

Discussion Paper: 2015/05

When is it really justifiable to ignore explanatory variable endogeneity in a regression model?

Jan F. Kiviet

www.ase.uva.nl/uva-econometrics

Amsterdam School of Economics

Roetersstraat 11
1018 WB AMSTERDAM
The Netherlands

UvA  UNIVERSITEIT VAN AMSTERDAM



When is it really justifiable to ignore explanatory variable endogeneity in a regression model?

Jan F. Kiviet*

Version of 10 May 2016

JEL Classifications: *C2; C13; O5*

Keywords:

sensitivity analysis; simultaneity; asymptotic expansions; least-squares; growth regression

Abstract

A procedure that aims to pinpoint the sensitivity of ordinary least-squares based inferences regarding the degree of endogeneity of some regressors has been put forward in Ashley and Parmeter, *Economics Letters*, 137 (2015) 70-74. Here it is demonstrated that this procedure is based on an incorrect and systematically too optimistic asymptotic approximation to the variance of inconsistent least-squares. Therefore, and because the suggested sensitivity findings pertain to a random set of estimated endogeneity correlations, the claimed significance levels are misleading. For a very basic one coefficient model it is demonstrated why much more sophisticated asymptotic expansions under a stricter set of assumptions are required. This enables to replace some of the flawed sensitivity analysis results for an empirical growth model by findings based on a proper limiting distribution for a feasible inconsistency corrected least-squares estimator, given an adopted fixed range for the correlation between the endogenous regressor and the disturbance. Finally it is discussed how similar results could be achieved for models with several possibly endogenous regressors estimated by possibly endogenous instruments.

1. Introduction

An attempt is made in Ashley and Parmeter (2015a), henceforth APLS, to assess the sensitivity of ordinary least-squares (OLS) based tests with respect to simultaneity of an arbitrary number of regressors. The methods used in APLS are basically a specialization

*Emeritus Professor of Econometrics, Amsterdam School of Economics, University of Amsterdam, PO Box 15867, 1001 NJ Amsterdam, The Netherlands (j.f.kiviet@uva.nl). Most of this paper was written while I was enjoying hospitality as Visting Professor at the Division of Economics, School of Humanities and Social Sciences, Nanyang Technological University, Singapore. Comments by a referee are gratefully acknowledged.

of an approach put forward in Ashley and Parmeter (2015b), henceforth APIV, which aims to assess the sensitivity of instrumental variables (IV) based inference in situations where instruments are in fact endogenous. The APLS results are obtained by simply adopting in the APIV approach the regressors as instruments. The APIV study builds on Ashley (2009).

In Kiviet (2013), henceforth KLS, we developed an asymptotically valid inference method on the basis of inconsistent OLS results for models with one endogenous regressor. This would not be possible without making a further identifying assumption. In the KLS approach identification is achieved by making a nonzero moment condition, instead of the habitual zero moment conditions exploited by consistent IV. This approach enables to assess the sensitivity of OLS based t -test inferences over intervals nested in $(-1, +1)$ which should contain the true value of the correlation between the single endogenous regressor and the disturbance term.

APLS uses a similar, yet different, conversion of OLS inference than put forward in KLS. However, both approaches depart from an assessment of the limiting distribution of an unfeasible estimator which corrects OLS for its inconsistency. In this paper we will just focus on the results in APLS for which KLS offers an alternative, i.e. for models with just one endogenous regressor. We will show that the APLS approach suffers from two fundamental flaws. First, its asymptotic analysis is much too naive, which leads to three shortcomings: (i) it does not highlight that numerical assumptions on third and fourth moments are required; (ii) it overlooks that a choice has to be made regarding conditioning or not on exogenous regressors; (iii) the asymptotic variance adopted is systematically too small. Second, the suggested approach leads to an assessment of sensitivity over an adopted set of values for the covariance between regressors and disturbances. To interpret this for the dataset under study, the correlation between regressors and disturbances is estimated on the basis of the adopted covariance and the sample variances of the regressors and the residuals. However, the randomness of these latter estimates is not taken into account, whereas it is in the KLS approach. Therefore, test results obtained by the KLS approach are either tight or conservative, so they can claim asymptotic significance at a level equal to or not exceeding some chosen value, whereas the APLS results cannot.

By analyzing in all detail a very simple case we will show in Section 2 that obtaining the required asymptotic results is much more involved than suggested in APLS. In Section 3 we will compare for an empirical growth model with just one endogenous regressor the sensitivity results obtained in APLS by those produced by the KLS approach. In the concluding Section 4 we put the foregoing into perspective, and indicate some further literature that seems useful for achieving the goals of assessing the sensitivity of OLS and IV with respect to invalid orthogonality conditions for models with possibly more

than just one endogenous regressor.

2. Examination of a simple case

Consider the very simple linear regression model with just one random regressor and i.i.d. observations

$$y = x\beta + \varepsilon, \text{ with } x \sim (0, \sigma_x^2 I) \text{ and } \varepsilon \sim (0, \sigma_\varepsilon^2 I), \quad (2.1)$$

where both y and x are $n \times 1$ observed data vectors. Regressor x could be endogenous. Therefore we suppose that

$$x = \xi + \lambda\varepsilon, \quad (2.2)$$

where ε and $\xi \sim (0, \sigma_\xi^2 I)$ are independent. For $i = 1, \dots, n$ we denote $\sigma_{x\varepsilon} = E(x_i\varepsilon_i) = \rho\sigma_x\sigma_\varepsilon = \lambda\sigma_\varepsilon^2$, thus endogeneity correlation $\rho = \lambda\sigma_\varepsilon/\sigma_x$ with $|\rho| < 1$. Of course, $\sigma_x^2 = \sigma_\xi^2 + \lambda^2\sigma_\varepsilon^2$, thus $\sigma_\xi^2 = (1 - \rho^2)\sigma_x^2$. For $\lambda = 0$ (which implies $\rho = 0$) regressor x is exogenous and endogenous otherwise.

We make standard regularity assumptions, involving that the unknown scalars β , λ , $\sigma_x > 0$ and $\sigma_\varepsilon > 0$ are all finite. So $x'x > 0$ and the ordinary least-squares (OLS) estimator for β exists and is given by

$$b = x'y/x'x = \beta + x'\varepsilon/x'x. \quad (2.3)$$

We shall examine its limiting behavior for $n \rightarrow \infty$.

2.1. Expanding the OLS estimator

For $w_{x\varepsilon} = x'\varepsilon/n - \sigma_{x\varepsilon}$ we find $E(w_{x\varepsilon}) = 0$ and $Var(w_{x\varepsilon}) = n^{-2}\sum_{i=1}^n Var(x_i\varepsilon_i - \sigma_{x\varepsilon}) = n^{-2}\sum_{i=1}^n E[\xi_i\varepsilon_i + \sigma_{x\varepsilon}\sigma_\varepsilon^{-2}(\varepsilon_i^2 - \sigma_\varepsilon^2)]^2$. Assuming further that ε_i has a symmetric distribution with finite kurtosis coefficient κ (hence $E\varepsilon_i^3 = 0$ and $E\varepsilon_i^4 = \kappa\sigma_\varepsilon^4$) then $Var(w_{x\varepsilon}) = n^{-1}[\sigma_\xi^2\sigma_\varepsilon^2 + (\kappa - 1)\sigma_{x\varepsilon}^2] = O(n^{-1})$. Thus, the order of probability of $w_{x\varepsilon}$ is $O_p(n^{-1/2})$, as $n^{1/2}w_{x\varepsilon}$ has a finite distribution. Likewise, assuming $E(x_i^3) = 0$ and $E(x_i^4) = \kappa\sigma_x^4$, we find for $w_{xx} = x'x/n - \sigma_x^2$ that it is $O_p(n^{-1/2})$. In fact,

$$w_{x\varepsilon} \sim (0, n^{-1}[1 + (\kappa - 2)\rho^2]\sigma_\varepsilon^2\sigma_x^2) \text{ and } w_{xx} \sim (0, n^{-1}(\kappa - 1)\sigma_x^4). \quad (2.4)$$

For the inverse of $x'x/n = \sigma_x^2(1 + \sigma_x^{-2}w_{xx})$ we find

$$\begin{aligned} (x'x/n)^{-1} &= \sigma_x^{-2}(1 + \sigma_x^{-2}w_{xx})^{-1} \\ &= \sigma_x^{-2}(1 - \sigma_x^{-2}w_{xx} + \sigma_x^{-4}w_{xx}^2 - \sigma_x^{-6}w_{xx}^3 + \dots) \\ &= \sigma_x^{-2} - \sigma_x^{-4}w_{xx} + O_p(n^{-1}), \end{aligned} \quad (2.5)$$

where we took it for granted that all the omitted terms, which are of decreasing order, have a sum of the same order as the first and largest omitted term, which is $\sigma_x^{-6}w_{xx}^2 = O_p(n^{-1})$.

Substituting (2.5) we find for the OLS estimator (2.3)

$$\begin{aligned} n^{1/2}(b - \beta) &= n^{1/2}(x'x/n)^{-1}x'\varepsilon/n \\ &= (\sigma_x^{-2} - \sigma_x^{-4}w_{xx})n^{1/2}(\sigma_{x\varepsilon} + w_{x\varepsilon}) + O_p(n^{-1/2}) \\ &= n^{1/2}\sigma_{x\varepsilon}/\sigma_x^2 + n^{1/2}(\sigma_x^{-2}w_{x\varepsilon} - \sigma_{x\varepsilon}\sigma_x^{-4}w_{xx}) + O_p(n^{-1/2}), \end{aligned}$$

where the term $-n^{1/2}\sigma_x^{-4}w_{xx}w_{x\varepsilon}$ could be absorbed in the remainder term because $w_{xx}w_{x\varepsilon} = O_p(n^{-1})$. This yields the expansion

$$n^{1/2}(b - \sigma_{x\varepsilon}/\sigma_x^2 - \beta) = n^{1/2}(\sigma_x^{-2}w_{x\varepsilon} - \sigma_{x\varepsilon}\sigma_x^{-4}w_{xx}) + O_p(n^{-1/2}), \quad (2.6)$$

for the infeasible consistent estimator $b - \sigma_{x\varepsilon}/\sigma_x^2$.

2.2. Limiting distributions

Because $\sigma_x^{-2}w_{x\varepsilon} - \sigma_{x\varepsilon}\sigma_x^{-4}w_{xx}$ can be written as the sample average of n i.i.d. random drawings with zero mean and finite variance a standard Central Limit Theorem can be applied. Therefore the limiting distribution of (2.6) will be normal. Its variance is given by the variance of the leading term of (2.6), which has a finite distribution. It can be found from (2.4) and from

$$\begin{aligned} E(w_{x\varepsilon}w_{xx}) &= E[(x'\varepsilon/n - \sigma_{x\varepsilon})(x'x/n - \sigma_x^2)] \\ &= n^{-2}E(x'\varepsilon x') - n^{-1}\sigma_{x\varepsilon}E(x'x) - n^{-1}\sigma_x^2E(x'\varepsilon) + \sigma_{x\varepsilon}\sigma_x^2 \\ &= n^{-1}[2\rho + (\kappa - 3)\rho^3]\sigma_x^3\sigma_\varepsilon, \end{aligned} \quad (2.7)$$

as $E(x'\varepsilon) = n\sigma_{x\varepsilon}$, $E(x'x) = n\sigma_x^2$ and

$$\begin{aligned} E(x'\varepsilon x') &= E[\sum_{i=1}^n(\xi_i\varepsilon_i + \lambda\varepsilon_i^2)\sum_{j=1}^n(\xi_j^2 + 2\lambda\xi_j\varepsilon_j + \lambda^2\varepsilon_j^2)] \\ &= 2\lambda n\sigma_\xi^2\sigma_\varepsilon^2 + \lambda n^2\sigma_\varepsilon^2\sigma_\xi^2 + \lambda^3[n(n-1) + \kappa n]\sigma_\varepsilon^4 \\ &= \rho(1 - \rho^2)(2n + n^2)\sigma_x^3\sigma_\varepsilon + \rho^3[n^2 + (\kappa - 1)n]\sigma_x^3\sigma_\varepsilon \\ &= n[(n + 2)\rho + (\kappa - 3)\rho^3]\sigma_x^3\sigma_\varepsilon. \end{aligned}$$

This yields $Var(\sigma_x^{-2}w_{x\varepsilon} - \sigma_{x\varepsilon}\sigma_x^{-4}w_{xx}) = n^{-1}[1 + (2\kappa - 5)\rho^2 - (\kappa - 3)\rho^4]\sigma_\varepsilon^2\sigma_x^{-2}$, and thus, when $\kappa = 3$ (so especially when x_i and ε_i are jointly normal), then

$$n^{1/2}(b - \sigma_{x\varepsilon}/\sigma_x^2 - \beta) \xrightarrow{d} \mathcal{N}(0, (1 - \rho^2)\sigma_\varepsilon^2/\sigma_x^2). \quad (2.8)$$

It seems that APLS in their formula (7), when specialized to the one regressor case, use the same infeasible inconsistency expression $\sigma_{x\varepsilon}/\sigma_x^2$ for the correction of the OLS

estimator.¹ However, this is not the case.² They consider in their formula (6) an estimator that in the present simple one regressor context is equivalent to

$$\tilde{b} = (x'y - n\sigma_{x\varepsilon})/x'x = \beta + w_{x\varepsilon}/(x'x/n). \quad (2.9)$$

Employing (2.5) we find the expansion

$$\begin{aligned} n^{1/2}(\tilde{b} - \beta) &= n^{1/2}(\sigma_x^{-2} - \sigma_x^{-4}w_{xx})w_{x\varepsilon} + O_p(n^{-1/2}) \\ &= n^{1/2}\sigma_x^{-2}w_{x\varepsilon} + O_p(n^{-1/2}). \end{aligned} \quad (2.10)$$

Since $\text{var}(n^{1/2}\sigma_x^{-2}w_{x\varepsilon}) = [1 + (\kappa - 2)\rho^2]\sigma_\varepsilon^2/\sigma_x^2$, for $\kappa = 3$ this yields limiting distribution

$$n^{1/2}(\tilde{b} - \beta) \xrightarrow{d} \mathcal{N}(0, (1 + \rho^2)\sigma_\varepsilon^2/\sigma_x^2). \quad (2.11)$$

Hence, the by APLS naively chosen expression $\sigma_\varepsilon^2/\sigma_x^2$ for the limiting variance of $n^{1/2}(\tilde{b} - \beta)$ is wrong, and under simultaneity it is systematically too small when $\kappa > 2$. Note that choosing (2.8) as a starting point seems much more attractive because its limiting distribution has a systematically smaller variance.

Due to skipping a formal derivation of (2.11) APLS do not realize either that assumptions on third and fourth moments are required. Moreover, they do not recognize that it matters whether or not one conditions on the random exogenous component ξ of the vector x . In KLS it has been shown that when conditioning on ξ , while assuming normality of x and ε , one obtains

$$n^{1/2}(b - \sigma_{x\varepsilon}/\sigma_x^2 - \beta) \xrightarrow{d} \mathcal{N}(0, (1 - \rho^2)(1 - 2\rho^2 + 2\rho^4)\sigma_\varepsilon^2/\sigma_x^2), \quad (2.12)$$

which is even more attractive than (2.8), because $1 - 2\rho^2 + 2\rho^4 \leq 1$ for $|\rho| < 1$. However, in what follows we will focus on the unconditional case; note that this will yield results which are conservative (over-cautious) for the conditional case.

For the unconditional case it has been derived in KLS as well that for the feasible consistent estimator b_ρ^* , which is obtained by correcting b using the actual value of ρ and consistently estimating σ_ε^2 and σ_x^2 , one has

$$n^{1/2}(b_\rho^* - \beta) \xrightarrow{d} \mathcal{N}(0, \sigma_\varepsilon^2/\sigma_x^2), \quad (2.13)$$

where

$$\begin{aligned} b_\rho^* &= b - n^{1/2}\rho(1 - \rho^2)^{-1/2}se(b), \quad \text{with} \\ se(b) &= s/(x'x)^{1/2} \quad \text{and} \quad s^2 = (y - xb)'(y - xb)/(n - k). \end{aligned} \quad (2.14)$$

¹Note that when $\Sigma_{X\varepsilon}$ is introduced in (6) of APLS there is a confusing typo: $E(X'_i\varepsilon_i)$ should read $E(X'\varepsilon)$.

²The term $-E_{XX}^{-1}\Sigma_{X\varepsilon}$ should be removed from formula (7) in APLS, because estimator $\tilde{\gamma}$ has already been corrected. However, this has been corrected by $-(X'X/n)^{-1}\Sigma_{X\varepsilon}$.

Here $se(b)$ is the expression for the usual OLS standard error estimate of b as produced under assumed exogeneity of x . In the present special model k (the number of regressors) is one, but for the asymptotic result to hold a degrees of freedom correction is irrelevant, of course. Estimator b_ρ^* is consistent for β , because $\sigma_{x\varepsilon}/\sigma_x^2 = \rho(\sigma_\varepsilon^2/\sigma_x^2)^{1/2}$, whereas $x'x/n$ is consistent for σ_x^2 and, as has been proved in KLS, $s^2/(1-\rho^2)$ is a consistent estimator of σ_ε^2 . Note that $n^{1/2}se(b) = O_p(1)$ and therefore the inconsistency correction term in b_ρ^* is finite too and vanishes only for $\rho = 0$.

One may find it inappropriate that above estimator b_ρ^* is addressed as a feasible estimator, whereas in practice the actual value of ρ is commonly unknown. A not unreasonable response to that is: OLS/IV estimators are usually not labelled as unfeasible either, whereas their underlying (just) identifying orthogonality conditions simply adopt the value zero for the correlation between regressor(s)/instrument(s) and disturbance, for which it is equally difficult or sheer impossible to find statistical evidence, see Kiviet (2016).

Result (2.13) is quite remarkable, not because consistent feasible estimator b_ρ^* is found to have, due to estimating $(\sigma_\varepsilon^2/\sigma_x^2)^{1/2}$, a larger asymptotic variance than the unfeasible estimator $b - \sigma_{x\varepsilon}/\sigma_x^2$, but because this increment exactly leads to the same asymptotic variance as b has when it is consistent. However, this equivalence only holds (unconditionally) under the special assumptions made regarding skewness and kurtosis for x and ε ; it can be shown that the asymptotic variance of b_ρ^* will be larger in case of excess kurtosis. It is remarkable too that estimating $(\sigma_\varepsilon^2/\sigma_x^2)^{1/2}$ for obtaining consistent estimator b_ρ^* increases the limiting variance of $b - \sigma_{x\varepsilon}/\sigma_x^2$ less than estimating just σ_x^2 for obtaining consistent estimator \tilde{b} .

Summarizing: Unlike KLS, APLS do not fix ρ but $\sigma_{x\varepsilon}$, which seems much less practicable, because this is not dimensionless like ρ is. Moreover, it requires to employ the less attractive limiting distribution (2.11) of consistent estimator \tilde{b} . Using simply, as APLS do, the incorrect standard expression $\mathcal{N}(0, \sigma_\varepsilon^2/\sigma_x^2)$, which does apply to b_ρ^* , will lead to systematically too optimistic inferences.

2.3. Sensitivity of OLS with respect to endogeneity

Asymptotically valid inference on β based on (2.13) can be obtained from the asymptotic approximation

$$b_\rho^* \overset{a}{\sim} \mathcal{N}(\beta, s^2/[(1-\rho^2)x'x]), \quad (2.15)$$

which for $\rho = 0$ simplifies to the standard result $b \overset{a}{\sim} \mathcal{N}(\beta, s^2/x'x)$. Let us focus now on testing $\mathcal{H}_0 : \beta = \beta_0$ against $\mathcal{H}_1 : \beta \neq \beta_0$ for known numerical value β_0 . The usual statistic $t(\beta_0) = (b - \beta_0)/se(b)$ is asymptotically standard normal under \mathcal{H}_0 , provided

$\rho = 0$. From (2.15) it follows that under simultaneity

$$t_\rho(\beta_0) = (1 - \rho^2)^{1/2}(b_\rho^* - \beta_0)/se(b) = (1 - \rho^2)^{1/2}t(\beta_0) - n^{1/2}\rho \quad (2.16)$$

is asymptotically standard normal under \mathcal{H}_0 . This result can be used for a sensitivity analysis as follows. Imagine we find a t -value of 3 in a regression where $n = 100$. Then by solving $3(1 - \rho^2)^{1/2} - 10\rho > 1.96$, which gives $\rho \leq 0.102$, we learn that rejection against a right-hand side alternative at the asymptotic significance level of 2.5% requires a negative or just a very small positive simultaneity correlation.

Note that for finite $t(\beta_0)$ and $n > 4$ we find that for ρ getting closer to +1 we have $t_\rho(\beta_0) < -c$ and for ρ getting closer to -1 we have $t_\rho(\beta_0) > c$, for c some positive critical value. This means that for any OLS test result and any chosen nominal significance level we can always find a range of values for the degree of simultaneity for which the test is asymptotically significant and also one for which it is insignificant. This clearly demonstrates the inadequacy of the APLS approach, because they report in their Table 1 that they did establish OLS inferences which were found to be robust to any degree of simultaneity ($r_{\min} = 1$).

As we understand it, the essential elements of the APLS sensitivity analysis in the context of the above simple one coefficient model involves the following. Suppose one has found for a particular data set that $|t(\beta_0)| > z_{1-\alpha/2}$, where z_p denotes for $0 < p < 1$ the p^{th} quantile of the standard normal distribution. So \mathcal{H}_0 is rejected at level α , presupposing exogeneity. Now the method seeks to establish the values ρ for which \mathcal{H}_0 would still be rejected. This is done as follows. The set of values for $\sigma_{x\varepsilon}$ (here scalar), say $\mathcal{S}(\sigma_{x\varepsilon})$, is assessed (by a random search procedure), giving $\mathcal{S}(\sigma_{x\varepsilon}) = \left\{ \sigma_{x\varepsilon} \in \mathbb{R} : z_{\alpha/2} < (\tilde{b} - \beta_0)/se(\tilde{b}) < z_{1-\alpha/2} \right\}$, where

$$\begin{aligned} \tilde{b} &= b - \sigma_{x\varepsilon}/(x'x/n) \text{ with} \\ se(\tilde{b}) &= \hat{\sigma}_{\sigma_{x\varepsilon}}/(x'x)^{1/2} \text{ and } \hat{\sigma}_{\sigma_{x\varepsilon}}^2 = (y - x\tilde{b})'(y - x\tilde{b})/n. \end{aligned} \quad (2.17)$$

When the numerical problem to assess $\mathcal{S}(\sigma_{x\varepsilon})$ has been solved, the corresponding set of values for $\hat{\rho}$ given by $\mathcal{S}(\hat{\rho}) = \left\{ \hat{\rho} = \sigma_{x\varepsilon}/[\hat{\sigma}_{\sigma_{x\varepsilon}}^2(x'x/n)]^{1/2} : \sigma_{x\varepsilon} \in \mathcal{S}(\sigma_{x\varepsilon}) \right\}$ is assessed, and the conclusion is drawn that rejection of \mathcal{H}_0 is robust with respect to endogeneity for all values covered by $\mathcal{S}(\hat{\rho})$.

Hence, a crucial difference with KLS is that a choice is made regarding $\sigma_{x\varepsilon}$ and not with respect to ρ directly. The price for that is that APLS find a random set $\mathcal{S}(\hat{\rho})$ for ρ , and omit to discuss the consequences of this randomness (which is not due to the random search, but to the dependence of $\hat{\rho}$ on $\hat{\sigma}_{\sigma_{x\varepsilon}}^2$ and on $x'x$). Moreover, as shown in the foregoing section, the formula used for $se(\tilde{b})$ in (2.17) is systematically too small.

3. Empirical illustration

The methods developed in APLS have been applied to a particular empirical growth model presented in Mankiw et al. (1992) where possible endogeneity of regressors has not been taken into account. This model for 98 countries can be represented as $y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \varepsilon_i$, where y_i is per capita output, x_{i2} is the rate of human capital, x_{i3} is investment in physical capital and x_{i4} is the logarithm of the sum of the population growth rate, the growth rate in technology and the depreciation rate. The focus in the APLS analysis is on testing $\beta_2 = 0$ and $\beta_2 + \beta_3 + \beta_4 = 1$ (constant returns to scale). APLS find OLS results as given in their equation (8), which are slightly different from those in Mankiw et al. (1992, p.420, Table II, first column). APLS perform a sensitivity analysis for four different scenarios, which allow either x_{i2} , or x_{i3} or x_{i4} to be endogenous, or x_{i2} and x_{i3} are jointly assumed endogenous.

The KLS results for regression models with just one explanatory variable that may be endogenous can also serve the situation where the model has some further exogenous regressors, which have all been partialled out. This means that only for the first scenario of APLS (x_{i2} endogenous) we can produce KLS results relevant for testing β_2 . For these access to the observations on the regressand and regressors is not required. The OLS coefficient estimates and standard errors presented in APLS equation (8) suffice to analyze $t_\rho(\beta_0)$ of (2.16). Just for illustrative purposes we also present in Table 1 results relevant on testing β_3 allowing x_{i3} to be endogenous and on β_4 allowing x_{i4} to be endogenous.

Table 1: KLS sensitivity of significance at 2.5% of growth coefficients

alternative hypothesis	$\beta_2 > 0$	$\beta_3 > 0$	$\beta_4 < 0$
admissible values for ρ	$\rho \leq 0.572$	$\rho \leq 0.311$	$\rho \geq -0.222$

The parallel finding in APLS regarding β_2 is: $r_{\min} = 0.425$ at 5% (from their text it is not clear whether they confront the test statistic with critical value 1.96 or 1.64; when we test the significance of β_2 at 5% critical value 1.64 the result is $\rho \leq 0.591$). Note that there is also a range of ρ values for which the observed OLS t -ratio's imply significance of opposite sign. For $\rho \geq 0.782$, which would lead to a huge positive bias of b and thus of $t(\beta_0)$, the empirical results corroborate with $\beta_2 < 0$ at 2.5%.

4. How to tackle more general cases?

From the above it should be clear that the KLS approach has yet been developed for only very few special simple cases, whereas the APLS approach has serious flaws in various of its underpinnings. Even when employing sound asymptotics, the latter will lead

to findings which are hard to interpret because the resulting r_{\min} vector is random and obtaining an asymptotic approximation to its distribution to assess its accuracy seems far from easy. On the other hand, further development of the KLS approach for more general cases seems possible, but requires an asymptotic analysis which certainly cannot be characterized as in APLS (just above their section 2.2) as "easy" and "straightforward", because of the following three reasons.

First, although the corrected estimator (6) in APLS is clearly consistent, it is not self-evident what its limiting distribution will look like unless one embarks on a proper derivation. Doing so reveals that extra assumptions are required regarding the numerical values of third and fourth moments of the regressors and disturbance. Second, an inconsistent estimator has a bias which is $O(1)$, and hence any useful random assessment of this bias will be $O_p(1)$, and thus employing such an assessment to correct the inconsistent estimator will affect its limiting distribution and require a further expansion.³ Third, a further complicating issue is that whereas the limiting distribution of standard consistent estimators is similar whether or not one conditions on exogenous variables, this situation apparently changes when consistency is achieved by correcting an inconsistent estimator.

Many aspects of these complications for regression models with an arbitrary number of endogenous and exogenous regressors have already been addressed in Kiviet and Niemczyk (2012) with respect to OLS estimation and in Kiviet and Niemczyk (2014) with respect to IV estimation when invalid instruments may have been used. A next step should be to obtain for these more general settings the limiting distributions of inconsistency corrected estimators which are feasible in the sense of KLS, and next exploit these to produce inference which is valid over a credible set of values for all endogeneity correlations. Only after this has been achieved for empirically relevant models it seems justifiable to provide a fully satisfactory answer to the question posed in the titles of APLS and of this study. Other approaches to achieve a similar goal in one way or another can be found in Murray (2006), Small (2007), Ebbes et al. (2009), Conley et al. (2012) and Kraay (2012).

References

Ashley, R., 2009. Assessing the credibility of instrumental variables inference with imperfect instruments via sensitivity analysis. *Journal of Applied Econometrics* 24, 325-337.

³When a consistent estimator is biased, its bias is usually $O(n^{-1})$. The bias of the estimator can be reduced by subtracting an $O_p(n^{-1})$ assessment of this bias. In that case the corrected estimator retains the same limiting distribution as the uncorrected estimator, because the correction just affects higher-order asymptotic aspects.

Ashley, R.A., Parmeter, C.F., 2015a. When is it justifiable to ignore explanatory variable endogeneity in a regression model? *Economics Letters* 137, 70-74.

Ashley, R.A., Parmeter, C.F., 2015b. Sensitivity analysis for inference in 2SLS estimation with possibly-flawed instruments. *Empirical Economics* 49, 1153-1171.

Conley, T., Hansen, C, Rossi, P., 2012. Plausibly exogenous. *The Review of Economics and Statistics* 94, 260-272.

Ebbes, P., Wedel, M., Bockenholt, U., 2009. Frugal IV alternatives to identify the parameter for an endogenous regressor. *Journal of Applied Econometrics* 24, 446-468.

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. Discriminating between (in)valid external instruments and (in)valid exclusion restrictions. UvA-Econometrics working paper 2015/04, revised 5 May 2016. Forthcoming in the *Journal of Econometric Methods*.

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., Niemczyk, J., 2014. On the limiting and empirical distribution of IV estimators when some of the instruments are actually endogenous, pp.425-490 in *Advances in Econometrics, Volume 33; Essays in Honor of Peter C.B. Phillips* (Eds. Yoosoon Chang, Thomas B. Fomby, Joon Y. Park), Emerald Group Publishing Limited, Bingley, UK.

Kraay, A., 2012. Instrumental variables regressions with uncertain exclusion restrictions: A Bayesian approach. *Journal of Applied Econometrics* 27, 108-128.

Mankiw, N.G., Romer, D., Weil, D.N., 1992. A contribution to the empirics of economic growth. *Quarterly Journal of Economics* 107, 407-437.

Murray, M. P., 2006. Avoiding invalid instruments and coping with weak instruments. *Journal of Economic Perspectives* 20, 111-132.

Small, D. S., 2007. Sensitivity analysis for instrumental variables regression with overidentifying restrictions. *Journal of the American Statistical Association* 102, 1049-1058.