Likelihood Ratio based Joint Test for the Exogeneity and the Relevance of Instrumental Variables*

DUKPA KIM† YOONSEOK LEE‡

September 2009

Abstract

This paper develops a joint test for the exogeneity and the relevance of instrumental variables using an approach similar to Vuong's (1989) model selection test. The test statistic is derived from the likelihood ratio of two competing models: one with exogenous and possibly relevant instruments and the other with irrelevant and even possibly endogenous instruments. The joint test is asymptotically pivotal under the null hypothesis that the instruments are exogenous and irrelevant, and is consistent against the alternative hypothesis that the instruments are exogenous and relevant. In contrast, the probability of rejecting the null decreases as the endogeneity of the instruments gets severer. Hence, non-rejection of the joint test should be taken as an evidence suggesting instruments of poor quality. Another salient feature of the test is that its asymptotic null distribution is the same under both the conventional and the weak instruments asymptotic frameworks, which implies it has better size control than other tests employing only one particular asymptotic framework such as the overidentifying restrictions test.

Keywords: Joint test, exogeneity, relevance, instrumental variables, likelihood ratio.

JEL Classifications: C12; C30

---

*First draft: March 2009.

†Department of Economics, University of Virginia, Monroe Hall, McCormick Rd., Charlottesville, VA 22903; *email*: `dukpa@virginia.edu`

‡Corresponding Author. Department of Economics, University of Michigan, 611 Tappan Street, Ann Arbor, MI 48109; *email*: `yoolee@umich.edu`

1 Introduction

A set of instrumental variables is said to be *relevant* if they are correlated with the endogenous regressors and *exogenous* if uncorrelated with the errors. It is a common practice in empirical researches to check these two conditions since the standard inference results on the structural parameters hold only when these conditions hold. The overidentifying restrictions (OID) test (e.g., Anderson and Rubin, 1949; Sargan, 1958; Basmann, 1960) is widely used for the exogeneity condition, and the first stage F and Wald tests are typically used for the relevance condition (or for the weak instruments). Hall, Rudebusch and Wilcox (1996) and Stock and Yogo (2005) are more recently developed relevance tests. All aforementioned testing procedures are, however, designed for only one of the two conditions and there is no test considering both conditions simultaneously. Surprisingly, it has not been discussed much in the econometrics literature how to combine and interpret these two types of tests.[1]

Importantly, the null distribution of the OID test is approximated by the Chi-square distribution under the implicit assumption that the instruments are relevant and strong.[2] As Staiger and Stock (1997) point out, however, the Chi-square distribution is a good approximation only if the instruments are strongly correlated with the endogenous regressors. Therefore, without the knowledge of the relevance of the instruments, we cannot be sure about the legitimacy of the Chi-square approximation. One may consider a two-stage testing procedure−testing for the (weak) relevance first and continuing to test for the exogeneity if the first stage relevance test rejects no or weak relevance. However, the distribution of the OID test conditional on the rejection of the relevance test can be quite different from the unconditional distribution and the literature has not determined, even asymptotically, what the exact error probability is when we use the conventional critical values. This paper

[1] Recently, Moreira (2003) among others, propose an inferential method that is robust to arbitrarily weak instruments. This weak instruments robust inference, however, requires exogenous instruments and the necessity to check the exogeneity of instruments still remains (e.g., Doko and Dufour, 2008).

[2] In addition, it is also required that the number of instruments is fixed or small relative to the sample size. Lee and Okui (2009), for example, study nonstandard asymptotic properties of the Sargan test with many instruments.

presents some Monte Carlo experiment results related to this issue.

The main contribution of this paper is to develop a testing procedure considering the relevance and exogeneity conditions at the same time. We employ an approach similar to Vuong's (1989) model selection test. From a structural equation and its associated reduced form equation, we consider two competing models: one imposing the instruments to be exogenous and the other imposing the instruments to be completely irrelevant. Assuming normality, we show that the likelihood ratio of these two models is equivalent to the difference of the standard first stage Wald statistic and the OID test statistic. We propose a new Q_{IV} test based on this likelihood ratio, and it can be viewed as providing a formal way of interpreting the difference between the commonly used relevance test and the OID test.

More precisely, we set the null hypothesis as the intersection of the two models described above so that the instruments are exogenous and irrelevant. Then, the Q_{IV} statistic is shown to be asymptotically pivotal under the null hypothesis, whereas it diverges to positive infinity when the instruments are relevant and exogenous. This implies that the probability of rejecting the Q_{IV} test with a large positive value approaches one as the sample size grows if the instruments are indeed exogenous and relevant. In contrast, for any given level of relevance, the probability of rejecting the null hypothesis decreases toward zero as the endogeneity of the instruments gets severer. Hence, non-rejection of the joint test should be taken as an evidence suggesting instruments of poor quality.

Another salient feature of our test statistic is that its asymptotic null distribution is invariant to the asymptotic framework: the limiting distribution is the same under both the conventional and the Staiger and Stock (1997) weak instruments asymptotic framework. This is a very important property because it implies that our test has better size control than other tests employing only one particular asymptotic framework such as the OID test.

One caveat is that the Q_{IV} statistic could diverge to positive infinity even when the instruments are not strictly exogenous to the structural error. A leading case of this instance is when the instruments slightly violate the exogeneity condition while they retain strong correlation with the endogenous regressors, so that the asymptotic bias of the instrumental variables estimator is smaller than that of the ordinary least squares estimator. If a set of instruments indeed reduces the asymptotic bias relative to the OLS estimator, it is deemed to be of good quality and the Q_{IV} test concludes likewise. Nevertheless, it should be noted that rejection of the Q_{IV} test with a large positive value should not be taken as a strong evidence of completely exogenous and relevant instruments, while non-rejection of the Q_{IV} test should still be taken as an evidence that the set of instruments is of poor quality.

This paper is organized as follows. Section 2 describes the model and the new test statistic Q_{IV}. Section 3 presents the asymptotic properties of the Q_{IV} test under the null and the alternative hypotheses. Section 4 contains Monte Carlo experiment results. Section 5 concludes with some remarks. All the technical proofs and simulation results are provided in Appendix.

2 Model and Test Statistic

We consider a structural equation and an associated reduced form equation given by

$$y = Y\beta + X\alpha + \varepsilon \quad (1)$$

$$Y = Z\Pi + X\Phi + V, \quad (2)$$

where y is a $T \times 1$ vector, Y is a $T \times n$ matrix of n endogenous variables, X is a $T \times K_1$ matrix of (included) exogenous variables, and Z is a $T \times K_2$ matrix of (excluded) exogenous variables to be used as instruments. The number of instruments, K_2, satisfies $n < K_2 < T$ and it is assumed to be fixed. ε and V are, respectively, a $T \times 1$ vector and a $T \times n$ matrix of random disturbances.

In the standard setup, the set of instrumental variables, Z, is assumed to be uncorrelated with both ε and V, so that Y is correlated with ε only through the correlation between ε and V. We, however, allow for a more general framework using the following structure, under which Z and ε could be also correlated.[3]

[3] Since we consider a linear system given by (1) and (2), we naturally assume linear correlation between

Assumption 1 *$\varepsilon = Z\omega + u$ in (1), where u is correlated with V.*

Using Assumption 1, we can rewrite the structural equation (1) as (e.g., Basmann, 1960)

$$y = Y\beta + X\alpha + Z\omega + u. \tag{3}$$

In this specification, the set of instrumental variables Z is *exogenous* if $\omega = 0$ so that it is orthogonal to the structural error ε. Z is *relevant* if $\Pi \neq 0$ or more precisely Π is of full column rank, and thus correlations between the instruments and the endogenous regressors are nonzero. Based on (3), we can rewrite the model in a system of equations given by

$$\bar{Y}\begin{pmatrix} 1 & 0 \\ -\beta & I_n \end{pmatrix} = \bar{Z}\begin{pmatrix} \omega & \Pi \\ \alpha & \Phi \end{pmatrix} + \bar{V}, \tag{4}$$

where I_n is the identity matrix with rank n, $\bar{Y} = [y, Y]$, $\bar{V} = [u, V]$ and $\bar{Z} = [Z, X]$. If we define $P_W = W(W'W)^{-1}W'$ and $M_W = I - P_W$ for any matrix W, we can project out X from (2) and (3) to have $y^{\perp} = Y^{\perp}\beta + Z^{\perp}\omega + u^{\perp}$ and $Y^{\perp} = Z^{\perp}\Pi + V^{\perp}$, where $A^{\perp} = M_X A$ for any matrix A.

The main interest of this paper is to develop a joint test for the exogeneity and the relevance of a set of instruments by considering the two conditions at the same time: $\omega = 0$ and $\Pi \neq 0$.[4] Obviously, it is difficult to test such a composite hypothesis in the standard testing framework. We instead take an approach similar to the model selection test of Vuong (1989).

Z and ε here. However, this assumption is not overly restrictive because we assume no correlation under the null hypothesis and the linearity does not play any role in deriving the asymptotic null distribution of our test statistic.

[4]It may also be interesting to test for the exogeneity (i.e., $\omega = 0$) and strength of instruments. The distinction between weak and strong instruments is typically formulated under the local-to-zero asymptotics and testing for the strength should also be done under the same asymptotics (e.g., Stock and Yogo, 2005). However, this approach induces nuisance parameters in the limiting distribution of the test statistic that we propose and makes the test unattractive.

To this end, we consider the following two non-nested models:

$$\text{Model } \mu_\omega \quad : \quad \bar{Y}\begin{pmatrix} 1 & 0 \\ -\beta & I_n \end{pmatrix} = \bar{Z}\begin{pmatrix} 0 & \Pi \\ \alpha & \Phi \end{pmatrix} + \bar{V};$$

$$\text{Model } \mu_\Pi \quad : \quad \bar{Y}\begin{pmatrix} 1 & 0 \\ -\beta & I_n \end{pmatrix} = \bar{Z}\begin{pmatrix} \omega & 0 \\ \alpha & \Phi \end{pmatrix} + \bar{V}.$$

The first model μ_ω is (4) with a restriction $\omega = 0$, whereas the second model μ_Π is (4) with a restriction $\Pi = 0$. Note that under the first model μ_ω, the instruments are exogenous though its relevance is left unspecified. Under the second model μ_Π, the instruments are irrelevant and could be even endogenous depending on the value of ω. The key idea is that the likelihood ratio between the two models, μ_ω and μ_Π, can be used as a model selection test. Formally, we set the null hypothesis as

$$H_0 : \omega = 0 \text{ and } \Pi = 0, \tag{5}$$

which implies that two specifications are equally close to the true data generating model. When the null hypothesis is rejected in favor of the alternative hypothesis

$$H_1 : \omega = 0 \text{ and } \Pi \neq 0 \tag{6}$$

(i.e., the first specification μ_ω is closer to the true model), we may expect the set of instruments to be exogenous and relevant or likely to be so. It should be noted that there is an important difference between Vuong's (1989) approach and ours. For Vuong's (1989) model selection test, two competing models are equally important, and each model is required to have a unique value of the model parameter vector that minimizes the Kullback-Leibler distance to the true distribution. Because of this condition, Vuong's (1989) test can select the model closer to the true distribution with probability approaching one as the sample size grows. In our setup, the model μ_Π does not satisfy this condition because some parameters are not properly identified. However, this lack of identification in the model μ_Π is not critical because it is of little interest to differentiate the model μ_Π from the null hypothesis H_0. In other words, the main purpose of the test in this paper is not selecting a model between μ_ω and μ_Π but to reject the null hypothesis H_0 in favor of the model μ_ω.

To derive a test statistic, we assume that $(u_t, V_t')'|Z_t, X_t \sim i.i.d.\mathcal{N}(0, \Sigma)$,[5] where

$$\Sigma = \begin{pmatrix} \sigma_{uu} & \Sigma_{uV} \\ \Sigma_{Vu} & \Sigma_{VV} \end{pmatrix}$$

with $\Sigma_{uV}(= \Sigma_{Vu}') \neq 0$ and the partition is conformable with $(u_t, V_t')'$. The likelihood function is denoted as $L(\theta)$ with $\theta = (\beta', \omega', vec(\Pi)', vec(\Phi)', vec(\Sigma)')'$. Then, the likelihood ratio between the non-nested models μ_ω and μ_Π can be derived as

$$\begin{aligned} 2LR &= 2 \max_{\theta:\omega=0} \log L(\theta) - 2 \max_{\theta:\Pi=0} \log L(\theta) \\ &= T\left(\log\left|I_n + \frac{1}{T}G_T\right| - \log(1 + \frac{1}{T}\phi(\hat{\beta}_{LIML}))\right) \\ &\simeq tr\,(G_T) - \phi(\hat{\beta}_{LIML}), \end{aligned} \tag{7}$$

where $tr\,(\cdot)$ is the trace operator, $\hat{\beta}_{LIML}$ is the standard LIML estimator,

$$G_T = \hat{\Sigma}_{VV}^{-1/2}\left(Y^{\perp\prime}P_{Z^\perp}Y^\perp\right)\hat{\Sigma}_{VV}^{-1/2} \quad \text{and} \quad \phi(\hat{\beta}_{LIML}) = \frac{\hat{\varepsilon}^{\perp\prime}P_{Z^\perp}\hat{\varepsilon}^\perp}{\hat{\varepsilon}^{\perp\prime}M_{Z^\perp}\hat{\varepsilon}^\perp/T} \tag{8}$$

with $\hat{\Sigma}_{VV} = Y^{\perp\prime}M_{Z^\perp}Y^\perp/T$, $\hat{\varepsilon}^\perp = y^\perp - Y^\perp\hat{\beta}_{LIML}$. The detailed derivation of (7) is given in Appendix. Notice that the first component of (7), $tr\,(G_T)$, is nothing but the Wald statistic testing for $\Pi = 0$. The commonly used first stage F statistic is equivalent to this statistic when there is only one endogenous regressor; Hall, Rudebusch and Wilcox's (1996), and Stock and Yogo's (2005) statistics are its variants. On the other hand, the second component of (7), $\phi(\hat{\beta}_{LIML})$, is the standard overidentifying restrictions (OID) test statistic. For example, it is Anderson-Rubin (1949) statistic when the true value of β is used

[5] The normality assumption is not needed for our main asymptotic results presented in the next section.

instead of $\hat{\beta}_{LIML}$ and is the Basmann's (1960) OID test when the two stage least squares (TSLS) estimator $\hat{\beta}_{TSLS}$ is used.

From the model selection point of view, a large positive value of the LR statistic in (7) indicates that the model μ_ω has a Kullback-Leibler distance to the true model smaller than that of the model μ_Π. In addition, as the sample size grows, we can show the LR statistic tends to positive infinity when the model μ_ω is true with $\Pi \neq 0$, whereas it shifts to the opposite direction if the model μ_Π is true with $\omega \neq 0$. Under the intersection of the two models (i.e., both $\omega = 0$ and $\Pi = 0$), the LR statistic is asymptotically pivotal as Theorem 1 in the following section. It is thus natural to consider a testing procedure which concludes that a given set of instruments is closer to being exogenous and relevant (i.e., of good quality) when the LR in (7) takes a large positive value.

The new joint test statistic developed in this paper has basically the same structure as the LR statistic in (7). Specifically, the test statistic (on the quality of instrumental variables: Q_{IV}) that we consider is defined as[6]

$$Q_{IV} = \lambda_{\min}\left(G_T^0\right) - \phi(\hat{\beta}(k_T)), \tag{9}$$

where $\lambda_{\min}(\cdot)$ is the minimum eigenvalue of a given matrix and

$$\begin{aligned} G_T^0 &= \tilde{\Sigma}_{VV}^{-1/2}\left(Y^{\perp\prime}P_{Z^\perp}Y^\perp\right)\tilde{\Sigma}_{VV}^{-1/2} \ \text{ with } \tilde{\Sigma}_{VV} = \frac{1}{T}Y^{\perp\prime}Y^\perp, \\ \phi(\hat{\beta}(k_T)) &= \frac{\tilde{\varepsilon}^{\perp\prime}P_{Z^\perp}^\perp\tilde{\varepsilon}^\perp}{\tilde{\varepsilon}^{\perp\prime}M_{Z^\perp}\tilde{\varepsilon}^\perp/T} \ \text{ with } \tilde{\varepsilon}^\perp = y^\perp - Y^\perp\hat{\beta}(k_T). \end{aligned}$$

Here we compute the covariance matrix of the reduced form error V with assuming $\Pi = 0$. $\hat{\beta}(k_T)$ is the standard k-class estimator defined as

$$\hat{\beta}(k_T) = (Y^{\perp\prime}(I_T - k_T M_{Z^\perp})Y^\perp)^{-1}Y^{\perp\prime}(I_T - k_T M_{Z^\perp})y^\perp, \tag{10}$$

[6]It should be noted that we need more number of instruments than the number of endogenous regressors because the test statistic Q_{IV} uses the OID test statistic.

in which $k_T = 1$ for the TSLS estimator; $k_T = \hat{k}_T$ for the LIML estimator; and $k_T = \hat{k}_T - 1/(T - K_1 - K_2)$ for the Fuller-k estimator with $\hat{k}_T$ being the smallest root satisfying $|\bar{Y}' M_X \bar{Y} - \hat{k}_T \bar{Y}' M_{\bar{Z}} \bar{Y}| = 0$. Note that G_T^0 and G_T are asymptotically equivalent if $\Pi = 0$, and $\lambda_{\min}(G_T)$ is the test statistic for instrument weakness suggested by Stock and Yogo (2005), which is based on Cragg and Donald's (1993) statistic. An interesting point is that the new test statistic Q_{IV} can be interpreted as the difference between a relevant (or weak) instrument test statistic and the standard OID test statistic. Our test procedure can also be viewed as providing a formal way of interpreting the difference between the commonly used relevance test and the OID test.

Remark 1 One reason to consider the minimum eigenvalue in (9) instead of the sum of all eigenvalues given by the trace operator in (7) is, as discussed by Hall, Rudebusch and Wilcox (1996), that testing the significance of $\lambda_{\min}(G_n)$ is equivalent to testing the significance of the smallest canonical correlation between Y and Z. If the smallest canonical correlation is not significantly different from zero then the predicted Y from the first stage regression (2) is likely to be rank deficient, which will yield lack of identification in the second stage regression (1).

3 Asymptotic Results

We first derive the asymptotic distribution of the Q_{IV} statistic (9) under the null hypothesis (5). We let $\rho = \Sigma_{VV}^{-1/2} \Sigma_{Vu} \sigma_{uu}^{-1/2}$ and $\Omega = S_{ZZ} - S_{ZX} S_{XX}^{-1} S_{XZ}$, where

$$S = \mathbb{E}\left(\bar{Z}_t \bar{Z}_t'\right) = \begin{pmatrix} S_{ZZ} & S_{XZ} \\ S_{ZX} & S_{XX} \end{pmatrix}$$

with $\bar{Z}_t'$ being the t-th row of $\bar{Z}$ and $S_{ZX} = S_{XZ}'$. We make the high level assumptions following Staiger and Stock (1997).

Assumption 2 **(a)** *$T^{-1}\bar{V}'\bar{V} \xrightarrow{p} \Sigma$ and $T^{-1}\bar{Z}'\bar{Z} \xrightarrow{p} S$ as $T \to \infty$, where both Σ and S are positive definite and finite.* **(b)** *$(X'u, Z'u, X'V, Z'V)/\sqrt{T} \xrightarrow{d} (\Psi_{Xu}, \Psi_{Zu}, \Psi_{XV}, \Psi_{ZV})$ as $T \to \infty$, where $(\Psi'_{Xu}, \Psi'_{Zu}, vec(\Psi_{XV})', vec(\Psi_{ZV})')' \sim \mathcal{N}(0, \Sigma \otimes S)$.*

Based on Assumption 2, we also define Gaussian random matrices $z_u = \Omega^{-1/2}(\Psi_{Zu} - S_{ZX}S_{XX}^{-1}\Psi_{Xu})\sigma_{uu}^{-1/2}$ and $z_V = \Omega^{-1/2}(\Psi_{ZV} - S_{ZX}S_{XX}^{-1}\Psi_{XV})\Sigma_{VV}^{-1/2}$ so that $(z_u', vec(z_V)')' \sim \mathcal{N}\left(0, \overline{\Sigma} \otimes I_{K_2}\right)$ with $\overline{\Sigma} = \begin{pmatrix} 1 & \rho' \\ \rho & I_n \end{pmatrix}$. The first theorem derives the asymptotic distribution of the Q_{IV} statistic under the null hypothesis (5).

Theorem 1 *We suppose $\omega = 0$ and $\Pi = 0$. Under Assumptions 1 and 2, $\kappa_T = T(\hat{k}_T - 1) \xrightarrow{d} \kappa^*$ as $T \to \infty$, where κ^* is the smallest root satisfying $|(\eta, z_V)'(\eta, z_V) - \kappa^* I_{n+1}| = 0$ and $\eta = (z_u - z_V\rho)/\sqrt{1 - \rho'\rho}$ so that $(\eta', vec(z_V)')' \sim \mathcal{N}\left(0, I_{(n+1)K_2}\right)$. Furthermore,*

$$Q_{IV} \xrightarrow{d} Q_{IV}^0 = \lambda_{\min}(z_V' z_V) - \frac{\eta'(I_{K_2} - z_V(z_V'z_V - \kappa I_n)^{-1}z_V')^2\eta}{1 + \eta' z_V(z_V'z_V - \kappa I_n)^{-2}z_V'\eta},$$

where $\kappa = 0$ for the TSLS estimator; $\kappa = \kappa^*$ for the LIML estimator; and $\kappa = \kappa^* - 1$ for the Fuller-k estimator.

Although the Q_{IV} test does not conform to the standard maximum likelihood ratio test, its limiting null distribution is nuisance parameter free (i.e., the Wilks phenomenon). It depends only on the number of instrumental variables (K_2) and the number of endogenous regressors (n). Tables 1.A to 1.C in Appendix report the relevant quantiles of Q_{IV}^0. It is also evident from Theorem 1 that the LR statistic in (7) is asymptotically pivotal. The following theorem derives the asymptotic behavior of the Q_{IV} statistic under various hypotheses including the alternative hypothesis (6).

Theorem 2 *Under Assumptions 1 and 2, as $T \to \infty$ we have the following asymptotic results:*

(i) If $\omega = 0$ and $\Pi \neq 0$, then $Q_{IV} \xrightarrow{p} \infty$.

(ii) If $\omega \neq 0$ and $\Pi = 0$, then $\kappa_T = T(\hat{k}_T - 1) = O_p(1)$. Moreover, let $\kappa_T \xrightarrow{d} \kappa^*$, then

$$Q_{IV} \xrightarrow{d} Q_{IV}^{\omega} = \lambda_{\min}(z_V' z_V) - \frac{\omega' \Omega^{1/2\prime} \left(I_{K_2} - z_V (z_V' z_V - \kappa I_n)^{-1} z_V' \right)^2 \Omega^{1/2} \omega}{\omega' \Omega^{1/2\prime} z_V (z_V' z_V - \kappa I_n)^{-2} z_V' \Omega^{1/2} \omega}, \tag{11}$$

where $\kappa = 0$ for the TSLS estimator; $\kappa = \kappa^*$ for the LIML estimator; and $\kappa = \kappa^* - 1$ for the Fuller-k estimator.

(iii) If $\omega \neq 0$ and $\Pi \neq 0$, then $\hat{k}_T \xrightarrow{p} k^*$ with k^* being the smallest root satisfying $|\Theta - (k^* - 1)\Sigma| = 0$ with $\Theta = \left(\begin{smallmatrix} \omega'\Omega\omega & \omega'\Omega\Pi \\ \Pi'\Omega\omega & \Pi'\Omega\Pi \end{smallmatrix}\right)$, and

$$\text{plim}_{T\to\infty} \frac{Q_{IV}}{T} = \lambda_{\min} \left((\Pi'\Omega\Pi + \Sigma_{VV})^{-1} \Pi'\Omega\Pi \right) - \frac{(\omega - \Pi b(k))' \Omega (\omega - \Pi b(k))}{\sigma_{uu} + b(k)' \Sigma_{VV} b(k) - 2\Sigma_{uV} b(k)}, \tag{12}$$

where $b(k) = \text{plim}_{T\to\infty} (\hat{\beta}(k_T) - \beta) = (\Pi'\Omega\Pi - k\Sigma_{VV})^{-1} (\Pi'\Omega\omega - k\Sigma_{Vu})$ with $k = 0$ for the TSLS estimator; and $k = k^* - 1$ for the LIML and the Fuller-k estimators.

Theorem 2-(i) indicates that the Q_{IV} statistic diverges to positive infinity with exogenous and relevant instruments (i.e., under H_1 in (6)), and thus the probability to reject the null hypothesis approaches one as the sample size grows. This result is the basic building block of our new test Q_{IV}: we reject H_0 (i.e., $\omega = 0$ and $\Pi = 0$) in favor of H_1 (i.e., $\omega = 0$ and $\Pi \neq 0$) if Q_{IV} is large enough. Furthermore, Monte Carlo experiments indicate that the distribution of Q_{IV} shifts to the left when $\omega \neq 0$ and $\Pi = 0$ (i.e., Q_{IV}^{ω} in (11) has smaller quantiles than Q_{IV}^{0} uniformly over ω) and the probability to erroneously reject H_0 in favor of H_1 remains under the controlled level, though it does not diverge to the negative infinity as the sample size grows. In addition, Q_{IV}^{ω} does not depend on the degree of endogeneity ρ. Theorem 2-(iii) shows that, for any given Π, $\text{plim}_{T\to\infty} T^{-1} Q_{IV}$ decreases as $||\omega||$ increases except for some pathological cases. Then the implication from the view point of the Bahadur efficiency is that the probability of rejecting the null hypothesis decreases toward zero as the endogeneity of the instruments gets severer. See Section 4 for the relevant simulation results.

One caveat is that Theorem 2-(iii) implies the Q_{IV} statistic could diverge to positive infinity if $\text{plim}_{T\to\infty} T^{-1} Q_{IV} > 0$, even when the instruments are correlated with the structural error ($\omega \neq 0$) as long as they are correlated strongly enough with the endogenous regressors ($\Pi \neq 0$). Therefore, it is important to emphasize that rejection of the Q_{IV} test with a large positive value should not be taken as a strong evidence of good instruments. Rather, non-rejection of the Q_{IV} test should be taken as an evidence that the set of instruments is of poor quality.

However, the sign of $\text{plim}_{T\to\infty} T^{-1} Q_{IV}$ is not completely arbitrary. For example, we consider the TSLS estimator $\hat{\beta}_{TSLS}$ for $\hat{\beta}(k_T)$. Then, $\text{plim}_{T\to\infty} T^{-1} Q_{IV} > 0$ if and only if

$$\begin{aligned} \frac{\lambda_{\min}(\Lambda'\Lambda)}{1+\lambda_{\min}(\Lambda'\Lambda)} &> \frac{\omega'\Omega\omega + b(0)'\Pi'\Omega\Pi b(0) - 2\omega'\Omega\Pi b(0)}{\sigma_{uu} + b(0)'\Sigma_{VV} b(0) - 2\Sigma_{uV} b(0)} \\ &= \frac{\omega'\Omega\omega - \omega'\Omega\Pi(\Pi'\Omega\Pi)^{-1}\Pi'\Omega\omega}{\sigma_{uu} - 2\Sigma_{uV}(\Pi'\Omega\Pi)^{-1}\Pi'\Omega\omega + 2\omega\Omega\Pi(\Pi'\Omega\Pi)^{-1}\Sigma_{VV}(\Pi'\Omega\Pi)^{-1}\Pi'\Omega\omega} \\ &= \frac{\xi' M_\Lambda \xi}{1 - 2\rho'(\Lambda'\Lambda)^{-1}\Lambda'\xi + \xi'\Lambda(\Lambda'\Lambda)^{-2}\Lambda'\xi} \end{aligned} \tag{13}$$

from (12), where $\Lambda = \Omega^{1/2}\Pi\Sigma_{VV}^{-1/2}$, $\xi = \Omega^{1/2}\omega\sigma_{uu}^{-1/2}$ and $M_\Lambda = I_{K_2} - \Lambda(\Lambda'\Lambda)^{-1}\Lambda'$. Note that, given Ω and Σ, the inequality (13) is more likely to hold when $||\xi||$ is very small (i.e., ω is close to zero) and the smallest eigenvalue of the concentration matrix $\Lambda'\Lambda = \Sigma_{VV}^{-1/2}\Pi'\Omega\Pi\Sigma_{VV}^{-1/2}$ is large enough (i.e., Π is of full rank and its elements considerably deviate from zero). In other words, Q_{IV} is more likely to have large positive value when the violation of the exogeneity is not that severe (if any) and the instruments are strong enough to offset the endogeneity problem. Therefore, in this case, the Q_{IV} test concludes that the given set of instruments is good to use for the TSLS. On the other hand, when ω considerably deviates from zero or $\lambda_{\min}(\Lambda'\Lambda)$ is very small (i.e., the exogeneity is severely violated or the instruments are too weak), the inequality (13) is very unlikely. Of course, if we need to find some conditions to make the inequality hold in such cases, the range of ρ needs to be very much restricted.

Remark 2 The sign of $\text{plim}_{T\to\infty} T^{-1} Q_{IV}$ is roughly linked to the relative magnitude of the biases between the instrument-based estimator $\hat{\beta}(k_T)$ and the OLS estimator $\hat{\beta}_{OLS}$:

$\text{plim}_{T\to\infty} T^{-1} Q_{IV}$ is more likely positive when the asymptotic bias of $\hat{\beta}(k_T)$ is smaller than that of $\hat{\beta}_{OLS}$ in absolute values (i.e., the instruments works properly to reduce the endogeneity problem). More precisely, for $\text{plim}_{T\to\infty}\hat{\beta}_{OLS} - \beta = (\Pi'\Omega\Pi + \Sigma_{VV})^{-1}(\Pi'\Omega\omega + \Sigma_{Vu})$, the relative magnitude of the asymptotic biases between $\hat{\beta}_{TSLS}$ and $\hat{\beta}_{OLS}$ with respect to $\Pi'\Omega\Pi$ is given by[7]

$$
\begin{aligned}
\mathcal{R}_b &= \text{plim}_{T\to\infty} \frac{(\hat{\beta}_{TSLS} - \beta)'(\Pi'\Omega\Pi)(\hat{\beta}_{TSLS} - \beta)}{(\hat{\beta}_{OLS} - \beta)'(\Pi'\Omega\Pi)(\hat{\beta}_{OLS} - \beta)} \\
&= \frac{\omega'\Omega\Pi(\Pi'\Omega\Pi)^{-1}\Pi'\Omega\omega}{(\omega'\Omega\Pi + \Sigma_{Vu}')(\Pi'\Omega\Pi + \Sigma_{VV})^{-1}(\Pi'\Omega\Pi)(\Pi'\Omega\Pi + \Sigma_{VV})^{-1}(\Pi'\Omega\omega + \Sigma_{Vu})} \\
&= \frac{\xi' P_\Lambda \xi}{\rho' B_\Lambda \rho + 2\rho' B_\Lambda \Lambda'\xi + \xi'\Lambda B_\Lambda \Lambda'\xi}
\end{aligned}
\tag{14}
$$

where $P_\Lambda = \Lambda\left(\Lambda'\Lambda\right)^{-1}\Lambda'$ and $B_\Lambda = (\Lambda'\Lambda + I_n)^{-1}(\Lambda'\Lambda)(\Lambda'\Lambda + I_n)^{-1}$. If we suppose none of $\xi' P_\Lambda \xi$ and $\xi' M_\Lambda \xi$ is zero, then both expressions in (13) and (14) are more likely to be small as $||\xi|| \to 0$. It thus suggests that for a given pair of (ρ, Λ), small-enough ξ will satisfy the inequality in (13), which is also likely to satisfy $\mathcal{R}_b < 1$ as well (i.e., TSLS has a smaller bias than OLS) since $\lambda_{\min}(\Lambda'\Lambda)/(1 + \lambda_{\min}(\Lambda'\Lambda)) < 1$. In this sense, the sign of $\text{plim}_{T\to\infty} T^{-1} Q_{IV}$ is related to the relative bias $\mathcal{R}_b$. Therefore, when a set of instruments slightly violates the exogeneity condition but its correlation with the endogenous regressors remains strong enough, so that the asymptotic bias of the instrumental variables estimator is smaller than that of the ordinary least squares estimator, the quality of the instruments should be deemed to be good enough and the Q_{IV} test concludes to that direction.

Remark 3 As extreme cases, if $\Lambda'\xi = 0$, $\hat{\beta}_{TSLS}$ has no asymptotic bias but we do not necessarily conclude that the instruments are valid. This implies that the Q_{IV} test is unnecessarily tough. If ξ is in the range space of Λ (or equivalently $\omega = \Pi c$ for some vector c), however, we always conclude that the instruments are valid even when $\omega \neq 0$ regardless of the relative bias. The standard OID tests share the same feature so that it has no power

[7] It is similar to the relative magnitude of the biases with respect to $Y^{\perp\prime}Y^{\perp}$ (e.g., Stock and Yogo, 2005). Note that $Y^{\perp\prime}Y^{\perp}/T \xrightarrow{p} \Pi'\Omega\Pi + \Sigma_{VV}$ and $\Pi'\Omega\Pi$ corresponds to the pure signal from the instruments.

against such a violation of the exogeneity condition (e.g., Newey, 1985).

Remark 4 One may think that the null hypothesis in (5) is also testable by testing for the significance of the coefficient from the regression of $y^{\perp}$ on $Z^{\perp}$, because $y^{\perp} = Z^{\perp}(\Pi\beta + \omega) + V^{\perp}\beta + u^{\perp}$ from (1) and (2). However, this test delivers far less information on the quality of the instruments than the Q_{IV} test does. First, the exact null hypothesis of the significance test is $\Pi\beta + \omega = 0$ in this case, which can hold even when the null hypothesis (5) does not hold. Second, when $\Pi = 0$ and $\omega \neq 0$, the significance test is consistent and this is an extremely undesirable property. Third, when $\Pi \neq 0$ and $\omega \neq 0$, the effect of ω on the power of the significance test is obscure. Hence, both the rejection and non-rejection of the significance test are difficult to interpret in relation to the quality of the instruments.

Remark 5 Rejection of the Q_{IV} test is suggestive of the correlation between the instruments and the endogenous variables. A similar conclusion may be obtained from the first stage F test. However, an important difference is that the rejection of the first stage F test implies the correlation without any consideration of the endogeneity while rejection of the Q_{IV} test implies the correlation that is strong enough to offset the endogeneity.

Finally, in order to investigate the local power property of the test statistic Q_{IV}, we further assume the following local-to-zero assumptions similarly as Staiger and Stock (1997).

Assumption 3 $\omega = d/\sqrt{T}$ *and* $\Pi = C/\sqrt{T}$ *for some* $0 < d, C < \infty$.

One of the novelties of the Q_{IV} statistic is that its limiting distribution under $\omega = 0$ and $\Pi = 0$ is invariant to the asymptotic framework. That is, the limiting distribution is the same either under the conventional or under Staiger and Stock's weak instruments framework. This is a very important feature because the error probabilities of our test is controlled much better than other tests employing only one particular asymptotic framework such as the OID test.

Theorem 3 *Under Assumptions 1, 2 and 3, we have* $Q_{IV} \xrightarrow{d} Q_{IV}^{loc}$ *as* $T \to \infty$, *where*

$$Q_{IV}^{loc} \equiv \lambda_{\min}((z_V + \Lambda_C)'(z_V + \Lambda_C)) - \frac{(z_u - (z_V + \Lambda_C)\Delta_\xi(\kappa))'(z_u - (z_V + \Lambda_C)\Delta_\xi(\kappa))}{1 - 2\rho'\Delta_\xi(\kappa) + \Delta_\xi(\kappa)'\Delta_\xi(\kappa)},$$

$\Delta_\xi(\kappa) = ((z_V + \Lambda_C)'(z_V + \Lambda_C) - \kappa I_n)^{-1}[(z_V + \Lambda_C)'(z_u + \xi_d) - \kappa\rho]$, $\Lambda_C = \Omega^{1/2}C\Sigma_{VV}^{-1/2}$ *and* $\xi_d = \Omega^{1/2}d\sigma_{uu}^{-1/2}$. $\kappa = 0$ *for the TSLS estimator;* $\kappa = \kappa^*$ *for the LIML estimator; and* $\kappa = \kappa^* - 1$ *for the Fuller-k estimator with* κ^* *being the smallest root satisfying* $\left|(z_u + \xi_d, z_V + \Lambda_C)'(z_u + \xi_d, z_V + \Lambda_C) - \kappa^*\overline{\Sigma}\right| = 0$.

Obviously, Q_{IV}^{loc} depends on a nuisance parameter ρ unless $\omega = 0$ and $\Pi = 0$. Furthermore, if there are multiple endogenous variables ($n \geq 2$), Q_{IV}^{loc} depend on all the eigenvalues of $\Lambda_C'\Lambda_C$, as Stock and Yogo (2005) point out. Therefore, this asymptotic distribution cannot be directly used for inferences. See Section 4 where we report the local power of the Q_{IV} test obtained from simulating Q_{IV}^{loc}.

4 Monte Carlo Simulation

4.1 Null rejection probability of the standard OID test

In this subsection, we demonstrate via Monte Carlo experiments the difficulties arising when the relevance test and exogeneity test are used in the conventional manner. First, we show the dependence of the limiting null distribution of the standard OID test on the correlation between the instruments and endogenous variables. In particular, the size of the OID test can depart from the nominal level by a large margin when the correlation between the instruments and endogenous variables is weak. One may consider a two-stage testing procedure−testing for the relevance first and continuing to test for the exogeneity if the first stage relevance test rejects no or weak relevance. In this case, we show that the dependence on the instrumental strength gets intensified, which causes even larger size distortion of the OID test.

Let $\phi(\hat{\beta}(k_T))$ and $\lambda_{\min}(G_T)$ be the standard OID test and Stock and Yogo's (2005) weak instruments test, respectively, as defined in (8), where $\hat{\beta}(k_T)$ is the k-class estimator as in (10). From Theorem 3, the limit expressions for these statistics are

$$\phi(\hat{\beta}(k_T)) \xrightarrow{d} \phi_\infty \equiv \frac{(z_u - (z_V + \Lambda_C)\Delta_\xi(\kappa))'(z_u - (z_V + \Lambda_C)\Delta_\xi(\kappa))}{1 - 2\rho'\Delta_\xi(\kappa) + \Delta_\xi(\kappa)'\Delta_\xi(\kappa)} \quad (15)$$

$$\lambda_{\min}(G_T) \xrightarrow{d} g_\infty \equiv \lambda_{\min}((z_V + \Lambda_C)'(z_V + \Lambda_C)) \quad (16)$$

with $\xi_d = 0$. We simulate the limiting quantities ϕ_∞ and g_∞ in order to avoid any other finite sample complications.

Tables 2.A and 2.B. in Appendix report the rejection probabilities of the OID test, where Tables 2.A is based on the TSLS estimator and Tables 2.B is based on the Fuller-k estimator. The case of one endogenous regressor ($n = 1$) and 3, 9 instrumental variables ($K_2 = 3, 9$) are presented but the results remain qualitatively unchanged for other values of n and K_2. For each value of K_2, there are two columns: "rej." and "uncond." Each of these two columns corresponds to the rejection probabilities of the OID test ϕ_∞ conditional on the rejection of the first-stage weak IV test g_∞ (i.e., $\mathbb{P}(\phi_\infty > \chi^2_{K_2-n,0.05} | g_\infty \text{ rejects})$) and unconditionally (i.e., $\mathbb{P}(\phi_\infty > \chi^2_{K_2-n,0.05})$), respectively. Note that the first-stage weak IV tests using (16) are based on the 10% TSLS/Fuller-k bias (see Stock and Yogo, 2005, for the precise definition) at the 5% significance level. The second-stage OID tests using (15) are based on the Chi-square distribution. Three values of $\Lambda_C'\Lambda_C$ are simulated:[8] 0.5, 0.8, and 1.2 times of $\lambda^*_{\min}$, where $\lambda^*_{\min}$ is the boundary value for the weak instruments set based on the 10% TSLS/Fuller-k bias. Hence, the cases of 0.5 and 0.8 correspond to the weak instruments, while 1.2 to the strong instruments. For each value of $\Lambda_C'\Lambda_C$, different degrees of endogeneity $\rho = 0.1$, 0.3, 0.5, 0.8 and 1.0 are considered.

The first observation in Table 2.A is that the size distortion of ϕ_∞ increases (e.g., see the "uncond." columns) as the instruments get weaker, which shows the danger of applying the standard OID test without knowing the strength of the instruments. Note that the actual sizes vary from less than 4% to more than 20% when the instruments are weakly correlated

[8] $\Lambda_C'\Lambda_C$ is the weak instrument limit of the concentration matrix $\Sigma_{VV}^{-1/2}\Pi' Z^{\perp\prime} Z^{\perp}\Pi\Sigma_{VV}^{-1/2}$.

with the endogenous variables. The second observation is that the sequential procedure creates more size distortion in that the rejection probabilities of ϕ_∞ conditional on the rejection of g_∞ is always greater than their unconditional counterparts. Also, the size of ϕ_∞ is very liberal near $\rho = 1$, while it is mildly conservative near $\rho = 0$. Table 2.B reports the results obtained using the Fuller-k estimator instead of the TSLS estimator. Overall, Table 2.B exhibits a great deal of similarity to Table 2.A and the general conclusions from Table 2.A remain valid. One notable difference is that the largest size distortion is not associated with $\rho = 1$ but with $\rho = 0.1$.

4.2 Size and power properties of the Q_{IV} test

We also conduct Monte Carlo experiments for the finite sample size and power of the Q_{IV} test. The model we simulate is based on (2) and (3) without X:

$$y = Y\beta + Z\omega + u \quad \text{and} \quad Y = Z\Pi + V.$$

We consider the number of endogenous variables $n = 1$, 2, 3 and the number of instruments $K_2 = n + 1$, $n + 3$, $n + 5$. The errors $(u_t, V_t')'$ are specified as

$$u_t = e_t + \Sigma_{Vu}' E_t \quad \text{and} \quad V_t = E_t,$$

where $(e_t, E_t')' \sim i.i.d.\mathcal{N}(0, I_{n+1})$ and Σ_{Vu} is a vector of ones multiplied by $0.5/\sqrt{n}$. Z_t is from a multivariate normal with unit mean and identity variance covariance matrix. Also, we let $\beta = 0$ since the Q_{IV} test is exactly invariant to the value of β. The number of replications is 5,000. The sample size is $T = 50$, 100, 200 and 300.

For the finite sample size simulation, we assume $\omega = 0$ and $\Pi = 0$. The results are reported in Tables 3.A and 3.B. For any sample size T, the Q_{IV} test shows that actual sizes are very close to the nominal 5% whether it is based on the TSLS or Fuller-k estimator.

For the finite sample power simulation, we set $\omega = 0$ and $\Pi = 0.25\Xi_{n,K_2}$, where the columns of Ξ_{n,K_2} are a set of orthonormalized vectors which are randomly selected from a uniform distribution for each value of n and K_2. Tables 4.A and 4.B report the results. They show that the probability rejecting the null hypothesis ($H_0 : \omega = 0$ and $\Pi = 0$) with a large positive value of Q_{IV} quickly approaches one as the sample size grows. The last power experiments assume $\Pi = 0$ while ω is a vector of zeros except for the last element, which is equal to 0.5. This particular shape of ω reflects that only one or two instruments violate the exogeneity condition, which is very likely if the researcher is careful enough. Tables 5.A and 5.B show that the rejection probabilities are much smaller than the nominal 5% for all n and K_2. The last row in Table 5.A corresponds to the limiting case which are obtained from simulating Q_{IV}^{ω} given in Theorem 2. Note that the power does not collapse to zero because Q_{IV} does not diverge to the negative infinity.

Finally, we simulate Q_{IV}^{loc} to see the local power of the Q_{IV} test. Q_{IV}^{loc} depends on ρ and we consider three cases: ρ is proportional to a vector of ones with $||\rho|| = 0.2$, 0.5 and 0.8. We present only the results of the Fuller-k estimator since the TSLS and LIML estimators give very similar results. Also, the results are quite stable across different pairs of (n, K_2) and we report three cases $(n, K_2) = (1, 3)$, $(2, 5)$, and $(3, 7)$. The number of replications is 20,000.

Figure 1.A shows the results when $\omega = 0$ but $\Pi = C/\sqrt{T}$ as C gets away from zero. More precisely, we let $\xi_d = 0$ and $\Lambda_C = c\Xi_{n,K_2}$, where c varies from 0 to 5. In all cases, the power increases toward one as $|c|$ increases. Figure 1.B, on the other hand, shows the results when $\Pi = 0$ but $\omega = d/\sqrt{T}$ as d gets away from zero. We let $\Lambda_C = 0$ and $\xi_d = (0, \ldots, 0, \xi_{K_2})'$, where ξ_{K_2} varies from 0 to 5. In all cases, the power decreases toward zero as $|\xi_{K_2}|$ increases. Lastly, Figure 1.C shows the results when $\omega = d/\sqrt{T}$ and $\Pi = C/\sqrt{T}$ as both d and C get away from zero. We again set $\Lambda_C = c\Xi_{n,K_2}$ and $\xi_d = (0, \ldots, 0, \xi_{K_2})'$, where c varies from 0 to 5 and ξ_{K_2} from 0 to 25. Since it is somewhat difficult to display multiple surfaces distinctively, we report only the case when $||\rho|| = 0.5$. But the results have great resemblance across the values of $||\rho||$. The power decreases toward zero as ξ_{K_2} increases for any given value of c, while it increases toward one as c increases for any given value of ξ_{K_2}. Also, the larger c is, the larger value of ξ_{K_2} is required for the local

asymptotic power to be non-trivial. This points out the fact that the Q_{IV} tests rejects the null hypothesis only when the correlation between the instrumental variables and the endogenous variables is strong enough to offset the correlation between the instrumental variables and the structural errors.

5 Conclusion

A joint test for the IV relevance and the exogeneity conditions is proposed using an approach similar to Vuong's (1989) model selection test. In particular, the test statistic is derived from two competing models: one imposing the instruments to be exogenous and the other imposing the instruments to be completely irrelevant. The likelihood ratio of these two models is shown to be equivalent to the difference of the standard first stage Wald statistic and the OID test statistic.

The proposed Q_{IV} test is based on this likelihood ratio. The null hypothesis is set to be the intersection of the two models described above so that the instruments are exogenous and irrelevant. Then, the Q_{IV} statistic is shown to be asymptotically pivotal under the null hypothesis, and is consistent against the alternative hypothesis that the instruments are relevant and exogenous. On the contrary, the probability of rejecting the null decreases as the endogeneity of the instruments gets severer. This result implies that non-rejection of the Q_{IV} test should be taken as an evidence that the set of instruments is of questionable quality. It should be noted that, however, the Q_{IV} statistic could reject the null hypothesis even if the instruments are not strictly exogenous to the structural error. Therefore, rejection of the Q_{IV} test should not be taken as a strong evidence of exact exogeneity, while the rejection still tells the instruments to be of good quality in the sense that the TSLS estimator based on these instruments is likely to have smaller asymptotic bias relative to the OLS estimator.

Appendix

A.1 Mathematical Proofs

Derivation of the Likelihood Ratio (7) First, we impose $\omega = 0$ and write the log-likelihood as

$$\begin{aligned} \log L(\theta) &= -\frac{T}{2}\log|\sigma_{uu}| - \frac{1}{2\sigma_{uu}}\varepsilon^{\perp\prime}\varepsilon^{\perp} \\ &\quad -\frac{T}{2}\log\left|\Sigma_{V|u}\right| - \frac{1}{2}tr\left[\left(Y^{\perp} - Z^{\perp}\Pi - \varepsilon^{\perp}\delta'\right)'\Sigma_{V|u}^{-1}\left(Y^{\perp} - Z^{\perp}\Pi - \varepsilon^{\perp}\delta'\right)\right], \end{aligned}$$

where $\varepsilon^{\perp} = y^{\perp} - Y^{\perp}\beta$, $\delta = \Sigma_{Vu}/\sigma_{uu}$ and $\Sigma_{V|u} = \Sigma_{VV} - \Sigma_{Vu}\Sigma_{uV}/\sigma_{uu} = \Sigma_{VV}^{1/2}\left(I_n - \rho\rho'\right)\Sigma_{VV}^{1/2}$. We denote the estimates obtained imposing $\omega = 0$ with a subscript 0. From the first order conditions, we obtain $\hat{\sigma}_{uu,0} = \hat{\varepsilon}^{\perp\prime}\hat{\varepsilon}^{\perp}/T$, where $\hat{\varepsilon}^{\perp} = y^{\perp} - Y^{\perp}\hat{\beta}_{LIML}$, and

$$\begin{aligned} \hat{\delta}_0 &= \left(Y^{\perp} - Z^{\perp}\hat{\Pi}_0\right)'\hat{\varepsilon}^{\perp}(\hat{\varepsilon}^{\perp\prime}\hat{\varepsilon}^{\perp})^{-1} = Y^{\perp\prime}M_{Z^{\perp}}\hat{\varepsilon}^{\perp}(\hat{\varepsilon}^{\perp\prime}M_{Z^{\perp}}\hat{\varepsilon}^{\perp})^{-1} \\ \hat{\Pi}_0 &= (Z^{\perp\prime}Z^{\perp})^{-1}Z^{\perp\prime}(Y^{\perp} - \hat{\varepsilon}^{\perp}\hat{\delta}_0') = (Z^{\perp\prime}Z^{\perp})^{-1}Z^{\perp\prime}(Y^{\perp} - \hat{\varepsilon}^{\perp}\hat{\varepsilon}^{\perp\prime}M_{Z^{\perp}}Y^{\perp}(\hat{\varepsilon}^{\perp\prime}M_{Z^{\perp}}\hat{\varepsilon}^{\perp})^{-1}) \\ \hat{\Sigma}_{V|u,0} &= \frac{1}{T}\left(Y^{\perp} - Z^{\perp}\hat{\Pi}_0 - \hat{\varepsilon}^{\perp}\hat{\delta}_0'\right)'\left(Y^{\perp} - Z^{\perp}\hat{\Pi}_0 - \hat{\varepsilon}^{\perp}\hat{\delta}_0'\right) \\ &= \frac{1}{T}Y^{\perp}M_{Z^{\perp}}Y^{\perp} - \frac{1}{T}Y^{\perp\prime}M_{Z^{\perp}}\hat{\varepsilon}^{\perp}(\hat{\varepsilon}^{\perp\prime}M_{Z^{\perp}}\hat{\varepsilon}^{\perp})^{-1}\hat{\varepsilon}^{\perp\prime}M_{Z^{\perp}}Y^{\perp}. \end{aligned}$$

Note that

$$\left|\hat{\Sigma}_{V|u,0}\right| = \frac{1}{T^n}\left|\hat{\Gamma}'\bar{Y}^{\perp\prime}M_{Z^{\perp}}\bar{Y}^{\perp}\hat{\Gamma}\right|\frac{1}{\hat{\varepsilon}^{\perp\prime}M_{Z^{\perp}}\hat{\varepsilon}^{\perp}} = \frac{1}{T^n}\left|\bar{Y}^{\perp\prime}M_{Z^{\perp}}\bar{Y}^{\perp}\right|\frac{1}{\hat{\varepsilon}^{\perp\prime}M_{Z^{\perp}}\hat{\varepsilon}^{\perp}},$$

where $\hat{\Gamma} = \left(\begin{smallmatrix} 1 & 0 \\ -\hat{\beta}_{LIML} & I_n \end{smallmatrix}\right)$. We then have (up to a constant addition)

$$\begin{aligned} \max_{\theta:\omega=0}\log L(\theta) &= -\frac{T}{2}\log\left(|\hat{\sigma}_{uu,0}|\left|\hat{\Sigma}_{V|u,0}\right|\right) \\ &= -\frac{T}{2}\log\left(\frac{\hat{\varepsilon}^{\perp\prime}\hat{\varepsilon}^{\perp}}{\hat{\varepsilon}^{\perp\prime}M_{Z^{\perp}}\hat{\varepsilon}^{\perp}}\frac{1}{T^{n+1}}\left|\bar{Y}^{\perp\prime}M_{Z^{\perp}}\bar{Y}^{\perp}\right|\right) \\ &= -\frac{T}{2}\log(1 + \frac{1}{T}\phi(\hat{\beta}_{LIML})) - \frac{T}{2}\log\left|\frac{1}{T}\bar{Y}^{\perp\prime}M_{Z^{\perp}}\bar{Y}^{\perp}\right| \end{aligned}$$

Now, we impose $\Pi = 0$ and write

$$\begin{aligned}\log L(\theta) &= -\frac{T}{2}\log\left|\sigma_{u|V}\right| - \frac{1}{2\sigma_{u|V}}\left(y^{\perp} - Z^{\perp}\omega - Y^{\perp}\varphi\right)'\left(y^{\perp} - Z^{\perp}\omega - Y^{\perp}\varphi\right) \\ &\quad -\frac{T}{2}\log\left|\Sigma_{VV}\right| - \frac{1}{2}tr\left[\Sigma_{VV}^{-1}Y^{\perp\prime}Y^{\perp}\right],\end{aligned}$$

where $\varphi = \gamma + \beta$, $\sigma_{u|V} = \sigma_{uu} - \Sigma_{uV}\Sigma_{VV}^{-1}\Sigma_{Vu}$ and $\gamma = \Sigma_{VV}^{-1}\Sigma_{Vu}$. We use a subscript 1 to all estimates obtained with $\Pi = 0$ restriction. Similarly as the first case, we have

$$\begin{aligned}\hat{\Sigma}_{VV,1} &= \frac{1}{T}Y^{\perp\prime}Y^{\perp} \\ \hat{\omega}_1 &= (Z^{\perp\prime}Z^{\perp})^{-1}Z^{\perp\prime}\left(y^{\perp} - Y^{\perp}\hat{\varphi}_1\right) = (Z^{\perp\prime}Z^{\perp})^{-1}Z^{\perp\prime}\left(y^{\perp} - Y^{\perp}(Y^{\perp\prime}M_{Z^{\perp}}Y^{\perp})^{-1}Y^{\perp\prime}M_{Z^{\perp}}y^{\perp}\right) \\ \hat{\varphi}_1 &= (Y^{\perp\prime}Y^{\perp})^{-1}Y^{\perp\prime}\left(y^{\perp} - Z^{\perp}\hat{\omega}_1\right) = (Y^{\perp\prime}M_{Z^{\perp}}Y^{\perp})^{-1}Y^{\perp\prime}M_{Z^{\perp}}y^{\perp} \\ \hat{\sigma}_{u|V,1} &= \frac{1}{T}\left(y^{\perp} - Z^{\perp}\hat{\omega}_1 - Y^{\perp}\hat{\varphi}_1\right)'\left(y^{\perp} - Z^{\perp}\hat{\omega}_1 - Y^{\perp}\hat{\varphi}_1\right) \\ &= \frac{1}{T}y^{\perp\prime}M_{Z^{\perp}}y^{\perp} - \frac{1}{T}y^{\perp\prime}M_{Z^{\perp}}Y^{\perp}(Y^{\perp\prime}M_{Z^{\perp}}Y^{\perp})^{-1}Y^{\perp\prime}M_{Z^{\perp}}y^{\perp} = \frac{1}{T}\frac{\left|\bar{Y}^{\perp\prime}M_{Z^{\perp}}\bar{Y}^{\perp}\right|}{\left|Y^{\perp\prime}M_{Z^{\perp}}Y^{\perp}\right|}.\end{aligned}$$

Moreover,

$$\begin{aligned}\left|\hat{\Sigma}_{VV,1}\right| &= \frac{1}{T^n}\left|Y^{\perp\prime}Y^{\perp}\right| = \frac{1}{T^n}\left|Y^{\perp\prime}M_{Z^{\perp}}Y^{\perp} + Y^{\perp\prime}P_{Z^{\perp}}Y^{\perp}\right| \\ &= \frac{1}{T^n}\left|(Y^{\perp\prime}M_{Z^{\perp}}Y^{\perp})^{1/2}(I_n + \frac{1}{T}G_T)(Y^{\perp\prime}M_{Z^{\perp}}Y^{\perp})^{1/2}\right| \\ &= \frac{1}{T^n}\left|I_n + \frac{1}{T}G_T\right|\left|Y^{\perp\prime}M_{Z^{\perp}}Y^{\perp}\right|,\end{aligned}$$

which yields (up to a constant addition)

$$\begin{aligned}\max_{\theta:\Pi=0}\log L(\theta) &= -\frac{T}{2}\log\left(\left|\hat{\sigma}_{u|V,1}\right|\left|\hat{\Sigma}_{VV,1}\right|\right) \\ &= -\frac{T}{2}\log\left(\left|\frac{1}{T}\bar{Y}^{\perp\prime}M_{Z^{\perp}}\bar{Y}^{\perp}\right|\right) - \frac{T}{2}\log\left(\left|I_n + \frac{1}{T}G_T\right|\right).\end{aligned}$$

Therefore, the LR statistic can be derived as

$$\begin{aligned}2LR &= 2\max_{\theta:\omega=0}\log L(\theta) - 2\max_{\theta:\Pi=0}\log L(\theta)\\ &= T\left(\log\left|I_n + \frac{1}{T}G_T\right| - \log(1+\frac{1}{T}\phi(\hat{\beta}_{LIML}))\right)\\ &\simeq tr\,(G_T) - \phi(\hat{\beta}_{LIML}),\end{aligned}$$

where the last approximation is valid since, by construction, $\|G_T\|$ is small under $\Pi = 0$ and so is $\phi(\hat{\beta}_{LIML})$ under $\omega = 0$. ■

Proof of Theorem 1 This is a special case of Theorem 3 with $C = 0$ and $d = 0$. Note that $z_u = (1-\rho'\rho)^{1/2}\eta + z_V\rho$ and

$$\begin{aligned}\Delta_0(\kappa) &= (z_V'z_V - \kappa I)^{-1}(z_V'z_u - \kappa\rho)\\ &= \rho + (1-\rho'\rho)^{1/2}(z_V'z_V - \kappa I)^{-1}z_V'\eta,\end{aligned}$$

which implies

$$\begin{aligned}z_u - z_V\Delta_0(\kappa) &= (1-\rho'\rho)^{1/2}(I - z_V(z_V'z_V-\kappa I)^{-1}z_V')\eta\\ \Delta_0(\kappa)'\Delta_0(\kappa) &= \rho'\rho + (1-\rho'\rho)\eta' z_V(z_V'z_V-\kappa I)^{-2}z_V'\eta + 2(1-\rho'\rho)^{1/2}\rho'(z_V'z_V-\kappa I)^{-1}z_V'\eta\\ \rho'\Delta_0(\kappa) &= \rho'\rho + (1-\rho'\rho)^{1/2}\rho'(z_V'z_V-\kappa I)^{-1}z_V'\eta.\end{aligned}$$

Therefore,

$$Q_{IV} \xrightarrow{d} \lambda_{\min}(z_V'z_V) - \frac{(z_u - z_V\Delta_0(\kappa))'(z_u - z_V\Delta_0(\kappa))}{1 - 2\rho'\Delta_0(\kappa) + \Delta_0(\kappa)'\Delta_0(\kappa)},$$

in which the second component is given by

$$\frac{(z_u - z_V\Delta_0(\kappa))'(z_u - z_V\Delta_0(\kappa))}{1 - 2\rho'\Delta_0(\kappa) + \Delta_0(\kappa)'\Delta_0(\kappa)} = \frac{(1-\rho'\rho)\eta'(I - z_V(z_V'z_V-\kappa I)^{-1}z_V')^2\eta}{1-\rho'\rho + (1-\rho'\rho)\eta' z_V(z_V'z_V-\kappa I)^{-2}z_V'\eta},$$

where $\kappa = 0$ for the TSLS estimator; $\kappa = \kappa^*$ for the LIML estimator; and $\kappa = \kappa^* - 1$ for the Fuller-k estimator with κ^* being the smallest root satisfying $|(z_u, z_V)'(z_u, z_V) - \kappa^*\bar{\Sigma}| = 0$. Note that $|(z_u, z_V)'(z_u, z_V) - \kappa^*\bar{\Sigma}| = |D'(\eta, z_V)'(\eta, z_V)D - \kappa^* D'D| = |(\eta, z_V)'(\eta, z_V) - \kappa^* I_{n+1}| = 0$, where $D = \left(\begin{smallmatrix}(1-\rho'\rho)^{1/2} & 0\\ \rho & I_n\end{smallmatrix}\right)$. ■

Proof of Theorem 2 Part (i) is trivial and omitted. For part (ii), the limit of the first component $\lambda_{\min}(G_T^0)$ is the same as in Theorem 1. For the second component, since the roots of $\left|\bar{Y}^{\perp\prime}\bar{Y}^{\perp} - k_T\bar{Y}^{\perp\prime}M_{Z^\perp}\bar{Y}^{\perp}\right| = 0$ are the same as those of $\left|J'\bar{Y}^{\perp\prime}\bar{Y}^{\perp}J - k_T J'\bar{Y}^{\perp\prime}M_{Z^\perp}\bar{Y}^{\perp}J\right| =$

0 with $J = \begin{pmatrix} 1 & 0 \\ -\beta & I_n \end{pmatrix}$, we consider

$$\begin{aligned} 0 &= |J'\bar{Y}^{\perp\prime}\bar{Y}^{\perp}J - \hat{k}_T J'\bar{Y}^{\perp\prime}M_{Z^\perp}\bar{Y}^{\perp}J| \\ &= |J'\bar{Y}^{\perp\prime}P_{Z^\perp}\bar{Y}^{\perp}J - T(\hat{k}_T - 1)\frac{1}{T}J'\bar{Y}^{\perp\prime}M_{Z^\perp}\bar{Y}^{\perp}J| \\ &= |\hat{\Sigma}^{-1/2\prime}J'\bar{Y}^{\perp\prime}P_{Z^\perp}\bar{Y}^{\perp}J\hat{\Sigma}^{-1/2} - \kappa_T I_{n+1}|, \end{aligned}$$

where κ_T is the smallest root and $\hat{\Sigma} = T^{-1}J'\bar{Y}^{\perp\prime}M_{Z^\perp}\bar{Y}^{\perp}J$. Under $\omega \neq 0$ and $\Pi = 0$, we have $\hat{\Sigma} = T^{-1}J'\bar{Y}^{\perp\prime}M_{Z^\perp}\bar{Y}^{\perp}J \xrightarrow{p} \Sigma$ and we can write

$$P_{Z^\perp}\bar{Y}^{\perp}J = P_{Z^\perp}[y^\perp, Y^\perp]J = [Z^\perp\omega, 0] + [P_{Z^\perp}u^\perp, P_{Z^\perp}V^\perp] \equiv A_1 + A_2.$$

Note that $\lambda_{\min}(\hat{\Sigma}^{-1/2\prime}A_1'A_1\hat{\Sigma}^{-1/2}) = 0$ for all T. We denote by ν_{A_1} the eigenvector of $\hat{\Sigma}^{-1/2\prime}A_1'A_1\hat{\Sigma}^{-1/2}$ corresponding to the zero eigenvalue. Also, $\hat{\Sigma}^{-1/2\prime}A_2'A_2\hat{\Sigma}^{-1/2} = O_p(1)$, which implies that the largest eigenvalue of $\hat{\Sigma}^{-1/2\prime}A_2'A_2\hat{\Sigma}^{-1/2}$ is $O_p(1)$ as well. We let ν_{A_2} be the eigenvector of $\hat{\Sigma}^{-1/2\prime}A_2'A_2\hat{\Sigma}^{-1/2}$ corresponding to the largest eigenvalue. Let ν be the eigenvector of $\hat{\Sigma}^{-1/2\prime}J'\bar{Y}^{\perp\prime}P_{Z^\perp}\bar{Y}^{\perp}J\hat{\Sigma}^{-1/2}$ associated with κ_T. Then, for a nonzero vector x,

$$\begin{aligned} \kappa_T &= \min_{\|x\|=1} x'\hat{\Sigma}^{-1/2\prime}J'\bar{Y}^{\perp\prime}P_{Z^\perp}\bar{Y}^{\perp}J\hat{\Sigma}^{-1/2}x \\ &= \nu'\hat{\Sigma}^{-1/2\prime}J'\bar{Y}^{\perp\prime}P_{Z^\perp}\bar{Y}^{\perp}J\hat{\Sigma}^{-1/2}\nu \\ &= \nu'\hat{\Sigma}^{-1/2\prime}\left(A_1'A_1 + A_2'A_2 + A_1'A_2 + A_2'A_1\right)\hat{\Sigma}^{-1/2}\nu \\ &\leq \nu_{A_1}'\hat{\Sigma}^{-1/2\prime}\left(A_1'A_1 + A_2'A_2 + A_1'A_2 + A_2'A_1\right)\hat{\Sigma}^{-1/2}\nu_{A_1} \\ &= 0 + \nu_{A_1}'\hat{\Sigma}^{-1/2\prime}A_2'A_2\hat{\Sigma}^{-1/2}\nu_{A_1} + 2\nu_{A_1}'\hat{\Sigma}^{-1/2\prime}A_1'A_2\hat{\Sigma}^{-1/2}\nu_{A_1} \\ &\leq \nu_{A_2}'\hat{\Sigma}^{-1/2\prime}A_2'A_2\hat{\Sigma}^{-1/2}\nu_{A_2} + 2(\nu_{A_1}'\hat{\Sigma}^{-1/2\prime}A_1'A_1\hat{\Sigma}^{-1/2}\nu_{A_1})^{1/2}(\nu_{A_1}'\hat{\Sigma}^{-1/2\prime}A_2'A_2\hat{\Sigma}^{-1/2}\nu_{A_1})^{1/2} \\ &= O_p(1) + 0. \end{aligned}$$

Now, for $\hat{\beta}(k_T) = \beta + (V^{\perp\prime}(I_T - k_T M_{Z^\perp})V^\perp)^{-1}V^{\perp\prime}(I_T - k_T M_{Z^\perp})(Z^\perp\omega + u^\perp)$, we have

$$T^{-1/2}(\hat{\beta}(k_T) - \beta) \xrightarrow{d} \Sigma_{VV}^{-1/2}(z_V'z_V - \kappa I_n)^{-1}z_V'\Omega^{1/2}\omega$$

because

$$\begin{aligned} V^{\perp\prime}(I_T - k_T M_{Z^\perp})V^\perp &= V^{\perp\prime}V^\perp - k_T V^{\perp\prime}M_{Z^\perp}V^\perp \\ &= V^{\perp\prime}P_{Z^\perp}V^\perp - (k_T - 1)V^{\perp\prime}M_{Z^\perp}V^\perp \xrightarrow{d} \Sigma_{VV}^{1/2\prime}z_V'z_V\Sigma_{VV}^{1/2} - \kappa\Sigma_{VV} \end{aligned}$$

and

$$
\begin{aligned}
& T^{-1/2}V^{\perp\prime}(I - k_T M_{Z^\perp})(Z^\perp\omega + u^\perp) \\
= \; & T^{-1/2}V^{\perp\prime}(Z^\perp\omega + u^\perp) - T^{-1/2}k_T V^{\perp\prime}M_{Z^\perp}(Z^\perp\omega + u^\perp) \\
= \; & T^{-1/2}V^{\perp\prime}P_{Z^\perp}(Z^\perp\omega + u^\perp) - T(k_T - 1)T^{-3/2}V^{\perp\prime}M_{Z^\perp}(Z^\perp\omega + u^\perp) \xrightarrow{d} \Sigma_{VV}^{1/2\prime} z_V' \Omega^{1/2}\omega,
\end{aligned}
$$

where $\kappa = 0$ for the TSLS estimator, $\kappa = \kappa^*$ for the LIML estimator and $\kappa = \kappa^* - 1$ for the Fuller-k estimator if we let $\kappa_T \xrightarrow{d} \kappa^*$. In addition, $\tilde{\varepsilon}^\perp = y^\perp - Y^\perp\hat{\beta}(k_T) = Z^\perp\omega + u^\perp - V^\perp(\hat{\beta}(k_T) - \beta)$ since $Y^\perp = V^\perp$. We thus have

$$
\begin{aligned}
T^{-1}\tilde{\varepsilon}^{\perp\prime}P_{Z^\perp}\tilde{\varepsilon}^\perp = & \; T^{-1}\omega' Z^{\perp\prime}Z^\perp\omega + T^{-1}(\hat{\beta}(k_T) - \beta)' V^{\perp\prime}P_{Z^\perp}V^\perp(\hat{\beta}(k_T) - \beta) \\
& - 2T^{-1}\omega' Z^{\perp\prime}V^\perp(\hat{\beta}(k_T) - \beta) + o_p(1) \\
\xrightarrow{d} & \; \omega'\Omega\omega + \omega\Omega^{1/2\prime} z_V(z_V' z_V - \kappa I)^{-1} z_V' z_V (z_V' z_V - \kappa I)^{-1} z_V' \Omega^{1/2}\omega \\
& - 2\omega\Omega^{1/2\prime} z_V(z_V' z_V - \kappa I)^{-1} z_V' \Omega^{1/2}\omega
\end{aligned}
$$

and

$$
\begin{aligned}
T^{-2}\tilde{\varepsilon}^{\perp\prime}M_{Z^\perp}\tilde{\varepsilon}^\perp &= T^{-2}\tilde{\varepsilon}^{\perp\prime}\tilde{\varepsilon}^\perp + o_p(1) \\
&= T^{-2}(\hat{\beta}(k_T) - \beta)' V^{\perp\prime}V^\perp(\hat{\beta}(k_T) - \beta) + o_p(1) \\
&\xrightarrow{d} \omega'\Omega^{1/2\prime} z_V' (z_V' z_V - \kappa I)^{-2} z_V' \Omega^{1/2}\omega.
\end{aligned}
$$

Therefore,

$$
\phi(\hat{\beta}(k_T)) = \frac{\tilde{\varepsilon}^{\perp\prime}P_{Z^\perp}^{\perp}\tilde{\varepsilon}^\perp}{\tilde{\varepsilon}^{\perp\prime}M_{Z^\perp}\tilde{\varepsilon}^\perp / T} \xrightarrow{d} \frac{\omega'\Omega^{1/2\prime}\left(I - z_V(z_V' z_V - \kappa I)^{-1} z_V'\right)^2 \Omega^{1/2}\omega}{\omega'\Omega^{1/2\prime} z_V' (z_V' z_V - \kappa I)^{-2} z_V' \Omega^{1/2}\omega}.
$$

For part (iii), note that Assumption 2 implies that $T^{-1}V^{\perp\prime}V^\perp \xrightarrow{p} \Sigma_{VV}$, $T^{-1}V^{\perp\prime}u^\perp \xrightarrow{p} \Sigma_{Vu}$, $T^{-1}u^{\perp\prime}u^\perp \xrightarrow{p} \sigma_{uu}$, $T^{-1}Z^{\perp\prime}Z^\perp \xrightarrow{p} \Omega$, $T^{-1}Z^{\perp\prime}Y^\perp \xrightarrow{p} \Omega\Pi$, $T^{-1}Z^{\perp\prime}y^\perp \xrightarrow{p} \Omega\omega$, and $\tilde{\Sigma}_{VV} \xrightarrow{p} \Pi'\Omega\Pi + \Sigma_{VV}$. Hence, $\lambda_{\min}(G_T^0/T) \xrightarrow{p} \lambda_{\min}((\Pi'\Omega\Pi + \Sigma_{VV})^{-1}\Pi'\Omega\Pi)$. Recall that $\hat{\beta}(k_T) = (Y^{\perp\prime}(I_T - k_T M_{Z^\perp})Y^\perp)^{-1}Y^{\perp\prime}(I_T - k_T M_{Z^\perp})y^\perp$, where $k_T = 1$ for the TSLS estimator; $k_T = \hat{k}_T$ for the LIML estimator; and $k_T = \hat{k}_T - 1/(T - K_1 - K_2)$ for the Fuller-k estimator with $\hat{k}_T$ being the smallest root of $|\bar{Y}' M_X \bar{Y} - \hat{k}_T \bar{Y}' M_{\bar{Z}} \bar{Y}| = 0$. For any T, the roots of $|\bar{Y}' M_X \bar{Y} - \hat{k}_T \bar{Y}' M_{\bar{Z}} \bar{Y}| = 0$ are the same as those of $|T^{-1}\bar{Y}' M_X \bar{Y} - \hat{k}_T T^{-1}\bar{Y}' M_{\bar{Z}} \bar{Y}| = 0$. Thus, $\text{plim}_{T\to\infty}\hat{k}_T = k^*$ where k^* is the smallest root of $|\text{plim}_{T\to\infty}T^{-1}\bar{Y}' M_X \bar{Y} - k^*\text{plim}_{T\to\infty}T^{-1}\bar{Y}' M_{\bar{Z}} \bar{Y}| = |\Theta + \Sigma - k^*\Sigma| = |\Theta - (k^* - 1)\Sigma| = 0$ with Θ as defined in the theorem, and $\text{plim}_{T\to\infty} k_T = k^*$ for both the LIML and Fuller-k estimators. Hence,

$$\hat{\beta}(k_T) = (Y^{\perp\prime}P_{Z^\perp}Y^\perp - (k_T-1)Y^{\perp\prime}M_{Z^\perp}Y^\perp)^{-1}(Y^{\perp\prime}P_{Z^\perp}y^\perp - (k_T-1)Y^{\perp\prime}M_{Z^\perp}y^\perp)$$
$$\xrightarrow{p} \beta + (\Pi'\Omega\Pi - (k^*-1)\Sigma_{VV})^{-1}(\Pi'\Omega\omega - (k^*-1)\Sigma_{Vu}).$$

Now, $\hat{\varepsilon}^\perp = y^\perp - Y^\perp\hat{\beta}(k_T) = Z^\perp\omega + u^\perp - Y^\perp(\hat{\beta}(k_T)-\beta)$ and

$$\frac{\phi(\hat{\beta}(k_T))}{T} = \frac{\hat{\varepsilon}^{\perp\prime}P_{Z^\perp}\hat{\varepsilon}^\perp/T}{\hat{\varepsilon}^{\perp\prime}M_{Z^\perp}\hat{\varepsilon}^\perp/T} \xrightarrow{p} \frac{\omega'\Omega\omega + b\,(k)'\,\Pi'\Omega\Pi b\,(k) - 2\omega'\Omega\Pi b\,(k)}{\sigma_{uu} + b\,(k)'\,\Sigma_{VV}b\,(k) - 2\Sigma_{uV}b\,(k)}. \quad \blacksquare$$

Proof of Theorem 3 Obvious from Theorems 1 and 3 in Staiger and Stock (1997) and thus omitted. ■

A.2 Simulation Results

Table 1.A. Critical Values for the Q_{IV} test, TSLS

	$n=1$				$n=2$				$n=3$			
	.90	.95	.975	.99	.90	.95	.975	.99	.90	.95	.975	.99
$K_2=2$	3.97	5.35	6.74	8.60								
3	4.93	6.46	7.99	10.06	1.88	2.57	3.24	4.16				
4	5.70	7.37	9.03	11.21	2.43	3.24	4.01	5.03	1.21	1.65	2.11	2.72
5	6.45	8.28	10.09	12.40	2.90	3.82	4.72	5.83	1.62	2.19	2.74	3.45
6	7.03	8.99	10.89	13.19	3.31	4.34	5.30	6.57	1.94	2.59	3.22	3.99
7	7.60	9.69	11.64	14.14	3.68	4.85	5.93	7.33	2.22	2.98	3.70	4.59
8	8.13	10.39	12.45	15.05	4.01	5.28	6.51	7.92	2.48	3.33	4.15	5.14
9	8.59	10.94	13.12	15.83	4.35	5.70	6.98	8.57	2.71	3.67	4.55	5.63
10	9.02	11.42	13.66	16.30	4.56	6.00	7.37	9.07	2.91	3.94	4.89	5.97
11	9.49	11.99	14.33	17.26	4.83	6.41	7.80	9.56	3.10	4.23	5.28	6.50
12	9.89	12.49	14.89	17.80	5.11	6.78	8.26	10.04	3.28	4.51	5.61	6.91
13	10.28	13.02	15.50	18.59	5.34	7.07	8.63	10.48	3.40	4.72	5.90	7.32
14	10.65	13.50	16.09	19.21	5.55	7.38	9.06	11.00	3.57	4.95	6.18	7.69
15	11.06	13.96	16.58	19.74	5.76	7.68	9.37	11.46	3.63	5.12	6.39	7.87
16	11.38	14.36	17.14	20.27	5.97	7.98	9.78	11.87	3.76	5.33	6.73	8.33
17	11.71	14.81	17.61	20.88	6.15	8.17	10.05	12.35	3.88	5.50	6.95	8.63
18	12.04	15.23	18.11	21.61	6.29	8.48	10.38	12.66	4.00	5.72	7.20	9.01
19	12.36	15.55	18.47	22.05	6.53	8.76	10.72	13.03	4.08	5.88	7.45	9.26
20	12.71	16.04	18.99	22.54	6.69	9.01	11.01	13.41	4.18	6.09	7.79	9.62
21	12.98	16.39	19.50	23.24	6.78	9.20	11.29	13.78	4.20	6.13	7.90	9.84
22	13.16	16.60	19.80	23.53	7.05	9.49	11.65	14.19	4.40	6.44	8.16	10.24
23	13.50	17.02	20.29	23.94	7.18	9.67	11.89	14.54	4.38	6.48	8.23	10.35
24	13.84	17.46	20.67	24.82	7.28	9.91	12.14	14.84	4.42	6.60	8.48	10.71
25	14.17	17.86	21.08	25.11	7.41	10.04	12.37	15.17	4.62	6.82	8.74	11.02

Note: n is the number of endogenous regressors and K_2 is the number of instruments.

Table 1.B. Critical Values for the Q_{IV} test, LIML

	$n=1$				$n=2$				$n=3$			
	.90	.95	.975	.99	.90	.95	.975	.99	.90	.95	.975	.99
$K_2=2$	3.99	5.37	6.76	8.60								
3	4.96	6.49	8.01	10.08	1.90	2.58	3.25	4.17				
4	5.76	7.42	9.05	11.23	2.48	3.28	4.05	5.05	1.23	1.67	2.13	2.73
5	6.52	8.33	10.12	12.42	2.99	3.89	4.76	5.87	1.68	2.23	2.78	3.47
6	7.13	9.06	10.94	13.24	3.43	4.44	5.37	6.64	2.03	2.66	3.28	4.04
7	7.73	9.78	11.71	14.20	3.87	4.98	6.05	7.40	2.37	3.10	3.79	4.68
8	8.28	10.49	12.53	15.11	4.25	5.45	6.64	8.04	2.70	3.49	4.28	5.23
9	8.77	11.07	13.22	15.89	4.64	5.90	7.13	8.70	3.00	3.88	4.71	5.75
10	9.25	11.57	13.75	16.40	4.91	6.28	7.57	9.21	3.28	4.21	5.09	6.11
11	9.73	12.16	14.46	17.34	5.26	6.72	8.02	9.74	3.55	4.57	5.54	6.71
12	10.17	12.66	15.03	17.91	5.61	7.13	8.52	10.25	3.82	4.88	5.90	7.15
13	10.57	13.23	15.68	18.70	5.92	7.47	8.94	10.73	4.04	5.19	6.25	7.56
14	10.97	13.71	16.26	19.35	6.17	7.84	9.41	11.27	4.29	5.48	6.59	8.02
15	11.41	14.21	16.78	19.88	6.47	8.19	9.76	11.76	4.47	5.71	6.87	8.25
16	11.77	14.63	17.34	20.42	6.75	8.53	10.20	12.18	4.70	6.03	7.28	8.74
17	12.14	15.10	17.85	21.06	6.99	8.80	10.51	12.68	4.92	6.28	7.54	9.09
18	12.50	15.56	18.35	21.79	7.24	9.12	10.93	13.05	5.12	6.55	7.88	9.53
19	12.85	15.92	18.74	22.27	7.49	9.48	11.28	13.46	5.31	6.80	8.17	9.79
20	13.23	16.37	19.25	22.81	7.73	9.75	11.63	13.89	5.53	7.12	8.56	10.23
21	13.52	16.77	19.76	23.47	7.93	10.03	11.95	14.27	5.66	7.26	8.76	10.53
22	13.75	17.04	20.12	23.77	8.26	10.39	12.36	14.75	5.93	7.58	9.11	10.96
23	14.10	17.43	20.59	24.17	8.44	10.60	12.63	15.08	6.06	7.73	9.27	11.11
24	14.47	17.92	21.00	25.07	8.63	10.88	12.96	15.45	6.23	7.95	9.55	11.50
25	14.80	18.31	21.50	25.36	8.83	11.11	13.19	15.85	6.49	8.25	9.88	11.87

Note: n is the number of endogenous regressors and K_2 is the number of instruments.

Table 1.C. Critical Values for the Q_{IV} test, Fuller-k

	$n=1$				$n=2$				$n=3$			
	.90	.95	.975	.99	.90	.95	.975	.99	.90	.95	.975	.99
$K_2=2$	3.96	5.35	6.74	8.60								
3	4.94	6.47	8.00	10.07	1.80	2.52	3.20	4.12				
4	5.73	7.40	9.04	11.23	2.40	3.22	4.00	5.01	1.06	1.55	2.03	2.65
5	6.50	8.32	10.12	12.41	2.92	3.84	4.73	5.84	1.54	2.13	2.70	3.42
6	7.11	9.05	10.93	13.24	3.37	4.39	5.34	6.61	1.91	2.57	3.21	3.99
7	7.71	9.77	11.70	14.19	3.81	4.94	6.02	7.38	2.25	3.02	3.73	4.63
8	8.26	10.48	12.52	15.10	4.19	5.42	6.61	8.02	2.60	3.42	4.22	5.18
9	8.76	11.06	13.21	15.89	4.59	5.87	7.11	8.68	2.91	3.81	4.66	5.72
10	9.23	11.56	13.75	16.39	4.87	6.25	7.54	9.20	3.19	4.15	5.04	6.08
11	9.72	12.15	14.45	17.34	5.21	6.69	8.00	9.72	3.47	4.51	5.49	6.67
12	10.15	12.65	15.02	17.91	5.57	7.10	8.50	10.23	3.73	4.83	5.86	7.12
13	10.55	13.22	15.67	18.69	5.88	7.44	8.92	10.72	3.97	5.13	6.21	7.53
14	10.96	13.71	16.25	19.35	6.12	7.81	9.40	11.26	4.22	5.43	6.56	7.99
15	11.40	14.20	16.77	19.88	6.43	8.16	9.74	11.74	4.41	5.66	6.83	8.22
16	11.76	14.63	17.34	20.41	6.71	8.51	10.18	12.16	4.64	5.99	7.26	8.72
17	12.13	15.09	17.84	21.05	6.95	8.77	10.49	12.67	4.86	6.25	7.51	9.07
18	12.49	15.55	18.35	21.79	7.21	9.10	10.91	13.04	5.06	6.51	7.85	9.51
19	12.84	15.91	18.73	22.27	7.46	9.45	11.26	13.45	5.24	6.75	8.14	9.76
20	13.22	16.37	19.24	22.81	7.70	9.73	11.61	13.87	5.47	7.08	8.53	10.20
21	13.51	16.76	19.76	23.47	7.89	10.01	11.93	14.25	5.60	7.22	8.73	10.50
22	13.74	17.03	20.12	23.77	8.23	10.37	12.35	14.75	5.88	7.54	9.08	10.94
23	14.09	17.43	20.59	24.17	8.41	10.58	12.61	15.07	6.01	7.69	9.23	11.09
24	14.46	17.92	21.00	25.07	8.60	10.87	12.94	15.44	6.18	7.92	9.53	11.48
25	14.79	18.31	21.49	25.36	8.80	11.09	13.18	15.84	6.44	8.22	9.85	11.84

Note: n is the number of endogenous regressors and K_2 is the number of instruments.

Table 2.A. Rejection Probabilities of the Overidentifying Restrictions Test, TSLS

$n = 1$		$K_2 = 3$		$K_2 = 9$	
$\Lambda_C'\Lambda_C/\lambda^*_{\min}$	ρ	rej.	uncond.	rej.	uncond.
0.5	0.1	0.0550	0.0365	0	0.0450
	0.3	0.0661	0.0411	0	0.0506
	0.5	0.0752	0.0531	0.1667	0.0660
	0.8	0.1229	0.1143	0.1667	0.1179
	1.0	0.3321	0.2016	0.3333	0.1880
$\Lambda_C'\Lambda_C/\lambda^*_{\min}$	ρ	rej.	uncond.	rej.	uncond.
0.8	0.1	0.0419	0.0395	0.0442	0.0465
	0.3	0.0498	0.0438	0.0549	0.0508
	0.5	0.0613	0.0547	0.0869	0.0619
	0.8	0.0972	0.0985	0.1631	0.0948
	1.0	0.2328	0.1494	0.3125	0.1347
$\Lambda_C'\Lambda_C/\lambda^*_{\min}$	ρ	rej.	uncond.	rej.	uncond.
1.2	0.1	0.0473	0.0419	0.0486	0.0475
	0.3	0.0516	0.0455	0.0542	0.0508
	0.5	0.0600	0.0542	0.0663	0.0587
	0.8	0.0996	0.0841	0.1046	0.0810
	1.0	0.1808	0.1161	0.1609	0.1058

Note: n is the number of endogenous regressors; K_2 is the number of instruments; ρ represents the degree of endogeneity; $\Lambda_C'\Lambda_C$ corresponds to the limit of the concentration matrix; and $\lambda^*_{\min}$ is the boundary value of the minimum eigenvalue for the weak instruments set based on the 10% TSLS bias. The column "rej." shows $\mathbb{P}(\phi_\infty > \chi^2_{K_2-n,0.05} | g_\infty \text{ rejects})$ and "uncond." $\mathbb{P}(\phi_\infty > \chi^2_{K_2-n,0.05})$.

Table 2.B. Rejection Probabilities of the Overidentifying Restrictions Test, Fuller-k

$n=1$		$K_2=3$		$K_2=9$	
$\Lambda_C'\Lambda_C/\lambda^*_{\min}$	ρ	rej.	uncond.	rej.	uncond.
0.5	0.1	0.1027	0.1739	0.1887	0.2823
	0.3	0.1655	0.1815	0.3032	0.2888
	0.5	0.2240	0.1789	0.4218	0.2779
	0.8	0.2924	0.1449	0.5782	0.1907
	1.0	0.2753	0.1074	0.5040	0.0780
$\Lambda_C'\Lambda_C/\lambda^*_{\min}$	ρ	rej.	uncond.	rej.	uncond.
0.8	0.1	0.0906	0.1440	0.1667	0.2420
	0.3	0.1309	0.1402	0.2427	0.2319
	0.5	0.1666	0.1277	0.3304	0.2060
	0.8	0.2115	0.0900	0.4147	0.1147
	1.0	0.1939	0.0702	0.3380	0.0578
$\Lambda_C'\Lambda_C/\lambda^*_{\min}$	ρ	rej.	uncond.	rej.	uncond.
1.2	0.1	0.0933	0.1192	0.1520	0.2023
	0.3	0.1183	0.1091	0.2044	0.1824
	0.5	0.1420	0.0945	0.2495	0.1505
	0.8	0.1572	0.0665	0.2670	0.0796
	1.0	0.1466	0.0560	0.2172	0.0520

Note: n is the number of endogenous regressors; K_2 is the number of instruments; ρ represents the degree of endogeneity; $\Lambda_C'\Lambda_C$ corresponds to the limit of the concentration matrix; and $\lambda^*_{\min}$ is the boundary value of the minimum eigenvalue for the weak instruments set based on the 10% TSLS bias. The column "rej." shows $\mathbb{P}(\phi_\infty > \chi^2_{K_2-n,0.05}|g_\infty \text{ rejects})$ and "uncond." $\mathbb{P}(\phi_\infty > \chi^2_{K_2-n,0.05})$.

Table 3.A. Finite Sample Sizes of the Q_{IV} test, TSLS

	$n = 1$			$n = 2$			$n = 3$		
K_2	2	4	6	3	5	7	4	6	8
$T = 50$	.0428	.0374	.0324	.0504	.0560	.0478	.0560	.0552	.0476
100	.0524	.0526	.0402	.0518	.0494	.0478	.0550	.0516	.0500
200	.0442	.0498	.0458	.0542	.0544	.0538	.0560	.0504	.0530
300	.0458	.0488	.0458	.0500	.0432	.0474	.0568	.0520	.0514

Table 3.B. Finite Sample Sizes of the Q_{IV} test, Fuller-k

	$n = 1$			$n = 2$			$n = 3$		
K_2	2	4	6	3	5	7	4	6	8
$T = 50$	.0430	.0374	.0324	.0492	.0554	.0478	.0532	.0536	.0482
100	.0522	.0532	.0396	.0510	.0488	.0468	.0552	.0526	.0500
200	.0438	.0502	.0468	.0528	.0552	.0528	.0550	.0502	.0530
300	.0460	.0488	.0454	.0490	.0436	.0476	.0578	.0516	.0516

Table 4.A. Finite Sample Powers of the Q_{IV} test, $\omega = 0$ and $\Pi \neq 0$, TSLS

	$n = 1$			$n = 2$			$n = 3$		
K_2	2	4	6	3	5	7	4	6	8
$T = 50$	.7102	.3568	.1642	.4856	.2152	.2780	.3790	.1978	.1758
100	.9624	.4042	.4724	.7594	.5534	.4808	.6376	.5252	.4130
200	.8978	.8006	.8326	.9474	.8888	.8040	.9172	.8116	.7262
300	1.000	.9690	.9966	.9872	.9602	.9472	.9838	.9426	.8948

Note: T is the sample size, n is the number of endogenous regressors and K_2 is the number of instruments.

Table 4.B. Finite Sample Powers of the Q_{IV} test, $\omega = 0$ and $\Pi \neq 0$, Fuller-k

	$n=1$			$n=2$			$n=3$		
K_2	2	4	6	3	5	7	4	6	8
$T=50$	.7112	.3602	.1668	.4884	.2200	.2904	.3798	.1998	.1850
100	.9642	.4106	.4794	.7626	.5662	.5018	.6434	.5416	.4350
200	.8996	.8104	.8450	.9502	.8996	.8202	.9238	.8320	.7628
300	1.000	.9732	.9980	.9890	.9676	.9532	.9860	.9586	.9156

Table 5.A. Finite Sample Powers of the Q_{IV} test, $\omega \neq 0$ and $\Pi = 0$, TSLS

	$n=1$			$n=2$			$n=3$		
K_2	2	4	6	3	5	7	4	6	8
$T=50$	.0186	.0070	.0036	.0350	.0178	.0102	.0450	.0270	.0152
100	.0186	.0092	.0020	.0274	.0126	.0044	.0384	.0184	.0118
200	.0174	.0044	.0036	.0294	.0134	.0038	.0362	.0180	.0076
300	.0192	.0040	.0016	.0304	.0114	.0054	.0404	.0140	.0110
∞	.0168	.0034	.0000	.0290	.0099	.0037	.0344	.0118	.0058

Table 5.B. Finite Sample Powers of the Q_{IV} test, $\omega \neq 0$ and $\Pi = 0$, Fuller-k

	$n=1$			$n=2$			$n=3$		
K_2	2	4	6	3	5	7	4	6	8
$T=50$	.0186	.0080	.0046	.0310	.0192	.0124	.0352	.0278	.0178
100	.0198	.0114	.0032	.0254	.0132	.0080	.0310	.0178	.0130
200	.0184	.0060	.0048	.0244	.0154	.0072	.0258	.0200	.0090
300	.0194	.0060	.0032	.0266	.0122	.0076	.0296	.0152	.0150

Note: T is the sample size, n is the number of endogenous regressors and K_2 is the number of instruments.

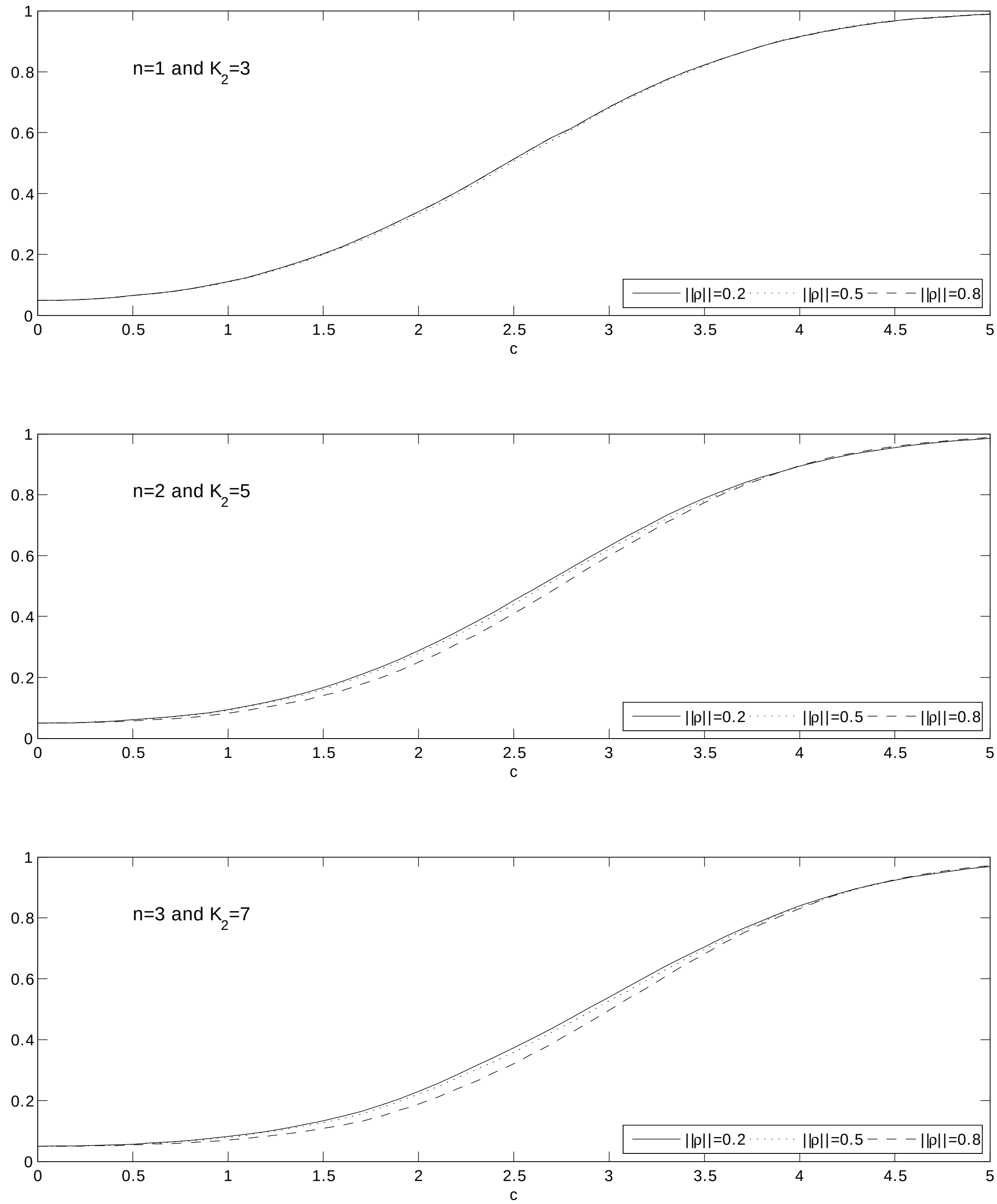

Figure 1.A. Local Power of the Q_{IV} test, $\Lambda_C \neq 0$ and $\xi_d = 0$.

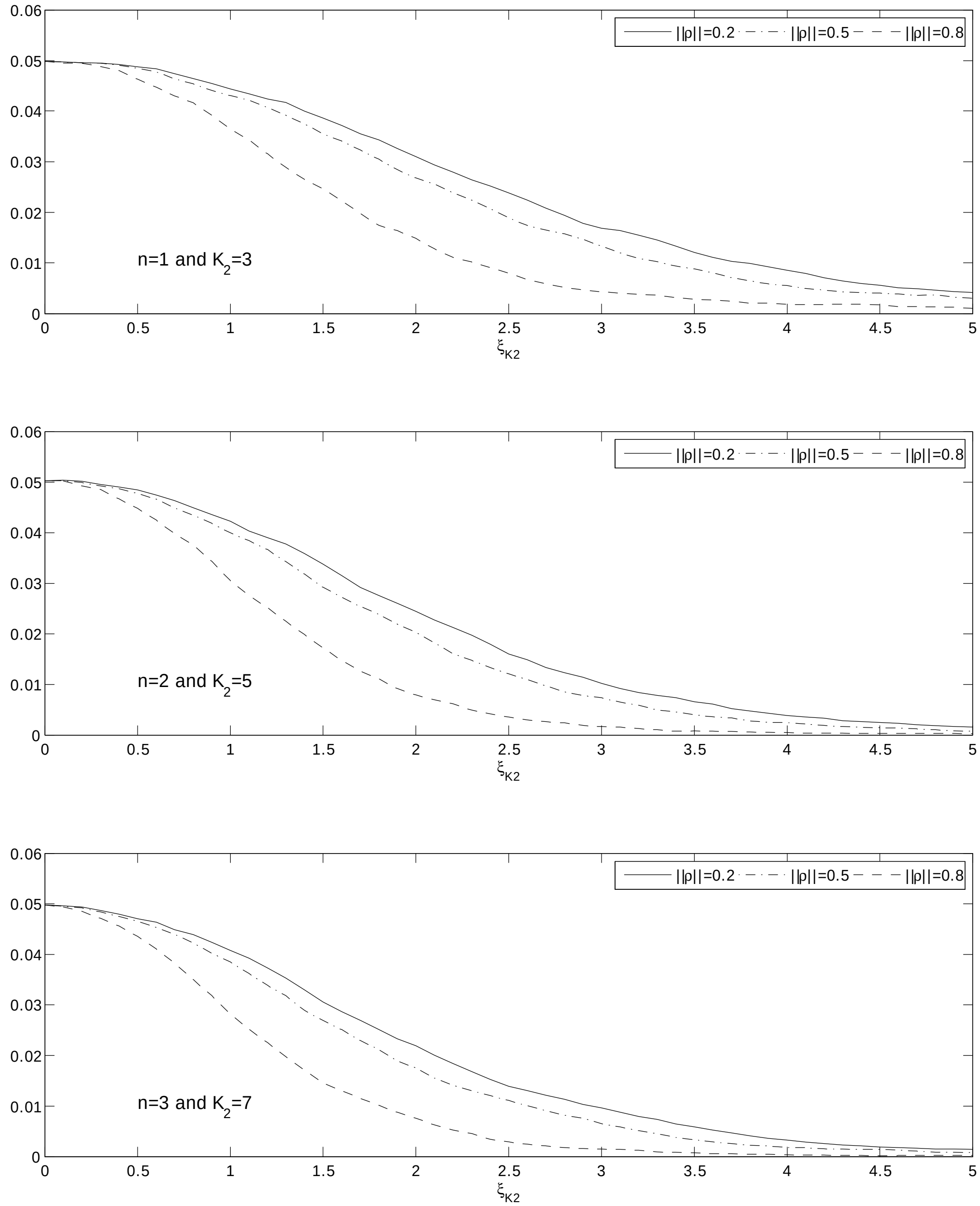

Figure 1.B. Local Power of the Q_{IV} test, $\Lambda_C = 0$ and $\xi_d \neq 0$.

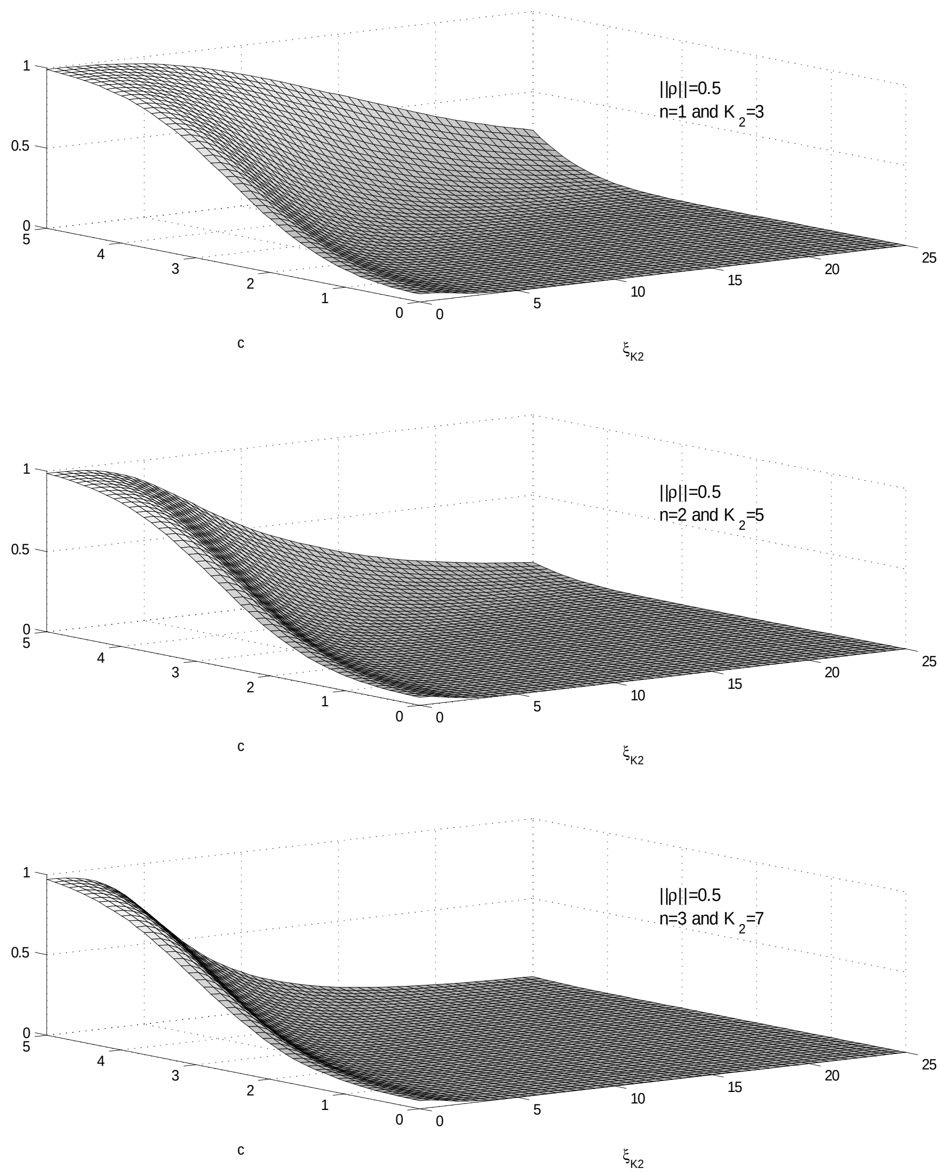

Figure 1.C. Local Power of the Q_{IV} test, $\Lambda_C \neq 0$ and $\xi_d \neq 0$.

References

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

[2] Basmann, R. L. (1960). On Finite Sample Distributions of Generalized Classical Linear Identifiability Test Statistics, *Journal of the American Statistical Association*, 55, 650-659.

[3] Cragg, J. G. and S. G. Donald (1993). Testing Identifiability and Specification in Instrumental Variable Models, *Econometric Theory*, 9, 222-240.

[4] Doko, F. and J-M. Dufour (2008). Instrumental Endogeneity and Identification-robust Tests: Some Analytical Results, *Journal of Statistical Planning and Inference*, 138, 2649-2661.

[5] Hall, A. R., G. D. Rudebusch and D. W. Wilcox (1996). Judging Instrument Relevance in Instrumental Variables Estimation, *International Economic Review*, 37, 283-298.

[6] Lee, Y. and R. Okui (2009). Specification Test for Instrumental Variables Regression with Many Instruments, *mimeo.*

[7] Moreira, M. J. (2003). A Conditional Likelihood Ratio Test for Structural Models, *Econometrica*, 71, 1027-1048.

[8] Newey, W. K. (1985). Maximum likelihood specification testing and conditional moment tests, *Econometrica*, 53, 1047–1070.

[9] Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables, *Econometrica*, 26, 393–415.

[10] Staiger, D. and J. H. Stock (1997). Instrumental Variables Regression with Weak Instruments, *Econometrica*, 65, 557-586.

[11] Stock, J. H. and M. Yogo (2005). Testing for Weak Instruments in Linear IV Regression, in Identification and Inference for Econometric Models, *Essays in Honor of Thomas Rothenberg*, ed. by D. W. K. Andrews and J. H. Stock, 80-107.

[12] Vuong, Q. H. (1989). Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses, *Econometrica*, 57, 307-333