Author/s:
Hirschberg, J;Lye, J

Title:
Inverting the indirect—The ellipse and the boomerang: Visualizing the confidence intervals of the structural coefficient from two-stage least squares

Date:
2017-08-01

Citation:
Hirschberg, J. & Lye, J. (2017). Inverting the indirect—The ellipse and the boomerang: Visualizing the confidence intervals of the structural coefficient from two-stage least squares. Journal of Econometrics, 199 (2), pp.173-183. https://doi.org/10.1016/j.jeconom.2017.05.008.

Persistent Link:
https://hdl.handle.net/11343/295958

Inverting the Indirect - The Ellipse and the Boomerang: Visualising the Confidence Intervals of the Structural Coefficient from Two-Stage Least Squares. [1]

Abstract:

In 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) employed Fieller's 1932 result to derive the density function of the estimator. In this paper we present a novel graphical exposition of Fieller's 1954 technique to approximate the confidence interval for the *2SLS* estimator. We use this approach to examine how the degree of endogeneity and instrument relevance influences the correspondence between the Fieller and traditional asymptotic confidence intervals for the estimator.

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

JEL: C12, C26, C36, C18

Authors:

Joe Hirschberg, Economics, University of Melbourne, Australia.
j.hirschberg@unimelb.edu.au

Jenny Lye, Economics, University of Melbourne, Australia.
Jnlye@unimelb.edu.au

---

[1] We have greatly benefited by the comments of the referees for this paper. The usual caveat holds.

1. Introduction

Basmann (1974) found that when the structural equation is exactly identified, 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. 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 1954, Fieller developed a technique to approximate the confidence interval (CI) for the ratio that has become the most widely employed alternative to the usual asymptotic estimate. In this paper we construct a new graphic method for comparing the Fieller and usual asymptotic CIs.

We demonstrate how the approximate CIs 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) CI and the Fieller interval.[2] 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 CI 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.

---

[2] 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.

The estimation of the just-identified model is 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 review the equivalence of the indirect least squares and the *2SLS* estimators to demonstrate that the usual *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 graphical method for the comparison of the Delta to the Fieller CIs for the ratio. A number of simulated cases are then presented to demonstrate how the degree of endogeneity and instrument relevance influences the relationship between these CIs. We then present a simple empirical application along with two alternative graphical methods for comparing these intervals. We conclude with a summary of our findings and suggestions for further research in this area.

2. Tests of Significance and CIs 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]

$$Y = X\beta + u \tag{1}$$

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. In this case instrumental variable estimation employs a T by 1 vector Z that satisfies two

[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)

conditions: (*i*) it is uncorrelated with u (a property referred to as: exogeneity) and (*ii*) it is correlated with X (a property referred to as: relevance).

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 - also referred to as the first stage – defined by:

$$X = Z\gamma + \varepsilon \tag{2}$$

where,

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

and the correlation between u and ε is defined as $\rho = \frac{\sigma_{u\varepsilon}}{\sigma_\varepsilon \sigma_u}$. The reduced form equation corresponding to (1) can be written as;

$$\begin{aligned} Y &= Z\gamma\beta + v, \text{ or} \\ Y &= Z\theta + v \end{aligned}, \tag{4}$$

where,

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

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

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

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

The *2SLS* estimator for β in (1) is

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

where $P_Z = Z\left(Z^\top Z\right)^{-1} Z^\top$.

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,

$$\sqrt{T}\left(b_s-\beta\right)\xrightarrow{d}N\left(0,\sigma_u^2\gamma^{-2}\left(V_{ZZ}\right)^{-1}\right),\ V_{ZZ}=\underset{T\to\infty}{plim}\ \frac{1}{T}Z^{\mathsf{T}}Z>0\,. \tag{9}$$

In the exactly identified case it is well known that the Indirect Least Squares (*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 θ and γ. Here we review the corresponding result that the Delta method estimate of the variance of the ratio of the reduced form parameters is equivalent to the usual *2SLS* and *IV* variance estimates.

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

$$b_I=\tfrac{h}{g}=\frac{\left(\left(Z^{\mathsf{T}}Z\right)^{-1}Z^{\mathsf{T}}Y\right)}{\left(\left(Z^{\mathsf{T}}Z\right)^{-1}Z^{\mathsf{T}}X\right)}=\left(Z^{\mathsf{T}}X\right)^{-1}Z^{\mathsf{T}}Y \tag{10}$$

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

$$\operatorname{cov}(h,g)=S_Z{}^{-1}\Omega \tag{11}$$

where $S_Z=Z^{\mathsf{T}}Z$ is the sum of the squared elements of *Z*. 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 *ILS* estimate can be defined as:[4]

$$\operatorname{Var}(b_I)\approx S_Z{}^{-1}\begin{bmatrix}\frac{\partial b_I}{\partial h} & \frac{\partial b_I}{\partial g}\end{bmatrix}\Omega\begin{bmatrix}\frac{\partial b_I}{\partial h}\\ \frac{\partial b_I}{\partial g}\end{bmatrix} \tag{12}$$

[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.

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

$$\text{Var}(b_I) \approx g^{-4} S_Z^{-1} \left(g^2 \sigma_v^2 + h^2 \sigma_\varepsilon^2 - 2\sigma_{v\varepsilon} gh \right) \tag{13}$$

From (6) we have $\sigma_u^2 = \sigma_v^2 - \beta\sigma_\varepsilon^2 - 2\beta\sigma_{v\varepsilon}$. Substituting the consistent estimate b_I for β to obtain $\tilde{\sigma}_u^2$:

$$\tilde{\sigma}_u^2 = \sigma_v^2 - b_I \sigma_\varepsilon^2 - 2b_I \sigma_{v\varepsilon} \tag{14}$$

Thus we find that:

$$\text{Var}(b_I) \approx g^{-2} S_Z^{-1} \tilde{\sigma}_u^2, \tag{15}$$

If we estimate $\tilde{\sigma}_u^2$ by $\hat{\sigma}_u^2 = \frac{1}{T-1}\left(Y - Xb_I\right)^{\mathsf{T}}\left(Y - Xb_I\right)$ then 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 is equivalent to the usual *2SLS* and *IV* estimate of the variance of the estimate for β.[5]

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

$$W = \frac{\left(b_I - \beta_0\right)^2 X^{\mathsf{T}} P_Z X}{\hat{\sigma}_u^2} \tag{16}$$

has an asymptotic $\chi^2(1)$ distribution where $\hat{\sigma}_u^2$ converges in probability to σ_u^2. Then *W* is the square of the asymptotic *t*-ratio which is commonly used in applied work. Thus the equivalent Wald CIs 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

[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.

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 CIs for instrumental variable estimation with weak instruments, that is, when instrument relevance is low. Staiger and Stock (1997) show that under weak instruments *2SLS* 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 CIs 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 CIs in the presence of weak instruments has been illustrated in a number of Monte Carlo studies. Zivot, Startz and Nelson (1998) show that in the presence of weak instruments and strong endogeneity Delta CIs 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 (1950) statistic (*AR*) provides an alternative approach to hypothesis tests and constructing CIs. The CIs based on the *AR* statistic can be shown to be equivalent to the Fieller intervals in the just-identified model. The *AR* CI can be bounded, unbounded or cover the real line. The unbounded interval occurs if the *t*-statistic for testing $\gamma = 0$ in (2) cannot be rejected at the $t_{1-\alpha}$ 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 *t*-statistic for testing $\gamma = 0$ (a test of instrument relevance) in (2) is less than $\sqrt{10} \approx 3.163$, (see also Stock and Yogo 2005 for formal tests of weak instruments).

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). Moreira (2001) shows that the *AR* statistic is UMP unbiased for the just-identified model when the errors are *iid* homoscedastic normal and Ω is known. Zivot, Startz and Nelson (1998) investigate properties of 95% CIs 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 CIs 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).

3. Employing a Constrained Optimization to compare 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 CIs. 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 − α)% 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 scalar δ defined as a linear combination of a $k \times 1$ vector $\mathbf{\Pi}^{\mathrm{T}} = \{\pi_1, \pi_2, \cdots, \pi_k\}$ defined as $\delta = \mathbf{a}^{\mathrm{T}}\mathbf{\Pi} + c$, 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:

$$\mathcal{L} = \left(\mathbf{a}^{\mathrm{T}}\mathbf{\Pi} + c\right) - \lambda\left(\left(\mathbf{P} - \mathbf{\Pi}\right)^{\mathrm{T}} \mathbf{\Phi}^{-1}\left(\mathbf{P} - \mathbf{\Pi}\right) - z^2\right) \tag{17}$$

where $\mathbf{P}_{k\times 1} \sim N\left(\mathbf{\Pi}_{k\times 1}, \mathbf{\Phi}_{k\times k}\right)$, $\mathbf{P}^{\mathrm{T}} = \{p_1, p_2, \cdots, p_k\}$, $\mathbf{\Phi}$ is the covariance matrix of the p_is 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 Distribution.[6]

Von Luxburg and Franz (2004) show that in the case where $k = 2$ Fieller's bounds for the ratio of two normally distributed variables $d = p_1/p_2$ as an estimate of $\delta = \pi_1/\pi_2$ where:

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \sim N\left\{\begin{bmatrix} \pi_1 \\ \pi_2 \end{bmatrix}, \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{bmatrix}\right\} \tag{18}$$

can be found as the maximum and minimum from the constrained optimization with the same constraint as in (17) which is defined by:[7]

$$\mathcal{L} = \left(\frac{\pi_1}{\pi_2}\right) - \lambda\left(\left[\left(p_1 - \pi_1\right) \quad \left(p_2 - \pi_2\right)\right]\begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{bmatrix}^{-1}\begin{bmatrix} \left(p_1 - \pi_1\right) \\ \left(p_2 - \pi_2\right) \end{bmatrix} - z^2\right). \tag{19}$$

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 find that the corresponding Delta method CI 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 p_1/p_2 and the estimate of ϕ_{12} (denoted by $\hat{\phi}_{12}$). We conclude that when the signs of p_1/p_2 and $\hat{\phi}_{12}$ are the same and we can reject the null that $\pi_2 = 0$ with a high degree of confidence the Fieller and Delta CIs will be

---

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

7 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 (17) and (19).

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 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 CIs.

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 reparameterize the optimization problem in order to construct a constraint shape in two dimensions, where one dimension is the parameter of interest δ. Thus we can determine the CI by the limits of the constraint shape.

In general if we consider a function of two parameters $\delta = f(\pi_1, \pi_2)$ that is subject to a quadratic constraint

$$\mathcal{L} = f(\pi_1, \pi_2) - \lambda\left(\begin{bmatrix} (p_1 - \pi_1) & (p_2 - \pi_2) \end{bmatrix} \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{bmatrix}^{-1} \begin{bmatrix} (p_1 - \pi_1) \\ (p_2 - \pi_2) \end{bmatrix} - z^2 \right). \tag{20}$$

and if we can solve this function for one of the parameters as in the form : $\pi_1 = g(\delta, \pi_2)$, then we can substitute for π_1 in the function defined by:

$$\mathcal{L} = \delta - \lambda\left(\begin{bmatrix} (p_1 - g(\delta, \pi_2)) & (p_2 - \pi_2) \end{bmatrix} \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{bmatrix}^{-1} \begin{bmatrix} (p_1 - g(\delta, \pi_2)) \\ (p_2 - \pi_2) \end{bmatrix} - z^2 \right) \tag{21}$$

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

$$\begin{bmatrix} (p_1 - g(\delta, \pi_2)) & (p_2 - \pi_2) \end{bmatrix} \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{bmatrix}^{-1} \begin{bmatrix} (p_1 - g(\delta, \pi_2)) \\ (p_2 - \pi_2) \end{bmatrix} = z^2 \tag{22}$$

Thus (22) defines a constraint shape in (δ, π_2) space and the limits of this shape projected on either the δ or the π_2 axis define the 100(1-α)% CI for the corresponding estimates d and p_2 as determined by the value of *z* (i.e. $z = 1.96$ when $\alpha = .05$).

In the case of the Fieller approach applied to the ratio of random normal variables we define the function $f(\pi_1, \pi_2)$ as $\delta = \pi_1 / \pi_2$. Thus we can solve for $\pi_1 = \delta\pi_2$ and the constraint shape is provided by:

$$\begin{bmatrix} (p_1 - \delta\pi_2) & (p_2 - \pi_2) \end{bmatrix} \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{bmatrix}^{-1} \begin{bmatrix} (p_1 - \delta\pi_2) \\ (p_2 - \pi_2) \end{bmatrix} = z^2 \tag{23}$$

For the equivalent Delta CI we linearize the ratio $d = p_1 / p_2$ using a first order Taylor series approximation around the estimates b_1 and b_2 as:

$$\delta \approx d + \begin{bmatrix} \frac{\partial d}{\partial p_1} & \frac{\partial d}{\partial p_2} \end{bmatrix} \begin{bmatrix} (\pi_1 - p_1) \\ (\pi_2 - p_2) \end{bmatrix} = \frac{p_1}{p_2} - \frac{1}{p_2}(\pi_1 - p_1) + \frac{p_1}{p_2^2}(\pi_2 - p_2) \tag{24}$$

When solving for π_1 as a linear function of δ and π_2:

$$\pi_1 = p_2\left(\delta + \frac{p_1}{p_2^2}(\pi_2 - p_2)\right), \tag{25}$$

and substitute (25) for π_1 into (22) we obtain the formula for a constraint shape as:

$$\begin{bmatrix} \left(-\frac{1}{p_2}\left(\delta p_2^2 - 2p_1p_2 + \pi_2 p_1\right)\right) & (p_2 - \pi_2) \end{bmatrix} \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{12} & \phi_{22} \end{bmatrix}^{-1} \begin{bmatrix} \left(-\frac{1}{p_2}\left(\delta p_2^2 - 2p_1p_2 + \pi_2 p_1\right)\right) \\ (p_2 - \pi_2) \end{bmatrix} = z^2 \tag{26}$$

4 The Fieller and Delta intervals in the Case of the Just-identified Simultaneous Equations Model.

In order to make comparisons between the Delta and Fieller CIs for the parameter β from the structural equation $Y = X\beta + u$ we consider the parameters from the reduced form equations (2) and (4). β is estimated by *ILS* as the ratio of the estimates of θ as *h* and γ as *g*, where $g = \left(Z^\top Z\right)^{-1} Z^\top X$ and $h = \left(Z^\top Z\right)^{-1} Z^\top Y$. Recall from (5) we have that $\begin{pmatrix} v_i \\ \varepsilon_i \end{pmatrix} \sim$ i.i.d. $(0, \Omega)$.

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.

Their simulations assume $\sigma_\varepsilon^2 = 1$, $\sigma_u^2 = 1$, and define the instrument as an independent standard normal variable $\left(Z \sim N(\mathbf{0}, \mathbf{I}_T)\right)$ thus:

$$\begin{bmatrix} h \\ g \end{bmatrix} \sim \left[\begin{pmatrix} \theta \\ \gamma \end{pmatrix}, \frac{1}{T} \times \begin{pmatrix} \beta + 2\beta\rho + 1 & \beta + \rho \\ \beta + \rho & 1 \end{pmatrix} \right], \tag{27}$$

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 (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 β and ρ changing the implied values of g (and by implication h), we can investigate the comparison between the Delta and Fieller CIs using the implied shapes defined above in (23) and (26).

In these cases the inverse of Ω is defined as

$$\Omega^{-1} = \tfrac{1}{\rho^2 - 1} \begin{pmatrix} -1 & \beta + \rho \\ \beta + \rho & -\left(\beta^2 + 2\rho\beta + 1\right) \end{pmatrix} \tag{28}$$

and the quadratic constraint in the Lagrangian from (21) is defined as:

$$\left(\rho^2 - 1\right)^{-1} \left[(h - \theta) \quad (g - \gamma) \right] \begin{pmatrix} -1 & \beta + \rho \\ \beta + \rho & -\left(\beta^2 + 2\rho\beta + 1\right) \end{pmatrix} \begin{bmatrix} (h - \theta) \\ (g - \gamma) \end{bmatrix} = \frac{z^2}{T}, \tag{29}$$

which is equivalent to the constraint shape defined by the implicit equation:

$$(g - \gamma)\left(\left(\rho + \tfrac{1}{g} h\right) \tfrac{h - \theta}{\rho^2 - 1} - \tfrac{g - \gamma}{\rho^2 - 1}\left(\tfrac{1}{g^2} h^2 + \tfrac{2}{g} h\rho + 1\right)\right) - (h - \theta)\left(\tfrac{h - \theta}{\rho^2 - 1} - \left(\rho + \tfrac{1}{g} h\right) \tfrac{g - \gamma}{\rho^2 - 1}\right) = \tfrac{1}{T} z^2. \tag{30}$$

4.1 *The Ellipse that defines the Fieller and Delta Intervals*

Under the assumption that $T = 100$, $g = .5$, $h = .5$, $\beta = 1$, and $z = 1.96$, the constraint shape defined in (30) can be drawn in (θ, γ) space as an ellipse as shown in Figure 1 for $\rho = .01$ and $\rho = .99$.

(Figure 1)

The univariate 95% CI for γ can be found from the limits of the ellipse projected onto the horizontal axis and the equivalent interval for θ from the projection onto the vertical axis. Because the variance for θ is a function of ρ we note that values of ρ influence the CI 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)

Figure 2 follows the graphic method proposed in HL 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 shape we use the definition of θ in terms of β and γ as defined in (4) by $\theta = \gamma\beta$. In this case the expression (23) can be written as:

$$-(g-\gamma)\left(\tfrac{g-\gamma}{\rho^2-1}\left(b^2+2\rho b+1\right)-\tfrac{b+\rho}{\rho^2-1}\left(h-\beta\gamma\right)\right)-\left(\tfrac{1}{\rho^2-1}\left(h-\beta\gamma\right)-\left(b+\rho\right)\tfrac{g-\gamma}{\rho^2-1}\right)\left(h-\beta\gamma\right)=\tfrac{1}{T}z^2, \quad (31)$$

which defines constraint shapes in (β, γ) space as shown in Figure 3.

(Figure 3)

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 CI 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 CI 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 to zero 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.[8]

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 with a similar shape to 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:

$$\beta = b + \begin{bmatrix} \frac{\partial b}{\partial h} & \frac{\partial b}{\partial g} \end{bmatrix} \begin{bmatrix} (\theta - h) \\ (\gamma - g) \end{bmatrix} = \tfrac{1}{g^2}\left(g\theta - h\gamma + gh\right) \qquad (32)$$

Now we solve (32) for θ and we obtain:

$$\theta = g\beta + \tfrac{h}{g}\gamma - h \qquad (33)$$

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 (30) we obtain the following constraint ellipse in (β, γ) space.

$$\begin{pmatrix} -\left(\frac{1}{\rho^2-1}\left(g\beta - 2h + \frac{1}{g}h\gamma\right) + (b+\rho)\frac{g-\gamma}{\rho^2-1}\right)\left(g\beta - 2h + \frac{1}{g}h\gamma\right) \\ -(g-\gamma)\left(\frac{g-\gamma}{\rho^2-1}\left(b^2 + 2\rho b + 1\right) + \frac{b+\rho}{\rho^2-1}\left(g\beta - 2h + \frac{1}{g}h\gamma\right)\right) \end{pmatrix} = \tfrac{1}{T}z^2 \qquad (34)$$

From Figure 4 we can see that unlike the case for the Fieller interval, the value of ρ does not

---

[8] 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.

influence the confidence bounds we find for the Delta estimate. Note that these CIs match the result we obtained from Figure 2.[9]

(Figure 4)

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

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 and clearly demonstrates how the Fieller interval is strongly influenced by the level of endogeneity.

(Figure 5)

In Figure 6 we change the sample size from 100 to 50 and assume a 99% confidence level. 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 ρ = .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 1% level of significance 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 widens so that the lower-bound of the Fieller interval approaches zero.

(Figure 6)

A characteristic of the Fieller method is that the resulting CI may not necessarily have finite limits when (31) defines an open constraint shape. The CI 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

---

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

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 $\rho = .9$ implying the presence of a high degree 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)

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 four 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.[10] 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 two plots for the cases where $F = 4$ and $F = 1$ result in open constraints shapes for the Fieller. This indicates that in these cases the Fieller lower-bounds are not finite.

(Figure 8)

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 "average index of protection against expropriation risk from 1985 to 1995" or *IPAE* 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:

$$\log(y_i) = \mu + \beta R_i + W_i'\lambda + u_i \tag{35}$$

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 variable. Thus the reduced form equation for R (the first equation) is defined as:

$$R_i = \eta + \gamma \log\left(M_i\right) + W_i'\delta + \varepsilon_i \tag{36}$$

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

$$\log(y_i) = \lambda + \theta \log(M_i) + W_i'\phi + v_i \tag{37}$$

[10] 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.

Thus this model is just-identified with $\log(M_i)$ being excluded from equation (35).

(Table 1)

Table 1 reports *2SLS* estimates when the dependent variable is Log GDP per capita.[11] Table 2 provides the equivalent estimates of the two reduced form equations (36) and (37) along with the estimate of the correlation between the estimates of θ and γ.

(Table 2)

5.1 Example Fieller and Delta Constraint Shapes

In order to compare the Fieller and the Delta intervals for the specifications defined above, we can plot the corresponding surface for a given value of α for each specification. Figure 9 plots the Fieller Boomerang and the Delta Ellipse constraint shapes for the results in Table 2 for three values of α. In this example we find that the correlation is positive and approximately .6 between the reduced form coefficient estimates. 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. However, the Fieller constraint shape when $\alpha = .05$ is open to the higher values of β. 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. Thus we conclude that the 95% interval ($\alpha = .05$) for the Fieller is $(.603, \infty)$ and the Delta is (.223, 1.99). 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).

[11] These results are equivalent to column 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. The Stata and SAS computer code for the generation of the figures and results presented in this section are available in Hirschberg and Lye (2017).

(Figure 9)

5.2 *Alternate Graphical Representations of the Fieller and Delta Intervals*

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. These approaches have previously been suggested to be used for comparisons between the Fieller and Delta CIs. However, these plots only provide a means for comparison of the alternative CIs and they do not indicate how the estimated parameters influence the comparison. First we present the line plot approach that can be used to determine the nature of the Fieller solution when the roots of the implied quadratic are complex. This alternative has been presented in Hirschberg et al 2008, Hirschberg and Lye 2010a, 2017 for ratios and other functions of parameters from regression equations. In the second we present the inferences that may be drawn with different values of α and can be useful to highlight the divergence between the Fieller and Delta CIs as proposed by Cook and Weisberg (1990).

5.3 *The Implicit Linear Function Plot*

The implicit linear function or line plot method for the definition of the Fieller intervals is based on the interpretation of the Fieller 's 1954 approximation method as the inversion of a test for a linear function of normally distributed random variables. In Hirschberg and Lye (2010a) we demonstrate this method for the Fieller case and discuss the generalization of this method in Lye and Hirschberg (2012) for the related problem of function turning points in high order polynomials.

Consider the linear combination *l* of the estimated parameters defined as:

$$l = h - b_0 g \qquad (38)$$

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)% CI. Thus when $l = 0$, $h - b_I g = 0$ and b_0 is equal to the *2SLS* estimate for β. The CI for this line is defined by:

$$\mathrm{CI}\left(l \mid b_0\right) = h - b_0 g \pm t_{\alpha/2}\sqrt{\hat{\sigma}_h^2 - 2b_0\hat{\sigma}_{gh} + \hat{\sigma}_g^2 b_0^2 g^2} \tag{39}$$

where $t_{\alpha/2}$ is the value from the *t* distribution with an α level of significance and $T-k$ degrees of freedom, *k* is the number of included covariates including the intercept term, and $\hat{\sigma}_h^2, \hat{\sigma}_{gh}$, and $\hat{\sigma}_g^2$ are the estimated elements of cov(*h*,*g*). The inverse confidence bounds for this estimate are defined by values of b_0 when the CI for *l* cuts the zero axis as defined by the two possible values b_1 and b_2 where:

$$h - b_1 g - t_{\alpha/2}\sqrt{\hat{\sigma}_h^2 - 2b_0\hat{\sigma}_{gh} + \hat{\sigma}_g^2 b_0^2 g^2} = 0 \text{ and } h - b_2 g + t_{\alpha/2}\sqrt{\hat{\sigma}_h^2 - 2b_0\hat{\sigma}_{gh} + \hat{\sigma}_g^2 b_0^2 g^2} = 0 \tag{40}$$

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

$$b_i^2\left(g^2 - t_{\alpha/2}^2 g\hat{\sigma}_g^2\right) - b_i\left(2g - t_{\alpha/2}^2\hat{\sigma}_{gh}\right) + \left(h^2 - t_{\alpha/2}^2\hat{\sigma}_h^2\right) = 0 \tag{41}$$

In Figure 10 we plot the line defined by (38) along with the 100(1-α)% CI as defined by (39), it is possible to obtain the Fieller interval by locating the values of b_0 where they cross the zero axis. One useful feature of the line plot method is that it can always be constructed, thus it can detect those circumstances when the Fieller interval does not provide finite CI bounds and the real roots of (41) may not exist.

(Figure 10)

The plot in Figure 10 confirms the same interval obtained in Figure 9 for $\alpha = .01$ in the case where the lower-bound is present but not an upper-bound. 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 interval (41) are real but only one is meaningful. This plot displays the advantage of the line plot since it can be drawn in cases where the roots of (41) 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).

5.4 Comparisons of CI Plots by variation in α

The implied CIs for the Fieller and the Delta vary with different levels of significance as noted from Figure 9. Although this figure demonstrates how the constraint shapes differ for various values of α they do not provide a means for making easy comparisons by values of α. Cook and Weisberg (1990) propose that such a comparison can be made with these plots. Figure 11 presents such a plot that is based on the parameter estimates from our empirical example.

From Figure 11 one can note that the divergence of the two intervals becomes most pronounced at the lower values of α. By definition the CI for $\alpha \to 1$ will converge at the point estimate b_I. It can be noted from this figure that the upper-bound for the Fieller CI is not defined for values of α greater than .1, however the lower-bound is defined as a finite value above zero for α < .01. In addition, these plots also demonstrate the highly asymmetrical nature of the Fieller CIs.

(Figure 11)

6. Conclusions

We have demonstrated that the Delta and Fieller CIs 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 CIs agree or diverge without the need for a simulation.

Our findings indicate that when the degree of endogeneity increases, the Fieller CI 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 $\alpha = .05$ 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., 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 R. Blundell, W. K. Newey, and T. Persson, (Eds.), *Advances in Economics and Econometrics: Theory and Applications*, Ninth World Congress, Vol. 3, Cambridge University Press, Cambridge, pp. 122–173.

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

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., Richardson, D. H. and Rohr, R. J., 1974, "An Experimental Study of Structural Estimators and Test Statistics Associated with Dynamical Econometric Models, *Econometrica*, 42, 717-730.

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

Chernozhukov, V and Hansen, C., 2008, "The reduced form: A simple approach to inference with weak instruments", *Economics Letters*, 100, 68–71.

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

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: B. Baltagi, (Ed.) *Companion to Theoretical Econometrics*, Blackwell, Oxford, pp. 494-519.

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. ,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.

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. and Lye, J., 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. and Lye, J., 2017, "Alternative Graphical Representations of the Confidence Intervals for the Structural Coefficient from Exactly Identified Two-State Least Squares", *Department of Economics*, *University of Melbourne, Working Paper*.

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

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.

Lye, J. and Hirschberg, J., 2012, "Inverse Test Confidence Intervals for Turning-Points: A demonstration with Higher Order Polynomials.", in: D. Terrell and D. Millimet, (Eds.), D. *Advances in Econometrics*, 30th Anniversary Edition (Volume 30), Emerald Group Publishing Limited, Bingley, 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.

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.

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

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, Tübingen.

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.

Figures

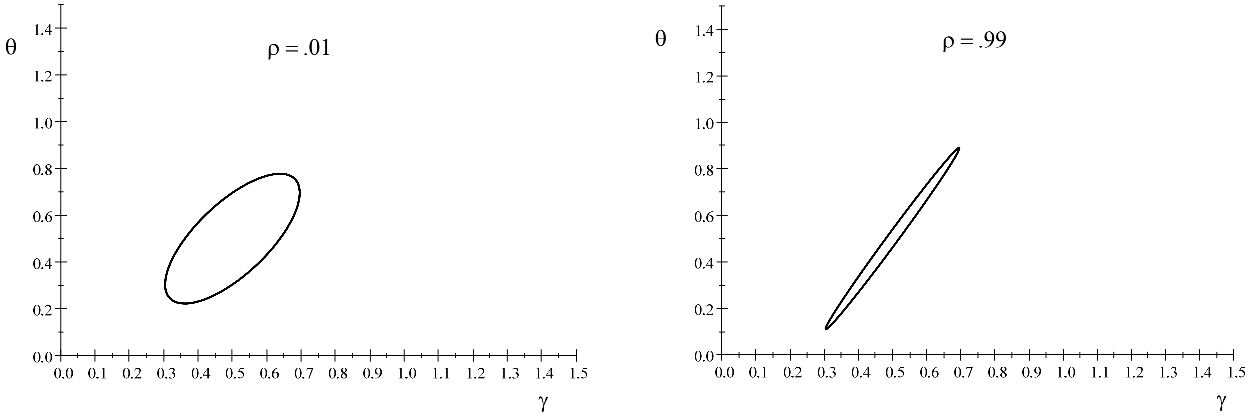


Figure 1 The implied ellipse constraint in (θ,γ) space when $T = 100$, $g = .5$, $h = .5$, $\beta = 1$, and $z = 1.96$ for $\rho = .01$ and $\rho = .99$.

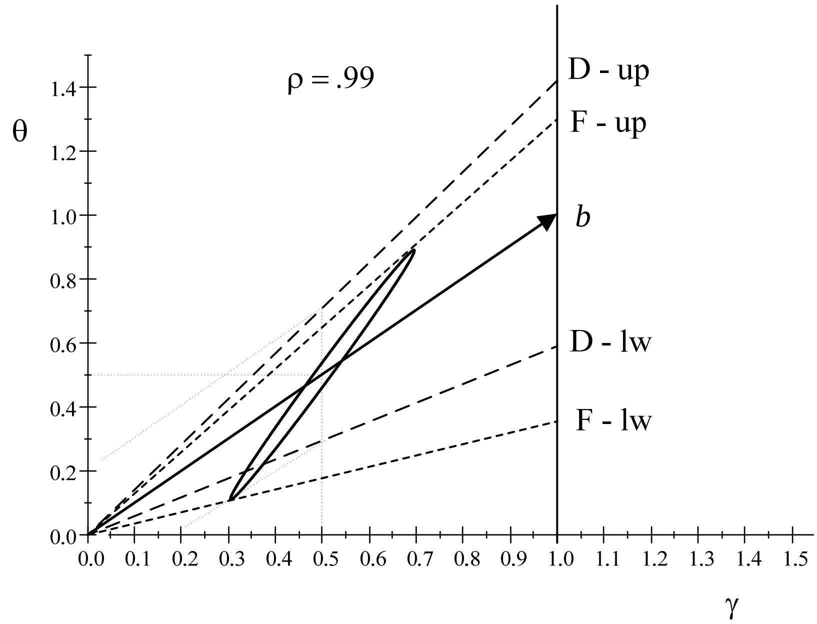
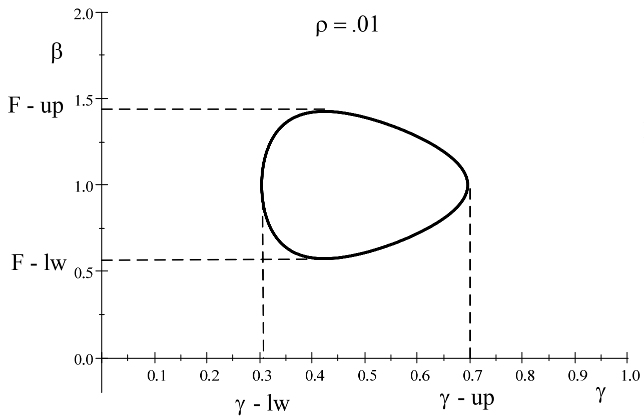


Figure 2 The construction of the Fieller bounds ((F-up, F-lw) and Delta bounds (D-up, D-lw) for β defined from the constraint ellipse for $T = 100$, $g = .5$, $h = .5$, $\beta = 1$, $z = 1.96$, and when $\rho = .99$ based on the method proposed in HL.



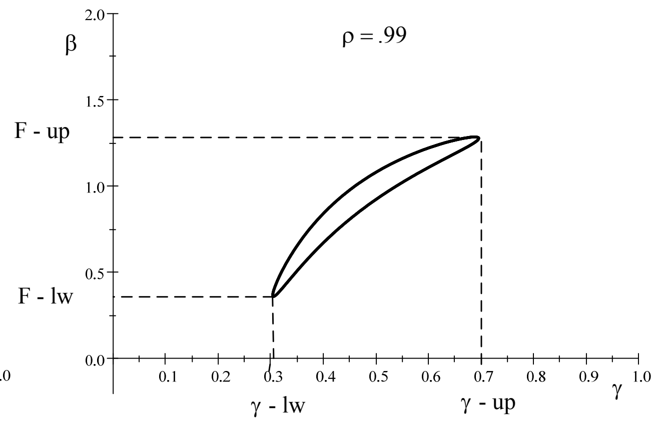


Figure 3 The Fieller 95% confidence bounds for β (F-up, F-lw) are defined as the limits of the implied constraint shape for $T = 100$, $g = .5$, $h = .5$, $\beta = 1$, and $z = 1.96$. When $\rho = .99$ and $\rho = .01$.

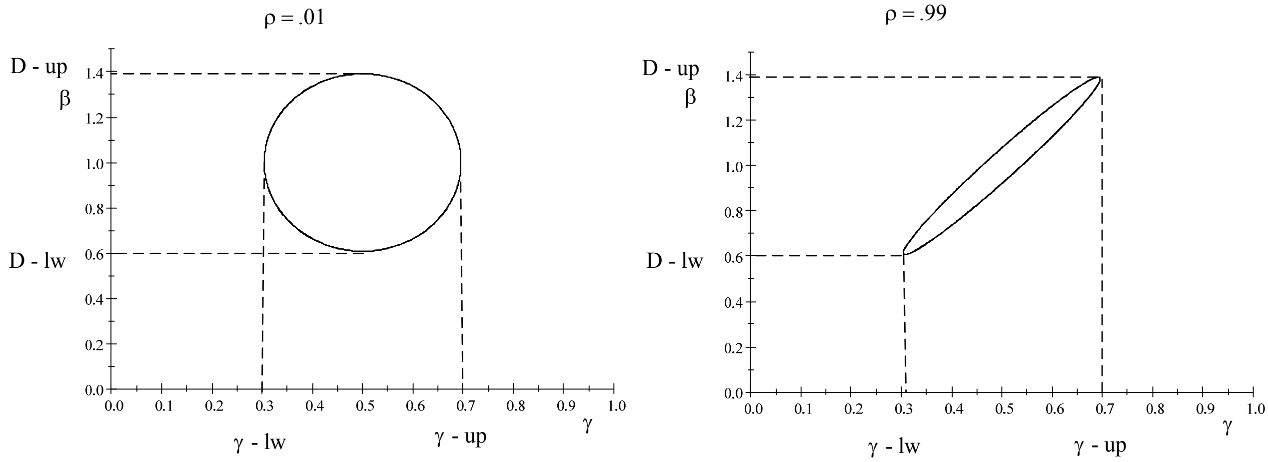


Figure 4 The Delta method 95% confidence bounds (D-up, D-lw) for β are defined as the limits of the implied constraint shape (in this case an ellipse) for $T = 100, g = .5, h = .5, \beta = 1,$ and $z = 1.96$. When $\rho = .99$ and $\rho = .01$.

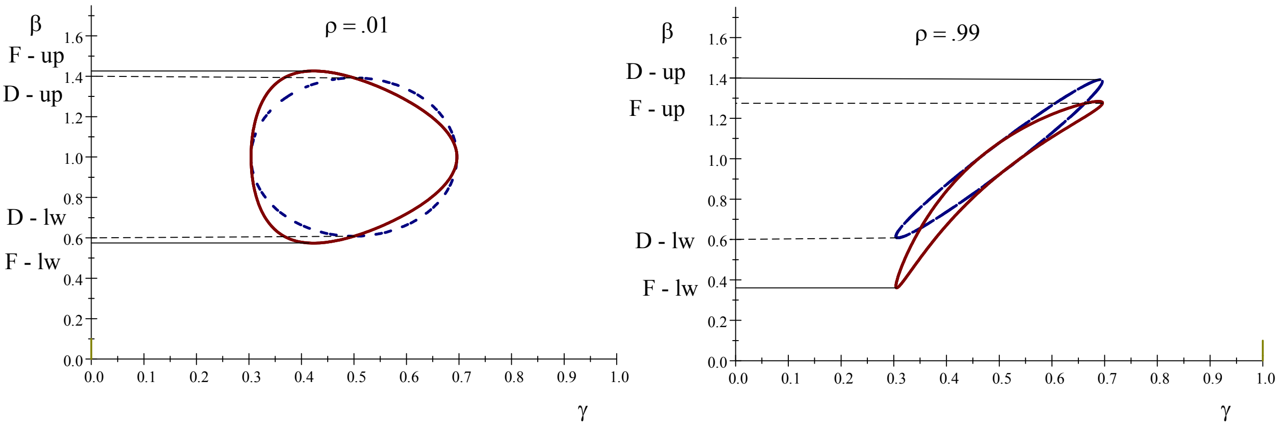


Figure 5 The Delta (dashed line) and Fieller (solid line) 95% CIs for β implied by the two constraint shapes for $T = 100$, $g = .5$, $h = .5$, $\beta = 1$, and $z = 1.96$. When $\rho = .99$ and $\rho = .01$.

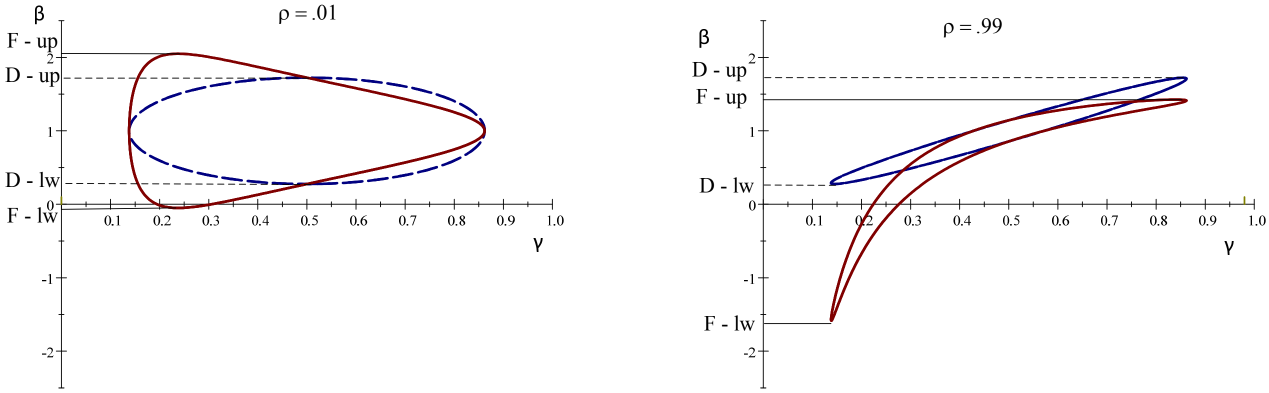


Figure 6 The Delta (dashed line) and Fieller (solid line) 95% CIs for β implied by the two constraint shapes for $T = 50$, $g = .5$, $h = .5$, $\beta = 1$, and $z = 1.96$. When $\rho = .99$ and $\rho = .01$.

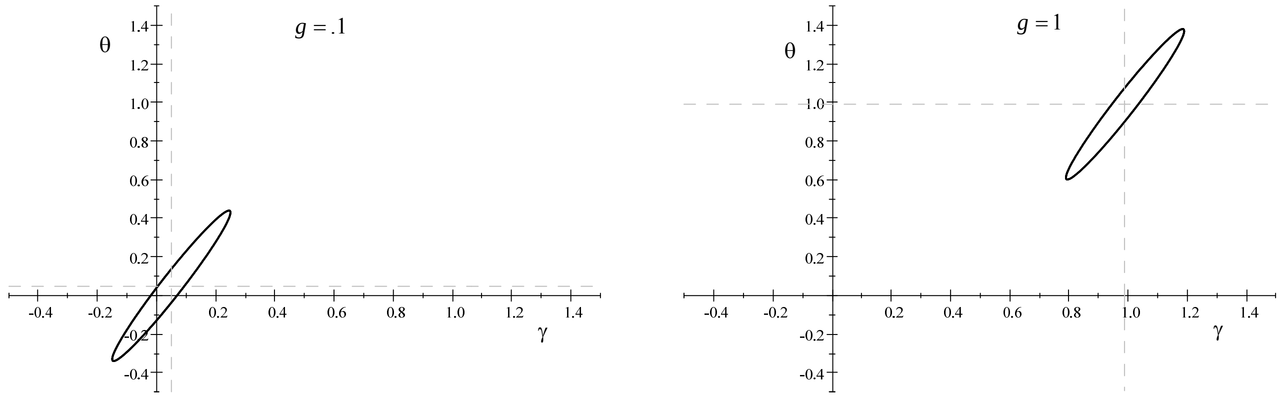


Figure 7. The 95% constraint ellipse in (θ, γ) space for $T = 100$, $g = h$, $\beta = 1$, $\rho = .9$, $z = 1.96$. When $g = 1$ and $g = .1$.

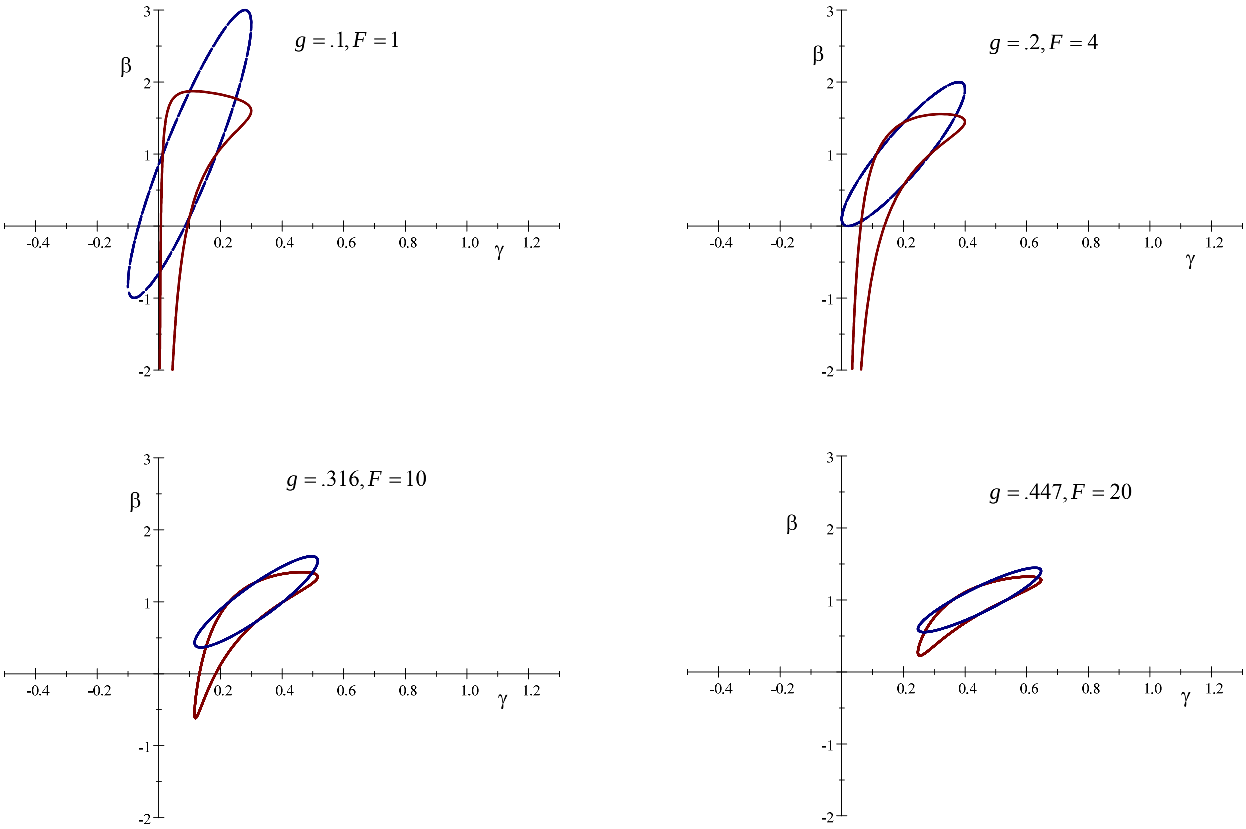


Figure 8. The Fieller and Delta constraint shapes in (γ,β) space for $T = 100$, $g = h$, $\beta = 1$, $\rho = .9$, $z = 1.96$. When g takes the values .1, .2, .316, and .447 that imply the equivalent for four values of the F-statistic for the first stage regression of 1, 4, 10 and 20.

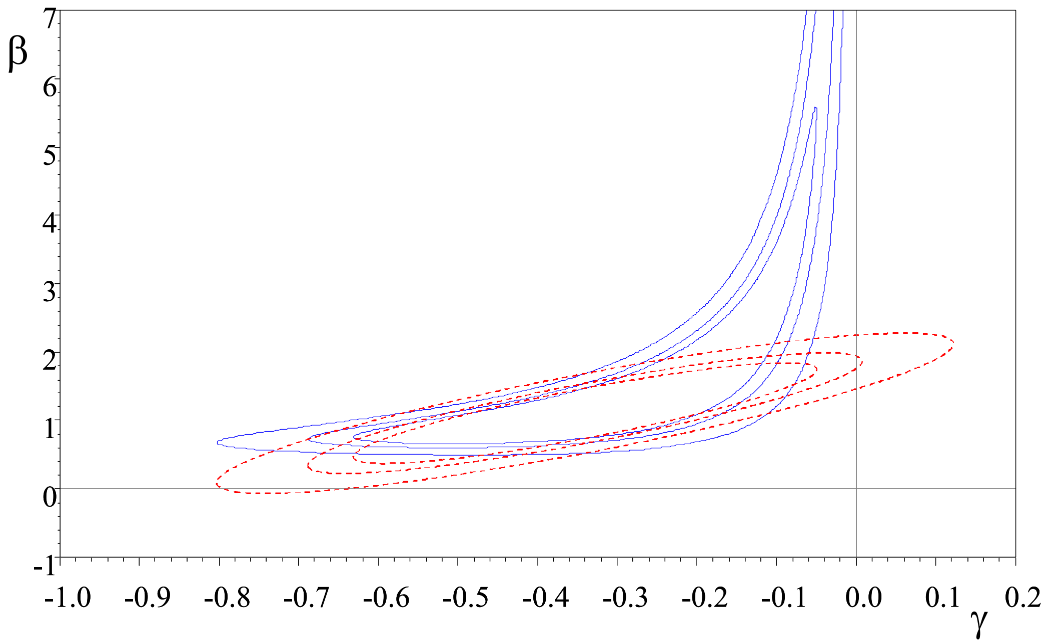


Figure 9. The Fieller (solid) and Delta (dotted) constraint shapes in (γ,β) space for varying values of α = .01, .05 and .10 as implied from the estimates reported in Table 2.

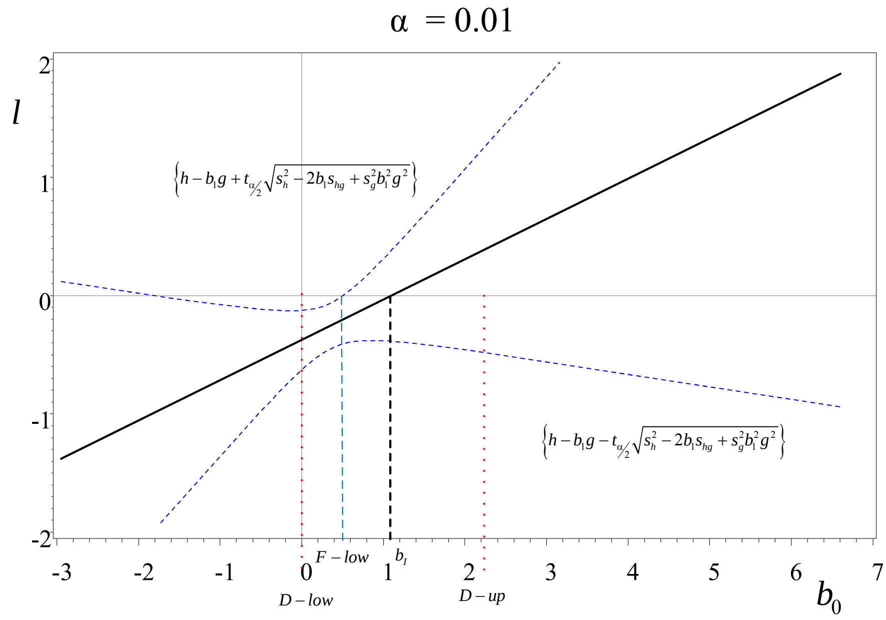


Figure 10. A line plot when α = .01 showing both the Fieller and Delta intervals for the results reported in Table 2.

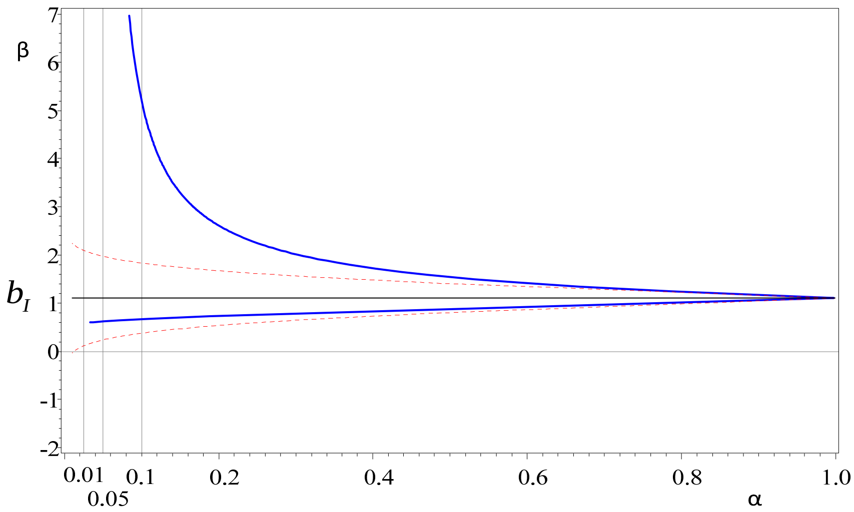


Figure 11 The confidence bounds for the Fieller (solid line) and Delta (dotted line) as a function of the value of α for the results reported in Table 2.

Tables

VARIABLES	
IPAE	1.107
	(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	1.440
	(0.533)
Observations	64

Table 1 The *2SLS* estimates for (35).

VARIABLES	*Log GDP per capita*	*IPAE*
Log European settler mortality	-0.377	-0.340
	(-3.868)	(-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	9.997	7.729
	(19.636)	(8.485)
Observations	64	64
R-squared	.584	.328
$\hat{\rho}_{hg}$	.609	

Table 2 The reduced form equation estimates for (36) and (37).