On the Power Curves of the Conditional Likelihood Ratio and Related Tests for Instrumental Variables Regression with Weak Instruments*

Nicolas Van de Sijpe[†] Frank Windmeijer[‡]

August 4, 2020

Abstract

We show that the Likelihood Ratio (LR) statistic for testing the value of the coefficient β in a linear instrumental variables model with a single endogenous variable is identical to the $t_0(\widehat{\beta}_L)^2$ statistic as proposed by Mills, Moreira, and Vilela (2014), where $\widehat{\beta}_L$ is the LIML estimator. This implies the equivalence of their conditional versions that are robust to weak instruments. From this result, properties of the power of the Conditional LR (CLR) test can be understood; in particular the asymmetric nature of the power curve as a function of the true value of β when testing H_0: $\beta = \beta_0$ for fixed β_0, when the instruments are weak and the variance matrix of the structural and first-stage errors is held constant. Power curves of the CLR and related tests have often been presented for a design where instead the variance matrix of the reduced-form and first-stage errors has been held constant. This latter design changes the endogeneity features at each value of β and results in a power curve that is close to the points with maximum power in the design with fixed variance of the structural and first-stage errors. As the results for the design with fixed variance of the structural and first-stage errors are informative for the behaviour of the test-based confidence intervals, it seems more natural to consider this design. We find that LIML- and Fuller-based conditional Wald and conditional $t_0(\widehat{\beta}_{Full})^2$ tests, which are not unbiased tests, are more powerful than the CLR test when the degree of endogeneity is low to moderate.

JEL Classification: C12, C26

Keywords: Weak Instruments, Robust Inference, Conditional Likelihood Ratio Test, Power

---

*We thank Marcelo Moreira and Yanos Zylberberg for very helpful comments and suggestions.

[†]Dept. of Economics, University of Sheffield, UK, n.vandesijpe@sheffield.ac.uk.

[‡]Dept. of Statistics and Nuffield College, University of Oxford, UK, frank.windmeijer@stats.ox.ac.uk, corresponding author

1 Introduction

For the linear instrumental variables (IV) model with one endogenous explanatory variable, $y_i = x_i\beta + u_i$, the conditional Likelihood Ratio (CLR) test of Moreira (2003) and related tests, like the AR (Anderson and Rubin, 1949), LM, and conditional Wald (CW) tests are tests for the hypothesis H$_0$: $\beta = \beta_0$. They are robust to weak instruments in the sense that they have correct size when instruments are weak, with CLR, AR, and LM unbiased, similar tests, whereas the CW tests are not unbiased.

For the evaluation of the power of these tests, two different designs have been used in the literature. Let the first-stage model be given by $x_i = z_i'\pi + v_i$. Then one design keeps the variance of the structural and first-stage errors,

$$\Sigma = Var\left((u_i\, v_i)'\right) = \begin{bmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{bmatrix},$$

fixed when varying the value of β, whereas the other design keeps the variance of the reduced form errors, $\Omega = Var\left((v_{yi}\, v_i)'\right)$, fixed, where the reduced form for y_i is given by $y_i = z_i'\pi\beta + u_i + \beta v_i = z_i'\pi_y + v_{yi}$. The simulations in Kleibergen (2002), Moreira (2003, 2009), and Stock, Wright, and Yogo (2002), amongst others, are based on the fixed Σ design, whereas Andrews, Moreira, and Stock (2006, 2007), Mills, Moreira, and Vilela (2014) and Moreira and Moreira (2019) are examples of simulations based on the fixed Ω design.

Poskitt and Skeels (2008) discuss these two designs and show that simulation results can differ substantially between them, but do not provide an explanation for these differences. Davidson and MacKinnon (2008) highlight that a design with Ω fixed changes Σ when changing the value of β and conclude that Ω is “not a sensible quantity to keep fixed” (Davidson and MacKinnon 2008, p 455). Andrews, Marmer, and Yu (2019) make the same observation and propose a design where the value of β is fixed, but the value of β_0 in H$_0$: $\beta = \beta_0$ is varied instead. Their motivation for this design is its direct link to the formation of confidence intervals based on inverting test statistics. The main focus of the analysis in Andrews et al. (2019) is on the probability of obtaining infinite length confidence intervals, and on the difference of the CLR power curve from the two-sided power envelope for extreme values of β_0. However, this design where β is kept fixed but β_0 is varied is essentially the same as the fixed Σ design, with the test statistics for

H_0: $\beta = \beta_0$ when $\beta = \beta^*$ in the fixed Σ design then identical to the test statistics for H_0: $\beta = -\beta^*$ when $\beta = -\beta_0$ in the Andrews et al. (2019) design when varying the value for β^*. Therefore, the power functions of the design with β fixed and varying β_0 are the mirror images around 0 of those of the fixed Σ design.

In this paper, we make the following contributions. We directly compare and explain the behaviour of the power curves of the CLR test in the two designs. We are better able to explain the behaviour of the CLR test in the fixed Σ design by first showing, in Proposition 1 in Section 2, that the LR statistic is the same as the $t_0(\widehat{\beta}_L)^2 \equiv W_0(\widehat{\beta}_L)$ statistic proposed by Mills et al. (2014), where $\widehat{\beta}_L$ is the LIML estimator of β. The only difference between $W_0(\widehat{\beta}_L)$ and the standard LIML-based Wald test is the estimator for σ_u^2 in the denominator of the test. For $W_0(\widehat{\beta}_L)$ this variance is estimated under the null and denoted $\widehat{\sigma}_0^2$. Under weak instruments, the power of the CLR test is then boosted by low values of $\widehat{\sigma}_0^2$ and can reach one even with very weak instruments when $\rho_{uv} = \sigma_{uv}/\sigma_u\sigma_v$ approaches 1 or -1, and for certain values of β, such that $\widehat{\sigma}_0^2 \to 0$. We can thus explain the asymmetry of the power of the CLR test under weak instruments, as documented before by e.g. Stock et al. (2002) and Davidson and MacKinnon (2008).

The fixed Σ design keeps the structural endogeneity features constant, in particular the degree of endogeneity ρ_{uv}. As mentioned above, this is not the case for the fixed Ω design, where σ_v^2 is kept fixed, but both $\sigma_u^2(\beta)$ and $\sigma_{uv}(\beta)$ are changing with the true value of β in the DGP in a very specific way, as shown in Figure 1 in Section 3. In particular, the correlation $\rho_{uv}(\beta)$ approaches 1 for large negative values of β, and approaches -1 for large positive values of β. We therefore compare the fixed Ω power curve with various fixed $\Sigma(\beta^*)$ power curves for different values of β^*, and where $\sigma_u^2 = \sigma_u^2(\beta^*)$ and $\sigma_{uv} = \sigma_{uv}(\beta^*)$, matching the values to those of the fixed Ω design at β^*. We find that, in weak instruments setups, the power curve for the fixed Ω design crosses the various paths of the fixed $\Sigma(\beta^*)$ designs at points close to maximum power, as displayed in Figure 3 in Section 3. The fixed Ω design therefore exploits the asymmetry of the power curve of the CLR test in the fixed Σ design. This explains why, for a large part of the parameter space for β, the fixed Ω design shows much higher power of the CLR test than the fixed Σ design under weak instruments.

As a final contribution, we compare the behaviours of the LIML- and Fuller-based CW tests and the Fuller-CW_0 test to that of the CLR test in the fixed Σ design for different degrees of endogeneity ρ_{uv}. As far as we are aware these tests have only been compared in the fixed Ω design, see Andrews et al. (2007) and Mills et al. (2014). We find that for low to moderate degrees of endogeneity, these conditional tests are more powerful than the CLR test. Even for medium to high values of ρ_{uv} the Fuller-CW_0 test is well behaved, with higher power than the CLR test for part of the parameter space.

As a final remark, it is sometimes argued to keep Ω fixed as it is known, i.e. can be estimated consistently, unlike Σ under weak instruments, see e.g. the discussion in Andrews et al. (2019). In all our simulations we follow the asymptotic design of Andrews et al. (2006) with Ω known, also for the fixed Σ design, where $\Omega(\beta)$ is known but changes with the value of β. Given the link of the fixed Σ power curve with the behaviour of the test-statistics-based confidence intervals, it seems more natural to consider this design.

2 Model and Tests

The linear IV model specification for a sample $\{y_i, x_i, z_i'\}_{i=1}^n$ is given by

$$\begin{aligned} y_i &= x_i\beta + u_i \\ x_i &= z_i'\pi + v_i, \end{aligned} \tag{1}$$

where z_i is the k_z vector of instrumental variables. The instruments satisfy $E(z_iu_i) = 0$. Standard assumptions on the data, see e.g. Assumption M in Stock and Yogo (2005), needed for limiting normal distributions and consistent estimation of variance matrices are assumed to hold. The explanatory variable x_i is endogenous as $E(x_iu_i) = E(u_iv_i) \neq 0$. Other exogenous explanatory variables, including the constant, have been partialled out. The errors are assumed to be conditionally homoskedastic, with

$$Var\left((u_i\ v_i)'\,|z_i\right) = \begin{bmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{bmatrix} \equiv \Sigma, \tag{2}$$

and correlation $\rho_{uv} = \sigma_{uv}/\sigma_u\sigma_v$.

The reduced form for y_i is given by

$$\begin{aligned} y_i &= z_i'\pi\beta + u_i + v_i\beta \\ &= z_i'\pi_y + v_{yi}, \end{aligned} \tag{3}$$

and the reduced form error variance of $(v_{yi} \;\; v_i)'$ is given by

$$\Omega(\beta) = \begin{bmatrix} \sigma_u^2 + 2\beta\sigma_{uv} + \beta^2\sigma_v^2 & \sigma_{uv} + \beta\sigma_v^2 \\ \sigma_{uv} + \beta\sigma_v^2 & \sigma_v^2 \end{bmatrix}. \tag{4}$$

Let $b = (1 \;\; -\beta)'$, and $a = (\beta \;\; 1)'$, then it follows that $b'\Omega(\beta)\,b = \sigma_u^2 = |\Omega(\beta)|\, a'(\Omega(\beta))^{-1} a$, and $|\Omega(\beta)| = \sigma_u^2\sigma_v^2(1-\rho_{uv}^2)$.

Let y and x be the n-vectors (y_i) and (x_i) and Z the $n \times k_z$ matrix of instruments. The standard 2SLS estimator for β is given by

$$\widehat{\beta}_{2sls} = \frac{x'P_Z y}{x'P_Z x},$$

where $P_Z = Z(Z'Z)^{-1}Z'$. The 2SLS estimator is based on the OLS estimator for π, given by $\widehat{\pi} = (Z'Z)^{-1}Z'x$. Let $\widehat{x} = Z\widehat{\pi}$, then $\widehat{\beta}_{2sls} = \widehat{x}'y/(\widehat{x}'x) = \widehat{x}'y/(\widehat{x}'\widehat{x})$.

Dropping notationally the dependence of Ω on β for ease of exposition, an estimator for

$$\Omega = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{12} & \omega_{22} \end{bmatrix}$$

is $\widehat{\Omega} = W'M_Z W/n$, where $W = [y \;\; x]$ and $M_Z = I_n - P_Z$. The LIML estimator for β is then given by

$$\widehat{\beta}_L = \frac{x'P_Z y - n\widehat{\kappa}\widehat{\omega}_{12}}{x'P_Z x - n\widehat{\kappa}\widehat{\omega}_{22}},$$

where $\widehat{\kappa}$ is the minimum eigenvalue,

$$\widehat{\kappa} = \min \operatorname{eval}\left(\left(n^{-1}W'P_Z W\right)\widehat{\Omega}^{-1}\right).$$

Let $a_L = (\widehat{\beta}_L \;\; 1)'$. The definition of the LIML estimator for π as used in Moreira (2003) is given by

$$\widehat{\pi}_L = (Z'Z)^{-1}Z'W\widehat{\Omega}^{-1}a_L\left(a_L'\widehat{\Omega}^{-1}a_L\right)^{-1}. \tag{5}$$

Alternative expressions for $\widehat{\pi}_L$ are,

$$\begin{aligned} \widehat{\pi}_L &= (Z'M_{\widehat{u}_L}Z)^{-1}Z'M_{\widehat{u}_L}x & (6) \\ &= \widehat{\pi} - \frac{(Z'Z)^{-1}Z'\widehat{u}_L\left(\widehat{\omega}_{12} - \widehat{\beta}_L\widehat{\omega}_{22}\right)}{\widehat{\sigma}_L^2}, & (7) \end{aligned}$$

where $\widehat{u}_L = y - x\widehat{\beta}_L$ and $\widehat{\sigma}_L^2 = \widehat{u}_L' M_Z \widehat{u}_L / n = b_L' \widehat{\Omega} b_L$, with $b_L = (1 \ \ -\widehat{\beta}_L)'$. Expression (6) is the standard expression as given in e.g. Bowden and Turkington (1984, p 108), from which (7) can be derived, see also Hausman (1983, p 424).

Let $\widehat{x}_L = Z\widehat{\pi}_L$, then $\widehat{\beta}_L = \widehat{x}_L' y / (\widehat{x}_L' x) = \widehat{x}_L' y / (\widehat{x}_L' \widehat{x}_L)$, see Windmeijer (2018) for the latter equality. It follows that the score for the LIML estimator of β is given by $\widehat{x}_L' \left(y - x\widehat{\beta}_L \right) = 0$.

Consider testing the null H_0: $\beta = \beta_0$ against the two-sided alternative H_1: $\beta \neq \beta_0$. The distributional properties of the AR, LM, and LR tests as described below are exact under fixed instruments, known Ω and normally distributed errors. Instrument strength is then determined by $\lambda = \pi' Z' Z \pi / \sigma_v^2$. The limiting distributions of the tests are the same when relaxing these assumptions and using $\widehat{\Omega}$ as an estimator for Ω, see Moreira (2003) and Kleibergen (2002). Weak instrument asymptotics then imply $\pi = \pi_n = c/\sqrt{n}$, where c is a vector of constants, with instrument strength then determined by $\lambda = \mathrm{plim}\left(\pi_n' Z' Z \pi_n / \sigma_v^2 \right) = c' Q_{zz} c / \sigma_v^2$, where $Q_{zz} = \mathrm{plim}\left(Z'Z/n \right)$.

Let $u_0 = y - x\beta_0$. The Anderson-Rubin test statistic is given by

$$AR = \frac{u_0' P_Z u_0}{\widehat{\sigma}_0^2},$$

where $\widehat{\sigma}_0^2 = b_0' \widehat{\Omega} b_0 = u_0' M_Z u_0 / n$, with $b_0 = (1 \ \ -\beta_0)'$. AR has a limiting $\chi^2_{k_z}$ distribution under the null, independent of the strength of the instruments. AR is a test for overidentifying restrictions in model (1), imposing the null.

Let $\widehat{\pi}_{L0}$ be the LIML estimator of π under the null, given by

$$\widehat{\pi}_{L0} = \left(Z'Z \right)^{-1} Z' W \widehat{\Omega}^{-1} a_0 / \left(a_0' \widehat{\Omega}^{-1} a_0 \right),$$

with $a_0 = (\beta_0 \ \ 1)'$. Then the score of the LIML estimator, evaluated under the null is given by $\widehat{x}_{L0}' \left(y - x\beta_0 \right) = \widehat{x}_{L0}' u_0$, where $\widehat{x}_{L0} = Z\widehat{\pi}_{L0}$, and the LM test statistic is given by

$$LM = \frac{u_0' \widehat{x}_{L0} \left(\widehat{x}_{L0}' \widehat{x}_{L0} \right)^{-1} \widehat{x}_{L0}' u_0}{\widehat{\sigma}_0^2}. \tag{8}$$

Under the null, LM has a limiting χ_1^2 distribution, again independent of the strength of the instruments.

An interesting link between AR and the estimators $\widehat{\pi}_{L0}$ and $\widehat{\pi}$ is that the AR statistic is equal to the Hausman test statistic,

$$H = (\widehat{\pi} - \widehat{\pi}_{L0})' \left(V\widehat{a}r(\widehat{\pi}) - V\widehat{a}r(\widehat{\pi}_{L0}) \right)^{-1} (\widehat{\pi} - \widehat{\pi}_{L0}) = AR. \tag{9}$$

This follows as $\widehat{\pi}_{L0}$ can alternatively be expressed as

$$\widehat{\pi}_{L0} = \widehat{\pi} - \frac{(Z'Z)^{-1} Z'u_0 (\widehat{\omega}_{12} - \beta_0 \widehat{\omega}_{22})}{\widehat{\sigma}_0^2},$$

following (7), linking the definitions of Moreira (2003) and Kleibergen (2002). Further,

$$V\widehat{a}r(\widehat{\pi}) - V\widehat{a}r(\widehat{\pi}_{L0}) = \left(\widehat{\omega}_{22} - \left(a_0' \widehat{\Omega}^{-1} a_0 \right)^{-1} \right) (Z'Z)^{-1},$$

and $a_0' \widehat{\Omega}^{-1} a_0 = \widehat{\sigma}_0^2 / \left| \widehat{\Omega} \right|$. It follows that $\widehat{\omega}_{22} - \left(a_0' \widehat{\Omega}^{-1} a_0 \right)^{-1} = (\widehat{\omega}_{12} - \beta_0 \widehat{\omega}_{22})^2 / \widehat{\sigma}_0^2$ and hence result (9) follows.

The Likelihood Ratio test we consider here is the test denoted LR_1 in Moreira (2003), which is a criterion difference test. This LR statistic is given by

$$\begin{aligned} LR &= \frac{u_0' P_Z u_0}{\widehat{\sigma}_0^2} - \frac{\widehat{u}_L' P_Z \widehat{u}_L}{\widehat{\sigma}_L^2} \\ &= AR - B\left(\widehat{\beta}_L \right), \end{aligned} \tag{10}$$

where $B\left(\widehat{\beta}_L \right)$ is the Basmann (1960) test for overidentifying restrictions in model (1), with $B\left(\widehat{\beta}_L \right) = n\widehat{\kappa}$. Under standard strong instrument asymptotics, LR has a limiting χ_1^2 distribution. However, under weak instruments, its distribution is not invariant with respect to the value of $\pi_n = c/\sqrt{n}$, unlike AR and LM. As Moreira (2003) showed, the asymptotic conditional distribution of LR under the null, conditional on the value of the $\widehat{\pi}_{L0}$-based Wald test statistic for testing H$_0$: $\pi = 0$,

$$\tau_0 = \widehat{\pi}_{L0}' \left(V\widehat{a}r(\widehat{\pi}_{L0}) \right)^{-1} \widehat{\pi}_{L0} = \frac{a_0' \widehat{\Omega}^{-1} W' P_Z W \widehat{\Omega}^{-1} a_0}{a_0' \widehat{\Omega}^{-1} a_0}, \tag{11}$$

is given by

$$f(LR|\tau_0) = \frac{1}{2} \left(\xi_1 + \xi_{k_z - 1} - \tau_0 + \sqrt{(\xi_1 + \xi_{k_z - 1} + \tau_0)^2 - 4\xi_{k_z - 1}\tau_0} \right),$$

where ξ_1 and ξ_{k-1} are independent χ^2_1 and $\chi^2_{k_z-1}$ distributed random variables. Conditional critical values for the LR test can then be simulated, or the conditional p-values calculated by numerical integration, Moreira (2003), Andrews et al. (2007), and Mikusheva and Poi (2006), resulting in correct size for this conditional LR (CLR) test, also when instruments are weak or uninformative.

Conditional tests with correct size under the null in weakly identified models can also be obtained for standard Wald tests, for example based on 2SLS, LIML, Fuller and bias-corrected 2SLS estimators, see Andrews, Moreira, and Stock (2007) and Mills, Moreira, and Vilela (2014). Mills et al. (2014) provide the details for obtaining the distributions of these tests conditional on τ_0, and they also considered one-sided conditional t-tests. A LIML-based test considered by Mills et al. (2014) is given by

$$W_0\left(\widehat{\beta}_L\right) = t_0(\widehat{\beta}_L)^2 = \frac{\left(\widehat{\beta}_L - \beta_0\right)^2 (x'P_Z x - n\widehat{\kappa}\widehat{\omega}_{22})}{\widehat{\sigma}_0^2}. \tag{12}$$

The difference with the standard LIML-based Wald test is the use of the restricted estimator $\widehat{\sigma}_0^2$ instead of the unrestricted $\widehat{\sigma}_L^2$. We show that the $W_0\left(\widehat{\beta}_L\right)$ statistic is identical to the LR statistic, as stated in the following proposition.

Proposition 1. *Let LR be as defined in (10) and $W_0(\widehat{\beta}_L)$ be as defined in (12). Then*

$$W_0\left(\widehat{\beta}_L\right) = LR.$$

Proof. See Appendix A. □

It follows from Proposition 1 and the results in Mills et al. (2014) that the conditional $W_0(\widehat{\beta}_L)$ ($CW_0(\widehat{\beta}_L)$) and CLR tests are also equivalent. For the just-identified case, $k_z = 1$, it follows that $W_0\left(\widehat{\beta}_{IV}\right) = LR = AR = LM$, where $\widehat{\beta}_{IV} = (z'x)^{-1} z'y$. The equivalence of $W_0\left(\widehat{\beta}_{IV}\right)$ and AR was derived by Feir, Lemieux, and Marmer (2016), see also Lee, McCrary, Moreira, and Porter (2020).

We explain features of the fixed Σ design power curve for the CLR test on the basis of the representation of the LR statistic as $W_0(\widehat{\beta}_L)$ in the next section. Before we do so, we discuss an alternative version of the LR test statistic, which has the same conditional properties, but which does not seem to have been used in applied work.

2.1 Sargan Version of the CLR Test

The LIML-based Sargan (1958) test for overidentifying restrictions is given by

$$S\left(\widehat{\beta}_L\right) = \frac{\widehat{u}_L' P_Z \widehat{u}_L}{\widehat{u}_L' \widehat{u}_L / n},$$

and the Stock and Wright (2000) S statistic is

$$S\left(\beta_0\right) = \frac{u_0' P_Z u_0}{u_0' u_0 / n}.$$

Then an alternative LR test statistic is given by the criterion difference

$$LR_S = S\left(\beta_0\right) - S\left(\widehat{\beta}_L\right).$$

This statistic is obtained by replacing $\widehat{\Omega}$ above by $\widehat{\Sigma}_W = W'W/n$. A conditional CLR_S statistic can be defined in a similar manner by conditioning on the $\widehat{\pi}_{L0,\widehat{\Sigma}_W}$-based Wald test for H$_0$: $\pi = 0$. The CLR_S test has the same asymptotic properties as the CLR test.

3 Power Curves

As Andrews et al. (2019, p 466) observe, if Ω is held fixed then Σ varies with β in the following way

$$\Sigma\left(\beta\right) = \left[\begin{array}{cc} \omega_{11} - 2\beta\omega_{12} + \beta^2\omega_{22} & \omega_{12} - \beta\omega_{22} \\ \omega_{12} - \beta\omega_{22} & \omega_{22} \end{array}\right]. \tag{13}$$

It is common for simulations based on the fixed Ω design to set $\omega_{11} = \omega_{22} = 1$ and $\omega_{12} = \rho_\Omega$, from which it follows that

$$\sigma_u^2\left(\beta\right) = 1 - 2\beta\rho_\Omega + \beta^2, \tag{14}$$

$$\sigma_{uv}\left(\beta\right) = \rho_\Omega - \beta. \tag{15}$$

For testing H$_0$: $\beta = 0$, it follows that under the null, $\Sigma\left(0\right) = \Omega$, and so ρ_Ω is then an indicator of the degree of endogeneity in the null model only.

Figure 1 displays the values of $\rho_{uv}\left(\beta\right) = \sigma_{uv}\left(\beta\right)/\sigma_u\left(\beta\right)$ and $\sigma_u^2\left(\beta\right)$ as a function of β for values of $\rho_\Omega = 0$, 0.5 and 0.95. The latter two values have often been used in simulations. As is clear from the formulae (14) and (15), and highlighted by Figure 1, for

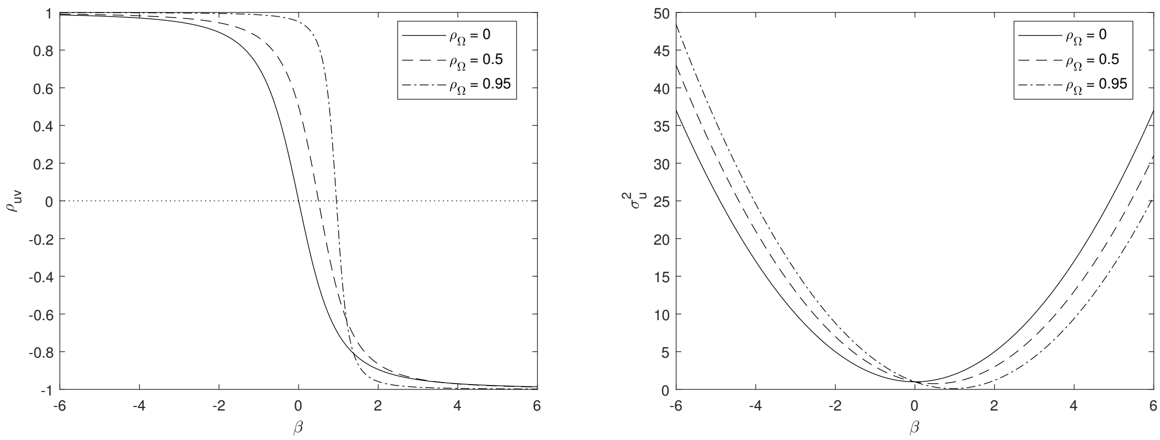

Figure 1: Values of ρ_{uv} and σ_u^2 as a function of β when holding $\Omega = \begin{bmatrix} 1 & \rho_\Omega \\ \rho_\Omega & 1 \end{bmatrix}$ constant.

every value of β the endogeneity and variance properties of the structural model change. The correlations ρ_{uv} are positive for $\beta < \rho_\Omega$ and negative for $\beta > \rho_\Omega$, approaching 1 and -1 quite rapidly, especially for $\rho_\Omega = 0.95$.

The representation of the LR statistic as $W_0\left(\widehat{\beta}_L\right)$ enables us to better understand the weak instrument power properties of the LR and CLR tests in the fixed Σ design, and to link them to those of the fixed Ω design via the representation in (13). Holding Σ fixed while varying the value of β does affect the location of the LIML estimator $\widehat{\beta}_L$, but not its variance. From the weak instrument asymptotic results of Stock and Yogo (2005), we find that the LIML estimator is median unbiased but is skewed to the left if $\rho_{uv} > 0$ and skewed to the right when $\rho_{uv} < 0$. From this it follows that the power function of an infeasible LIML-based Wald test for H_0: $\beta = \beta_0$ that uses the true unknown σ_u^2 instead of the variance estimator $\widehat{\sigma}_L^2$ is asymmetric when $\rho_{uv} \neq 0$. For $\delta > 0$, if $\rho_{uv} > 0$ then the power of this unfeasible test is higher for $\beta = \beta_0 - \delta$ than for $\beta = \beta_0 + \delta$ and vice versa when $\rho_{uv} < 0$.

For the fixed Σ design, treating $\Omega(\beta)$ known, we have for $\sigma_0^2(\beta)$ in the denominator of $W_0\left(\widehat{\beta}_L\right)$ in (12), that

$$\sigma_0^2(\beta) = b_0'\Omega(\beta)\, b_0 = \sigma_u^2 - 2(\beta_0 - \beta)\,\sigma_{uv} + (\beta_0 - \beta)^2\,\sigma_v^2.$$

It follows that $\sigma_0^2(\beta)$ is minimised at $\beta = \beta_\sigma = \beta_0 - \frac{\sigma_{uv}}{\sigma_v^2} = \beta_0 - \rho_{uv}\frac{\sigma_u}{\sigma_v}$, with $\sigma_0^2(\beta_\sigma) = \sigma_u^2 - \frac{\sigma_{uv}^2}{\sigma_v^2} = \sigma_u^2(1 - \rho_{uv}^2)$. Therefore, if $\rho_{uv} > 0$, it follows that $\sigma_0^2(\beta)$ is minimised for a value of $\beta < \beta_0$, and vice versa. Together with the skewness of the LIML estimator, it is then clear that the asymptotic power function of the LR test for H_0: $\beta = \beta_0$ is asymmetric when $\rho_{uv} \neq 0$. For $\delta > 0$, the power is higher when $\beta = \beta_0 - \delta$ than when $\beta = \beta_0 + \delta$ if $\rho_{uv} > 0$, and vice versa if $\rho_{uv} < 0$.

It follows that $\sigma_0^2(\beta_\sigma) \to 0$ for $|\rho_{uv}| \to 1$. As Davidson and MacKinnon (2015, pp 831-832) show, for $\lambda > 0$, $n\widehat{\kappa} \xrightarrow{d} \chi^2_{k_z - 1}$ when $|\rho_{uv}| \to 1$. Further, the distribution of $x'P_z x$ and the value of $\omega_{22} = \sigma_v^2$ are not affected by ρ_{uv}. It therefore follows that $LR \to \infty$ when $\beta = \beta_\sigma$ and $|\rho_{uv}| \to 1$, resulting in the power of the LR test then going to 1. When $\rho_{uv} \to 1$, the point where the power of the LR test goes to 1 lies to the left of β_0 at $\beta_\sigma = \beta_0 - \frac{\sigma_u}{\sigma_v}$. This is the situation we depict in the right panel of Figure 2 below.

For the CLR test, the critical values depend on the value of τ_0, which changes with different values of β. They range from the critical values of the χ_1^2 distribution as $\tau_0(\beta) \to \infty$ to the critical values of the $\chi^2_{k_z}$ distribution as $\tau_0(\beta) \to 0$ (Moreira, 2003). It therefore follows that the power of the CLR test also approaches 1 when $\beta = \beta_\sigma$ and $|\rho_{uv}| \to 1$.

The non-centrality parameter of the limiting non-central chi-squared distribution of $\tau_0(\beta)$ is given by

$$nc_{\tau_0(\beta)} = \sigma_v^2 \lambda \frac{\left(a'\left(\Omega(\beta)\right)^{-1} a_0\right)^2}{a_0'\left(\Omega(\beta)\right)^{-1} a_0} = \frac{\lambda}{\sigma_u^2\left(1 - \rho_{uv}^2\right)} \frac{\left(\sigma_u^2 - \sigma_{uv}(\beta_0 - \beta)\right)^2}{\sigma_0^2(\beta)}.$$

For $\lambda > 0$, it follows that $nc_{\tau_0}(\beta_{nc}) = 0$ for $\beta_{nc} = \beta_0 - \frac{\sigma_u^2}{\sigma_{uv}} = \beta_0 - \frac{1}{\rho_{uv}} \frac{\sigma_u}{\sigma_v}$. Further, for $\beta \neq \{\beta_\sigma, \beta_{nc}\}$ it follows that $nc_{\tau_0(\beta)} \to \infty$ for $|\rho_{uv}| \to 1$, whereas $nc_{\tau_0}(\beta_\sigma) = \lambda$. The maximum of $nc_{\tau_0(\beta)}$ is obtained at $\beta = \beta_0$, with $nc_{\tau_0(\beta_0)} = \frac{\lambda}{1 - \rho_{uv}^2}$. As the critical values of the CLR test are a decreasing function of $\tau_0(\beta)$ and hence of $nc_{\tau_0(\beta)}$, and because $nc_{\tau_0(\beta)}$ is asymmetric around β_0, it is not a priori clear whether the weak instruments asymmetry of the power function of the LR test is carried over to that of the CLR test, but in the simulations in the literature this has been found to be the case, as confirmed below.

Whilst the standard 2SLS-based Wald tests have their largest weak instrument size distortions at $|\rho_{uv}| = 1$, see Stock and Yogo (2005), this is not the case for the LR test. In fact, for $\lambda > 0$, the LR test has *no* size distortion when $|\rho_{uv}| \to 1$, as then $nc_{\tau_0(\beta_0)} = \frac{\lambda}{1 - \rho_{uv}^2} \to \infty$, see also Andrews et al. (2019). The non-centrality parameter $nc_{\tau_0(\beta_0)}$ is minimised at $\rho_{uv} = 0$, and hence the weak instrument size distortion of the LR test is maximum at $\rho_{uv} = 0$, see Figure B.1 in Appendix B for an illustration.

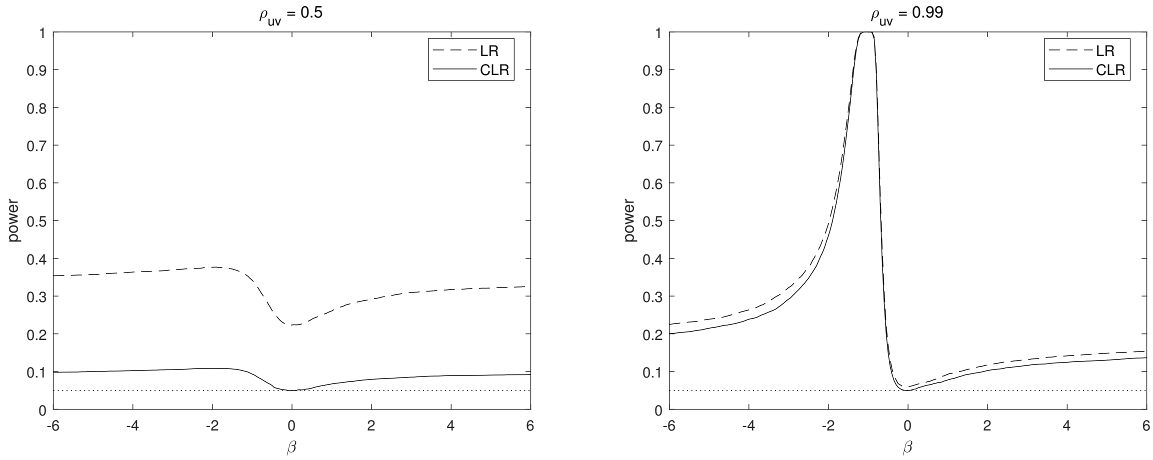

Figure 2: Asymptotic power of LR and CLR tests holding $\Sigma = \begin{bmatrix} 1 & \rho_{uv} \\ \rho_{uv} & 1 \end{bmatrix}$ constant, $\lambda = 1$, $k_z = 5$.

Figure 2 shows the asymptotic power curves, calculated as in Andrews et al. (2006), for the fixed Σ design with $\sigma_u^2 = \sigma_v^2 = 1$ and for the values of $\rho_{uv} = 0.5$ and $\rho_{uv} = 0.99$. This is for a very weak instrument setting of $\lambda = 1$ and for $k_z = 5$. Throughout, the significance level is 5%. The results confirm the asymmetry of the LR test as derived above, which is maintained by the CLR test. Especially for the high endogeneity case of $\rho_{uv} = 0.99$ the asymmetry is severe, as expected. When $\rho_{uv} = 0.5$, the CLR test corrects the quite large size distortion of the LR test, but it has low power everywhere, due to the weakness of the instruments. When $\rho_{uv} = 0.99$, the differences between the LR and CLR tests are small as expected, but even in this very weak instrument setting, the power of the CLR test is equal to 1 when $\beta = \beta_\sigma = -0.99$, confirming the derivations above. These power curves of the CLR test are very similar to the ones presented by Stock et al. (2002) and Davidson and MacKinnon (2008).

The asymmetry of the power curve for the fixed Σ design, together with the relationship between $\Sigma(\beta)$ and Ω in the fixed Ω design as depicted in Figure 1, explains the behaviour of the power curve for the fixed Ω design. In Figure 3, we overlay the power curve for the fixed Ω design, with $\omega_{11} = \omega_{22} = 1$ and $\omega_{12} = \rho_\Omega$, to those of fixed $\Sigma(\beta^*)$ designs for different values of β^*, such that $\sigma_u^2(\beta^*) = 1 - 2\beta^*\rho_\Omega + \beta^{*2}$, $\sigma_{uv}(\beta^*) = \rho_\Omega - \beta^*$, and $\sigma_v^2 = 1$, satisfying (13) at $\beta = \beta^*$, but then holding these constant whilst varying the value of β and testing H$_0$: $\beta = \beta_0$. Hence, for any given choice of β^*, the fixed Ω and fixed $\Sigma(\beta^*)$ power curves coincide at $\beta = \beta^*$. At other values of β the difference is that the fixed $\Sigma(\beta^*)$ design holds Σ constant at $\Sigma(\beta^*)$, while in the fixed Ω design ρ_{uv} and σ_u^2 vary with β as shown in Figure 1.

Differences in the behaviour of σ_0^2 in the fixed Ω and fixed $\Sigma(\beta^*)$ designs are an important, and tractable, element in understanding why the two designs yield very different power curves. In the fixed Ω design σ_0^2 is not a function of β, and with $\beta_0 = 0$ we have here that $\sigma_0^2 = b_0'\Omega b_0 = 1$. In contrast, in the fixed $\Sigma(\beta^*)$ design, $\sigma_0^2(\beta)$ does vary with β. More specifically, in the fixed $\Sigma(\beta^*)$ design, for a given choice of β^*, we have that $\sigma_0^2(\beta) = \sigma_u^2(\beta^*) - 2(\beta_0 - \beta)\sigma_{uv}(\beta^*) + (\beta_0 - \beta)^2\sigma_v^2$. It follows, for $\beta_0 = 0$, that in this design $\sigma_0^2(\beta)$ is minimised at $\beta = \beta_\sigma^* = \beta^* - \rho_\Omega$, where $\sigma_0^2(\beta_\sigma^*) = 1 - \rho_\Omega^2$. Hence, for $\rho_\Omega = 0$, σ_0^2 in the fixed Ω design is always – regardless of the value of β – at the minimum value of 1 attained by $\sigma_0^2(\beta)$ in the fixed $\Sigma(\beta^*)$ design. For $\rho_\Omega = 0$, the power curve of the fixed Ω design then crosses those of the fixed $\Sigma(\beta^*)$ designs very close to the points of maximum power of the fixed $\Sigma(\beta^*)$ designs, as shown in the left panel of Figure C.1 in Appendix C.

When $\rho_\Omega \neq 0$, the Ω design is not on the minimum $\sigma_0^2(\beta_\sigma^*)$ path of the fixed $\Sigma(\beta^*)$ designs, but it is not far away from it. The left panel of Figure 3 shows the power curve for the fixed Ω design with $\rho_\Omega = 0.5$, again for $\lambda = 1$ and $k_z = 5$. Even in this very weak instrument setting, power approaches 1 for large absolute values of β, very unlike the power curve for the fixed Σ design with $\rho_{uv} = 0.5$. In the right panel, this fixed Ω power curve is overlayed by the fixed $\Sigma(\beta^*)$ power curves, for values of $\beta^* = -4, -2, 1, 3, 5$. It is clear that the fixed Ω power curve goes through the points close to the maxima of the fixed $\Sigma(\beta^*)$ power curves, in this case especially for positive and large absolute values of β. A similar pattern is found for $\rho_\Omega = 0.95$ as shown in right panel of Figure C.1 in Appendix C.

4 LIML- and Fuller-Based CW and CW_0 Tests

Moreira (2003) compared the behaviour of the conditional 2SLS-based Wald test to that of the CLR test in a fixed Σ design. Andrews et al. (2007) compared the behaviours of the 2SLS-, LIML- and Fuller(1)-based CW tests to that of the CLR test in the fixed Ω design. They find that the CW-Fuller test performs best of the three conditional Wald tests, but that its performance is, overall, "very poor relative to the CLR test"

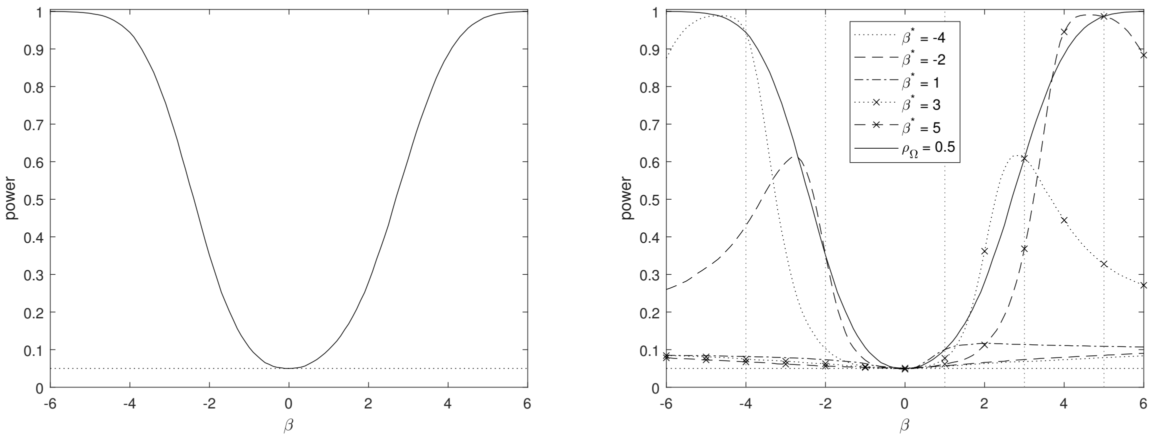

Figure 3: Asymptotic power of CLR test, $\lambda = 1$, $k_z = 5$. Left panel holding Ω constant, $\rho_\Omega = 0.5$. Right panel holding $\Sigma(\beta^*)$ fixed for various values of β^*, with $\sigma_u^2(\beta^*) = 1 - 2\beta^*\rho_\Omega + \beta^{*2}$ and $\sigma_{uv}(\beta^*) = \rho_\Omega - \beta^*$, $\rho_\Omega = 0.5$.

(Andrews et al., 2007, p 131). Mills et al. (2014) compared the conditional t-test versions for one-sided tests in the fixed Ω design and found the 2SLS and Fuller versions to have good performance. They also analysed the performances of $CW_0\left(\widehat{\beta}_L\right)$ and $CW_0\left(\widehat{\beta}_{Full}\right)$ for two-sided tests, but again for the fixed Ω design only. As far as we are aware, the performances of these conditional tests have not been evaluated under the fixed Σ design, which is what we do here. We compare the performances of the CW-LIML, CW-Fuller, CLR/CW_0-LIML and CW_0-Fuller tests, keeping Σ fixed for different values of β, with $\sigma_u^2 = \sigma_v^2 = 1$ and for $\rho_{uv} = 0, 0.25, 0.50, 0.75$, testing H_0: $\beta = 0$. As above, $k_z = 5$.

Figure 4 shows the power curves for an instrument strength of $\lambda/k_z = 2$. Appendix D further presents the power curves for $\lambda/k_z = 0.5$, 1 and 4. We follow the practice in the literature to report the rejection frequencies of the tests as a function of $\beta\sqrt{\lambda}$. At low levels of endogeneity, $\rho_{uv} = 0$ and $\rho_{uv} = 0.25$, the behaviour of the CW-LIML and CW-Fuller tests are virtually identical and they are the most powerful across the range of values of $\beta\sqrt{\lambda}$ when $\rho_{uv} = 0$. For $\rho_{uv} = 0.25$ they are also most powerful, but for a small bias of the tests for small negative values of $\beta\sqrt{\lambda}$. The bias of the CW tests increases with increasing values of ρ_{uv}. The CW_0-Fuller test is less biased than the CW tests. Its power dominates that of the CLR test at the lower endogeneity levels $\rho_{uv} = 0$ and $\rho_{uv} = 0.25$. At the higher level of $\rho_{uv} = 0.5$, the power of the CW_0-Fuller test also dominates that of the CLR test except for some negative values of $\beta\sqrt{\lambda}$ close to 0, and where the difference in power between the two tests is small. At the higher level of endogeneity, $\rho_{uv} = 0.75$,

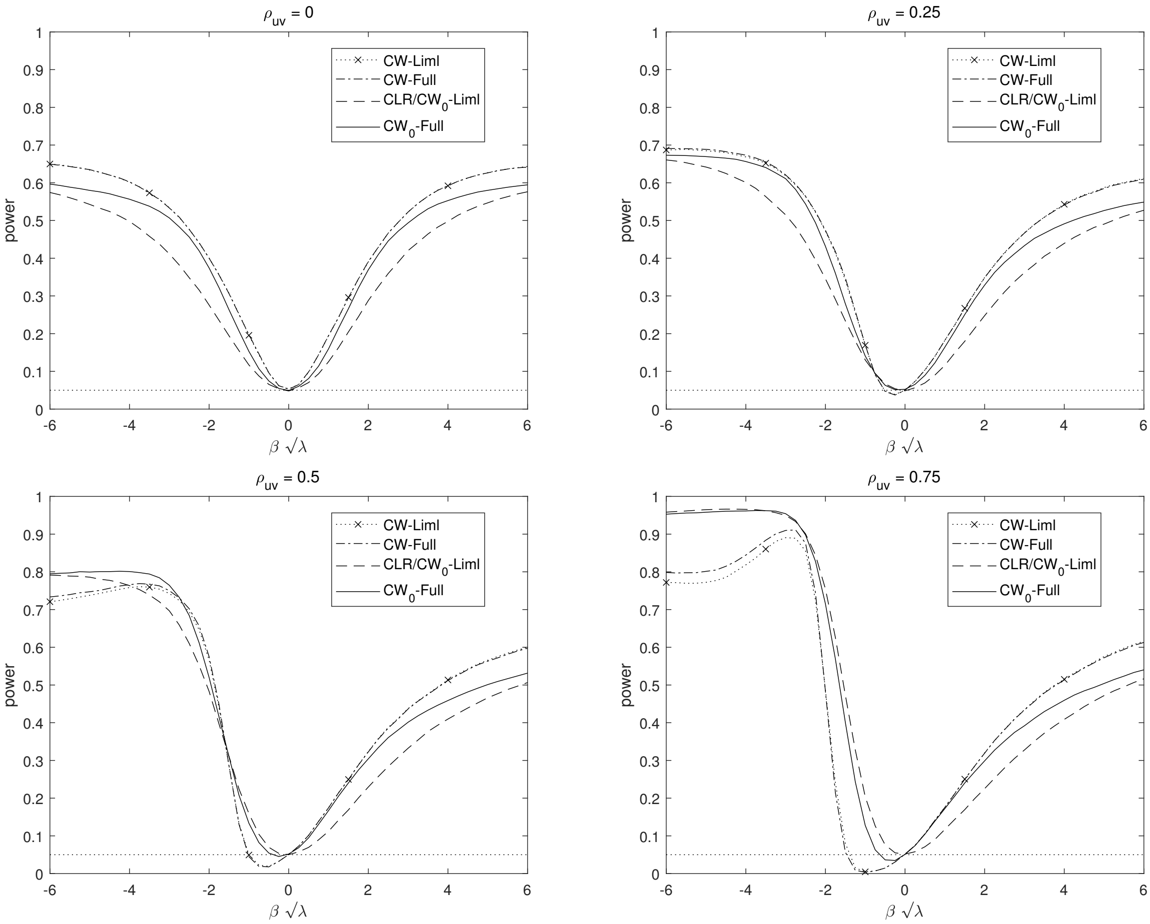

Figure 4: Power curves for CW-LIML, CW-Fuller, CLR/CW_0-LIML, and CW_0-Fuller tests for fixed Σ design, $k_z = 5$ and $\lambda/k_z = 2$, for different values of ρ_{uv}.

the CW-tests and the CW_0-test have more power than the CLR test for positive values of $\beta\sqrt{\lambda}$, whereas the CLR test dominates at negative values of $\beta\sqrt{\lambda}$.

Whilst the CLR test has been shown to have power close to the two-sided power envelope for unbiased tests (Andrews et al., 2007), the results here show that the biased CW and CW_0-Fuller tests can have more power than the CLR test in low to moderate endogeneity environments, in which case there is also only a small to moderate bias in these tests. This seems an important observation, as this is a situation that may well be encountered in practice.

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*, 20, 46–63.

Andrews, D. W. K., V. Marmer, and Z. Yu (2019): "On Optimal Inference in the Linear IV Model," *Quantitative Economics*, 10, 457–485.

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

——— (2007): "Performance of Conditional Wald Tests in IV Regression With Weak Instruments," *Journal of Econometrics*, 139, 116–132.

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

Bowden, R. J. and D. A. Turkington (1984): *Instrumental Variables*, Cambridge University Press.

Davidson, R. and J. G. MacKinnon (2008): "Bootstrap Inference in a Linear Equation Estimated by Instrumental Variables," *Econometrics Journal*, 11, 443–477.

——— (2015): "Bootstrap Tests for Overidentification in Linear Regression Models," *Econometrics*, 3, 825–863.

Feir, D., T. Lemieux, and V. Marmer (2016): "Weak Identification in Fuzzy Regression Discontinuity Designs," *Journal of Business & Economic Statistics*, 34, 185–196.

Hausman, J. A. (1983): "Specification and Estimation of Simultaneous Equation Models," in *Handbook of Econometrics*, ed. by Z. Griliches and M. D. Intriligator, Elsevier, vol. I, 391–448.

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

Lee, D. S., J. McCrary, M. J. Moreira, and J. Porter (2020): "Which t for IV?" Mimeo, Princeton University.

Mikusheva, A. and B. P. Poi (2006): "Tests and Confidence Sets with Correct Size when Instruments are Potentially Weak," *The Stata Journal*, 6, 335–347.

Mills, B., M. J. Moreira, and L. P. Vilela (2014): "Tests Based on t-Statistics for IV Regression with Weak Instruments," *Journal of Econometrics*, 182, 351–363.

Moreira, H. and M. J. Moreira (2019): "Optimal Two-Sided Tests for Instrumental Variables Regression with Heteroskedastic and Autocorrelated Errors," *Journal of Econometrics*, 213, 398–433.

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

——— (2009): "Tests with Correct Size when Instruments can be Arbitrarily Weak," *Journal of Econometrics*, 152, 131–140.

Poskitt, D. and C. Skeels (2008): "Conceptual Frameworks and Experimental Design in Simultaneous Equations," *Economics Letters*, 100, 138–142.

Sargan, J. D. (1958): "The Estimation of Economic Relationships using Instrumental Variables," *Econometrica*, 26, 393–415.

Stock, J. H. and J. H. Wright (2000): "GMM with Weak Identification," *Econometrica*, 68, 1055–1096.

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

Stock, J. H. and M. Yogo (2005): "Testing for Weak Instruments in Linear IV Regression," in *Identification and Inference for Econometric Models*, ed. by D. W. K. Andrews and J. H. Stock, Cambridge University Press, 80–108.

Windmeijer, F. (2018): "Testing Over- and Underidentification in Linear Models, with Applications to Dynamic Panel Data and Asset-Pricing Models," Discussion Paper 18/696, Dept. of Economics, University of Bristol.

Appendix

A Proof of Proposition 1

It follows from (7) that

$$\begin{aligned} \widehat{u}_L' P_Z x &= \widehat{u}_L' Z \widehat{\pi} = \widehat{u}_L' Z \widehat{\pi}_L + \frac{\widehat{u}_L' P_Z \widehat{u}_L}{\widehat{\sigma}_L^2} \left(\widehat{\omega}_{12} - \widehat{\beta}_L \widehat{\omega}_{22} \right) \\ &= B\left(\widehat{\beta}_L\right) \left(\widehat{\omega}_{12} - \widehat{\beta}_L \widehat{\omega}_{22} \right), \end{aligned}$$

as $\widehat{u}_L' Z \widehat{\pi}_L = \widehat{u}_L' \widehat{x}_L = 0$.

As $u_0 = \widehat{u}_L + x\left(\widehat{\beta}_L - \beta_0\right)$, it follows that

$$\begin{aligned} \frac{u_0' P_Z u_0}{\widehat{\sigma}_0^2} &= \frac{\widehat{u}_L' P_Z \widehat{u}_L + 2\widehat{u}_L' P_Z x \left(\widehat{\beta}_L - \beta_0\right) + \left(\widehat{\beta}_L - \beta_0\right)^2 x' P_Z x}{\widehat{\sigma}_0^2} \\ &= \frac{\widehat{u}_L' P_Z \widehat{u}_L}{\widehat{\sigma}_0^2} + \frac{2B\left(\widehat{\beta}_L\right)\left(\widehat{\beta}_L - \beta_0\right)}{\widehat{\sigma}_0^2} \left(\widehat{\omega}_{12} - \widehat{\beta}_L \widehat{\omega}_{22} \right) + \frac{\left(\widehat{\beta}_L - \beta_0\right)^2 x' P_Z x}{\widehat{\sigma}_0^2}. \end{aligned}$$

Further

$$\begin{aligned} \frac{\widehat{u}_L' P_Z \widehat{u}_L}{\widehat{\sigma}_0^2} - \frac{\widehat{u}_L' P_Z \widehat{u}_L}{\widehat{\sigma}_L^2} &= \frac{\widehat{u}_L' P_Z \widehat{u}_L}{\widehat{\sigma}_L^2 \widehat{\sigma}_0^2} \left(\widehat{\sigma}_L^2 - \widehat{\sigma}_0^2 \right) \\ &= \frac{B\left(\widehat{\beta}_L\right)}{\widehat{\sigma}_0^2} \left(\left(\widehat{\beta}_L^2 - \beta_0^2 \right) \widehat{\omega}_{22} - 2\left(\widehat{\beta}_L - \beta_0\right) \widehat{\omega}_{12} \right). \end{aligned}$$

As

$$\left(\widehat{\beta}_L^2 - \beta_0^2 \right) \widehat{\omega}_{22} = -\left(\widehat{\beta}_L - \beta_0\right)^2 \widehat{\omega}_{22} + 2\widehat{\beta}_L \left(\widehat{\beta}_L - \beta_0\right) \widehat{\omega}_{22},$$

and $B\left(\widehat{\beta}_L\right) = n\widehat{\kappa}$, it follows that

$$\begin{aligned} \frac{u_0' P_Z u_0}{\widehat{\sigma}_0^2} - \frac{\widehat{u}_L' P_Z \widehat{u}_L}{\widehat{\sigma}_L^2} &= \frac{\left(\widehat{\beta}_L - \beta_0\right)^2 \left(x' P_Z x - B\left(\widehat{\beta}_L\right) \widehat{\omega}_{22} \right)}{\widehat{\sigma}_0^2} \\ &= \frac{\left(\widehat{\beta}_L - \beta_0\right)^2 \left(x' P_Z x - n\widehat{\kappa} \widehat{\omega}_{22} \right)}{\widehat{\sigma}_0^2} \\ &= W_0\left(\widehat{\beta}_L\right). \end{aligned}$$

B Size Distortion of the LR Test

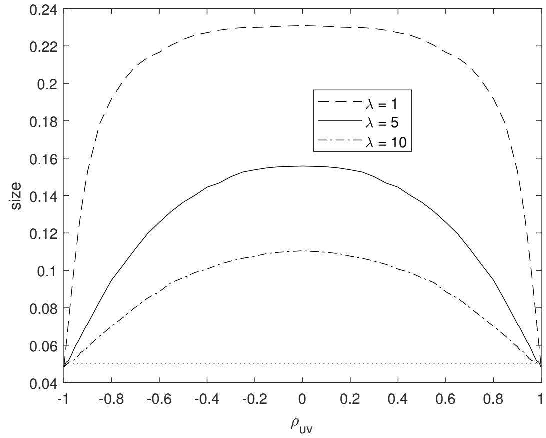

Figure B.1: Size properties of LR test for different values of ρ_{uv} and instrument strength λ. $k_z = 5$.

C Power Curves of CLR Test, Fixed Ω Design and Fixed $\Sigma(\beta^*)$ Designs

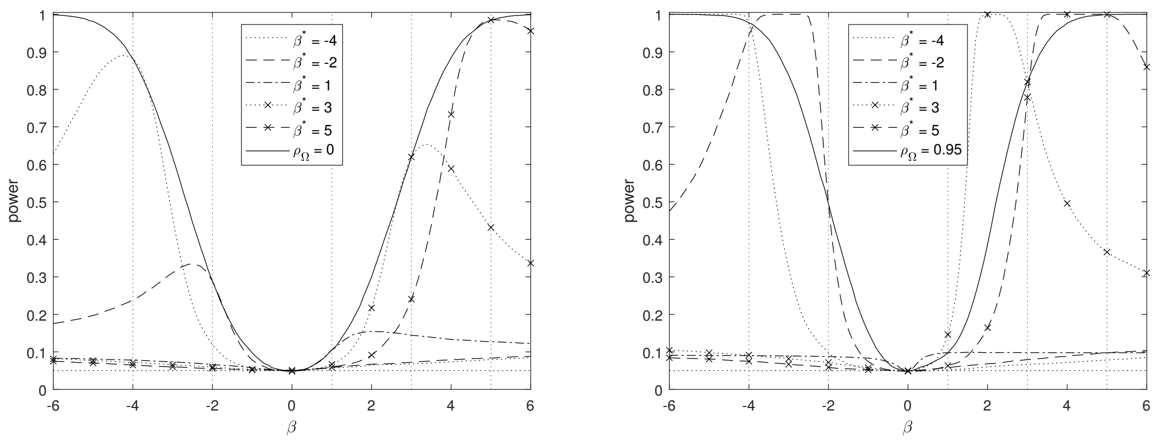

Figure C.1: Asymptotic power of CLR test, $\lambda = 1$, $k_z = 5$. Fixed Ω design, left panel $\rho_\Omega = 0$, right panel $\rho_\Omega = 0.95$ and fixed $\Sigma(\beta^*)$ designs for various values of β^*, with $\sigma_u^2(\beta^*) = 1 - 2\beta^*\rho_\Omega + \beta^{*2}$ and $\sigma_{uv}(\beta^*) = \rho_\Omega - \beta^*$.

D Power Curves of CLR, CW-Fuller and CW_0-Fuller Tests

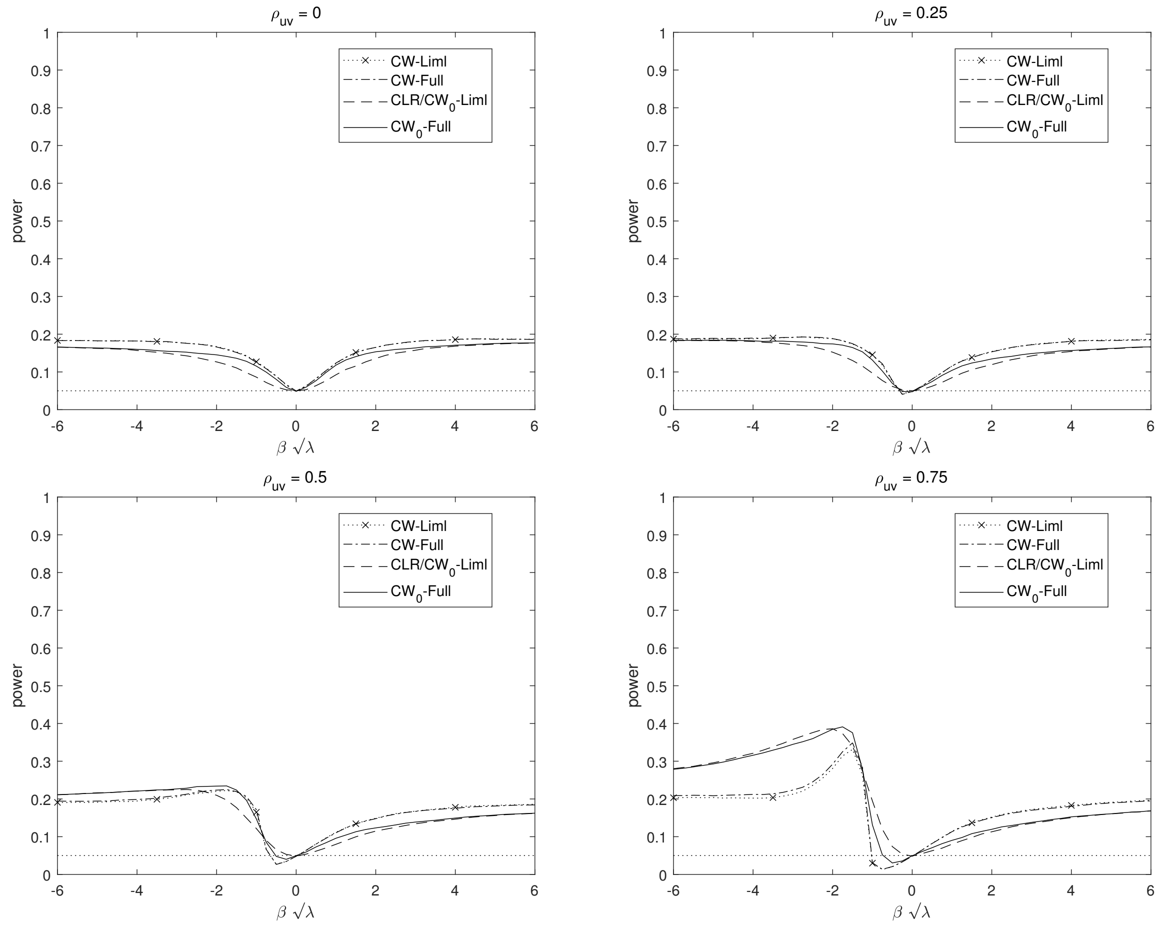

Figure D.1: $\lambda/k_z = 0.5$, $k_z = 5$.

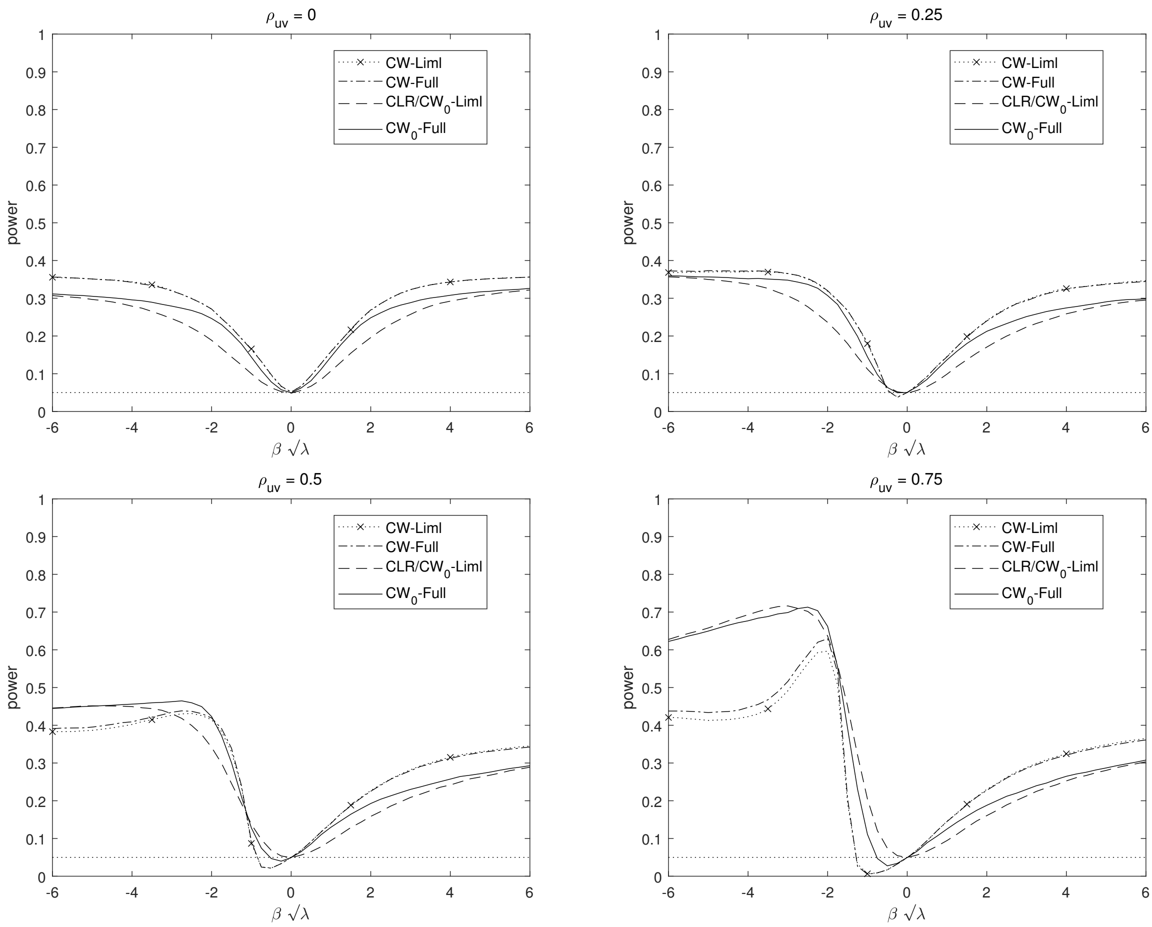

Figure D.2: $\lambda/k_z = 1$, $k_z = 5$.

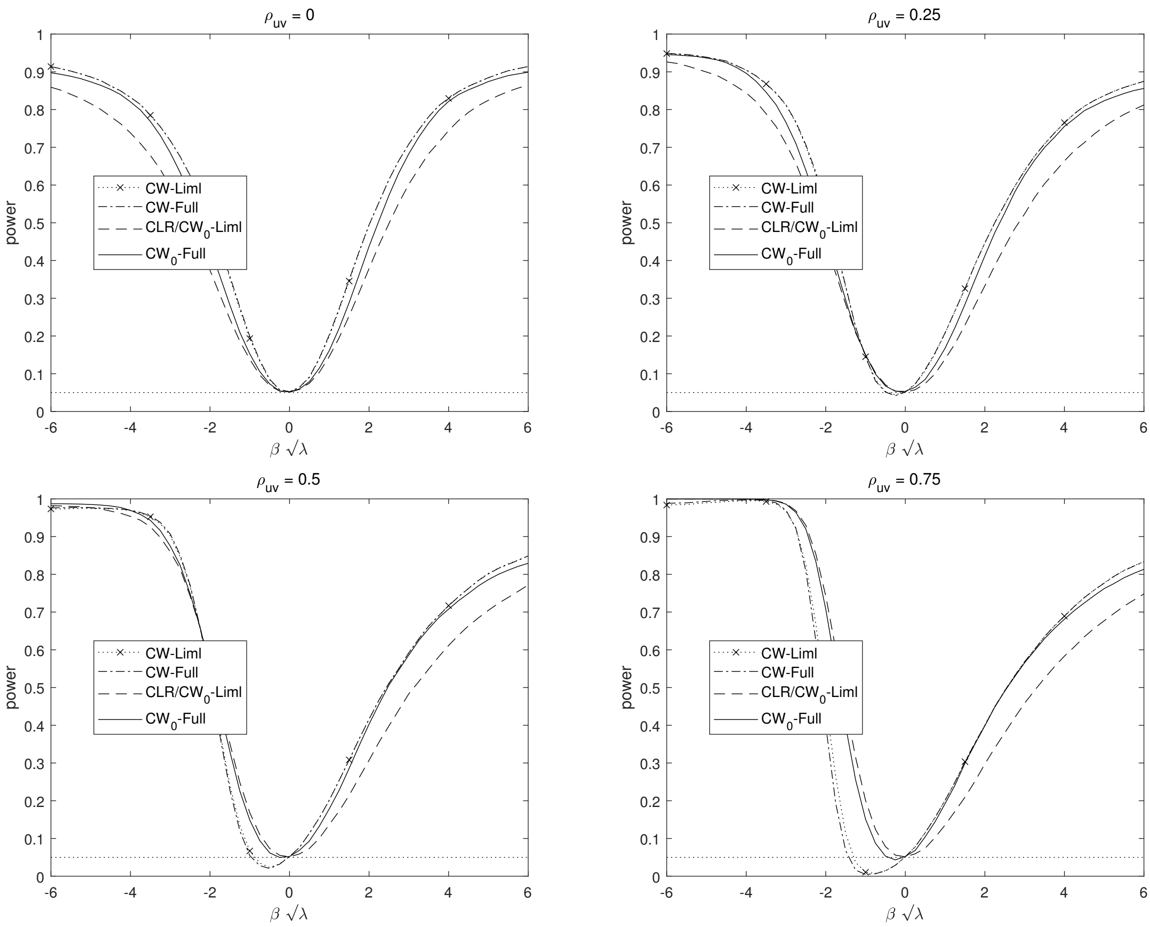

Figure D.3: $\lambda/k_z = 4$, $k_z = 5$.