Alternative Graphical Representations of the Confidence Intervals for the Structural Coefficient from Exactly Identified Two-Stage Least Squares.

Joe Hirschberg &Jenny Lye
Economics, University of Melbourne, Australia.

Jan 2017

Abstract:

In the case of the just identified model the exact distribution of the two-stage least squares (2SLS) estimator of the coefficient of the endogenous regressor is a ratio of two normally distributed random variables. Robert Basmann (1960, 1961, 1974) used Fieller's 1932 result to derive the density function of the estimator. In this paper we employ a novel graphical exposition of Fieller's subsequent 1954 technique to approximate the confidence interval for the ratio. This approach involves the construction of a constraint shape that provides an insight as to how the characteristics of the reduced form estimates influences the comparison of the Delta and the Fieller confidence intervals. In particular, the degree of endogeneity and the relevance of the instrument can be shown to have a direct influence on these shapes. An example application of this approach is then applied to consider two specifications of an exactly identified model.

Key words: Indirect Least Squares, Inverse Test, Fieller Method, Anderson and Rubin Test, Delta Method

JEL: C12, C26, C36, C18

1. Introduction

In this paper we demonstrate how the confidence intervals for the parameter estimate from a just identified two-stage least squares (2SLS) estimate can be examined with a single diagram of a two dimensional shape that captures the primary characteristics of the estimating relationships. This graphic method allows for a direct comparison between the usual asymptotic (Delta) confidence interval and the Fieller interval.[1] This graphic representation can be used to verify many of the results previously obtained from Monte Carlo studies that have investigated the properties of the 2SLS estimate in the just identified case.

Although numerous diagrammatic approaches to the Fieller interval exist (see Fieller 1932, 1954; Creasy 1954; Guiard 1989; von Luxburg and Franz 2004; Hirschberg and Lye 2010a, 2010b, 2010c), our contribution is to construct an equivalent diagram for the Delta approximation of the bounds of the confidence interval for the ratio of parameters which can be used to make comparisons between the two methods that incorporates the features of the just-identified 2SLS case.

Drawing inferences from the ratio of regression coefficients is elemental in a number of statistical applications and numerous comparisons between Delta and Fieller confidence intervals have been conducted in a variety of applications. These include: cost–effectiveness ratios; willingness-to-pay measures; bioassay applications for the median dose response; the validation ratio for surrogate endpoint evaluations; and long-run parameters from dynamic regression models. For example, Polsky et al. (1997) examine the performance of the Delta and Fieller intervals for Cost-Effectiveness ratios. For intervals based on Delta method they found only 2% of the errors occurred because the lower limit of the confidence interval exceeded the true ratio whereas 98% of the intervals occurred because the upper limit of the interval fell below the true ratio. Thus this implies that the upper limits of the confidence intervals using the Delta method were often too low. Errors were relatively symmetric however for intervals based on the Fieller method.

[1] As we will discuss below, Anderson and Rubin, profile likelihood, and Fieller intervals all coincide in the case of the just identified two-stage least squares and IV estimator considered here.

Applications in the economics literature are also common. Dufour (1997) proposed that for ratios of regression parameter problems confidence intervals based on the Fieller type methods should be used. Fieller estimates have been used to calculate confidence bounds: for long-run elasticities in dynamic energy demand models (Bernard et al. 2005); mean elasticities obtained from linear regression models (Valentine 1979, Hirschberg et al 2008); non-accelerating inflation rate of unemployment, the NAIRU (Staiger et al. 1997); steady state coefficients in models with lagged dependent variables (Blomqvist 1973) and the extremum of a quadratic model (Hirschberg and Lye 2004, Lye and Hirschberg 2012).

The focus of this paper is the just identified model that is estimated via instrumental variables. This model has been widely used in applied practice. In a survey of applications in which instrumental variable estimation is used for papers published in *The American Economic Review*, *Journal of Political Economy* and *Quarterly Journal of Economics* from 1999 to 2004, Chernozhukov and Hansen (2008) found that the bulk of estimated instrumental variables models employed exactly as many instruments as endogenous regressors. In addition, the advice from Angrist and Pischke (2009) for practitioners to avoid difficulties with weak instruments is to pick a single 'best' instrument and estimate the just-identified model.

This paper proceeds as follows. First we discuss the contribution of Robert Basmann in introducing the use of Fieller's method to examine the finite sample properties of the 2SLS estimator. We then provide a review of the equivalence of the indirect least squares and the 2SLS estimators so that we can examine the nature of the joint distribution of the parameter estimates and show that the 2SLS variance estimate is equivalent to the application of the Delta method for the estimate of the variance of a ratio of random variables. This is followed by the development of a unique method for the comparison of the Delta to the Fieller. A number of cases are then discussed with a simple empirical application. We conclude with a summary of our findings and suggestions for further research in this area.

2. Robert Basmann's Contribution

Robert Basmann has been a pioneer in the estimation and testing methods for simultaneous equations models and finite sample distribution theory. As discussed in Basmann (1993), both Basmann (1957) and Theil (1953) are widely credited with the development of the 2SLS estimator of the coefficients of one structural equation in a simultaneous equations model.[2]

Robert Basmann developed the finite sample distribution theory for the 2SLS in a number of papers. In 1961 Basmann presented the exact distribution of the 2SLS estimator of the estimated endogenous regressor for two particular models. The first case corresponded to a single over-identifying restriction where he proved that the variance of the 2SLS estimator did not exist. The 2SLS estimator was shown to have a finite variance but no higher order moments in the second case when an additional over-identifying restriction was allowed. This led to his conjecture, and later verified by others including Kinal (1980), Hillier et al. (1984) and Phillips (1983), that integer moments of the 2SLS estimator exist up to the degree of over-identification. These results had important implications for the construction and evaluation of Monte Carlo experiments and comparisons of the 2SLS estimator using criteria such as mean-squared error.

When the structural equation is exactly identified, Basmann (1974) showed that the 2SLS estimator of the coefficient of the endogenous regressor is a ratio of two normally distributed random variables and used Fieller's result (1932) to derive the density function of the estimator. One of his motivations for deriving the exact distribution function was to be able to compute numerical approximations. Again using results based on Fieller (1932) he showed that the exact distribution functions could be approximated using a normal distribution under certain conditions (Basmann 1960).

In addition, Robert Basmann was among the first to consider properties of the t-ratio in the simultaneous equations model. Basmann, Richardson and Rohr (1974) reports on Monte Carlo simulations conducted on the *t*-ratio in a simultaneous equations model that contains lagged

[2] Note Anderson and Rubin (1950) also claim to have derived the 2SLS estimator by deriving the asymptotic distribution of the limited information maximum likelihood (LIML) estimator by finding the asymptotic distribution of essentially the 2SLS estimator see Anderson (2005).

endogenous variables as regressors. They conclude that the exact distribution was not affected by the presence of lagged endogenous regressors. In this paper we examine the use of "Fieller's theorem" as presented in Fieller's 1940, 1944 and 1954 contributions, to construct a novel graphical exposition of the confidence intervals based on the Fieller Method.

3. Tests of Significance and Confidence Intervals in the Exactly Identified Case: A Recap

Suppose we are interested in estimating the single parameter β from a structural equation in a simultaneous equation model, [3]

$$Y = X\beta + u \qquad (1)$$

where Y is a T by 1 vector of the dependent variable of interest, X is a T by 1 endogenous regressor vector and u is the T by 1 vector of i. i. d. errors with zero expected value and variance σ_u^2. Instrumental variables estimation requires a T by r matrix Z $(r \geq 1)$ that satisfies two conditions: (*i*) they are uncorrelated with u (a property referred to as: exogeneity) and (*ii*) and are correlated with X (a property referred to as: relevance). If $r = 1$ we say that (1) is exactly identified otherwise it is over-identified.

One of the most frequently used instrumental variables estimators is 2SLS where the instruments are formed as the fitted values from the reduced form equation,

$$X = Z\gamma + \varepsilon \qquad (2)$$

where,

$$\begin{pmatrix} u_i \\ \varepsilon_i \end{pmatrix} \sim iid\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \sigma_u^2 & \sigma_{u\varepsilon} \\ \sigma_{u\varepsilon} & \sigma_\varepsilon^2 \end{bmatrix}\right) \sim iid\left(0, \Sigma\right) \qquad (3)$$

In addition, the reduced form equation corresponding to (1) is written,

$$Y = Z\gamma\beta + v, \qquad (4)$$

where,

[3] Note that (1) and (2) can also include RHS predetermined or exogenous variables but we have just assumed they have been "partialed out" of the specification for expository purposes (see Chernozhukov and Hansen 2008)

$$\begin{pmatrix} v_i \\ \varepsilon_i \end{pmatrix} \sim iid\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \sigma_v^2 & \sigma_{v\varepsilon} \\ \sigma_{v\varepsilon} & \sigma_\varepsilon^2 \end{bmatrix}\right) \sim iid\left(0, \Omega\right) \tag{5}$$

Note that we can define the relationship between the elements of Σ and Ω implied by $v = u + \varepsilon\beta$:

$$\Sigma = \begin{bmatrix} \sigma_u^2 & \sigma_{u\varepsilon} \\ \sigma_{u\varepsilon} & \sigma_\varepsilon^2 \end{bmatrix} = \begin{bmatrix} \sigma_v^2 - \beta\sigma_\varepsilon^2 - 2\beta\sigma_{v\varepsilon} & \sigma_{v\varepsilon} - \beta\sigma_\varepsilon^2 \\ \sigma_{v\varepsilon} - \beta\sigma_\varepsilon^2 & \sigma_\varepsilon^2 \end{bmatrix} \tag{6}$$

$$\Omega = \begin{bmatrix} \sigma_v^2 & \sigma_{v\varepsilon} \\ \sigma_{v\varepsilon} & \sigma_\varepsilon^2 \end{bmatrix} = \begin{bmatrix} \beta\sigma_\varepsilon^2 + 2\beta\sigma_{\varepsilon u} + \sigma_u^2 & \beta\sigma_\varepsilon^2 + \sigma_{u\varepsilon} \\ \beta\sigma_\varepsilon^2 + \sigma_{u\varepsilon} & \sigma_\varepsilon^2 \end{bmatrix} \tag{7}$$

The 2SLS estimator for β in (1) is

$$b_s = \left(X^\mathsf{T} P_Z X\right)^{-1}\left(X^\mathsf{T} P_Z Y\right), \tag{8}$$

where $P_Z = Z\left(Z^\mathsf{T} Z\right)^{-1} Z^\mathsf{T}$, for *Z* full rank. The second equation (2) is known as the '*first stage regression*' that relates the endogenous regressor to a vector of *r* exogenous variables called instruments *Z*.

3.1 2SLS/IV Estimation

Here we summarize the 2SLS and equivalent IV estimator for β in the exactly identified case ($r = 1$). The formula for the more general IV estimator can be defined by pre-multiplying (1) by the transpose of the instrumental variable *Z*.

$$Z^\mathsf{T} Y = Z^\mathsf{T} X\beta + Z^\mathsf{T} u\,. \tag{9}$$

The estimate of β is obtained by pre-multiplying both sides by $\left(Z^\mathsf{T} X\right)^{-1}$, thus the 2SLS estimator is defined as: $b_v = \left(Z^\mathsf{T} X\right)^{-1} Z^\mathsf{T} Y$ and the estimate of the variance is defined as:

$$\mathrm{Var}(b_v) = \left(Z^\mathsf{T} X\right)^{-1} Z^\mathsf{T} E\left[uu^\mathsf{T}\right] Z \left(Z^\mathsf{T} X\right)^{-1} \tag{10}$$

And using the assumption that $E\left[uu^\mathsf{T}\right] = \sigma_u^2 I_n$ $\mathrm{Var}(b_v) = \sigma_u^2 \left(Z^\mathsf{T} X\right)^{-1} Z^\mathsf{T} Z \left(Z^\mathsf{T} X\right)^{-1}$. From equation (2) we can define the least squares estimate of γ as $g = \left(Z^\mathsf{T} Z\right)^{-1} Z^\mathsf{T} X$ thus we can define $\left(Z^\mathsf{T} X\right)^{-1} = \left(Z^\mathsf{T} Z\right)^{-1} g^{-1}$ consequently:

$$\text{Var}(b_v) \approx \sigma_u^2 \left(Z^\top Z\right)^{-1} g^{-2} \tag{11}$$

substituting $S_Z^{-1} = \left(Z^\top Z\right)^{-1}$ and the estimated value of σ_u^2 as s_u^2 we get the estimated variance of b_v as:

$$\widehat{\text{Var}(b_v)} \approx g^{-2} S_Z^{-1} s_u^2 \tag{12}$$

3.2 Indirect Least Squares (ILS) Estimation

In the exactly identified case it is well known that the ILS estimate for the parameter β is the same as the 2SLS and IV solutions and can be found as the ratio of the least squares parameter estimates for $\theta = \gamma\beta$. Here we review the derivation of this result and the corresponding result that the Delta method estimate of the variance of the ratio of the reduced form parameters is equivalent to the asymptotic 2SLS and IV variance estimates.

Define the least squares estimates of θ and γ by *h* and *g*, where $g = \left(Z^\top Z\right)^{-1} Z^\top X$ and $h = \left(Z^\top Z\right)^{-1} Z^\top Y$. The indirect estimate of β is defined as:

$$b_I = \tfrac{h}{g} = \frac{\left(\left(Z^\top Z\right)^{-1} Z^\top Y\right)}{\left(\left(Z^\top Z\right)^{-1} Z^\top X\right)} = \left(Z^\top X\right)^{-1} Z^\top Y \tag{13}$$

Given the covariance of the estimated reduced form parameters is defined by:

$$\text{cov}(h, g) = S_Z^{-1}\Omega \tag{14}$$

And we define the estimated value of $\text{cov}(h, g)$ by:

$$\text{cov}(h, g) = S_Z^{-1} \begin{bmatrix} \sigma_v^2 & \sigma_{v\varepsilon} \\ \sigma_{v\varepsilon} & \sigma_\varepsilon^2 \end{bmatrix} \tag{15}$$

Applying the Delta method to estimate the variance of the ratio of parameters (see for example Casella and Berger (2002), page 240), the approximate variance of the indirect least squares can be defined as:[4]

[4] Note that when estimating the reduced form equations to construct the indirect least squares estimate it is necessary to employ a seemingly unrelated regression type covariance estimate that allows for a non-zero covariance between the errors in the regression and the estimate of $s_{v\varepsilon}$ is computed from the covariance of the residuals from the two reduced form equations.

$$\text{Var}(b_I) \approx \begin{bmatrix} \frac{\partial b_I}{\partial h} & \frac{\partial b_I}{\partial g} \end{bmatrix} S_Z^{-1} \hat{\Omega} \begin{bmatrix} \frac{\partial b_I}{\partial h} \\ \frac{\partial b_I}{\partial g} \end{bmatrix} \tag{16}$$

Thus we have that $b_I = \frac{h}{g}, \frac{\partial b_I}{\partial h} = \frac{1}{g}, \frac{\partial b_I}{\partial g} = -\frac{1}{g^2} h$

$$\text{Var}(b_I) \approx g^{-4} S_Z^{-1} \left(g^2 \sigma_v^2 + h^2 \sigma_\varepsilon^2 - 2\sigma_{v\varepsilon} gh \right) \tag{17}$$

Using the equivalence between the covariance matrices from (6) we have $\sigma_u^2 = \sigma_v^2 - \beta\sigma_\varepsilon^2 - 2\beta\sigma_{v\varepsilon}$ and thus:

$$\sigma_u^2 = \sigma_v^2 - b_I \sigma_\varepsilon^2 - 2 b_I \sigma_{v\varepsilon} \tag{18}$$

Thus we find that:

$$\text{Var}(b_I) \approx g^{-2} S_Z^{-1} \sigma_u^2, \tag{19}$$

which means that the approximation is the same as (11).

To summarize, in the exactly identified case the ILS estimate for β is the same as the 2SLS and IV estimates and the Delta approximation of the variance of the ILS estimate equivalent to the asymptotic 2SLS and IV estimate of the variance of the estimate for β.[5]

3.3 Properties of the Test Statistics for the 2SLS Estimate[6]

Under standard regularity conditions and when both conditions (*i*) instrument exogeneity and (*ii*) instrument relevance are satisfied the 2SLS estimate has an approximate normal distribution in large sample sizes, that is,

$$\sqrt{T}\left(b - \beta\right) \to N\left(0, \sigma_u^2 \left(\gamma^{\mathsf{T}} V_{ZZ} \gamma\right)^{-1}\right),\ V_{ZZ} = \underset{T\to\infty}{plim}\ \frac{1}{T} Z^{\mathsf{T}} Z > 0 .$$

To test the hypothesis $H_0 : \beta = \beta_0$, under standard conditions the Wald statistic,

$$W = \frac{\left(b - \beta_0\right)^2 X^{\mathsf{T}} P_Z X}{\hat{\sigma}_u^2} \tag{20}$$

[5] One implication of the equivalence of the 2SLS and the ILS result via the Delta is the equivalence between the Delta and non-linear least squares as shown by Mikulich et al (2003). This implies that the 2SLS results can also be obtained via SUR estimation of systems of non-linear equations.

[6] To simplify our exposition we will refer to 2SLS or IV or ILS as 2SLS.

has an asymptotic $\chi^2(1)$ distribution where $\hat{\sigma}_u^2$ converges in probability to σ_u^2. If

$\hat{\sigma}_u^2 = \frac{1}{T-r}\left(Y - X\beta\right)^{\mathrm{T}}\left(Y - X\beta\right)$ then W is the square of the asymptotic t-ratio which is commonly used in applied work. Thus confidence intervals can be formed using the 2SLS parameter estimate plus or minus a multiple of the asymptotic error and we will refer to this interval as the Delta method interval. There is now a large literature that examines the finite sample properties of this approach as well as suggesting alternative techniques.

For (1) above, Richardson and Rohr (1971) derive an exact distribution of a t-ratio under the null hypothesis. However, the t-ratio is based on an inconsistent estimator of the error variance. Basmann, Richardson and Rohr (1974) perform Monte Carlo simulations on the t-ratio in a simultaneous equations model that contains lagged endogenous variables as regressors. They conclude that the exact distribution was not affected by the presence of lagged endogenous regressors.

More recently, emphasis in the literature has been given to investigating the properties of t-statistics and confidence intervals for instrumental variable estimation with weak instruments, that is, when instrument relevance is low. Staiger and Stock (1997) show that under weak instruments $\hat{\beta}$ is not consistent. Dufour (1997) shows that if the correlation between the instruments and the regressor are arbitrary close to zero as in the case of weak instruments, any robust testing procedure must produce confidence intervals of infinite length with positive probability. This implies that the standard Delta Wald test for testing $H_0 : \beta = \beta_0$ in (1) cannot be robust to weak instruments since the corresponding confidence set is finite with probability one.

The poor performance of Delta confidence intervals in the presence of weak instruments has been illustrated in a number of Monte Carlo experiments. Zivot, Startz and Nelson (1998) show that in the presence of weak instruments and strong endogeneity Delta confidence intervals lead to the wrong conclusion and the probability they reject the null hypothesis is far greater than their nominal size. Kiviet (2013) using Monte-Carlo experiments show that for a weak instrument in the just-identified equation that the Wald test based on the 2SLS estimated variance can both be conservative (when the simultaneity is moderate) and yield under coverage (when the simultaneity is more serious). Similar results have also been shown in Hall et al. (1996) and Dufour and Khalaf (2001).

The Anderson-Rubin statistic (*AR*) provides an alternative approach to hypothesis tests and constructing confidence intervals. Taking (1) and adding and subtracting $X\beta_0$ from both sides and substituting for *X* from (2) results in:

$$\left(Y - X\beta_0\right) = Z\gamma\left(\beta - \beta_0\right) + v\left(\beta - \beta_0\right) + u = Z\gamma\varphi + u^* \tag{21}$$

where $u^* = v\left(\beta - \beta_0\right) + u$. To test $H_0 : \beta = \beta_0$ in (1) corresponds to testing $H_0 : \varphi = 0$ in (21) which can be tested using a standard *F*-statistic,

$$AR = \frac{\left(Y - X\beta_0\right)^{\mathsf{T}} P_Z \left(Y - X\beta_0\right) / r}{\left(Y - X\beta_0\right)^{\mathsf{T}} M_Z \left(Y - X\beta_0\right) / \left(T - r\right)}, \tag{22}$$

where $M_Z = I - P_Z$. The *AR* statistic is distributed exactly as $F\left(r, T - r\right)$ if the identifying restrictions that exclude *Z* from (1) are true and if $\left(u_i, v_i\right) \sim iid\ N\left(0, \Sigma\right)$. Otherwise, Staiger and Stock (1997) show that *AR* is distributed $\chi^2\left(r\right)/r$ under fairly general assumptions about the disturbances and quality of the instruments. The confidence interval based on the *AR* statistic consists of all values of β_0 that satisfy,

$$\frac{\left(Y - X\beta_0\right)^T P_Z \left(Y - X\beta_0\right)}{\left(Y - X\beta_0\right)^T M_Z \left(Y - X\beta_0\right)} \leq 1 + F_{1-\alpha}\left(r, T - r\right)\frac{r}{T - r} \tag{23}$$

where $F_{1-\alpha}\left(r, T - r\right)$ denotes the $\left(1 - \alpha\right)$ quantile of the *F* distribution. The *AR* confidence interval can be bounded, unbounded or cover the real line. Furthermore, for overidentified models it can be empty. The unbounded interval occurs if the *F*-statistic for testing $\gamma = 0$ in (2) cannot be rejected at the $F_{1-\alpha}\left(r, T - r\right)$ level of significance, in other words when the '*first stage regression'* is not significant. This condition is related to the detection of weak instruments. Staiger and Stock (1997) provide a '*rule of thumb*' that instruments are weak if the *F*-statistic for testing $\gamma = 0$ (a test of instrument relevance) in (2) is less than 10, (see also Stock and Yogo 2005 for formal tests of weak instruments) or in the case of an exactly identified regressor when the *t*-statistic is less than $\sqrt{10} \approx 3.163$.

The null distribution of the *AR* statistic does not depend on γ and so it is considered to be fully robust to weak instrumental variables. However, it is not robust to heteroskedasticity and/or autocorrelation of the structural error *u* (see Andrews and Stock 2006).

When $r = 1$ the confidence intervals based on the AR statistic can be shown to be equivalent to the Fieller intervals. Moreira (2001) shows that it is UMP unbiased when $r = 1$ and errors are *iid* homoscedastic normal and Ω is known. Zivot, Startz and Nelson (1998) investigate properties of 95% confidence intervals obtained in this case using Monte Carlo simulations. A range of cases were considered that covered a weak to strong instrument and they conclude that in all cases the confidence intervals have empirical coverage close to 95%. Chernozhukov and Hansen (2008) show how to make the *AR* statistic in the just-identified case robust to heteroskedasticity and autocorrelation through the use of conventional robust covariance matrix estimators. For over-identified models however, the power of the AR statistic has been shown to be not so good and the literature has sought more powerful tests that are robust to weak instruments (see Andrews and Stock 2006).

4. The Boomerang and the Ellipse: A Comparison of the Fieller and Delta Intervals

In Hirschberg and Lye (2010b) (henceforth referred to as HL) we present a method for the graphical comparison of the Fieller and the Delta confidence intervals. In that paper we demonstrate that the intervals from those two methods could be found by a graphic solution to a constrained optimization problem as proposed by Durand (1954), Scheffé (1959, appendix III) and in the econometrics literature by Leamer (1978, Theorem 5.4). These earlier contributions demonstrate that $100(1-\alpha)\%$ confidence bounds for a univariate linear combination of a vector of parameters whose estimates are multivariate normally distributed, are the maximum and minimum solutions from a constrained extrema problem. Specifically, given a linear combination of a $k \times 1$ vector $\boldsymbol{\beta}^{\mathrm{T}} = \{\beta_1, \beta_2, \cdots, \beta_k\}$ defined as $\theta = \mathbf{a}^{\mathrm{T}}\boldsymbol{\beta} + c$, in which $\mathbf{a}$ is a $k \times 1$ vector and *c* is a scalar, the constrained extrema problem is defined by the Lagrangian function:

$$\mathcal{L} = \left(\mathbf{a}^{\mathrm{T}}\boldsymbol{\beta} + c\right) - \lambda\left(\left(\mathbf{B} - \boldsymbol{\beta}\right)^{\mathrm{T}} \boldsymbol{\Sigma}^{-1}\left(\mathbf{B} - \boldsymbol{\beta}\right) - z^2\right) \tag{24}$$

In (24), $\mathbf{B}_{k\times1} \sim N\left(\boldsymbol{\beta}_{k\times1}, \boldsymbol{\Sigma}_{k\times k}\right)$, $\mathbf{B}^{\mathrm{T}} = \{b_1, b_2, \cdots, b_k\}$ and λ is the Lagrange multiplier. The proof of this proposition is a straightforward application of matrix calculus. Note that *z* is the 100(1-α)% univariate bound from a standard normal.[7]

Von Luxburg and Franz (2004) show that in the case where *k = 2* Fieller's bounds for the ratio of two normally distributed variables b_1/b_2 where:

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \sim N\left\{\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}, \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}\right\} \tag{25}$$

can be found as the maximum and minimum from the constrained optimization with the same constraint as in (24) which is defined by:[8]

$$\mathcal{L} = \left(\frac{\beta_1}{\beta_2}\right) - \lambda\left(\left[(b_1 - \beta_1) \quad (b_2 - \beta_2)\right]\begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}^{-1}\begin{bmatrix} (b_1 - \beta_1) \\ (b_2 - \beta_2) \end{bmatrix} - z^2\right). \tag{26}$$

In HL we demonstrate that the maximum and minimum values for the ratio of these variables can be found from a graph of the ellipse implied by the constraint using a construction of rays from the origin and that these intervals are equivalent to the Fieller interval. In addition, we also show that the corresponding Delta method confidence interval can be determined from this diagram. In that paper we show how the relationship between the two intervals is a function of the univariate confidence bound of the denominator and the correspondence of the signs of b_1/b_2 and σ_{12}. We conclude that when the signs are the same and we can reject the null that $\beta_2 = 0$ with a

---

[7] In the remainder of this paper we will use *z* although in small samples the equivalent *t* may be used.

[8] See Hirschberg and Lye (2010c) for the proof that the confidence intervals of the linear function and from the Fieller method are solutions to the optimizations defined by (24) and (26).

high degree of confidence the Fieller and Delta confidence intervals will be closely in agreement. However, when this is not the case they will not necessarily agree. These results are demonstrated by Figures 4 and 5 in HL. We show that as the *p*-value for the test that $\beta_2 = 0$ increases or equivalently the absolute value of the *t*-statistic for the numerator becomes lower in magnitude - the agreement between the signs becomes more important in determining the differences between the Fieller and Delta confidence intervals.

Unfortunately, this graphic construction of the Fieller and Delta intervals becomes rather complex and can be cumbersome to employ for comparisons between the Fieller and the Delta intervals. In order to facilitate the interpretation of the differences between these methods we reparameterized the optimization problem so that we can construct a new constraint in two dimensions. By reparameterizing the constraint in terms of the estimated value of the ratio we find that the limits of the two dimensional shape define the confidence interval.

In general if we consider a function of two parameters $\theta = f(\beta_1, \beta_2)$ that is subject to a quadratic constraint

$$\mathcal{L} = f(\beta_1, \beta_2) - \lambda \left(\begin{bmatrix} (b_1 - \beta_1) & (b_2 - \beta_2) \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}^{-1} \begin{bmatrix} (b_1 - \beta_1) \\ (b_2 - \beta_2) \end{bmatrix} - z^2 \right). \tag{27}$$

In addition if we can solve this function for one of the parameters as in the form : $\beta_1 = g(\theta, \beta_2)$. Then we can substitute for β_1 in the function defined by:

$$\mathcal{L} = \theta - \lambda \left(\begin{bmatrix} (b_1 - g(\theta, \beta_2)) & (b_2 - \beta_2) \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}^{-1} \begin{bmatrix} (b_1 - g(\theta, \beta_2)) \\ (b_2 - \beta_2) \end{bmatrix} - z^2 \right) \tag{28}$$

This implies that the solution is to find the limits of the implied equation defined by the constraint shape defined by the expression:

$$\begin{bmatrix} (b_1 - g(\theta, \beta_2)) & (b_2 - \beta_2) \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}^{-1} \begin{bmatrix} (b_1 - g(\theta, \beta_2)) \\ (b_2 - \beta_2) \end{bmatrix} = z^2 \tag{29}$$

Thus we can plot this constraint in (θ, β_2) space and observe the limits to find the appropriate

100(1-α)% interval.

In the case of the Fieller approach applied to the ratio of random normal variables we define the function $f(\beta_1, \beta_2)$ as $\theta = \beta_1 / \beta_2$. Thus we can solve for $\beta_1 = \theta\beta_2$ and the constraint shape is provided by:

$$\begin{bmatrix}(b_1 - \theta\beta_2) & (b_2 - \beta_2)\end{bmatrix}\begin{bmatrix}\sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22}\end{bmatrix}^{-1}\begin{bmatrix}(b_1 - \theta\beta_2) \\ (b_2 - \beta_2)\end{bmatrix} = z^2 \tag{30}$$

In the case of the Delta approach we user the definition of the estimate of the ratio as $\hat{\theta} = b_1 / b_2$ we can define the Delta estimate as the first order Taylor series approximation around the estimates b_1 and b_2 as:

$$\begin{aligned}\theta &= \hat{\theta} + \begin{bmatrix}\frac{\partial\hat{\theta}}{\partial b_1} & \frac{\partial\hat{\theta}}{\partial b_2}\end{bmatrix}\begin{bmatrix}(\beta_1 - b_1) \\ (\beta_2 - b_2)\end{bmatrix} \\ &= \tfrac{b_1}{b_2} - \tfrac{1}{b_2}(\beta_1 - b_1) + \tfrac{b_1}{b_2^2}(\beta_2 - b_2)\end{aligned} \tag{31}$$

When solving for β_1 we derive a linear function of θ and β_2 :

$$\beta_1 = b_2\left(\theta + \tfrac{b_1}{b_2^2}(\beta_2 - b_2)\right), \tag{32}$$

use (32) to substitute for β_1 into (30) we obtain the formula for a constraint shape as:

$$\begin{bmatrix}\left(-\tfrac{1}{b_2}\left(\theta b_2^2 - 2b_1b_2 + \beta_2 b_1\right)\right) & (b_2 - \beta_2)\end{bmatrix}\begin{bmatrix}\hat{\sigma}_{11} & \hat{\sigma}_{12} \\ \hat{\sigma}_{12} & \hat{\sigma}_{22}\end{bmatrix}^{-1}\begin{bmatrix}\left(-\tfrac{1}{b_2}\left(\theta b_2^2 - 2b_1b_2 + \beta_2 b_1\right)\right) \\ (b_2 - \beta_2)\end{bmatrix} = z^2 \tag{33}$$

4.1 The Just Identified Simultaneous Equations Model.

In order to make comparisons between the Delta and Fieller confidence intervals for the parameter β from the structural equation $Y = X\beta + u$ we consider the parameters from the reduced form equations (2) and (4). Where β is estimated by ILS as the ratio of the estimates of θ as *h* and γ as *g*, where $g = (Z^{\top}Z)^{-1} Z^{\top}X$ and $h = (Z^{\top}Z)^{-1} Z^{\top}Y$. From (5) we have:

$$\begin{pmatrix}v_i \\ \varepsilon_i\end{pmatrix} \sim \text{i.i.d. } (0, \Omega)$$

Thus via the properties of the reduced form estimates obtained via OLS we have the asymptotic result:

$$\begin{bmatrix} h \\ g \end{bmatrix} \underset{a}{\sim} \left[\begin{pmatrix} \theta \\ \gamma \end{pmatrix}, \left(Z^{\mathrm{T}} Z \right)^{-1} \Omega \right] \tag{34}$$

In a series of Monte Carlo studies Zivot, Startz and Nelson (1998) consider this model in order to examine the implications of various scenarios with respect to the parameters of this model. In particular they examine the case where $\sigma_\varepsilon^2 = 1$ and $\sigma_u^2 = 1$, and defining the instrument as an independent standard normal variable $\left(Z \sim N(\mathbf{0}, \mathbf{I}_T) \right)$ thus the expected value of the sum of squares of the instrument would be equal to the number of observations. Thus:

$$\begin{bmatrix} h \\ g \end{bmatrix} \sim \left[\begin{pmatrix} \theta \\ \gamma \end{pmatrix}, \frac{1}{T} \times \begin{pmatrix} \beta + 2\beta\rho + 1 & \beta + \rho \\ \beta + \rho & 1 \end{pmatrix} \right], \tag{35}$$

where $\rho = \frac{\sigma_{\varepsilon u}}{\sigma_\varepsilon \sigma_u}$ is the correlation between the error ε in the first stage equation (2) and the structural equation error u in specification (1). The value of ρ can be viewed as a measure of the endogeneity of X, thus the larger the value of ρ the greater the problem of endogeneity in the estimation of the structural equation (1). By assuming a value for b and ρ changing the implied values of g (and by implication h), we can investigate the comparison between the Delta and Fieller confidence intervals using the implied shapes defined above in (30) and (33).

In these cases we assume we can obtain the inverse of Ω as

$$\Omega^{-1} = \tfrac{1}{\rho^2 - 1} \begin{pmatrix} -1 & \beta + \rho \\ \beta + \rho & -\left(\beta^2 + 2\rho\beta + 1 \right) \end{pmatrix} \tag{36}$$

and define the quadratic constraint in the Lagrangian from (28) as:

$$\left(\rho^2 - 1 \right)^{-1} \left[(h - \theta) \quad (g - \gamma) \right] \begin{pmatrix} -1 & \beta + \rho \\ \beta + \rho & -\left(\beta^2 + 2\rho\beta + 1 \right) \end{pmatrix} \begin{bmatrix} (h - \theta) \\ (g - \gamma) \end{bmatrix} = \frac{z^2}{T}, \tag{37}$$

which is equivalent to:

$$(g - \gamma)\left(\left(\rho + \tfrac{1}{g} h \right) \tfrac{h - \theta}{\rho^2 - 1} - \tfrac{g - \gamma}{\rho^2 - 1} \left(\tfrac{1}{g^2} h^2 + \tfrac{2}{g} h\rho + 1 \right) \right) - (h - \theta)\left(\tfrac{h - \theta}{\rho^2 - 1} - \left(\rho + \tfrac{1}{g} h \right) \tfrac{g - \gamma}{\rho^2 - 1} \right) = \tfrac{1}{T} z^2. \tag{38}$$

Under the assumption that $T = 100$, $g = .5$, $h = .5$, $b = 1$, and $z = 1.96$, this constraint can be drawn in (θ,γ) space as an ellipse as shown in Figure 1 for $\rho=.01$ and $\rho=.99$.

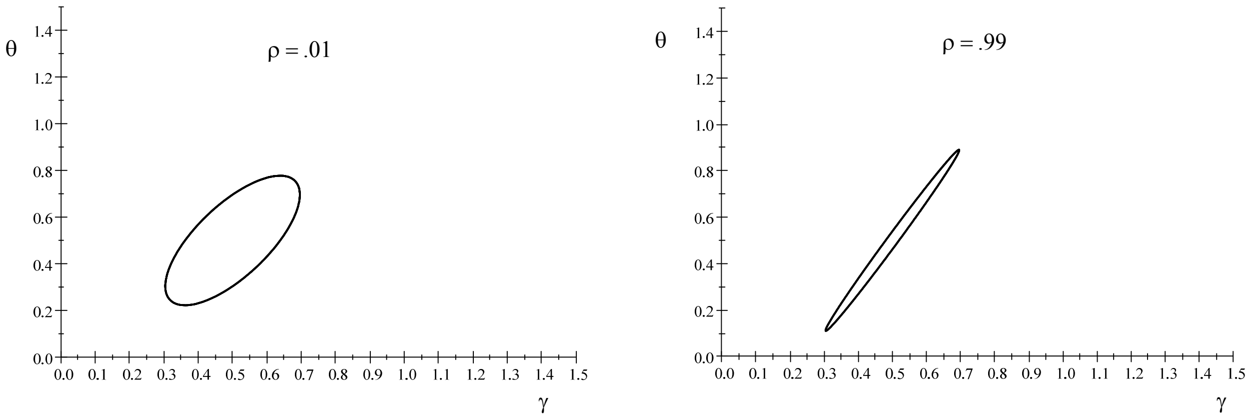


Figure 1 The constraint ellipse when $T=100$, $g=.5$, $h=.5$, $b=1$, $z=1.96$ for $\rho=.01$ and $\rho=.99$.

Note that the univariate 95% confidence interval for γ can be found from the limits of the ellipse on the horizontal axis and the equivalent interval for θ on the vertical axis. Because the variance for θ is a function of ρ we note that values of ρ influence the confidence interval for θ but the interval for γ remains unaffected by ρ. From Figure 1 it can also be noted that the significance of the estimate for γ is directly related to the distance of the ellipse from the origin.

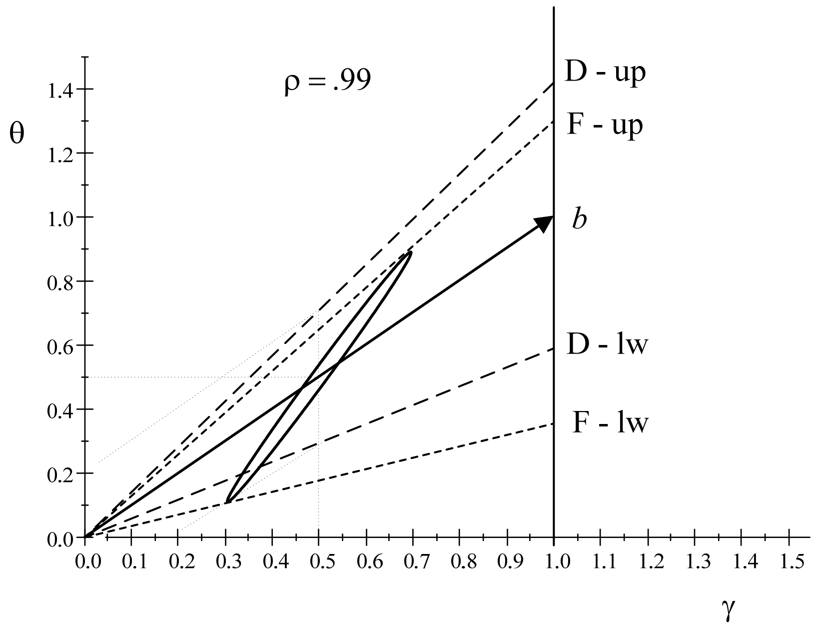


Figure 2 The construction of the Fieller bounds ((F-up, F-lw) and Delta bounds (D-up, D-lw) for β defined from the constraint ellipse when $T=100$, $g=.5$, $h=.5$, $b=1$, $z=1.96$ for $\rho=.99$

Figure 2 follows the graphic method proposed in HL for this case to compare the Delta and Fieller intervals for the case when $\rho=.99$. From this figure it can be seen that the Fieller and Delta intervals are not the same. The Delta interval is always symmetric by construction and in this case both the Fieller interval lower bound and upper bound are lower than the corresponding Delta bounds. The interpretation of this figure is complicated by the number of extra lines needed to define the Delta interval as well as the Fieller. In the next section we propose a diagrammatic method that provides a cleaner comparison of the two sets of intervals

4.2 The Egg and the Boomerang Constraint Shapes for the Fieller Interval

In order to produce a simple diagram from which the Fieller limits can be read directly from the constraint surface we use the definition of θ in terms of the ratio of the parameters β and the denominator γ as defined by $\theta = \gamma\beta$. In this case the expression (30) can be written as:

$$-(g-\gamma)\left(\tfrac{g-\gamma}{\rho^2-1}\left(b^2+2\rho b+1\right)-\tfrac{b+\rho}{\rho^2-1}\left(h-\beta\gamma\right)\right)-\left(\tfrac{1}{\rho^2-1}\left(h-\beta\gamma\right)-\left(b+\rho\right)\tfrac{g-\gamma}{\rho^2-1}\right)\left(h-\beta\gamma\right)=\tfrac{1}{T}z^2, \quad (39)$$

which defines a constraint shape in (β,γ) space.

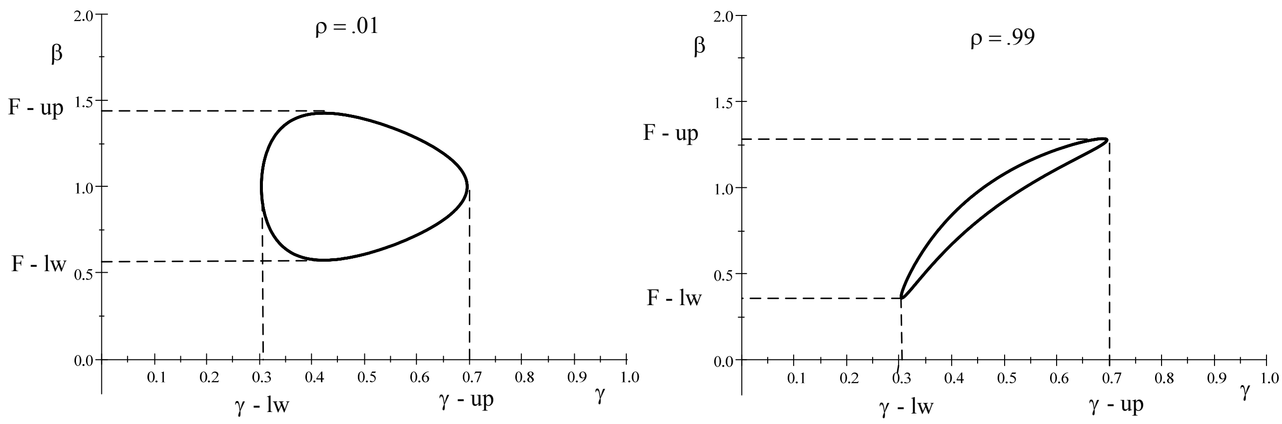


Figure 3 The Fieller 95% confidence bounds for β (F-up, F-lw) are defined from the constraint surface when $T=100$, $g=.5$, $h=.5$, $b=1$, $z=1.96$ for $\rho=.99$ and .01.

Figure 3 demonstrates the case when the Fieller bounds are found to be finite since the constraint shapes in these cases are closed. These two cases indicate how the symmetry and limits of the confidence interval are influenced by ρ when the location of the estimate of γ and the implied test of instrument relevance is held constant. Note that the limits for the confidence interval for the denominator coefficient (γ) can also be found from this graph by the relative distance of the shape from the origin. Thus the distance of the lower limit of the confidence bound for γ is directly related to the significance of the parameter for the instrumental variable in the first stage regression which can be interpreted as a measure of the relevance of the instrument and the shape is indicative of the implied endogeneity.[9]

4.3 The Delta Interval Ellipse Constraint Shape

The Delta method approximates the ratio with a linear combination of the numerator and denominator. Using this linear combination we find that the resulting constraint shape in the (β, γ) space is an ellipse just like the original quadratic constraint defined in (θ, γ) space. Consider the Delta approximated value which is based on a first order Taylor series expansion around the estimated values of θ and γ defined by *h* and *g*. Given the function for the estimate $b = h / g$ we can estimate β by:

$$\beta = b + \begin{bmatrix} \frac{\partial b}{\partial h} & \frac{\partial b}{\partial g} \end{bmatrix} \begin{bmatrix} (\theta - h) \\ (\gamma - g) \end{bmatrix} = \tfrac{1}{g^2}\left(g\theta - h\gamma + gh\right) \qquad (40)$$

Now we solve (40) for θ and we obtain:

$$\theta = g\beta + \tfrac{h}{g}\gamma - h \qquad (41)$$

thus θ can be defined as a linear function of β and γ where the estimated coefficients determine the slope and the constants in the equation. Now to substitute this expression into the quadratic constraint defined in (38) we obtain the following constraint ellipse in (β, γ) space.

$$\begin{pmatrix} -\left(\frac{1}{\rho^2-1}\left(g\beta - 2h + \frac{1}{g}h\gamma\right) + (b+\rho)\frac{g-\gamma}{\rho^2-1}\right)\left(g\beta - 2h + \frac{1}{g}h\gamma\right) \\ -(g-\gamma)\left(\frac{g-\gamma}{\rho^2-1}\left(b^2 + 2\rho b + 1\right) + \frac{b+\rho}{\rho^2-1}\left(g\beta - 2h + \frac{1}{g}h\gamma\right)\right) \end{pmatrix} = \tfrac{1}{T}z^2 \qquad (42)$$

From Figure 4 we can see that unlike the case for the Fieller interval, the value of ρ does not influence the confidence bounds we find for the Delta estimate. Note that these confidence intervals match the result we obtained from Figure 2.[10]

---

[9] Conversely, if the value of γ is estimated as negative this would be indicated by the distance from the upper limit of the confidence interval to zero.

[10] When setting the $\sigma_\varepsilon^2 = 1$ and $\sigma_u^2 = 1$ the Delta/2SLS variance for β is can be shown to be $N^{-1}g^{-1}$ thus it is not a function of ρ.

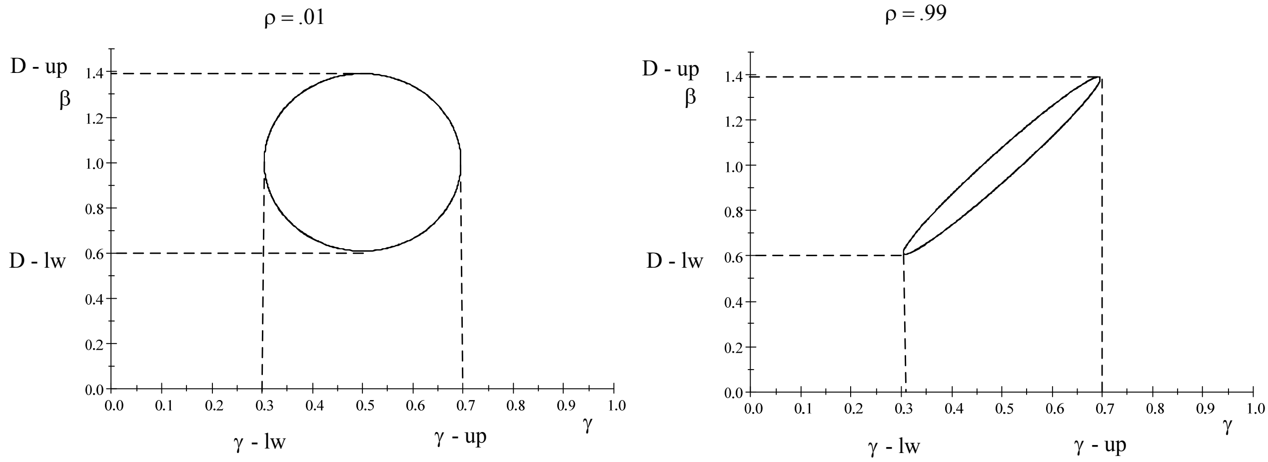


Figure 4 The Delta method 95% confidence bounds (D-up, D-lw) for β defined from the constraint surface when $T = 100, g = .5, h = .5, b = 1, z = 1.96,$ and $\rho = .99, \rho = .01$.

4.4 A Comparison of the Delta and Fieller Constraint Shapes When ρ Varies.

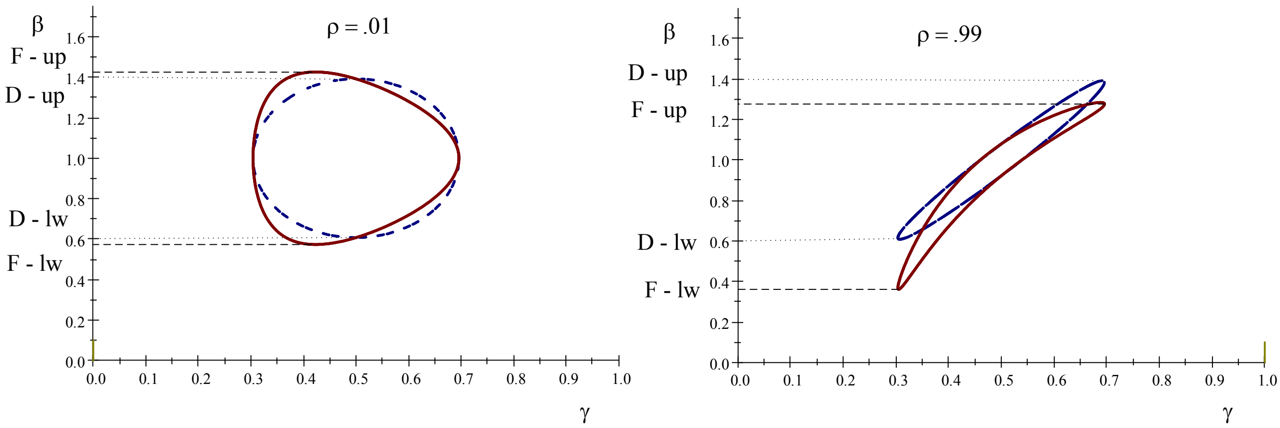


Figure 5 The Delta and Fieller methods 95% confidence bounds for β defined from the two constraint surfaces when $T = 100$, $g = .5$, $h = .5$, $b = 1$, $z = 1.96$ for $\rho = .99$ and .01.

Figure 5 provides a comparison of the Delta and Fieller method 95% confidence bounds when the measure of endogeneity (ρ) equals .01 and .99 by overlaying Figures 3 and 4. This comparison plot allows the consideration of the differences between these methods.

In Figure 6 we assume a sample size of 50 and a 99% confidence level which implies that the reduced form estimates are estimated with greater standard errors. In this case the comparisons of the two intervals indicate that the lower bound for the Fieller would be less than zero for the two values of ρ. In the case when $\rho = .99$ the Fieller method indicates a lower positive upper bound than the Delta method and a lower lower bound. Thus if one were to base a hypothesis test using the Delta interval they would be able to reject $H_0 : \beta = 0$ versus $H_1 : \beta \neq 0$ at the 99% level of confidence while the Fieller interval would not allow a rejection of the two-sided alternative. In the case that ρ = .01 we find that the Fieller constraint shape becomes wider and wider as the lower bound for the estimate of γ approaches zero.

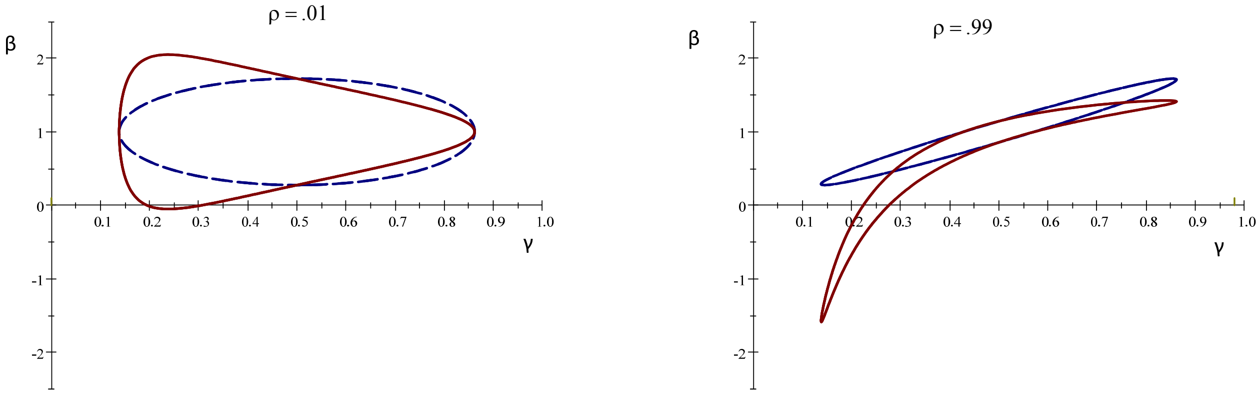


Figure 6 The Delta and Fieller methods 99% constraint surfaces when $T = 50$, $g = .5$, $h = .5$, $b = 1$, $z = 1.96$ for $\rho = .99$ and .01.

A characteristic of the Fieller method is that the resulting confidence interval may not necessarily have finite limits when (39) defines an open constraint shape. The confidence interval may also be the complement of a finite interval or the whole real line (e.g. Scheffé 1970). When the constraint shape is closed in only one direction, it may only be possible to find a finite lower bound but no finite upper bound for a particular confidence level or vice versa. This case can be shown within the examples examined here by assuming a smaller sample size or a different confidence level. (See Figure 8 for examples of open Fieller confidence shapes).

4.5 A Comparison of Delta and Fieller Confidence Bounds when the relevance of the instrument varies.

In these cases we alter the value of the estimate of the parameter γ in the first stage regression. A number of previous studies have examined the impact of the case when the relevance of the instrument as measured by the *F*-statistic for the instrument parameter in the first stage regression is low and thus indicates that the instrument may have a low degree of relevance and thus be weak. Here we assume the value of ρ = .9 implying the presence of endogeneity.

Figure 7 shows the variation in the constraint ellipse in (θ, γ) space when the value of the estimate for γ varies. Note we continue to assume that the estimate of β (*b*) = 1 thus we change the value of g in order to correspond to the value of h by ensuring that h is equal to g. Thus the ellipse defined by the estimates of θ and γ will move toward the origin as g approaches zero.

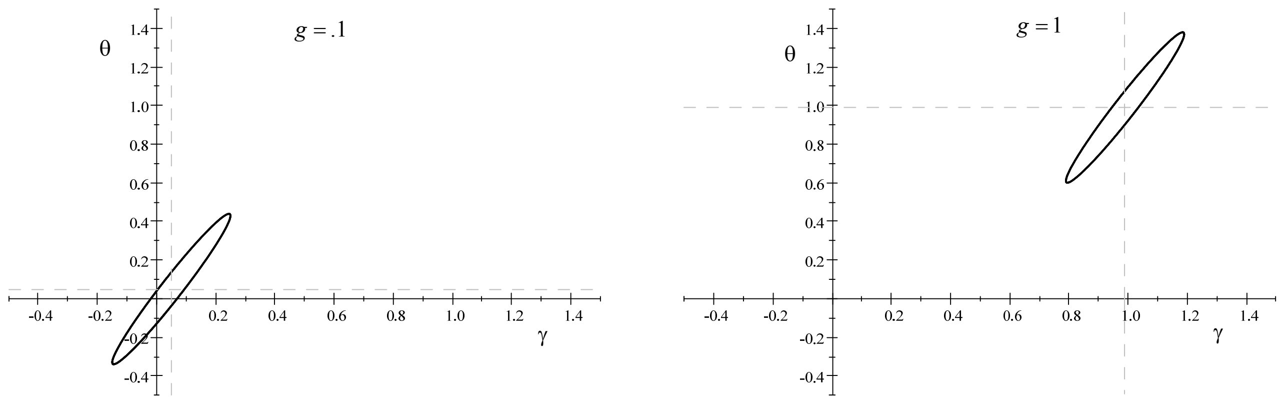


Figure 7. The 95% constraint ellipse in (θ,γ) space when $T=100$, $g=h$, $b=1$, $\rho=.9$, $z=1.96$ for $g=1$ and .1.

Figure 8 is the equivalent to Figure 6 where we examine the constraint for the Fieller and Delta methods in (γ,β) space. In this case the value of g implies the relevance of the instrument. Note that in this case we have assumed the variance of g is defined as $1/T$ when T = 100 the standard error of g is .1 and the value of g at the limit of the weak instruments as suggested by an F = 10 would be when g = .316. In Figure 8 we have plotted for cases where the instrument has varying degrees of relevance based on F-statistics of 1, 4, 10 and 20.

From Figure 8 it can be noted that Fieller method results in a finite upper bound in all of these cases however the lower bound may not be finite. Even when the F-statistic = 1 we obtain a finite upper bound away from zero that is considerably lower than the Delta equivalent.[11] From these cases we find that although the Delta 95% CI for the estimate of β is well away from zero the Fieller interval indicates that the upper bound is lower but the lower bound is less than zero. The three plots for the cases where $F < 10$ result in open constraints shapes for the Fieller. This indicates that in these cases the Fieller lower bounds are not finite.

[11] In this case we have the complementary surface for the Fieller where an alternative lower bound appears but is above the upper bound that has already been found.

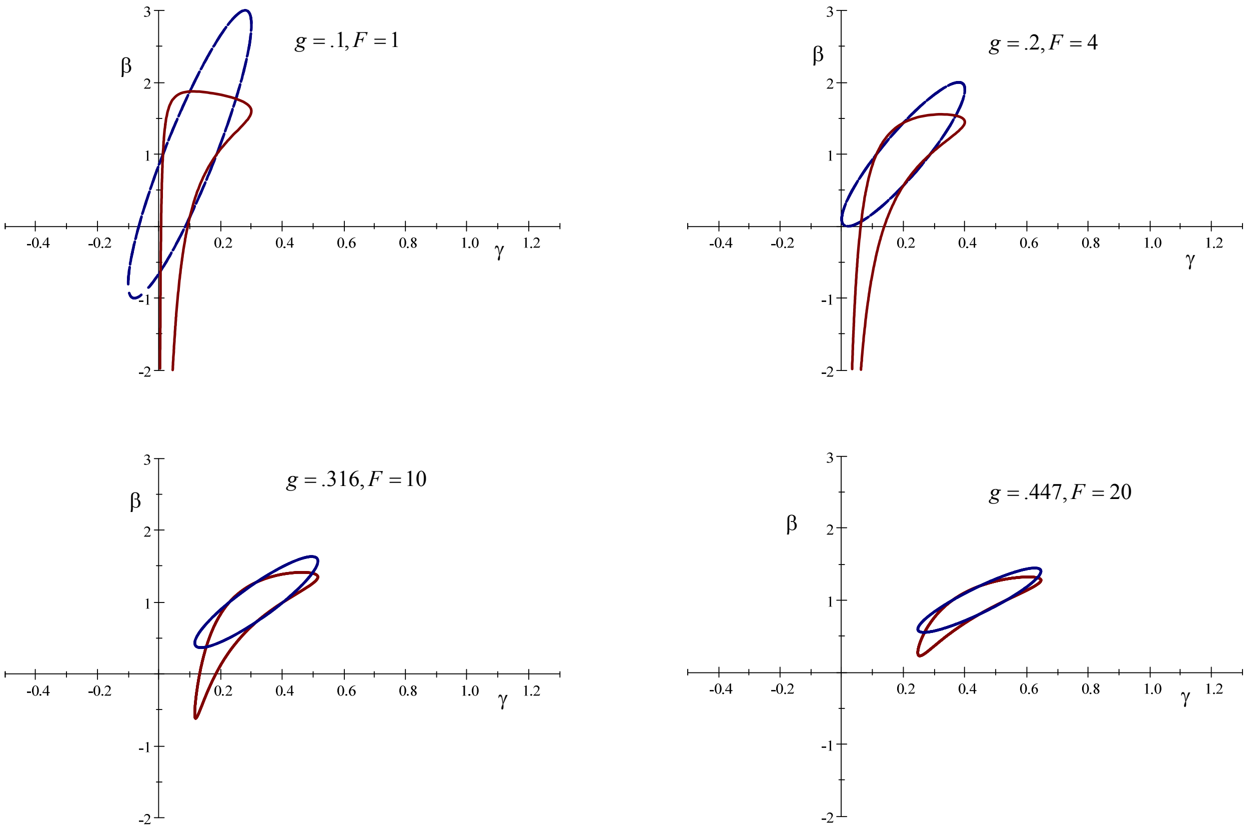


Figure 8. The 95% constraint shapes in (γ,β) space for the Fieller and Delta intervals when $T=100$, $g=h$, $b=1$, $\rho=.9$, $z=1.96$ for four values of g and the implied values of the F-statistic for the first stage regression.

5. An Empirical Example

A widely cited example of a case where the implications of a just identified system of equations are of interest is the paper by Acemoglu et al. (2001) in which they employ a cross country data set to examine the effect of institutions on economic performance. In order to account for the endogeneity of the "index of protection against expropriation" or *IPAE*[12] in explaining a nation's per capita income, they use the mortality rates among European colonists as an instrumental variable.

The structural equation is defined by:

$$\log(y_i) = \mu + \beta R_i + W_i\lambda + u_i \tag{43}$$

where y_i is income per capita in country i, R_i is the *IPAE* for country i, W_i is a vector of location indicators, and λ is the vector of parameters. The *IPAE* variable (R_i) is treated as an endogenous

12 See Coplin et al. (1991) for details of this variable.

variable. Thus the reduced form is:

$$R_i = \eta + \gamma \log(M_i) + W_i'\delta + \varepsilon_i \tag{44}$$

where M_i is the settler mortality rate used as the instrument. The reduced form equation for log income per capita is

$$\log(y_i) = \lambda + \theta \log(M_i) + W_i'\phi + v_i \tag{45}$$

Thus this model is just-identified with $\log(M_i)$ being excluded from equation (43). Table 1 reports 2SLS estimates for two specifications when the dependent variable is Log GDP per capita.[13] In the first specification only one regressor is assumed and the second also includes the location covariates.

	Specification	
VARIABLES	*1*	*2*
IPAE	0.944	1.107
	(6.129)	(2.509)
Latitude		-1.178
		(-0.705)
Africa dummy		-0.437
		(-1.083)
Asia dummy		-1.047
		(-2.097)
"Other" continent dummy		-0.990
		(-1.042)
Constant	1.910	1.440
	(1.890)	(0.533)
Observations	64	64
R-squared	0.187	0.011

t-statistics in parentheses based on asymptotic standard error estimates

Table 1 The 2SLS estimates for the two specifications.

Table 2 provides the equivalent estimates of the two reduced form equation for the two specifications along with the estimate of the correlation between the estimates of θ and γ. Again as shown above the indirect least squares estimate is identical to the 2SLS estimate for the parameter on the endogenous variable of interest.

13 These results are equivalent to columns 1 and 8 in Table 4 reported in Acemoglu et al (2001) and are computed using Stata 14 based on the data available from http://economics.mit.edu/faculty/acemoglu/data/ajr2001.

	Specification			
VARIABLES	*1*	*2*	*1*	*2*
	Log GDP per capita		*IPAE*	
Log European settler mortality	-0.573	-0.377	-0.607	-0.340
	(-7.644	(-3.868)	(-4.867)	(-1.953)
Latitude		1.046		2.009
		(1.414)		(1.518)
Africa dummy		-0.723		-0.258
		(-3.317)		(-0.662)
Asia dummy		-0.525		0.472
		(-1.96)		(0.986)
"Other" continent dummy		0.185		1.062
		(0.412)		(1.322)
Constant	10.731	9.997	9.341	7.729
	(29.692)	(19.636)	(15.541)	(8.485)
Observations	64	64	64	64
R-squared	.478	.584	.270	.378
$\hat{\rho}_{hg}$	.643	.609	.643	.609

t-statistics in parentheses based on asymptotic variance estimates

Table 2 The reduced form estimates for the two specifications.

5. Fieller and Delta Constraint Shapes

In order to compare the Fieller and the Delta intervals we can plot the corresponding surface for a given value of α for each specification. Figures 9 and 10 show the plots of the constraint shapes of the two specifications for three values of α. In this example we find that the correlation is positive and approximately .6 between the reduced form coefficient estimates for both specifications.

In Figure 9 we note that the parameter estimate for the instrumental variable from the first stage regression specified by (45) has a t-statistic of -4.867 that implies an *F*-statistic of 23.6 thus indicating that the denominator of the ratio is significantly different from zero. This is shown by the confidence interval implied by the location of the both the Fieller and Delta constraint shapes away from the zero axis. Note that the constraint shapes for both the Delta and the Fieller are largest for $\alpha = .01$ and smallest for $\alpha = .10$. Since the lower bound for both the Fieller and Delta constraint shapes are above zero on the vertical axis we could reject the null hypothesis that $\beta = 0$ at the .01 level.

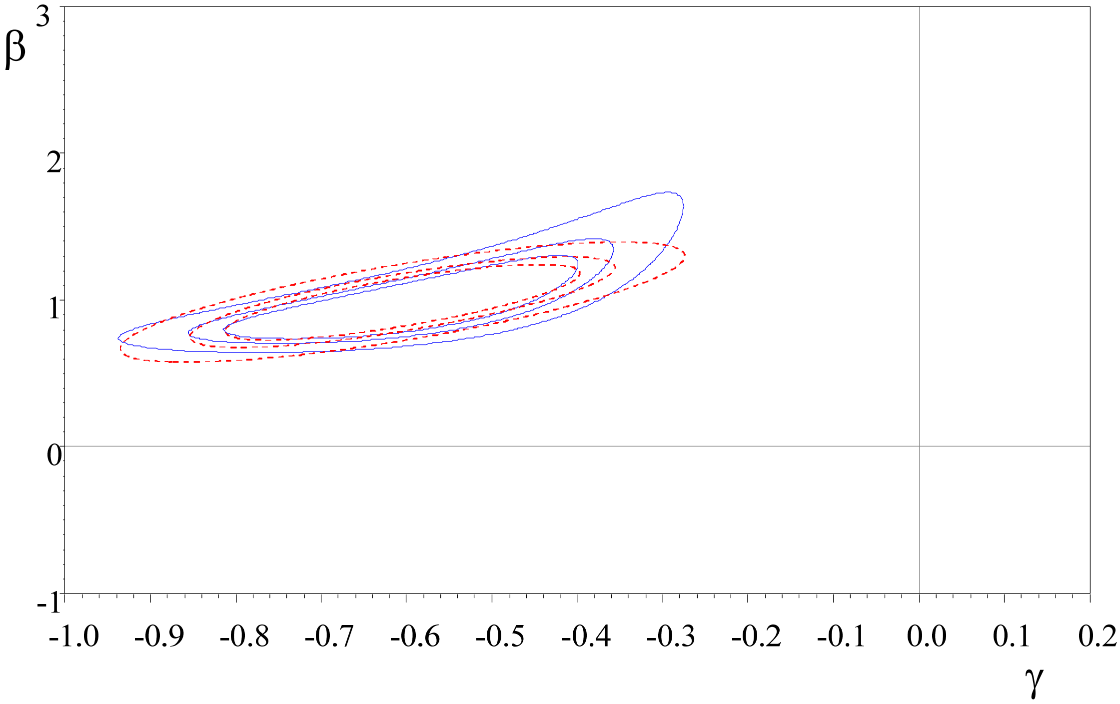


Figure 9. The Fieller (solid) and Delta (dotted) constraint shapes in (γ, β) space for varying values of $\alpha = .01$, .05 and .10 for specification 1.

Figure 10 plots the Fieller Boomerang and the Delta Ellipse constraint shapes for the results from specification (2) that includes covariates determined by the country's location. In this case the estimate of γ has a t-statistic of -1.953 which implies an F-statistic of 3.81 and thus indicates that there would be a greater difference between the Fieller and Delta intervals. In this case the Fieller constraint shape for $\alpha = .05$ is open to the north. Thus, when $\alpha \leq .05$ the lower bound is above the lower bound indicated by the Delta constraint ellipse while the upper bound for the Fieller constraint shape is not finite. This implies that there is sufficient evidence to conclude that the structural parameter (β) is greater than zero with a high degree of confidence and that the Fieller interval indicates that the parameter is further from zero than indicated by the equivalent Delta interval. In the case of specification (2) we find that the 95% interval ($\alpha = .05$) for the Fieller is $(.603, \infty)$ and the Delta is (.223, 1.99). This would indicate a much greater lower bound based on the Fieller. If we allow for a 90% ($\alpha = .10$) interval the Fieller bounds are (.660, 5.656) while the corresponding Delta bounds are (.369, 1.845).

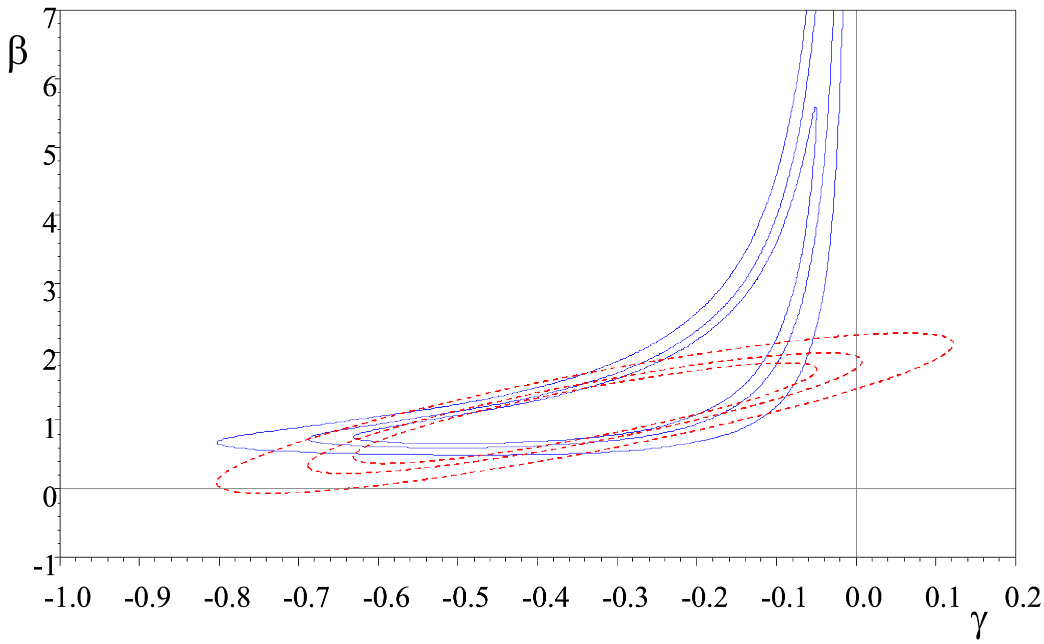


Figure 10. he Fieller (solid) and Delta (dotted) constraint shapes in (γ,β) space for varying values of α = .01, .05 and .10 for specification 2.

6 Previous Graphical Representations of the Fieller Interval

In this section we provide some alternative plots for this example that have been proposed to demonstrate the relationship between the Fieller and the Delta in the previous literature. They have been included to put the proposed plots in perspective. Note that although these plots provide a comparison of the two methods they do not indicate how the estimated parameters influence the comparison. First we present the line plot approach that does not provide any insight into the underlying covariance of the two parameter estimates but can be used to determine the nature of the Fieller solution when the roots of the implied quadratic are not real bounds. In the second we present the inferences that may be drawn with different values of α and can be used to define intervals when α may be greater than .05 or when an asymmetric rejection region is used.

The line plot method for the definition of the Fieller intervals is based on the interpretation of the Fieller method as the inversion of a test for a linear function of normally distributed random variables. Consider the linear combination l of the estimated parameters defined as:

$$l = h - b_0 g \tag{46}$$

where h and g are the estimates of the reduced form parameters θ and γ, and b_0 is a potential value for the 2SLS estimate of β in the exactly identified case. The value of the line l can be plotted as a function of b_0 along with a 100(1−α)% confidence interval. Thus when $l = 0$, $h - b_l g = 0$ and b_0 is equal to the 2SLS estimate for β. The confidence interval for this line is defined by:

$$CI\left(l \mid b_0\right) = h - b_0 g \pm t_{\alpha/2}\sqrt{s_h^2 - 2b_0 s_{gh} + s_g^2 b_0^2 g^2} \tag{47}$$

where $t_{\alpha/2}$ is the value from the t distribution with an α level of significance and $T - k$ degrees of freedom where k is the number of included covariates including the intercept term. The inverse confidence bounds for this estimate are defined by values of b_0 when the confidence interval for l cuts the zero axis as defined by the two possible values b_1 and b_2 where:

$$h - b_1 g - t_{\alpha/2}\sqrt{s_h^2 - 2b_0 s_{gh} + s_g^2 b_0^2 g^2} = 0 \text{ and } h - b_2 g + t_{\alpha/2}\sqrt{s_h^2 - 2b_0 s_{gh} + s_g^2 b_0^2 g^2} = 0 \tag{48}$$

which is equivalent to finding the roots of the quadratic equation defined by:

$$b_i^2\left(g^2 - t_{\alpha/2}^2 g s_g^2\right) - b_i\left(2g - t_{\alpha/2}^2 s_{gh}\right) + \left(h^2 - t_{\alpha/2}^2 s_h^2\right) = 0 \tag{49}$$

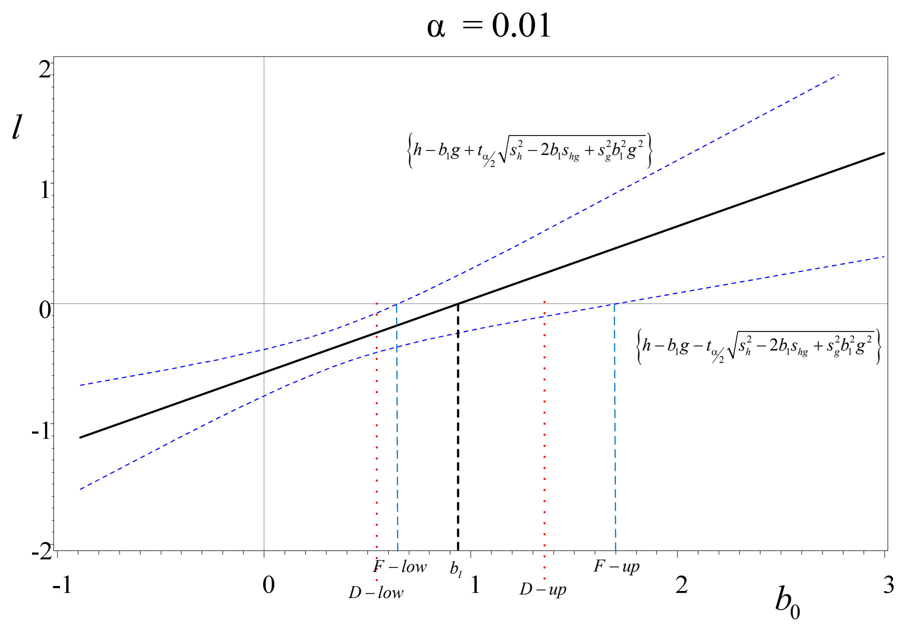


Figure 11. A line plot for specification (1) when α = .01 showing both the Fieller and Delta intervals .

By plotting the line defined by (46) along with the 100(1-α)% confidence interval as defined by (47), it is possible to obtain the Fieller interval by locating the values of b_0 where they cross the zero axis.[14] One useful feature of the line plot method is that it can always be used thus it can detect those circumstances when the Fieller interval does not provide finite confidence interval bounds and the real roots of (49) may not exist.

From Figure 11 one can locate the Fieller bounds from the points where the confidence interval for the line defined by $l = h - b_0 g$ cuts the zero axis line. Note that the Fieller upper bound is shown to be much larger than the Delta. It can also be seen from this diagram that if the α is smaller implying a greater value of $t_{\alpha/2}$ the lower confidence bound for l might not cut the zero axis at all. This would imply that upper limit of the confidence bound may not be present.

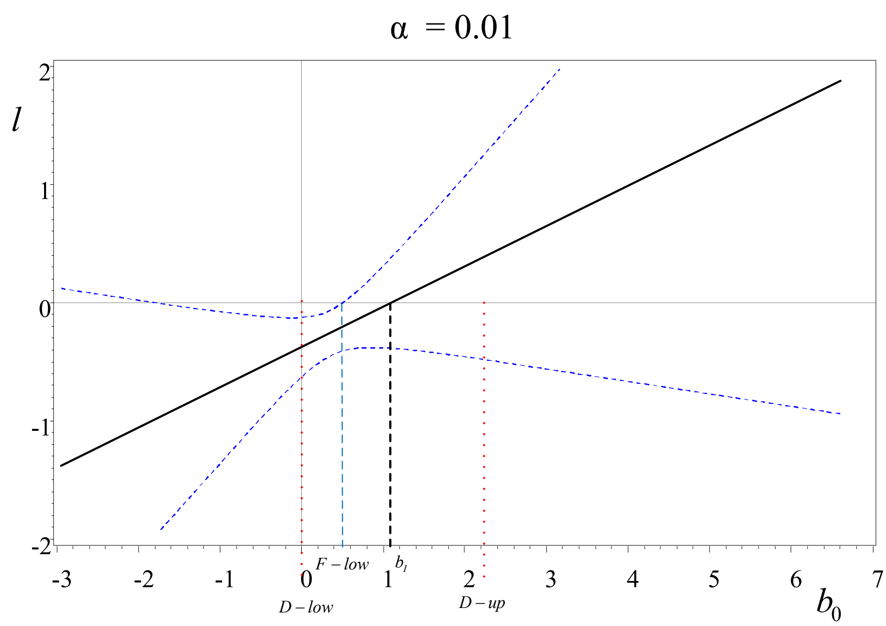


Figure 12. A line plot for specification (2) when α = .01 showing both the Fieller and Delta intervals.

The case where the lower bound is present but not an upper bound can be seen in Figure 12 where we examine the same interval with α = .01 for the results obtained using specification (2). Note that the lower bound is now defined by the second cross of the upper bound of the line and the first point where this bound crosses the zero line is not considered as a bound. Also the lower bound of the line never crosses the zero line thus indicating that there is no finite upper bound in this case. This is an example of the case where both roots of the quadratic that defines the Fieller

[14] See Hirschberg and Lye 2010a for more detail on this procedure when applied to the ratio of parameters from a single regression equation.

interval (49) are real but only one is meaningful. This plot displays the advantage of the line plot since it can be drawn in cases were the roots of (49) are not real. In that case neither bound of the line *l* will cross the zero line. These results imply that even in the presence of weak instruments we still may be able to make useful inferences. In this case the lower bound appears to be positive and different from 0 which suggests that institutions do matter for GDP. A similar conclusion was made by Chernozhukov and Hansen (2008).

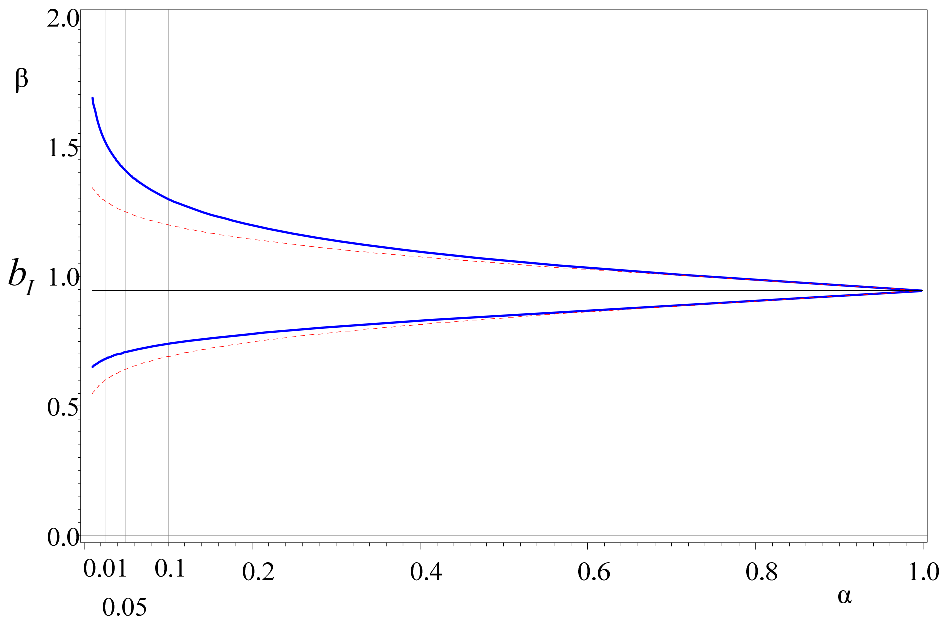


Figure 13 The confidence bounds for the Fieller (solid line) and Delta (dotted line) methods by value of α for Specification 1

Figures 13 and 14 are similar to the figures in Cook and Weisberg (1990) where the confidence intervals are given for both the Delta and Fieller methods by value of α. Note that the divergence of the two intervals becomes most pronounced at the lower values of α. Here it can be noted that the Fieller interval intervals for Specification (2) indicate that the upper bound is not defined for values of α greater than .1, however for both specifications we find that the lower bound is above zero. These plots can be used to define intervals where the rejection regions are not symmetric. This is pronounced in Figure 14 where a finite upper bound may only be defined for a larger rejection region.

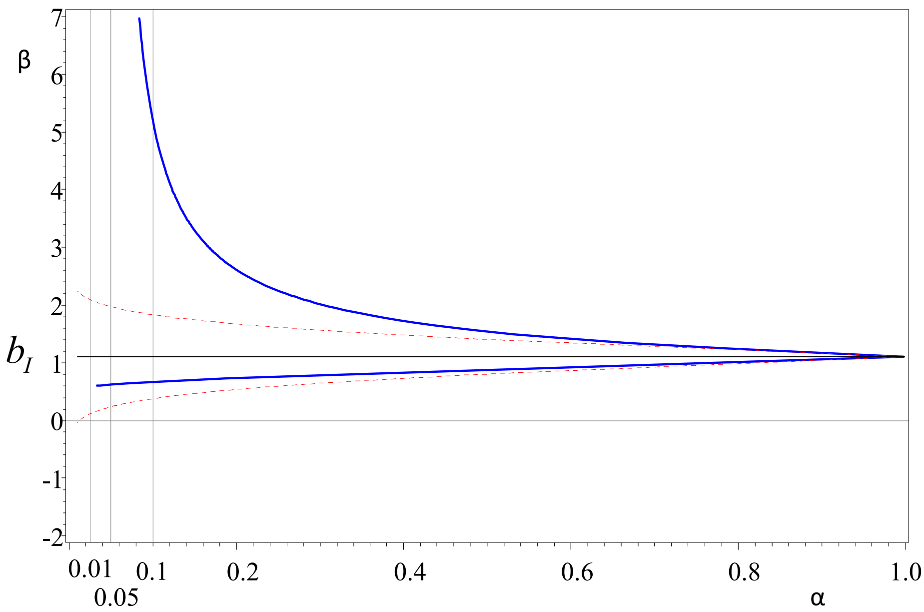


Figure 14 The confidence bounds for the Fieller (solid line) and Delta (dotted line) methods by value of α for Specification 2

7. Conclusions

We have demonstrated that the Delta and Fieller confidence intervals for structural equation parameter in the just identified simultaneous equations can be compared by the examination of two appropriately defined constraint shapes. By controlling for the levels of endogeneity and the relevance of the instrument, it is possible to demonstrate how these two characteristics influence the degree to which these estimated confidence intervals agree or diverge without the need for a simulation.

Our findings indicate that when the degree of endogeneity increases, the Fieller confidence interval becomes more asymmetric. Furthermore, this asymmetry becomes more pronounced when the instrument is found to be less relevant as measured by the *t*-statistic for the first stage regression. Thus the typical case of weak instruments, when the t-statistic in the first stage on the instrument is also low and the level of endogeneity is high, may result in cases where the Fieller and the Delta diverge the most. And in some cases, we find indications that the finite Fieller bound may be to the interior of the Delta bound while the other bound may not be finite.

We then examine an example of a just identified model in the presence of a weak instrument. In this case we found that the upper Fieller bound is infinite for α = .05 while the lower

Fieller bound is greater than the lower bound of Delta bound. In order to demonstrate the additional value of the constraint shapes presented here we also included two alternative methods for the graphic comparison of the Fieller and Delta bounds that have appeared in the literature.

References

Acemoglu, D., Johnson, S. and Robinson, A. (2001), "The Colonial Origins of Comparative Development: An Empirical Investigation", *American Economic Review*, 91, 1369-1401.

Anderson, T.W. (2005), "Origins of the limited information maximum likelihood and two-stage least squares estimators", *Journal of Econometrics*, 127, 1–16.

Anderson, T.W., Rubin, H. (1950), "The asymptotic properties of estimates of the parameters of a single equation in a complete system of stochastic equations", *Annals of Mathematical Statistics*, 21, 570–582.

Andrews, D. W. K. and Stock, J. H. (2006), "Inference with weak instruments", in ed. Blundell, R., Newey, W. K. and Persson, T., *Advances in Economics and Econometrics: Theory and Applications*, Ninth World Congress, Vol. 3, 122–173. Cambridge: Cambridge University Press.

Angrist, J.D. and Pischke, J., (2009), *Mostly harmless econometrics: an empiricist's companion*, Princeton : Princeton University Press.

Basmann, R. L. (1957), "A Generalized Classical Method of Linear Estimation of Coefficients in a Structural Equation," *Econometrica*, 25, 77-83.

Basmann, R.L. (1960), "An Expository Note on Estimation of Simultaneous Structural Equations", *Biometrics*, 16, 464-480.

Basmann, R.L. (1961), "A note on Finite Sample Distributions of Generalised Classical Linear Identifiability Test Statistics", *Journal of the American Statistical Association*, 16, 650-659.

Basmann, R.L. (1974), "Exact Finite Sample Frequency Functions of Generalised Classical Linear Estimators in Two Leading Over-Identified Cases", *Journal of the American Statistical Association*, 56, 650-659.

Basmann, R. L. (1993), "'Discovery' of two-stage least squares: the predicament of a living primary source", in ed R. Hébert, *Themes on Economic Discourse, Method, Money and Trade: selected papers from the history of economics*, 171-191, Aldershot, England: Edward Elgar.

Basmann, R. L., Richardson, D. H. and Rohr, R. J. (1974), "An Experimental Study of Structural

Estimators and Test Statistics Associated with Dynamical Econometric Models, *Econometrica*, 42, pp. 717-730.

Bernard, J.T., Idoudi, N., Khalaf, L. and C. Yélou (2005), “Finite Sample Inference Methods for Dynamic Energy Demand Models”, *Economics, Laval University,Working Paper* 2005-3.

Blomqvist, A. G. (1973), “Hypothesis Tests and Confidence Intervals for Steady-State Coefficients in Models with Lagged Dependent Variables: Some Notes on Fieller’s Method”, *Oxford Bulletin of Economics and Statistics*, 35, 69-74.

Casella, G., and Berger, R. L. (2002), *Statistical Inference* (2nd ed.), Pacific Grove, CA.: Duxbury.

Chernozhukov, V and Hansen, C. (2008), “The reduced form: A simple approach to inference with weak instruments”, *Economics Letters*, 68–71.

Cook, R. D., and Weisberg, S. (1990), “Confidence Curves in Nonlinear Regression”, *Journal of the American Statistical Association*, 85, 544-551.

Coplin, W. D, O’Leary, M. K. and Sealy, T. (1991), *A business guide to political risk for international decisions*, 2nd ed., New York: Political Risk Services.

Creasy, M. A. (1954), “Limits for the Ratio of Means”, *Journal of the Royal Statistical Society, Series B*, 16, 186-194.

Dufour, J.-M. (1997), “Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models”, *Econometrica,* 65, 1365-1387.

Dufour, J.-M. and Khalaf, L., (2001), Monte Carlo Test Methods in Econometrics, in ed. Badi Baltagi, *Companion to Theoretical Econometrics*, Chapter 23, 494-519, Blackwell, Oxford, U.K.

Durand, D. (1954), “Joint Confidence Regions for Multiple Regression Coefficients”, *Journal of the American Statistical Association*, 49, 130-146.

Fieller, E. C. (1932), “The Distribution of the Index in a Normal Bivariate Population”, *Biometrika*, 24, 428-440.

Fieller, E. C. (1940), “The Biological Standardization of Insulin”, supplement to the *Journal of the*

Royal Statistical Society, 7, 1-64.

Fieller, E. C. (1944), "A Fundamental Formula in the Statistics of Biological Assay, and Some Applications", *Quarterly Journal of Pharmacy and Pharmacology,* 17, 117-123.

Fieller, E. C. (1954), "Some Problems in Interval Estimation", *Journal of the Royal Statistical Society, Series B*, 16, 174-185.

Guiard, V. (1989), "Some Remarks on the Estimation of the Ratio of the Expected Values of a Two-dimensional Normal Random Variable (Correction of the Theorem of Milliken)", *Biometrical Journal*, 31, 681-697.

Hall, A. R., Rudebusch, G. D. and Wilcox, D. W., (1996), "Judging Instrument Relevance in Instrumental Variables Estimation", *International Economic Review*, 37, 283-298.

Hillier, G. H., Kinal, T. W., Srivastava, V. K. (1984), "On the moments of ordinary least squares and instrumental variables estimators in a general structural equation", *Econometrica* 52, 185–202.

Hirschberg, J. and Lye, J. (2004), "Inferences for the Extremum of Quadratic Regression Models", *Department of Economics*, *University of Melbourne,Working Paper* 906.

Hirschberg, J. and Lye, J. (2010a), "Two geometric representations of confidence intervals for ratios of linear combinations of regression parameters: An application to the NAIRU", *Economics Letters*, 108, 73-76.

Hirschberg, J. and Lye, J. (2010b), "A Geometric Comparison of the Delta and Fieller Confidence Intervals", *The American Statistician*, 64, 234-241.

Hirschberg, J. G., and J. N. Lye, (2010c), Notes on the Construction of Geometric Representations of Confidence Intervals of Ratios using Stata, Gauss and Eviews,*Supplementary Materials for: "*A Geometric Comparison of the Delta and Fieller Confidence Intervals", http://amstat.tandfonline.com/doi/suppl/10.1198/tast.2010.08130/suppl_file/utas_a_10713128_sm0001.zip

Hirschberg, J., Lye, J. and Slottje, D. (2008), "Inferential methods for elasticity estimates", *Journal of Econometrics*, 147, 299-315.

Kinal, T. W. (1980), "The Existence of Moments of *k*-Class Estimators", *Econometrica*, 48, 241-249.

Kiviet, J.F. (2013), "Identification and inference in a simultaneous equation under alternative information sets and sampling schemes", *The Econometrics Journal*, 16, S24–S59.

Leamer, E. E., (1978), Specification Searches: Ad Hoc Inference with Nonexperimental Data", John Wiley & Sons, New York, NY.

Lye, J. N. and J. G. Hirschberg, (2012), "Inverse Test Confidence Intervals for Turning-Points: A demonstration with Higher Order Polynomials.", Dek Terrell, Daniel Millimet, in (ed.) 30th Anniversary Edition (*Advances in Econometrics*, Volume 30), Emerald Group Publishing Limited, pp. 59 - 95

Mikulich, S., Zerbe, G., Jones, R. and Crowley, T. (2003), "Comparing Linear and Nonlinear Mixed Model Approaches to Cosinor Analysis", Statistics in Medicine, 22, 3195-3211.

Moreira, M.J., (2009), "Tests with correct size when instruments can be arbitrarily weak", *Journal of Econometrics*, 152, 131140.

Phillips, P. C. B. (1983), "Exact small sample theory in the simultaneous equations model", in: Griliches, Z., Intriligator, M. D., eds., *Handbook of Econometrics*, volume 1, chapter 8, North Holland, Amsterdam.

Polsky, D., Glick, H. A., Willke, R., and Schulman, K. (1997), "Confidence Intervals for Cost-Effectiveness Ratios: A Comparison of Four Methods", *Health Economics*, 6, 243-252.

Richardson, D.H. and Rohr, R.J. (1971), "Distribution of a Structural t-Statistic for the Case of Two Included Endogenous Variables", *Journal of the American Statistical Association*, 66, 375-382.

Scheffé, H. (1959), *The Analysis of Variance*, New York, NY: John Wiley and Sons.

Scheffé, H. (1970), "Multiple Testing versus Multiple Estimation. Improper Confidence Sets.

Estimation of Directions and Ratios", *The Annals of Mathematical Statistics*, 41, 1-29.

Staiger, D. and Stock, J. H. (1997), "Instrumental Variables Regressions with Weak Instruments", *Econometrica*, 65, 557-586.

Staiger, D., Stock, J. H. and Watson, M. (1997), "The NAIRU, Unemployment and Monetary Policy", *The Journal of Economic Perspectives*, Vol. 11, 33-49.

Stock, J. H. and Yogo, M. (2005), "Testing for weak instruments in linear IV regression", in ed. Andrews, D. W. K. and Stock, J. H., *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, 80–108, New York: Cambridge University Press.

Theil, H. (1953), "Repeated Least-Squares Applied to Complete Equation Systems," Mimeo. The Hague: Central Planning Bureau.

Valentine, T.J. (1979), "Hypothesis Tests and Confidence Intervals for Mean Elasticities Calculated from linear regression models", *Economics Letters*, 4, 363-367.

Von Luxburg, U., and Franz, V. (2004), "Confidence Sets for Ratios: A Purely Geometric Approach to Fieller's Theorem", *Technical Report N0. TR-133*, Max Planck Institute for Biological Cybernetics.

Zivot; E., Startz; R. and Nelson, C. R. (1998), "Valid Confidence Intervals and Inference in the Presence of Weak Instruments", *International Economic Review*, 39, 1119-1144.

Appendix A: Stata routine for application results providing the line plots.

Stata routine to estimate an example of 2SLS and demonstrate the correspondences between IV, 2SLS, Nonlinear SUR, and the Delta applied to the ratio of the parameter estimates. This routine also provides a plot equivalent to the line plot comparisons of the Fieller to the Delta intervals as shown in Figures 11 and 12. In addition this code demonstrates the equivalence of the estimation via non-linear SUR, 2SLS and Delta Method.

A.1 Stata code:

```
/*
       Example of the estimation of 2stage least squares and nonlinear system of
       equations providing the same answer for the Acemoglu, Johnson and Robinson
       (2001) AER paper.
*/
       use "C:\Users\josephh\Dropbox\submissions\JOE_Basmann\maketable4.dta", clear

       drop if baseco ~= 1

       ivregress 2sls logpgp95 africa asia lat_abst other_cont (avexpr = logem4)

       condivreg logpgp95 africa asia lat_abst other_cont (avexpr = logem4)

       nlsur (logpgp95 = {a11} + {beta}*{b12}* logem4 + {b13} * africa + ///
       {b14}*asia + {b15}*lat_abst + {b16}* other_cont) ///
       (avexpr = {a2} + {b12}*logem4 +  {b23} * africa + {b24}*asia + ///
       {b25}*lat_abst + {b26}* other_cont)


/*
       Use a SUR estimation of the reduced form equations that are equivalent to
       the 2SLS estimate.
*/

       sureg (logpgp95 avexpr  = logem4 lat_abst africa asia other_cont) /* specification 2 */
*      sureg (logpgp95 avexpr  = logem4 ) /* specification 1*/
 /* parsimonious specification based on the t_tests from the complete system estimated */
*      sureg (logpgp95  = logem4  africa asia ) (avexpr  = logem4 )
       estimates save surf, replace
       scalar b1 = _b[logpgp95:logem4]
       scalar b2 = _b[avexpr:logem4]
/*
       Compute the estimate of the parameter from the structural equation using
       the Delta Method on the ratio which is the same as 2SLS
*/
       nlcom (beta: _b[logpgp95:logem4]/_b[avexpr:logem4]), post
/*
       Save the estimate and the standard error so that we can construct the bounds for
       the estimation based on 6 times the se
       alfa is the offset for the 95% Delta interval line
*/
       scalar se = _se[beta]
       scalar ratio = _b[beta]
       scalar bottom = ratio - 6*se
```

```
       scalar in1 = (12*se) / 25
       scalar alfa =  1.96*se*b2

/*
       Compute the Fieller confidence interval using the line-plot approach
       by evaluating the linear combination at the points determined by the values
       of in1.  Then set up the data set meffect to plot the linear combination
       define the variables in this data set as me sme and Beta.
*/
       tempname meffect
       tempfile mefile
       postutil clear
       postfile meffect me sme Beta  using mefile, replace
/*
       loop over 25 points in the 12 times the Delta/2SLS standard error to plot
       the linear combination at each value of beta
*/

       forvalues i = 1/25    {
              scalar a1`i' = (`i'* in1) + bottom
              estimates use surf
              quietly nlcom (mee: _b[logpgp95:logem4] -_b[avexpr:logem4]*a1`i'  ), post
              scalar me = _b[mee]
              scalar sme = _se[mee]
              post meffect (me) (sme) (a1`i')
                                      }
       postclose meffect
/*
       Use the data set with the linear combination at each potential value of beta
*/
       use mefile, clear

       summarize

       gen up = me + 1.96 * sme
       gen lw = me - 1.96 * sme
       gen dup = me + alfa
       gen dlw = me - alfa
/*
       Plot the linear combination b1 - Beta*b2 and determine where it is equal to
       zero along with the 95% confidence interval for the Feiller interval
       Also plot the implied line for the Delta interval.
*/
       line me lw up dup dlw Beta, pstyle(p2 ci ci p2 p2 ) ///
       lwidth(thin medium medium thin thin ) ///
       lpattern(solid solid solid dash dash ) ///
       sort yscale(range(0 .05)) yline(0) ///
       saving(Fieller,replace) title(Fieller and Delta Intervals for Beta)
```

A.2 Results

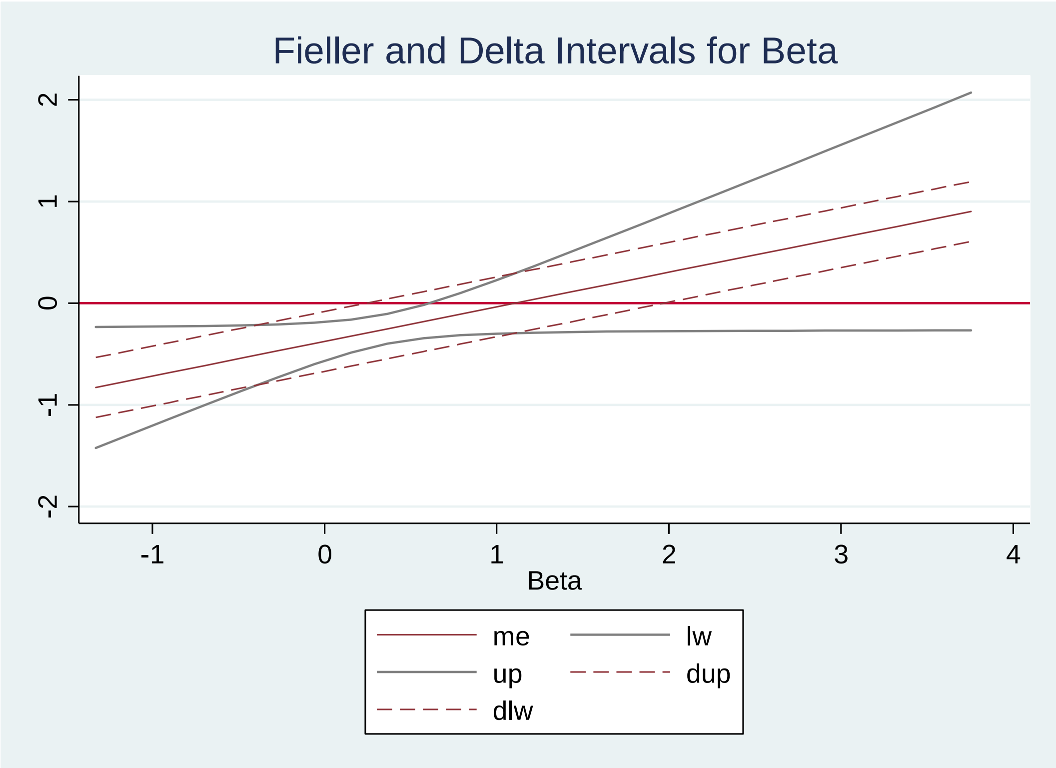
```
. do "C:\Users\josephh\AppData\Local\Temp\STD00000000.tmp"

.         use "C:\Users\josephh\Dropbox\submissions\JOE_Basmann\maketable4.dta", clear

.
.         drop if baseco ~= 1
(99 observations deleted)

.
.         ivregress 2sls logpgp95 africa asia lat_abst other_cont (avexpr = logem4)

Instrumental variables (2SLS) regression          Number of obs   =         64
                                                  Wald chi2(5)    =      37.78
                                                  Prob > chi2     =     0.0000
                                                  R-squared       =     0.0108
                                                  Root MSE        =     1.0296

------------------------------------------------------------------------------
    logpgp95 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      avexpr |   1.107077   .4413078     2.51   0.012     .2421296    1.972024
      africa |  -.4372669    .403835    -1.08   0.279    -1.228769    .3542352
        asia |  -1.047085    .499367    -2.10   0.036    -2.025826   -.0683438
    lat_abst |  -1.178178   1.671129    -0.71   0.481    -4.453531    2.097174
  other_cont |  -.9904014   .9500526    -1.04   0.297     -2.85247    .8716674
       _cons |   1.440453   2.703206     0.53   0.594    -3.857732    6.738639
------------------------------------------------------------------------------
Instrumented:  avexpr
Instruments:   africa asia lat_abst other_cont logem4

.
end of do-file

. do "C:\Users\josephh\AppData\Local\Temp\STD00000000.tmp"

.         condivreg logpgp95 africa asia lat_abst other_cont (avexpr = logem4)

Instrumental variables (2SLS) regression

First-stage results                                    Number of obs =      64
-----------------------                                F(  5,    58) =    6.85
F(  1,    58) =    3.46                                Prob > F      =  0.0000
Prob > F      =  0.0681                                R-squared     =  0.0108
R-squared     =  0.3277                                Adj R-squared = -0.0745
Adj R-squared =  0.2698                                Root MSE      =   1.082

------------------------------------------------------------------------------
    logpgp95 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      avexpr |   1.107077   .4635724     2.39   0.020      .179136    2.035018
      africa |  -.4372669   .4242091    -1.03   0.307    -1.286414    .4118798
        asia |  -1.047085   .5245608    -2.00   0.051    -2.097108    .0029374
    lat_abst |  -1.178178    1.75544    -0.67   0.505    -4.692073    2.335716
  other_cont |  -.9904014   .9979841    -0.99   0.325    -2.988084    1.007281
       _cons |   1.440453   2.839586     0.51   0.614    -4.243596    7.124502
------------------------------------------------------------------------------
Instrumented:  avexpr
Instruments:   africa asia lat_abst other_cont logem4
Confidence set and p-value for avexpr are based on normal approximation
------------------------------------------------------------------------------



------------------------------------------------------------------------------
                Coverage-corrected confidence set and p-value
                        for Ho: _b[avexpr] = 0
                   LIML estimate of _b[avexpr] = 1.107077

------------------------------------------------------------------------------
 Test                            Confidence Set                     p-value
------------------------------------------------------------------------------
 Conditional LR        (-inf, -11.27459] U [ .5898383, +inf)          0.0003
------------------------------------------------------------------------------

.
.         nlsur (logpgp95 = {a11} + {beta}*{b12}* logem4 + {b13} * africa + ///
```

```
>         {b14}*asia + {b15}*lat_abst + {b16}* other_cont) ///
>         (avexpr = {a2} + {b12}*logem4 +  {b23} * africa + {b24}*asia + ///
>         {b25}*lat_abst + {b26}* other_cont)
(obs = 64)

Calculating NLS estimates...
Iteration 0:  Residual SS =  126.5609
Iteration 1:  Residual SS =  119.8903
Iteration 2:  Residual SS =  119.8902
Iteration 3:  Residual SS =  119.8902
Calculating FGNLS estimates...
Iteration 0:  Scaled RSS =       128

FGNLS regression
-----------------------------------------------------------------------
       Equation |        Obs   Parms       RMSE      R-sq     Constant
----------------+------------------------------------------------------
 1     logpgp95 |         64       7   .6677511    0.5839          a11
 2       avexpr |         64       6   1.194735    0.3277           a2
-----------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        /a11 |   9.997378   .5091475    19.64   0.000     8.999467    10.99529
       /beta |   1.107077   .4413078     2.51   0.012     .2421296    1.972024
        /b12 |  -.3403178   .1742797    -1.95   0.051    -.6818998    .0012641
        /b13 |  -.7228463   .2179443    -3.32   0.001    -1.150009   -.2956833
        /b14 |  -.5245579    .267608    -1.96   0.050     -1.04906   -.0000559
        /b15 |   1.046296   .7400363     1.41   0.157    -.4041483    2.496741
        /b16 |   .1847984   .4489049     0.41   0.681    -.6950391    1.064636
         /a2 |   7.729295   .9109629     8.48   0.000     5.943841     9.51475
        /b23 |  -.2579581   .3899443    -0.66   0.508    -1.022235    .5063187
        /b24 |   .4719882   .4788021     0.99   0.324    -.4664467    1.410423
        /b25 |   2.009322   1.324067     1.52   0.129    -.5858021    4.604447
        /b26 |   1.061534   .8031773     1.32   0.186    -.5126647    2.635733
------------------------------------------------------------------------------

.
end of do-file

. do "C:\Users\josephh\AppData\Local\Temp\STD00000000.tmp"

. /*
>         Use a SUR estimation of the reduced form equations that are equivalent to
>         the 2SLS estimate.
> */
.
.         sureg (logpgp95 avexpr  = logem4 lat_abst africa asia other_cont) /* specification 2 */

Seemingly unrelated regression
----------------------------------------------------------------------
Equation          Obs  Parms        RMSE    "R-sq"       chi2        P
----------------------------------------------------------------------
logpgp95           64      5    .6677511    0.5839      89.81   0.0000
avexpr             64      5    1.194735    0.3277      31.20   0.0000
----------------------------------------------------------------------

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
logpgp95     |
      logem4 |  -.3767581   .0974069    -3.87   0.000    -.5676721    -.185844
    lat_abst |   1.046296   .7400363     1.41   0.157    -.4041483    2.496741
      africa |  -.7228463   .2179443    -3.32   0.001    -1.150009   -.2956833
        asia |  -.5245579    .267608    -1.96   0.050     -1.04906   -.0000559
  other_cont |   .1847984   .4489049     0.41   0.681    -.6950391    1.064636
       _cons |   9.997378   .5091475    19.64   0.000     8.999467    10.99529
-------------+----------------------------------------------------------------
avexpr       |
      logem4 |  -.3403178   .1742797    -1.95   0.051    -.6818998    .0012641
    lat_abst |   2.009322   1.324067     1.52   0.129    -.5858021    4.604447
      africa |  -.2579581   .3899443    -0.66   0.508    -1.022235    .5063187
        asia |   .4719882   .4788021     0.99   0.324    -.4664467    1.410423
  other_cont |   1.061534   .8031773     1.32   0.186    -.5126647    2.635733
       _cons |   7.729295   .9109629     8.48   0.000     5.943841     9.51475
------------------------------------------------------------------------------

. *       sureg (logpgp95 avexpr  = logem4 ) /* specification 1*/
.  /* parsimonious specification based on the t_tests from the complete system estimated */
```

```
. *       sureg (logpgp95  = logem4  africa asia ) (avexpr  = logem4 )
.         estimates save surf, replace
file surf.ster saved

.         scalar b1 = _b[logpgp95:logem4]

.         scalar b2 = _b[avexpr:logem4]

. 
end of do-file

. do "C:\Users\josephh\AppData\Local\Temp\STD00000000.tmp"

.         nlcom (beta: _b[logpgp95:logem4]/_b[avexpr:logem4]), post

        beta:  _b[logpgp95:logem4]/_b[avexpr:logem4]

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        beta |   1.107077   .4413078     2.51   0.012     .2421296    1.972024
------------------------------------------------------------------------------

. 
end of do-file

. do "C:\Users\josephh\AppData\Local\Temp\STD00000000.tmp"

.         scalar se = _se[beta]

.         scalar ratio = _b[beta]

.         scalar bottom = ratio - 6*se

.         scalar in1 = (12*se) / 25

.         scalar alfa =  1.96*se*b2

. 
end of do-file

. do "C:\Users\josephh\AppData\Local\Temp\STD00000000.tmp"

.         tempname meffect

.         tempfile mefile

.         postutil clear

.         postfile meffect me sme Beta  using mefile, replace

. /*
>         loop over 25 points in the 12 times the Delta/2SLS standard error to plot
>         the linear combination at each value of beta
> */
. 
.         forvalues i = 1/25      {
  2.                scalar a1`i' = (`i'* in1) + bottom
  3.                estimates use surf
  4.                quietly nlcom (mee: _b[logpgp95:logem4] -_b[avexpr:logem4]*a1`i'  ), post
  5.                scalar me = _b[mee]
  6.                scalar sme = _se[mee]
  7.                post meffect (me) (sme) (a1`i')
  8.                                        }

.         postclose meffect

. /*
>         Use the data set with the linear combination at each potential value of beta
> */
.         use mefile, clear

. 
.         summarize

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
          me |        25    .0360444    .5305589  -.8290207   .9011095
         sme |        25    .2724874    .1591923   .0746308   .5964864
        Beta |        25    1.212991     1.55901  -1.328942   3.754924
```

```
.
.         gen up = me + 1.96 * sme

.         gen lw = me - 1.96 * sme

.         gen dup = me + alfa

.         gen dlw = me - alfa

. /*
>         Plot the linear combination b1 - Beta*b2 and determine where it is equal to
>         zero along with the 95% confidence interval for the Feiller interval
>         Also plot the implied line for the Delta interval.
> */
.         line me lw up dup dlw Beta, pstyle(p2 ci ci p2 p2 ) ///
>         lwidth(thin medium medium thin thin ) ///
>         lpattern(solid solid solid dash dash ) ///
>         sort yscale(range(0 .05)) yline(0) ///
>         saving(Fieller,replace) title(Fieller and Delta Intervals for Beta)
(file Fieller.gph saved)

.
end of do-file
```

A.3 Resulting Plot



.

Appendix B: SAS Proc IML routine for simulations

This routine was written in Proc IML in the SAS computer language to simulate the intervals from the paper by Zivot Startz and Nelson, IER 1998 as plotted in the Ellipse and other intervals.

```
*
      Program to simulate the results from Zivot Startz and Nelson, IER 1998.
;
      proc iml;
*==================================================================================================;
      start Fieller_old( b1, b2, cov,t);     ;
      t2 = t**2 ; ratio = b1/b2 ;
      denom = ((t2*cov[2,2])-(b2*b2) );
      p = j(3,1);
      p[1] = (b2**2 - t2*cov[2,2]);
      p[2] = -2*(b1*b2- t2*cov[1,2]);
      p[3] = b1**2 - t2*cov[1,1];
      r = polyroot(p);
      up1 = max (r[1,1],r[2,1] ) ; lw1 = min (r[1,1],r[2,1] ) ; return( lw1 || up1 ) ;
      if r[1,2] > 0 then do ; up = -99999 ; lw = -99999 ; return( lw || up ); end;
      if denom > 0 then do ; return( lw1 || up1 ); end;
      d1 = abs(up1 - ratio) ; d2 = abs(lw1-ratio) ; * print d1 d2 denom lw1 up1 ratio ;
      if ( d1 > d2 ) & (up1 < ratio ) then do; lw = up1 ; up = -99999 ; return (lw || up) ; end;
      if ( d1 > d2 ) & (up1 > ratio ) then do; lw = -99999 ; up = up1 ; return (lw || up) ; end;
      if ( d1 < d2 ) & (lw1 < ratio ) then do; lw = lw1 ; up = -99999 ; return (lw || up) ; end;
      if ( d1 < d2 ) & (lw1 > ratio ) then do; lw = -99999 ; up = lw1 ; return (lw || up) ; end;
      if (d1 = 0) & (d2 = 0) then do ; up = -99999 ; lw = -99999 ; return( lw || up ); end;

      finish;
*==================================================================================================;
*==================================================================================================;
      start Fieller( b1, b2, cov,t); * with check for complimentary roots ;
      ratio = b1 / b2 ;
      t2 = t**2 ; up = -99998 ; lw = -99999 ;
      rad = (-t2*cov[1,1]*cov[2,2])-(2*b1*b2*cov[1,2])  + (t2*cov[1,2])+ (b2*b2*cov[1,1]) +
      (b1*b1*cov[2,2]) ;

        if (rad <= 0) then return( lw || up) ; * reject the imaginary roots ;

      denom = ((t2*cov[2,2])-(b2*b2) ); p2 = t2*cov[1,2]-b1*b2 ;
      p1 = t*sqrt(rad) ;
      lmx1 = ( p2 + p1 ) / denom;
      lmx2 = ( p2 - p1 ) / denom;
      lm1 = max(lmx1,lmx2) ; lm2 = min(lmx1,lmx2) ; *print lm2 ratio lm1 denom ;

      if (lm2 <= ratio <= lm1 )  then return( (lm2 || lm1) ) ;

      d1 = abs(lm1 - ratio) ; d2 = abs(lm2-ratio) ; * compute distance to the ratio ;

      if (d1 < d2) & ( lm1 < ratio) then do ; lw = lm1 ; up = -99998 ; end ;
      if (d1 < d2) & ( lm1 > ratio) then do ; up = lm1 ; lw = -99999 ; end ;
      if (d1 > d2) & ( lm2 < ratio) then do ; lw = lm2 ; up = -99998 ; end ;
      if (d1 > d2) & ( lm2 > ratio) then do ; up = lm2 ; lw = -99999 ; end ;

      return ((lw || up) ) ;

      finish;
*==================================================================================================;
```

```
*==========================================================================================;
      start Delta( b1, b2, cov, t);   ;
      ratio = b1/b2 ;
      r1 = (1/b2) || (-b1/(b2##2)) ;
      se = sqrt(r1 * cov * r1`) ;
      lw_up = ( ratio - t # se) || ( ratio + t # se) ;
      return( lw_up );
      finish;
*==========================================================================================;
*==========================================================================================;
      /** H_Bisection: find root on bracketing interval [a,b] for Hinkley CI when critical value =
      limit1.
      dy = limit on probability dx = accuracy of the limiting value
     **/
      start H_bisection(a, b, limit1, b1, b2, sss);
      dx = 1e-6; dy = 1e-6;

        do i1 = 1 to 100; /** max iterations **/
        c = (a+b)/2;
        fc = func_hinkley(c, b1, b2, sss) ;
               if abs( fc - limit1 )  < dy | (b-a)/2 < dx then  return(c);
        fa = func_hinkley(a, b1, b2, sss) ;
        if ( fa - limit1 ) # ( fc - limit1 )  > 0 then a = c;
                                                                  else b = c;
        end;
      print "H_bisection - Max iterations" a c b fa fc ;
      return (.); /** no convergence **/
      finish;

*==========================================================================================;
*==========================================================================================;
      /**
        use the Hinkley function
      */
      start func_hinkley(j1, b1, b2, sss ) ;
      s1 = sqrt(sss[1,1]) ; s2 = sqrt(sss[2,2]) ; s12 = sss[1,2] ; rh = s12/(s1#s2) ;
      aa = sqrt(abs( (j1##2/s1##2) - (2#j1#rh/(s1#s2)) + (1/s2##2)  ) );
      h = ((b2#j1)-b1)/(s1#s2#aa); k = b2/s2 ; r = (s2#j1 - rh#s1)/(s1#s2#aa);
      h = round(h,1e-6) ; k = round(k,1e-6) ;
      if abs(r) >= 1 then r = .99 *(((r < 0)*-1)+(r >= 0))  ;
      if (abs(h) < 1.e-6) | (h = 0 ) then do;* print h k r "hinkley" ;
      if (abs(h) < 1.e-6) | (h = 0 ) then h = .001 * (((h < 0)*-1)+(h >= 0)) ; * stay away from zero
      ;
      if (abs(k) < 1.e-6) | (k = 0 ) then k = .001 * (((k < 0)*-1)+(k >= 0)) ; * print aa h k r
      "hinkley fixed" ; end;
      if ( probnorm(h) > .00001 ) | ( probnorm(k) > .00001 ) then l1 =  probbnrm(h,k,r); else l1 = 0
      ;
      mh = -h ; mk = -k ;
      if (abs(mh) < 1.e-6) | (mh = 0 ) then do; * print mh mk r "hinkley" ;
      if (abs(mh) < 1.e-6) | (mh = 0 )  then mh = .001 * (((mh < 0)*-1)+(mh >= 0)) ; * stay away
      from zero ;
      if (abs(mk) < 1.e-6) | (mk = 0 )  then mk = .001 * (((mk < 0)*-1)+(mk >= 0))  ; *  print aa mh
      mk r "hinkley fixed" ; end;

      if ( probnorm(mh)> .00001 ) | ( probnorm(mk) > .00001 ) then l2 =  probbnrm(mh,mk,r); else l2
      = 0 ;

      fdist2 = ( l1 + l2 );  * Hinkley prob at q[ii] ;
      return(fdist2) ;
      finish;

*==========================================================================================;
```

```
*=========================================================================================;
       /** F_Bisection: find root on bracketing interval [a,b]. for Fieller CI when critical value
      = limit1.
       dy = limit on probability dx = accuracy of the limiting value
     **/
      start f_bisection(a, b, limit1,  b1, b2, sss);
      dx = 1e-6; dy = 1e-6; n_its = 100 ;

       do i1 = 1 to n_its; /** max iterations **/
       c = (a+b)/2;
       fc = func_fieller1(c, b1, b2, sss) ;
              if (abs( fc - limit1 ) < dy) | ((b-a)/2 < dx)  then return(c);
       fa = func_fieller1(a, b1, b2, sss) ;
       if (fa - limit1 )#(fc-limit1) > 0 then a = c;
                                                       else b = c;
       end;
      print "F_bisection - Max iterations" a c b fa fc ;
      return (.); /** no convergence **/
      finish;

*=========================================================================================;
*=========================================================================================;
      /**
       Estimate Fieller
      */
      start func_fieller(j1, b1, b2, sss) ;
      s1 = sqrt(sss[1,1]) ; s2 = sqrt(sss[2,2]) ; s12 = sss[1,2] ; rh = s12/(s1#s2) ;
      aa = sqrt(abs( (j1##2/s1##2) - (2#j1#rh/(s1#s2)) + (1/s2##2)  ) );
      h = ((b2#j1)-b1)/(s1#s2#aa); k = b2/s2 ; r = (s2#j1 - rh#s1)/(s1#s2#aa);
      h = round(h,1e-8) ; k = round(k,1e-8) ;
      if abs(r) >= 1 then r = .99 *(((r < 0)*-1)+(r >= 0))  ;
      if (abs(h) < 1.e-6) | (h = 0 ) then do; *print h k r "Fieller" ;
      if (abs(h) < 1.e-6) | (h = 0 ) then h = .001 * (((h < 0)*-1)+(h >= 0)) ; * stay away from zero
      ;
      if (abs(k) < 1.e-6) | (k = 0 ) then k = .001 * (((k < 0)*-1)+(k >= 0)) ;  *print aa h k r
      "Fieller fixed" ; end;
      if ( probnorm(h) > .00001 ) | ( probnorm(k) > .00001 ) then fdist1 =  probbnrm(h,k,r);
      return(fdist1) ;
      finish;

*=========================================================================================;
*=========================================================================================;
      /**
       Estimate Fieller using the prob normal for a single parameter estimate
       inversion method using the linear combination approach
      */
      start func_fieller1(j1, b1, b2, sss) ;
      mu = b1 - j1*b2; tmp = (1 || -1#j1) ; var = tmp * sss * tmp` ; sig = sqrt(var);
      w = mu/sig;
      fdist = max(1-probnorm(w),0);  * Fieller prob at q[ii];
      return(fdist) ;
      finish;

*=========================================================================================;
*=========================================================================================;
      start DFH( b1, b2, sss,n, alpha = .05, pi1, rho1, beta1, flagzz,iteration);

*  Flags for finding limiting values and set starting and increments for cdfs;

      lw_prob = alpha/2 ; up_prob = 1 - (alpha/2) ; tpq = b1/b2;
      df = n ;
*     t = tinv(up_prob,df); t = probit(up_prob) ; * use probit for normal ;

      fg1 = 0 ; fg2 = 0  ; fg3 = 0 ; fg4 = 0 ; fg5 = 0 ; fg6 = 0 ;
```

```
      n_int = 100 ;  * evaluation points ;

      s1 = sqrt(sss[1,1]) ; s2 = sqrt(sss[2,2]) ; s12 = sss[1,2] ; rh = s12/sqrt(s1#s2) ;
*
      Define limits
;
      dlimits = Delta( b1, b2, sss,2); * get the Delta limits for t = 2 ;

      r1 = (1/b2) || (-b1/b2##2) ;
      d_se = sqrt(r1 * sss * r1`) ;

      d_range = dlimits[1,2] - dlimits[1,1] ;
      inc = 4#(d_range)/ n_int ;
      start = dlimits[1,1] - (d_range) ; * Use the minimum of  Delta minus half the range of Delta;

      q = start + (1:n_int)`# inc ; * define the points for the evaluation based on increment and
      start.;

*
      cycle over values to determine the cdf based on the Delta, Fieller and Hinkley distributions
      see Koti JOurnal of BIOpharmaceutical Statistics 2007
;
      lw1 = -99999 ; lw2 = -99999 ; up1 = -99999 ; up2 = -99999 ; med1 = -99999 ; med2 = -99999 ;
      exp_val = 0 ; wgt_val = 0 ; med = .5 ;
      lw_prob = alpha/2 ; up_prob = 1 - (alpha/2) ;

        do ii = 1 to n_int;

        fdist = func_fieller1(q[ii], b1, b2, sss) ;  * Fieller use the traditional inverse method;
        ddist = probnorm((q[ii]-tpq )/d_se) ; * Delta prob at q[ii] ;
        fdist1 = func_fieller(q[ii], b1, b2, sss) ;  * Fieller using bivariate normal ;
        fdist2 = func_Hinkley(q[ii], b1, b2, sss) ;  * Hinkley prob at q[ii] ;
*
      Locate the 1-alpha CI and median for Hinkley distribution
;

        rt =  fdist2 || ddist || pi1|| rho1 || beta1 || q[ii] ||  fdist || fdist1 ;
        if ncol(rt) = 8 then res2 = res2 // rt ;

        if ii > 1 then do ;

               if fdist1 > .5 & fg1 = 0 then do ; fg1 = 1 ; med1 = f_bisection(q[ii-1],q[ii],med ,
      b1, b2, sss); end;
               if fdist1 > lw_prob & fg2 = 0 then do ; fg2 = 1 ; lw1 = f_bisection(q[ii-
      1],q[ii],lw_prob , b1, b2, sss); end;
               if fdist1 > up_prob & fg3 = 0 then do ; fg3 = 1 ; up1 = f_bisection(q[ii-
      1],q[ii],up_prob , b1, b2, sss); end;

               if fdist2 > .5 & fg4 = 0 then do ; fg4 = 1 ; med2 = h_bisection(q[ii-1],q[ii],med ,
      b1, b2, sss); end;
               if fdist2 > lw_prob & fg5 = 0 then do ; fg5 = 1 ; lw2 = h_bisection(q[ii-
      1],q[ii],lw_prob , b1, b2, sss); end;
               if fdist2 > up_prob & fg6 = 0 then do ; fg6 = 1 ; up2 = h_bisection(q[ii-
      1],q[ii],up_prob , b1, b2, sss); end;

               exp_val = exp_val +  (q[ii]-(inc/2))#(fdist2-lfdist2)/ inc ;
               wgt_val = wgt_val +  (fdist2-lfdist2)/ inc ;
               end;

        lfdist2 = fdist2 ;

        end;

        if wgt_val > 0 then exp_val = exp_val / wgt_val ; else exp_val = -99999 ;
```

```
       rnames = {"hdist" "ddist" "pi" "rho" "beta" "q" "fdist" "fdist1" } ;
       if flagzz = 1 then do ; create dens from res2[colname=rnames] ; end ;
       if iteration = 1 then do; setout dens ; * append from res2 ; end ;
       *print res2[colname = rnames] ; free res2 ;


       return( (lw1 || up1)||(lw2 || up2)||med2 ||exp_val);
       finish;
*==========================================================================================;
*==========================================================================================;
*
      routine to generate points to plot the line graph for differen confidence intervals
;
      start FP( b1, b2, sss,n, alpha = .05, pi1, rho1, beta1, flagyy , iteration);

*  Flags for finding limiting values and set starting and increments for cdfs;

      lw_prob = alpha/2 ; up_prob = 1 - (alpha/2) ; tpq = b1/b2;
      df = n ;
*     t = tinv(up_prob,df); t = probit(up_prob) ; * use probit for normal ;

      fg1 = 0 ; fg2 = 0  ; fg3 = 0 ; fg4 = 0 ; fg5 = 0 ; fg6 = 0 ;
      n_int = 40 ;  * evaluation points ;

      s1 = sqrt(sss[1,1]) ; s2 = sqrt(sss[2,2]) ; s12 = sss[1,2] ; rh = s12/sqrt(s1#s2) ;
*
      Define limits based on the Delta limits times 4
;
      dlimits = Delta( b1, b2, sss,2); * get the Delta limits for t = 2 ;

      r1 = (1/b2) || (-b1/b2##2) ;

      d_se = sqrt(r1 * sss * r1`) ;

      d_range = dlimits[1,2] - dlimits[1,1] ;

      inc = 4#(d_range)/ n_int ;
      start = dlimits[1,1] - (d_range) ; * Use the minimum of  Delta minus half the range of Delta;

      q = start + (1:n_int)`# inc ; * define the points for the evaluation based on increment and
      start.;

*
      cycle over values to plot predicted linear relationships
;
      pvs = { .70 .85 .90 .95 .975 .995 } ;
      do jj = 1 to ncol(pvs) ;

      p_value = 2#(1-pvs[1,jj]) ;
      t = probit(pvs[1,jj]) ; * use probit for normal ;
      dlim = Delta( b1, b2, sss,t);
      d_range = 3#(dlim[1,2] - dlim[1,1]) ;
      inc = (d_range)/ (n_int-1) ;
      start = tpq - (d_range/2) ; * Use the minimum of  Delta minus half the range of Delta;
      q = start + (1:(n_int-1))`# inc ; * define the points for the evaluation based on increment
      and start.;
      q = q // (b1/b2) ;

       do ii = 1 to n_int;

       lc = ((1)||(-1#q[ii]));
       cv = lc*sss*lc` ;
       test = b1 - q[ii]*b2 ;
```

```
        up = test + t * sqrt(cv) ;
        lw = test - t * sqrt(cv) ;
        f_stat = (b1*b1)/sss[2,2] ;
        rt = ( pi1||f_stat || rho1|| beta1|| p_value || t || q[ii] || up || test|| lw || dlim || cv
      ) ;
        res = res // rt ;


        end;
        end ;


        rnames = { "pi" "f_stat" "rho" "beta" "p_value" "t" "q" "up" "test" "lw" "dlw" "dup" "var"}
      ;
        if flagyy = 1 then do ; create lineplot from res[colname=rnames] ; end ;
        if iteration = 1 then do; setout lineplot ;  append from res ; end ;
        *print res2[colname = rnames] ; free res ;

        return((nrow(res)));
        finish;
*==================================================================================================;
*
      Set sample size values of gamma and beta,
      correlation between the errors in each equation is r1
      v1 and v2 are corresponding variances for each equations error
;
*     Define sample size (n), num of parameters (k), variance of u (v_u) variance of v (v_v) ;
      N = 100;  k = 1 ;  v_u = 1 ; v_v = 1 ;

      pis = { .5 .7 .8 .9 .95} ; * values of pi ;
      *pis = { .1 .5  1} ; * values of pi ;
      *rhos = { -.99 -.5 0 .1 .5 .9 .99 } ;
      rhos = {.5 .7 .8 .9 .99 } ;
      *betas = {-1 0 1} ;
      betas = { 1} ;

      flag = 1 ; zflag = 1 ; flagzz = 1 ; flagxx = 1 ;  flagyy = 1 ;
      call randseed(4325);
*
      Define the instrumental variable
;
      z = rannor(j(n,k,989775)) ;  * z = ranuni(j(n,k,898765)) - .5;
      z = z # sqrt(n /ssq(z)) ; * normalize the sum of squares of z to be = n ;
      zpz = z`*z ; izpz = inv(z`*z) ; print "initial values " n k v_u v_v izpz zpz pis rhos betas;

*IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII;

      do iteration = 1 to 1 ; * use for simulations ;

      do beta_ii = 1 to ncol(betas) ; beta = betas[1,beta_ii] ;
      do pi_ii = 1 to ncol(pis) ; pi = pis[1,pi_ii]  ;
      do rho_ii = 1 to ncol(rhos) ; rho = rhos[1,rho_ii] ;
      theta = beta # pi ;
*
      Define the covariance matrix based on theoretical values
;

      tss = ( (beta##2 + 2#rho#beta +1 ) || (beta + rho ) ) //
         ( (beta + rho )              || (1)           )  ;
      tss = tss / n ;

*     set up the distribution of the errors ;
      c_uv = rho # sqrt(v_v#v_u) ;
      Mean = {0, 0};  Cov =  ( v_u || c_uv )//(c_uv || v_v ) ;
```

```
*     generate normal random deviates based on mean and cov and check correlation ;

      X = RandNormal(N, Mean, Cov);
      corr = corr(x);  * print  corr;

      u = x[,1] ; v = x[,2] ;
*
      Generate y1 and the endogenous regressor y2
;
      y2 = pi * z + v ;
      y1 = beta * y2 + u ;


/*    new = j(n,1,iteration) || j(n,1,pi) || j(n,1,rho) || y2 || y1 || z || u || v ;
      new_names = { "iteration" "pi" "rho" "y2" "y1" "z" "u" "v"} ;
      if zflag = 1 then do ; create new from new[colname=new_names] ; zflag = 0 ;  end ;
      setout new ;  append from new ; free new ;
*/
*
      Check the correlation between the endogenous variable and the instrument,
      the instrument and the error in structural equation and the correlation between y1 and y2
;
      corr1 = (corr(y2||z)) [1,2] ; corr2 = (corr(u||z)) [1,2] ; corr3 = (corr(y1||y2)) [1,2] ;

*     print "corr(y2,z) = " corr1 "corr(e1,z) = " corr2 "corr(y2,y1) = " corr2 ;
*
      estimate the equations separately
;

*     reduced form equation for y1 ;
      cf1 = izpz*z`*y1 ; py1 = z*cf1 ; sig1 = ssq(y1-py1)/(n-k) ; se_c1 = sqrt(sig1#izpz) ;
      t_c1 = cf1/se_c1 ; Rs_1 = (corr(y1||py1)##2) [1,2]; *print cf1 sig1 se_c1 t_c1 rs_1 ;
*     reduced form equation for y1 ;
      cf2 = izpz*z`*y2 ; py2 = z*cf2 ; sig2 = ssq(y2-py2)/(n-k) ; se_c2 = sqrt(sig2#izpz) ;
      t_c2 = cf2/se_c2 ; Rs_2 = (corr(y2||py2)##2) [1,2]; *print cf2 sig2 se_c2 t_c2 rs_2;

      sig12 = ((y1-py1)`*(y2-py2)) / (n-k) ; *Sur covariance of the errors ;
      cov_red = ((sig1 ||sig12)//(sig12 ||sig2) );
      cov2sur = izpz # cov_red ; corr_red = sig12/sqrt(sig1#sig2) ;

      *print beta pi rho tss cov2sur (100#(cov2sur - tss )/ tss) ;


*
      Intervals using Fieller, Delta and Hinkley
;
*
      Use theoretical values and generate the confidence bound limits by values of
      1-alpha
;
      cf1 = theta ; cf2 = pi ; cov2sur = tss ;  print cf1 cf2 cov2sur rho pi beta ;

      do a_value = .01 to .99 by .005 ;
        up_prob = 1 - (a_value/2) ; t = probit(up_prob) ; * use probit for normal ;

        f_int =  Fieller_old( cf1, cf2, cov2sur  ,t) ;
        f_int1 =  Fieller( cf1, cf2, cov2sur  ,t) ;
        d_int = Delta( cf1, cf2, cov2sur , t);
        h_int = DFH( cf1 , cf2, cov2sur , n, a_value, pi, rho, beta, flagzz,iteration)  ;
        if flagzz = 1 then flagzz = 0 ; * reset flag after the first time ;
        med_hink = h_int[1,5] ; exp_hink = h_int[1,6] ;
        h_int1 = h_int[1,3:4] ; f_int2 = h_int[1,1:2] ;
        r_new = beta|| pi || rho || cf1 ||cf2 || a_value ||t || f_int || d_int || h_int1 || f_int2
      || f_int1 ;
```

```
       rn_names = {"beta" "pi" "rho" "cf1" "cf2" "alpha" "t_value"  "fint1" "fint2" "dint1" "dint2"
      "hint1"
                              "hint2" "fint1a" "fint2a" "fint1b" "fint2b" } ;

       if flagxx = 1  then do; create r_new from r_new[colname=rn_names] ; flagxx = 0 ; end;
       setout r_new ; append from r_new ;
      end ;

      p_int = FP( cf1 , cf2, cov2sur , n, .02, pi, rho, beta, flagyy,iteration)  ;
      if flagyy = 1 then flagyy = 0 ; * reset flag after the first time ;
*
       Estimate Ordinary least squares of inconsistent estimates
;
      beta_ols = inv(y2`*y2)*Y2`*y1 ;
*
       Estimate 2stage least squares
;
      inv1 = inv((y2`*z)*izpz*(z`*y2));
      beta_2s = inv1 * ((y2`*z)*izpz*(z`*y1)) ;
      py1_2s = y2*beta_2s ; sigma2s = ssq(y1-py1_2s) /( n-1) ;
      cov_2s = sigma2s * inv1 ; Rs_2s = (corr(y1||py1_2s)##2) [1,2];
*     print beta_2s cov_2s rs_2s ;
      t_int = (beta_2s - 1.96#sqrt(cov_2s[1,1])) || (beta_2s + 1.96#sqrt(cov_2s[1,1])) ;
*     print h_int f_int d_int t_int ;
*
      Compute the r-squares for the two equations
;
       Rs1 = (corr(y1||py1)##2) [1,2];
       rs2 = (corr(y2||py2)##2) [1,2];

*
      Compute the t-stats for each equation
;

*      print "reduced form coefficients = "(cf1//cf2) "t_stats = "(t_c1//t_c2)  ;
*
      estiamte beta as the ratio of the reduced form estimated parameters
;
      e_beta = cf1/cf2 ;
      * print (r_sqrs#N) rs_2s  beta beta_2s e_beta med_hink exp_hink ;

      sims = sims // ( pi || cf1 ||cf2 || t_c1 || t_c2|| sig1 || sig2 || sig12|| corr_red ||  rho ||
      h_int || f_int || d_int ||
                       t_int || (rs2#N) || rs1 || rs2|| beta || beta_2s || e_beta || beta_ols ||
      med_hink || exp_hink) ;

      end ;
      end;
      end ;
      cnames2 = {"pi" "cf1" "cf2" "t_c1" "t_c2" "sig1" "sig2" "sig12" "corr_red" "rho" "hint1"
                 "hint2" "fint1" "fint2" "dint1" "dint2" "tint1" "tint2" "r_sqrstN" "rs1" "rs2"
                  "beta_set" "beta_2s" "beta_ratio" "beta_ols" "med_hink" "exp_hink"} ;

      if flag = 1 then do ; create sims from sims[colname=cnames2] ; flag = 0 ;
      print sims[colname=cnames2] ; end ;
      setout sims ; append from sims ; free sims ;
end ;

*IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII;

quit;
run;
run;
*
```

```
        Use the results of the simulation to plot the Fieller figures
;
        proc sort data=r_new ; by beta pi rho ; run;

        proc print data=sims ; run;

        proc sort data=lineplot ; by  beta pi rho p_value q ;

        data lineplot1 ; set lineplot ; by beta pi rho  p_value ;
        d_inc = dup - beta ;
        d_up = test + d_inc * pi ;
        d_lw = test - d_inc * pi ;
        first =  first.p_value; run ;

        proc print data=lineplot1(where=( (first) & ( .04 < p_value < .06 ) ) );
        id beta f_stat pi rho p_value ; var dup dlw ; run;

        proc gplot data=lineplot1(where=((  p_value < .04 ) ));  by beta pi rho  ;
        symbol1 v=none i=join color=blue l=2;
        symbol2 v=none i=join color=black l=1;
        symbol3 v=none i=join color=blue l=2;
        symbol4 v=none i=join color=red l=3;
        symbol5 v=none i=join color=red l=3;
        axis1 order=0 to 3 by .5 ;
        *plot var*q ;
        plot (up test lw d_up d_lw ) * q / overlay vref = 0  href = 1  ; run;


         data r_new1 ; set r_new ;
        array ints fint1 fint2 hint1 hint2 fint1a fint2a fint1b fint2b;
        do over ints ; if ints < -9999 then ints = . ; end ;
        f_low = fint1 ; f_hi = fint2 ;
        f_lowa = fint1a ; f_hia = fint2a ;
        f_lowb = fint1b ; f_hib = fint2b ;
        h_low =  hint1; h_hi =hint2 ;
        d_low =  min(dint1,dint2) ; d_hi = max(dint1,dint2) ;
        run;

        proc gplot data=r_new1;*(where=((beta=1)&(pi = 1 ) ));  by beta pi rho  ;
        symbol1 v=none i=join color=blue l=2;
        symbol2 v=none i=join color=blue l=2;
        symbol3 v=none i=join color=red l=3;
        symbol4 v=none i=join color=red l=3;
        symbol5 v=none i=join color=green l=5;
        symbol6 v=none i=join color=green l=5;

        axis1 order=0 to 1 by .2 ;
        *plot (  f_low f_hi ) * (alpha) /overlay href= .01 .025 .05  vref = -1 0 1 ;
        *plot (  f_low f_hi d_low d_hi h_low h_hi) * (alpha) /overlay href= .01 .025 .05  vref =  1 ;
        plot (  f_low f_hi d_low d_hi h_low h_hi) * (t_value) /overlay href= 1.96  vref =  1 ;
        *plot ( d_low d_hi h_low h_hi ) * alpha /overlay href= .01 .025 .05  vref = -1 0 1 ;
         run;

         proc gplot data=r_new1;* (where=((beta=1)&(pi = 1 )&(rho > .5) ));  by beta pi  ;
        symbol1 v=none i=join color=black l=1;
        symbol2 v=none i=join color=black l=1;
        symbol3 v=none i=join color=blue l=2;
        symbol4 v=none i=join color=blue l=2;
        symbol5 v=none i=join color=red l=3;
        symbol6 v=none i=join color=red l=3;

        axis1 order=0 to 1 by .2 ;
        plot (dint1 dint2)*alpha  = rho  (fint1 fint2) * alpha = rho /overlay href= .01 .025 .05  vref
        = -1 0 1 ;
```

```
run;
```

Appendix C: SAS Proc IML code for contour plots for example application.

This routine was used to generate the contour plots of the constraint shapes as provided for the example application. Note that it is setup to plot the shapes for the specification #2.

```
PROC IMPORT OUT= WORK.data1
            DATAFILE= "C:\Users\josephh\Dropbox\submissions\JOE_Basmann\maketable4.dta"
            DBMS=STATA REPLACE;

RUN;
*
      Program to perform a 2 stage least squares and plot the confidence intervals
      Using the inverse test
;


      data new ; set data1 ;
        if baseco = 1 ;
        y2 = avexpr ;
        y1 = logpgp95 ;
        x1 = logem4 ;
        x2 = africa ;
        x3 = asia ;
        x4 = lat_abst ;
        x5 = other_cont ;
      * keep y2--x3 ;
      run ;
/*    data news ;
        num = 50 ;
        do ii = 1 to num;
        y2 = (ranuni(989875)*10) ; z = ranuni(9890753)*1;
        y1 = 5*y2 + z + 10 + rannor(979875)*10 ;;
        x1 = y2 + rannor(909875)*30 ;
        x2 =  (y2)+ rannor(987453)*20 ;
        x3 = (y2) + rannor(989745)*50 ;
        output ;
        end;
*/
       proc princomp data=new    out=new2  ; var x1-x5; run;

      data new2 ; set new2 ;
      array xx z1-z5 ; array pr prin1-prin5 ;
      do over xx ; xx = pr ; end ;
       run;

      Proc standard data=new2  mean=0 std=1 out=new3 ; var x1-x5 ; run;

      proc corr data=new3 ; var x1-x5 prin1-prin5 ; run;

/*    proc syslin data=new ;
      endogenous y2 ;
      instruments x1 x2 x3 ;
      m1: model y1 = y2 ;
      *m2: model y2 = x1 x2 x3 ;
      run; */


      proc syslin data=new3 2sls outest=est_2sls outcov ;
      endogenous y2 ;
      instruments x1 x2 x3 x4 x5;
      m1: model y1 = y2 x2 x3 x4 x5 ;
```

```
      run;

/*    proc syslin data=new outest=est_ols outcov ; ;
      m1: model y1 = x1 x2 x3 x4 x5;
      m2: model y2 = x1 x2 x3 x4 x5 ;
      run;
`     data new ; set new ; x1a = x1 ;
      proc syslin data=new outest=est_sur outcov sur ; ;
      m1: model y1 = x1 ;
      m2: model y2 = x1a ;
      run;*/

      data coeff ; set est_2sls(where=(_name_ ="")); run;
      data cov ; set est_2sls(where=(_name_ ~="")); run;

      proc iml ;

*================================================================;
      start Fieller( b1, b2, cov,t1);
      ratio = b1 / b2 ;
      t = t1 ;
      t2 = t**2 ;
      rad = (-t2*cov[1,1]*cov[2,2])-(2*b1*b2*cov[1,2])  + (t2*cov[1,2])+ (b2*b2*cov[1,1]) +
      (b1*b1*cov[2,2]) ;
      denom = ((t2*cov[2,2])-(b2*b2) ); p2 = t2*cov[1,2]-b1*b2 ;
      if (rad > 0) then do;
        p1 = t*sqrt(rad) ;
        lmx1 = ( p2 + p1 ) / denom;
        lmx2 = ( p2 - p1 ) / denom;
        lm1 = max(lmx1,lmx2) ; lm2 = min(lmx1,lmx2) ;
        if denom > 0 then do ; up = lm1 ; lw = lm2 ; end;
                else do ;
                d1 = abs(lm1 - ratio) ; d2 = abs(lm2-ratio) ;
                if (d1 < d2) then do ; lw = lm1 ; up = -99998 ; end ;
                                  else do ; up = lm2 ; lw = -99999 ; end ;
                end;
      end;
        else do ; up = -99998 ; lw = -99999 ; end;
/*    p = j(3,1);
      p[1,1] = (b2*b2 - t2*cov[2,2]);
      p[2,1] = -2*(b1*b2- t2*cov[1,2]);
      p[3,1] = b1*b1 - t2*cov[1,1];
      r = polyroot(p);
      if r[1,2] > 0 then do; r[1,1] = -99998 ; r[2,1] = -99999 ; end;

      lw_up = (min(r[1,1],r[2,1])) || ( max(r[1,1],r[2,1]) ) ; print b1 b2 ratio p r lw_up t up lw
      lmx1 lmx2 ; */
      lw_up = lw || up;
      return( lw_up );
      finish;
*================================================================;
*================================================================;
      start Delta( b1, b2, cov,t);    ;
      r1 = (1/b2) || (-b1/b2##2) ;
      se = sqrt(r1 * cov * r1`) ;
      lw_up = ( (b1/b2) - t # se) || ( (b1/b2) + t # se) ;
      return( lw_up );
      finish;
*================================================================;

*==================================================;
      start Cont_E(est, sss,  t, nPts=100);
      h = est[1,1] ; g = est[1,2] ;
      s1 = sss[1,1] ; s2 = sss[2,2] ; cv = sss[1,2] ;
```

```
      isss = inv(sss) ; w1 = isss[1,1] ; w2 = isss[2,2] ; w12 = isss[1,2] ;

*     subroutine to evaluate the  ellipse via a contour plot

      est = 1 by 2 vector of numerator then denominator
      sss = 2 by 2 matrix of covariance
      h = Numerator parameter estimate
      g = Denominator parameter estimate
      s1 = numerator estiamted variance
      s2 = denominator estimated variance
      cv = estimated covariance
      t = t-statistic assumed
      nPts = The number of points where evaluated
      ;
      upg = g + t*sqrt(s2) ;
      lwg = g - t*sqrt(s2) ;
      uph = h + t*sqrt(s1) ;
      lwh = h - t*sqrt(s1) ;

      diffg = (upg-lwg)/nPts ;
      diffh = (uph-lwh)/nPts ;
      vec1 = est*isss ;
      const1 = vec1*(est`) ;

       do gam = lwg to upg by diffg ;
       do theta = lwh to uph by diffh ;
              value = theta || gam ;
              tv = (const1 - 2#vec1*value` + value*isss*value`) ;
              if tv <= 0 then tv = -9999 ; else tv = sqrt(tv) ;
              tt = tt // ( tv || value );
       end ;
       end ;

      return( tt) ;
       finish ;

*==================================================;
*==================================================;
      start Cont_F(est, sss,  t, nPts=100);
      h = est[1,1] ; g = est[1,2] ;
      s1 = sss[1,1] ; s2 = sss[2,2] ; cv = sss[1,2] ;
      isss = inv(sss) ; w1 = isss[1,1] ; w2 = isss[2,2] ; w12 = isss[1,2] ;

*     subroutine to evaluate the  Fieller via a contour plot

      est = 1 by 2 vector of numerator then denominator
      sss = 2 by 2 matrix of covariance
      h = Numerator parameter estimate
      g = Denominator parameter estimate
      s1 = numerator estiamted variance
      s2 = denominator estimated variance
      cv = estimated covariance
      t = t-statistic assumed
      nPts = The number of points where evaluated
      ;
      upg = g + t*sqrt(s2) ;
      lwg = g - t*sqrt(s2) ;
      do tt1 = t to .2 by -.2 ;
      lw_up = Fieller( h, g, sss ,tt1);  * print tt1 lw_up ;
      if lw_up[1,2] > -999 then go to e1;
      end;
e1:   upb = 2* lw_up[1,2] ;
      lwb = 2* lw_up[1,1];
      print "Fieller" upb lwb ;
```

```
      lw_up = Delta( h, g, sss ,t);
      upb1 = 2* lw_up[1,2] ;
      lwb1 = 2* lw_up[1,1];
      print "Delta" upb lwb ;
      upb = max(upb,upb1) ;
      lwb = min(lwb,lwb1) ;
      upb = 10 ; lwb = -1 ; ***********************************CHnage ************************;

      diffg = (upg-lwg)/nPts ;
      diffb = (upb-lwb)/nPts ;
      vec1 = est*isss ;
      const1 = vec1*(est`) ;

        do gam = lwg to upg by diffg ;
        do beta = lwb to upb by diffb ;
               value = (beta#gam) || gam ;
               tv = (const1 - 2#vec1*value` + value*isss*value`) ;
               if tv <= 0 then tv = -9999 ; else tv = sqrt(tv) ;
               tt = tt // ( tv || beta || gam );
        end ;
        end ;

      return( tt) ;
       finish ;

*==================================================;
*==================================================;
      start Cont_D(est, sss,  t, nPts=100);
      h = est[1,1] ; g = est[1,2] ;
      s1 = sss[1,1] ; s2 = sss[2,2] ; cv = sss[1,2] ;
      isss = inv(sss) ; w1 = isss[1,1] ; w2 = isss[2,2] ; w12 = isss[1,2] ;

*     subroutine to evaluate the  ellipse via a contour plot

      est = 1 by 2 vector of numerator then denominator
      sss = 2 by 2 matrix of covariance
      h = Numerator parameter estimate
      g = Denominator parameter estimate
      s1 = numerator estiamted variance
      s2 = denominator estimated variance
      cv = estimated covariance
      t = t-statistic assumed
      nPts = The number of points where evaluated
      ;
      upg = g + t*sqrt(s2) ;
      lwg = g - t*sqrt(s2) ;
      lw_up = Delta( h, g, sss ,t);
      upb = 2* lw_up[1,2] ;
      lwb = 2* lw_up[1,1];
      upb = 10 ; lwb = -1 ; ***********************************CHnage ************************;
      diffg = (upg-lwg)/nPts ;
      diffb = (upb-lwb)/nPts ;
      vec1 = est*isss ;
      const1 = vec1*(est`) ;

        do gam = lwg to upg by diffg ;
        do beta = lwb to upb by diffb ;
               value =( (beta#g)+(h/g)#gam - h) || gam ;
               tv = (const1 - 2#vec1*value` + value*isss*value`) ;
               if tv <= 0 then tv = -9999 ; else tv = sqrt(tv) ;
               tt = tt // ( tv || beta || gam );
        end ;
        end ;
```

```
      return( tt) ;
       finish ;

*==================================================;


      *  read the data  for the just identified case ;

       use new ;
       read all var {y1} into y1 ;
       read all var {y2} into y2 ;
       read all var {x1 x2 x3 x4 x5} into z ; * specification (2) ;
      * read all var {x1} into z ; * specification (1) ;


/*     y1 = y1 / y1[:,] ; y2 = y2 / y2[:,] ; z = z /z[:,] ;*/  * divide data by the mean to recentre
      problem ;
*
      estimate the equations separately with the same regressors - the first is the instrument in
      the
      just identified case.
;

       n = nrow(y1) ;
       k = ncol(z) + 1 ;
       z = z || j(n,1,1); * put the intercept at the end of the regressors ;
       izpz = inv(z`*z) ;

*     reduced form equation for y1 ;
      cf1 = izpz*z`*y1 ; py1 = z*cf1 ; sig1 = ssq(y1-py1)/(n) ;
      Rs_1 = (corr(y1||py1)##2) [1,2]; print cf1 sig1  rs_1 ;

*     reduced form equation for y1 ;

      cf2 = izpz*z`*y2 ; py2 = z*cf2 ; sig2 = ssq(y2-py2)/(n) ;
      Rs_2 = (corr(y2||py2)##2) [1,2]; print cf2 sig2  rs_2;

      sig12 = ((y1-py1)`*(y2-py2)) / (n) ; *Sur covariance of the errors divide by n instead of n-k;


      cov_red = ((sig1 ||sig12)//(sig12 ||sig2) );

      covsur = cov_red @ izpz ;

      indx = 1 || (k + 1) ;

      cov = covsur[indx,indx ] ; print cov_red cov covsur ;

      corr_sur = sig12/sqrt(sig1#sig2) ; * assume the single instrument is the second parameter ;

      /*cov2sur = 5#({1.3333  -0.66667} // {-0.66667 1.3333}) ; cf1 = {3 3}` ; cf2 = cf1 ; * use a
      known result ;
      cov2sur = inv(cov2sur) ; */

      ratios = cf1[1,1] / cf2[1,1] ;

      c_est = cf1[1,1] || cf2[1,1] ; *estimated coefficients ; print "variable estimates" ratios
      c_est cov ;

      flw_up = Fieller(cf1[1,1] , cf2[1,1],cov,2) ;
      dlw_up = Delta(cf1[1,1] , cf2[1,1],cov,2) ;

      print flw_up dlw_up ;
```

```
*
      Define the range of values for the 2sls CI as +- 4*se of the 2sls interval
      NPts = the number of points for evaluation
;

      nPts = 500 ;  * set the number of points for evaluation ;
      t_value = 3; * determine the range of values around the means ;


      do jj = 1 to 1 ;

      id0 = ( jj || ratios || cf2[2,1]  ) ;
      tt = Cont_e(c_est,cov , t_value , nPts );
      c_e = c_e // ( tt || (2#(1-probt(tt[,1],n))) || ( j(nrow(tt),1,1)*id0 ) );

      tf = Cont_f(c_est,cov ,  t_value , nPts );
      c_f = c_f // ( tf || (2#(1-probt(tf[,1],n))) || ( j(nrow(tf),1,1)*id0 ) );

      td = Cont_d(c_est,cov ,  t_value, nPts );
      c_d = c_d // ( td || (2#(1-probt(td[,1],n))) || ( j(nrow(td),1,1)*id0 ) );




      end;

      create cont_e from c_E[c={"t" "denom" "numer" "probt" "parameter" "ratio1" "beta2"  }]; append
      from c_E;
      create cont_f from c_f[c={"t" "ratio" "denom" "probt" "parameter" "ratio1" "beta2"  }]; append
      from c_f;
      create cont_d from c_d[c={"t" "ratio" "denom" "probt" "parameter" "ratio1" "beta2"  }]; append
      from c_d;

      stop;
      quit ;
      run;

      run;

      goptions reset ; run;

      proc gcontour data=cont_e(where=(t > 0)) ; title "ellipse for estimates" ;
      axis1 order = -1 to 0 by .1;
      axis2 order = -1 to 0 by .1;
      plot numer*denom=probt / levels = .01 .025 .05 .1 .15 .49  haxis=axis1 vaxis=axis2; run; ;
      run;

      proc gcontour data=cont_f(where=(t > .001)) ;  title "Fieller Interval" ;
      axis1 order = -1 to 5 by 1;
      axis2 order = -1 to .5 by .1;
      plot ratio*denom=probt / levels = .000001 .00001 .0001 .001 .01 .025 .05 .1 .15 .49 vref = 0
      href = 0 vaxis= axis1 haxis = axis2 ; run;
      proc gcontour data=cont_d ; Title "Delta Interval" ;
      symbol1 color=black h=.5 font="Arial Rounded MT Bold" ;
      axis1 order = -1 to 5 by 1;
      axis2 order = -1 to .5 by .1;
      plot ratio*denom=probt / levels = .000001 .00001 .0001 .001  .01 .025 .05 .1 .15 .49 vref = 0
      href = 0 vaxis= axis1 haxis = axis2 ; run;

      goptions reset=all ;
      Title 'Specification (2)'  ; run;

      ods rtf ;
```

```
    ods graphics /  OUTPUTFMT=EMF ; run ;

    proc gcontour data=cont_f(where=(t > .001)) ;  title2 "Fieller Interval p_value = .01" ;
    symbol1 color=blue w=1.5 font="Arial Rounded MT Bold" ;
    symbol2 color=blue w=1.5 font="Arial Rounded MT Bold" ;
*   axis1 order = -1 to 3 by 1 value=(h=2); * use for specification 1 ;
    axis1 order = -1 to 7 by 1 value=(h=2);
    axis2 order = -1 to .2 by .1 value=(h=2);
    plot ratio*denom=probt / levels = .01  .99 vref = 0 href = 0 vaxis= axis1 haxis = axis2 ; run;

    proc gcontour data=cont_d ; Title2 "Delta Interval p_value = .01" ;
    symbol1 color=red w=.5 font="Arial Rounded MT Bold" ;
    symbol2 color=red w=.5 font="Arial Rounded MT Bold" ;

    plot ratio*denom=probt /
    levels = .01  .99 vref = 0 href = 0 vaxis= axis1 haxis = axis2 llevels = 2  ; run;

    proc gcontour data=cont_f(where=(t > .001)) ;  title2 "Fieller Interval p_value = .05" ;
    symbol1 color=blue w=1.5 font="Arial Rounded MT Bold" ;
    symbol2 color=blue w=1.5 font="Arial Rounded MT Bold" ;

    plot ratio*denom=probt / levels = .05  .99 vref = 0 href = 0 vaxis= axis1 haxis = axis2 ; run;

    proc gcontour data=cont_d ; Title2 "Delta Interval p_value = .05" ;
    symbol1 color=red w=.5 font="Arial Rounded MT Bold" ;
    symbol2 color=red w=.5 font="Arial Rounded MT Bold" ;

    plot ratio*denom=probt /
    levels = .05  .99 vref = 0 href = 0 vaxis= axis1 haxis = axis2 llevels = 2  ; run;

    proc gcontour data=cont_f(where=(t > .001)) ;  title2 "Fieller Interval p_value = .1" ;
    symbol1 color=blue w=1.5 font="Arial Rounded MT Bold" ;
    symbol2 color=blue w=1.5 font="Arial Rounded MT Bold" ;
    plot ratio*denom=probt / levels = .1  .99 vref = 0 href = 0 vaxis= axis1 haxis = axis2 ; run;

    proc gcontour data=cont_d ; Title2 "Delta Interval  p_value = .1" ;
    symbol1 color=red w=.5 font="Arial Rounded MT Bold" ;
    symbol2 color=red w=.5 font="Arial Rounded MT Bold" ;
    plot ratio*denom=probt /
    levels = .1  .99 vref = 0 href = 0 vaxis= axis1 haxis = axis2 llevels = 2  ; run;

    proc gcontour data=cont_f(where=(t > .001)) ;  title2 "Fieller Interval" ;
    symbol1 color=blue w=2 h=2  ;*  font="Arial Rounded MT Bold" ;
    symbol2 color=blue w=2 h=2  ;* font="Arial Rounded MT Bold" ;
    symbol3 color=blue w=2 h=2  ;* font="Arial Rounded MT Bold" ;
    symbol4 color=blue w=2 h=2  ;*  font="Arial Rounded MT Bold" ;
    plot ratio*denom=probt / levels = .01 .05 .1  .99 vref = 0 href = 0
      vaxis= axis1 haxis = axis2 ; run;

    proc gcontour data=cont_d ; Title2 "Delta Interval " ;
    symbol1 color=red w=2 h=2 font="Arial Rounded MT Bold" ;
    symbol2 color=red w=2 h=2 font="Arial Rounded MT Bold" ;
    symbol3 color=red w=2 h=2 font="Arial Rounded MT Bold" ;
    symbol4 color=red w=2 h=2 font="Arial Rounded MT Bold" ;
    plot ratio*denom=probt /
    levels = .01 .05 .1 .99 vref = 0 href = 0 vaxis= axis1 haxis = axis2 llevels = 2 2 2 2 2 ;
    run;


    ods rtf  close ;

    proc gcontour data=cont_e(where=(t > 0)) ; title "ellipse for estimates" ;
    axis1 order = -2 to 2 by 1;
    axis2 order = -2 to 2 by 1;
```

```
plot numer*denom=t / levels = 1 to 6 ;*vref=0 3 href=0 1 3 levels = 1 to 6   haxis=axis1
vaxis=axis2; run; ; run;

proc gcontour data=cont_f(where=(t > 0)) ; title "Fieller Interval" ;
axis1 order = -2 to 2 by 1;
axis2 order = -2 to 2 by 1;
plot ratio*denom=t / levels = 1 to 6 ;*/ vref=0 1 href= 3  levels = 1 to 6   haxis=axis1
vaxis=axis2; run; ; run;

proc gcontour data=cont_d(where=(t > 0)) ;  Title "Delta Interval" ;
symbol1 color=black h=.2 font="Arial Rounded MT Bold" ;
axis1 order = -2 to 2 by 1;
axis2 order = -2 to 2 by 1;
plot ratio*denom=t / levels = 1 to 6 ;*/ vref=0 1 href= 3  levels = 1 to 6   haxis=axis1
vaxis=axis2;* autolabel ; run; ; run;
```