The Stata Journal (2018)
18, Number 3, pp. 663–691

Practical considerations for questionable IVs

Damian Clarke
Department of Economics
Universidad de Santiago de Chile
damian.clarke@usach.cl

Benjamín Matta
Department of Economics
Universidad de Santiago de Chile
benjamin.matta@usach.cl

Abstract. In this article, we examine several techniques that allow for the construction of bounds estimates based on instrumental variables, even when the instruments are not valid. We introduce the `plausexog` and `imperfectiv` commands, which implement methods described by Conley, Hansen, and Rossi (2012, *Review of Economics and Statistics* 94: 260–272) and Nevo and Rosen (2012b, *Review of Economics and Statistics* 94: 659–671). We examine the performance of these bounds under a range of circumstances, which leads to several practical results related to the informativeness of the bounds in different situations.

Keywords: st0538, plausexog, imperfectiv, instrumental variables, exclusion restrictions, invalidity, plausibly exogenous, imperfect IVs

1 Introduction

Instrumental variables (IVs) are a workhorse estimator in economics and in other fields concerned with the causal estimation of relationships of interest. Nonetheless, credible IVs are hard to come by. While finding variables that are correlated with an endogenous variable of interest ("relevant" in IV terms) is generally not a challenge, motivating and defending a zero correlation with unobserved error terms ("validity") is much less straightforward.[1]

As is well known, validity assumptions in an IV setting are untestable. While partial tests exist (Sargan 1958; Hansen 1982; Kitagawa 2015), these tests are necessary, rather than sufficient, to demonstrate instrumental validity. This often leads to the uncomfortable position where the best estimates for a parameter are based on a strong assumption for which no definitive proof can be offered.

In this article, we examine several recent methodologies for inference with instruments that (potentially) fail the typical IV validity assumption. In particular, we focus on two methods that provide bounds on an endogenous variable of interest with as few as one IV that does not necessarily have zero correlation with the unobserved error term. These methodologies—one from Conley, Hansen, and Rossi (2012) and one from Nevo and Rosen (2012b)—loosen IV assumptions in different ways and are relevant to different types of settings in which IVs are suspected not to hold precisely. As we lay out in further detail below, Conley, Hansen, and Rossi (2012) replace the (exact) exclusion restriction in an IV model with an assumption related to its support or distribution,

1. We lay out the classical IV model in section 2 and the traditional assumptions leading to consistent estimates of parameters of interest.

 st0538

whereas Nevo and Rosen (2012b) replace the zero correlation assumption between the instrument and the unobserved error term with an assumption related to the "sign" of the correlation.

IV bounds under weaker-than-standard assumptions are potentially of use in a wide range of applications. Much effort is often spent in empirical work to convincingly argue for the validity of instruments. Nonetheless, the validity of IVs are often questioned. Consider the survey paper of Rosenzweig and Wolpin (2000), which describes several "natural" IVs that are not under the control of humans and hence have been proposed to be valid IVs.[2] Among those listed, most have been questioned on various grounds. The use of season of birth (Angrist and Krueger 1991) was suggested to be potentially correlated with many relevant correlates (Bound, Jaeger, and Baker 1995) and then documented to be directly related to maternal characteristics in the United States (Buckles and Hungerman 2013). The use of twins (Rosenzweig and Wolpin 1980a,b) was later questioned based on birth spacing and parental responses (Rosenzweig and Zhang 2009) and parental behavior in utero (Bhalotra and Clarke 2016); the use of the gender mix of children (Angrist and Evans 1998) was shown to have other relevant effects on family behavior (Dahl and Moretti 2008).

However, often critiques of IVs imply minor, rather than major, correlations between instruments and unobserved behavior. In this article, we introduce two Stata commands that permit for the construction of valid bounds in circumstances precisely like this: The `plausexog` command, based on Conley, Hansen, and Rossi's (2012) plausibly exogenous inference, and the `imperfectiv` command, based on Nevo and Rosen's (2012b) imperfect IV inference. These methods allow for the construction of IV bounds under weaker-than-traditional assumptions. We lay out the basics of each methodology and the use of each command and discuss several factors to be considered when confronted with questionable IVs. As we show, the relative informativeness of `plausexog` and `imperfectiv` bounds depends on the particular context, with each being particularly suitable in different (invalid) IV circumstances. In the remainder of this article, we document the scope of each procedure and suggest that these commands should be considered as complements, rather than substitutes, in the applied researcher's toolbox.

2 Methodology

The habitual linear IVs model is laid out as follows,

$$\begin{aligned} \mathbf{Y} &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (1) \\ \mathbf{X} &= \mathbf{Z}\Pi + \mathbf{V} \quad (2) \end{aligned}$$

where $\mathbf{Y}$ is an outcome variable of interest, $\mathbf{X}$ a matrix of (potentially endogenous) treatment variables, and $\mathbf{Z}$ a matrix of instruments that are uncorrelated by assumption with the error term $\boldsymbol{\varepsilon}$. Presuming that $\mathbf{X}$ contains an endogenous variable (or variables),

2. In particular, they listed five outcomes arising from natural (biological or climate) processes that were potentially random and had been used as instruments. These were i) twin births, ii) human cloning (monozygotic twinning), iii) birth date, iv) gender, and v) weather events.

the parameter vector $\boldsymbol{\beta}$ is not consistently estimable with ordinary least squares (OLS). The existence of valid instruments $\mathbf{Z}$, which can be excluded from (1), thus drives the estimation of the structural parameters of interest $\boldsymbol{\beta}$.

Validity is typically presented in one of two formats. The first is in terms of the exclusion restriction: the instruments $\mathbf{Z}$ have no direct effect on $\mathbf{Y}$ once purged of their effect on $\mathbf{X}$. The second is in terms of correlations with unobservables: if $\mathbf{Z}$ is uncorrelated with $\boldsymbol{\varepsilon}$, instrumental validity is fulfilled. While either condition is appropriate to motivate consistent estimation of parameters in IV models,[3] we consider both here because they provide alternative approaches to conceptualize failures of the underlying assumption in IVs.[4] If it can be credibly argued that the validity assumption holds, two-stage least squares (2SLS) estimates of $\boldsymbol{\beta}$ from (1) are consistent.

However, as discussed in the introduction, this validity assumption is untestable given that it is related to the behavior of the unobservable $\boldsymbol{\varepsilon}$. Even if instruments are shown to be unrelated to many observable factors, or to pass overidentification tests, this does not provide definitive proof of their validity. This has given rise to a modern literature focused on relaxing these assumptions. Work by Manski and Pepper (2000, 2009) loosened the validity assumption, replacing strict equalities with (weak) inequalities. This work has been extended by, among others, Conley, Hansen, and Rossi (2012); Nevo and Rosen (2012b) propose linear[5] models in an IV framework but with the absence of the traditional IV validity assumption. Rather than driving estimation and inference from dogmatic priors, which require strict equalities in the exclusion restriction or correlations, it has been shown that bounds on parameters can be estimated under considerably loosened conditions.

While both Conley, Hansen, and Rossi (2012) and Nevo and Rosen (2012b) suggest ways of loosening traditional assumptions to form IV bounds with as few as one (invalid)

3. Indeed, their implications are equivalent in the simultaneous equations framework laid out here (additional discussion related to their difference in the potential-outcomes interpretation of the Rubin [1974] casual model can be found in Angrist and Pischke [2009, 85–91]). If we consider two structural equations of the form

$$y = \boldsymbol{\beta}_0^a + \boldsymbol{\beta}_1^a \mathbf{X} + \boldsymbol{\varepsilon}^a$$

and

$$y = \boldsymbol{\beta}_0^b + \boldsymbol{\beta}_1^b \mathbf{X} + \boldsymbol{\beta}_2^b \mathbf{Z} + \boldsymbol{\varepsilon}^b$$

failure of the exclusion restriction means that $\boldsymbol{\beta}_2^b$ is not equal to zero. However, a nonzero value of $\boldsymbol{\beta}_2^b$ also implies that $\rho_{Z,\varepsilon^a} \neq 0$ (where ρ is the covariance) and, by definition, $\rho_{y,\varepsilon^a} > 0$. Thus, assuming that the exclusion restriction holds in this setup is equivalent to assuming that $\mathbf{Z}$ is uncorrelated with the structural error term. And vice versa, once the conditional correlation between the instrument and the error term is assumed to be zero, the exclusion restriction assumption is superfluous.

4. In this article, we do not consider in much length the relevance assumption. This assumption is testable, and considerable literature exists on this topic.

5. While these methods are exclusively presented in terms of linear IV in this article, the underlying logic can extend to nonlinear models. One particular method is provided in Conley, Hansen, and Rossi (2008), who show a nonlinear extension to relax the exclusion restriction assumption. A benefit of restricting our analysis to a linear IV setup here is that this allows for bounds to be produced with clear links to frequently used (linear) models such as 2SLS and OLS and the `regress` and `ivregress` commands available within Stata.

IV,[6] the precise manner in which this is undertaken in each case is different. The suggestion of Conley, Hansen, and Rossi (2012) is to relax the exclusion restriction, where rather than assuming that it is exactly equal to zero, some range is allowed for the coefficient on the instrument in the structural equation. They allow the exclusion restriction to fail but proceed with estimation by restricting the failure to some range. Nevo and Rosen (2012b), on the other hand, document that assuming a "direction" for the covariance between the instrument and the stochastic error $\boldsymbol{\varepsilon}$ can result in two-sided bounds for the parameter of interest $\boldsymbol{\beta}$. We consider each method, as well as the resulting bounds below, before turning to the practicalities of estimation later in this article.

Relaxing the exclusion restriction assumption The classical IV system of equations defined in (1) and (2) is a restricted version of the following:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \tag{3}$$

$$\mathbf{X} = \mathbf{Z}\Pi + \mathbf{V} \tag{4}$$

We arrive at (1) and (2) by imposing the (strong) prior that $\boldsymbol{\gamma} = \mathbf{0}$, resulting in point estimates of the parameter vector of interest $\boldsymbol{\beta}$. One way to loosen the IV assumptions is to remove the assumption that $\boldsymbol{\gamma}$ is precisely equal to zero. A range of literature seeks to restrict the range of this unidentified parameter (or parameter vector) $\boldsymbol{\gamma}$ without assuming that it is exactly equal to zero. Manski and Pepper (2000) document inference in IV settings where the strict equality in $\boldsymbol{\gamma} = \mathbf{0}$ is replaced by a weak inequality, giving "monotone IVs".[7] Earlier work by Hotz, Mullin, and Sanders (1997) proposes bounding in an IV setting where the exclusion restriction is assumed to hold for some part of the population and not hold for others, requiring an estimate or assumption regarding the degree of contamination of the IV. More recent extensions, including Small (2007) and Conley, Hansen, and Rossi (2012), seek to further restrict the range of values for $\boldsymbol{\gamma}$ while still allowing the exclusion restriction to fail either by searching for plausible parameters in overidentified systems (Small 2007) or by allowing researchers to specify priors for $\boldsymbol{\gamma}$ in a range of flexible ways (Conley, Hansen, and Rossi 2012).

In what remains, when considering relaxations of the exclusion restriction, we will follow the procedure implemented by Conley, Hansen, and Rossi (2012). This procedure allows for valid inference using an IV (or variables) even when the exclusion restriction does not hold precisely. They document several procedures that can be followed, depending on a researcher's prior belief regarding the degree of failure of the exclusion restriction and the amount of structure that the researcher is willing to place on this violation. In particular, assumptions can be made regarding the range of values that

6. However, there is also an alternative set of methodologies proposing inference in an IV framework without strict validity assumptions, using more than one (invalid) IV. For example, Small (2007) proposes a case with as few as two instruments, and Kolesár et al. (2015) and Kang et al. (2016) describe estimation procedures with many invalid or invalid and valid instruments.

7. Strictly speaking, Manski and Pepper's approach does not require (3) and (4), because it is based in a nonparametric setting, where instruments are assumed to monotonically impact conditional expectations, and so involves conditional means rather than covariances.

$\boldsymbol{\gamma}$ can take in (3), regarding the entire distribution for $\boldsymbol{\gamma}$, or a fully Bayesian approach can be undertaken, in which, as well as a prior for the $\boldsymbol{\gamma}$ term, priors for each model parameter and the distribution of error terms must be provided.

The first of these approaches consists of simply replacing the original exclusion restriction assumption of $\boldsymbol{\gamma} = \mathbf{0}$ with an assumption regarding the minimum and maximum values that $\boldsymbol{\gamma}$ may take. This allows for circumstances in which $\boldsymbol{\gamma}$ can be assumed to be entirely positive or negative, or alternatively, overlapping zero. Estimation thus consists of producing confidence intervals on $\boldsymbol{\beta}$ for a range of models of the following form, where $\boldsymbol{\gamma}_0$ refers to values from an (appropriately binned) range $[\boldsymbol{\gamma}_{\min}, \boldsymbol{\gamma}_{\max}]$.

$$(\mathbf{Y} - \mathbf{Z}\boldsymbol{\gamma}_0) = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

In each case, the above model can be fit by 2SLS using the transformed dependent variable $\mathbf{Y} - \mathbf{Z}\boldsymbol{\gamma}_0$. Conley, Hansen, and Rossi (2012) name this approach the union of confidence intervals (UCI), because in practice, bounds consist of the union of all confidence intervals in the assumed range of $\boldsymbol{\gamma}_0 \in [\boldsymbol{\gamma}_{\min}, \boldsymbol{\gamma}_{\max}]$. In the case that more than one plausibly exogenous IV exists, the above procedure is followed with priors over $\boldsymbol{\gamma}_0$ for each instrument; thus, $\boldsymbol{\gamma}_0$ is a vector rather than a scalar. More importantly, nothing restricts these priors over $\boldsymbol{\gamma}_0$ to be identical for different instruments, either in magnitude or in sign.

Additional structure can be placed on assumptions regarding $\boldsymbol{\gamma}$ to relax the exclusion restriction. If, rather than assuming simple maximum and minimum values for $\boldsymbol{\gamma}$, one makes a distributional assumption, bounds on the parameter $\boldsymbol{\beta}$ can be calculated using the entire assumed distribution for $\boldsymbol{\gamma}$. This allows, among other things, for more or less weight to be placed on values of $\boldsymbol{\gamma}$, which are perceived to be more or less likely, for example, by placing more weight on values of $\boldsymbol{\gamma}$ close to zero and less weight on values of $\boldsymbol{\gamma}$ further away.[8] As Conley, Hansen, and Rossi (2012) document, replacing the assumption that $\boldsymbol{\gamma} = \mathbf{0}$ with an assumption that $\boldsymbol{\gamma} \sim F$ (where F is some arbitrary distribution) implies the following approximate distribution for $\widehat{\boldsymbol{\beta}}$:

$$\widehat{\boldsymbol{\beta}} \overset{a}{\sim} \mathcal{N}(\boldsymbol{\beta}, V_{2\text{SLS}}) + A\boldsymbol{\gamma} \tag{5}$$

Here the original 2SLS asymptotic distribution is inflated by a second term, where $A = \{X'X(Z'Z)^{-1}Z'X\}^{-1}(X'Z)$ and $\boldsymbol{\gamma}$ is assumed to follow some arbitrary distribution F, assumed independent of $\mathcal{N}(\boldsymbol{\beta}, V_{2\text{SLS}})$. This approach is called the local to zero (LTZ) approximation and treats uncertainty regarding $\boldsymbol{\gamma}$ and sampling uncertainty as of a similar magnitude.

Practically, estimating bounds on $\boldsymbol{\beta}$ using the result in (5) can proceed in multiple ways. A simulation-based approach can be used that allows for any type of distribution for $\boldsymbol{\gamma}$, or if $\boldsymbol{\gamma}$ is assumed to have a Gaussian distribution, this leads to a convenient

8. Conley, Hansen, and Rossi (2012) also discuss how this can be housed in the union of confidence interval approach discussed above by giving more or less weight to certain values in the $[\boldsymbol{\gamma}_{\min}, \boldsymbol{\gamma}_{\max}]$ range. However, the present approach allows for the flexibility to easily include any distributions for $\boldsymbol{\gamma}$, and so we focus on this here.

analytical bounds formula for $\boldsymbol{\beta}$. In the case that $\boldsymbol{\gamma}$ is assumed to follow a Gaussian distribution, $\mathcal{N}(\boldsymbol{\mu}_\gamma, \boldsymbol{\Omega}_\gamma)$, bounds on $\boldsymbol{\beta}$ from (5) simplify to

$$\widehat{\boldsymbol{\beta}} \overset{a}{\sim} \mathcal{N}(\boldsymbol{\beta} + A\boldsymbol{\mu}_\gamma, V_{2\text{SLS}} + A\boldsymbol{\Omega}_\gamma A')$$

As in the UCI case, if multiple instruments are available, both $\boldsymbol{\mu}_\gamma$ and $\boldsymbol{\Omega}_\gamma$ refer to the distributional assumptions for each $\boldsymbol{\gamma}$ term, where particular priors over the violation of the exclusion restriction are allowed to vary for different instruments.[9] If a non-Gaussian prior for $\boldsymbol{\gamma}$ is assumed, Conley, Hansen, and Rossi (2012) outline a simulation algorithm for calculating bounds on $\boldsymbol{\beta}$. This procedure consists of generating draws of the following quantity, which calculates deviations of $\widehat{\boldsymbol{\beta}}$ from $\boldsymbol{\beta}$, where draws from the assumed $\boldsymbol{\gamma}$ distribution are included in the second part of the formula:

$$\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, V_{2\text{SLS}}) + A\boldsymbol{\gamma}$$

In practice, with numerous draws of $\boldsymbol{\eta}$ in hand, confidence intervals on $\boldsymbol{\beta}$ can be found by subtracting desired quantiles of the $\boldsymbol{\eta}$ distribution from $\widehat{\boldsymbol{\beta}}$ in (5). Both the exact and simulation-based method can be implemented using the `plausexog` command described in further detail later in this article.[10]

Finally, even further structure can be placed on the exclusion restriction if rather than simply assuming a range of values for $\boldsymbol{\gamma}$ (UCI), or a distribution for $\boldsymbol{\gamma}$ (LTZ), one follows a full Bayesian procedure. This requires assuming not only a distribution for $\boldsymbol{\gamma}$ but also a prior for error terms and other model parameters. We do not go into additional detail regarding this Bayesian procedure here; however, we direct interested readers to Conley, Hansen, and Rossi (2012) and computational implementations (in R) as *bayesm* (Rossi 2017).

In the methods described by Conley, Hansen, and Rossi (2012), prior beliefs over the violation of the exclusion restriction play an important role in the eventual bounds estimates. Deciding precisely which values to indicate as priors is an empirical consideration and will vary considerably depending on the plausibility of IVs and posited reasons why an exclusion restriction may not hold. As Conley, Hansen, and Rossi suggest, these beliefs are likely to vary by researchers, pointing to the importance of sensitivity analyses related to estimated bounds. While it is not possible to provide a general rule for setting priors related to the exclusion restriction, it is often the case that researchers do hold subjective beliefs about the exclusion restriction and hypotheses about why it may

9. If multiple instruments are used, there is no limit on the way that priors for $\boldsymbol{\gamma}$ need be specified. This includes cases where multiple instruments may be thought to suffer different failures of the exclusion restriction in sign or in magnitude (by varying parameters in the $\boldsymbol{\mu}_\gamma$ vector) or where the degree of uncertainty for one instrument may be more than the degree of uncertainty for another instrument (varying variance terms in the $\boldsymbol{\Omega}_\gamma$ matrix).

10. While it is preferable to use the exact result if a Gaussian prior is assumed for the distribution of $\boldsymbol{\gamma}$, a Gaussian prior can also be included using the simulation-based algorithm described in Conley, Hansen, and Rossi (2012), and assuming that a large enough number of draws of $\boldsymbol{\eta}$ are taken, these two approaches return identical bounds. By default, `plausexog` draws 5,000 realizations of $\boldsymbol{\eta}$, and this generally leads to similar bounds in the simulated and closed-form approaches with a Gaussian prior. The number of draws of $\boldsymbol{\eta}$ can be changed by users. Where possible, more draws should always be preferred.

not hold precisely. Several cases show how such priors may be formed using economic logic. Bound, Jaeger, and Baker (1995), for example, perform a back-of-the-envelope calculation related to the direct effect of season of birth (a proposed IV) on educational outcomes, which Conley, Hansen, and Rossi (2012) use to form a prior. In examining selectiveness of twin births, Bhalotra and Clarke (2016) aim to directly estimate the degree of violation of the exclusion restriction using additional data, auxiliary to their main analysis. An alternative approach to estimating (rather than assuming) $\boldsymbol{\gamma}$ is suggested by van Kippersluis and Rietveld (2017) by focusing on particular subsamples.

Relaxing IV correlation assumptions The classical IV approach described in (1) and (2) produces consistent estimates of $\boldsymbol{\beta}$ based on the (unobservable) validity assumption $E(Z\boldsymbol{\varepsilon}) = 0$. Bounds inference in an IV setting can proceed with weaker-than-classical assumptions by replacing the validity (zero covariance) assumption with an assumption on the sign of the covariance. Nevo and Rosen (2012b) proceed with a linear IV model in which the zero covariance assumption is loosened in this way. Their results extend an earlier line of research from Leamer (1981); Klepper and Leamer (1984); Bekker, Kapteyn, and Wansbeek (1987); and Manski and Pepper (2000). Nevo and Rosen (2012b) document that replacing the demanding zero covariance assumption with an assumption regarding the sign of the covariance between an IV and the stochastic error leads to convenient and easily estimable bounds in the linear IV model.

To define these bounds, we follow Nevo and Rosen (2012b) in using $\rho_{x\varepsilon}$ to signify correlation and $\sigma_{x\varepsilon}$ to signify covariance and σ_x to signify standard deviation, where subscripts make clear the random variables considered. The traditional IV validity assumption is thus denoted $\rho_{z\varepsilon} = 0$. Nevo and Rosen (2012b) replace this validity assumption with an assumption regarding only the "direction" of correlation between an instrument Z and the stochastic error term ε in (1):

$$\rho_{x\varepsilon}\rho_{z\varepsilon} \geq 0 \tag{6}$$

This assumption (Nevo and Rosen's [2012b] "assumption 3"[11]) thus states that the instrument has (weakly) the same direction of correlation with the omitted error term as the endogenous variable X.

This assumption, combined with a fourth assumption, gives the definition of an "imperfect instrumental variable" as an IV that has the same direction of correlation with the unobserved error term as the endogenous variable of interest x but is less endogenous than x:

$$|\rho_{x\varepsilon}| \geq |\rho_{z\varepsilon}| \tag{7}$$

Based on (7), we can define a quantity denoting the relative degree of correlation between the instrument and the error term compared with the same correlation between the original endogenous variable and the stochastic error term. This quantity, which captures how much less flawed the instrument is than the endogenous variable, $\lambda^* = \rho_{z\varepsilon}/\rho_{x\varepsilon}$, is not known without further assumptions; however, it is clearly bounded between 0, in

11. Nevo and Rosen (2012b) make a series of standard assumptions regarding the sampling process and any exogenous covariates included in the model, as assumptions 1 and 2.

the case that the traditional IV assumption holds and 1 in the case where (7) holds with equality.

Ignoring for now that λ^* is unknown, if it were known, a new valid compound instrument could be constructed by forming $\sigma_X \mathbf{Z} - \lambda^* \sigma_Z \mathbf{X}$. The logic behind this instrument is that the endogenous components of the original endogenous variable $\mathbf{X}$ and the (less) endogenous $\mathbf{Z}$ can be canceled out, and hence,

$$E\{(\sigma_X \mathbf{Z} - \lambda^* \sigma_Z \mathbf{X})\boldsymbol{\varepsilon}\} = \sigma_X \sigma_{Z\varepsilon} - \lambda^* \sigma_Z \sigma_{X\varepsilon} = 0$$

is a valid instrument. Nevo and Rosen's (2012b) proposal is to replace this above the valid instrument, denoted $V(\lambda^*) = (\sigma_X \mathbf{Z} - \lambda^* \sigma_Z \mathbf{X})$, with $V(1) = (\sigma_X \mathbf{Z} - 1\sigma_Z \mathbf{X})$, the instrument in the limit case implied by (7). While this will not give point estimates on the parameter of interest $\boldsymbol{\beta}$, it will allow for the construction of bounds in certain circumstances discussed below.

Consider now the probability limits of three different estimators: $\boldsymbol{\beta}^{\text{OLS}}$, the original estimand of $\boldsymbol{\beta}$ using endogenous $\mathbf{X}$ in a standard linear regression; $\boldsymbol{\beta}_z^{\text{IV}}$, the 2SLS estimator using the imperfect IV; and $\boldsymbol{\beta}_{v(1)}^{\text{IV}}$, the 2SLS estimator of the transformed variable described above. Based on the above two assumptions in (6) and (7), these parameters are not guaranteed to bound the true parameter $\boldsymbol{\beta}$. However, if the instrument is negatively correlated with the endogenous variable, $\sigma_{xz} < 0$, this allows for the construction of upper and lower bounds on the true parameter $\boldsymbol{\beta}$. These bounds are described in panel A of table 1. The right-hand panel describes the case in which Nevo and Rosen's (2012b) assumption 4 is not maintained; hence, $\boldsymbol{\beta}_{v(1)}^{\text{IV}}$ is not used. In this case, the original $\boldsymbol{\beta}_z^{\text{IV}}$ parameter and the OLS estimate $\boldsymbol{\beta}^{\text{OLS}}$ bound $\boldsymbol{\beta}$, with the upper and lower bounds depending on the assumed correlation between $\mathbf{X}$ and $\boldsymbol{\varepsilon}$ (and hence $\mathbf{Z}$ and $\boldsymbol{\varepsilon}$).[12] However, if the correlation between $\mathbf{X}$ and $\mathbf{Z}$ is positive, only one-sided bounds can be formed. In the case that assumption 4 (7) is maintained, this leads to a further tightening of the bounds, given that the inconsistent $\boldsymbol{\beta}^{\text{OLS}}$ parameter can be replaced by the less inconsistent $\boldsymbol{\beta}_{v(1)}^{\text{IV}}$ parameter.[13] Once again, however, if the correlation between the endogenous variable and the instrument is not negative, informative bounds cannot be formed, leading to only one-sided bounds for $\boldsymbol{\beta}$. Both bounds with and without assumption 4 can be produced by the `imperfectiv` command described later in this article.

12. To see why the IV and OLS parameters bound the true parameter $\boldsymbol{\beta}$, note that in the simple linear model described in (1) and (2), we can write $\boldsymbol{\beta}^{\text{OLS}} = \boldsymbol{\beta} + (\sigma_{x\varepsilon}/\sigma_x^2)$ and $\boldsymbol{\beta}_z^{\text{IV}} = \boldsymbol{\beta} + (\sigma_{z\varepsilon}/\sigma_{xz})$. Given that σ_{xz} is assumed negative (a testable assumption) and σ_x^2 is positive, these two parameters bound $\boldsymbol{\beta}$.
13. To see why $\boldsymbol{\beta}_z^{\text{IV}}$ and $\boldsymbol{\beta}_{v(1)}^{\text{IV}}$ bound the true parameter, we can start from $\boldsymbol{\beta}_z^{\text{IV}}$ and $\boldsymbol{\beta}^{\text{OLS}}$, which we know provide bounds. Given that $\boldsymbol{\beta}_{v(1)}^{\text{IV}}$ is a weighted average of $\boldsymbol{\beta}_z^{\text{IV}}$ and $\boldsymbol{\beta}^{\text{OLS}}$ assuming $\lambda = 1$ (see Nevo and Rosen [2012b] for full details), this estimate will remove part of the bias from the $\boldsymbol{\beta}^{\text{OLS}}$ parameter, moving estimates toward the $\boldsymbol{\beta}^{\text{IV}}$ parameter. However, given that z is less endogenous than x, the contribution of z to the compound instrument $\boldsymbol{\beta}_{v(1)}^{\text{IV}}$ will never be sufficient to completely reverse the direction of the bias of the original $\boldsymbol{\beta}^{\text{OLS}}$ estimate, and so $\boldsymbol{\beta}_z^{\text{IV}}$ and $\boldsymbol{\beta}_{v(1)}^{\text{IV}}$ still provide (potentially tighter) two-sided bounds.

In the discussion up to this point, we have justified the relaxation of the instrumental-validity assumption when one imperfect IV is present. However, Nevo and Rosen (2012b) demonstrate that if more than one imperfect IV is available, this result can be used to potentially generate tighter bounds[14] and, under an auxiliary assumption, produce two-sided bounds where previously only one-sided bounds were observed. In the simplest case, without further restrictions on the nature of each imperfect IV (beyond the fact that they each meet assumptions 3 and 4), the bounding procedure consists of a search among all imperfect IVs and the OLS estimate to generate the tightest set of bounds possible given the assumptions maintained in assumptions (6) and (7). This can be seen as a generalization of panel A of table 1, where each $\boldsymbol{\beta}^{\text{IV}}$ parameter is replaced with its min (for upper bounds) or max (for lower bounds). In the case that various candidates exist for upper or lower bounds, inference in the Nevo and Rosen procedure must account for uncertainty in various coefficients. As laid out in Nevo and Rosen (2012b, 665–666), this is based on a variant of Chernozhukov, Lee, and Rosen's (2013) intersection bounds. This inference procedure is performed by default in the `imperfectiv` command when multiple similar bound candidates exist.

Finally, Nevo and Rosen (2012b) show that if more than one instrument is available and if one instrument is assumed to be better than another in both relevance and validity, then two-sided bounds can be produced, even if the original imperfect IVs are positively correlated with the endogenous variable X. Consider two imperfect IVs, Z_1 and Z_2, where assumption 6 is assumed to hold, $\sigma_{xz_1} > \sigma_{xz_2}$ (Z_1 is more relevant than Z_2), and it is assumed that $\sigma_{\varepsilon z_1} < \sigma_{\varepsilon z_2}$ (Z_1 is less endogenous than Z_2). Then, the production of a new instrument,

$$\omega(\gamma) = \gamma Z_2 - (1-\gamma) Z_1$$

will lead to two-sided bounds so long as $\sigma_{\omega(\gamma)\varepsilon} \geq 0$ and $\sigma_{\omega(\gamma)x} < 0$. These bounds are described in panel B of table 1 and are summarized as Nevo and Rosen's proposition 5. In practice, Nevo and Rosen (2012b) suggest using a value of $\gamma = 0.5$ to form the reweighted imperfect IV. In the `imperfectiv` command, $\gamma = 0.5$ is used by default, and a "better" and "worse" imperfect IV must be indicated by the user to produce bounds in this case.

14. Recent work from Wiseman and Sørensen (2017) suggests under an alternative (implicit) assumption that Nevo and Rosen's bounds can, in some cases, be further tightened, especially when instruments are weak.

Table 1. Summary of the propositions of Nevo and Rosen (2012b)

	Assumption 4		No assumption 4	
	$\sigma_{xz} < 0$	$\sigma_{xz} > 0$	$\sigma_{xz} < 0$	$\sigma_{xz} > 0$
Panel A: One instrument				
$\rho_{x\varepsilon} > 0$	$\beta_z^{\mathrm{IV}} \le \beta \le \beta_{v(1)}^{\mathrm{IV}}$	$\beta \le \min\{\beta_{v(1)}^{\mathrm{IV}}, \beta_z^{\mathrm{IV}}\}$	$\beta_z^{\mathrm{IV}} \le \beta \le \beta^{\mathrm{OLS}}$	$\beta \le \min\{\beta^{\mathrm{OLS}}, \beta_z^{\mathrm{IV}}\}$
$\rho_{x\varepsilon} < 0$	$\beta_{v(1)}^{\mathrm{IV}} \le \beta \le \beta_z^{\mathrm{IV}}$	$\beta \ge \max\{\beta_{v(1)}^{\mathrm{IV}}, \beta_z^{\mathrm{IV}}\}$	$\beta^{\mathrm{OLS}} \le \beta \le \beta_z^{\mathrm{IV}}$	$\beta \ge \max\{\beta^{\mathrm{OLS}}, \beta_z^{\mathrm{IV}}\}$
Panel B: Multiple instruments with "proposition 5"				
$\rho_{x\varepsilon} > 0$	-	$\beta_\omega^{\mathrm{IV}} \le \beta \le \min\{\beta_{z_1}^{\mathrm{IV}}, \beta_{z_2}^{\mathrm{IV}}, \beta^{\mathrm{OLS}}\}$	-	$\beta_\omega^{\mathrm{IV}} \le \beta \le \min\{\beta_{z_1}^{\mathrm{IV}}, \beta_{z_2}^{\mathrm{IV}}, \beta_{v_1(1)}^{\mathrm{IV}}, \beta_{v_2(1)}^{\mathrm{IV}}, \beta_{v*(1)}^{\mathrm{IV}}\}$
$\rho_{x\varepsilon} < 0$	-	$\beta_\omega^{\mathrm{IV}} \ge \beta \ge \min\{\beta_{z_1}^{\mathrm{IV}}, \beta_{z_2}^{\mathrm{IV}}, \beta^{\mathrm{OLS}}\}$	-	$\beta_\omega^{\mathrm{IV}} \ge \beta \ge \min\{\beta_{z_1}^{\mathrm{IV}}, \beta_{z_2}^{\mathrm{IV}}, \beta_{v_1(1)}^{\mathrm{IV}}, \beta_{v_2(1)}^{\mathrm{IV}}, \beta_{v*(1)}^{\mathrm{IV}}\}$

NOTES: Full details of the bounding procedure and assumptions are available in Nevo and Rosen (2012b). Notation is defined in section 2 of this article. The final case in panel B for $\rho_{x\varepsilon} < 0$ is not shown in Nevo and Rosen (2012b); however, the $\rho_{x\varepsilon} > 0$ case is shown to hold without loss of generality in Nevo and Rosen (2008), giving the reverse case also shown here. In each case, the statement $\rho_{x\varepsilon} > 0$ implies $\rho_{z\varepsilon} > 0$, and $\rho_{x\varepsilon} < 0$ implies $\rho_{z\varepsilon} < 0$; that is, assumption 6 is always assumed to hold.

3 Stata commands

Below we describe the basic syntax of two commands that implement the estimators described in the previous section. These two commands are `plausexog`, which implements Conley, Hansen, and Rossi's (2012) bounds relaxing the exclusion restriction, and `imperfectiv`, which implements Nevo and Rosen's (2012b) bounding procedure by relaxing the traditional validity assumption. We examine the commands in turn in sections 3.1 and 3.2. We also provide extended examples of their syntax and their use by replicating empirical examples from Nevo and Rosen (2012b) and Conley, Hansen, and Rossi (2012) in appendix A. In both cases, the syntax is presented following a linear IV model, and constant coefficients are assumed.[15]

3.1 The plausexog command

Syntax

`plausexog` is closely related to Stata's IV regression command, `ivregress`, with arguments describing the prior expectation of the degree of the violation of the exclusion restriction. The generic syntax of the command is as follows:

`plausexog` *method depvar* [*varlist1*] `(`*varlist2* `=` *varlist_iv*`)` [*if*] [*in*] [*weight*]
[`, level(`*#*`) vce(`*vcetype*`) gmin(`*numlist*`) gmax(`*numlist*`) grid(`*#*`)`
`mu(`*numlist*`) omega(`*numlist*`) distribution(`*name*`,` *params*`) seed(`*#*`)`
`iterations(`*#*`) graph(`*varname*`) graphmu(`*numlist*`) graphomega(`*numlist*`)`
`graphdelta(`*numlist*`)` *graph_options*]

method must be specified as either `uci` (union of confidence intervals) or `ltz` (local to zero), depending on the desired estimator. The remainder of the syntax follows Stata's `ivregress` syntax, where first any exogenous variables are specified as *varlist1*, then the endogenous variables are specified as *varlist2*, and finally "plausibly exogenous" instruments are specified as *varlist_iv*.

Options

`level(`*#*`)` sets the confidence level. The default is `level(0.95)`.

`vce(`*vcetype*`)` determines the type of standard error reported in the estimated regression model and allows standard errors that are robust to certain types of misspecification. *vcetype* may be `robust`, `cluster` *clustvar*, `bootstrap`, `jackknife`, `unadjusted`, or `hac` *kernel*.

15. Nevertheless, in the case of `plausexog`, estimated parameter bounds can also be interpreted as the bounds on the average treatment effect assuming heterogeneous treatment effects. Discussion of this is provided in Conley, Hansen, and Rossi (2012, 261).

`gmin(`*numlist*`)` specifies the minimum values for γ on plausibly exogenous variables (specified with `uci` only). One `gmin()` value must be specified for each plausibly exogenous variable, and these values likely vary for each plausibly exogenous IV.

`gmax(`*numlist*`)` specifies the maximum values for γ on plausibly exogenous variables (specified with `uci` only). One `gmax()` value must be specified for each plausibly exogenous variable, and these values likely vary for each plausibly exogenous IV.

`grid(`*#*`)` specifies the number of points (in [`gmin()`, `gmax()`]) at which to calculate bounds (specified with `uci` only). The default is `grid(2)`.

`mu(`*numlist*`)` specifies the mean value for the prior distribution of γ, assuming a Gaussian prior and the `ltz` method. One `mu()` value must be specified for each plausibly exogenous variable, and these values likely vary for each plausibly exogenous IV.

`omega(`*numlist*`)` specifies the variance value for the prior distribution of γ, assuming a Gaussian prior and the `ltz` method. One `omega()` value must be specified for each plausibly exogenous variable, and these values likely vary for each plausibly exogenous IV.

`distribution(`*name*`,` *params*`)` allows for non-Gaussian priors for the distribution of gamma. When one uses the `distribution()` option, the `mu()` and `omega()` options do not need to be specified. Bounds based on nonnormal distributions for gamma are calculated using the simulation-based algorithm described in Conley, Hansen, and Rossi (2012, 265) and section 2. *name* may be `normal`, `uniform`, `chi2`, `poisson`, `t`, `gamma`, and `special`. When one specifies any of the first six *name*s, *params* must be specified along with each of these distributions. For `normal`, parameters are the assumed mean and standard deviation; for `uniform`, the parameters are the minimum and maximum; for `chi2` (χ^2), it is the degrees of freedom; for `poisson`, it is the distribution mean; for `t`, it is the degrees of freedom; and for `gamma`, it is the shape and scale of the assumed distribution. For any assumed distribution of gamma that is not contained in the previous list, `special` can be specified, and a variable can be passed containing analytical draws from this distribution. If more than one plausibly exogenous variable is used, the relevant parameters must be specified for each plausibly exogenous variable. Note that although a Gaussian prior is allowed in this format, if a Gaussian prior is assumed, it is preferable to use the `mu()` and `omega()` options because these give an exact, rather than approximate (simulated), set of bounds.

`seed(`*#*`)` sets the seed for simulation-based calculations when using a non-Gaussian prior with `ltz`. This option is required when the `distribution()` option is specified.

`iterations(`*#*`)` determines the number of iterations for simulation-based calculations when using a non-Gaussian prior with `ltz`. The default is `iterations(5000)`. In Stata/IC and Small Stata, the number of iterations cannot exceed the maximum matrix size permitted by Stata. As such, these are set to 800 and 100, respectively. The `distribution()` option should be used with care in these flavors of Stata.

`graph(`*varname*`)` indicates that a graph should be produced of bounds over a range of assumptions related to the failure of the exclusion restriction. *varname* indicates the name of the endogenous variable (from *varlist2*) that the user wishes to graph. In the `uci` method, confidence intervals will be graphed, whereas in the `ltz` method, both confidence intervals and a point estimate will be graphed over a range of gamma values.

`graphmu(`*numlist*`)` (specified with `ltz` when a graph is desired) provides the values for a series of `mu()` values for each point desired on the graph. Each point refers to the mean value of γ assuming a Gaussian prior.

`graphomega(`*numlist*`)` (specified with `ltz` when a graph is desired) provides the values for a series of `omega()` values for each point desired on the graph. Each point corresponds to the value in the `graphmu()` list and specifies the variance of the Gaussian prior at each point.

`graphdelta(`*numlist*`)` allows for the plotting of values on the graph. If not specified, the values in `graphmu()` will be plotted on the horizontal axis.

graph_options are any other options documented in [G-3] ***twoway_options***. This overrides default graph options such as title and axis labels.

Stored results

`plausexog` stores the following in `e()`:

Scalars

`e(lb_`*endogname*`)`	lower-bound estimate for each (plausibly) instrumented variable
`e(ub_`*endogname*`)`	upper-bound estimate for each (plausibly) instrumented variable

In the case where Conley, Hansen, and Rossi's (2012) `ltz` method is used with an assumption of normality, the following are also returned:

Matrices

`e(b)`	coefficient vector under plausible exogeneity
`e(V)`	variance–covariance matrix of the estimators under plausible exogeneity

These are the coefficient vector and variance–covariance matrix of the estimated parameters based on the plausibly exogenous model.

3.2 The imperfectiv command

Syntax

The `imperfectiv` command is also closely related to Stata's IV regression command, with arguments describing correlation between the endogenous variable and the unobservable to replace the validity assumption of the instruments. The generic syntax of the command is as follows:

`imperfectiv` *depvar* [*varlist1*] `(`*varlist2* `=` *varlist_iv*`)` [*if*] [*in*] [*weight*] [`,`
`level(`#`)` `vce(`*vcetype*`)` `ncorr` `prop5` `noassumption4` `exogvars(`*varlist*`)`
`bootstraps(`#`)` `seed(`#`)` `verbose`]

The syntax follows Stata's `ivregress` syntax, where first any exogenous variables are specified as *varlist1*, then the endogenous variables are specified as *varlist2*, and finally "imperfect" instruments are specified as *varlist_iv*.

Options

`level(`#`)` sets the confidence level. The default is `level(0.95)`.

`vce(`*vcetype*`)` determines the type of standard error reported in the estimated regression model and allows standard errors that are robust to certain types of misspecification. *vcetype* may be `robust`, `cluster` *clustvar*, `bootstrap`, or `jackknife`.

`ncorr` specifies that the correlation between the endogenous variable and the unobservable error is assumed negative. By default, this correlation is assumed to be positive.

`prop5` specifies that proposition 5 of Nevo and Rosen (2012b) should be used in the estimation of bounds. If the correlation between the endogenous variable and each imperfect instrument is positive, the result of the estimation is an interval with only one bound. If there is more than one imperfect instrument, then proposition 5 of Nevo and Rosen can be used to generate two-sided bounds. If `prop5` is specified, the first two instruments specified in `varlist_iv` are used, and it is assumed that the "better" instrument is listed first. Additional discussion is provided in section 2.

`noassumption4` specifies that assumption 4 of Nevo and Rosen (2012b) does not hold. By default, this assumption is assumed to hold. Assumption 4 states that the correlation between the imperfect instrument and the unobservable is less than the correlation between the endogenous variable and the unobservable.

`exogvars(`*varlist*`)` specifies to display bounds on exogenous variables included in *varlist1* (if present). By default, only bounds on the endogenous variable of interest in the model are presented. Any variable indicated in `exogvars()` must be included in the original model in *varlist1*.

`bootstraps(#)` specifies the number of bootstrap replications. Wherever possible, a larger number of bootstraps should be specified. In the case when multiple candidates exist for upper or lower bounds, inference procedures consider uncertainty in each estimate that is close to binding using a bootstrap procedure (refer to Nevo and Rosen [2012b, 666] for full details).

`seed(#)` allows for the seed to be set to permit replicability of the bootstrap procedure. This option is relevant only when multiple candidates for upper or lower bounds exist.

`verbose` causes additional output be produced during the running of the command. This option is relevant when large datasets are used and multiple bounds are considered, because the bootstrap procedure may take some time to complete.

Stored results

`imperfectiv` stores the following in `e()`:

Scalars

`e(lb_`*endogname*`)`	lower-bound point estimate for the endogenous instrumented variable
`e(ub_`*endogname*`)`	upper-bound point estimate for the endogenous instrumented variable
`e(CIlb_`*endogname*`)`	lower-bound confidence interval for the endogenous variable
`e(CIub_`*endogname*`)`	upper-bound confidence interval for the endogenous variable

In this case, these values refer to point estimates identifying bound endpoints. The confidence intervals associated with these estimates (and hence the bounds) are returned as `e(CIlb_`*endogname*`)` and `e(CIub_`*endogname*`)`.

Matrices

`e(LRbounds)`	upper and lower bounds on each endogenous variable and all exogenous covariates

`e(LRbounds)` returns both the point estimates at each end of the bounds, as well as the confidence interval on these estimates.

4 Performance under simulation

We demonstrate the usage of the `imperfectiv` and `plausexog` commands under a series of simulations. These simulations allow us to examine the behavior of bounds on the (known) endogenous parameter of interest under a series of different assumptions. In particular, we can compare the behavior of bounds using the `uci` and `ltz` methods of Conley, Hansen, and Rossi (2012) and with and without the use of Nevo and Rosen's (2012b) assumption 4.

We aim to examine performance of bounds under a wide range of situations. To do so, we consider a linear model in which we allow the correlation between an endogenous variable of interest $\mathbf{x}$ and the unobserved error term $\boldsymbol{\varepsilon}$ to vary (that is, varying the degree of endogeneity of the parameter of interest) and in which the correlation between the "instrument" **z** and the unobserved compound error term varies (varying the quality of the instrument). In particular, we allow for this in the following two-stage setup:

$$\begin{pmatrix} z \\ \varepsilon \\ \nu \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right]$$

$$\mathbf{x} = \pi\mathbf{z} + \mu\varepsilon + \boldsymbol{\nu} \tag{8}$$

$$\mathbf{y} = \beta\mathbf{x} + \gamma\mathbf{z} + \boldsymbol{\varepsilon}$$

Here **y** is a dependent variable, **x** an endogenous variable of interest, and **z** an imperfect, or plausibly exogenous, instrumental variable. In all simulations presented here, we consider the case where one instrument exists; however, we provide an illustration with multiple instruments in appendix B. Provided that $\mu \neq 0$, β cannot be estimated consistently with an OLS regression, and provided that $\gamma \neq 0$, IV estimates of β will not be consistent under standard assumptions. The instrument **z** and error terms $\boldsymbol{\varepsilon}$ and $\boldsymbol{\nu}$ are simulated from independent normal distributions. In traditional 2SLS, γ is assumed to be 0; thus, $\gamma\mathbf{z}$ is omitted from the final equation. This leads to a compound error term $(\gamma\mathbf{z} + \boldsymbol{\varepsilon})$, which we refer to as $\boldsymbol{\eta}$ below.

Using this structure, we examine the use of and performance of `imperfectiv` and `plausexog` by varying γ (the degree of instrumental invalidity) and μ (the degree of endogeneity). We fix π at -0.6 in all simulations, ensuring that the instrument is not weak. The performance of `plausexog` following this data-generating process (DGP) is documented below:

```
. set obs 1000
number of observations (_N) was 0, now 1,000
. set seed 1716
. foreach var in u z v w {
  2.     generate `var´ = rnormal()
  3. }
. generate x  = -0.6*z + 0.33*u + v
. generate y1 =  3.0*x + 0.10*z + u
. plausexog uci y1 (x=z), gmin(0) gmax(0.2)
Estimating Conely et al.´s uci method
Exogenous variables:
Endogenous variables: x
Instruments: z
Conley et al (2012)´s UCI results                    Number of obs =      1000

Variable    Lower Bound     Upper Bound

x           2.7463585       3.2788045
_cons       -.05578396      .07826073
```

```
. plausexog ltz y1 (x=z), mu(0.1) omega(0.01)
Estimating Conely et al.´s ltz method
Exogenous variables:
Endogenous variables: x
Instruments: z

Conley et al. (2012)´s LTZ results                  Number of obs =     1000

                   Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]

          x     3.017955    .1722409   17.52   0.000     2.680369    3.355541
      _cons     .0079718    .0339273    0.23   0.814    -.0585244    .0744681
```

Above we document the use of `plausexog` with the `uci` and `ltz` methods. In each case, we "correctly" specify the prior over the violation of the exclusion restriction. In the `uci` example, the exclusion restriction is allowed to have support $\in [0, 0.2]$, with the true value simulated being 0.1. In the `ltz` example, the exclusion restriction is specified to fall within a normal distribution mean of 0.1 and variance of 0.01. In each case, bounds on the endogenous variable `x` contain the true parameter $\beta = 3$.

Below we document the use of `imperfectiv` using the same DGP. We first specify that bounds be calculated without assuming that the instrument is "less endogenous" than the endogenous variable; in the second case, we add this assumption:

```
. imperfectiv y1 (x=z), noassumption4

Nevo and Rosen (2012)´s Imperfect IV bounds           Number of obs =     1000

Variable         Lower Bound(CI)  LB(Estimator) UB(Estimator) Upper Bound(CI)

x                  [2.7463585      (2.8548516    3.1942966)     3.243692]

. imperfectiv y1 (x=z)

Nevo and Rosen (2012)´s Imperfect IV bounds           Number of obs =     1000

Variable         Lower Bound(CI)  LB(Estimator) UB(Estimator) Upper Bound(CI)

x                  [2.7463585      (2.8548516    3.0819896)     3.1396404]
```

These examples document performance of `plausexog` and `imperfectiv` under one particular DGP. In table 2 below, we consider a range of DGPs where we vary γ (within each panel) and μ (across each panel). Here γ refers to the failure of the exclusion restriction with which Conley, Hansen, and Rossi (2012) are concerned, and the resulting correlations between $\mathbf{x}$ and $\boldsymbol{\eta}$ (the compound error term) and $\mathbf{z}$ and $\boldsymbol{\eta}$ with which Nevo and Rosen (2012b) are concerned are displayed in subsequent columns. Bounds are then documented under two cases in Conley, Hansen, and Rossi (2012) (the `uci` and `ltz` methods, each with correctly specified priors) and two cases in Nevo and Rosen (2012b) (with and without assumption 4). In the case of Nevo and Rosen (2012b), the assumptions for "No A4" will be met provided that the sign on $\rho_{x,\eta}$ and $\rho_{z,\eta}$ is the same and will be met for "Assumption 4" only if $\rho_{x,\eta} \geq \rho_{z,\eta}$.

Table 2. Performance of various bounds under Monte Carlo simulation

γ	$\rho_{x,\eta}$	$\rho_{z,\eta}$	Plausibly exogenous: UCI	Plausibly exogenous: LTZ $\mathcal{N}(\mu,\sigma^2)$	Imperfect IV: No A4	Imperfect IV: Assumption 4
Panel A: Minor correlation between $\mathbf{x}$ and $\boldsymbol{\varepsilon}$						
0.1	0.27	0.10	[2.744 3.247]	[2.684 3.318]	[2.744 3.268]	[2.744 3.150]
0.2	0.22	0.19	[2.582 3.397]	[2.563 3.439]	[2.582 3.229]	[2.582 3.074]
0.3	0.16	0.28	[2.419 3.550]	[2.468 3.534]	[2.419 3.190]	[2.419 2.997]
0.4	0.11	0.36	[2.256 3.706]	[2.388 3.614]	[2.256 3.152]	[2.256 2.922]
Panel B: Moderate correlation between $\mathbf{x}$ and $\boldsymbol{\varepsilon}$						
0.1	0.61	0.10	[2.735 3.238]	[2.681 3.321]	[2.735 3.429]	[2.735 3.277]
0.2	0.56	0.19	[2.565 3.380]	[2.558 3.443]	[2.565 3.405]	[2.565 3.217]
0.3	0.51	0.28	[2.394 3.528]	[2.462 3.540]	[2.394 3.381]	[2.394 3.157]
0.4	0.46	0.36	[2.223 3.681]	[2.380 3.622]	[2.223 3.357]	[2.223 3.097]
Panel C: Major correlation between $\mathbf{x}$ and $\boldsymbol{\varepsilon}$						
0.1	0.91	0.10	[2.706 3.208]	[2.670 3.332]	[2.706 3.292]	[2.706 3.224]
0.2	0.88	0.19	[2.508 3.336]	[2.538 3.463]	[2.508 3.288]	[2.508 3.196]
0.3	0.84	0.28	[2.309 3.515]	[2.433 3.569]	[2.309 3.283]	[2.309 3.169]
0.4	0.80	0.36	[2.110 3.710]	[2.341 3.661]	[2.110 3.279]	[2.110 3.142]

We display 95% confidence intervals associated with the parameter β in square parentheses. The true value of β is 3 in the DGP described in (8). The value of γ in each case is displayed in the left-hand column (between 0.1 and 0.4), and the correlation between $\mathbf{x}$ and $\boldsymbol{\eta}$ and $\mathbf{z}$ and $\boldsymbol{\eta}$ inferred in each case is listed in subsequent columns. Here $\boldsymbol{\eta}$ refers to the compound error term that causes endogeneity and instrumental invalidity. We use 1,000 simulated observations. Different panels allow the correlation between the endogenous variable $\mathbf{x}$ and the $\boldsymbol{\varepsilon}$ term to vary, making $\mathbf{x}$ more or less endogenous. Confidence intervals for the plausibly exogenous UCI case are based on a support assumption implying that the true value of γ is at the mean and hence is $[0, 2\gamma]$. In the LTZ case, the distribution for γ is assumed to be normal, with mean equal to the value of gamma, and variance equal to $\gamma/10$. Confidence intervals for imperfect IV estimates are based on assumptions that $\rho_{x,\eta} > 0$ and $\rho_{z,\eta} > 0$ in the "No A4" case and that $\rho_{x,\eta} \geq \rho_{z,\eta} > 0$ in the "Assumption 4" case. The veracity of each assumption can be determined from displayed correlations in columns 2 and 3.

The bounds produced in each case on the endogenous variable of interest are presented in table 2. In nearly all simulations, the bounds include the true value of $\beta = 3$. The only cases in which this is not seen is with those in the right-most columns at the bottom of panel A. This is to be expected, given that in this case, the assumptions underlying the bounds (assumption 4 of Nevo and Rosen [2012b]) are not met; hence, the `imperfectiv` command should correctly have been run with the `noassumption4` option. In each circumstance, the Conley, Hansen, and Rossi (2012) bounds contain the true parameter, but this is dependent on correctly specifying the prior over γ, as we ensure in table 2. Given that in practice, knowing the true prior for γ is an empirical challenge (see, for example, Bhalotra and Clarke [2016] as well as additional discussion in section 2 of this article), researchers may prefer conservative assumptions on γ.

In general, although the procedures of both Nevo and Rosen (2012b) and Conley, Hansen, and Rossi (2012) allow the strong assumptions relating to unobservables in an IV setting to be loosened, bounds estimates still rely on a willingness to specify something about the relationship between instruments and unobservables. Ideally, these assumptions should be well founded in a theory related to the nature of failure of IV validity. In the case of Nevo and Rosen (2012b), a willingness to assume that an instrument is positively or negatively related to unobservables may reflect some underlying model of selection into an instrument or of behavioral response to a particular draw of the instrumental variable. Consider briefly two well-known instruments in models of human fertility: the gender mix of children and the occurrence of twin births. In the case of gender mix of births, Dahl and Moretti (2008) document a "demand for sons", suggesting that investments following sons may depend positively on this particular realization of the IV. In the case of twins, Bhalotra and Clarke (2016) document a cross-cutting positive selection of twin births, where many (positive) maternal health behaviors in utero increase the likelihood of giving live births to twins (even if twin conception is random). Here assumptions relating to a positive correlation between the instrument and unobservables seems reasonable based on positive correlations between the instrument and many hard-to-measure and frequently unobserved variables.[16] As discussed above, the willingness to assume a particular range or distribution for the failure of the exclusion restriction is also an empirical challenge. While in the case of Conley, Hansen, and Rossi (2012), bounds are constructed based on stronger assumptions than just the sign of the correlation, a benefit of this approach is that it allows for the sign to be indeterminate, for example, if one is concerned that instruments may only be "close" to exogenous but not certain of the direction in which failures of validity occur. We return to these considerations below.

Abstracting now from why identifying assumptions may be met, we see that table 2 offers several lessons regarding the relative performance of Conley, Hansen, and Rossi and Nevo and Rosen bounds. First, the bounds on the endogenous parameter using Conley, Hansen, and Rossi's (2012) `plausexog` procedure are approximately constant across panels (given a particular value for γ) because the degree of endogeneity of $\mathbf{x}$ does not impact the estimated bounds. In the case of Nevo and Rosen (2012b), all else remaining constant, bounds are more or less wide when the independent variable of interest is more or less exogenous. This is because Nevo and Rosen (2012b) use information from the original endogenous variable to form one side of their bounds (when two-sided bounds are formed). In the limit case when assumption 4 is not assumed, the bound on the OLS estimate of β itself is used. When both cases are examined using the methods of Nevo and Rosen (2012b), the lower bound consists of the original IV estimate, which agrees with the lower bound determined by the Conley, Hansen, and Rossi (2012) UCI approach. This is not always the case in Conley, Hansen, and Rossi's methods and only occurs when the lower limit of γ is fixed at 0. Then, the IV would be valid, and the lower bound becomes the unaltered IV estimate.

16. More generally, often intuitively, the likely direction of correlation between an observed variable and an unobserved error term is assumed in empirical applications. For example, in simple linear models, the well-known omitted variable bias in OLS can be signed if the correlation between an included variable and the unobservable error is assumed. In Nevo and Rosen's imperfect IV application, we are concerned with similar correlations between IVs and unobserved errors. Whether a reasonable assumption regarding the potential correlation between an instrument and the error term exists depends entirely on the phenomenon under study.

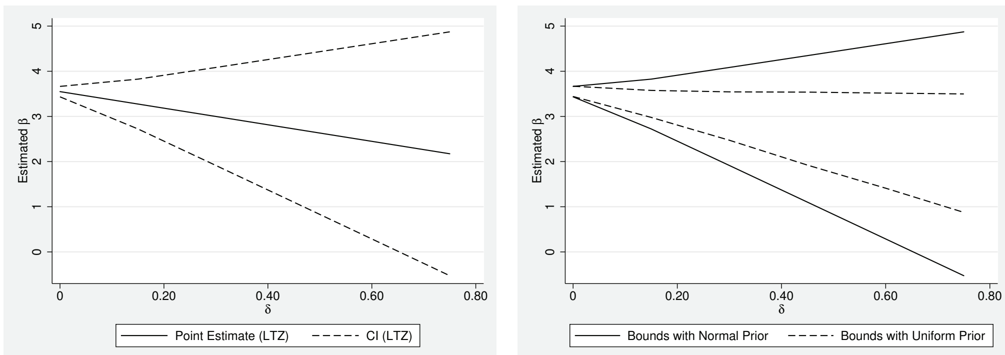
Second, we note that bounds from Nevo and Rosen (2012b) are always tighter when assumption 4 is used (in the case shown in table 2, the upper bound on β always falls). Of course, this is not free but rather a direct result of the assumption that $\mathbf{z}$ is less endogenous than $\mathbf{x}$. In the case that this is true, bounds are both tighter and contain the true parameter. When assumption 4 is not met, bounds are tighter but do not contain the true parameter.

Finally, we note that in this case, adding additional structure in the Conley, Hansen, and Rossi (2012) bounding procedure with the LTZ approach actually results in wider bounds in some cases. This is a direct result of the parameters assumed in each case. In the UCI case, we allow for a support of $[0, 2 \times \gamma]$ for each implementation, while the LTZ case assumes that $\gamma \sim \mathcal{N}(\gamma, \gamma/10)$, which often results in a probability distribution for γ that has a considerable probability mass outside the values allowed in the UCI approach. This should not be seen as necessarily representative of the use of the UCI and LTZ approaches. Frequently, the LTZ approach leads to tighter bounds, given the additional structure placed on the prior for γ. Indeed, in the above simulations, if we were to use a Gaussian prior in the LTZ approach with an identical variance of a uniform spanning the UCI γ_{min} and γ_{max} values, bounds in the LTZ approach would be tighter than those in the UCI approach. This is a direct result of placing greater weight on values closer to the true value of γ when using the normal prior. Unlike the Nevo and Rosen (2012b) method, the Conley, Hansen, and Rossi (2012) method allows for a prior that the instrument may be positively related, negatively related, or unrelated with the unobserved error term. However, the additional flexibility of the Conley, Hansen, and Rossi (2012) method also comes with the caveat that rather than knowing the sign of the correlation between the instrument and the error term, we must assume something about the magnitude of the failure of the exclusion restriction.

While Nevo and Rosen (2012b) are based on two assumptions and no further priors are required (as documented in the two columns of table 2), Conley, Hansen, and Rossi (2012) bounds are based on parametric priors that can take an unlimited range of values. Thus, if a researcher uses Conley, Hansen, and Rossi (2012) bounds, it may be useful or illustrative to visualize bounds based on a range of values for a particular parametric prior.[17] This can be achieved using the graphing options of `plausexog`. We document an example of this code below, which produces figure 1a.[18]

17. A comprehensive example of this procedure is provided in the original Conley, Hansen, and Rossi (2012, 267) article. We show how to replicate a portion of their results using the `plausexog` command in appendix A.2.
18. All code in the article is made available on one of the author's website, currently at http://www.damianclarke.net/replication/.

```
. generate x  = 0.33*u + 0.6*z + v
. generate y3 = Beta*x + 0.3*z + u
. quietly plausexog ltz y3 (x=z), omega(0.01) mu(0.3) graph(x)
> graphomega(0 0.0225 0.09 0.2025 0.36 0.5625)
> graphmu(0 0.15 0.3 0.45 0.6 0.75)
> graphdelta(0 0.15 0.3 0.45 0.6 0.75) scheme(sj)
> ytitle(Estimated {&beta}) xtitle({&delta})
> xlabel(0 "0" 0.2 "0.20" 0.4 "0.40" 0.6 "0.60" 0.8 "0.80")
> legend(order(1 "Point Estimate (LTZ)" 2 "CI (LTZ)")) ylabel(0(1)5);
```



(a) Analytical Gaussian (b) Gaussian versus uniform

Figure 1. Plausibly exogenous bounds varying prior assumptions

Figure 1a assumes a Gaussian (normal) prior for γ in the LTZ approach of Conley, Hansen, and Rossi (2012) but varies the mean and variance. Bounds at each point on the graph are based on the assumption that $\gamma \sim \mathcal{N}(\delta, \delta^2)$. Figure 1b compares the bounds from the Gaussian prior with bounds based on a uniform prior that assumes $\gamma \sim U(0, 2 \times \delta)$. The true value for γ is 0.3, and the true value for β is 3. This allows for the comparison of the bounds estimator over a range of priors for γ. We observe (in figure 1a) that the true parameter is contained in the bounds only when the mean of the exclusion restriction is sufficiently high to approach the true value and that as in table 2, the bounds grow as the prior allows for additional probability mass on more extreme values of the violation of the exclusion restriction. In each case, classical IV imposing the exact assumption that $\gamma = 0$ would result in confidence intervals considerably above the true population parameter.

5 Conclusion

In this article, we discussed several issues involved in the estimation of bounds when examining a causal relationship in the presence of endogenous variables. These types of bounding procedures are likely to be particularly useful given the difficulties inherent in IV estimation and challenges in convincingly arguing for IV validity or the exclusion restriction in an IV model.

We introduced two procedures for estimating bounds in Stata: `imperfectiv` for Nevo and Rosen (2012b)'s "imperfect instrumental variable" procedure and `plausexog` for Conley, Hansen, and Rossi (2012)'s "plausible exogeneity". In documenting these procedures, we laid out numerous considerations when implementing each bounding process.

Nevo and Rosen (2012b) bounds are particularly appropriate when one is convinced of the direction of correlation of an IV with an unobserved error term but not necessarily its magnitude. The Conley, Hansen, and Rossi (2012) procedure, on the other hand, is well suited for situations in which the direction of correlation need not be known (but can be known) but in which the practitioner has some belief over the magnitude of the IV's importance in the system of interest. All else constant, Nevo and Rosen (2012b) bounds perform relatively better when the endogenous variable is less correlated with unobservables, while Conley, Hansen, and Rossi (2012) bounds perform equally well regardless of the correlation between the endogenous variable of interest and unobservables. Finally, while Conley, Hansen, and Rossi (2012) bounds are often based on more parametric or otherwise stronger assumptions related to the unobservable behavior of IVs, it is simple to undertake sensitivity testing of estimated bounds' stability to changes in these assumptions, and such sensitivity tests are encouraged when dealing with questionable IVs.

Given that these methodologies loosen IV assumptions in different ways and are well suited to different types of (classically invalid) IVs, we suggest that these methodologies should be seen as a complement, rather than a substitute, in the empirical researcher's toolbox. The ease of use of each methodology and their ability to recover parameters under a broad range of failures of IV assumptions suggest that these procedures should act as a go-to consistency test in the increasingly large number of cases where concerns exist regarding the veracity of IVs.

6 Acknowledgments

We are grateful to Christian Hansen, Adam Rosen, and Marc F. Bellemare for very useful comments on code and the exposition of this article. We also thank users of the code whose feedback has resulted in improvements and extensions in `plausexog`.

7 References

Angrist, J. D., and W. N. Evans. 1998. Children and their parents' labor supply: Evidence from exogenous variation in family size. *American Economic Review* 88: 450–477.

Angrist, J. D., and A. B. Krueger. 1991. Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics* 106: 979–1014.

Angrist, J. D., and J.-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Bekker, P., A. Kapteyn, and T. Wansbeek. 1987. Consistent sets of estimates for regressions with correlated or uncorrelated measurement errors in arbitrary subsets of all variables. *Econometrica* 55: 1223–1230.

Bhalotra, S. R., and D. Clarke. 2016. The twin instrument. IZA Discussion Paper No. 10405, Institute for the Study of Labor (IZA). http://ftp.iza.org/dp10405.pdf.

Bound, J., D. A. Jaeger, and R. M. Baker. 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogeneous explanatory variable is weak. *Journal of the American Statistical Association* 90: 443–450.

Buckles, K. S., and D. M. Hungerman. 2013. Season of birth and later outcomes: Old questions, new answers. *Review of Economics and Statistics* 95: 711–724.

Chernozhukov, V., S. Lee, and A. M. Rosen. 2013. Intersection bounds: Estimation and inference. *Econometrica* 81: 667–737.

Conley, T. G., C. B. Hansen, and P. E. Rossi. 2008. Plausibly exogenous. Working Paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=987057.

———. 2012. Plausibly exogenous. *Review of Economics and Statistics* 94: 260–272.

Dahl, G. B., and E. Moretti. 2008. The demand for sons. *Review of Economic Studies* 75: 1085–1120.

Hansen, L. P. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50: 1029–1054.

Hotz, V. J., C. H. Mullin, and S. G. Sanders. 1997. Bounding causal effects using data from a contaminated natural experiment: Analysing the effects of teenage childbearing. *Review of Economic Studies* 64: 575–603.

Kang, H., A. Zhang, T. T. Cai, and D. S. Small. 2016. Instrumental variables estimation with some invalid instruments and its application to Mendelian randomization. *Journal of the American Statistical Association* 111: 132–144.

Kitagawa, T. 2015. A test for instrument validity. *Econometrica* 83: 2043–2063.

Klepper, S., and E. E. Leamer. 1984. Consistent sets of estimates for regressions with errors in all variables. *Econometrica* 52: 163–184.

Kolesár, M., R. Chetty, J. Friedman, E. Glaeser, and G. W. Imbens. 2015. Identification and inference with many invalid instruments. *Journal of Business and Economic Statistics* 33: 474–484.

Leamer, E. E. 1981. Is it a demand curve, or is it a supply curve? Partial identification through inequality constraints. *Review of Economics and Statistics* 63: 319–327.

Manski, C. F., and J. V. Pepper. 2000. Monotone instrumental variables: With an application to the returns to schooling. *Econometrica* 68: 997–1010.

———. 2009. More on monotone instrumental variables. *Econometrics Journal* 12: S200–S216.

Nevo, A., and A. Rosen. 2008. Identification with imperfect instruments. Working Paper CWP16/08, Cemmap, Institute for Fiscal Studies. http://www.cemmap.ac.uk/wps/cwp1608.pdf.

———. 2012a. Replication data for: Identification with imperfect instruments. Harvard Dataverse. http://hdl.handle.net/1902.1/18721.

———. 2012b. Identification with imperfect instruments. *Review of Economics and Statistics* 94: 659–671.

Rosenzweig, M. R., and K. I. Wolpin. 1980a. Testing the quantity-quality fertility model: The use of twins as a natural experiment. *Econometrica* 48: 227–240.

———. 1980b. Life-cycle labor supply and fertility: Causal inferences from household models. *Journal of Political Economy* 88: 328–348.

———. 2000. Natural "natural experiments" in economics. *Journal of Economic Literature* 38: 827–874.

Rosenzweig, M. R., and J. Zhang. 2009. Do population control policies induce more human capital investment? Twins, birth weight and China's "one-child" policy. *Review of Economic Studies* 76: 1149–1174.

Rossi, P. 2017. *bayesm: Bayesian inference for marketing/micro-econometrics.* R package version 3.1-0.1. https://cran.r-project.org/web/packages/bayesm/index.html.

Rossi, P. E., T. G. Conley, and C. B. Hansen. 2012. Replication data for: Plausibly exogenous. Harvard Dataverse. http://hdl.handle.net/1902.1/18022.

Rubin, D. B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 66: 688–701.

Sargan, J. D. 1958. The estimation of economic relationships using instrumental variables. *Econometrica* 26: 393–415.

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

van Kippersluis, H., and N. Rietveld. 2017. Beyond plausibly exogenous. Tinbergen Institute Discussion Paper TI2017-096/V, Tinbergen Institute. https://papers.tinbergen.nl/17096.pdf.

Wiseman, N., and T. A. Sørensen. 2017. Bounds with imperfect instruments: Leveraging the implicit assumption of intransitivity in correlations. IZA Discussion Paper No. 10646, Institute of Labor Economics (IZA). http://ftp.iza.org/dp10646.pdf.

About the authors

Damian Clarke is an associate professor in the department of economics at the Universidad de Santiago de Chile and a research associate at the Centre for the Study of African Economies, Oxford.

Benjamín Matta is a final year master student in the Master in Economic Sciences at the Universidad de Santiago de Chile.

A Empirical examples using original data

We illustrate the performance of each of the `imperfectiv` and `plausexog` commands in Stata by replicating empirical examples from Nevo and Rosen (2012b) and Conley, Hansen, and Rossi (2012). These examples use data from the original articles[19] and the syntax of each command as explained in section 3.

A.1 Nevo and Rosen's (2012b) demand for cereal example

Below we replicate the bounds calculated by Nevo and Rosen (2012b) in their empirical application examining the demand for cereal. We use the `imperfectiv` command described above to calculate bounds. This syntax replicates the results in table 2 of Nevo and Rosen (2012b, 667) and in particular columns 3 and 4, where the imperfect IV methodology is used.

We first show the case where "assumption 4" is not imposed and output bounds on both the endogenous and each exogenous variable and then replicate the results assuming that "assumption 4" holds. In the second case, we display only the bounds on the endogenous variable of interest and one exogenous variable as presented in Nevo and Rosen (2012b), using the `exogvars()` option to simplify output. We note that in each case, the results displayed here (for the confidence intervals only) are slightly different from those reported in the article. Results displayed for the estimators themselves are identical. This difference owes to the simulation-based procedure followed for inference, described in Nevo and Rosen (2012b, 665–666).

```
. use nevorosen2012, clear
(Nevo and Rosen´s (2012) REStat cereal demand example)
. replace addv=addv/10
(986 real changes made)
```

19. Both of these datasets are available for public download from the Harvard Dataverse; refer to Rossi, Conley, and Hansen (2012) and Nevo and Rosen (2012a) for full details.

```
. local w addv bd1 bd2 bd3 bd4 bd5 bd6 bd7 bd8 bd9 bd10 bd11 bd12 bd13 bd14
> bd15 bd16 bd17 bd18 bd19 bd20 bd21 bd22 bd23 bd24 dd2 dd3 dd4 dd5 dd6 dd7 dd8
> dd9 dd10 dd11 dd12 dd13 dd14 dd15 dd16 dd17 dd18 dd19 dd20 sfdum

. generate qavgpo=p_bs

. replace qavgpo=p_sf if city==7
(495 real changes made)

. imperfectiv y `w´ (price=qavgp qavgpo), prop5 noassumption exogvars(`w´)
Nevo and Rosen (2012)´s Imperfect IV bounds          Number of obs =    990
```

Variable	Lower Bound(CI)	LB(Estimator)	UB(Estimator)	Upper Bound(CI)
price	[-11.374594	(-8.6880097	-4.0775182)	-2.0114159]
addv	[.16464915	(.27997955	.2984391)	.42634903]
bd1	[-.04658989	(.31879272	.92096513)	1.2107664]
bd2	[.3517848	(.52184143	.76148481)	.91496402]
bd3	[.29607864	(.52176292	.86312992)	1.057621]
bd4	[.09445771	(.36348807	.79478678)	1.0123517]
bd5	[.2194114	(.39085272	-1.2152841)	-.10164419]
bd6	[-.23430896	(-.0602687	.19185713)	.34736924]
bd7	[.07644934	(.19387363	.31510206)	.43441841]
bd8	[-.65082539	(-.52044026	-.36450968)	-.23924725]
bd9	[-.51739615	(-.40278766	-.28907719)	-.17017835]
bd10	[.79419792	(.95862659	1.1773915)	1.3330057]
bd11	[.33849583	(.53183548	.81935473)	.98935052]
bd12	[-.80965492	(-.55555294	-.14994102)	.05513481]
bd13	[.6066799	(.70922607	.76195243)	.87482299]
bd14	[-.12946468	(.00740146	.17926588)	.30958453]
bd15	[-.16953307	(-.07601362	-.06468287)	.04021644]
bd16	[-.17496545	(.08218297	-2.6424916)	-.69030523]
bd17	[-.64602115	(-.52156319	-.38020246)	-.25973429]
bd18	[-.19268084	(-.08270381	.01721605)	.13047166]
bd19	[-.3321358	(-.15811579	.09357173)	.24547956]
bd20	[.31778867	(.41761932	.29539792)	.48054201]
bd21	[-.86633982	(-.68524192	-.42802781)	-.27180957]
bd22	[-.00272639	(.09304845	-.07916141)	.11525372]
bd23	[-.7086475	(-.48185292	-.12731672)	.06207178]
bd24	[-.19987399	(.06647916	.49375976)	.70912936]
dd2	[-.10411673	(-.01868476	-.01359145)	.07957804]
dd3	[-.06364777	(.02241417	-.05978335)	.09186974]
dd4	[-.1838241	(-.09690161	-.28861194)	-.10026549]
dd5	[.10764943	(.1961398	-.07747668)	.15686027]
dd6	[.0822766	(.17019612	-.11133631)	.12466767]
dd7	[.1548558	(.24748077	-.18745247)	.14113948]
dd8	[.07232434	(.16762425	-.33668355)	.04177717]
dd9	[.04546349	(.14267331	-.40659777)	-.00498925]
dd10	[.02818531	(.124874	-.40942752)	-.02356694]
dd11	[-.01469474	(.08069586	-.42489117)	-.05381292]
dd12	[-.12544516	(-.02937618	-.55191795)	-.17312484]
dd13	[-.02303931	(.07812767	-.55719801)	-.10269805]
dd14	[.04028046	(.14216755	-.50754673)	-.04849103]
dd15	[.03570445	(.13436028	-.44665285)	-.0222374]
dd16	[-.08954305	(.00839237	-.55593204)	-.14027698]
dd17	[.04099528	(.13694727	-.38298065)	.00307687]
dd18	[.05112385	(.14315789	-.27322378)	.03860742]
dd19	[.05634496	(.15135756	-.34413822)	.02660973]
dd20	[-.06413548	(.03281487	-.51054805)	-.10844663]
sfdum	[-.20866809	(-.13629732	-.90239378)	-.35730793]

```
. imperfectiv y `w´ (price=qavgp qavgpo), prop5 exogvars(addv)
Nevo and Rosen (2012)´s Imperfect IV bounds          Number of obs =    990

Variable          Lower Bound(CI)  LB(Estimator) UB(Estimator) Upper Bound(CI)

price               [-11.374594     (-8.6880097   -5.9886321)    -3.5545173]
addv                [.16464915      (.27997955    .29078735)     .42368452]
```

A.2 Conley, Hansen, and Rossi's (2012) 401(k) example

Below we replicate the plausibly exogenous bounds calculated by Conley, Hansen, and Rossi (2012) in their empirical application examining the effect of participation in a 401(k) on asset accumulation. We use the `plausexog` command described in section 3 to calculate the LTZ bounds.

```
. use conleyetal2012
(Conely et al´s (2012) REStat for 401(k) participation)
. local xvar i2 i3 i4 i5 i6 i7 age age2 fsize hs smcol col marr twoearn db pira
> hown
. plausexog ltz net_tfa `xvar´ (p401 = e401), omega(25000) mu(0) level(.99)
> vce(robust) graph(p401) graphmu(0 1000 2000 3000 4000 5000)
> graphomega(0 333333.33 1333333.3 3000000 5333333.3 8333333.3)
> graphdelta(0 2000 4000 6000 8000 10000)
Estimating Conely et al.´s ltz method
Exogenous variables: i2 i3 i4 i5 i6 i7 age age2 fsize hs smcol col marr twoearn
> db pira hown
Endogenous variables: p401
Instruments: e401
```

Conley et al. (2012)´s LTZ results Number of obs = 9915

	Coef.	Std. Err.	z	P>|z|	[99% Conf.	Interval]
p401	13222.14	1926.609	6.86	0.000	8259.53	18184.76
i2	962.1541	700.6402	1.37	0.170	-842.5755	2766.884
i3	2190.277	992.1113	2.21	0.027	-365.2329	4745.786
i4	5313.626	1420.208	3.74	0.000	1655.411	8971.84
i5	10400.47	2017.663	5.15	0.000	5203.31	15597.62
i6	21859.43	2239.623	9.76	0.000	16090.55	27628.32
i7	62464.83	5871.894	10.64	0.000	47339.83	77589.82
age	-1811.558	536.1392	-3.38	0.001	-3192.561	-430.555
age2	28.68893	6.712006	4.27	0.000	11.39995	45.97791
fsize	-724.4649	378.4213	-1.91	0.056	-1699.214	250.2839
hs	2761.253	1244.257	2.22	0.026	-443.742	5966.247
smcol	2750.739	1643.95	1.67	0.094	-1483.795	6985.273
col	5161.979	1926.959	2.68	0.007	198.461	10125.5
marr	4453.186	1853.123	2.40	0.016	-320.1425	9226.515
twoearn	-15051.59	2125.758	-7.08	0.000	-20527.18	-9576.003
db	-2750.19	1207.883	-2.28	0.023	-5861.49	361.1102
pira	31667.72	1730.29	18.30	0.000	27210.79	36124.65
hown	4200.889	767.7217	5.47	0.000	2223.369	6178.409
_cons	18929.86	9755.124	1.94	0.052	-6197.679	44057.39

The output displayed above documents bounds on each model parameter. Bounds on the endogenous variable of interest (p401) are displayed at the top of the output table and agree with those displayed in figure 2 of Conley, Hansen, and Rossi (2012).

Below we display the output from replicating the full figure 2 of Conley, Hansen, and Rossi (2012) with bounds across a range of priors using the LTZ approach and the graphing capabilities of plausexog.

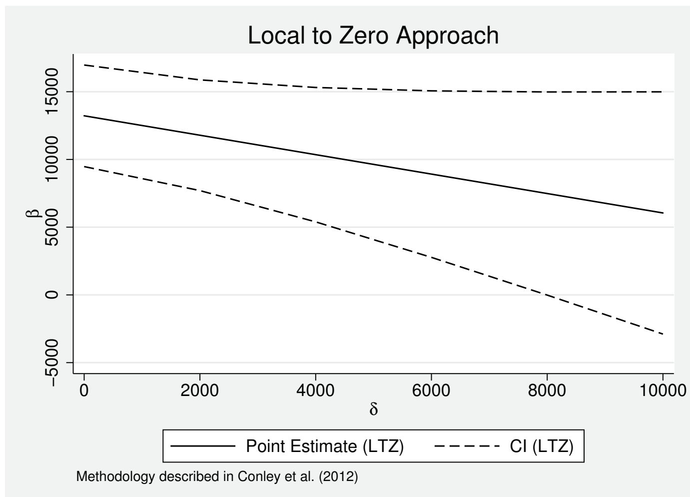


Figure A.1. Replicating figure 2 for 95% confidence intervals with positive prior

B A simple simulated example with multiple instruments

In simulations presented in section 4 and described in the system of equations (8), we considered a case where one plausibly exogenous or imperfect IV ($\mathbf{z}$) exists. To see how this situation is generalized to multiple IV cases, we document a situation below with two IVs suffering similar problems to those described in the article. In this case, two IVs ($\mathbf{z}_1$ and $\mathbf{z}_2$) exist, both of which do not satisfy the exclusion restriction. The violation of the exclusion restriction is larger for the second instrument, and so in both the UCI and LTZ implementations of plausexog, the priors over the sign of $\boldsymbol{\gamma}$ for each IV capture this DGP. In the case of the UCI method, this is accommodated with various (different) values provided in the gmax() option, allowing the violation of the exclusion restriction to reach up to 0.2 for the first instrument and up to 0.4 for the second instrument. Similarly, in the LTZ approach, a normal prior is assumed, and the mean value is assumed to be 0.1 for the first instrument, while up to 0.2 for the second instrument. In the case of the imperfectiv examples, no special considerations need be made, because we simply assume that both instruments are correlated in the same (in this case positive) direction with the unobserved error term. Full output in each of the four columns examined in table 2 is provided below.

```
. set obs 1000
number of observations (_N) was 0, now 1,000

. foreach var in u z1 z2 v w {
  2.     generate `var´=rnormal()
  3. }

. generate x  = -0.6*z1 - 0.40*z2 + 0.33*u  + v

. generate y1 =  3.0*x  + 0.10*z1 + 0.20*z2 + u

. plausexog uci y1 (x = z1 z2), gmin(0 0) gmax(0.2 0.4)
Estimating Conely et al.´s uci method
Exogenous variables:
Endogenous variables: x
Instruments: z1 z2

Conley et al (2012)´s UCI results                  Number of obs =      1000
-----------------------------------------------------------------------------

Variable    Lower Bound     Upper Bound
-----------------------------------------------------------------------------

x           2.6148176       3.3474023
_cons       -.09554232      .05332347
-----------------------------------------------------------------------------

. plausexog ltz y1 (x = z1 z2), mu(0.1 0.2) omega(0.01 0.02)
Estimating Conely et al.´s ltz method
Exogenous variables:
Endogenous variables: x
Instruments: z1 z2

Conley et al. (2012)´s LTZ results                 Number of obs =    1000
-----------------------------------------------------------------------------
             |     Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+---------------------------------------------------------------
           x |  2.988505   .1709607    17.48   0.000     2.653428    3.323582
       _cons | -.0196366   .0363677    -0.54   0.589    -.0909159    .0516428
-----------------------------------------------------------------------------

. imperfectiv y1 (x=z1 z2), noassumption4

Nevo and Rosen (2012)´s Imperfect IV bounds          Number of obs =    1000
-----------------------------------------------------------------------------
Variable         Lower Bound(CI)  LB(Estimator) UB(Estimator) Upper Bound(CI)
-----------------------------------------------------------------------------
x                  [2.6646757      (2.788265     3.1014073)     3.1493633]
-----------------------------------------------------------------------------

. imperfectiv y1 (x=z1 z2)

Nevo and Rosen (2012)´s Imperfect IV bounds          Number of obs =    1000
-----------------------------------------------------------------------------
Variable         Lower Bound(CI)  LB(Estimator) UB(Estimator) Upper Bound(CI)
-----------------------------------------------------------------------------
x                  [2.6636537      (2.788265     2.9651037)     3.0330018]
-----------------------------------------------------------------------------
```