Reassessment of classic case studies in labor economics with new instrument-free methods

Jan F. Kiviet and Sebastian Kripfganz

5 July 2020

Appendices

The following appendices provide further background to the Monte Carlo simulation study and to the options to cope with omitted relevant regressors by TSLS and KLS:

Appendix A gives a detailed description of the graphically presented findings from the simulation study.

Appendix B provides full technical details on the chosen design of the simulation experiments.

Appendix C indicates the technical requirements for consistent estimation of the direct causal effect of particular regressors by TSLS or by KLS in linear regression models with omitted regressors.

A. Findings from the simulation study

The data generating process we used in the Monte Carlo experiments, presented in all its details in Appendix B, is a generalization of those used in the earlier studies on KLS. It concerns a linear regression model for a dependent variable y with an intercept and one slope coefficient β for the single possibly endogenous regressor x. The very simple model $y = c + \beta x + u$ has i.i.d. (independent and identically distributed) disturbances u. The correlation of the regressor and the disturbance, indicated by ρ_{xu}, can be controlled in the experiments. Next to the internal instrument established by the constant, there are two external instrumental variables, z_1 and z_2. For these their correlation (strength/weakness) with the single regressor can be controlled by ρ_{z_1x} and ρ_{z_2x} respectively. Moreover, their correlation (validity/invalidity) with the disturbance can be controlled by ρ_{z_1u} and ρ_{z_2u}.

For various interesting combinations of ρ_{xu}, ρ_{z_1x}, ρ_{z_2x}, ρ_{z_1u}, ρ_{z_2u} and sample size n we will examine: (i) the rejection probability of the Sargan test at nominal significance level α, where we shall consider $0.01 \leq \alpha \leq 0.5$; and (ii) the estimation errors $\hat{\beta} - \beta$ for various estimators of the slope coefficient, namely OLS, IV (just using the external instrument z_1), TSLS (using both z_1 and z_2) and the new instrument-free estimator KLS. In Appendix B it is proved that all presented findings are invariant with respect to the actual values of the intercept c and slope β, and also to the means of x, z_1 and z_2. Therefore, without loss of generality, we fixed these all at zero. The results are also invariant regarding the variance of z_1 and of z_2. Therefore we gave σ_{z_1} and σ_{z_2} value unity. In the graphs below we present the quartiles of the distribution (as assessed from 100,001 replications of the experiments) of the various estimation errors for the case $\sigma_u/\sigma_x = 1$. Outcomes for different σ_u/σ_x ratios can be obtained simply by adapting the scale on the vertical axis accordingly. For the Sargan test we present the rejection frequency over all replications of the simulation for different values of α. These frequencies are in fact not only invariant with respect to β, σ_{z_1} and σ_{z_2}, but also to σ_u and σ_x.

Not all values smaller than one in absolute value for the five correlation coefficients are compatible. For instance, it is self-evidently impossible to have $\rho_{z_1u} = 0$, whereas both ρ_{xu} and ρ_{z_1x} are close to unity. Close to boundary values, and to notoriously problematic cases such as $\rho_{z_1x} \to 0$, $\rho_{xu} \to 1$, or n very small, instrumental variable estimators may show pathological behavior. It is not our intention here to demonstrate that such cases exist and are also problematic for the Sargan test.[1] Our primary aim is here to demonstrate that serious problems occur as well for parameter combinations which at first sight seem pretty harmless. Therefore we start to examine a reasonably large sample ($n = 250$) and rather middle of the road combinations of the correlations arising from:

$$\begin{aligned} \rho_{z_1 x} &\in \{0.3, 0.6\}, \qquad & \rho_{z_2 x} &\in \{0.1, 0.4\}, \\ \rho_{z_1 u} &\in \{0.0, 0.1\}, \qquad & \rho_{z_2 u} &\in \{0.0, 0.2\}, \\ \rho_{xu} &\in \{0.3, 0.6\}. \end{aligned} \tag{A.1}$$

Hence, the instruments will not be chosen ultra-weak, nor extremely invalid. The estimator error quartiles will be examined over the whole range $0 \leq \rho_{xu} \leq 0.9$, but the Sargan test rejection frequency only for the two ρ_{xu} values given in (A.1). The artificial samples drawn are typical for cross-section data, because for all series their n observations are drawn independently. Moreover, we took all of them from the Gaussian distribution.

[Figure A.1 here]

For the various indicated specific situations the four left-hand panels of Figure A.1 contain rejection frequencies of the Sargan test, and the four right-hand panels present the quartiles of the distributions of the four estimation methods compared here. Each row of panels concerns a specific situation regarding instrument (in)validity. Each left-hand panel presents rejection frequencies over a range of nominal significance levels α for the same eight different situations regarding degree of simultaneity and strength of the two instruments. Each right-hand panel presents over a range of values of the endogeneity correlation the three quartiles of the estimation error of the slope coefficient for each of four estimation techniques, and for different situations –when relevant– regarding instrument strength. Therefore, every right-hand panel contains three similarly colored/marked lines for the same eight different estimators/cases as indicated in the legend. Of course, for each triple of lines the central one is the median. The other two lines give an impression of the dispersion of the distribution of the estimation errors around the median. Their vertical distance represents the interquartile range: of the generated estimation errors 50% landed within these two lines for each ρ_{xu} value.

[1] This is one of the main objectives in: Davidson, R., MacKinnon, J.G., 2015. Bootstrap tests for overidentification in linear regression models. *Econometrics* 3, 825-863.

In the top-row of panels both instruments are valid. The top-left panel shows that for the examined eight cases mentioned in the legend the Sargan test shows no noteworthy size problems: the actual probability of type I errors is extremely close to the nominal significance level for all α values examined. In the top-right panel, for all eight estimators/cases represented, except OLS, the three lines are found to be almost horizontal. Thus, these distributions are hardly determined by endogeneity of x, and they suggest median unbiasedness, especially for moderate values of ρ_{xu}. On the other hand, the estimation errors of OLS seem proportional to the degree of endogeneity. Given the relatively small dispersion of OLS and KLS, the graphs show the increasing effects on the dispersion of using weaker and fewer instruments. Note that the dispersion of OLS improves for higher ρ_{xu} and is not beaten by any of the other estimators, although KLS comes close. It is striking that the KLS estimator beats all other estimators when taking both median bias and interquartile range into account. Note, though, that this is the unfeasible version of KLS, which uses full knowledge of the actual value ρ_{xu}. However, one should realize that the instrument based estimators build on assuming ρ_{z_1u} and ρ_{z_2u} both being zero, whereas in practice their true values are in fact unknowable too.

The second graph on the Sargan test shows what the effects are on its rejection probability when one of the two instruments is mildly invalid. When the valid instrument is relatively weak we see that the Sargan test will not very often detect the instrument invalidity. The situation is slightly better when the valid instrument is stronger. However, when using $\alpha = 0.05$ then instrument invalidity will be detected with probability 0.3 at most (for the sample size and correlation combinations examined), so the type II error probability is at least 0.7. The adjacent panel shows that nondetected instrument invalidity (of just $\rho_{z_1u} = 0.1$) is devastating for the estimators based on instruments, especially for the IV estimator just using the invalid instrument. The TSLS estimators based on a valid and an invalid instrument are also seriously affected and for most their interquartile range does no longer overlap with that of KLS. For ρ_{xu} small OLS is in fact to be preferred to IV or TSLS. Note that the OLS and KLS results are similar in all four rows of panels, because they are invariant to the properties of the two instruments.

In the third row of panels instrument z_2 is invalid ($\rho_{z_2u} = 0.2$), so the IV results are similar to those in the top panel. Using $\alpha = 0.5$ would lead for all cases examined to detection of the invalidity with a probability above 0.9, and above 0.5 when using $\alpha = 0.05$. The effect on the estimation errors of TSLS is more determined by the strength of the valid instrument than by the strength of the invalid instrument.

In the fourth row both instruments are invalid, and here we clearly note the perils of the Sargan test not being consistent. For some cases the rejection probability is quite high, but for two of them it hardly exceeds the nominal significance level. These are the two cases where $\rho_{z_1x} = 0.3$ and $\rho_{z_2x} = 0.4$, so the most seriously invalid instrument is also the strongest. The area in the parameter space where the Sargan test will lack power for the chosen data generating process can be derived analytically, see (B.17). The bottom-right panel dramatically undermines trust in instrument-based methods, as this shows that the TSLS estimator for these often Sargan-approved cases is very badly biased over the whole range of ρ_{xu} values. Note that the KLS results are always the most attractive in all four right-hand panels, simply because they are not based on instruments and thus do not require the doubtful approval by the Sargan test.

[Figure A.2 here]

Figure A.2 presents some results for a much smaller and for a much larger sample size than 250. We just cover the cases to be compared with the second and fourth rows of panels of Figure A.1. The size control in this simple cross-section model (not presented in the figure) was found to be close to perfect, irrespective of the sample size (whereas it has been established that size problems for the Sargan/Hansen test are serious in the context of dynamic panel data models). As is to be expected, the detection probability of instrument invalidity is generally lower in smaller samples and larger in bigger samples. For most cases it is (almost) one when $n = 2500$, but even then (and not surprisingly also for $n = 50$) for the same particular cases as in Figure A.1 the detection probability is alarmingly low, with devastating consequences for inference on β as the right-hand graphs show. Note that after a rejection by the Sargan test, producing inference on β requires a further search to find valid instruments.

[Figure A.3 here]

Next we examine the vulnerability of KLS regarding an incorrect assessment r_{xu} of the true endogeneity correlation ρ_{xu}. Figure A.3 presents the quartiles of $\hat{\beta}_{KLS}(r_{xu})$ for all compatible combinations of $\rho_{xu} = -0.9(0.1)0.9$ and $r_{xu} - \rho_{xu} = -0.3(0.1)0.3$, so that $|r_{xu}| < 1$. Results are given for $n = 50$, 250, 2500 and 25000. The four panels clearly show that the median of the various distributions of the estimation errors seems invariant to sample size. Apparently the median bias in finite samples is for $n \geq 50$ simply given by the inconsistency of KLS. This inconsistency is only zero when the correct ρ_{xu} value has been used. Of course, the sample size does have a mitigating effect on the interquartile range, which is close to zero when n is very large. We see that, roughly, when $n = 250$ and $|\rho_{xu}| \leq 0.4$ an error of ± 0.3 may give rise to a shift of the quartiles of up to about $0.5 \times \sigma_u/\sigma_x$, and about half of that for errors of ± 0.2. The latter vulnerability, although substantial, seems more limited than that of IV/TSLS when using mildly weak and/or mildly invalid instruments.

Figure A.1 Simulation results for $n = 250$; $\sigma_x/\sigma_u = 1$; and correlations (A.1)

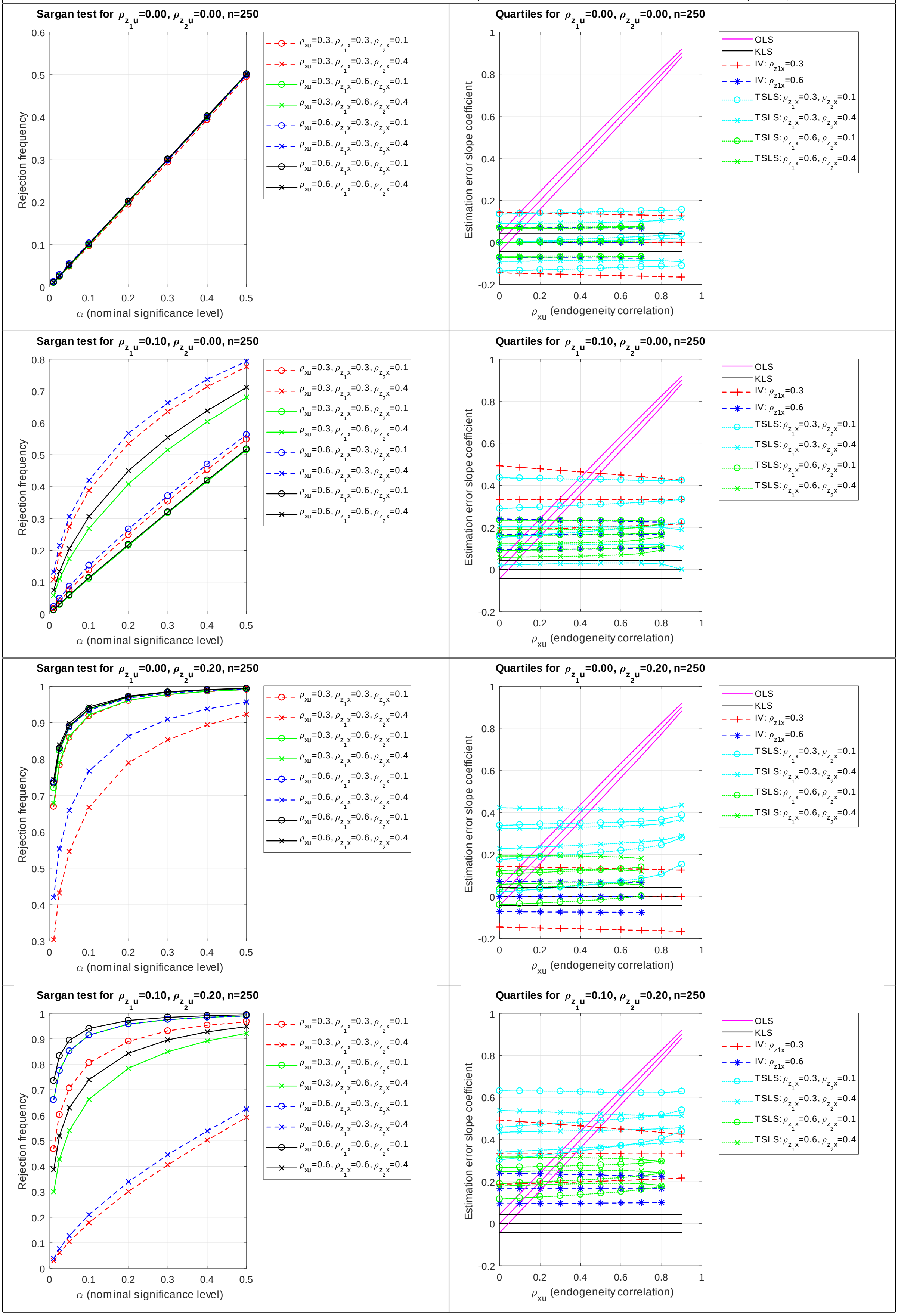


Figure A.2 Simulation results for $n = 50$, 2500; $\sigma_x/\sigma_u = 1$; and correlations (A.1)

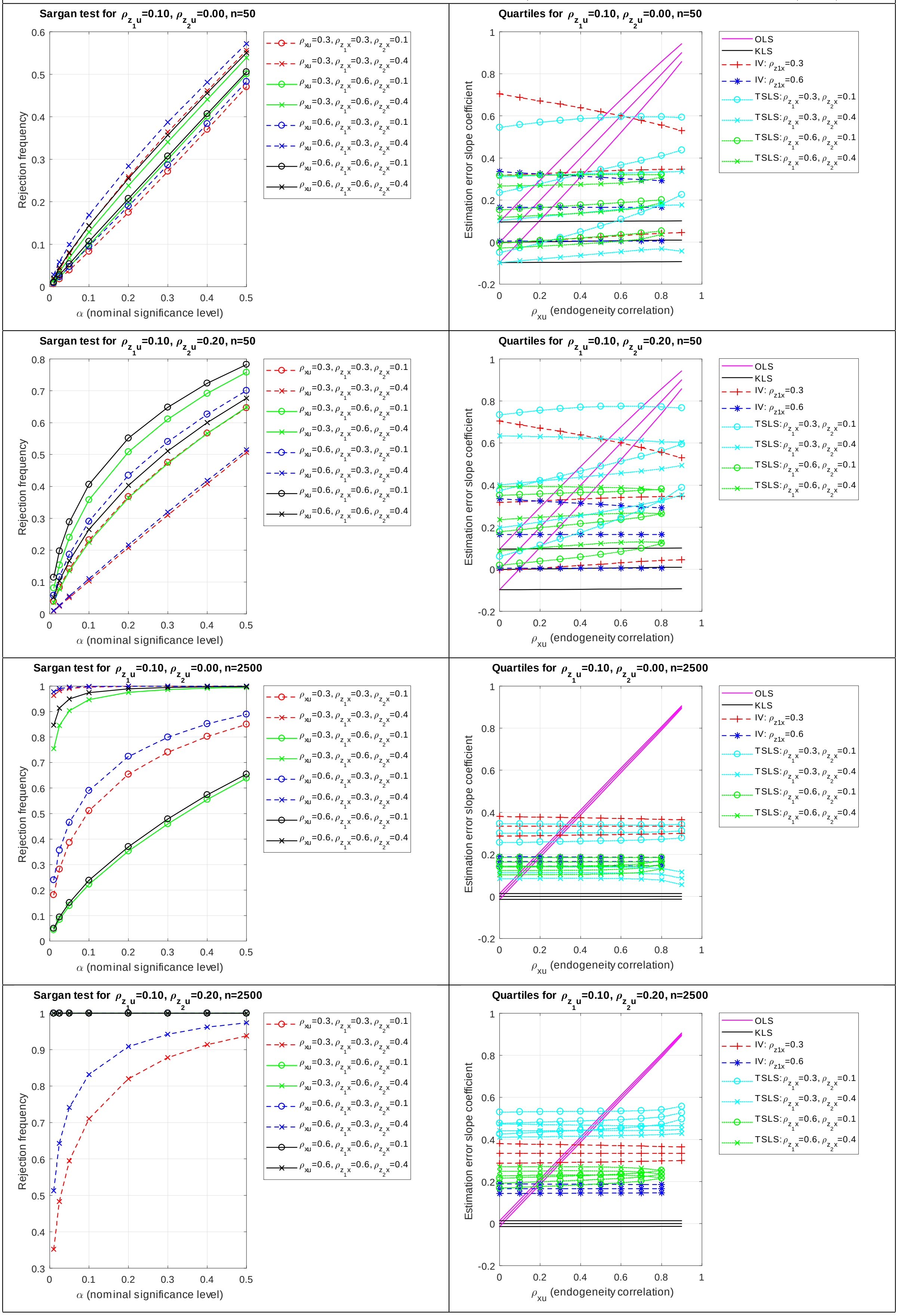
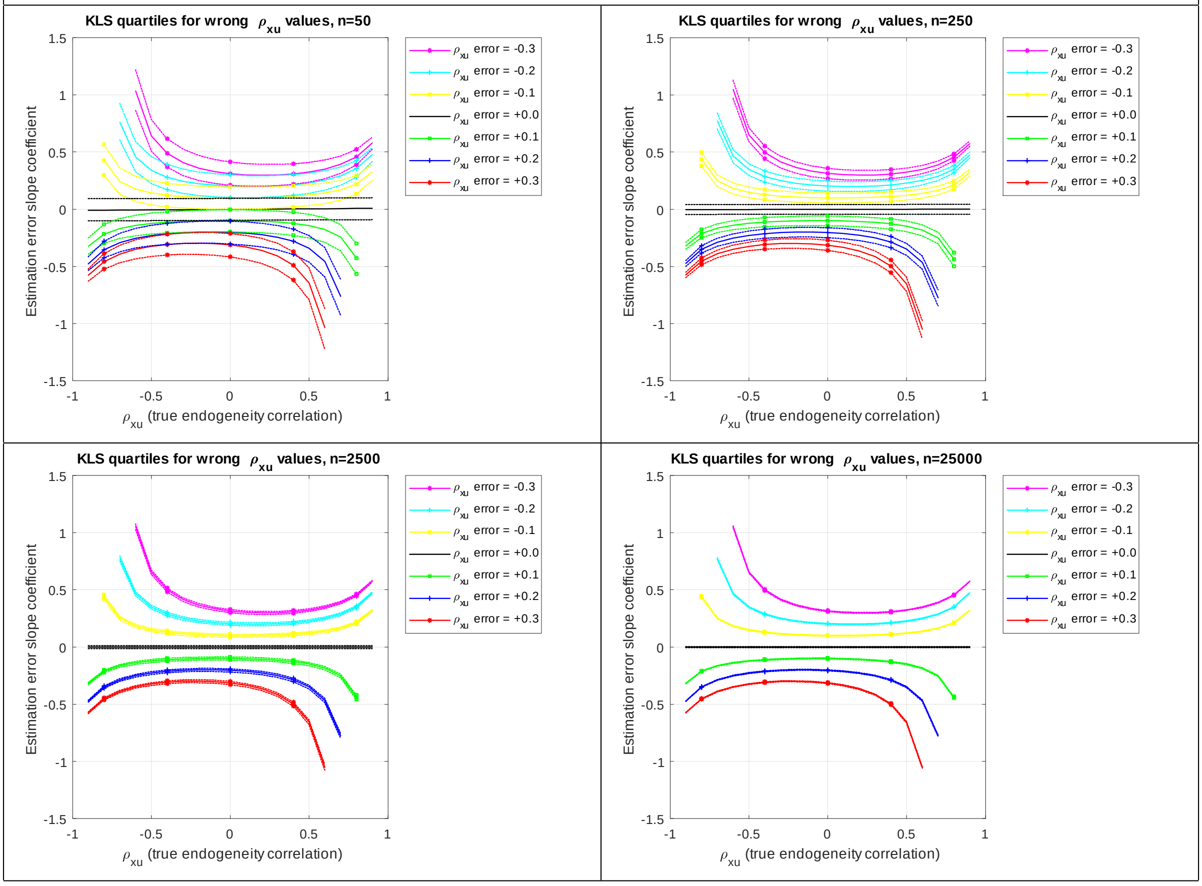


Figure A.3 KLS quantiles using a wrong ρ_{xu} for $n = 50, 250, 2500, 25000$.



B. Further details on the chosen simulation design

The Monte Carlo design used in Appendix A is defined as follows. Let ε_i, ξ_i, ζ_{i1} and ζ_{i2} be four mutually independent series ($i = 1, ..., n$) of identically distributed independent drawings with mean zero and unit variance. From these we construct the four series

$$u_i = \sigma_u \varepsilon_i \sim iid(0, \sigma_u^2), \tag{B.1}$$

$$x_i = \sigma_x[(1 - \rho_{xu}^2)^{1/2}\xi_i + \rho_{xu}\varepsilon_i] \sim iid(0, \sigma_x^2), \tag{B.2}$$

$$z_{ij} = \sigma_{z_j}(\rho_{z_j\zeta_j}\zeta_{ji} + \rho_{z_j\xi}\xi_i + \rho_{z_ju}\varepsilon_i) \sim iid(0, \sigma_{z_j}^2) \text{ for } j = 1, 2, \tag{B.3}$$

where all ρ coefficients do not exceed 1 in absolute value; moreover,

$$\rho_{z_j\zeta_j}^2 + \rho_{z_j\xi}^2 + \rho_{z_ju}^2 = 1 \text{ for } j = 1, 2. \tag{B.4}$$

Obviously, $\sigma_{xu} = \rho_{xu}\sigma_x\sigma_u$, $\sigma_{z_ju} = \rho_{z_ju}\sigma_{z_j}\sigma_u$ and $\sigma_{z_jx} = \sigma_{z_j}\sigma_x[\rho_{z_j\xi}(1 - \rho_{xu}^2)^{1/2} + \rho_{z_ju}\rho_{xu}]$, hence $\rho_{z_jx} = \rho_{z_j\xi}(1 - \rho_{xu}^2)^{1/2} + \rho_{z_ju}\rho_{xu}$, which yields

$$\rho_{z_j\xi} = (\rho_{z_jx} - \rho_{z_ju}\rho_{xu})(1 - \rho_{xu}^2)^{-1/2}, \tag{B.5}$$

for $\rho_{xu}^2 < 1$. From (B.4) we also have

$$\rho_{z_j\zeta_j} = (1 - \rho_{z_j\xi}^2 - \rho_{z_ju}^2)^{1/2}. \tag{B.6}$$

Hence, when values for $\sigma_u > 0$, $\sigma_x > 0$, $\sigma_{z_j} > 0$, $|\rho_{xu}| < 1$, $\left|\rho_{z_jx}\right| \leq 1$ and $\left|\rho_{z_ju}\right| \leq 1$ are chosen, we can generate series for u_i and x_i and find from (B.5) matching values for $\rho_{z_j\xi}$ and for $\rho_{z_j\zeta}$ from (B.6), so that series z_{i1} and z_{i2} can be generated as well. However, the choices for ρ_{xu}, ρ_{z_jx} and ρ_{z_ju} are only compatible if they yield values for $\rho_{z_j\xi}^2$ and $\rho_{z_j\zeta_j}^2$ such that $0 \leq \rho_{z_j\xi}^2 \leq 1$ and $0 \leq \rho_{z_j\zeta_j}^2 \leq 1$. This requires

$$0 \leq \frac{(\rho_{z_jx} - \rho_{z_ju}\rho_{xu})^2}{(1 - \rho_{xu}^2)} \leq 1 \tag{B.7}$$

and

$$0 \leq 1 - \frac{(\rho_{z_jx} - \rho_{z_ju}\rho_{xu})^2}{(1 - \rho_{xu}^2)} - \rho_{z_ju}^2 \leq 1. \tag{B.8}$$

The latter implies

$$0 \leq 1 - \rho_{xu}^2 - (\rho_{z_jx} - \rho_{z_ju}\rho_{xu})^2 - \rho_{z_ju}^2(1 - \rho_{xu}^2) \leq 1 - \rho_{xu}^2$$

or

$$0 \leq (1 - \rho_{xu}^2)(1 - \rho_{z_ju}^2) - (\rho_{z_jx} - \rho_{z_ju}\rho_{xu})^2 \leq 1 - \rho_{xu}^2,$$

giving the two requirements

$$(\rho_{z_j x} - \rho_{z_j u}\rho_{xu})^2 \leq (1-\rho_{xu}^2)(1-\rho_{z_j u}^2), \tag{B.9}$$

and

$$-\rho_{z_j u}^2(1-\rho_{xu}^2) - (\rho_{z_j x} - \rho_{z_j u}\rho_{xu})^2 \leq 0.$$

The latter will always be satisfied, whereas restriction (B.9) implies that (B.7) will also be satisfied.

So, by choosing values $|\rho_{xu}| < 1$, $\left|\rho_{z_j x}\right| \leq 1$ and $\left|\rho_{z_j u}\right| \leq 1$, which obey (B.9), we have two instruments with correlation

$$\begin{aligned}\rho_{z_1 z_2} &= \rho_{z_1 \xi}\rho_{z_2 \xi} + \rho_{z_1 u}\rho_{z_2 u} \\ &= (\rho_{z_1 x} - \rho_{z_1 u}\rho_{xu})(\rho_{z_2 x} - \rho_{z_2 u}\rho_{xu})(1-\rho_{xu}^2)^{-1} + \rho_{z_1 u}\rho_{z_2 u}.\end{aligned}$$

For each realization of the series u_i, x_i and z_{ij} in the simulation replications, we may first subtract their respective sample average from each observation. In that way we cover a model with one slope coefficient and an arbitrary intercept, to be estimated by OLS, KLS, IV or TSLS, because there are next to the intercept two (possibly invalid) instruments, each with a possibly non-zero arbitrary mean which has been partialled out. The dependent variable is generated by the model

$$y_i = x_i\beta + u_i. \tag{B.10}$$

This can be estimated by

$$\begin{aligned}\hat{\beta}_{OLS} &= \frac{x'y}{x'x} = \beta + \frac{x'u}{x'x}, \\ \hat{\beta}_{KLS} &= \hat{\beta}_{OLS} - \rho_{xu}\left(\frac{\hat{u}_{OLS}'\hat{u}_{OLS}}{x'x}\right)^{1/2}, \text{ where } \hat{u}_{OLS} = y - x\hat{\beta}_{OLS},\ \hat{u}_{OLS}'\hat{u}_{OLS} = u'u - \frac{(u'x)^2}{x'x}, \\ \hat{\beta}_{IV}^{(j)} &= \frac{z_j'y}{z_j'x} = \beta + \frac{z_j'u}{z_j'x},\ j = 1,2, \\ \hat{\beta}_{TSLS} &= \frac{x'Z(Z'Z)^{-1}Z'y}{x'Z(Z'Z)^{-1}Z'x} = \beta + \frac{x'P_Z u}{x'P_Z x}, \text{ with } Z = (z_1, z_2) \text{ and } P_Z = Z(Z'Z)^{-1}Z'.\end{aligned}$$

For the estimation errors we find, writing $\xi_i^* = (1-\rho_{xu}^2)^{1/2}\xi_i + \rho_{xu}\varepsilon_i$,

$$\hat{\beta}_{OLS} - \beta = \frac{\sigma_u \Sigma_i \xi_i^* \varepsilon_i}{\sigma_x \Sigma_i \xi_i^{*2}}, \tag{B.11}$$

$$\begin{aligned} \hat{\beta}_{KLS} - \beta &= \frac{\sigma_u \Sigma_i \xi_i^* \varepsilon_i}{\sigma_x \Sigma_i \xi_i^{*2}} - \rho_{xu} \left\{ \frac{\sigma_u^2 \Sigma_i \varepsilon_i^2}{\Sigma_i x_i^2} - \frac{\sigma_u^2 [\Sigma_i \varepsilon_i x_i]^2}{[\Sigma_i x_i^2]^2} \right\}^{1/2} \\ &= \frac{\sigma_u}{\sigma_x} \left\{ \frac{\Sigma_i \xi_i^* \varepsilon_i}{\Sigma_i \xi_i^{*2}} - \rho_{xu} \left[\frac{\Sigma_i \varepsilon_i^2}{\Sigma_i \xi_i^{*2}} - \frac{\Sigma_i \xi_i^* \varepsilon_i}{(\Sigma_i \xi_i^{*2})^2} \right]^{1/2} \right\}, \end{aligned} \tag{B.12}$$

$$\hat{\beta}_{IV}^{(j)} - \beta = \frac{\sigma_u}{\sigma_x} \frac{\Sigma_i (\rho_{z_j\zeta}\zeta_{ij} + \rho_{z_j\xi}\xi_i + \rho_{z_j u}\varepsilon_i)\varepsilon_i}{\Sigma_i (\rho_{z_j\zeta_j}\zeta_{ij} + \rho_{z_j\xi}\xi_i + \rho_{z_j u}\varepsilon_i)\xi_i^*} \tag{B.13}$$

$$\hat{\beta}_{TSLS} - \beta = \frac{x' P_Z u}{x' P_Z x}. \tag{B.14}$$

Because P_Z is invariant with respect to the scale of the vectors $z^{(1)}$ and $z^{(2)}$ the estimation errors of $TSLS$, like those of IV are invariant with respect to σ_{z_1} and σ_{z_2}, so without loss of generality we may fix these at value 1.

It is easily seen that all the estimation errors are also invariant regarding β and are all a multiple of σ_u/σ_x. Hence, without loss of generality we may choose in the simulations $\beta = 0$, $\sigma_{z_1} = \sigma_{z_2} = 1$ and $\sigma_x = 1$. Then the dispersion of all estimators can be regulated by varying σ_u. However, their **relative** differences will be invariant with respect to σ_u. So, by just choosing $\sigma_u = 1$ all relevant information will be obtained, through choosing relevant compatible values for the remaining design parameters: n, ρ_{xu}, $\rho_{z_j x}$ and $\rho_{z_j u}$, where the latter two determine $\rho_{z_j\xi}$ and $\rho_{z_j\zeta_j}$. Changing the sign of any of the correlations while keeping their absolute value fixed has simple (anti-)symmetric effects just on the sign of the estimation errors. Therefore we shall mostly just investigate nonnegative values for ρ_{xu}, $\rho_{z_j x}$ and $\rho_{z_j u}$.

For the TSLS residuals we find

$$\hat{u}_{TSLS} = y - \hat{\beta}_{TSLS} x = u - \frac{x' Z (Z'Z)^{-1} Z' u}{x' Z (Z'Z)^{-1} Z' x} x = u - \frac{x' P_Z u}{x' P_Z x} x, \tag{B.15}$$

and for the Sargan test statistic

$$\begin{aligned} S &= n \frac{\hat{u}_{TSLS}' Z (Z'Z)^{-1} Z' \hat{u}_{TSLS}}{\hat{u}_{TSLS}' \hat{u}_{TSLS}} \\ &= n \frac{u' P_Z u - \frac{(x' P_Z u)^2}{x' P_Z x}}{u'u - 2u'x \frac{x' P_Z u}{x' P_Z x} + x'x \frac{(x' P_Z u)^2}{(x' P_Z x)^2}} \\ &= n \frac{u' P_Z u (x' P_Z x)^2 - (x' P_Z u)^2 x' P_Z x}{u'u (x' P_Z x)^2 - 2u'x (x' P_Z u) x' P_Z x + x'x (x' P_Z u)^2}. \end{aligned} \tag{B.16}$$

It is obvious that this is invariant with respect to β and to all scale factors, because all individual terms, both in the numerator and in the denominator, are multiples of $\sigma_u^2 \sigma_x^4$.

It is well known that the Sargan test is equivalent to literally testing over-identification exclusion restrictions. In the present design, this amounts to estimating the model $y_i = \beta x_i + \delta z_{ij} + u_i$, where j is either 1 or 2, using both instruments, and next testing the significance of δ. One easily finds that the probability limit of the estimator for δ is a multiple of $\rho_{z_1x}\rho_{z_2u} - \rho_{z_2x}\rho_{z_1u}$, which is zero when

$$\rho_{z_1u}/\rho_{z_1x} = \rho_{z_2u}/\rho_{z_2x}. \tag{B.17}$$

Indeed, when running simulations with parameter values obeying (B.17) with values of ρ_{z_1u} and ρ_{z_2u} far away from zero, we always found rejection probabilities of the Sargan test similar to the nominal significance level, with TSLS inference of course being seriously corrupted.

C. Coping with omitted variables by TSLS or KLS

We consider estimating the model

$$y = X\beta + u, \tag{C.1}$$

where $X\beta = X_1\beta_1 + X_2\beta_2$, $u = \gamma + \varepsilon$ and $\gamma = X_3\beta_3$, as defined in (4.1) and (4.2). Employing an $n \times L$ instrument matrix Z with $L \geq K$, such that

$$E[Z'(\gamma + \varepsilon)] = 0 \text{ and } rank(Z'X) = K_1 + K_2, \tag{C.2}$$

yields

$$\hat{\beta}_{TSLS} = (X'P_Z X)^{-1} X'P_Z y, \text{ where } P_Z = Z(Z'Z)^{-1}Z'. \tag{C.3}$$

Under (C.2) and some further standard regularity conditions estimator (C.3) is consistent for $\beta = (\beta_1, \beta_2)'$.

Using the regressors in X_2 as instruments, so specializing $Z = (Z_1, X_2)$, will thus yield a consistent estimator for β if

$$E(Z_1'\gamma) + E(Z_1'\varepsilon) = 0 \text{ and } E(X_2'\gamma) + E(X_2'\varepsilon) = 0$$

or, substituting $\gamma = X_3\beta_3$ and assuming $E(X_2'\varepsilon) = 0$, if

$$E(Z_1'X_3)\beta_3 + E(Z_1'\varepsilon) = 0 \text{ and } E(X_2'X_3)\beta_3 = 0. \tag{C.4}$$

The conditions (C.4) can be satisfied for particular β_3 while $E(Z_1'X_3) \neq O$, $E(Z_1'\varepsilon) \neq 0$ and $E(X_2'X_3) \neq O$. However, such β_3 are then functionally related to the three covariances in (C.4), which seems an empirically most uncommon situation. Conditions (C.4) will be satisfied for any β_3 if

$$\text{(i) } E(Z_1'\varepsilon) = 0, \text{ (ii) } E(Z_1'X_3) = O, \text{ and (iii) } E(X_2'X_3) = O. \tag{C.5}$$

So, in studies where TSLS is used to obtain –despite the omission of the regressors X_3– consistent estimates of the direct effects β of all the regressors in X, one should pay attention explicitly to these three conditions. Condition (i) entails that the regressors Z_1 are rightly excluded from the theory model (4.1). Condition (ii) is explicitly addressed when arguing or testing whether Z_1 establishes valid external instruments for the empirical model (C.1), which is the case when the variables in Z_1 are all uncorrelated with γ. Condition (iii) of (C.5) requires that variables X_2, which are exogenous in theory model (4.1), will still be exogenous in underspecified model (C.1). Obviously, they are not, when they are correlated with γ.

The above should make clear that, if (i) from (C.5) is satisfied, this does certainly not imply that (ii) will be satisfied as well. External instruments Z_1 should be chosen such that they do not have a direct effect on y, but through their required association with X_1 (to avoid underidentification) just an indirect effect (provided $\beta_1 \neq 0$). In the applications, Z_1 contains variables such as presence of a college in the neighborhood or quarter of birth. If these are correlated with γ, for instance because both Z_1 and γ depend on variables from X_2, then requirement (ii) is at risk. Condition (iii) will be at risk if the subjective and possibly discriminative assessment (by those who determine an individual's wage) of the multidimensional characteristic "ability" is blurred by the objective characteristics in X_2.

The serious difficulties to ascertain (ii) and (iii) can partly be avoided if one is willing to redefine the variables X_3, and give up obtaining a consistent estimator for the direct effects β_2 of X_2. This can be achieved as follows. In the linear world adopted here, any correlation between γ and X_2 can be removed from γ, by defining

$$X_3^* = (I - P_{X_2})X_3, \text{ where } P_{X_2} = X_2(X_2'X_2)^{-1}X_2'. \tag{C.6}$$

Then

$$\begin{aligned} X_2\beta_2 + X_3\beta_3 &= X_2\beta_2 + X_3^*\beta_3 + X_2(X_2'X_2)^{-1}X_2'X_3\beta_3 \\ &= X_2\beta_2^* + \gamma^*, \end{aligned} \tag{C.7}$$

where $\beta_2^* = \beta_2 + (X_2'X_2)^{-1}X_2'\gamma$, $\gamma^* = X_3^*\beta_3$, $X_2'X_3^* = O$, so $X_2'\gamma^* = 0$. By defining

$$Z_1^* = (I - P_{X_2})Z_1 \tag{C.8}$$

as well, we can make use of the fact that the matrices $Z = (Z_1, X_2)$ and $Z^* = (Z_1^*, X_2)$ span the same column space, so that $P_Z = P_{Z^*}$, implying that the TSLS coefficient estimator for model

$$y = X_1\beta_1 + X_2\beta_2^* + (\gamma^* + \varepsilon) \tag{C.9}$$

does not change when we replace Z_1 by Z_1^*. From (C.5) it follows that these estimators are consistent for β_1 and β_2^*, provided

$$\text{(i) } E(Z_1^{*\prime}\varepsilon) = 0, \text{ (ii) } E(Z_1^{*\prime}X_3^*) = O, \text{ and (iii) } E(X_2'X_3^*) = O. \tag{C.10}$$

Now (iii) is satisfied by definition, and (i) still entails that the external instruments should be correctly excluded from the causal model (4.1). Condition (ii) focusses now on requiring that the instruments Z_1^* are also validly excluded from the redefined empirical model (C.9), which means that one can argue that the variables in Z_1 and X_3, after projecting out any association they may have with the variables X_2, which yields Z_1^* and X_3^*, results in mutually uncorrelated variables.

The above analysis highlights that as a rule instrument matrix Z should not contain the variables X_1 from X for which one wants to assess exclusively their direct effect β_1 on y, in isolation from any indirect effects X_1 may have on y. The remaining variables X_2 may be used as internal instruments, but their estimated coefficients will represent both their direct effect and any indirect effect they may have via the omitted variables X_3.

For identification one needs at least as many external instruments Z_1 as there are variables X_1. These external instruments Z_1 should meet three criteria. Two with respect to their validity, namely: (a) Z_1 should have no direct causal effect on y; and (b), after removing from Z_1, or from omitted component γ, any mutual linear association with X_2, they should be uncorrelated, i.e. $Z_1'(I - P_{X_2})\gamma = 0$. And Z_1 should meet a third criterion, namely regarding relevance, being: (c) one should consider $Z_1'(I - P_{X_2})X_1$, because the external instruments Z_1 should be sufficiently strong, so the correlations of the variables in Z_1 with those of X_1, after netting out their relation with X_2, should not be small.

Selecting which of the regressors belongs to X_1 or to X_2, is also a crucial issue in the instrument-free approach. It determines which elements of ρ_{xu} can simply be set at zero, and for which credible numerical (interval) assumptions have to be made. Of course, the other crucial issues in a TSLS context, indicated in the previous paragraph by (a), (b) and (c), are irrelevant for an instrument-free approach. When avoiding assumptions on instruments, these have to be replaced by an assumption on credible numerical values of $\rho_{x^{(1)}u}$, which is associated with the covariance $E(X_1'u) = E(X_1'\gamma^*)$. The j^{th} element of $\rho_{x^{(1)}u}$ can be obtained by dividing the corresponding element of $E(X_1'u)$ by $n\sigma_u\sigma_j$, where σ_j is the standard deviation of the j^{th} regressor in X_1.