsilon'_{\Pi_2^*}\Xi\beta_0\right)}{1+\beta_0'\beta_0} \\
&\overset{d}{=} \frac{1}{1+\beta_0'\beta_0}\mathcal{M}\chi^2_{m,\beta_0'\Xi' P_{\left(\Pi_2^*+\Xi\right)}\Xi\beta_0}.
\end{aligned}
$$

∎

Proof of Lemma 4

Proof. (i). Let $\xi_i \sim_{iid} \mathcal{N}(0,1)$. $\forall x > 0$,

$$
\varPsi(x; k+1, 0) = \mathbb{P}\left(\sum_{i=1}^{k+1}\xi_i^2 \geq x\right) > \mathbb{P}\left(\sum_{i=1}^{k}\xi_i^2 \geq x\right) = \varPsi(x; k, 0).
$$

Then, $\forall \lambda > 0$,

$$
\begin{aligned}
\varPsi(x; k+1, \lambda) &= \sum_{j=0}^{\infty}\frac{e^{-\frac{\lambda}{2}}\left(\frac{\lambda}{2}\right)^j}{j!}\varPsi(x; k+1+2j, 0) \\
&\geq \sum_{j=0}^{\infty}\frac{e^{-\frac{\lambda}{2}}\left(\frac{\lambda}{2}\right)^j}{j!}\varPsi(x; k+2j, 0) = \varPsi(x; k, \lambda).
\end{aligned}
$$

(ii). $\forall x > 0, k \in \mathbb{N}_+$,

$$
\begin{aligned}
\frac{\partial}{\partial\lambda}\varPsi(x; k, \lambda) &= -\frac{1}{2}e^{-\frac{\lambda}{2}}\varPsi(x; k, 0) + \sum_{j=1}^{\infty}\frac{e^{-\frac{\lambda}{2}}\left(\frac{\lambda}{2}\right)^{j-1}}{j!}\left(\frac{j}{2} - \frac{\lambda}{4}\right)\cdot\varPsi(x; k+2j, 0) \\
&= \frac{1}{2}\sum_{j=1}^{\infty}\frac{e^{-\frac{\lambda}{2}}\left(\frac{\lambda}{2}\right)^{j-1}}{(j-1)!}\varPsi(x; k+2j, 0) - \frac{1}{2}\sum_{j=0}^{\infty}\frac{e^{-\frac{\lambda}{2}}\left(\frac{\lambda}{2}\right)^{j}}{j!}\varPsi(x; k+2j, 0) \\
&= \frac{1}{2}\left[\varPsi(x; k+2, \lambda) - \varPsi(x; k, \lambda)\right] \qquad (21) \\
&> 0 \quad \text{by (i)}.
\end{aligned}
$$

(iii). As $\varPsi(x; k+1, 0) > \varPsi(x; k, 0)\ \forall x > 0,\ \forall k \in \mathbb{N}_+$, we have, $\forall\alpha \in (0,1)$,

$$
q_{k+1}^{1-\alpha} > q_k^{1-\alpha}.
$$

For any even $k \in \mathbb{N}_+$, define

$$g_2(\lambda) := \Psi\left(q_{k+2}^{1-\alpha}; k+2, \lambda\right) - \Psi\left(q_k^{1-\alpha}; k, \lambda\right).$$

Taking the first-order derivative at $\lambda = 0$, by (21),

$$\begin{aligned} g_2^{'}(0) &= \frac{1}{2}\left[\Psi\left(q_{k+2}^{1-\alpha}; k+4, 0\right) - \alpha - \Psi\left(q_k^{1-\alpha}; k+2, 0\right) + \alpha\right] \\ &= \frac{1}{2}\left[\frac{\Gamma\left(\frac{k+4}{2}, \frac{q_{k+2}^{1-\alpha}}{2}\right)}{\Gamma\left(\frac{k+4}{2}\right)} - \frac{\Gamma\left(\frac{k+2}{2}, \frac{q_k^{1-\alpha}}{2}\right)}{\Gamma\left(\frac{k+2}{2}\right)}\right], \end{aligned}$$

where $\Gamma(s, x)$ denotes the upper incomplete gamma function. For integer s, we have the finite series representation

$$\Gamma(s, x) = (s-1)! \cdot e^{-x} \sum_{j=0}^{s-1} \frac{x^j}{j!}.$$

Using a slight abuse of notation, it is frequently convenient in the following to suppress the index $1-\alpha$ in $q_k^{1-\alpha}$ and instead use a superscript on q_k to denote powers of q_k as in: $q_k^x \equiv \left(q_k^{1-\alpha}\right)^x$. However, whenever α becomes relevant in derivations, the index $1-\alpha$ will be retained in $q_k^{1-\alpha}$. With this notation in mind, for even k, we have

$$\begin{aligned} g_2^{'}(0) &= \frac{1}{2}\left[\frac{\left(\frac{k}{2}+1\right)! \cdot e^{-\frac{q_{k+2}}{2}} \sum_{j=0}^{\frac{k}{2}+1} \frac{q_{k+2}^j}{2^j j!}}{\left(\frac{k}{2}+1\right)!} - \frac{\left(\frac{k}{2}\right)! \cdot e^{-\frac{q_k}{2}} \sum_{j=0}^{\frac{k}{2}} \frac{q_k^j}{2^j j!}}{\left(\frac{k}{2}\right)!}\right] \\ &= \frac{1}{2}\left[e^{-q_{k+2}} \sum_{j=0}^{k/2} \frac{q_{k+2}^j}{2^j j!} + e^{-q_{k+2}} \frac{q_{k+2}^{\frac{k}{2}+1}}{2^{\frac{k}{2}+1}\left(\frac{k}{2}+1\right)!} - e^{-q_k} \sum_{j=0}^{k/2-1} \frac{q_k^j}{2^j j!} - e^{-q_k} \frac{q_k^{\frac{k}{2}}}{2^{\frac{k}{2}}\left(\frac{k}{2}\right)!}\right] \\ &= \frac{1}{2}\left[\alpha + e^{-\frac{q_{k+2}}{2}} \frac{q_{k+2}^{\frac{k}{2}+1}}{2^{\frac{k}{2}+1}\left(\frac{k}{2}+1\right)!} - \alpha - e^{-\frac{q_k}{2}} \frac{q_k^{\frac{k}{2}}}{2^{\frac{k}{2}}\left(\frac{k}{2}\right)!}\right] \\ &= \frac{1}{2}\left[e^{-\frac{q_{k+2}}{2}} \frac{q_{k+2}^{\frac{k}{2}+1}}{2^{\frac{k}{2}+1}\left(\frac{k}{2}+1\right)!} - e^{-\frac{q_k}{2}} \frac{q_k^{\frac{k}{2}}}{2^{\frac{k}{2}}\left(\frac{k}{2}\right)!}\right] \\ &= \frac{1}{2^{\frac{k}{2}+2}\left(\frac{k}{2}+1\right)!}\left[e^{-\frac{q_{k+2}}{2}} q_{k+2}^{\frac{k}{2}+1} - (k+2)\, e^{-\frac{q_k}{2}} q_k^{\frac{k}{2}}\right], \end{aligned} \tag{22}$$

where the third line follows from the observation that:

$$\alpha = \Psi(q_k; k, 0) = e^{-\frac{q_k}{2}} \sum_{j=0}^{\frac{k}{2}-1} \frac{q_k^j}{2^j j!}. \tag{23}$$

Note that the ratio of the two terms in the square bracket in (22) is

$$\frac{e^{-\frac{q_{k+2}}{2}} q_{k+2}^{\frac{k}{2}+1}}{(k+2)\, e^{-\frac{q_k}{2}} q_k^{\frac{k}{2}}} = \left(\frac{q_{k+2}}{q_k}\right)^{\frac{k}{2}} \cdot \frac{q_{k+2}}{k+2} \cdot \frac{1}{e^{\frac{1}{2}(q_{k+2}-q_k)}}.$$

Taking logarithms, we have

$$\begin{aligned}
\log\left(\frac{e^{-\frac{q_{k+2}}{2}} q_{k+2}^{\frac{k}{2}+1}}{(k+2)\, e^{-\frac{q_k}{2}} q_k^{\frac{k}{2}}}\right) &= \frac{k}{2}\left(\log q_{k+2} - \log q_k\right) + \log q_{k+2} - \log(k+2) + \frac{1}{2}\left(q_k - q_{k+2}\right) \\
&\le \frac{k}{2}\left(q_{k+2} - q_k\right) \cdot \frac{1}{q_k} + \log q_{k+2} - \log(k+2) + \frac{1}{2}\left(q_k - q_{k+2}\right) \\
&= \log q_{k+2} - \log(k+2) - \frac{1}{2}\left(q_{k+2} - q_k\right)\left(1 - \frac{k}{q_k}\right),
\end{aligned}$$

where the inequality follows from the mean-value theorem and the fact that $q_{k+2} > q_k$. By Inglot (2010, Proposition 5.1), for $k \ge 2$ and $\alpha \le 0.17$, we have the inequality

$$q_k \ge k + 2\log\frac{1}{\alpha} - \frac{5}{2},$$

so that

$$\log\left(\frac{e^{-\frac{q_{k+2}}{2}} q_{k+2}^{\frac{k}{2}+1}}{(k+2)\, e^{-\frac{q_k}{2}} q_k^{\frac{k}{2}}}\right) \le \log q_{k+2} - \log(k+2) - \frac{1}{2}\left(q_{k+2} - q_k\right)\left(1 - \frac{k}{2k + 4\log\frac{1}{\alpha} - 5}\right), \quad (24)$$

the last term of which

$$1 - \frac{k}{2k + 4\log\frac{1}{\alpha} - 5} \to 1 \quad \text{as } \alpha \searrow 0.$$

Now, recall that $\alpha e^{\frac{q_{k+2}^{1-\alpha}}{2}} = \sum_{j=0}^{\frac{k}{2}} \frac{\left(q_{k+2}^{1-\alpha}\right)^j}{2^j j!}$ by (23). Taking derivatives with respect to α on both sides, we have

$$\begin{aligned}
e^{\frac{q_{k+2}^{1-\alpha}}{2}} + e^{\frac{q_{k+2}^{1-\alpha}}{2}} \cdot \frac{1}{2}\frac{\partial}{\partial\alpha} q_{k+2}^{1-\alpha} &= \sum_{j=0}^{\frac{k}{2}} \frac{j\left(q_{k+2}^{1-\alpha}\right)^{j-1}}{2^j j!} \cdot \frac{\partial}{\partial\alpha} q_{k+2}^{1-\alpha} = \frac{1}{2}\frac{\partial}{\partial\alpha} q_{k+2}^{1-\alpha} \cdot \sum_{j=0}^{\frac{k}{2}-1} \frac{\left(q_{k+2}^{1-\alpha}\right)^j}{2^j j!} \\
&= \frac{1}{2}\frac{\partial}{\partial\alpha} q_{k+2}^{1-\alpha} \cdot \left[\alpha e^{\frac{q_{k+2}^{1-\alpha}}{2}} - \frac{\left(q_{k+2}^{1-\alpha}\right)^{\frac{k}{2}}}{2^{\frac{k}{2}}\left(\frac{k}{2}\right)!}\right],
\end{aligned}$$

which implies that

$$\begin{aligned}
\frac{\partial}{\partial\alpha} q_{k+2}^{1-\alpha} = \frac{e^{\frac{q_{k+2}^{1-\alpha}}{2}}}{\alpha e^{\frac{q_{k+2}^{1-\alpha}}{2}} - \frac{\left(q_{k+2}^{1-\alpha}\right)^{\frac{k}{2}}}{2^{\frac{k}{2}}\left(\frac{k}{2}\right)!} - e^{\frac{q_{k+2}^{1-\alpha}}{2}}} &= -\frac{1}{1 - \alpha + \frac{e^{-\frac{q_{k+2}^{1-\alpha}}{2}}\left(q_{k+2}^{1-\alpha}\right)^{\frac{k}{2}}}{2^{\frac{k}{2}}\left(\frac{k}{2}\right)!}} \\
&= -\frac{1}{1 - \alpha + 2\psi\left(q_{k+2}^{1-\alpha}; k+2, 0\right)}, \qquad (25)
\end{aligned}$$

where $\psi(x; k, 0)$ denotes the pdf of χ_k^2. By Lemma 5, which is stated and proved below, we have

$$0 < \psi\left(q_{k+2}^{1-\alpha}; k+2, 0\right) < \psi\left(q_k^{1-\alpha}; k, 0\right), \tag{26}$$

and so by (25)

$$0 > \frac{\partial}{\partial\alpha} q_k^{1-\alpha} > \frac{\partial}{\partial\alpha} q_{k+2}^{1-\alpha}. \tag{27}$$

Next consider

$$h(\alpha) := \log q_{k+2} - \log(k+2) - \frac{1}{2}(q_{k+2} - q_k), \tag{28}$$

and

$$\begin{aligned} h'(\alpha) &= \frac{\partial}{\partial\alpha}\left(\log q_{k+2}^{1-\alpha} - \log(k+2) - \frac{1}{2}\left(q_{k+2}^{1-\alpha} - q_k^{1-\alpha}\right)\right) \\ &= \frac{1}{q_{k+2}^{1-\alpha}} \cdot \frac{\partial}{\partial\alpha} q_{k+2}^{1-\alpha} + \frac{1}{2}\frac{\partial}{\partial\alpha} q_k^{1-\alpha} - \frac{1}{2}\frac{\partial}{\partial\alpha} q_{k+2}^{1-\alpha} \\ &= \frac{1}{2}\frac{\partial}{\partial\alpha} q_k^{1-\alpha} - \left(\frac{1}{2} - \frac{1}{q_{k+2}^{1-\alpha}}\right) \cdot \frac{\partial}{\partial\alpha} q_{k+2}^{1-\alpha} \\ &> 0. \end{aligned} \tag{29}$$

By Chen and Rubin (1986), noting that $\frac{1}{2}\chi_k^2 \sim Gamma\left(\frac{k}{2}, 1\right)$, we have the inequality

$$\frac{1}{2}k - \frac{1}{3} < \frac{1}{2} q_k^{\frac{1}{2}} < \frac{1}{2}k, \quad \text{or } k - \frac{2}{3} < q_k^{\frac{1}{2}} < k,$$

so that

$$\begin{aligned} h\left(\frac{1}{2}\right) &= \log q_{k+2}^{\frac{1}{2}} - \log(k+2) - \frac{1}{2}\left(q_{k+2}^{\frac{1}{2}} - q_k^{\frac{1}{2}}\right) \\ &< \log(k+2) - \log(k+2) - \frac{1}{2}((k+2-1) - k) = -\frac{1}{2} \\ &< 0. \end{aligned}$$

As $h(\alpha)$ is strictly increasing on $\left(0, \frac{1}{2}\right)$ by (29), we have

$$h(\alpha) < 0, \ \forall \alpha \in \left(0, \frac{1}{2}\right).$$

Hence, following equation (24), for α small enough, we have

$$\log\left(\frac{e^{-\frac{q_{k+2}}{2}} q_{k+2}^{\frac{k}{2}+1}}{(k+2)\, e^{-\frac{q_k}{2}} q_k^{\frac{k}{2}}}\right) < 0,$$

i.e. $e^{-\frac{q_{k+2}}{2}} q_{k+2}^{\frac{k}{2}+1} - (k+2)\, e^{-\frac{q_k}{2}} q_k^{\frac{k}{2}} < 0$, and thus by (22) it follows that

$$g_2'(0) < 0.$$

Hence, locally in a neighborhood of 0, i.e., $\forall\lambda \in \left(0, \overline{\lambda}\right)$ for some $\overline{\lambda} > 0$, we have

$$\Psi\left(q_{k+2}^{1-\alpha}; k+2, \lambda\right) < \Psi\left(q_{k}^{1-\alpha}; k, \lambda\right),$$

as required. ∎

Lemma 5. *For any $k \geq 2$ and small enough $\alpha \in (0,1)$,*

$$\psi\left(q_{k}^{1-\alpha}; k, 0\right) > \psi\left(q_{k+2}^{1-\alpha}; k+2, 0\right),$$

where $\psi(x; k, 0)$ denotes the pdf for χ_k^2.

Remark. We have numerically verified the stronger inequality $\psi\left(q_{k}^{1-\alpha}; k, 0\right) > \psi\left(q_{k+1}^{1-\alpha}; k+1, 0\right)$ for all combinations of $\alpha \in \{0.01, 0.05, 0.1\}$ and $k \leq 1000$.[4]

Proof. First recall that $\psi(x; k, 0) = \frac{e^{-\frac{x}{2}} x^{\frac{k}{2}-1}}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}$, so

$$\frac{\psi\left(q_k + x;\ k, 0\right)}{\psi\left(q_{k+2} + x;\ k+2, 0\right)} = \frac{\frac{e^{-\frac{q_k+x}{2}}(q_k+x)^{\frac{k}{2}-1}}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}}{\frac{e^{-\frac{q_{k+2}+x}{2}}(q_{k+2}+x)^{\frac{k}{2}}}{2^{\frac{k}{2}+1}\Gamma\left(\frac{k}{2}+1\right)}} = \frac{e^{\frac{1}{2}(q_{k+2}-q_k)}}{k} \cdot \left(\frac{(q_k + x)^{\frac{k}{2}-1}}{(q_{k+2} + x)^{\frac{k}{2}}}\right). \tag{30}$$

Taking logarithms of the fraction in parenthesis, we obtain

$$f(x) := \log\left(\frac{(q_k + x)^{\frac{k}{2}-1}}{(q_{k+2} + x)^{\frac{k}{2}}}\right) = \left(\frac{k}{2} - 1\right)\log\left(q_k + x\right) - \frac{k}{2}\log\left(q_{k+2} + x\right),$$

and

$$f'(x) = \frac{(k-2)\, q_{k+2} - k q_k - 2x}{2\left(q_k + x\right)\left(q_{k+2} + x\right)}. \tag{31}$$

By Proposition 5.1 of Inglot (2010), and by Laurent and Massart (2000), cited in Inglot (2010) as Theorem A, for $k \geq 2$ and $\alpha \leq 0.17$, we have

$$\begin{aligned} q_{k+2} &\leq k + 2 + 2\log\frac{1}{\alpha} + 2\sqrt{(k+2)\log\frac{1}{\alpha}}, \\ q_k &\geq k + 2\log\frac{1}{\alpha} - \frac{5}{2}, \end{aligned}$$

[4]Computations were performed in MatLab with a machine epsilon of approximately 2×10^{-16}. The minimum difference $\psi\left(q_{k}^{1-\alpha}; k, 0\right) - \psi\left(q_{k+1}^{1-\alpha}; k+1, 0\right)$ was found to be a magnitude larger than 10^{-7}, and the minimum ratio $\psi\left(q_{k}^{1-\alpha}; k, 0\right) / \psi\left(q_{k+1}^{1-\alpha}; k+1, 0\right)$ was found to be a magnitude larger than $1 + 10^{-4}$.

and thus

$$
\begin{aligned}
&(k-2)\,q_{k+2}-kq_k\\
\leq\,&(k-2)\left(k+2+2\log\frac{1}{\alpha}+2\sqrt{(k+2)\log\frac{1}{\alpha}}\right)-k\left(k+2\log\frac{1}{\alpha}-\frac{5}{2}\right)\\
=\,&2\,(k-2)\sqrt{k+2}\sqrt{\log\frac{1}{\alpha}}-4\log\frac{1}{\alpha}+\frac{5}{2}k-4\\
<\,&0 \text{ for small enough } \alpha. \qquad (32)
\end{aligned}
$$

Plugging the inequality (32) into (31), we have

$$
f'(x)\leq\frac{(k-2)\,q_{k+2}-kq_k}{2\,(q_k+x)\,(q_{k+2}+x)}<0,\quad \forall x\in(0,\infty)\,.
$$

Hence, $f(x)$ is decreasing in x, and so is $\frac{\psi(q_k+x;\;k,0)}{\psi(q_{k+2}+x;\;k+2,0)}$ by (30).

Now, suppose that $\psi\left(q_k^{1-\alpha};k,0\right)\leq\psi\left(q_{k+2}^{1-\alpha};k+2,0\right)$, i.e.,

$$
\frac{\psi\left(q_k;\;k,0\right)}{\psi\left(q_{k+2};\;k+2,0\right)}\leq 1.
$$

Then, by the above, for $\alpha\leq 0.17$, we have

$$
\frac{\psi\left(q_k+x;\;k,0\right)}{\psi\left(q_{k+2}+x;\;k+2,0\right)}<1,\quad \forall x\in(0,\infty)\,. \qquad (33)
$$

which implies

$$
\begin{aligned}
\alpha=\Psi\left(q_{k+2}^{1-\alpha};k+2,0\right)&=\int_0^\infty\psi\left(q_{k+2}^{1-\alpha}+x;k+2,0\right)dx\\
&<\int_0^\infty\psi\left(q_k^{1-\alpha}+x;k,0\right)dx=\Psi\left(q_k^{1-\alpha};k,0\right)=\alpha,
\end{aligned}
$$

giving a contradiction. ∎

Proof of Theorem 3

Proof. By Theorem 2, under the null $H_0:\beta=\beta_0$, $W_{UR}^{\beta_0}\xrightarrow{d}\chi_K^2$, so the critical value for a size-α test of H_0 is given by $q_K^{1-\alpha}$. In the following we prove the proposition for the three cases with strong, irrelevant and weak instruments separately.

(i). Strong instruments: For the URRF test, under $H_1:\beta=\tilde{\beta}_n:=\beta_0+\frac{1}{\sqrt{n}}\left(\beta_1-\beta_0\right)$,

$$
\begin{aligned}
\sqrt{n}\left(\hat{\pi}_1-\hat{\Pi}_2\beta_0\right)&=\sqrt{n}\left(\hat{\pi}_1-\pi_1\right)-\sqrt{n}\left(\hat{\Pi}_2-\Pi_2\right)\tilde{\beta}_n+\hat{\Pi}_2\left(\beta_1-\beta_0\right)\\
&\xrightarrow{a.s.}\xi-\Xi\beta_0+\Pi_2\left(\beta_1-\beta_0\right),
\end{aligned}
$$

so

$$
\begin{aligned}
W_{UR}^{\beta_0} &\xrightarrow{a.s.} \frac{\left[\xi - \Xi\beta_0 + \Pi_2\left(\beta_1 - \beta_0\right)\right]' \left[\xi - \Xi\beta_0 + \Pi_2\left(\beta_1 - \beta_0\right)\right]}{1 + \beta_0'\beta_0} \\
&\overset{d}{=} \chi^2_{K,\left(\beta_1 - \beta_0\right)'\Pi_2'\Pi_2\left(\beta_1 - \beta_0\right)/\left(1+\beta_0'\beta_0\right)},
\end{aligned}
$$

and thus

$$
\mathcal{K}_{UR}\left(\beta_1; \beta_0\right) = \Psi\left(q_K^{1-\alpha};\ K,\ \frac{\left(\beta_1 - \beta_0\right)'\Pi_2'\Pi_2\left(\beta_1 - \beta_0\right)}{1 + \beta_0'\beta_0}\right).
$$

Under the Gaussianity assumption,

$$
\begin{aligned}
\sqrt{n}\left(\hat{\pi}_1 - \hat{\Pi}_2\beta_0\right) &= \sqrt{n}\left(\hat{\pi}_1 - \pi_1\right) + \sqrt{n}\left(\Pi_2 - \hat{\Pi}_2\right)\tilde{\beta}_n + \hat{\Pi}_2\sqrt{n}\left(\tilde{\beta}_n - \beta_0\right) \\
&\overset{d}{=} \xi - \Xi\tilde{\beta}_n + \left(\Pi_2 + \frac{1}{\sqrt{n}}\Xi\right)\left(\beta_1 - \beta_0\right). \\
&= \xi - \Xi\beta_0 + \Pi_2\left(\beta_1 - \beta_0\right).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
W_{UR}^{\beta_0} &\overset{d}{=} \frac{\left[\xi - \Xi\beta_0 + \Pi_2\left(\beta_1 - \beta_0\right)\right]' \left[\xi - \Xi\beta_0 + \Pi_2\left(\beta_1 - \beta_0\right)\right]}{1 + \beta_0'\beta_0} \\
&\overset{d}{=} \chi^2_{K,\left(\beta_1 - \beta_0\right)'\Pi_2'\Pi_2\left(\beta_1 - \beta_0\right)/\left(1+\beta_0'\beta_0\right)},
\end{aligned}
$$

and thus

$$
\mathcal{K}_{UR}^{(n)}\left(\beta_1; \beta_0\right) = \Psi\left(q_K^{1-\alpha};\ K,\ \frac{\left(\beta_1 - \beta_0\right)'\Pi_2'\Pi_2\left(\beta_1 - \beta_0\right)}{1 + \beta_0'\beta_0}\right).
$$

For the PRRF test, as $W_{PR}^{\beta_0} \to \chi_m^2$ under H_0, so the critical value of the size-α is given by $q_m^{1-\alpha}$. Under $H_1 : \beta = \tilde{\beta}_n$,

$$
\begin{aligned}
\sqrt{n}\left(\tilde{\pi}_1 - \hat{\Pi}_2\beta_0\right) &= \hat{\Pi}_2\sqrt{n}\left(\beta_{IV} - \tilde{\beta}_n\right) + \hat{\Pi}_2\left(\beta_1 - \beta_0\right) \\
&= P_{\hat{\Pi}_2}\sqrt{n}\left(\hat{\pi}_1 - \pi_1\right) - P_{\hat{\Pi}_2}\sqrt{n}\left(\hat{\Pi}_2 - \Pi_2\right)\tilde{\beta}_n + \hat{\Pi}_2\left(\beta_1 - \beta_0\right) \\
&\xrightarrow{a.s.} P_{\Pi_2}\left(\xi - \Xi\beta_0\right) + \Pi_2\left(\beta_1 - \beta_0\right),
\end{aligned}
$$

so

$$
\begin{aligned}
W_{PR}^{\beta_0} &\xrightarrow{a.s.} \frac{\left[P_{\Pi_2}\left(\xi - \Xi\beta_0\right) + \Pi_2\left(\beta_1 - \beta_0\right)\right]' \left[P_{\Pi_2}\left(\xi - \Xi\beta_0\right) + \Pi_2\left(\beta_1 - \beta_0\right)\right]}{1 + \beta_0'\beta_0} \\
&\overset{d}{=} \chi^2_{m,\left(\beta_1 - \beta_0\right)'\Pi_2'\Pi_2\left(\beta_1 - \beta_0\right)/\left(1+\beta_0'\beta_0\right)}.
\end{aligned}
$$

Note that, under the conditions for Lemma 4 (iii), we have

$$\begin{aligned}\mathcal{K}_{PR}\left(\beta_1;\beta_0\right) &= \varPsi\left(q_m^{1-\alpha};\ m,\ \frac{\left(\beta_1-\beta_0\right)'\Pi_2'\Pi_2\left(\beta_1-\beta_0\right)}{1+\beta_0'\beta_0}\right)\\ &> \varPsi\left(q_K^{1-\alpha};\ K,\ \frac{\left(\beta_1-\beta_0\right)'\Pi_2'\Pi_2\left(\beta_1-\beta_0\right)}{1+\beta_0'\beta_0}\right) = \mathcal{K}_{UR}\left(\beta_1;\beta_0\right),\end{aligned}$$

Under the Gaussianity assumption,

$$\begin{aligned}\sqrt{n}\left(\tilde{\pi}_1-\hat{\Pi}_2\beta_0\right) &\overset{d}{=} P_{A_\Xi}\left(\xi-\Xi\tilde{\beta}_n\right)+A_\Xi\left(\beta_1-\beta_0\right)\\ &= P_{A_\Xi}\left(\xi-\Xi\beta_0+\Pi_2\left(\beta_1-\beta_0\right)\right)\end{aligned}$$

so

$$W_{PR}^{\beta_0} \overset{d}{=} \frac{1}{1+\beta_0'\beta_0}\chi^2_{m,\kappa_n}$$

with

$$\kappa_n = \left(\Xi\beta_0-\Pi_2\left(\beta_1-\beta_0\right)\right)' P_{A_\Xi}\left(\Xi\beta_0-\Pi_2\left(\beta_1-\beta_0\right)\right).$$

Hence,

$$\mathcal{K}_{PR}^{(n)}\left(\beta_1;\beta_0\right) = \varPsi\left(\left(1+\beta_0'\beta_0\right)q_m^{1-\alpha};\ m,\ \kappa_n\right)$$

(ii). Irrelevant instruments: For the URRF test, under $H_1:\beta=\beta_1$,

$$\sqrt{n}\left(\hat{\pi}_1-\hat{\Pi}_2\beta_0\right) \xrightarrow{a.s.} \xi-\Xi\beta_0 \overset{d}{=} \mathcal{N}\left(\mathbf{0},\left(1+\beta_0'\beta_0\right)\cdot I_K\right)$$

so $W_{UR}^{\beta_0} \xrightarrow{d} \chi^2_K$, and thus

$$\mathcal{K}_{UR}\left(\beta_1;\beta_0\right)=\alpha.$$

For the PRRF test, the asymptotic distributions of $W_{PR}^{\beta_0}$ under H_0 and under H_1 are the same

$$W_{PR}^{\beta_0} \xrightarrow{d} \frac{1}{1+\beta_0'\beta_0}\chi^2_{m,\lambda_I(\Xi,\beta_0)},$$

so

$$\mathcal{K}_{PR}\left(\beta_1;\beta_0\right)=\alpha.$$

(iii). Weak instruments: For the URRF test, under $H_1:\beta=\beta_1\neq\beta_0$,

$$\begin{aligned}\sqrt{n}\left(\hat{\pi}_1-\hat{\Pi}_2\beta_0\right) &= \sqrt{n}\left(\hat{\pi}_1-\hat{\Pi}_2\beta_1\right)+\sqrt{n}\hat{\Pi}_2\left(\beta_1-\beta_0\right)\\ &\xrightarrow{a.s.} \xi-\Xi\beta_1+\left(\Pi_2^*+\Xi\right)\left(\beta_1-\beta_0\right)\\ &= \xi-\Xi\beta_0+\Pi_2^*\left(\beta_1-\beta_0\right),\end{aligned}$$

so

$$W_{UR}^{\beta_0} \xrightarrow{a.s.} \frac{\left(\xi - \Xi\beta_0 + \Pi_2^*\left(\beta_1 - \beta_0\right)\right)^{'}\left(\xi - \Xi\beta_0 + \Pi_2^*\left(\beta_1 - \beta_0\right)\right)}{1 + \beta_0^{'}\beta_0}$$
$$\stackrel{d}{=} \chi^2_{K,\left(\beta_1-\beta_0\right)^{'}\Pi_2^{*'}\Pi_2^*\left(\beta_1-\beta_0\right)/\left(1+\beta_0^{'}\beta_0\right)},$$

and thus

$$\mathcal{K}_{UR}\left(\beta_1;\beta_0\right) = \Psi\left(q_K^{1-\alpha};\ K,\ \frac{\left(\beta_1 - \beta_0\right)^{'}\Pi_2^{*'}\Pi_2^*\left(\beta_1 - \beta_0\right)}{1 + \beta_0^{'}\beta_0}\right).$$

For the PRRF test, under H_0, $W_{PR}^{\beta_0} \xrightarrow{d} \frac{1}{1+\beta_0^{'}\beta_0}\mathcal{M}\chi^2_{m,\lambda_W(\Xi,\beta_0)}$. Let $c_{\beta_0}^{1-\alpha}$ be the critical value of the size-α test such that

$$\mathbb{P}\left\{\frac{1}{1 + \beta_0^{'}\beta_0}\mathcal{M}\chi^2_{m,\beta_0^{'}\beta_0\cdot\chi_K^2} \geq c_{\beta_0}^{1-\alpha}\right\} = \alpha.$$

Then, under H_1,

$$\begin{aligned}
\sqrt{n}\left(\tilde{\pi}_1 - \hat{\Pi}_2\beta_0\right) &= \sqrt{n}\hat{\Pi}_2\left(\beta_{IV} - \beta_1\right) + \sqrt{n}\hat{\Pi}_2\left(\beta_1 - \beta_0\right) \\
&= P_{\sqrt{n}\hat{\Pi}_2}\left(\sqrt{n}\hat{\pi}_1 - \sqrt{n}\hat{\Pi}_2\beta_1\right) + \sqrt{n}\hat{\Pi}_2\left(\beta_1 - \beta_0\right) \\
&\xrightarrow{a.s.} P_{\left(\Pi_2^*+\Xi\right)}\left(\xi - \Xi\beta_1\right) + \left(\Pi_2^* + \Xi\right)\left(\beta_1 - \beta_0\right) \\
&\equiv P_{\left(\Pi_2^*+\Xi\right)}\left(\xi - \Xi\beta_0 + \Pi_2^*\left(\beta_1 - \beta_0\right)\right)
\end{aligned}$$

so

$$W_{PR}^{\beta_0} \xrightarrow{d} \frac{1}{1 + \beta_0^{'}\beta_0}\mathcal{M}\chi^2_{m,\left(\Xi\beta_0-\Pi_2^*(\beta_1-\beta_0)\right)^{'}P_{\left(\Pi_2^*+\Xi\right)}\left(\Xi\beta_0-\Pi_2^*(\beta_1-\beta_0)\right)},$$

and thus

$$\begin{aligned}
\mathcal{K}_{PR}\left(\beta_1;\beta_0\right) &= \mathbb{P}\left\{\frac{1}{1 + \beta_0^{'}\beta_0}\mathcal{M}\chi^2_{m,\left(\Xi\beta_0-\Pi_2^*(\beta_1-\beta_0)\right)^{'}P_{\Pi_2^*+\Xi}\left(\Xi\beta_0-\Pi_2^*(\beta_1-\beta_0)\right)} \geq c_{\beta_0}^{1-\alpha}\right\} \\
&= \mathbb{E}\left[\Psi\left(\left(1 + \beta_0^{'}\beta_0\right)c_{\beta_0}^{1-\alpha};\ m,\ \left(\Xi\beta_0 - \Pi_2^*\left(\beta_1 - \beta_0\right)\right)^{'}P_{\left(\Pi_2^*+\Xi\right)}\left(\Xi\beta_0 - \Pi_2^*\left(\beta_1 - \beta_0\right)\right)\right)\right].
\end{aligned}$$

For the special case of $\beta_0 = \mathbf{0}$, $c_{\beta_0}^{1-\alpha} = q_m^{1-\alpha}$, and $W_{PR}^{\mathbf{0}} \xrightarrow{d} \mathcal{M}\chi^2_{m,\beta_1^{'}\Pi_2^{*'}P_{\Pi_2^*+\Xi}\Pi_2^*\beta_1}$. Hence we have

$$\mathcal{K}_{PR}\left(\beta_1;\beta_0\right) = \mathbb{E}\left[\Psi\left(q_m^{1-\alpha};\ m,\ \beta_1^{'}\Pi_2^{*'}P_{\left(\Pi_2^*+\Xi\right)}\Pi_2^*\beta_1\right)\right].$$

∎

Proof of Theorem 4

Proof. We prove the result for different strengths of instruments.

(i). Strong instruments, against $H_1 : \beta = \tilde{\beta}_n := \beta_0 + \frac{1}{\sqrt{nK}} (\beta_1 - \beta_0)$: For the URRF test, under $H_0 : \beta = \beta_0$, $\tilde{W}_{UR}^{\beta_0} \xrightarrow{d} \mathcal{N}(0, 2)$, so the critical value for the size-α test under the null is $\sqrt{2}Q^{-1}(\alpha)$, where $Q(x) = 1 - \Phi(x)$ denotes the survival function for the standard normal distribution with cdf Φ and density φ. Under $H_1 : \beta = \tilde{\beta}_n$ and (A3')

$$\tilde{W}_{UR}^{\beta_0} = \sqrt{K}\left[\frac{\frac{1}{K}n\left(\hat{\pi}_1 - \hat{\Pi}_2\tilde{\beta}_n + \frac{1}{\sqrt{nK}}\hat{\Pi}_2(\beta_1 - \beta_0)\right)'\left(\hat{\pi}_1 - \hat{\Pi}_2\tilde{\beta}_n + \frac{1}{\sqrt{nK}}\hat{\Pi}_2(\beta_1 - \beta_0)\right)}{1 + \beta_0'\beta_0} - 1\right]$$

$$\xrightarrow{a.s.} \lim_{K\to\infty}\sqrt{K}\left[\frac{1}{K}\frac{\left(\xi - \Xi\beta_0 + \frac{1}{\sqrt{K}}\Pi_2(\beta_1 - \beta_0)\right)'\left(\xi - \Xi\beta_0 + \frac{1}{\sqrt{K}}\Pi_2(\beta_1 - \beta_0)\right)}{1 + \beta_0'\beta_0} - 1\right]$$

The term

$$\frac{1}{K}\frac{\left(\xi - \Xi\beta_0 + \frac{1}{\sqrt{K}}\Pi_2(\beta_1 - \beta_0)\right)'\left(\xi - \Xi\beta_0 + \frac{1}{\sqrt{K}}\Pi_2(\beta_1 - \beta_0)\right)}{1 + \beta_0'\beta_0} \tag{34}$$

is $\frac{1}{K}$ times a noncentral chi-squared distribution with K degrees of freedom and noncentrality parameter

$$\frac{(\beta_1 - \beta_0)'\frac{1}{K}\Pi_2'\Pi_2(\beta_1 - \beta_0)}{1 + \beta_0'\beta_0} \xrightarrow{a.s.} \frac{(\beta_1 - \beta_0)'\Omega_{\Pi_2}(\beta_1 - \beta_0)}{1 + \beta_0'\beta_0}.$$

The term (34) can also be interpreted as the sample average of K i.i.d. random variables drawn from the noncentral chi-squared distribution with one degree of freedom and noncentrality parameter

$$\frac{1}{K}\cdot\frac{(\beta_1 - \beta_0)'\frac{1}{K}\Pi_2'\Pi_2(\beta_1 - \beta_0)}{1 + \beta_0'\beta_0}.$$

The mean and variance of a $\chi^2_{1,\lambda}$ variate are $(1 + \lambda)$ and $2(1 + 2\lambda)$, so that under H_1 we have

$$\begin{aligned}\tilde{W}_{UR}^{\beta_0} &= \sqrt{K}\left(\frac{1}{K}W_U^{\beta_0} - 1 - \frac{1}{K}\frac{(\beta_1 - \beta_0)'\frac{1}{K}\Pi_2'\Pi_2(\beta_1 - \beta_0)}{1 + \beta_0'\beta_0}\right)\\ &\quad + \frac{1}{\sqrt{K}}\cdot\frac{(\beta_1 - \beta_0)'\frac{1}{K}\Pi_2'\Pi_2(\beta_1 - \beta_0)}{1 + \beta_0'\beta_0}\\ &\xrightarrow{d} \mathcal{N}(0, 2).\end{aligned}$$

It follows that the URRF test has trivial asymptotic power $\mathcal{K}_{UR}(\beta_1;\beta_0) = Q\left(Q^{-1}(\alpha)\right) = \alpha$ against local alternatives $H_1 : \beta = \beta_0 + \frac{1}{\sqrt{nK}}(\beta_1 - \beta_0)$. For the PRRF test, under $H_0 : \beta = \beta_0$ we have,

$$W_{PR}^{\beta_0} \xrightarrow{a.s.} \chi_m^2,$$

so the critical value is given by $q_m^{1-\alpha}$. Under (A3') and $H_1 : \beta = \beta_0 + \frac{1}{\sqrt{nK}}(\beta_1 - \beta_0)$, we have

$$\begin{aligned} W_{PR}^{\beta_0} &\xrightarrow{a.s.} \lim_{K\to\infty} \frac{\left[C'(\xi - \Xi\beta_0) + \frac{1}{\sqrt{K}}C'\Pi_2(\beta_1-\beta_0)\right]'\left[C'(\xi - \Xi\beta_0) + \frac{1}{\sqrt{K}}C'\Pi_2(\beta_1-\beta_0)\right]}{1+\beta_0'\beta_0} \\ &\overset{d}{=} \chi^2_{m,\lim_{K\to\infty}\frac{1}{K}(\beta_1-\beta_0)'\Pi_2'\Pi_2(\beta_1-\beta_0)/\left(1+\beta_0'\beta_0\right)} \equiv \chi^2_{m,(\beta_1-\beta_0)'\Omega_{\Pi_2}(\beta_1-\beta_0)/\left(1+\beta_0'\beta_0\right)}, \end{aligned}$$

so the PRRF test has nontrivial asymptotic power given by

$$\mathcal{K}_{PR}(\beta_1;\beta_0) = \Psi\left(q_m^{1-\alpha}; m, \frac{(\beta_1-\beta_0)'\Omega_{\Pi_2}(\beta_1-\beta_0)}{1+\beta_0'\beta_0}\right).$$

(ii). Strong instruments, against $H_1 : \beta = \beta_0 + \frac{1}{\sqrt{nK^{1/4}}}(\beta_1-\beta_0)$. A replication of the above analysis reveals that under H_1

$$\begin{aligned} \tilde{W}_{UR}^{\beta_0} &= \sqrt{K}\left(\frac{1}{K}W_U^{\beta_0} - 1 - \frac{1}{K}\frac{(\beta_1-\beta_0)'\frac{1}{\sqrt{K}}\Pi_2'\Pi_2(\beta_1-\beta_0)}{1+\beta_0'\beta_0}\right) + \frac{(\beta_1-\beta_0)'\frac{1}{K}\Pi_2'\Pi_2(\beta_1-\beta_0)}{1+\beta_0'\beta_0} \\ &\xrightarrow{d} \mathcal{N}\left(\frac{(\beta_1-\beta_0)'\Omega_{\Pi_2}(\beta_1-\beta_0)}{1+\beta_0'\beta_0}, 2\right) \end{aligned}$$

so that the URRF test has nontrivial asymptotic power

$$\mathcal{K}_{UR}(\beta_1;\beta_0) = Q\left(Q^{-1}(\alpha) - \frac{1}{\sqrt{2}}\frac{(\beta_1-\beta_0)'\Omega_{\Pi_2}(\beta_1-\beta_0)}{1+\beta_0'\beta_0}\right) > \alpha,$$

while the PRRF test has unit asymptotic power

$$\mathcal{K}_{PR}(\beta_1;\beta_0) = \lim_{K\to\infty}\Psi_m\left(q_m^{1-\alpha}; m, \sqrt{K}\frac{(\beta_1-\beta_0)'\Omega_{\Pi_2}(\beta_1-\beta_0)}{1+\beta_0'\beta_0}\right) = 1.$$

(iii). Weak instruments, against $H_1 : \beta = \beta_0 + \frac{1}{\sqrt{K}}(\beta_1-\beta_0)$: For the URRF test, under $H_0 : \beta = \beta_0$, $\tilde{W}_{UR}^{\beta_0} \xrightarrow{d} \mathcal{N}(0,2)$, so the critical value remains the same. Under (A3') and H_1, we have $\sqrt{n}\left(\hat{\pi}_1 - \hat{\Pi}_2\beta_0\right) \xrightarrow{a.s.} \xi - \Xi\beta_0 + \Pi_2^* \frac{1}{\sqrt{K}}\left(\beta_1 - \beta_0\right)$ and therefore

$$\begin{aligned}
\tilde{W}_{UR}^{\beta_0} &\xrightarrow{a.s.} \lim_{K\to\infty} \sqrt{K}\left[\frac{1}{K}\frac{\left(\xi - \Xi\beta_0 + \Pi_2^* \frac{1}{\sqrt{K}}\left(\beta_1 - \beta_0\right)\right)^{'}\left(\xi - \Xi\beta_0 + \Pi_2^* \frac{1}{\sqrt{K}}\left(\beta_1 - \beta_0\right)\right)}{1 + \beta_0^{'}\beta_0} - 1\right] \\
&\overset{d}{=} \lim_{K\to\infty} \sqrt{K}\left[\frac{1}{K}\frac{\left(\xi - \Xi\beta_0 + \frac{1}{\sqrt{K}}\Pi_2^*\left(\beta_1 - \beta_0\right)\right)^{'}\left(\xi - \Xi\beta_0 + \frac{1}{\sqrt{K}}\Pi_2^*\left(\beta_1 - \beta_0\right)\right)}{1 + \beta_0^{'}\beta_0} - 1 - \frac{\lambda_K}{K}\right] \\
&\quad + \frac{\lambda_K}{\sqrt{K}} \\
&\overset{d}{=} \mathcal{N}\left(0, 2\right).
\end{aligned}$$

with

$$\lambda_K := \frac{\left(\beta_1 - \beta_0\right)^{'}\frac{1}{K}\Pi_2^{*'}\Pi_2^*\left(\beta_1 - \beta_0\right)}{1 + \beta_0^{'}\beta_0} \xrightarrow{a.s.} \frac{\left(\beta_1 - \beta_0\right)^{'}\Omega_{\Pi_2^*}\left(\beta_1 - \beta_0\right)}{1 + \beta_0^{'}\beta_0},$$

and $\frac{1}{\sqrt{K}}\lambda_K \xrightarrow{a.s.} 0$. So the URRF test again has trivial asymptotic power $\mathcal{K}_{UR}\left(\beta_1; \beta_0\right) = \alpha$. For the PRRF test, under H_0,

$$\begin{aligned}
W_{PR}^{\beta_0} &\xrightarrow{a.s.} \lim_{K\to\infty} \frac{\left(\xi - \Xi\beta_0\right)^{'} P_{\left(\Pi_2^* + \Xi\right)}\left(\xi - \Xi\beta_0\right)}{1 + \beta_0^{'}\beta_0} \\
&\overset{d}{=} \lim_{K\to\infty} \frac{1}{\left(1 + \beta_0^{'}\beta_0\right)}\mathcal{M}\chi^2_{m, \beta_0^{'}\Xi P_{\left(\Pi_2^* + \Xi\right)}\Xi\beta_0}
\end{aligned}$$

For $\beta_0 = \mathbf{0}$, then $\tilde{W}_{PR}^{\mathbf{0}} = W_{PR}^{\mathbf{0}} \xrightarrow{d} \chi^2_m$, so the critical value is given by

$$c_{\mathbf{0}}^{1-\alpha} = q_m^{1-\alpha}.$$

For $\beta_0 \neq \mathbf{0}$,

$$\beta_0^{'}\Xi P_{\left(\Pi_2^* + \Xi\right)}\Xi\beta_0 = K \cdot \frac{1}{K}\Xi P_{\left(\Pi_2^* + \Xi\right)}\Xi\beta_0 \xrightarrow{a.s.} \infty.$$

By Muirhead (1982, p. 46, Problem 1.18) ,

$$\frac{\chi^2_{m,\lambda} - m - \lambda}{\sqrt{2\left(m + 2\lambda\right)}} \xrightarrow{d} \mathcal{N}\left(0, 1\right) \quad \text{as } \lambda \to \infty.$$

Hence, conditional on the sigma algebra $\sigma\left(\left(\Xi_{(k)}\right)_{k=1}^{\infty}\right)$, and expanding the probability space as needed to replace weak convergence by $a.s.$ convergence we have

$$\frac{\left(\xi - \Xi\beta_0\right)^{'} P_{\left(\Pi_2^* + \Xi\right)}\left(\xi - \Xi\beta_0\right) - m - \beta_0^{'}\Xi^{'} P_{\left(\Pi_2^* + \Xi\right)}\Xi\beta_0}{\sqrt{m + 2\beta_0^{'}\Xi P_{\left(\Pi_2^* + \Xi\right)}\Xi\beta_0}} \xrightarrow{a.s.} \mathcal{N}\left(0, 2\right).$$

As the limit distribution does not depend on Ξ, the unconditional limit distribution will also be given by $\mathcal{N}(0,2)$. As

$$\frac{1}{K}\Xi'\Xi \xrightarrow{a.s.} I_m, \quad \frac{1}{K}\Pi_2^{*\prime}\Pi_2^* \xrightarrow{a.s.} \Omega_{\Pi_2^*}, \quad \frac{1}{K}\Pi_2^{*\prime}\Xi \xrightarrow{a.s.} \mathbf{0},$$

we have

$$\begin{aligned}\frac{1}{K}\beta_0'\Xi P_{(\Pi_2^*+\Xi)}\Xi\beta_0 &= \frac{1}{K}\beta_0'\Xi'\left(\Pi_2^*+\Xi\right)\left[\frac{1}{K}\left(\Pi_2^*+\Xi\right)'\left(\Pi_2^*+\Xi\right)\right]^{-1}\frac{1}{K}\left(\Pi_2^*+\Xi\right)'\Xi\beta_0 \\ &\xrightarrow{a.s.} \beta_0'\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\beta_0.\end{aligned}$$

Let $(\Xi'\Xi)_{ij}$ denote the ij-th element of the $m\times m$ matrix $\Xi'\Xi$. Since $(\Xi'\Xi)_{ii}\sim\chi_K^2$,

$$\sqrt{K}\left(\frac{1}{K}(\Xi'\Xi)_{ii}-1\right)\xrightarrow{d}\mathcal{N}(0,2).$$

Also, for $i\neq j$, $(\Xi'\Xi)_{ij}=\sum_{k=1}^K\Xi_{ki}\Xi_{kj}$ with $\mathbb{E}\left[\Xi_{ki}\Xi_{kj}\right]=0$ and $\mathbb{E}\left[\Xi_{ki}^2\Xi_{kj}^2\right]=1$, so

$$\frac{1}{\sqrt{K}}\left(\Xi'\Xi\right)_{ij}\xrightarrow{d}\mathcal{N}(0,1).$$

It is easily shown that $Cov\left[\left(\Xi'\Xi\right)_{ij},\left(\Xi'\Xi\right)_{hl}\right]=0 \quad \forall(i,j)\neq(h,l)$. Let $\Omega_{\Xi'\Xi}$ denote the asymptotic variance-covariance matrix as described above. Then, working in the expanded probability space, the limiting matrix variate

$$\mathit{\Delta}_{\Xi'\Xi}:=\lim_{K\to\infty}\sqrt{K}\left(\frac{1}{K}\Xi'\Xi-I_m\right)\overset{d}{=}\mathcal{N}_{m,m}\left(\mathbf{0},\Omega_{\Xi'\Xi}\right).$$

is well-defined almost surely with nondegenerate limit distribution. Also, notice that

$$\frac{1}{\sqrt{K}}\Pi_2^{*\prime}\Xi\sim\mathcal{MN}_{m,m}\left(\mathbf{0},I_m\otimes\frac{1}{K}\Pi_2^{*\prime}\Pi_2^*\right)\xrightarrow{d}\mathcal{N}\left(\mathbf{0},I_m\otimes\Omega_{\Pi_2^*}\right),$$

and so we may similarly define $\Delta_{\Pi_2^{*\prime}\Xi}:=\lim_{K\to\infty}\frac{1}{\sqrt{K}}\Pi_2^{*\prime}\Xi$. By direct algebraic decomposition, we can write

$$\begin{aligned}&\frac{1}{\sqrt{K}}\Xi'P_{(\Pi_2^*+\Xi)}\Xi-\sqrt{K}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1} \\ =&\frac{1}{K}\Xi'\left(\Pi_2^*+\Xi\right)\left[\frac{1}{K}\left(\Pi_2^*+\Xi\right)'\left(\Pi_2^*+\Xi\right)\right]^{-1}\left(\frac{1}{\sqrt{K}}\Pi_2^{*\prime}\Xi+\sqrt{K}\left(\frac{1}{K}\Xi'\Xi-I_m\right)\right) \\ &+\left(\frac{1}{\sqrt{K}}\Xi'\Pi_2^*+\sqrt{K}\left(\frac{1}{K}\Xi'\Xi-I_m\right)\right)\left[\frac{1}{K}\left(\Pi_2^*+\Xi\right)'\left(\Pi_2^*+\Xi\right)\right]^{-1} \\ &+\sqrt{K}\left[\frac{1}{K}\left(\Pi_2^*+\Xi\right)'\left(\Pi_2^*+\Xi\right)\right]^{-1}-\sqrt{K}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1} \qquad (35)\end{aligned}$$

Define the limiting matrix variate $\Delta_{\Pi_2^{*\prime}\Pi_2^*} := \lim_{K\to\infty} \sqrt{K}\left(\frac{1}{K}\Pi_2^{*\prime}\Pi_2^* - \Omega_{\Pi_2^*}\right)$ and let

$$M_K = \frac{1}{K}\left(\Pi_2^* + \Xi\right)'\left(\Pi_2^* + \Xi\right),\ M := \Omega_{\Pi_2^*} + I_m.$$

By virtue of the above limit theory we have

$$\begin{aligned}\sqrt{K}\left(M_K - M\right) &= \sqrt{K}\left(\frac{1}{K}\Pi_2^{*\prime}\Pi_2^* - \Omega_{\Pi_2^*}\right) + \sqrt{K}\left(\frac{1}{K}\Xi'\Xi - I_m\right)\\ &\quad + \frac{1}{\sqrt{K}}\Pi_2^{*\prime}\Xi + \frac{1}{\sqrt{K}}\Xi'\Pi_2^*\\ &\xrightarrow{a.s.} \Delta_{\Pi_2^{*\prime}\Xi} + \mathit{\Delta}_{\Xi'\Xi} + \Delta'_{\Pi_2^{*\prime}\Xi} + \mathit{\Delta}_{\Pi_2^{*\prime}\Pi_2^*} =: \Delta.\end{aligned}$$

Since $M^{-1} > 0$ and $M_K > 0$ *a.s.*, we deduce by standard matrix delta methods that $\sqrt{K}\left(M_K^{-1} - M^{-1}\right) \xrightarrow{a.s.} -M^{-1}\Delta M^{-1}$, i.e.,

$$\sqrt{K}\left[\frac{1}{K}\left(\Pi_2^* + \Xi\right)'\left(\Pi_2^* + \Xi\right)\right]^{-1} - \sqrt{K}\left(\Omega_{\Pi_2^*} + I_m\right)^{-1} \xrightarrow{a.s.} -\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\Delta\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}$$

Then, in view of (35), we have

$$\begin{aligned}&\frac{1}{\sqrt{K}}\Xi' P_{\left(\Pi_2^*+\Xi\right)}\Xi - \sqrt{K}\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\\ &\xrightarrow{a.s.} \left(I_m + \Omega_{\Pi_2^*}\right)^{-1}\left(\Delta_{\Pi_2^{*\prime}\Xi} + \Delta_{\Xi'\Xi}\right) + \left(\Delta_{\Pi_2^{*\prime}\Xi} + \Delta_{\Xi'\Xi}\right)'\left(I_m + \Omega_{\Pi_2^*}\right)^{-1}\\ &\quad - \left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\Delta\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}.\end{aligned} \tag{36}$$

Hence,

$$\begin{aligned}\tilde{W}_{PR}^{\beta_0} &:= \frac{1}{\sqrt{K}}\left(W_{PR}^{\beta_0} - \frac{m + K\beta_0'\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\beta_0}{1 + \beta_0'\beta_0}\right)\\ &= \frac{\sqrt{\frac{m}{K} + 2\cdot\frac{1}{K}\beta_0'\Xi P_{\left(\Pi_2^*+\Xi\right)}\Xi\beta_0}}{1 + \beta_0'\beta_0}\cdot\frac{\left(\xi - \Xi\beta_0\right)' P_{\Pi_2^*+\Xi}\left(\xi - \Xi\beta_0\right) - m - \beta_0'\Xi P_{\left(\Pi_2^*+\Xi\right)}\Xi\beta_0}{\sqrt{m + 2\beta_0'\Xi P_{\left(\Pi_2^*+\Xi\right)}\Xi\beta_0}}\\ &\quad + \frac{\beta_0'\left[\frac{1}{\sqrt{K}}\Xi' P_{\left(\Pi_2^*+\Xi\right)}\Xi - \sqrt{K}\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\right]\beta_0}{1 + \beta_0'\beta_0}\\ &\xrightarrow{a.s.} \frac{1}{1 + \beta_0'\beta_0}\left[N\left(0, 4\beta_0'\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\beta_0\right) + 2\beta_0'\left(I_m + \Omega_{\Pi_2^*}^i\right)^{-1}\left(\Delta_{\Pi_2^{*\prime}\Xi} + \Delta_{\Xi'\Xi}\right)\beta_0\right.\\ &\quad \left. -\beta_0'\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\Delta\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\beta_0\right]\\ &\stackrel{d}{=} \mathcal{N}\left(0, \mathbb{AV}\mathrm{ar}\left[\tilde{W}_{PR}^{\beta_0}\right]\right).\end{aligned}$$

The asymptotic normality of $\tilde{W}_{PR}^{\beta_0}$ follows from the fact that the component variates $\Delta_{\Pi_2^{*\prime}\Xi}, \Delta_{\Xi'\Xi}, \Delta$ are jointly all normally distributed with zero mean. However, the asymptotic variance is complicated (though analytically derivable) due to the correlation between the random variables. The critical value is then given by

$$c_{\beta_0}^{1-\alpha} = \begin{cases} q_m^{1-\alpha}, & \beta_0 = \mathbf{0} \\ \sqrt{\mathbb{AV}\mathrm{ar}\left[\tilde{W}_{PR}^{\beta_0}\right]} \cdot Q^{-1}(\alpha), & \beta_0 \neq \mathbf{0} \end{cases}.$$

Under $H_1 : \beta = \tilde{\beta}_n = \beta_0 + \frac{1}{\sqrt{K}}(\beta_1 - \beta_0)$, we have as $n \to \infty$

$$\sqrt{n}\left(\tilde{\pi}_1 - \hat{\Pi}_2\beta_0\right) \xrightarrow{a.s.} P_{\left(\Pi_2^*+\Xi\right)}\left(\xi - \Xi\beta_0 + \Pi_2^* \frac{1}{\sqrt{K}}(\beta_1 - \beta_0)\right)$$

For $\beta_0 = \mathbf{0}$, the noncentrality parameter λ_K for the asymptotic distribution of $\tilde{W}_{PR}^{\mathbf{0}} = W_{PR}^{\mathbf{0}}$ is

$$\lambda_K = \frac{1}{K}\beta_1' \Pi_2^{*\prime} P_{\left(\Pi_2^*+\Xi\right)} \Pi_2^* \beta_1 \xrightarrow{a.s.} \beta_1' \Omega_{\Pi_2^*}\left(\Omega_{\Pi_2^*} + I_m\right)^{-1} \Omega_{\Pi_2^*}\beta_1,$$

and so

$$\tilde{W}_{PR}^{\mathbf{0}} \xrightarrow{a.s.} \chi^2_{m, \beta_1' \Omega_{\Pi_2^*}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\Omega_{\Pi_2^*}\beta_1}.$$

Hence, the PRRF test has nontrivial power:

$$\mathcal{K}_{PR}(\beta_1; \mathbf{0}) = \Psi\left(q_m^{1-\alpha}; m, \beta_1' \Omega_{\Pi_2^*}\left(\Omega_{\Pi_2^*} + I_m\right)^{-1} \Omega_{\Pi_2^*}\beta_1\right) \geq \alpha.$$

For $\beta_0 \neq \mathbf{0}$, note that now the noncentrality parameter λ_K for the asymptotic distribution of $W_{PR}^{\beta_0}$ is such that

$$\begin{aligned}
&\frac{1}{\sqrt{K}}\lambda_K - \sqrt{K}\beta_0\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\beta_0 \\
&= \left[\frac{1}{\sqrt{K}}\beta_0' \Xi' P_{\left(\Pi_2^*+\Xi\right)} \Xi\beta_0 - \sqrt{K}\beta_0\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\beta_0\right] \\
&\quad - 2\left[\frac{1}{\sqrt{K}}\beta_0' \Xi' P_{\left(\Pi_2^*+\Xi\right)} \Pi_2^* \frac{1}{\sqrt{K}}(\beta_1 - \beta_0)\right] \\
&\quad + \frac{1}{\sqrt{K}}\left[\frac{1}{\sqrt{K}}(\beta_1 - \beta_0)' \Pi_2^* P_{\left(\Pi_2^*+\Xi\right)} \Pi_2^* \frac{1}{\sqrt{K}}(\beta_1 - \beta_0)\right]
\end{aligned} \tag{37}$$

Notice that

$$\begin{aligned}
\frac{1}{K}\Pi_2^{*\prime} P_{\left(\Pi_2^*+\Xi\right)} \Pi_2^* &= \frac{1}{K}\Pi_2^{*\prime}\left(\Pi_2^* + \Xi\right)\left[\frac{1}{K}\left(\Pi_2^* + \Xi\right)'\left(\Pi_2^* + \Xi\right)\right]^{-1}\frac{1}{K}\left(\Pi_2^* + \Xi\right)'\Pi_2^* \\
&\xrightarrow{a.s.} \Omega_{\Pi_2^*}'\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\Omega_{\Pi_2^*} &(38) \\
\frac{1}{K}\Xi' P_{\left(\Pi_2^*+\Xi\right)} \Pi_2^* &\xrightarrow{a.s.} \left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\Omega_{\Pi_2^*} &(39)
\end{aligned}$$

Hence, applying (36), (38) and (39) to the corresponding terms in the three square brackets in (37), we have

$$
\begin{aligned}
&\frac{1}{\sqrt{K}}\lambda_K - \sqrt{K}\beta_0\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\beta_0 \\
\xrightarrow{a.s.} & 2\beta_0'\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\left(\Delta_{\Pi_2^*{}'\Xi} + \Delta_{\Xi'\Xi}\right)\beta_0 - \beta_0'\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\Delta\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\beta_0 \\
&- 2\beta_0'\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1 - \beta_0\right), \qquad (40)
\end{aligned}
$$

so that

$$
\begin{aligned}
\tilde{W}_{PR}^{\beta_0} &:= \frac{1}{\sqrt{K}}\left(W_{PR}^{\beta_0} - \frac{m + K\beta_0'\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\beta_0}{1 + \beta_0'\beta_0}\right) \\
&= \frac{\sqrt{\frac{m}{K} + 2\frac{\lambda_K}{K}}}{1 + \beta_0'\beta_0} \cdot \frac{\left(\xi - \Xi\beta_0 + \Pi_2^*\frac{1}{\sqrt{K}}\left(\beta_1 - \beta_0\right)\right)' P_{\left(\Pi_2^* + \Xi\right)}\left(\xi - \Xi\beta_0 + \Pi_2^*\frac{1}{\sqrt{K}}\left(\beta_1 - \beta_0\right)\right) - m - \lambda_K}{\sqrt{m + 2\lambda_K}} \\
&\quad + \frac{\frac{1}{\sqrt{K}}\lambda_K - \sqrt{K}\beta_0\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\beta_0}{1 + \beta_0'\beta_0} \\
&\xrightarrow{a.s.} \mathcal{N}\left(-\frac{2\beta_0'\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1 - \beta_0\right)}{1 + \beta_0'\beta_0}, \mathbb{AV}\mathrm{ar}\left[\tilde{W}_{PR}^{\beta_0}\right]\right)
\end{aligned}
$$

It follows that the power of the PRRF test for $\beta_0 \neq \mathbf{0}$ is given by

$$
\mathcal{K}_{PR}\left(\beta_1;\beta_0\right) = Q\left(Q^{-1}\left(\alpha\right) + \frac{2\beta_0'\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1 - \beta_0\right)}{\sqrt{\mathbb{AV}\mathrm{ar}\left[\tilde{W}_{PR}^{\beta_0}\right] \cdot \left(1 + \beta_0'\beta_0\right)}}\right),
$$

and $\mathcal{K}_{PR}\left(\beta_1;\beta_0\right) \geq \alpha$ if and only if

$$
\beta_0'\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1 - \beta_0\right) \leq 0.
$$

(iv). Weak instruments, against $H_1 : \beta = \beta_0 + \frac{1}{K^{1/4}}\left(\beta_1 - \beta_0\right)$: Note that the null distributions of the URRF and the PRRF test statistics both remain the same as in (iii). Under H_1, the URRF test statistic

$$
\begin{aligned}
\tilde{W}_{UR}^{\beta_0} &\xrightarrow{a.s.} \lim_{K\to\infty}\sqrt{K}\left[\frac{1}{K}\frac{\left(\xi - \Xi\beta_0 + \frac{1}{K^{1/4}}\Pi_2^*\left(\beta_1 - \beta_0\right)\right)'\left(\xi - \Xi\beta_0 + \frac{1}{K^{1/4}}\Pi_2^*\left(\beta_1 - \beta_0\right)\right)}{1 + \beta_0'\beta_0} - 1\right] \\
&\overset{d}{=} \lim_{K\to\infty}\sqrt{K}\left[\frac{1}{K}\frac{\left(\xi - \Xi\beta_0 + \frac{1}{K^{1/4}}\Pi_2^*\left(\beta_1 - \beta_0\right)\right)'\left(\xi - \Xi\beta_0 + \frac{1}{K^{1/4}}\Pi_2^*\left(\beta_1 - \beta_0\right)\right)}{1 + \beta_0'\beta_0} - 1 - \frac{\lambda_K}{K}\right] \\
&\quad + \frac{\lambda_K}{\sqrt{K}} \\
&\overset{d}{=} \mathcal{N}\left(\frac{\left(\beta_1 - \beta_0\right)'\Omega_{\Pi_2^*}\left(\beta_1 - \beta_0\right)}{1 + \beta_0'\beta_0}, 2\right)
\end{aligned}
$$

where

$$\lambda_K := \frac{(\beta_1 - \beta_0)' \frac{1}{\sqrt{K}} \Pi_2^{*\prime} \Pi_2^* (\beta_1 - \beta_0)}{1 + \beta_0' \beta_0},$$

$$\frac{1}{\sqrt{K}} \lambda_K = \frac{(\beta_1 - \beta_0)' \frac{1}{K} \Pi_2^{*\prime} \Pi_2^* (\beta_1 - \beta_0)}{1 + \beta_0' \beta_0} \xrightarrow{a.s.} \frac{(\beta_1 - \beta_0)' \Omega_{\Pi_2^*} (\beta_1 - \beta_0)}{1 + \beta_0' \beta_0}.$$

Hence, the asymptotic power of the URRF test is

$$\mathcal{K}_{UR}(\beta_1; \beta_0) = \Phi\left(\Phi^{-1}(\alpha) - \frac{(\beta_1 - \beta_0)' \Omega_{\Pi_2^*} (\beta_1 - \beta_0)}{\sqrt{2}\left(1 + \beta_0' \beta_0\right)}\right).$$

For the PRRF test,

$$\sqrt{n}\left(\tilde{\pi}_1 - \hat{\Pi}_2 \beta_0\right) \rightsquigarrow P_{(\Pi_2^* + \Xi)}\left(\xi - \Xi\beta_0 + \frac{1}{K^{1/4}} \Pi_2^* (\beta_1 - \beta_0)\right).$$

For $\beta_0 = \mathbf{0}$, the noncentrality parameter λ_K diverges to infinity:

$$\lambda_K = \sqrt{K} \cdot \frac{1}{K} \beta_1' \Pi_2^{*\prime} P_{(\Pi_2^* + \Xi)} \Pi_2^* \beta_1 \xrightarrow{a.s.} \infty,$$

so the PRRF test has unitary power:

$$\mathcal{K}_{PR}(\beta_1; \mathbf{0}) = 1.$$

For $\beta_0 \neq \mathbf{0}$, note that now the noncentrality parameter λ_K is such that

$$\begin{aligned}
&\frac{1}{\sqrt{K}} \lambda_K - \sqrt{K} \beta_0 \left(\Omega_{\Pi_2^*} + I_m\right)^{-1} \beta_0 \\
&= \left[\frac{1}{\sqrt{K}} \beta_0' \Xi' P_{(\Pi_2^* + \Xi)} \Xi \beta_0 - \sqrt{K} \beta_0 \left(\Omega_{\Pi_2^*} + I_m\right)^{-1} \beta_0\right] - 2\left[\frac{1}{\sqrt{K}} \beta_0' \Xi' P_{(\Pi_2^* + \Xi)} \Pi_2^* \frac{1}{K^{1/4}} (\beta_1 - \beta_0)\right] \\
&\quad + \left[\frac{1}{\sqrt{K}} \cdot \frac{1}{K^{1/4}} (\beta_1 - \beta_0)' \left(\Pi_2^* P_{(\Pi_2^* + \Xi)} \Pi_2^*\right) \frac{1}{K^{1/4}} (\beta_1 - \beta_0)\right] \qquad (41)
\end{aligned}$$

By (40), the term in the first square bracket converges to a nondegenerate random variable. By (38), the term in the third bracket converges to a finite constant:

$$(\beta_1 - \beta_0)' \Omega_{\Pi_2^*}' \left(\Omega_{\Pi_2^*} + I_m\right)^{-1} \Omega_{\Pi_2^*} (\beta_1 - \beta_0)$$

By (39),

$$\frac{1}{K} \beta_0' \Xi' P_{(\Pi_2^* + \Xi)} \Pi_2^* (\beta_1 - \beta_0) \xrightarrow{a.s.} \beta_0' \left(\Omega_{\Pi_2^*} + I_m\right)^{-1} \Omega_{\Pi_2^*} (\beta_1 - \beta_0).$$

If the right-hand side is nonzero, the term in the second bracket of (41) may diverge to positive or negative infinity, depending on the sign of $\beta_0^{'}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1-\beta_0\right)\neq 0$. Then:

$$\frac{1}{\sqrt{K}}\lambda_K-\sqrt{K}\beta_0\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\beta_0\xrightarrow{a.s.}\begin{cases}\infty, & \beta_0^{'}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1-\beta_0\right)<0,\\ -\infty, & \beta_0^{'}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1-\beta_0\right)>0.\end{cases}$$

In the particular case where

$$\beta_0^{'}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1-\beta_0\right)=0,$$

notice that

$$\begin{aligned}
&\frac{1}{\sqrt{K}}\beta_0^{'}\Xi^{'}P_{\left(\Pi_2^*+\Xi\right)}\Pi_2^*\frac{1}{K^{1/4}}\left(\beta_1-\beta_0\right)\\
&=\beta_0^{'}\frac{1}{K}\Xi^{'}\left(\Pi_2^*+\Xi\right)\left[\frac{1}{K}\left(\Pi_2^*+\Xi\right)^{'}\left(\Pi_2^*+\Xi\right)\right]^{-1}\frac{1}{K^{3/4}}\left(\Pi_2^{*'}\Pi_2^*+\Xi^{'}\Pi_2^*\right)\left(\beta_1-\beta_0\right)\\
&=K^{-\frac{1}{4}}\cdot\beta_0^{'}\frac{1}{K}\Xi^{'}\left(\Pi_2^*+\Xi\right)\left[\frac{1}{K}\left(\Pi_2^*+\Xi\right)^{'}\left(\Pi_2^*+\Xi\right)\right]^{-1}\\
&\times\left(\sqrt{K}\left(\frac{1}{K}\Pi_2^{*'}\Pi_2^*-\Omega_{\Pi_2^*}\right)+\frac{1}{\sqrt{K}}\Xi^{'}\Pi_2^*\right)\left(\beta_1-\beta_0\right)\\
&\quad+K^{-\frac{1}{4}}\cdot\beta_0^{'}\frac{1}{K}\Xi^{'}\left(\Pi_2^*+\Xi\right)\left[\frac{1}{K}\left(\Pi_2^*+\Xi\right)^{'}\left(\Pi_2^*+\Xi\right)\right]^{-1}\sqrt{K}\Omega_{\Pi_2^*}\left(\beta_1-\beta_0\right)\\
&=o_{a.s.}\left(1\right)-K^{-\frac{1}{4}}\beta_0^{'}\left(\frac{1}{\sqrt{K}}\Xi^{'}\Pi_2^*+\frac{1}{\sqrt{K}}\Xi^{'}\Xi\right)\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1-\beta_0\right)\\
&=o_{a.s.}\left(1\right)-o_{a.s}\left(1\right)-K^{-\frac{1}{4}}\beta_0^{'}\sqrt{K}\left(\frac{1}{K}\Xi^{'}\Xi-I\right)\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1-\beta_0\right)\\
&\quad+K^{-\frac{1}{4}}\sqrt{K}\beta_0^{'}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1-\beta_0\right)\\
&=o_{a.s.}\left(1\right)
\end{aligned}$$

since $\beta_0^{'}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1-\beta_0\right)=0$. In this particular case, by (41),

$$\begin{aligned}
&\frac{1}{\sqrt{K}}\lambda_K-\sqrt{K}\beta_0\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\beta_0\\
&\xrightarrow{a.s.}2\beta_0^{'}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\left(\Delta_{\Pi_2^{*'}\Xi}+\Delta_{\Xi^{'}\Xi}\right)\beta_0-\beta_0^{'}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\Delta\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\beta_0\\
&\quad+\left(\beta_1-\beta_0\right)^{'}\Omega_{\Pi_2^*}^{'}\left(\Omega_{\Pi_2^*}+I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1-\beta_0\right).
\end{aligned}$$

Hence,

$$
\begin{aligned}
\tilde{W}_{PR}^{\beta_0} &:= \frac{1}{\sqrt{K}}\left(W_{PR}^{\beta_0} - \frac{m + K\beta_0^{'}\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\beta_0}{1+\beta_0^{'}\beta_0}\right) \\
&= \frac{\sqrt{\frac{m}{K} + 2\frac{\lambda_K}{K}}}{1+\beta_0^{'}\beta_0} \cdot \frac{\left(\xi - \Xi\beta_0 + \Pi_2^* \frac{1}{\sqrt{K}}\left(\beta_1 - \beta_0\right)\right)^{'} P_{\Pi_2^*+\Xi}\left(\xi - \Xi\beta_0 + \Pi_2^* \frac{1}{\sqrt{K}}\left(\beta_1 - \beta_0\right)\right) - m - \lambda_K}{\sqrt{m + 2\lambda_K}} \\
&\quad + \frac{\frac{1}{\sqrt{K}}\lambda_K - \sqrt{K}\beta_0\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\beta_0}{1+\beta_0^{'}\beta_0} \\
&\xrightarrow{a.s.} \begin{cases} +\infty, & \beta_0^{'}\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1 - \beta_0\right) < 0, \\ \mathcal{N}\left(\mu, \mathbb{AV}\mathrm{ar}\left[\tilde{W}_{PR}^{\beta_0}\right]\right), & \beta_0^{'}\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1 - \beta_0\right) = 0, \\ -\infty, & \beta_0^{'}\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1 - \beta_0\right) > 0. \end{cases}
\end{aligned}
$$

where

$$
\mu := \frac{\left(\beta_1 - \beta_0\right)^{'}\Omega_{\Pi_2^*}^{'}\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1 - \beta_0\right)}{1+\beta_0^{'}\beta_0} \geq 0.
$$

Thus, the asymptotic power of the PRRF test is given by

$$
\mathcal{K}_{PR}\left(\beta_1;\beta_0\right) = \begin{cases} 1, & \beta_0^{'}\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1 - \beta_0\right) < 0, \\ Q\left(Q^{-1}\left(\alpha\right) - \frac{\mu}{\sqrt{\mathbb{AV}\mathrm{ar}\left[\tilde{W}_{PR}^{\beta_0}\right]}}\right) \geq \alpha, & \beta_0^{'}\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1 - \beta_0\right) = 0, \\ 0, & \beta_0^{'}\left(\Omega_{\Pi_2^*} + I_m\right)^{-1}\Omega_{\Pi_2^*}\left(\beta_1 - \beta_0\right) > 0. \end{cases}
$$

∎