Econometrics and Statistics 25 (2023) 1–22

Contents lists available at ScienceDirect

Econometrics and Statistics

journal homepage: www.elsevier.com/locate/ecosta

Instrument-free inference under confined regressor endogeneity and mild regularity

Jan F. Kiviet

[a] Amsterdam School of Economics, University of Amsterdam, PO Box 15867, 1001 NJ Amsterdam, The Netherlands
[b] Department of Economics, Stellenbosch University, South Africa

ARTICLE INFO

Article history:
Received 18 May 2021
Revised 27 December 2021
Accepted 28 December 2021
Available online 3 January 2022

JEL classification:
C12
C13
C21
C22
C26

Keywords:
endogeneity robust inference
exclusion restrictions test
replication studies
sensitivity analysis

ABSTRACT

The instrument-free approach adopts flexible bounds on the correlation between regressors and disturbances, instead of exploiting instruments presupposing their asymptotic uncorrelatedness with the model errors. Earlier findings on such instrument-free inference methods assumed the observations to be mesokurtic and independent and identically distributed. Adopting substantially weaker regularity, this alternative to Two-Stage Least-Squares (TSLS) is developed and simulated for general linear regression models, permitting time-dependent regressors with heterogeneous excess kurtosis. Replicating three prominent empirical studies TSLS is shown to be based on untenable exclusion restrictions, whereas instrument-free inference can arguably be more credible, while potentially producing narrower confidence intervals than (weak-instrument robust) TSLS.

1. Introduction

For rather specific regression models with endogenous explanatories Kiviet (2020) develops alternative inference techniques not requiring the use of instrumental variables. Instead of strict orthogonality assumptions on instrumental variables and the model errors, it requires bounds on the possible nonorthogonality of errors and regressors. Then, as long as the actual endogeneity respects these specified bounds, asymptotically valid instrument-free inference on coefficients can be produced. This inference is asymptotically conservative (actual asymptotic significance level is smaller than nominal significance level) to a degree increasing with the distance between the chosen bounds. For the sake of simplicity, it was assumed in the earlier derivations that for each subject in the sample its observed regressor variables and realized disturbance are: (i) mesokurtic (have the same kurtosis as the normal distribution), and (ii) are independently and identically distributed (iid). Evidently, (i) is not realistic in most practical situations, and (ii) excludes most time-series applications.

Here more general results are obtained, valid under milder regularity conditions, closely matching those adopted when standard TSLS (two-stage least-squares) is applied to cross-section data or to time-series relationships modeled by linear simultaneous autoregressive distributed lag models. In that popular context, for an instrument-free inconsistency corrected least-squares estimator its normal limiting distribution is obtained. This estimator is set-identified provided the unknown

E-mail address: j.f.kiviet@uva.nl

https://doi.org/10.1016/j.ecosta.2021.12.008
2452-3062/

endogeneity correlations of the regressors are confined to some postulated set. Varying this set discloses the sensitivity of inference regarding endogeneity, and provides much more flexibility than the claimed strict orthogonality of instruments and errors required for TSLS analysis. In the form of exclusion restrictions tests, this orthogonality can also be tested by the new techniques in a much more rigorous way than is possible by classic over-identification restrictions tests. By simulation it is demonstrated that asymptotic instrument-free inference can be remarkably accurate in finite samples, and often – though depending on the distance between the chosen bounds – yields much narrower and at the same time more trustworthy confidence intervals than standard or robustified TSLS, especially when instruments are weak, and certainly when they are invalid. The instrument-free estimator and its estimated variance are straight-forward functions of the sample data on the variables of the model and of the adopted endogeneity correlations. In three replication studies the new techniques are applied to the data used in earlier instrument-based publications. This reveals that some of the assumptions on which these studies have been built seem doubtful. For all case studies examined a new perspective regarding their empirical findings results.

When regressors are endogenous they are correlated with the model errors, which leads to serious bias of least-squares estimators, irrespective of the size of the sample. In such situations one usually reverts to applying method of moment estimators, which are built on the assumed orthogonality of so-called instrumental variables and the model errors. However, resulting inference may have serious impediments, associated with the proclaimed validity and relevance of the employed instrumental variables. See, for instance, Bound et al. (1995), Murray (2006, 2017), Kiviet and Niemczyk (2012), Andrews et al. (2019a), Andrews et al. (2019b), Doko Tchatoka and Dufour (2020), and many of the further references in those studies. The validity or orthogonality of instruments and errors can only very partially be vindicated on the basis of statistical evidence; the major justification of instrument validity depends as a rule just on subject matter specific rhetoric persuasiveness. External instruments can only be valid if they do not have a direct effect on the dependent variable, so their exclusion from the regression relationship should be true. Though, at the same time, in order to be relevant, they should have a relatively strong indirect effect on the dependent variable through their association with the endogenous regressors. If this association is weak then method of moment estimators may be as seriously biased as least-squares estimators are, and they will also be harmed by having an unattractive large dispersion and possibly a distinctly nonnormal distribution.

For the alternative instrument-free inference methods validity and relevance of instruments are not an issue, self-evidently. Their primary impediment is actual credibility regarding the chosen range of likely values of the degree of endogeneity of the individual regressors. A narrow range may yield seemingly more efficient but also unmistakably less credible inference; wide ranges will be more credible but will as a rule result in less pronounced statistical conclusions, as the applications will illustrate. These applications indicate that the new techniques provide a useful sensitivity analysis of instrument-based findings, revealing their vulnerability regarding presumptuous orthogonality conditions and robustness claims. They may also provide very attractive autonomous alternative inference on regression coefficients. The instrument-free technique is computationally not very demanding and available within Stata, see Kripfganz and Kiviet (2021). It involves the screening of asymptotically valid confidence intervals for regression coefficients over all compatible and credible values of the nuisance parameters established by the correlations between regressors and disturbances.

Identification of the parameters of single equations, or of the complete system to which they belong, has usually been obtained by exploiting normalization and exclusion restrictions or more general coefficient restrictions, see Koopmans et al. (1950) and Fisher (1959). Achieving identification by exploiting restrictions on the covariances of the disturbances has been introduced by Fisher (1963) and extended by Wegge (1965); more recently it has been specialized to exploiting heteroskedasticity for identification, see Lewbel (2012). In the approach developed here point-identification of a single structural equation is based on restricting yet other parameters, namely the correlations between regressors and disturbances. At first sight this may seem odd, because in current practice the actual sign and magnitude of these correlations are usually disregarded, except for the case of them being zero or not. Simulations in Kiviet and Niemczyk (2012) show, however, that these correlations are nuisance parameters which may, like the strength of instruments, seriously distort the finite sample distribution of TSLS based estimators and test statistics. Therefore, and because TSLS estimators are built on statistically unverifiable preconditions, as a rule statistical inference on the actual value of these endogeneity correlations will be highly unreliable. Below, however, it will be indicated that in many practical situations the theoretical arguments used to suggest a particular model specification implicitly entail assumptions on the sign and likely magnitude of endogeneity correlations. Moreover, the procedures suggested here do not require assumptions on the specific true values of these correlations. To achieve set-identification, as defined in Bontemps and Magnac (2017), they just require to specify intervals which should enclose these true correlation values.

In Section 2, it is first indicated how the role of the actual endogeneity correlation parameter is usually downplayed in current practice and how it can be made more explicit by reformulating the model. How this correlation depends on other parameters in the three basic empirically relevant situations that may give rise to endogeneity of regressors is spelled out in Appendix A, included in the Supplementary Material available in the online version of this article. The derived expressions facilitate to make credible assumptions on the likely sign and magnitude of any endogeneity of regressors. Next, the reformulated model, in which regressor endogeneity is directly parametrized, is used in the derivation of the asymptotic validity of the alternative instrument-free inference methods. Because these derivations are rather cumbersome for a model with an arbitrary number of regressors, from which probably more than one is endogenous, Section 3 first considers the simple model with just one regressor for which all matrix algebra can be avoided. This regressor may be endogenous, nonnormal, and also dependent on its own past, as is often the case for regressors in time-series relationships. In this section the instrument-free methods are also positioned with respect to more sophisticated instrument-based approaches, such as methods to cope with invalid exclusion restrictions (Nevo and Rosen, 2012; Conley et al., 2012) and to handle moment inequalities. The oversimplified model of Section 3 provides a helpful stepping stone towards the presentation of the results in Section 4 for single nonnormal possibly dynamic multiple linear regression models with an arbitrary number of endogenous explanatories. The technical derivations of the results presented in Sections 3 and 4 can be found in Appendices B through E included in the Supplementary Material. Section 5 provides simulation evidence on the accuracy of the proposed methods in finite samples. Those who are primarily interested in the actual practical achievements of the new approach may immediately jump to Section 6. This contains three empirical replication studies, where standard and non-standard instrumental variable based inferences are supplemented with instrument-free results. The latter reveal frailties in and provide alternatives to the earlier findings. Finally, Section 7 concludes by reviewing and confronting the aims, approach, achievements and hurdles of the instrument-based and instrument-free techniques.

2. Characteristics of the model

In social science, and especially in economics and business, relationships are usually modeled on the basis of so-called observational data (not stemming from controlled experiments) and under specification uncertainty. Then explanatory variables may be contemporaneously correlated with the model error. Like the dependent variable (the regressand), which is unavoidably contemporaneously correlated with the errors, such regressors are labeled endogenous. Apart from testing whether these correlations are zero or not, usually little attention is being paid to the actual non-zero values they may have. The classic instrument-based estimation techniques are developed using a formulation of the model that leaves the endogeneity correlations implicit. The resulting estimators are asymptotically invariant regarding the value of these correlations.

In the instrument-free approach, on the other hand, the endogeneity correlations are –like the model coefficients– parameters of interest. Therefore, in the derivations to follow, a reformulation of the model will be used in which these correlations are made explicit. Unlike the model coefficients, the endogeneity correlations will not be estimated. In order to trace their likely signs and actual value ranges, it is derived how they are functionally related to other parameters in the three fundamental situations that give rise to endogeneity of regressors, namely: (a) simultaneity, (b) errors in explanatories, and (c) wrongly omitted explanatories. One may argue that a fourth possibility is joint occurrence of autoregressive disturbances and lagged dependent variable regressors in a time-series regression. However, such endogeneity can be resolved in principle by including in the regression further lags of all regressors. So, in essence, this case is already covered by (c). Appendix A provides formulas by which in the various cases a rough assessment of the likely sign and magnitude of regressor endogeneity can be made.

Without the need to specify any reduced form equations, instrument-free inference methods can be developed for the slope coefficients β of the single multiple regression model

$$y_i = x_i'\beta + u_i, \tag{2.1}$$

where $K \times 1$ zero-mean regressor vector x_i may contain endogenous regressors, and

$$x_i \sim (0, \Sigma_{xx}) \text{ and } u_i \sim iid(0, \sigma_u^2). \tag{2.2}$$

In the derivations to follow the zero-mean assumption will prove to be convenient, while not leading to a loss of generality.

The linearity in β of the model implies that the endogenous regressors can be decomposed into two contemporaneously uncorrelated additive components (see Appendix A). One of these components is predetermined or exogenous, and the other is endogenous. The latter is simply a multiple of the error term and a factor which is proportional to the endogeneity correlation. This decomposition is denoted as

$$x_i = \xi_i + \lambda u_i, \tag{2.3}$$

where random $\xi_i \sim (0, \Sigma_{\xi\xi})$ and deterministic λ are both $K \times 1$ vectors, with $E(\xi_i u_i) = 0$ and $E(x_i u_i) = \lambda\sigma_u^2$. For $j, k = 1, ..., K$ denote $(\Sigma_{xx})_{jk} = \sigma_{jk} = E(x_{ij}x_{ik})$, $\sigma_k^2 = \sigma_{kk}$, $\sigma_k = |\sigma_{kk}|^{1/2}$, $\Sigma_x = \text{diag}(\sigma_1, ..., \sigma_K)$ and $\rho_k = E(x_{ik}u_i)/(\sigma_k\sigma_u)$, where $\sigma_u = \left|(\sigma_u^2)^{1/2}\right|$. Then, denoting $\rho_{xu} = (\rho_1, ..., \rho_K)'$,

$$\lambda = \sigma_u^{-1}\Sigma_x\rho_{xu}, \text{ with } \lambda_k = E(x_{ik}u_i)/\sigma_u^2 = \rho_k\sigma_k/\sigma_u. \tag{2.4}$$

The OLS (ordinary least-squares) estimator for β, given by $\hat{\beta}_{OLS} = (X'X)^{-1}X'y$, where $X = (x_1, ..., x_n)'$ is an $n \times K$ matrix and $y = (y_1, ..., y_n)'$ an $n \times 1$ vector, has (invoking the law of large numbers) probability limit given by

$$\begin{aligned} \text{plim}\, \hat{\beta}_{OLS} &= \beta + (\text{plim}\, n^{-1}X'X)^{-1}\, \text{plim}\, n^{-1}X'u = \beta + \sigma_u^2\Sigma_{xx}^{-1}\lambda \\ &= \beta + \sigma_u\Sigma_{xx}^{-1}\Sigma_x\rho_{xu}. \end{aligned} \tag{2.5}$$

Hence, in general, each element of $\hat{\beta}_{OLS}$ is inconsistent (thus biased, irrespective of the size of the sample) if any element of ρ_{xu} (or of λ) is nonzero. Each nonzero element undermines one of the K moments $E(x_i u_i) = \sigma_u^2\lambda = \sigma_u\Sigma_x\rho_{xu}$ to establish a valid orthogonality condition. Instead of invoking alternative valid orthogonality conditions based on often fanciful external instruments, (2.5) will be used to obtain an inconsistency-corrected least-squares estimator.

3. Instrument-free inference in a single regressor model

Kiviet (2013, 2016, 2020) focussed primarily on iid mesokurtic cross-section samples. Here time-series regressions with time-dependence between more heterogenous sample data should be covered too. The full proof of instrument-free inference under mild regularity for general linear multiple possibly dynamic regression models with some endogenous explanatory variables will be presented in Section 4. In this section the major issues will be presented first for model (2.1) when $K = 1$.

Thus, the single regressor may be serially dependent, i.e.

$$E(x_i x_t) \neq 0 \text{ for } \forall i, t, \tag{3.1}$$

and it may not only be correlated with the current but also with lagged disturbances, i.e.

$$E(x_i u_t) \neq 0 \text{ for } 1 \leq t \leq i \leq n, \tag{3.2}$$

while still assuming

$$E(x_i u_t) = 0 \text{ for } 1 \leq i < t \leq n. \tag{3.3}$$

Hence, although the regressor could be exogenous, namely when $E(x_i u_t) = 0\ \forall i, t$, or predetermined when $E(x_i u_t) = 0\ \forall i \leq t$, it could also be endogenous. Since it is adopted that

$$E(u_i \mid x_{i-1}, ..., x_1, u_{i-1}, ..., u_1) = 0 \text{ for } i \geq 2, \tag{3.4}$$

u_i is an innovation with respect to its own past and that of x_i, whereas x_i could depend on current and past u_i and on past x_i. So, the sample observations $\{y_i, x_i; i = 1, ..., n\}$ are not necessarily independent. For the sake of simplicity, though, they will be assumed to be identically distributed. Note that assumption (3.4) easily matches with iid cross-section applications. It does also with time-series regressions, in case the endogeneity stems from simultaneity or from errors in regressors, provided u_i and ε_i and η_i (introduced in Appendix A) are serially uncorrelated indeed. However, in case of wrongly omitted time-series regressors, the assumption that $u_i = \beta_2 x_i^{(2)} + \varepsilon_i$ is serially uncorrelated would require that omitted regressor $x_i^{(2)}$ is serially uncorrelated too. In many empirical time-series applications this seems highly unlikely.

In line with Section 2, it is assumed that the scalar regressor x_i can be decomposed as $x_i = \xi_i + \lambda u_i \sim (0, \sigma_x^2)$, where $\xi_i \sim (0, \sigma_\xi^2)$ and $E(u_i, \xi_t) = 0$ for $t = 1, ..., i$. So, component ξ_i is predetermined but could in fact be strictly exogenous. The endogeneity of the regressor can be expressed by the constant correlation

$$\rho_{xu} = \lambda \sigma_u / \sigma_x. \tag{3.5}$$

From $\sigma_x^2 = \sigma_\xi^2 + \lambda^2 \sigma_u^2 = \sigma_\xi^2 + \rho_{xu}^2 \sigma_x^2$ it follows that $\sigma_\xi^2 = (1 - \rho_{xu}^2)\sigma_x^2$. In this one-regressor model (all summations that follow are over the range $i = 1, ..., n$)

$$\hat{\beta}_{OLS} = (\Sigma x_i^2)^{-1} \Sigma x_i y_i = \beta + (\Sigma x_i^2)^{-1} \Sigma x_i u_i, \tag{3.6}$$

and (2.5) specializes into

$$\hat{\beta}_{OLS} = \beta + n^{-1} \Sigma x_i u_i / (n^{-1} \Sigma x_i^2) \to \beta + \rho_{xu} \sigma_u / \sigma_x. \tag{3.7}$$

where $\to$ indicates convergence in probability. Hence, $\hat{\beta}_{OLS}$ is inconsistent when the degree of endogeneity ρ_{xu} is nonzero.

Assuming for the moment that ρ_{xu} were known, then a consistent estimator of β could be obtained, provided consistent estimators for σ_u and σ_x can be found too. Since

$$\hat{\sigma}_x^2 = n^{-1} \Sigma x_i^2 \to \sigma_x^2, \tag{3.8}$$

$\hat{\sigma}_x \to \sigma_x$. From $\hat{u}_i = y_i - x_i \hat{\beta}_{OLS} = u_i - x_i(\hat{\beta}_{OLS} - \beta) = u_i - x_i \Sigma x_i u_i / \Sigma x_i^2$, it follows that $\Sigma \hat{u}_i^2 = \Sigma u_i^2 - (\Sigma x_i u_i)^2 / \Sigma x_i^2$, thus

$$n^{-1} \Sigma \hat{u}_i^2 = n^{-1} \Sigma u_i^2 - (n^{-1} \Sigma x_i u_i)^2 / n^{-1} \Sigma x_i^2 \to \sigma_u^2 - \rho_{xu}^2 \sigma_u^2 = \sigma_u^2 (1 - \rho_{xu}^2),$$

so

$$\hat{\sigma}_u^2(\rho_{xu}) = (1 - \rho_{xu}^2)^{-1} n^{-1} \Sigma \hat{u}_i^2 \to \sigma_u^2, \tag{3.9}$$

giving $\hat{\sigma}_u(\rho_{xu}) \to \sigma_u$. The above suggests the estimator

$$\hat{\beta}_{KLS}(\rho_{xu}) = \hat{\beta}_{OLS} - \rho_{xu} \hat{\sigma}_u(\rho_{xu}) / \hat{\sigma}_x \to \beta, \tag{3.10}$$

which in previous studies was addressed as kinky least-squares (KLS). It is consistent, although infeasible, unless ρ_{xu} is really known.

Appendix B shows that under the above indicated mild regularity conditions the limiting distribution of KLS is normal and has a variance which is not only determined by ρ_{xu}, but also by the kurtosis parameters $\kappa_u = E(u_i^4)/\sigma_u^4$ and $\kappa_x = E(x_i^4)/\sigma_x^4$, viz.

$$n^{1/2}[\hat{\beta}_{KLS}(\rho_{xu}) - \beta] \to \mathcal{N}[0, \theta(\rho_{xu}, \kappa_u, \kappa_x) \sigma_u^2 / \sigma_x^2], \tag{3.11}$$

$$\text{with}\quad \theta(\rho_{xu},\kappa_u,\kappa_x)=\frac{4+(\kappa_u+\kappa_x-14)\rho_{xu}^2-2(\kappa_u-5)\rho_{xu}^4}{4(1-\rho_{xu}^2)^2}.$$

So, the limiting variance is invariant regarding the sign of ρ_{xu}. Also any skewness of the distributions of x_i and u_i, and their fifth and higher-order moments, have no effect. When $\rho_{xu}=0$ KLS specializes to OLS, which has limiting distribution $\mathcal{N}(0,\sigma_u^2/\sigma_x^2)$, irrespective of the third and higher-order moments of the data. For the special case $\kappa_x=\kappa_u=3$ (which covers normality) and arbitrary ρ_{xu}, the KLS limiting distribution corresponds to that of OLS too, because $\theta(\rho_{xu},3,3)=1$. For small absolute values of ρ_{xu} the limiting variance is not very much affected by how much κ_u and κ_x differ from 3 (called their excess kurtosis); for both κ_x and κ_u smaller than 10 and $|\rho_{xu}|\leq 0.3$, factor $\theta(\rho_{xu},\kappa_u,\kappa_x)$ is smaller than 1.35, and for $|\rho_{xu}|\leq 0.5$ it does not exceed 2.5, giving a multiplicative boost to the KLS standard error, as obtained under zero excess kurtosis, of at most 1.16 and 1.58 respectively. However, it can also be shown that $\lim_{\rho_{xu}^2\uparrow 1}\theta(\rho_{xu},\kappa_u,\kappa_x)=\infty$, unless $\kappa_x=\kappa_u=3$.

The qualities of KLS in comparison to IV are illustrated by the following. When estimating the one-regressor model by IV the strongest possible valid though infeasible instrument would obviously be latent variable ξ_i. Its strength is expressed by $Corr(\xi_i,x_i)=\sigma_\xi/\sigma_x=(1-\rho_{xu}^2)^{1/2}$. Hence, the more serious the endogeneity is, the weaker even the strongest possible valid instrument must be. And, on the other hand: when a valid instrument is really very strong, this implies that the endogeneity cannot be very substantial at the same time. The variance of the limiting distribution of $\hat{\beta}_{IV(\xi)}=\Sigma(\xi_i y_i)/\Sigma(\xi_i x_i)$ is $\sigma_u^2/\sigma_\xi^2=(1-\rho_{xu}^2)^{-1}\sigma_u^2/\sigma_x^2$, whereas for KLS (in case $\kappa_u=\kappa_x=3$) this is σ_u^2/σ_x^2, which is never dominated by IV. It can easily be derived that in the simple one-regressor model, only for substantial excess kurtosis and limited endogeneity, infeasible but most efficient IV can be more efficient than infeasible KLS, namely when $\kappa_u+\kappa_\xi>10$ and $\rho_{xu}^2<1-4/(\kappa_u+\kappa_\xi-6)<1$.

For testing hypotheses on – or constructing confidence intervals for – β a consistent estimator of $\theta(\rho_{xu},\kappa_u,\kappa_x)\sigma_u^2/\sigma_x^2$ is required. Obviously, $Var[\hat{\beta}_{KLS}(\rho_{xu})]$ can be approximated in finite samples by

$$\widehat{Var}[\hat{\beta}_{KLS}(\rho_{xu})]=n\theta(\rho_{xu},\hat{\kappa}_u(\rho_{xu}),\hat{\kappa}_x)\hat{\sigma}_u^2(\rho_{xu})/\hat{\sigma}_x^2, \tag{3.12}$$

which uses the consistent kurtosis estimators

$$\left.\begin{aligned}\hat{\kappa}_u(\rho_{xu})&=n^{-1}\Sigma[y_i-x_i\hat{\beta}_{KLS}(\rho_{xu})]^4/\hat{\sigma}_u^4(\rho_{xu}),\\ \hat{\kappa}_x&=n^{-1}\Sigma x_i^4/\hat{\sigma}_x^4.\end{aligned}\right\} \tag{3.13}$$

In order to realize $\widehat{Var}[\hat{\beta}_{KLS}(0)]=\widehat{Var}(\hat{\beta}_{OLS})$, it makes sense to replace $\hat{\sigma}_u^2(\rho_{xu})$ in (3.12) by the asymptotically equivalent though degrees of freedom corrected expression

$$s_u^2(\rho_{xu})=\frac{1}{(1-\rho_{xu}^2)}\frac{\Sigma(y_i-x_i'\hat{\beta}_{OLS})^2}{n-K}, \tag{3.14}$$

where $K=1$, or larger when predetermined regressors have been partialled out. In Section 5 the accuracy in finite samples will be verified of the approximation $\hat{\beta}_{KLS}(\rho_{xu})\overset{a}{\sim}\mathcal{N}[\beta,\theta(\rho_{xu},\hat{\kappa}_u(\rho_{xu}),\hat{\kappa}_x)s_u^2(\rho_{xu})/x'x]$.

For true value ρ_{xu} and a chosen level α (with $0<\alpha<1$),

$$\mathcal{C}(\beta;\rho_{xu},\alpha)=[\hat{c}_{\alpha/2}(\rho_{xu}),\hat{c}_{1-\alpha/2}(\rho_{xu})] \tag{3.15}$$

represents the $(1-\alpha)100\%$ asymptotic confidence interval for β, where for $0<p<1$

$$\hat{c}_p(\rho_{xu})=\hat{\beta}_{KLS}(\rho_{xu})+\zeta_p\{\widehat{Var}[\hat{\beta}_{KLS}(\rho_{xu})]\}^{1/2},$$

with ζ_p denoting the p^{th} quantile of the standard normal distribution.

For any chosen real interval $\mathcal{I}=[r^L,r^U]$, where $-1<r^L\leq r^U<1$, such that $\rho_{xu}\in\mathcal{I}$, the interval

$$\mathcal{C}(\beta;\mathcal{I},\alpha)=[\hat{c}^L_{\alpha/2}(\mathcal{I}),\hat{c}^U_{1-\alpha/2}(\mathcal{I})],\ \text{with} \tag{3.16}$$

$$\hat{c}^L_{\alpha/2}(\mathcal{I})=\min_{r\in\mathcal{I}}\hat{c}_{\alpha/2}(r)\ \text{and}\ \hat{c}^U_{1-\alpha/2}(\mathcal{I})=\max_{r\in\mathcal{I}}\hat{c}_{1-\alpha/2}(r),$$

is an asymptotically conservative $(1-\alpha)\times 100\%$ confidence interval for β. This confidence interval also enables a conservative asymptotic test on β, assuming $\rho_{xu}\in\mathcal{I}$ holds. Interval (3.16) includes all real values β_0 for which at significance level α the null hypothesis $\beta=\beta_0$ should not be rejected against two-sided alternative.

The occurrence of the factor $1-\rho_{xu}^2$ in the denominators of $\theta(\rho_{xu},\kappa_u,\kappa_x)$ and $s_u^2(\rho_{xu})$ indicate that the closer one chooses r^L to -1 and/or r^U to +1 the wider the conservative confidence interval will be. Hence, as is always the case for realizing identification: some genuinely discriminative initial information should be adopted. Unlimited robustness of inference regarding regressor endogeneity is illusionary. Likewise, instrument-based inference cannot be made fully robust regarding unlimited weakness or invalidity of the instruments. However, for $r^L=r^U=\rho_{xu}$ the KLS estimator is point-identified, and for r^L and r^U not arbitrarily close to their extremes the identification set is bounded.

Some might like to replace the role of the interval $\mathcal{I}$ by adopting a prior distribution for ρ_{xu}. Such a Bayesian approach would form a natural alternative for the method advocated in Kraay (2012), where a prior is adopted for the degree of invalidity of external instruments. The approach based on $\mathcal{I}=[r^L,r^U]$ seems the most natural and a very flexible alternative to the standard frequentist instrument-based approach, where a strictly-zero assumption is adopted regarding the (asymptotic) correlation of instruments and errors. In the applications of Section 6 it will be shown that just one graph can serve to accommodate any choice regarding $\mathcal{I} = [r^L, r^U]$. A Bayesian approach would face the extra complication to specify a realistic prior for ρ_{xu}.

More alternatives for standard instrumental variables techniques have been suggested. Nevo and Rosen (2012) derive set estimates under assumptions on the signs and relative magnitudes of the simultaneity and possible instrument invalidity. Conley et al. (2012) augment the model with the external instruments and make assumptions on their coefficients (which would be zero under correct exclusion). This allows either frequentist or Bayesian methods to obtain inference allowing for instrument invalidity. Unlike the KLS approach, these alternatives still use instruments and may therefore suffer from bias and poor asymptotic approximations due to weakness of the instruments. An illustration in Kripfganz and Kiviet (2021) compares KLS with the frequentist variant of Conley et al. (2012).

If a chosen interval $\mathcal{I}$ could be represented by moment inequalities, this might enable to confront the inference technique suggested here with the achievements of the moment inequality literature; see, for instance, Andrews and Soares (2010) and Bugni et al. (2017). However, the conversion of two-sided restrictions on the dimensionless correlation ρ_{xu} into one-sided inequalities in not scale-free covariances between errors and regressors is not self-evident. Moreover, to date the moment inequality literature seems to have been developed just for iid samples, whereas here temporal dependence of regressors is allowed.

4. Instrument-free inference for general linear models

The results of the foregoing section (with $K = 1$) will be generalized here to instrument-free inference for multiple (either static or dynamic) linear regression models with an arbitrary number of endogenous regressors. First, in subsection 4.1, models for which all regressors are supposed to have the same known kurtosis parameter will be considered. This restriction greatly simplifies how to obtain the limiting distribution of the KLS estimator. Next, in subsection 4.2, it is shown how this result enables finding an upperbound to the limiting variance of the KLS estimator in models where each regressor may have arbitrary and unknown kurtosis.

4.1. Models where all regressors have equal kurtosis

Regarding the distribution of vector $(x_i', u_i)'$ for linear multiple regression model (2.1) it is assumed:

KLS Assumptions under equivalent regressor kurtosis

(a) First and second moments: The vectors $\{(x_i', u_i)'; i = 1, ..., n\}$ are identically (but not necessarily independently) distributed with zero mean and with second moments $E(x_i x_i') = \Sigma_{xx}$, $E(u_i^2) = \sigma_u^2$ and $E(x_i u_i) = \sigma_{xu}$ which are all finite. Scalar σ_{jk} denotes the typical element of Σ_{xx} and $\sigma_j = \left|(\sigma_{jj})^{1/2}\right|$ for $j, k = 1, ..., K$ with $\Sigma_x = \text{diag}(\sigma_1, ..., \sigma_K)$, hence σ_{xu} has typical element $\rho_j \sigma_j \sigma_u$, where $\sigma_u = \left|(\sigma_u^2)^{1/2}\right|$*and* ρ_j *is the typical element of vector $\rho_{xu} = \Sigma_x^{-1}\sigma_{xu}/\sigma_u$;*

(b) Fourth moments: $E(u_i^4) = \kappa_u \sigma_u^4$ and $E(x_{ik}^4) = \kappa_x \sigma_k^4$ for $k = 1, ..., K$, where κ_u and κ_x are both finite and (by definition) not smaller than unity;

(c) Time dependence: As $E(u_i u_t) = 0$ $\forall i \neq t$ and $E(x_i u_t) = 0$ for $t > i = 1, ..., n$ and arbitrary otherwise, the disturbances are serially uncorrelated and individual regressors may be either exogenous, predetermined or endogenous.

Note that no reduced form equations have been specified. That all regressors have zero mean will be helpful in the proof, but, as argued in Theorem 2 of Kiviet (2020), the findings will also apply to the slope coefficients of models with nonzero mean regressors that include an intercept.

The sample equivalents of Σ_{xx}, Σ_x^2 and Σ_x are given by $S_{xx} = n^{-1}\Sigma_{i=1}^n x_i x_i'$, S_x^2 (the matrix just containing the main diagonal of S_{xx}), and by the positive definite diagonal matrix S_x (for which $S_x S_x = S_x^2$) respectively. Appendix D, which uses some basic underlying derivations collected in Appendix C, proves:

KLS Theorem under equivalent regressor kurtosis

Under the above assumptions infeasible estimator

$$\hat{\beta}_{KLS}(\rho_{xu}) = \hat{\beta}_{OLS} - \hat{\sigma}_u(\rho_{xu}) S_{xx}^{-1} S_x \rho_{xu}, \tag{4.1}$$

where $\hat{\sigma}_u^2(\rho_{xu}) = \hat{\sigma}_{u,OLS}^2/(1 - \rho_{xu}' S_x S_{xx}^{-1} S_x \rho_{xu})$ and $\hat{\sigma}_{u,OLS}^2 = n^{-1}\Sigma_{i=1}^n (y_i - x_i'\hat{\beta}_{OLS})^2$, has limiting distribution

$$n^{1/2}[\hat{\beta}_{KLS}(\rho_{xu}) - \beta] \xrightarrow{d} \mathcal{N}[0, \sigma_u^2 V(\rho_{xu}, \kappa_u, \kappa_x)], \tag{4.2}$$

where $V(\rho_{xu}, \kappa_u, \kappa_x) = \Sigma_{xx}^{-1}\Theta\Sigma_{xx}^{-1}$, with

$$\begin{aligned} \Theta = {} & \Sigma_{xx} - (\Sigma_{xx}R^2 + R^2\Sigma_{xx}) + \theta^{-1}(\Phi - \Sigma_{xx}R^2\Sigma_{xx}^{-1}\Phi - \Phi\Sigma_{xx}^{-1}R^2\Sigma_{xx}) \\ & - \ 0.25(\kappa_u - 1)\theta^{-1}[R^2\Phi + \Phi R^2 - \theta^{-1}(1 - 2\rho_{xu}'R\Sigma_x\Sigma_{xx}^{-1}\Sigma_x R\rho_{xu})\Phi] \\ & + \ 0.25(\kappa_x - 1)(I + \theta^{-1}\Phi\Sigma_{xx}^{-1})\Sigma_x^{-1}R(\Sigma_{xx} \circ \Sigma_{xx})R\Sigma_x^{-1}(I + \theta^{-1}\Sigma_{xx}^{-1}\Phi), \end{aligned} \tag{4.3}$$

which uses $\theta = 1 - \rho_{xu}'\Sigma_x\Sigma_{xx}^{-1}\Sigma_x\rho_{xu} > 0,\ \Phi = \Sigma_x\rho_{xu}\rho_{xu}'\Sigma_x,\ R = \mathrm{diag}(\rho_1, ..., \rho_K)$ *and where* $\circ$ *denotes the Hadamard (element by element) product.*

This theorem for dependent data specializes for $\kappa_u = \kappa_x = 3$ to Theorem 1 of Kiviet (2020), which assumes independence of the mesokurtic sample data.

Estimator $\hat{\sigma}_u^2(\rho_{xu})$ – and thus KLS – only makes sense when $\rho_{xu}'S_xS_{xx}^{-1}S_x\rho_{xu} < 1$. This is in line with

$$\begin{aligned}\operatorname{plim}\hat{\sigma}_{u,OLS}^2 &= \operatorname{plim} n^{-1}u'[I - X(X'X)^{-1}X']u \\ &= \sigma_u^2 - \sigma_u^2\rho_{xu}'(\operatorname{plim} S_xS_{xx}^{-1}S_x)\rho_{xu} = \sigma_u^2\theta > 0.\end{aligned}$$

Thus, for vector ρ_{xu} one should only choose values that are confined to an ellipsoid enclosed by a unit sphere. For $K = 1$ this simply implies $\rho_1^2 < 1$, whereas for $K > 1$ values close to 1 for the absolute value of elements of vector ρ_{xu} may be infeasible. For $K = 2$ one finds $\rho_{xu}'\Sigma_x\Sigma_{xx}^{-1}\Sigma_x\rho_{xu} = (\rho_1^2 - 2\rho_1\rho_2\omega_{12} + \rho_2^2)/(1 - \omega_{12}^2) < 1$, where ω_{12} is the correlation between the two regressors. For $\rho_2 = 0$ this implies $\rho_1^2 < 1 - \omega_{12}^2 < 1$, which conveys that when the two regressors have a nonzero correlation, whereas one of them is predetermined, the other one cannot have an endogeneity correlation arbitrarily close to 1.

Multiple regression models with just one endogenous regressor occur very frequently in applied economics. Therefore it is useful to consider the special case where only the first column x_1 of the regressor matrix $X = (x_1, X_2)$ is endogenous, whereas $K \geq 2$. So, apart from ρ_1, all other elements of ρ_{xu} are zero. Then it follows from the theorem that the KLS estimator of β_1 can be expressed as

$$\begin{aligned}\hat{\beta}_{1,KLS}(\rho_1) &= \hat{\beta}_{1,OLS} - \rho_1\hat{\sigma}_{u,OLS}(1 - \rho_1^2e_1'S_xS_{xx}^{-1}S_xe_1)^{-1/2}e_1'S_{xx}^{-1}e_1(n^{-1}\Sigma x_{i1}^2)^{1/2} \\ &= \hat{\beta}_{1,OLS} - \rho_1[f_1/(1 - f_1\rho_1^2)]^{1/2}n^{1/2}SE(\hat{\beta}_{1,OLS}).\end{aligned} \tag{4.4}$$

Here $SE(\hat{\beta}_{1,OLS}) = \hat{\sigma}_{u,OLS}[e_1'(X'X)^{-1}e_1]^{1/2}$ is the usual (but when $\rho_1 \neq 0$ naive) estimate for the standard deviation of $\hat{\beta}_{1,OLS}$. Factor

$$f_1 = e_1'(X'X)^{-1}e_1\Sigma x_{i1}^2 \geq 1 \tag{4.5}$$

is also known (when $\rho_1 = 0$) as the 'variance inflation factor', being the ratio of $Var(\hat{\beta}_{1,OLS})$ and its hypothetical value if all regressors X_2 happened to be orthogonal to x_1. Scalar estimator $\hat{\beta}_{1,KLS}(\rho_1)$ is now only defined for

$$\rho_1^2 < 1/f_1 \leq 1. \tag{4.6}$$

Suppose for the moment that the values of κ_u and κ_x are known. Then an asymptotically valid estimator of the variance of (4.4), derived in Appendix E, is

$$\widehat{Var}[\hat{\beta}_{1,KLS}(\rho_1)] = s_u^2(\rho_1)\frac{4 - 8\rho_1^2 + (\kappa_u + \kappa_x - 6)\rho_1^2f_1 - 2(\kappa_u - 5)\rho_1^4f_1^2}{4(1 - \rho_1^2f_1)^2}\frac{f_1}{\Sigma_ix_{i1}^2}. \tag{4.7}$$

Using the degrees of freedom corrected $s_u^2(\rho_{xu})$ when estimating the variance of $\hat{\beta}_{KLS}(\rho_{xu})$ and the uncorrected $\hat{\sigma}_{u,OLS}$ in the expression for $\hat{\beta}_{KLS}(\rho_{xu})$ itself is deliberate, because in simulations these choices proved to be preferable in (very) small samples. Variance (4.7) increases with κ_x, and because $\rho_1^2f_1 - 2\rho_1^4f_1^2 = \rho_1^2f_1(1 - 2\rho_1^2f_1)$ (4.7) increases with κ_u if $\rho_1^2 < 0.5/f_1$ and decreases if $0.5/f_1 < \rho_1^2 < 1/f_1$. The dependence of the variance estimate on both kurtosis terms is of order ρ_1^2 and thus will be moderate for relatively small $|\rho_1|$.

4.2. Models where the regressors have unknown arbitrary kurtosis

For the much more general case, with an arbitrary number of endogenous regressors, and unknown and not necessarily equivalent kurtosis of all the regressor variables, the variance of KLS coefficient estimator (4.1) can, for given value of ρ_{xu}, be estimated conservatively by

$$\widehat{Var}[\hat{\beta}_{KLS}(\rho_{xu})] = ns_u^2(\rho_{xu})S_{xx}^{-1}\hat{\Theta}S_{xx}^{-1}, \tag{4.8}$$

where $\hat{\Theta}$ is obtained by replacing in expression (4.3) for Θ the matrices Σ_{xx} and Σ_x by S_{xx} and S_x respectively, and κ_u and κ_x by

$$\left.\begin{aligned}\hat{\kappa}_u(\rho_{xu}) &= n^{-1}\Sigma[y_i - x_i'\hat{\beta}_{KLS}(\rho_{xu})]^4/\hat{\sigma}_u^4(\rho_{xu}), \\ \hat{\kappa}_x &= \max_{j=1,...,K} n^{-1}\Sigma x_{ij}^4/\hat{\sigma}_{x_j}^4.\end{aligned}\right\} \tag{4.9}$$

The Schur Theorem on Hadamard products implies that the contribution to Θ of the term involving κ_x is positive-semidefinite. Hence, by taking for κ_x the maximum of the K individual kurtosis estimates it is avoided to underestimate the asymptotic variance. In calculations to follow, the actual contributions to the KLS variance of the two terms involving kurtosis will prove to be relatively small.

In the next three sub-subsections the following topics are addressed: (1) the actual generation of general feasible instrument-free inference, (2) some effects of partialling out regressors, and (3) how to implement tests for the invalidity of instruments without adopting implicitly the validity of identifying instruments.

4.2.1. Feasible instrument-free inference

Feasible asymptotically conservative (meaning cautious by securing that asymptotically type I error probabilities will never exceed the chosen significance level α) instrument-free inference regarding β on the basis of a test statistic $T(\rho_{xu})$ can now be obtained as follows. Test statistic $T(\rho_{xu})$, which will be a function of $\hat{\beta}_{KLS}(\rho_{xu})$, $\widehat{Var}[\hat{\beta}_{KLS}(\rho_{xu})]$ and of the null, should have a known limiting distribution under the null hypothesis. Then let $p_T(r)$ denote the p-values of statistic $T(r)$ for $K \times 1$ vectors r from a dense grid of vectors in a chosen region $\mathcal{I} \subset \mathbb{R}^K$, such that $r'S_xS_{xx}^{-1}S_xr < 1$, $\forall r \in \mathcal{I}$. Now assuming $\rho_{xu} \in \mathcal{I}$, the null hypothesis should be rejected if $\forall r \in \mathcal{I}$ one finds $p_T(r) \leq \alpha$, and not rejected if $\forall r \in \mathcal{I}$ one finds $p_T(r) > \alpha$; otherwise, if some $p_T(r)$ values exceed and some (for different r values in $\mathcal{I}$) do not exceed α, the test is inconclusive over $\mathcal{I}$. Note that it will be possible to construct a range of regions, say $\mathcal{I}_h$ (for $h = 1, 2, ...$), such that the test is conclusive over each separate subregion.

4.2.2. Partialling out some predetermined regressors

Often primary interest is in estimating (or testing a linear restriction on) just a subset, say β_1, of the K coefficients β. Suppose that the regressors and corresponding coefficients can be decomposed such that $X = (X_1, X_2)$ and $\beta_1' = (\beta_1', \beta_2')$ and that all endogenous regressors belong to $n \times K_1$ matrix X_1, possibly with some predetermined regressors as well. Hence, all regressors X_2 have zeroes in their corresponding elements of vector ρ_{xu}. It is well known from partitioned regression that vector $\hat{\beta}_{1,OLS} = H\hat{\beta}_{OLS}$, where $H = (I_{K_1}, O)$, can also be obtained by regressing $M_2y = y^*$ on $M_2X_1 = X_1^*$, where $M_2 = I - X_2(X_2'X_2)^{-1}X_2'$. Since the sum of squared residuals of the regressions of y on X and of y^* on X_1^* are equivalent, also $\hat{\sigma}^2_{u,OLS} = \hat{\sigma}^2_{u^*,OLS}$, where $u^* = M_2u$. This is in agreement with $\text{plim}\, n^{-1}u^{*\prime}u^* = \text{plim}\, n^{-1}u'M_2u = \text{plim}\, n^{-1}u'u$ from which it follows that $\sigma^2_{u^*} = \sigma^2_u$. This, in combination with,

$$\text{plim}\, n^{-1}X_1^{*\prime}u^* = \text{plim}\, n^{-1}X_1'[I - X_2(X_2'X_2)^{-1}X_2']u = \text{plim}\, n^{-1}X_1^{*\prime}u = \text{plim}\, n^{-1}X_1'u$$

yields $S_{x_1^*}\rho_{x_1^*u^*} = S_{x_1}\rho_{x_1u}$. Therefore, using a well-know result for the inverse of a partitioned symmetric matrix, $HS_{xx}^{-1}S_x\rho_{xu} = S^{-1}_{x_1^*x_1^*}S_{x_1}\rho_{x_1u} = S^{-1}_{x_1^*x_1^*}S_{x_1^*}\rho_{x_1^*u^*}$, and also $\rho_{xu}'S_xS_{xx}^{-1}S_x\rho_{xu} = \rho'_{x_1^*u^*}S_{x_1^*}S^{-1}_{x_1^*x_1^*}S_{x_1^*}\rho_{x_1^*u^*}$, thus $\hat{\sigma}_u(\rho_{xu}) = \hat{\sigma}_{u^*}(\rho_{x_1^*u^*})$, giving

$$\begin{aligned}\hat{\beta}_{1,KLS}(\rho_{xu}) &= H\hat{\beta}_{KLS}(\rho_{xu}) = \hat{\beta}_{1,OLS} - \hat{\sigma}_u(\rho_{xu})HS_{xx}^{-1}S_x\rho_{xu} \\ &= \hat{\beta}_{1,OLS} - \hat{\sigma}_{u^*}(\rho_{x_1^*u^*})S^{-1}_{x_1^*x_1^*}S_{x_1^*}\rho_{x_1^*u^*} = \hat{\beta}_{1,KLS}(\rho_{x_1^*u^*}). \end{aligned} \tag{4.10}$$

Hence, when the focus is just on β_1, its KLS results can also be obtained from the regression of y^* on X_1^*, under the understanding that for full correspondence of the KLS coefficient estimates the vector ρ_{x_1u} has to be replaced then by $\rho_{x_1^*u^*} = S^{-1}_{x_1^*}S_{x_1}\rho_{x_1u}$. Note that each individual element of $\rho_{x_1^*u^*}$ cannot be smaller than the corresponding element of ρ_{x_1u}, because $X_1'X_1 - X_1^{*\prime}X_1^* = X_1'X_2(X_2'X_2)^{-1}X_2'X_1$ is positive-semidefinite.

Result (4.10) can be useful to deal more satisfactorily with kurtosis of the regressors. Partialling out as many predetermined regressors (including any dummy variables) as possible then requires to make an assessment only of the maximum of the kurtosis of the K_1 variables in X_1^*. Remember that the KLS results suppose that the intercept has been partialled out already.

In the model with just one endogenous regressor, after partialling out all predetermined regressors, inference on the coefficient of the endogenous regressor can directly be obtained on the basis of the kurtotis of the single variable M_2x_1. Then taking the maximum of all K kurtosis estimates has been avoided, and the associated "conservativeness" problem circumvented. Note, though, that KLS inference will be conservative anyhow, due to not knowing ρ_{xu} and employing an interval $\mathcal{I}$.

4.2.3. Testing any subset of exclusion restrictions

An intriguing feature of tests based on KLS is that they allow verifying the validity of instruments by directly testing exclusion restrictions. Since KLS estimates are identified by some non-orthogonality conditions and not just by classic orthogonality conditions as in TSLS, each classic identifying restriction associated with an external instrument (not just the over-identifying ones!) can be tested, either on its own or in groups. Fundamental problems with such (Sargan/Hansen) tests are discussed in Kiviet (2017) and Kiviet and Kripfganz (2021). Let $y_i = \beta_1'x_{i1} + \beta_2'x_{i2} + u_i$, where $\rho_{x_2u} = 0$, with the variables in $K_1 \times 1$ vector x_{i1} possibly endogenous. For method of moments estimation at least K_1 external but valid instruments are required. Let $K_3 \times 1$ vector x_{i3} contain $K_3 \geq 1$ candidate external instruments. Augmenting the model and estimating $y_i = \beta_1'x_{i1} + \beta_2'x_{i2} + \beta_3'x_{i3} + u_i^*$ by KLS over a credible subspace $\mathcal{I}$ of postulated values r_1 for ρ_{x_1u}, and then testing the exclusion restrictions $\beta_3 = 0$ on the basis of $\hat{\beta}_{3,KLS}(r_1)$ may either endorse (or refute) the acceptability of variables x_{i3} as valid external instruments $\forall r_1 \in \mathcal{I}$ or be inconclusive. Note that this concerns all $K_3 \geq 1$ candidate external instruments, and not just the $K_3 - K_1 \geq 1$ potentially overidentifying ones, as in a Sargan test.

In the applications of the KLS instrument validity test to follow, a peculiar and confusing phenomenon emerges, namely: always the p-value of the exclusion restriction test is 1, or very close to 1, for r_1 close to the TSLS-based estimator of the endogeneity correlation estimate $\hat{\rho}_{x_1u}$. At first sight this seems to suggest that instruments x_3 are valid, especially for ρ_{x_1u} vectors close to $\hat{\rho}_{x_1u}$. Note, however, that the latter estimate is based on assuming validity of the very same instruments which are still under test.

In Appendix F it is demonstrated that interpreting these high p-values as supporting instrument validity entails a fallacy. It is shown that, when $K_1 = K_3 = 1$ and x_3 is a valid instrument, estimator $\hat{\beta}_{3,KLS}(\rho_1)$ evaluated at the true value ρ_1 tends to zero, as it should. This yields a high p-value for the test of $\beta_3 = 0$. Unfortunately, though, when x_3 is an invalid instrument, then estimator $\hat{\beta}_{3,KLS}(\hat{\rho}_1)$, evaluated at inconsistent estimator $\hat{\rho}_1 = n^{-1}x_1'\hat{u}_{TSLS}/(\hat{\sigma}_1\hat{\sigma}_{u,TSLS})$, tends to zero too. Hence, the test lacks power to detect instrument invalidity for postulated values r_1 which are close to endogeneity estimates obtained by invalid instruments. Therefore, one better just uses the test to unveil values ρ_1 for which external instruments seem invalid, due to low p-values, rather than claiming validity of the instruments in an area around $\hat{\rho}_1$, for which p-values of this test will never be small.

5. Simulation results on the accuracy of KLS inference

By executing controlled experiments it will be assessed whether the actual distribution of infeasible KLS (which uses the true ρ_{xu}) is well approximated in finite samples by the derived limiting distribution, and whether it behaves favorably in comparison to IV/TSLS. Only if both are the case, it seems worthwhile to further examine the actual achievements of feasible KLS inference (based on exploiting some putative endogeneity correlation set $\mathcal{I}$). For such simulation analyses first simple but representative families of DGPs (data generating processes) have to be designed. For simple Gaussian cross-section models reassuring results have already been obtained in Kiviet (2013, 2020). So, here it suffices to focus on the effects of nonnormality (and more in particular of regressors having different kurtosis) and of temporal dependence of the regressors.

In the next subsection the accuracy of KLS in the iid single regressor cross-section model of Section 3 will be examined, when the regressor and disturbance have unknown excess kurtosis or are skew. Next, an exogenous regressor will be added to the model and the consequences for inference will be examined when evaluating the variance formula of the KLS Theorem upon substituting for κ_x the maximum of the kurtosis estimates (4.9) of the two regressors. In subsection 5.2 the actual and asymptotic distribution of the KLS estimator in a dynamic time-series model with additional regressors will be investigated, and compared with OLS and TSLS, but now all the time sticking to cases where the variables are normal. All simulation results are based on 10^6 replications.

5.1. Results for nonnormal cross-section models

First the actual and approximated KLS distribution in finite samples will be examined for simple iid model (2.1) with $K = 1$. Drawings u_i and ξ_i from the standard normal $\mathcal{N}(0, 1)$ and from transformed Student(ν) and transformed $\chi^2(\nu)$ distributions are used, in order to examine the effects of excess kurtosis and skewness. Here ν indicates degrees of freedom. When η_i is Student(ν) with $\nu > 4$ then standardized drawings $\eta_i/[\nu/(\nu-2)]^{1/2}$, to be indicated by $St^*(\nu)$, have zero mean, unit variance, kurtosis $3 + 6/(\nu-4)$ and are symmetric. When $\psi_i \sim \chi^2(\nu)$ standardized drawings $(\psi_i - \nu)/(2\nu)^{1/2}$, indicated by $Chi^*(\nu)$, have zero mean, unit variance, kurtosis $3 + 12/\nu$, whereas the skewness is $(8/\nu)^{1/2}$. Next to the fully mesokurtic case, where $\kappa_u = \kappa_x = \kappa_\xi = 3$, situations will be considered where $\kappa_u = 9$ and/or $\kappa_\xi = 9$ by using drawings from $St^*(5)$ and $Chi^*(2)$. The latter has skewness 2.

Focus will be on the estimation errors $\hat{\beta}_{KLS} - \beta$, which are invariant regarding the actual value of β. Thus, without loss of generality, one may set $\beta = 0$ in the DGP. All findings are easily comparable by scaling such that $\sigma_u = \sigma_x = 1$, which requires $\sigma_\xi^2 = \sigma_x^2 - \lambda^2\sigma_u^2 = 1 - \rho_{xu}^2$. Although in theory ρ_{xu}^2 could have any value between 0 and 1, mostly just relatively small values will be considered for the following three reasons. Firstly, because it is expected that ρ_{xu}^2 will often have a moderate value in practice (see Appendix A); secondly, because a large ρ_{xu}^2 forces σ_ξ^2 to be small, implying that any available valid instruments are necessarily weak, which brings IV/TSLS in a disadvantageous position regarding KLS; and thirdly, to limit space.

In Figure 5.1 the actual and approximated KLS distributions are compared for 5 different combinations of underlying u_i and ξ_i distributions, all evaluated for $n = 100$, and choosing $\rho_{xu} = 0.2$ (left-hand panels) or $\rho_{xu} = 0.4$ (right-hand panels). Presented are the simulated actual densities (top-row panels) of the estimation errors, their asymptotic approximations (mid-row panels) given by $\mathcal{N}(0, n^{-1}[1 + \rho_{xu}^2(\kappa_u + \kappa_\xi - 6)/4])$, see Appendix B, and the discrepancies between the cumulative distributions of the actual and the approximated distributions (bottom-row panels). When both u_i and ξ_i are normal (black solid lines) the asymptotic approximation is found to be rather accurate for both ρ_{xu} values, as had been established already in Kiviet (2013, 2020) as far as tail probabilities are concerned. From the present results it follows that when $\kappa_u + \kappa_\xi$ increases up to 18, the accuracy of the asymptotic approximation to represent the actual distribution in finite samples gets notably worse, and nonsymmetry leads to some further deterioration of the accuracy of the asymptotic approximation at $n = 100$. However, the discrepancies occur especially in the central part of the distribution, and less in the tail areas. The latter are of course the major concern regarding size control of tests and coverage probabilities of confidence regions. From further calculations (not depicted) it was found that for $n = 500$ the symmetric asymptotic approximation is much more satisfactory.

Table 5.1 considers at $n = 100$ the same five combinations of u_i and ξ_i distributions. It gives for $\rho_{xu} = 0.0(0.2)0.6$ the resulting κ_x values, following from κ_ξ, κ_u and ρ_{xu}, and it presents specific results regarding KLS variance and actual coverage probability (CP) of one- and two-sided nominal 95% confidence intervals. The three variance assessments (asymptotic approximation, actual, estimated) are defined in the caption of the table. The effects of excess kurtosis on the actual variance

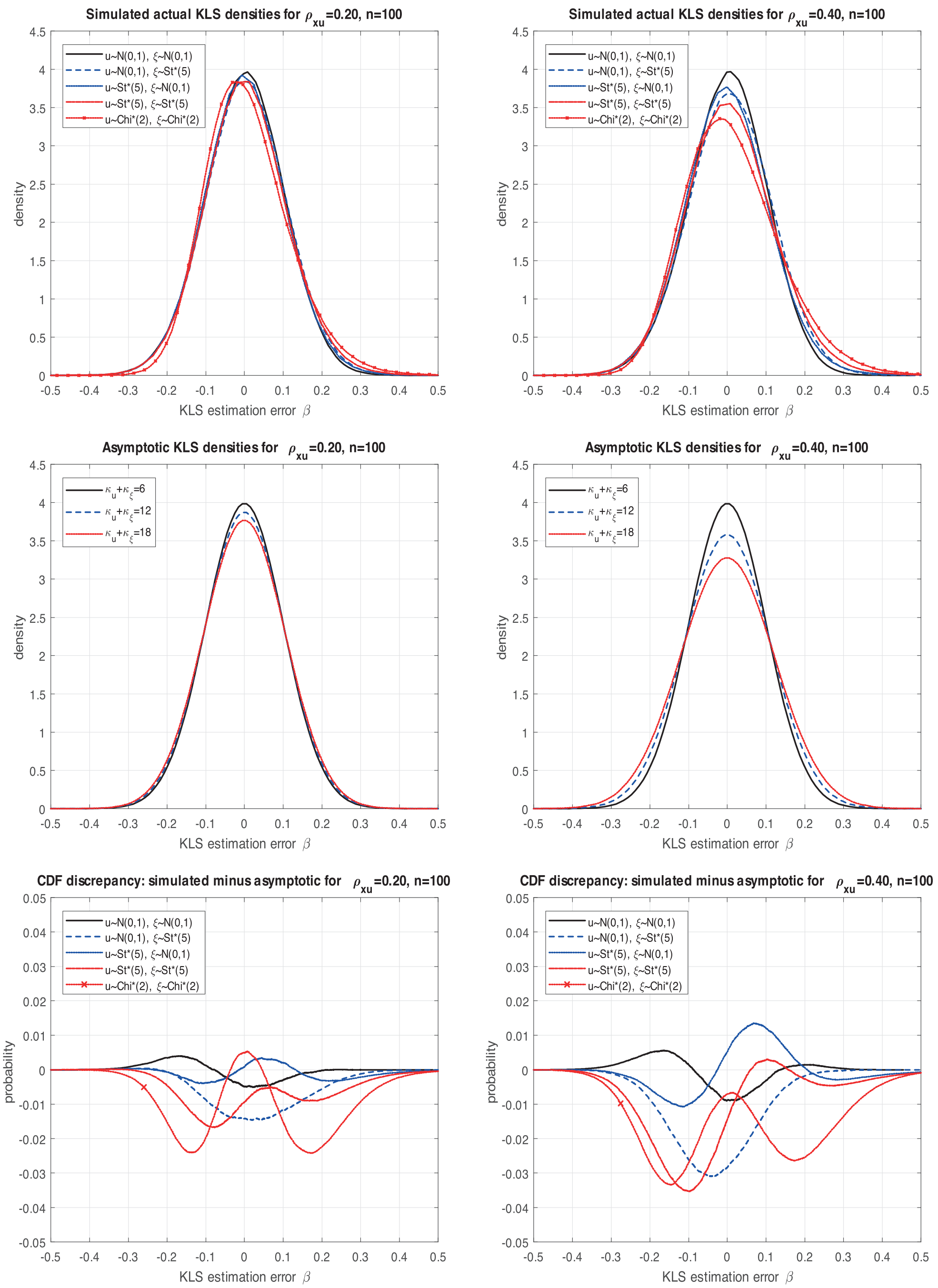

Fig. 5.1. KLS distribution for model (2.1) and (non)normal iid series u_i and ξ_i

of $\hat{\beta}_{KLS}(\rho_{xu})$ are found to be substantial and increasing with $\rho_{xu} > 0$. Nonsymmetry aggravates these effects. Although the first-order asymptotic approximation to the variance is invariant regarding skewness, the actual and estimated variance are found to increase due to positive skewness. For $\rho_{xu} > 0$ and positive excess kurtosis the estimated variance $\widehat{Var}[\hat{\beta}_{KLS}(\rho_{xu})]$ systematically underestimates the true variance. Therefore, it does not surprise that two-sided confidence intervals (constructed using percentiles from the Student distribution) are slightly liberal: they have true coverage probability (estimated from the Monte Carlo with standard error not exceeding 0.0003) a little smaller than the nominal level. For $\rho_{xu} > 0$ and symmetric u and ξ the one-sided intervals have a tendency to be slightly liberal when closed at the right-hand, and slightly conservative when closed at the left-hand side. This is aggravated when both u and ξ have positive skewness. For $\rho_{xu} = 0$

Table 5.1
$Var[\hat{\beta}_{KLS}(\rho_{xu})]$ and coverage probability (CP) in the static model with $K = 1$

	$u \sim N(0,1)$ $\xi \sim N(0,1)$	$u \sim N(0,1)$ $\xi \sim St^*(5)$	$u \sim St^*(5)$ $\xi \sim N(0,1)$	$u \sim St^*(5)$ $\xi \sim St^*(5)$	$u \sim Chi^*(2)$ $\xi \sim Chi^*(2)$
$\rho_{xu} = 0.0,\ n = 100$					
κ_x	3.00	9.00	3.00	9.00	9.00
Asymptotic app.	0.0100	0.0100	0.0100	0.0100	0.0100
Actual	0.0103	0.0107	0.0103	0.0107	0.0109
Estimated	0.0103	0.0107	0.0103	0.0107	0.0109
CP-right	0.950	0.950	0.950	0.950	0.941
CP-left	0.950	0.950	0.950	0.950	0.962
CP-2sided	0.950	0.950	0.950	0.950	0.951
$\rho_{xu} = 0.2,\ n = 100$					
κ_x	3.00	8.53	3.01	8.54	8.54
Asymptotic app.	0.0100	0.0106	0.0106	0.0112	0.0112
Actual	0.0103	0.0109	0.0109	0.0116	0.0120
Estimated	0.0103	0.0108	0.0106	0.0111	0.0116
CP-right	0.946	0.942	0.946	0.942	0.934
CP-left	0.954	0.956	0.951	0.954	0.964
CP-2sided	0.950	0.949	0.948	0.947	0.946
$\rho_{xu} = 0.4,\ n = 100$					
κ_x	3.00	7.23	3.15	7.39	7.39
Asymptotic app.	0.0100	0.0124	0.0124	0.0148	0.0148
Actual	0.0103	0.0117	0.0121	0.0137	0.0153
Estimated	0.0103	0.0113	0.0113	0.0124	0.0138
CP-right	0.942	0.930	0.943	0.932	0.923
CP-left	0.957	0.963	0.950	0.956	0.961
CP-2sided	0.949	0.944	0.945	0.942	0.937
$\rho_{xu} = 0.6,\ n = 100$					
κ_x	3.00	5.46	3.78	6.24	6.24
Asymptotic app.	0.0100	0.0154	0.0154	0.0208	0.0208
Actual	0.0103	0.0131	0.0138	0.0166	0.0200
Estimated	0.0101	0.0124	0.0125	0.0147	0.0174
CP-right	0.936	0.912	0.945	0.924	0.915
CP-left	0.959	0.969	0.943	0.958	0.957
CP-2sided	0.946	0.934	0.942	0.937	0.929

The three variance assessments have been obtained as follows: Asymptotic approximation: division of (3.11) or (B.9) by n; Actual: sample variance of the simulated coefficient estimates (3.10); Estimated: sample mean of the simulated variance estimates (3.12). All confidence intervals have nominal coverage probability 0.95.

KLS specializes to OLS. Then no effect of excess kurtosis is found on the coverage probability, whereas skewness has a minor effect on the one-sided intervals.

So, it was established that under the examined forms of nonnormality the accuracy of infeasible asymptotic KLS inference in the simple model deteriorates in finite samples for increasing $\rho_{xu} > 0$. However, for cases where endogeneity, skewness and kurtosis are moderate and the sample size not too small, KLS confidence intervals seem in fact reasonably precise. For $\rho_{xu} = 0.8$ (not included in the table) the 15 coverage probabilities for the five kurtosis/skewness combinations varied between 0.84 and 0.96 at $n = 100$, but this narrows at $n = 500$ to between 0.91 and 0.97.

To obtain evidence on the accuracy of KLS inference in cross-section models with several regressors that may have different kurtosis values, the simple model was augmented with one extra exogenous regressor. Here all experiments were based on symmetric distributions, and the exogenous regressor was generated independently from the endogenous regressor. Note that making them correlated would at the same time reduce their potential discrepancy with respect to kurtosis. Due to this independence the 2×2 matrices Σ_{xx} and Θ, which determine the asymptotic variance of KLS, will be diagonal, so that the KLS estimator of the coefficient of the exogenous regressor and its variance will not be affected by ρ_{xu} at all. Therefore, just the effects on inference regarding the coefficient of the endogenous regressor were examined. In Table 5.2 results are presented for eight different cases regarding kurtosis of u and the two series $\xi^{(1)}$ and $\xi^{(2)}$, underlying the endogenous and the exogenous regressors respectively. Now $\rho_{xu} = (\rho_1, \rho_2)'$, where $\rho_2 = 0$. Taking again $n = 100$, for $\rho_1 = 0.8$ excess kurtosis in both regressors or just the exogenous regressor leads to conservative confidence intervals, whereas excess kurtosis of just the disturbance yields liberal confidence intervals, also when the endogenous regressor has excess kurtosis too. Only when just the endogenous regressor has excess kurtosis the one-sided intervals are such that their opposite inaccuracies mitigate the inaccuracy of the two-sided interval. For $\rho_1 = 0.4$ much milder discrepancies were found between nominal and actual confidence levels.

Comparing with the results of Table 5.1, Table 5.2 shows that taking the maximum of the individual kurtosis estimates does indeed increase the coverage probability of the intervals, as is to be expected, because they will be asymptotically conservative. In fact, this procedure to deal with different kurtosis coefficients helps to alleviate the problem that in finite samples actual coverage probabilities may be smaller than the nominal level.

Table 5.2
Coverage probability (CP) for β_1 interval in the static model with $K = 2$

				CP-right	CP-left	CP-2sided
$\rho_1 = 0.4,\ \rho_2 = 0,\ n = 100$						
	$u \sim N(0,1)$	$\xi^{(1)} \sim N(0,1)$	$\xi^{(2)} \sim N(0,1)$	0.946	0.956	0.951
			$\xi^{(2)} \sim St^*(5)$	0.933	0.961	0.946
		$\xi^{(1)} \sim St^*(5)$	$\xi^{(2)} \sim N(0,1)$	0.954	0.964	0.961
			$\xi^{(2)} \sim St^*(5)$	0.940	0.965	0.953
	$u \sim St^*(5)$	$\xi^{(1)} \sim N(0,1)$	$\xi^{(2)} \sim N(0,1)$	0.947	0.948	0.947
			$\xi^{(2)} \sim St^*(5)$	0.935	0.955	0.942
		$\xi^{(1)} \sim St^*(5)$	$\xi^{(2)} \sim N(0,1)$	0.953	0.957	0.956
			$\xi^{(2)} \sim St^*(5)$	0.941	0.959	0.949
$\rho_1 = 0.8,\ \rho_2 = 0,\ n = 100$						
	$u \sim N(0,1)$	$\xi^{(1)} \sim N(0,1)$	$\xi^{(2)} \sim N(0,1)$	0.952	0.962	0.955
			$\xi^{(2)} \sim St^*(5)$	0.918	0.969	0.933
		$\xi^{(1)} \sim St^*(5)$	$\xi^{(2)} \sim N(0,1)$	0.988	0.993	0.992
			$\xi^{(2)} \sim St^*(5)$	0.972	0.992	0.982
	$u \sim St^*(5)$	$\xi^{(1)} \sim N(0,1)$	$\xi^{(2)} \sim N(0,1)$	0.949	0.903	0.905
			$\xi^{(2)} \sim St^*(5)$	0.930	0.936	0.918
		$\xi^{(1)} \sim St^*(5)$	$\xi^{(2)} \sim N(0,1)$	0.975	0.973	0.972
			$\xi^{(2)} \sim St^*(5)$	0.961	0.977	0.968

All confidence intervals have nominal coverage probability 0.95.

5.2. Results for a simple time-series model under normality

Next KLS will be examined for a simple stable synthetic dynamic regression relationship in stationary zero-mean variables, given for $i = 1, \dots, n$ by

$$y_i = \beta_1 x_i + \beta_2 y_{i-1} + u_i, \text{ with } |\beta_2| < 1, \tag{5.1}$$

$$\begin{aligned} &u_i \sim iid(0, \sigma_u^2), && \varepsilon_i \sim iid[0, (1-\pi^2)\sigma_\xi^2], \\ &x_i = \xi_i + \lambda_1 u_i, && \xi_i = \pi \xi_{i-1} + \varepsilon_i, \ |\pi| < 1. \end{aligned}$$

Hence ξ_i is an AR(1) process with $E(\xi_i \xi_{i-l}) = \pi^{|l|}\sigma_\xi^2$. So, if $\pi \neq 0$ then series x_i and thus y_i will be serially dependent, whereas x_i is endogenous as well, provided $\lambda_1 \neq 0$, and has variance $\sigma_x^2 = \sigma_\xi^2 + \lambda_1^2 \sigma_u^2$.

Scaling $\sigma_u^2 = 1$ goes without loss of generality. Primary interest is in moderately nonnegative values of β_2. By choosing $\beta_1 = 1 - \beta_2$ the long-run multiplier of y with respect to x will be kept constant at value unity, irrespective of the speed of the dynamic adjustment process determined by β_2. By choosing $\pi = 0.8$ regressor x_i mimics a smooth time-series. For ρ_1 moderate positive values will be examined. Given numerical values for σ_u, β_1,β_2, π and ρ_1 data can be generated as soon as a relevant value for σ_ξ has been set. The fact that relationships like (5.1) usually have a rather high coefficient of determination (low noise/signal ratio) is respected by imposing

$$1 - \sigma_u^2/\sigma_y^2 = R^*, \tag{5.2}$$

with $R^* = 0.9$. How this determines σ_ξ is explained in Appendix G.

That consistent estimates of dynamic models like (5.1) may show substantial bias in samples of finite size has aroused a rather massive literature. The magnitude of this bias has been assessed under normality, both by simulation and by analytical methods (higher-order asymptotic approximations), both for models in which x_i is exogenous and OLS has been examined, see Kiviet and Phillips (2012) and its references, and for models in which x_i is endogenous and TSLS has been examined, see Phillips and Liu-Evans (2016) and its references, and under nonnormality of the disturbances in Liu-Evans and Phillips (2018). Here the primary aim is simply to investigate the finite sample density of KLS estimators as obtained from simulation experiments, and compare these with competitors.

In Figure 5.2.1 densities are presented for the OLS, TSLS and KLS estimators of β_1 (left-hand panels) and of $\beta_2 = 1 - \beta_1$ (right-hand panels). Examined are $\beta_1 = 0.6$ and $\rho_1 = 0.3$ (top-row panels) and $\beta_1 = 0.8$ and $\rho_1 = 0.1$ (bottom-row panels), whereas $\kappa_\varepsilon = 3$, $\kappa_u = 3$ (in fact ε_i and u_i were drawn from the normal distribution) and $n = 50$. By choosing $\beta_1 + \beta_2 = 1$ (not imposed when estimating) the long-run multiplier of x with respect to y is unity. Taking $y_{-n} = 0$ and skipping n initial observations guarantees that the generated series are (sufficiently) stationary. The TSLS estimator is actually a simple IV estimator, for which x_{i-1} and y_{i-1} have been used as instruments. So, the degree of overidentification is zero, implying that formally the estimator has no finite moments thus its distribution may be fat tailed.

The inconsistent OLS estimator is found to be severely biased, especially for β_1, even for the relatively small value $\rho_1 = 0.3$. The consistent TSLS estimator is better centered around the true value, but it has substantially larger dispersion and it is also slightly skew for β_1. The KLS estimator clearly outperforms both alternatives, showing no substantial bias nor skewness, and visibly having the smallest mean squared error. Even for minor endogeneity correlation ($\rho_1 = 0.1$) and much slower

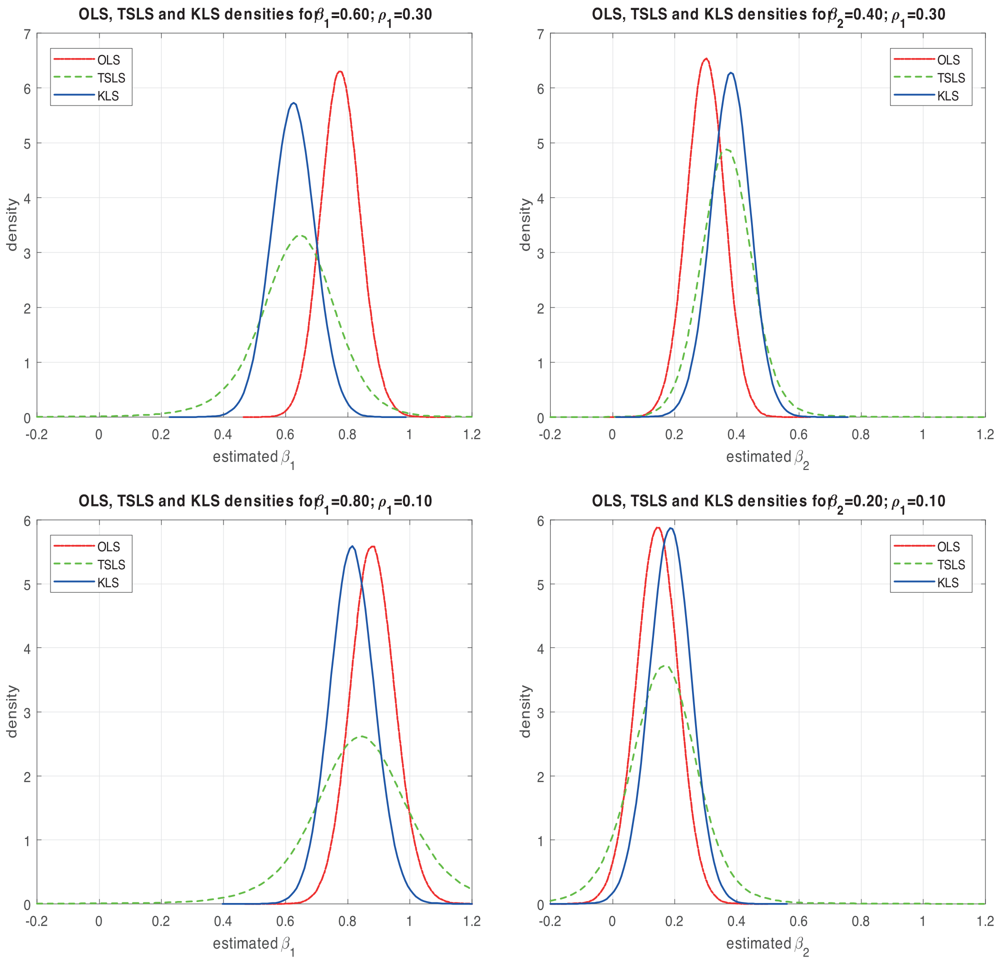

Fig. 5.2.1. Simulated densities for $n = 50$, $\pi = 0.8$, $R^{*} = 0.9$, $\kappa_{\varepsilon} = 3$, $\kappa_{u} = 3$

dynamic adjustment ($\beta_2 = 0.2$) the same patterns show up. Larger samples where $n = 250$ (not depicted) were examined too. Then all densities are more peaked, but still show the same distinctive properties.

Figure 5.2.2 depicts for both parameterizations examined in Figure 5.2.1 the strength of instrument x_{i-1} by presenting the density of its two-sided significance test (in F form) in the first-stage regression. In a nonnegligable number of replications this statistic turned out to be smaller than 10 (a criterion for weakness often used by practitioners), but on average it has been well above 10 in both cases.

In situations where validity of these moment conditions seems beyond dispute, one might (keeping the result proved in Appendix F in mind) attempt to exploit the KLS properties by substituting for the unknown ρ_1 the estimated correlation $\hat{\rho}_1$ between the TSLS residuals and the endogenous regressor. From simulations for $n = 50$ the distribution of such an estimator $\hat{\rho}$, although consistent (if not asymptotically weak) and reasonably well centered around its true value, was found to be very imprecise. Figure 5.2.3 presents for $n = 250$ its simulated density for both parameterizations examined in Figure 5.2.1. Although the estimated expectations of these ρ estimates seem most reasonable (0.294 and 0.091 respectively) the densities make clear that TSLS estimates of the degree of endogeneity are not very trustworthy due to their wide dispersion, let alone because of their always questionable consistency.

Of course, one should keep in mind that the KLS estimator as presented here is infeasible, because it uses, next to $E(y_{i-1}u_i) = 0$, the in practice unknown true value of ρ_1. However, that should not be seen as a disqualification. On the contrary, one can interpret KLS findings for a range of likely values of ρ_1, whereas interpretation of the estimators TSLS and OLS (which are feasible because their calculation just requires observables) is seriously hampered, because the strict orthogonality conditions they exploit cannot be vindicated statistically in practice. In this and earlier studies it has been established that, under various circumstances, infeasible KLS estimators often have much smaller mean squared errors than IV/TSLS, whereas its actual distribution is more accurately represented by its first-order asymptotic approximation. So, there seems substantial scope for feasible KLS (performed under confined regressor endogeneity) to provide more efficient and more robust inference.

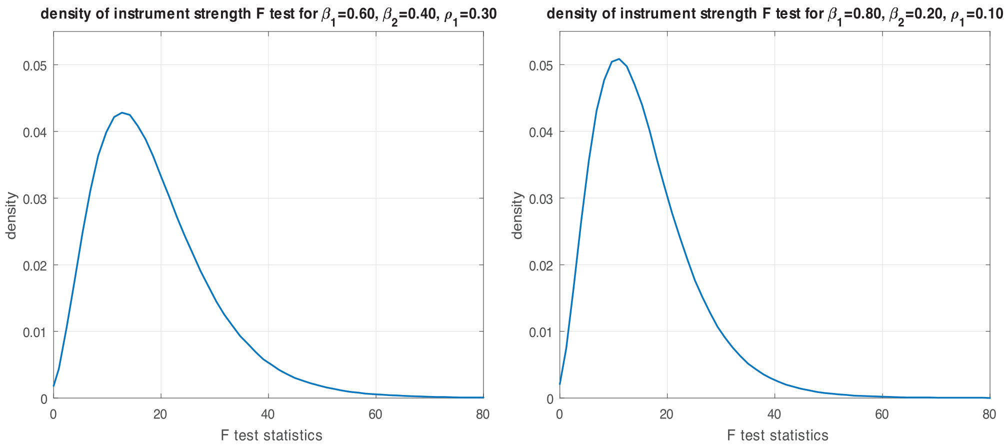

Fig. 5.2.2. Simulated densities of F test; $n = 50$, $\pi = 0.8$, $R^* = 0.9$, $\kappa_\varepsilon = 3$, $\kappa_u = 3$

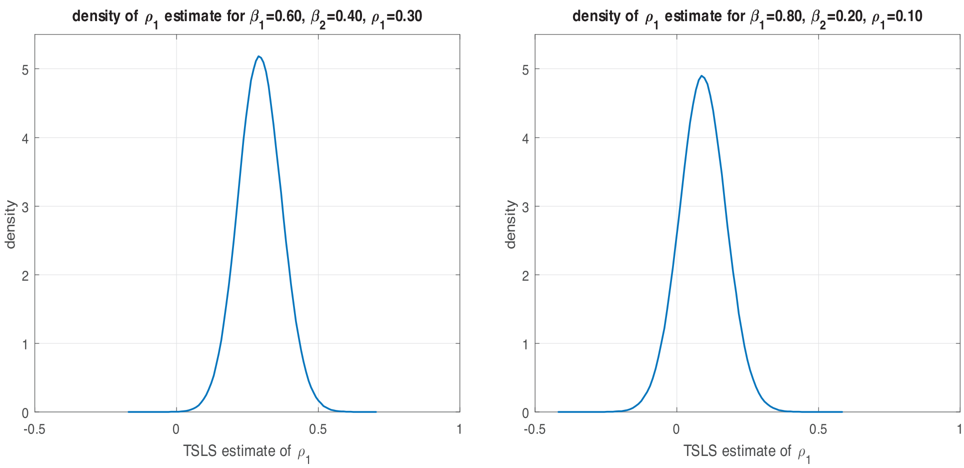

Fig. 5.2.3. Simulated densities of $\hat{\rho}_{xu}$; $n = 250$, $\pi = 0.8$, $R^* = 0.9$, $\kappa_\varepsilon = 3$, $\kappa_u = 3$

6. Three empirical illustrative replication studies

Below it will be illustrated how the techniques discussed here can be used in practice and can place earlier obtained results using instrumental variables into a new revealing perspective, either positively or negatively. The first is an international-macro application in which a country cross-section data set is re-analyzed from which the causal effect of international trade on income per capita has been assessed. In this illustration also the effect of nonzero excess kurtosis is investigated. In the second example a cross-sectional data set is re-analyzed on Vietnamese individuals examining the causal effect of personal income on risk aversion. Here a specification with one and also one with two endogenous regressors is investigated, comparing KLS findings with weak-instrument robust inference. In the third illustration a time-series data set is re-analyzed, paying extra attention to the fact that the obtained KLS formulas also apply to dynamic models based on temporally dependent time-series. All the time a nominal significance level of 5% has been used.

6.1. Effect of trade on growth

In a much referenced study (over 6000 citations according to Google) by Frankel and Romer (1999), below referred to as F&R, data for the year 1985 on 150 countries have been analyzed from which it has been concluded (F&R, Table 3, column 2) that a 1 percent-point raise of the trade share T (defined as the sum of exports and imports divided by GDP) leads to an increase of about 2% in income per person Y (the coefficient estimate in a regression of $\ln Y$ on T is 1.97 with standard error 0.99). This is the result of a linear IV analysis where trade share is the one and only endogenous regressor supplemented by an intercept and two exogenous covariates, namely log of population $\ln N$ and of area $\ln A$. The study uses one instrument, called the constructed trade share $\hat{T}$, which has been obtained by regressing trade share on a series of geographic characteristics. So, actually TSLS has been employed. However, by not providing the original set of instruments

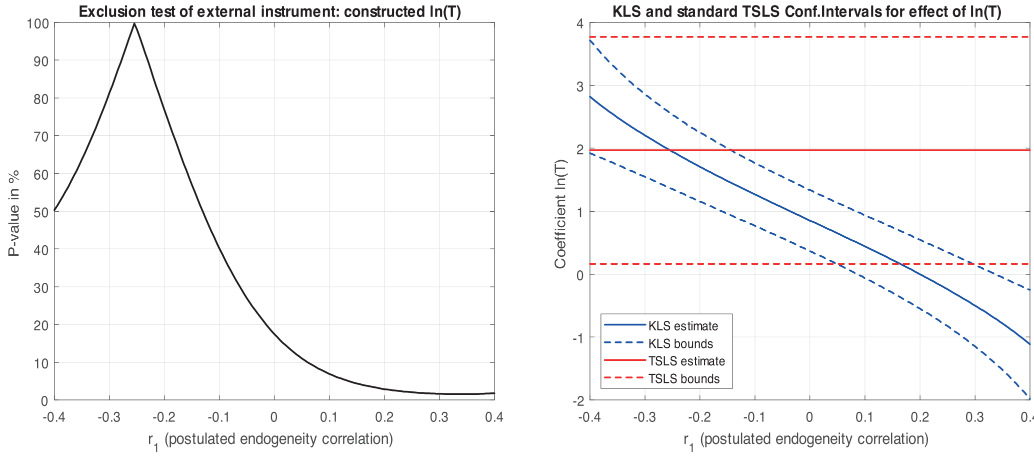

Fig. 6.1. KLS inference regarding the effect of trade on growth

used in the first stage, neither the Sargan test can be re-established from the provided second stage data in F&R's Table A1, nor the Anderson-Rubin confidence set on the coefficient of the endogenous regressor can be obtained.

The constructed single instrument $\hat{T}$ is not weak, but not very strong either, because the relevant first-stage F test statistic of $\hat{T}$ (in a regression of T on an intercept, $\ln A$, $\ln N$ and $\hat{T}$) is 13.1. Employing this test with critical value 10, suggested in Stock et al. (2002), is popular among practitioners, although it has shortcomings, see Hahn and Hausman (2003). The modest size of the F test may explain the relatively large standard error and consequently large confidence interval (0.03, 3.91) for the parameter of major interest. This interval has nominal confidence coefficient 95% but may in fact be highly unreliable. The OLS estimate of the coefficient of interest is only 0.85 with much smaller standard error 0.25. Of course, both coefficient estimates may be severely biased. By a Hausman test F&R establish that endogeneity of trade share is not significant (its t-value is 1.2), and therefore the substantial difference between the IV and OLS coefficient estimates is actually not a significant difference. In the end F&R conclude that the IV coefficient estimate being more than twice the size of the OLS estimate is simply due to random variation, also because they argue at length that they in fact expect that the IV estimate should be smaller than the OLS estimate, simply due to $\rho_1 > 0$.

KLS procedures can put some extra statistical evidence on the table. Postulating credible values r_1 for the unknown ρ_1 one can verify whether the constructed trade share seems an invalid instrument, by testing whether it seems wrongly omitted from the specified relationship. If this deprives trust in the TSLS findings then instrument-free inference on the parameter of interest, which is not infested with weak or invalid instrument problems, may be of more use. From the IV results an estimate for the degree of endogeneity of -0.25 is obtained, which may be very imprecise (as Section 5 indicates), or even inconsistent, if the instrument is invalid, possibly due to a poorly specified structural equation. So, in line with F&R's belief that $\rho_1 > 0$, one should certainly examine what the consequences would be of more credible positive values r_1.

The left-hand panel of Figure 6.1 shows that if ρ_1 were positive indeed, it seems highly unlikely that the instrument is valid, due to low p-values of the exclusion restriction test. Also note the high p-value around values for ρ_1 close to its IV estimate; this is the phenomenon proved in Appendix F: accepting an instrument as valid, yields estimates for the endogeneity correlation which approve the exclusion restriction. Note the circularity here; this does not provide evidence that ρ_1 is negative and neither that $\hat{T}$ is a valid instrument.

The right-hand panel of Figure 6.1 shows that for any value of $\rho_1 > -0.25$ KLS yields (much) lower values for the estimated effect of trade-share than IV produces, and also yields much narrower confidence intervals. If one chose to postulate that $-0.14 \leq \rho_1 \leq 0.05$ (and assuming that the model is adequately specified) one could conclude with a probability exceeding 95% that the effect of trade share is in between 0 and 2. Under the same proviso, KLS suggests that if $\rho_1 > 0.08$ then the coefficient of $\ln(T)$ is smaller than 1. Figure 6.1 does not disclose which values of r_1 are more credible. For ρ_1 close to -0.25 the IV and KLS estimates are almost similar, with the IV interval wider than the KLS interval. KLS refutes to assume jointly that the model is correctly specified, the instrument is valid, while ρ_1 is positive.

From the above one should conclude that the employed model has been rather poorly specified. It seems most likely that the current analysis is affected by the problems of reciprocal causality, measurement errors and omitted variables jointly. Rather than aiming to take that all into account by KLS methods, it seems that attempts should be undertaken first to formulate a better explanatory model of income in terms of trade share with further controls not available in the original data set.

Table 6.1 presents estimates of skewness and kurtosis of the relevant variables. Note the extremely large kurtosis of the external instrument $\hat{T}$. So, the KLS exclusion restriction test used 43.36 for the maximized value of κ_x, and 7.9 for the KLS inference on the effect of trade share. The graphs of Figure 6.1 were also constructed after partialling out the variables $\ln N$ and $\ln A$, and again by simply taking $\kappa_x = 3$; as this had hardly a visible effect, these results have not been presented. We

Table 6.1
Skewness and kurtosis of some F&R variables

	$\ln Y$	T	$\ln N$	$\ln A$	$\hat{T}$	IV-resid.
skewness	-0.19	1.64	-0.66	-0.11	5.14	-0.25
kurtosis	2.00	7.90	3.15	2.99	43.36	2.56

just mention this, because it indicates that the component of the KLS variance affected by the kurtosis coefficients seem of secundary importance.

6.2. Risk preferences in Vietnam

Tanaka et al. (2010) analyze cross-section data obtained by combining living standard survey data on individuals collected in 2002 with additional experiment-based direct measurements on risk and time preferences of the very same individuals living in nine particular communities. In their Section 2 the authors (below indicated as TCN) estimate relationships for two different dimensions of risk aversion. Below the focus will just be on the results for the dependent variable "concavity of the value function" (σ in the TCN notation), for which they examined two different specifications. In addition to a range of exogenous demographic control variables, TCN include as endogenous regressor(s) either just household income, or both relative income and mean income within the community. These two relationships are estimated by TSLS employing two external instrumental variables: rainfall and the dummy variable household head's ability to work. Hence, in one equation the degree of over-identification is one, and in the other it is zero. Although TCN use two variants of tests to verify the significance of the endogeneity of the income variables, no test of over-identification restrictions has been employed.

With respect to the validity of the two external instruments TCN just state that these are: "... unlikely to be correlated with preferences, as instrumental variables for income". This, however, is a most confusing remark, originating in the fact that TCN wrongly use the word correlation for equation coefficient (see also the caption of their Table 4). A relevant instrument should be (preferably substantially) correlated with the endogenous regressors, and will therefore (assuming the endogenous regressor has nonzero coefficient) also be correlated with the dependent variable. However, to be a valid instrument, it should be correctly excluded from the structural equation for the dependent variable and thus have coefficient zero in this equation. Therefore, its impact on the dependent variable should be solely indirect, via the endogenous regressor. Correct exclusion seems perhaps self-evident for rainfall, but is much less straight-forward regarding the dummy instrument ability to work, because one's preferences regarding risk taking may certainly be structurally affected in case of permanent disability. In their footnote 8 TCN remark that they "... tested several instrumental variables...". However, here they refer to F-tests in (reduced form) first-stage regressions, so testing the relevance (strength or weakness) of the instruments, but not their validity.

TCN obtain their empirical findings from a sample of 181 observations. Presupposing validity of the instruments and adequacy of the two specified equations (as TCN persistently do) the tests on endogeneity of the income variables reject consistency of the OLS results. TCN assume that the endogeneity is due to reciprocal causality. Given the findings in Appendix A this would imply the endogeneity correlations to be positive, unless $\gamma_0\beta_1 > 1$, which seems unlikely. In fact, however, for the equation with just one endogenous regressor the estimated correlation of the endogeneity is -0.52. Of course, apart from their substantial standard errors, these correlation estimates will only be consistent if the instruments are valid indeed.

TCN apply TSLS (see their Table 5) and find for the model with one endogenous regressor ($K_1 = 1$) for the coefficient of income 0.010 (standard error 0.006). Using their data set it is found that the Sargan test for the overidentification restriction has p-value 0.79, whereas testing the joint strength of the instruments yields F-value 5.96, so on the basis of these usual diagnostics the instruments seem valid (implicitly supposing that one instrument is valid anyhow), but seriously weak too.

Figure 6.2.1 presents KLS results for the relationship with $K_1 = 1$. The top-left panel (for the three curves KLS is defined for r_1 in absolute value smaller than 0.92, 0.94 and 0.92 respectively) shows that especially the "rainfall" dummy instrument has low p-values for $\rho_1 > -0.2$ and the joint exclusion of the two instruments finds certainly no support for positive ρ_1. Observing the peaks of the three p-value curves, it will not surprise that the TSLS based estimators for ρ_1 obtained when just using the instrument "rainfall" is -0.55, when just using the instrument "head unable to work" it is -0.42, and when using both instruments it is -0.52. Although the endogenous regressor has an estimated kurtosis of 31.7 this high value has little effect on the results in Figure 6.2.1, because substituting (incorrectly) $\kappa_x = 3$ in the variance formula was again found to have just minor effects.

The other panels (defined for $|r_1| \leq 0.95$) produce (asymptotic) 95% confidence intervals for the endogenous regressor Income, and for the exogenous regressors Gender and Education. Next to KLS and standard TSLS intervals they also present the weak-instrument robust intervals (robustC and robustU) constructed by Guggenberger et al. (2019). From the KLS findings on Income one may conclude that its effect is positive provided $\rho_1 < -0.2$. And, assuming that $\rho_1 > -0.6$, one can also conclude that the Income coefficient value does not exceed 0.02. For this coefficient, the robust – and thus much more trustworthy interval than the standard one – highlights that weak-instrument robust TSLS inference is actually not very efficient in comparison to KLS. For the coefficients of exogenous regressors this difference in width of the confidence intervals is less spectacular.

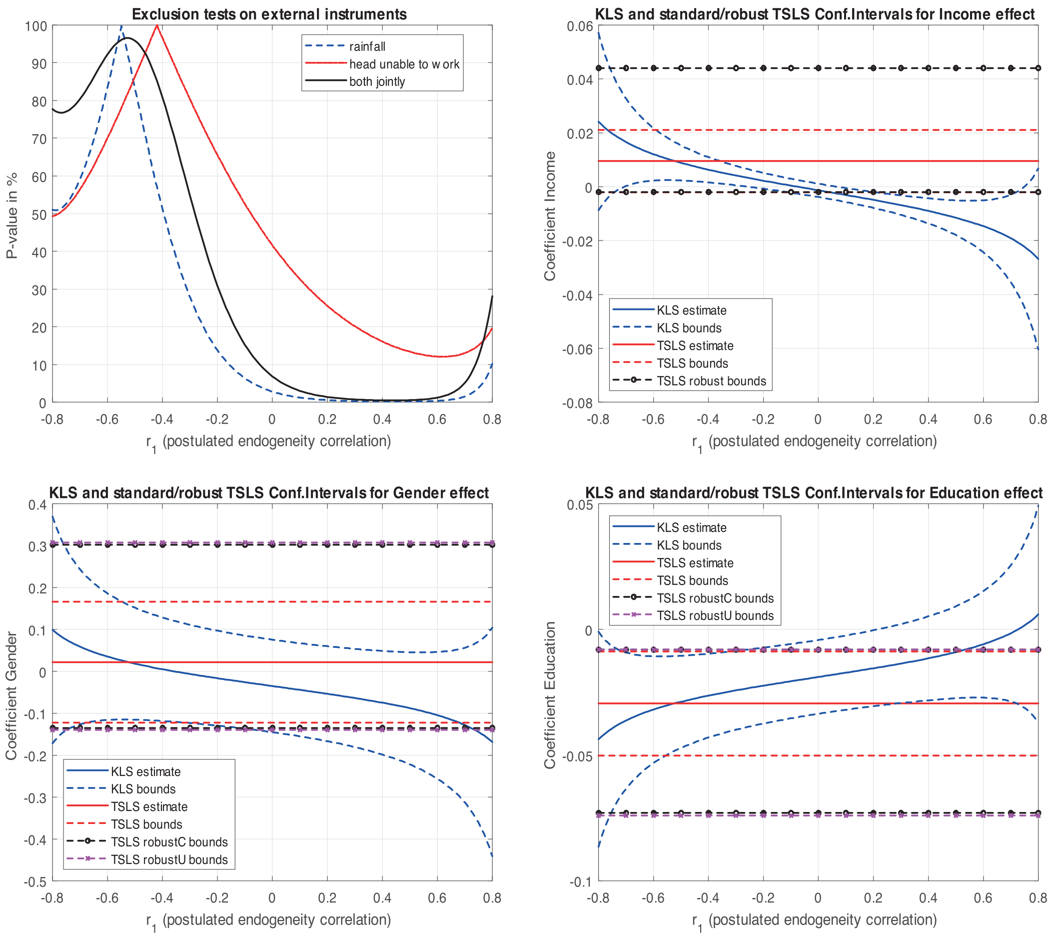

Fig. 6.2.1. KLS and (robust) TSLS inference for the model with $K_1 = 1$

For the model with two endogenous regressors it was found that the correlation estimates of the endogeneity of relative and mean income are -0.16 and -0.18 respectively. Figure 6.2.2 presents 2D contour plots of p-values for four different KLS-based tests over all feasible postulated values (r_1, r_2) for (ρ_1, ρ_2). Its North-West panel shows that for moderate absolute values of both endogeneity correlations the joint validity of the two instruments has to be rejected. The North-East and South-West panels show that it is again the "head unable to work" instrument which is to blame for this. Therefore, producing further TSLS inference on the coefficients of this relationship seems in vain. Assuming validity of both instruments, TCN and Guggenberger et al. conclude that the coefficient of relative income is insignificant, whereas that of mean income is –or is close to– significantly positive. Not using any external instruments, the South-East panel of Figure 6.2.2 presents the p-values (all larger than 0.9) for a KLS-based one-sided t-test. This shows that the hypothesis of a zero coefficient for mean income should not be rejected in favor of a positive value, irrespective of the values of the endogeneity correlations.

6.3. Demand at the Fulton fish market

To illustrate KLS procedures when applied to time-series observations, some studies will be used that examined the market for whiting based on 111 consecutive trading days at the Fulton fish market. These data originate from Graddy (1995). Angrist et al. (2000) produce in one of the columns of their Table 4 just-identified TSLS results for a base-line static linear in logs demand equation. These TSLS results represent the preferred specification in Graddy (2006, Table 4, column 2) and in Graddy and Kennedy (2010, Table 2, column 2). The latter study argues that supply is not just determined by the previous night's catch, but also by inventory changes, so that at this market quantity traded and its price are simultaneously determined indeed. Using this fish market example, Imbens (2014) provides a thorough explanation to convince especially statisticians that endogeneity of explanatory variables is often a reality in observational studies, despite the confusing fact that simultaneous equations cannot directly be simulated on a computer, without generating endogenous variables by their reduced form equations, or possibly (in case of the fish market) by integrating all relevant aspects of each individual transaction during the day.

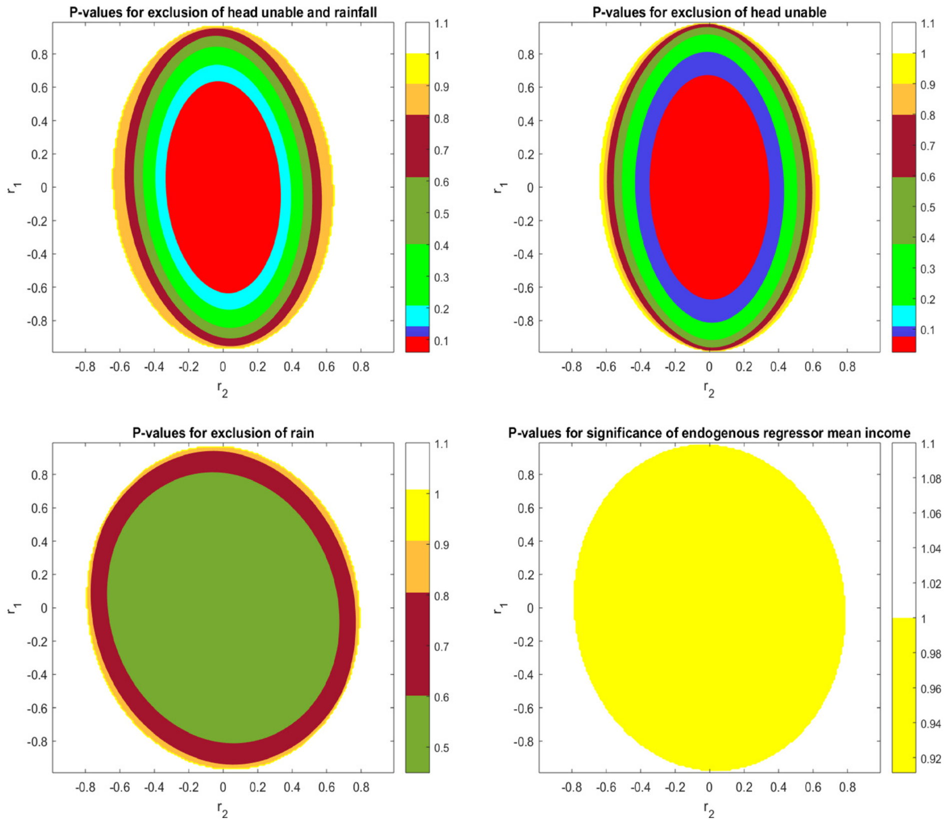

Fig. 6.2.2. KLS inference for the model with $K_1 = 2$ ($\kappa = 3$)

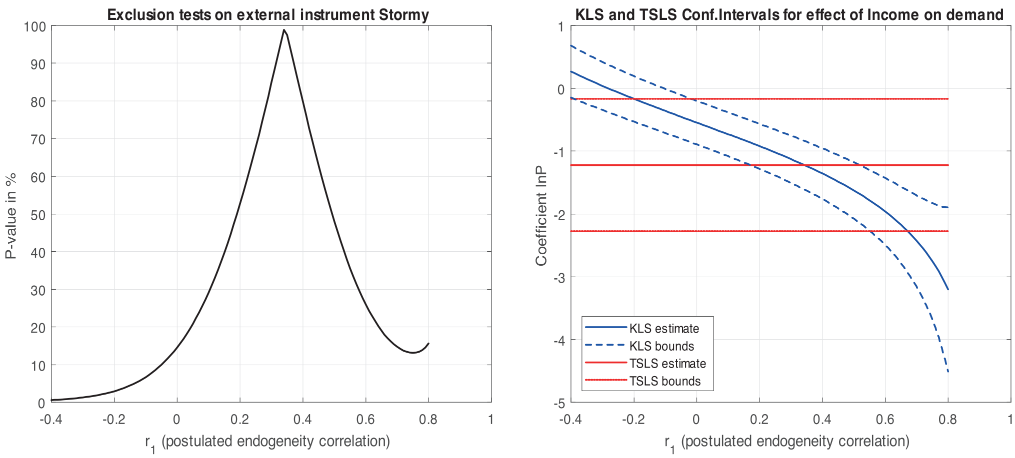

Fig. 6.3.1. KLS inference for the static demand equation

In the static demand equation examined in the just mentioned studies the regressand is $\log Q$ (quantity) and, according to (A.4), endogenous regressor $\log P$ (price) is expected to be positively correlated with the error term. In line with that OLS yields a larger (less negative) price coefficient estimate than IV. This equation also includes day of the week dummies and two variables ”cold” and ”rainy”, characterizing the weather on shore. It is just-identified, because weather variable at sea ”stormy” is supposed to be a determinant of supply, but not of demand.

Further calculations reveal the following about the static specification. As can be seen from the left panel of Figure 6.3.1 $\hat{\rho}_1 = 0.34$. Moreover, assuming for example $0.18 \leq \rho_1 \leq 0.48$, the exclusion restriction test has a p-value exceeding 0.5. So, for moderately positive values of ρ_1, the application of IV finds support. However, the first-stage F test for the external instrument is 5.85, so it is pretty weak. Therefore, the IV estimate of the price elasticity of demand is expected to be

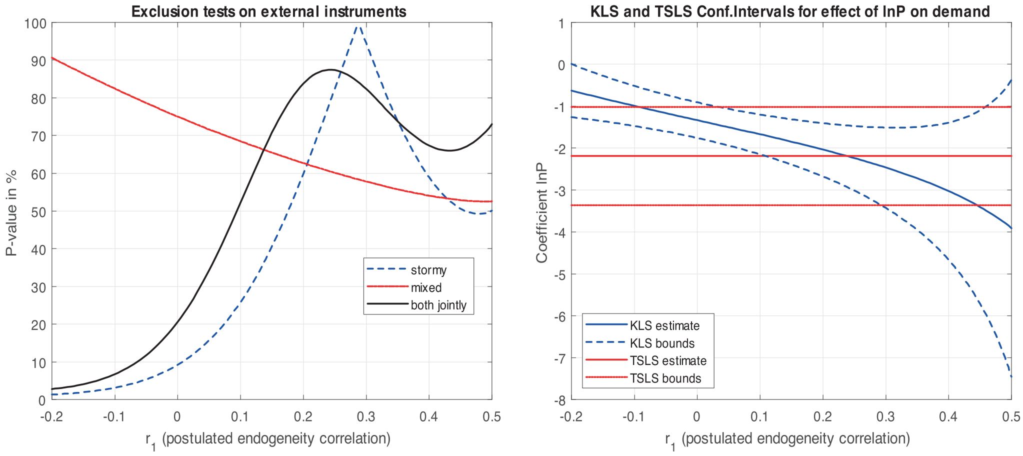

Fig. 6.3.2. KLS inference for over-identified dynamic specification (6.1)

biased in the direction of OLS, which is about -0.9. This would imply that the true value might be closer to -2 than to the obtained -1.22. However, KLS inference, which is not plagued by weak instrument problems and finite sample inaccuracies, indicates that the elasticity seems smaller than -2, provided this static specification of the demand equation is appropriate and $\rho_1 \leq 0.475$.

None of the studies presenting the investigated specification report any checks on its adequacy, apparently bedazzled by the ostensible correct signs of the coefficients despite their substantial standard errors. However, a test for 1-st order serial correlation produces a p-value of 0.00. This finding disqualifies all substantive inferences regarding this static specification.

Hendry and Nielsen (2006) analyze the same Fulton fish market data, adhering to a methodology which aims to fully respect the temporal-dependence in the observed variables. First they find single descriptive "congruent" time-series models for both $\log Q_t$ and $\log P_t$, and next they combine these in a VAR (vector autoregressive) model of order 1, which also includes two weather dummy variables, namely the earlier "stormy" and another called "mixed" (also used in further investigations by Angrist et al., 2000), together with a dummy "hol" which is unity at three particular dates close to holidays. This VAR specification passes an extensive mis-specification analysis. Next, using decomposition methods to the joint log-likelihood, these initial descriptive exercises inspire to formulate dynamic structural simultaneous equations for supply and demand. For the over-identified demand equation their maximum likelihood estimates are almost similar to the corresponding single equation TSLS estimates (see their Tables 15.5 and 15.9)

$$\underset{}{\log Q_t} = \underset{(0.07)}{8.52} - \underset{(0.59)}{2.19} \log P_t + \underset{(0.47)}{1.83} \log P_{t-1} - \underset{(0.36)}{1.89} hol_t. \tag{6.1}$$

Here "stormy" and "mixed" are used as external instruments. The test for the single over-identification restriction has p-value 0.60, and a test for 1st order serial correlation has p-value 0.71. The joint first stage F statistic of the two external instruments is 9.45, so although not strong, they are not very seriously weak either.

Application of KLS yields the following. The left-hand panel of Figure 6.3.2 supports the validity assumption regarding both instruments, assuming $\rho_1 > 0$ (the $\hat{\rho}_1$ estimates are 0.28, -0.31 and 0.24 for using stormy, mixed or both respectively as instruments). The right-hand panel produces (as seen before) for moderate r_1 values narrower confidence bands for the coefficient of the endogenous regressor ($\log P_t$), but not as dramatically as for cases where instruments are really weak.

This dynamic specification of demand implies a price elasticity which reacts immediately very strongly to price changes, but with a huge correction to that the next trading day. So, the long-run elasticity (which is attained already after one period) has TSLS estimate of only -0.36. By reformulating the model its TSLS standard error can be obtained, giving

$$\log Q_t = \underset{(0.07)}{8.52} - \underset{(0.59)}{2.19} \Delta \log P_t - \underset{(0.21)}{0.36} \log P_{t-1} - \underset{(0.36)}{1.89} hol_t. \tag{6.2}$$

So, the elasticity seems much smaller in absolute value than inferred before, and it does not even seem significantly negative. Figure 6.3.3 presents KLS inference on the coefficients of the predetermined regressors $\log P_{t-1}$ and hol_t. Note that the KLS results imply a significantly negative long-run demand elasticity provided $\rho_1 > 0.3$.

However, all the above inferences on the demand for whiting seem rash, because a TSLS-based Chow breakpoint test half-way the sample applied to model (6.2) yields a p-value of 0.00!

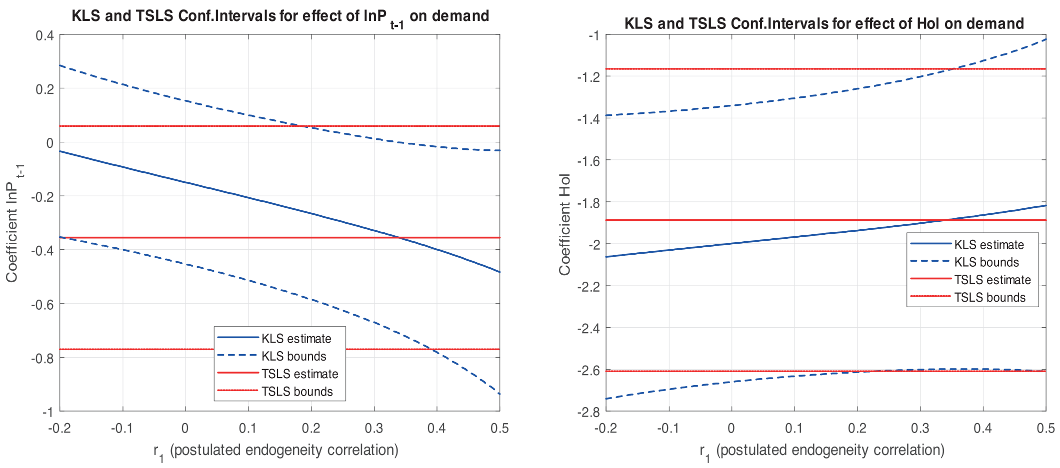

Fig. 6.3.3. Further KLS inference regarding over-identified specification (6.2)

7. Concluding remarks

Since the 1950's TSLS estimation has been the prominent workhorse in empirical econometrics for causal analysis of single linear relationships involving endogenous regressors. It is based on assuming uncorrelatedness between the error terms of the model and at least as many variables, not occurring in the model equation and therefore called external instruments, as there are endogenous explanatory variables in the model. The latter are or seem correlated with the error term due to either measurement errors, omitted relevant regressors or reverse causality, or a combination of these. If there are $K_1 > 0$ possibly endogenous regressors in the model, then, assuming that a subset of $L_{11} \geq K_1$ of the $L_1 > L_{11}$ external instruments is valid (uncorrelated with the errors indeed), the validity of any combination of the $L_1 - L_{11}$ candidate external instruments can be tested by including them as (non-endogenous) regressors into the model, applying TSLS, and testing whether their exclusion seems statistically acceptable. Thus, for K_1 instruments, testing their validity by TSLS is impossible, whereas their validity is a necessary requirement for testing the validity of any further candidate instruments. This is of course a very sad state of affairs.

In addition, for several decades it is widely known now, that when instruments are weak, meaning that when using them to fit the endogenous regressors this fit is actually rather poor, this weakness has three serious consequences, namely: (a) the TSLS estimators are seriously biased, even in large samples; (b) their variance will be large; (c) moreover, the usual estimate of their variance is much too optimistic. Only for the last mentioned problem the so-called weak-instrument asymptotic approaches have provided a cure. So, when instruments are weak, weak-instrument robust TSLS inference still suffers from bias and large variance, and often faces hard to refute criticism due to the untested proviso that all instruments are valid. Hence, a fundamentally different approach, based on out of the box thinking, seems to be called for.

To find the coefficient values of the regressors which explain the variation in the dependent variable, TSLS attempts to get rid of the variation in the endogenous regressors that is correlated with the error term, by employing just the variation in the endogenous regressors that can be explained by the available candidate instruments. Usually this does not establish all variation in the endogenous regressors that is uncorrelated with the errors, especially not when there are missing instruments or when instruments are weak, see Doko Tchatoka and Dufour (2020). And this procedure falls through completely if some instruments are actually invalid (correlated with the disturbances themselves). Therefore, the KLS technique, further developed in the present study, avoids using external instruments. It decomposes the variation in the endogenous regressors directly into two components, one proportional to the errors, and the remaining one. The former, not the latter, causes all the trouble due to endogeneity of regressors, so there the focus should be. TSLS aims to establish the latter, which fails when instruments are lacking or weak or invalid.

KLS uses an assessment of the component of each regressor that is infected by the error term to correct the inconsistency of the OLS estimator. This is only possible by making the far-fetched assumption that the investigator knows the degree of endogeneity of all regressors. Although this may be the case for some obviously predetermined or exogenous regressors (where it is zero), usually this is not the case for all explanatories. However, by varying the numerical assumptions regarding endogeneity over intervals thought reasonable, it is nevertheless possible to generate KLS inference that is robust regarding endogeneity and is not suffering from the problems faced by TSLS, since availability, strength and validity of instruments are not an issue.

It is striking how little attention users of the TSLS technique have mostly given to the degree of endogeneity of the regressors in their models. Usually two-sided test statistics are being used when testing the significance of endogeneity of regressors, because interest has always just been on absence or presence of endogeneity. From the illustrations in this paper it is clear that the authors of some of the articles used here for replication studies could have directly concluded from their estimates of ρ_{xu}, or from the sign of the difference between TSLS and OLS estimates, that their empirical findings in fact refute the theories and assumptions that they initially had embraced. If they had only noticed these incompatibilities, they should have felt a need to reformulate their assumptions, theories or specifications. Moreover, although the asymptotic distribution of TSLS is invariant regarding ρ_{xu}, its finite sample distribution is certainly not. So, focussing on ρ_{xu}, which is at the heart of KLS, should also bolster up instrument-based analysis.

KLS, which is simply a bias-corrected least-squares estimator, proves to be virtually unbiased in small samples from cross-sections, and their finite sample bias in dynamic models seems primarily due to predeterminedness (and not to endogeneity) of regressors. Since the endogeneity-bias is actually an inconsistency, the bias-correction complicates the derivation of the asymptotic variance of KLS. However, the obtained analytical KLS variance estimates, and the speed of convergence towards normality of the KLS coefficient estimates, are such, that inference can claim a high degree of accuracy, provided the model used represents the underlying data generating process well. Evidence supporting the specified model (or clues regarding its potential fruitful respecification) can be obtained by employing KLS-based tests regarding omitted regressors (including coefficient nonconstancy), serial correlation, heteroskedasticity, functional form (RESET) as provided and illustrated in Kripfganz and Kiviet (2021) in a Stata environment.

Because regressors of an econometric time-series model for which current observations are correlated with lagged disturbances are not predetermined but endogenous when the disturbances are serially correlated, robustification of KLS variance estimates with respect to unknown forms of serial correlation seems unattainable, as suggesting reasonable intervals for their endogeneity coefficients would require knowledge which will usually not be available. So, in a time-series context, KLS should in the end preferably be applied to models which are so complete, that serial correlation (and thus omitted relevant regressors) has been avoided. Regarding heteroskedasticity, Andrews et al. (2019b) urge that methods should be developed such that they achieve robustness regarding both weak instruments and unknown forms of heteroskedasticity. Likewise, instrument-free procedures, which at present still presuppose homoskedasticity, should ideally be robustified regarding heteroskedasticity too. On the other hand, Romano and Wolf (2017) argue that simply correcting estimated variances for unknown forms of heteroskedasticity may often yield relatively poor inferences, if not preceded by serious attempts to model the heteroskedasticity and weight the observations accordingly, in order to obtain first a more efficient coefficient estimator less affected by the remaining heteroskedasticity. In that light, the presently available KLS procedures should be employed as tools in a well-structured specification search. Its ultimate inferences should pertain to a model specification for which convincing evidence has been produced that its error terms may assumed to be iid. Only then trustworthy endogeneity robust inference may have been produced.

Declaration of Competing Interest

There are no conflicts of interest for the author of this paper.

Acknowledgment

I'm very grateful for constructive comments by four anonymous referees and an associate editor.

Supplementary material

Supplementary material associated with this article, grouped in Appendices A through G, can be found in the online version, at doi:10.1016/j.ecosta.2021.12.008

References

Andrews, D.W.K., Marmer, V., Yu, Z., 2019a. On optimal inference in the linear IV model. Quantitative Economics 10, 457–485.
Andrews, D.W.K., Soares, G., 2010. Inference for parameters defined by moment inequalities using generalized moment selection. Econometrica 78, 119–157.
Andrews, I., Stock, J., Sun, L., 2019b. Weak instruments in IV regression: Theory and practice. Annual Review of Economics 11, 727–753.
Angrist, J.D., Graddy, K., Imbens, G.W., 2000. The interpretation of instrumental variables estimators in simultaneous equations models with an application to the demand for fish. Review of Economic Studies 76, 499–527.
Bontemps, C., Magnac, T., 2017. Set identification, moment restrictions, and inference. Annual Review of Economics 9, 103–129.
Bound, J., Jaeger, D., Baker, R., 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variables is weak. Journal of the American Statistical Association 90, 443–450.
Bugni, F.A., Canay, I.A., Shi, X., 2017. Inference for subvectors and other functions of partially identified parameters in moment inequality models. Quantitative Economics 8, 1–38.
Conley, T.G., Hansen, C.B., Rossi, P.E., 2012. Plausibly exogenous. Review of Economics and Statistics 94, 260–272.
Doko Tchatoka, F., Dufour, J.M., 2020. Exogeneity tests, incomplete models, weak identification and non-gaussian distributions: Invariance and finite-sample distributional theory. Journal of Econometrics 218, 390–418.
Fisher, F.M., 1959. Generalisation of the rank and order conditions for identifiability. Econometrica 27, 431–447.
Fisher, F.M., 1963. Uncorrelated disturbances and identifiability criteria. International Economic Review 4, 134–152.
Frankel, J., Romer, D., 1999. Does trade cause growth? American Economic Review 89, 379–399.
Graddy, K., 1995. Testing for imperfect competition at the fulton fish market. RAND Journal of Economics 26, 75–92.
Graddy, K., 2006. Markets: The fulton fish market. Journal of Economic Perspectives 20, 207–220.
Graddy, K., Kennedy, P., 2010. When are supply and demand determined recursively rather than simultaneously? Eastern Economic Journal 36, 188–197.
Guggenberger, P., Kleibergen, F., Mavroeidis, S., 2019. A more powerful subvector anderson rubin test in linear instrumental variables regression. Quantitative Economics 10, 487–526.

Hahn, J., Hausman, J., 2003. Weak instruments: Diagnosis and cures in empirical econometrics. The American Economic Review: Papers and Proceedings 93 (2), 118–125.
Hendry, D.F., Nielsen, B., 2006. Econometric Modeling; A Likelihood Approach. Princeton University Press, Princeton, USA.
Imbens, G.W., 2014. Instrumental variables: An econometrician's perspective. Statistical Science 29, 323–358.
Kiviet, J.F., 2013. Identification and inference in a simultaneous equation under alternative information sets and sampling schemes. The Econometrics Journal 16, S24–S59.
Kiviet, J.F., 2016. When is it really justifiable to ignore explanatory variable endogeneity in a regression model? Economics Letters 145, 192–195.
Kiviet, J.F., 2017. Discriminating between (in)valid external instruments and (in)valid exclusion restrictions. Journal of Econometric Methods 6, 1–9.
Kiviet, J.F., 2020. Testing the impossible: identifying exclusion restrictions. The Journal of Econometrics 218, 294–316.
Kiviet, J.F., Kripfganz, S., 2021. Instrument approval by the sargan test and its consequences for coefficient estimation. Economics Letters 205, 109935.
Kiviet, J.F., Niemczyk, J., 2012. The asymptotic and finite sample (un)conditional distributions of OLS and simple IV in simultaneous equations. Journal of Computational Statistics and Data Analysis 56, 3567–3586.
Kiviet, J.F., Phillips, G.D.A., 2012. Higher-order asymptotic expansions of the least-squares estimation bias in first-order dynamic regression models. Journal of Computational Statistics & Data Analysis 56, 3705–3729.
Koopmans, T.C., Rubin, H., Leipnik, R.B., 1950. Measuring the equation systems of dynamic economics. In: Koopmans, T.C. (Ed.), Chapter II in *Statistical Inference in Dynamic Economic Models*. John Wiley and Sons, New York. Cowles Commission Monograph No. 10
Kraay, A., 2012. Instrumental variables regressions with uncertain exclusion restrictions: A bayesian approach. Journal of Applied Econometrics 27, 108–128.
Kripfganz, S., Kiviet, J.F., 2021. kinkyreg: Instrument-free inference for linear regression models with endogenous regressors. The Stata Journal 21, 772–813.
Lewbel, A., 2012. Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. Journal of Business & Economic Statistics 30, 67–80.
Liu-Evans, G., Phillips, G.D.A., 2018. On the use of higher order bias approximations for 2SLS and k-class estimators with non-normal disturbances and many instruments. Econometrics and Statistics 6, 90–105.
Murray, M.P., 2006. Avoiding invalid instruments and coping with weak instruments. Journal of Economic Perspectives 20, 111–132.
Murray, M.P., 2017. Linear model IV estimation when instruments are many or weak. Journal of Econometric Methods 6, 1–22.
Nevo, A., Rosen, A.M., 2012. Identification with imperfect instruments. The Review of Economics and Statistics 94, 659–671.
Phillips, G.D.A., Liu-Evans, G., 2016. Approximating and reducing bias in 2SLS estimation of dynamic simultaneous equation models. Journal of Computational Statistics & Data Analysis 100, 734–762.
Romano, J.P., Wolf, M., 2017. Resurrecting weighted least squares. Journal of Econometrics 197, 1–19.
Stock, J.H., Wright, J., Yogo, M., 2002. A survey of weak instruments and weak identification in generalized method of moments. Journal of Business & Economic Statistics 20, 518–529.
Tanaka, T., Camerer, C.F., Nguyen, Q., 2010. Risk and time preferences: Linking experimental and household survey data from vietnam. American Economic Review 100, 557–571.
Wegge, L.L., 1965. Identifiability criteria for a system of equations as a whole. Australian Journal of Statistics 7, 67–77.