NBER TECHNICAL WORKING PAPER SERIES

SOME FURTHER RESULTS ON THE EXACT SMALL SAMPLE PROPERTIES OF THE INSTRUMENTAL VARIABLE ESTIMATOR

Charles R. Nelson

Richard Startz

Technical Working Paper No. 68

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 1988

This research is part of NBER's research program in Financial Markets and Monetary Economics. Any opinions expressed are those of the authors not those of the National Bureau of Economic Research.

NBER Technical Working Paper #68
September 1988

SOME FURTHER RESULTS ON THE EXACT SMALL SAMPLE PROPERTIES OF THE INSTRUMENTAL VARIABLE ESTIMATOR

ABSTRACT

New results on the exact small sample distribution of the instrumental variable estimator are presented by studying an important special case. The exact closed forms for the probability density and cumulative distribution functions are given. There are a number of surprising findings. The small sample distribution is bimodal, with a point of zero probability mass. As the asymptotic variance grows large, the true distribution becomes concentrated around this point of zero mass. The central tendency of the estimator may be closer to the biased least squares estimator than it is to the true parameter value. The first and second moments of the IV estimator are both infinite. In the case in which least squares is biased upwards, and most of the mass of the IV estimator lies to the right of the true parameter, the mean of the IV estimator is infinitely negative. The difference between the true distribution and the normal asymptotic approximation depends on the ratio of the asymptotic variance to a parameter related to the correlation between the regressor and the regression error. In particular, when the instrument is poorly correlated with the regressor, the asymptotic approximation to the distribution of the instrumental variable estimator will not be very accurate.

Charles R. Nelson
Department of Economics, DK-30
University of Washington
Seattle, WA 98195

Richard Startz
Department of Economics, DK-30
University of Washington
Seattle, WA 98195

In this paper we present new results on the exact small sample distribution of the instrumental variable estimator. In particular, we compare the small sample distribution to the asymptotic distribution.

Among our findings are:

- The central tendency of the instrumental variable estimator is biased away from the true value.
- The central tendency is biased in the direction of the probability limit of the ordinary least squares estimator.
- The distribution is bimodal.
- The distribution has a point of zero probability weight between the modes.
- When the asymptotic variance is large, the distribution is concentrated around the point of zero probability weight.
- The distribution does not have finite first or second moments.
- When the right–hand side variable and the regression error are uncorrelated, that is when ordinary least squares is the appropriate estimator, then the asymptotic approximation to the distribution of the instrumental variable gives the exact distribution.
- The asymptotic distribution will be a poor approximation to the true distribution when the instruments are poor, in the sense of not being highly correlated with the regressor, and when the number of observations is small.

While we prove each of these claims for the case of one independent variable and one instrument, we see no reason why the intuition should not be of more general applicability.

The Model

Consider the model of n observations on $y_i = \beta x_i + u_i$, where x_i and u_i are joint normal and in general correlated. Assume that the marginal distribution of u is i.i.d. normal. Assume further that there exists an instrument vector z such that z_i is a nonstochastic sequence, that $plim\ (\frac{1}{n}z'u) = 0$, and that $plim\ (\frac{1}{n}z'x)$ exists and is nonsingular. The instrumental variable estimator of β is $\hat{\beta} = (z'x)^{-1}z'y$. The asymptotic distribution of $\hat{\beta}$ is $N(\beta, \sigma_u^2\ plim\ ((z'x)^{-1}z'z(x'z)^{-1})$.

The small sample properties of the instrumental variable estimator for this particular problem have been considered by, among others, Basmann (1974), who summarizes a large body of work with particular respect to Haavelmo's model of the marginal propensity to consume, by Mariano and McDonald (1979), who give the *pdf* for $\hat{\beta}$, and by Anderson (1982). who discusses approximations to the *cdf*. Basmann and Mariano and McDonald point out that $\hat{\beta}$ is the ratio of two correlated normal random variables and so that its distribution may be studied using Fieller's (1932) results. (See also Johnson and Kotz (1972), pp. 123–124, Hinkley (1969) and Marsaglia (1965)). This paper extends the work just cited by characterizing the *pdf* and *cdf* of instrumental variables, and comparing them to the asymptotic approximations, as the "quality" of the instruments varies. In addition, we present an intuitive explanation of why the instrumental variable estimator may be a very poor one in small samples.

The body of the paper consists of two sections. In the first, we present a useful parameterization of the model just described and then, since it is both straightforward and instructive, we derive directly the exact density function and cumulative distribution function of the instrumental variable estimator. In the second, we consider the behavior of the distribution for various parameter values and prove the propositions stated above. To facilitate discussion, consider the following model describing generation of the data matrix:

$$y = \beta x + u$$

$$x = \epsilon + \lambda^{-1} u$$

$$z = v + \gamma\epsilon$$

where without loss of generality $\beta = 0$. Think of the triple $\{u_i, \epsilon_i, v_i\}$ as being i.i.d. normal with zero means and variances $\sigma_u^2, \sigma_\epsilon^2$, and σ_v^2, respectively. The sequence $\{\epsilon_i, v_i\}$, for $i = 1 \ldots n$, is taken to be fixed in repeated samples and, as a mathematical convenience, to be generated such that the sample second moments achieve their population values. (The assumption that the sample second moments achieve their population values simplifies notation and is of no importance for any of the results). Therefore, within a sample there is only a single random variable, u, which accounts for the stochastic behavior of both y and x. Given the assumptions on the sequence $\{\epsilon_i, v_i\}$, z is fixed in repeated samples and the sequence of sample second moments of z, σ_z^2, equals $\frac{1}{n}\sum_{i=1}^{n} z_i^2 = \sigma_v^2 + \gamma^2\sigma_\epsilon^2$.

Note that

$$plim\ \left(\frac{1}{n}z'x\right) = \gamma\sigma_\epsilon^2 \quad and \quad plim\ \left(\frac{1}{n}z'z\right) = \sigma_z^2$$

If we write

$$V \equiv \frac{\sigma_u^2}{n}\frac{\sigma_z^2}{\gamma^2\sigma_\epsilon^4}$$

then

$$\hat{\beta} \sim AN(0, V).$$

Note that λ^{-1} is the (population value of) the regression coefficient of x on u. When the regressor and the error term are uncorrelated, $\lambda^{-1} = 0$, ordinary least squares is asymptotically unbiased. It may be useful in thinking about this parameterization to look at the relation between λ and the asymptotic bias of ordinary least squares.

$$b = \ plim \ (\hat{\beta}_{OLS} - \beta) = \lambda \cdot \frac{\sigma_u^2}{\lambda^2 \sigma_\epsilon^2 + \sigma_u^2}$$

Thus $0 < \ plim \ (\hat{\beta}_{OLS}) < \lambda$. Note that as σ_u^2 grows large, V grows large and the asymptotic bias of ordinary least squares converges to λ.

I. The Exact Density and Distribution Functions of Instrumental Variables

The instrumental variable estimator can be written as the ratio of two normal random variables, $z'u$ and $z'x$. For convenience, write the sample moment of $z'u$ as

$$m_{zu} = \frac{1}{n} \sum_{i=1}^{n} z_i u_i.$$

Note that

$$m_{zu} \sim N(0, \frac{\sigma_u^2}{n} \sigma_z^2) = N(0 \gamma^2 \sigma_\epsilon^4 V).$$

Since $z'x = z'\epsilon + \lambda^{-1} z'u$ we can write the instrumental variable estimator as

$$\hat{\beta} = \frac{m_{zu}}{\gamma \sigma_\epsilon^2 + \lambda^{-1} m_{zu}}.$$

Figure 1 displays a graph of $\hat{\beta}$ as a function of m_{zu}. Note that while $\hat{\beta}$ is neither continuous nor monotonic in m_{zu}, there is nonetheless a one–to–one and onto correspondence and that the function is differentiable everywhere except at the single discontinuity.

Since m_{zu} follows a normal density, it is straightforward to derive the density of $\hat{\beta}$ by change of variables. If m_{zu} has the density function $f_m(m_{zu})$, then $\hat{\beta}$ has the density function $f_m(m_{zu})\ \frac{dm}{d\hat{\beta}}$, where $\frac{dm}{d\hat{\beta}} = \frac{\gamma\sigma_\epsilon^2}{(1-\hat{\beta}/\lambda)^2}$. The density function for m_{zu} is given by

$$f(m_{zu}) = (2\pi\gamma^2\sigma_\epsilon^4 V)^{-\frac{1}{2}} exp\ [-\frac{1}{2\gamma^2\sigma_\epsilon^2 V} m_{zu}^2]$$

Therefore, the exact density of $\hat{\beta}$ is given by:

$$f(\hat{\beta}) = \frac{1}{(1-\hat{\beta}/\lambda)^2}\frac{1}{\sqrt{2\pi V}} \quad exp\ [-\frac{1}{2V}(\frac{\hat{\beta}}{1-\hat{\beta}/\lambda})^2] \qquad (1)$$

as compared to the asymptotic approximation given by:

$$f^A(\hat{\beta}) = \frac{1}{\sqrt{2\pi V}} \quad exp\ [-\frac{1}{2V}\hat{\beta}^2] \qquad (2)$$

Note, while the asymptotic distribution depends only on the asymptotic variance, V, the true distribution depends on two parameters, V and λ. In fact, the true distribution is homogeneous of degree one in λ, V, and $\hat{\beta}$. We call λ the "point of concentration," for reasons that will become apparent shortly.

Turn now to the derivation of the cumulative distribution function for $\hat{\beta}$. Since the mapping between $\hat{\beta}$ and m_{zu} is one–to–one and onto, the probability of $\hat{\beta}$ lying in a given interval is just the probability of m_{zu} lying in the corresponding interval of the normal. The *cdf* of $\hat{\beta}$ is defined piecewise according to whether $\hat{\beta}$ lies to the right of the singularity in Figure 1 ($\hat{\beta} > \lambda$) or to the left ($\hat{\beta} < \lambda$). If we write $\hat{\beta} = g(m_{zu})$, then, for

$\hat{\beta} < \lambda$, *prob* $(\hat{\beta} < \theta) =$ *prob* $(m_{zu} < g^{-1}(\theta)) -$ *prob* $(m_{zu} < -\lambda\gamma\sigma_\epsilon^2)$. For $\hat{\beta} > \lambda$, *prob* $(\hat{\beta} < \theta) =$ *prob* $(m_{zu} > -\lambda\gamma\sigma_\epsilon^2) +$ *prob* $(m_{zu} < g^{-1}(\theta))$. The *cdf* of m_{zu} is

$$\Phi(\frac{m_{zu}}{\sqrt{\gamma^2\sigma_\epsilon^4 V}})$$

where Φ is the standard normal *cdf*. Making the appropriate substitutions, the *cdf* of the instrumental variables estimator is as given in (3).

$$\begin{aligned} For \quad \beta < \lambda, F(\hat{\beta}) &= \Phi(\frac{\hat{\beta}}{1-\hat{\beta}/\lambda}\frac{1}{\sqrt{V}}) - \Phi(\frac{-\lambda}{\sqrt{V}}) \\ For \quad \beta > \lambda, F(\hat{\beta}) &= 1 - \Phi(\frac{-\lambda}{\sqrt{V}}) + \Phi(\frac{\hat{\beta}}{1-\hat{\beta}/\lambda}\frac{1}{\sqrt{V}}) \end{aligned} \tag{3}$$

In comparison, the *cdf* of the asymptotic distribution is $F^A(\hat{\beta}) = \Phi(\frac{\hat{\beta}}{\sqrt{V}})$.

II. Characterization of the Distribution of the Instrumental Variables Estimator

In this section we characterize the shape of the distribution of the instrumental variable estimator. The characterization takes two forms. We compare the actual distribution to the familiar centered–around–zero bell curve of the asymptotic distribution. We also prove a number of propositions which describe how the instrumental variable distribution changes when the parameters λ and V change.

Figure 2 shows a representative sample of the density of $\hat{\beta}$, with the corresponding asymptotic approximation drawn for comparison. The parameter values are $\lambda = 1$ and $V = 4$. The region shown is ± 2 asymptotic standard deviations around λ. For the particular parameter values, this region includes 62.47 percent of the mass of the asymptotic distribution and 82.52 percent of the mass of the true distribution.

The first striking characteristic about the picture of the true distribution is that it is bimodal and has a minimum at λ.

The derivative of the density of $\hat{\beta}$ is given in (4).

$$f'(\hat{\beta}) = \frac{f(\hat{\beta})}{1-\beta/\lambda} \cdot [\frac{2}{\lambda} - \frac{\beta}{V(1-\beta/\lambda)^2}]. \tag{4}$$

$\hat{\beta}$ has three critical points. There are two maxima, so the distribution is bimodal. The modes are given by setting $f' = 0$ and applying the quadratic formula

$$modes\ (\frac{\hat{\beta}}{\lambda}) = 1 + \frac{1}{4V} \pm \sqrt{(1+\frac{1}{4V})^2 - 1} \tag{5}$$

Note, from inspection of (5), that one mode occurs between zero and λ and the other to the right of λ. As V goes to ∞, the right–side of (5) goes to one, so both modes approach λ. As $V \to 0$, the modes go to zero and ∞ respectively.

If the position of the modes relative to λ depends on V, what does V depend on? First, for any fixed parameters, as the sample size grows, V goes to zero. If one envisions watching the modes spreading towards zero and ∞ as the sample size grows, then one sees the process of convergence in distribution of $f(\hat{\beta})$ to the asymptotic distribution. Second, in a finite sample, if the instrument is poorly correlated with the regressor, then V will be large. For example, in our parameterization as γ becomes smaller in absolute value, V grows without limit. The conventional wisdom is that with poor instruments V is large so that the asymptotic distribution of $\hat{\beta}$ is dispersed, as illustrated in Figure 2. However, with large V, the asymptotic approximation is a poor approximation. The distribution of $\hat{\beta}$ may be quite concentrated, though it has fat tails, around a point away from the true parameter value.

Figure 3 shows the distribution of $\hat{\beta}$ for a case in which V is large relative to λ. V is set to 16 in Figure 3, as compared to 4 in Figure 2. The region of $\hat{\beta}$ shown is the same in Figures 2 and 3. Asymptotic distribution theory suggests that the fraction of the probability mass shown in Figures 2 and 3 falls from .62 to .37. In fact, the fraction rises from .82 to .90. The small sample distribution is becoming more concentrated, not less, as V rises.

We can turn the question around and ask how much of the mass lies within $\pm$ 1.96 asymptotic standard deviations of zero. For the asymptotic distribution, the answer is always .95. The correct answer depends on both λ and V. In Figure 2 the true mass is .91, while in Figure 3 the true mass is .97.

Figure 4 shows the distribution of $\hat{\beta}$ for a small asymptotic variance ($V = \frac{1}{2}$). Here, the asymptotic approximation is better than in Figure 2. According to asymptotic distribution theory, the fraction of the probability mass shown rises from .62 to .92. In fact, the fraction falls from .82 to .78. In Figure 4, .79 of the true mass lies within $\pm$ 1.96 asymptotic standard deviations of zero.

In general, the accuracy of "confidence intervals" based on the asymptotic distribution depends on λ, V, and the location of the region under consideration.

We turn now to several lemmas which more precisely characterize the distribution of $\hat{\beta}$.

Lemma 1: When the regressor and the regression error are uncorrelated, so that ordinary least squares is the best linear asymptotically unbiased estimator, the asymptotic approximation is actually the true density.

Proof: This is the case where $\lambda^{-1} = 0$. The proof is by inspection of (1) and (2).

Lemma 2: The point of concentration, λ, is a point of zero mass.

Proof: As $\hat{\beta} \to \lambda, f(\hat{\beta}) \to 0$, by l'Hôpital's rule.

Note, by inspection of (4), that a minimum occurs at λ,where $f'(\lambda) = 0$.

Lemma 3: At the point of concentration, the actual and true *cdf*s are equal, $F(\lambda) = F^A(\lambda)$.

Proof: Consider the first line of (3). As $\hat{\beta} \to \lambda$ from the left,

$$\Phi(\frac{\hat{\beta}}{1-\hat{\beta}/\lambda}\frac{1}{\sqrt{V}}) - \Phi(\frac{-\lambda}{\sqrt{V}}) \to \Phi(\infty) = \Phi(\frac{-\lambda}{\sqrt{V}}) = \Phi(\frac{\lambda}{\sqrt{V}})$$

Corollary 3.1: $0 < \quad Median(\hat{\beta}) < \lambda$.

Proof: From the first line of (3), $F(0) = \Phi(0)(\frac{-\lambda}{\sqrt{V}}) < \frac{1}{2}$ and, by the preceding lemma, $F(\lambda) = \Phi(\frac{\lambda}{\sqrt{V}}) > \frac{1}{2}$.

Corollary 3.2: As $V \to \infty$, the median of $\hat{\beta} \to \lambda$.

Proof: By the preceding lemma, as $V \to \infty F(\lambda) = \Phi(\frac{\lambda}{\sqrt{V}}) \to \Phi(0) = \frac{1}{2}$.

Lemma 4: As $V \to \infty$, the distribution of $\hat{\beta}$ becomes concentrated around λ in the sense that $\lim_{V\to\infty} Prob\ (\lambda - \theta < \hat{\beta} < \lambda + \theta) = 1$, for all $\theta > 0$.

Proof:

$$\begin{aligned} Prob(\lambda - \theta < \hat{\beta} < \lambda + \theta) &= F_{\hat{\beta}>\lambda}(\lambda+\theta) - F_{\hat{\beta}<\lambda}(\lambda-\theta) \\ &= \{1 - \Phi(\frac{-\lambda}{V}) + \Phi(\frac{\lambda+\theta}{1-(\lambda+\theta)/\lambda}\frac{1}{\sqrt{V}})\} \\ &- \{\Phi(\frac{\lambda-\theta}{1-(\lambda-\theta)/\lambda}\frac{1}{\sqrt{V}}) - \Phi(\frac{-\lambda}{\sqrt{V}})\} \\ &= 1 + \Phi(-(\frac{\lambda^2}{\theta}+\lambda)\frac{1}{\sqrt{V}}) - \Phi((\frac{\lambda^2}{\theta}-\lambda)\frac{1}{\sqrt{V}}) \end{aligned}$$

But as $V \to \infty$, the values of each of the *cdfs* in the last expression go to one–half, cancelling one another, so the probability goes to one.

Lemma 5: The first and second moments of the distribution of $\hat{\beta}$ are infinite. In particular, the mean is infinitely negative.

Proof: We present the proof for the first moment first. Some intuition may be gained by looking at Figure 1. Just to the left of the discontinuity, $\hat{\beta}$ takes on infinite positive values. Just to the right of the discontinuity, $\hat{\beta}$ takes on infinite negative values. Since there is more mass of m_{zu} closer to zero, the negative values outweigh the positive ones.

The mean of $\hat{\beta}$ is $\int_{-\infty}^{\infty} \hat{\beta} f(\hat{\beta}) d\beta$. Define $\delta = \frac{\beta}{1-\beta/\lambda}$. Noting that $\beta = \frac{\delta\lambda}{\lambda+\delta}$, that $d\beta = [\frac{\lambda}{\lambda+\delta}]^2 d\delta$, and appropriately rearranging the limits of integration, we can make the change of variables to find the mean of $\hat{\beta}$ equals

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi V}} \frac{\lambda\delta}{\lambda+\delta} exp[\frac{-1}{2V}\delta^2] d\delta \tag{6}$$

Divide definite integral in (6) into three regions: $(-\infty, -2\lambda), (-2\lambda, 0)$, and $(0, \infty)$. Over the first and third region, the integrand is everywhere positive and $exp[-\frac{1}{2V}\delta^2]$ goes to zero rapidly, so the integral in these regions is positive and finite. Subdivide the middle region into the subregions $(-2\lambda, -\lambda)$ and $(-\lambda, 0)$. Over the first subregion perform the change of variable $s = -\lambda - \delta$ and over the second subregion perform the change of variable $s = -\lambda + \delta$. The integral becomes

$$\begin{aligned} &\int_{\lambda}^{0} \frac{1}{\sqrt{2\pi V}} \frac{\lambda(s+\lambda)}{s} exp[-\frac{1}{2V}(s+\lambda)^2](-1)ds \\ &+ \int_{0}^{\lambda} \frac{1}{\sqrt{2\pi V}} \frac{\lambda(s-\lambda)}{s} exp[-\frac{1}{2V}(s-\lambda)^2]ds \\ &= \frac{\lambda}{\sqrt{2\pi V}} \int_{0}^{\lambda} \frac{s+\lambda}{s} exp[-\frac{1}{2V}(s+\lambda^2)] \\ &+ \frac{(s-\lambda)}{s} exp[-\frac{1}{2V}(s-\lambda)^2]ds \end{aligned} \tag{7}$$

The first term in the integrand of (7) is positive within the limits of integration and the second term is negative. The first term is less than $\frac{s+\lambda}{s} exp[-\frac{1}{2V}\lambda^2]$, while the second term is greater in absolute value than $\frac{s-\lambda}{s} exp[-\frac{0}{2V}]$. Thus, the second, negative, integrand dominates. The definite integral of $-\frac{\lambda}{s}$ is

$$[-\lambda \; log \; s]_0^\lambda = -\lambda[log \; \lambda - log \; 0] = -\infty.$$

q.e.d.

The proof for the second moment is somewhat easier. The t^{th} moment of $\hat{\beta}$ is

$$\int_{-\infty}^{\infty} m(t) \; d\hat{\beta}$$

where

$$m(t) = \hat{\beta}^{t-2}(\frac{\hat{\beta}}{1-\hat{\beta}/\lambda})^2 \frac{1}{\sqrt{2\pi V}} exp \; [\frac{-1}{2V}(\frac{\hat{\beta}}{1-\hat{\beta}/\lambda})^2]$$

For $t = 2$, m is everywhere positive and approaches a constant as $\hat{\beta} \rightarrow \infty$. Therefore, the integral is infinite. (The argument holds for higher–order even moments, a fortiori.)

Intuition

We have, with regard to the location of the probability mass of $\hat{\beta}$, three sets of results which appear to be at odds with one another. First, asymptotic distribution theory asserts that the distribution of $\hat{\beta}$ is approximately bell–shaped and centered around zero. Second, the absence of finite moments of $\hat{\beta}$ suggests that the distribution is fat–tailed. Third, we have shown that as $V \rightarrow \infty$, the mass concentrates around λ. How can these three statements be reconciled?

Return to Figure 1, which shows the correspondence between m_{zu} and $\hat{\beta}$. Since m_{zu} is normal, its mass *is* bell–shaped and centered around zero. Suppose V is small, so the variance of m_{zu} is small, then most of the mass will be close to zero, say in the region marked AA. In this region, the mapping from m_{zu} to $\hat{\beta}$ is approximately linear, so a normal distribution on m_{zu} induces a normal distribution, centered around zero, on $\hat{\beta}$. Thus for a small asymptotic variance, the asymptotic distribution is a good approximation.

Suppose that V is somewhat larger, so that most of the mass of m_{zu} falls in the region marked BB. In this case, a significant portion of the mass of m_{zu} lies near the singularity, inducing values of $\hat{\beta}$ lying far out in the tails. This explains why the moments do not exist.

Finally, suppose that V is larger still, so that most of the mass of m_{zu} falls in the region marked CC. In this case, $\hat{\beta}$ is almost always close to the point of concentration, λ. Thus as V grows large, most of the mass of $\hat{\beta}$ concentrates around λ.

Conclusion

We have shown that the true distribution of the instrumental variable estimator looks very little like the asymptotic approximation. In the case we study, the distribution is bimodal, fat–tailed, and may be heavily concentrated around a point closer to the probability limit of least squares than to the true parameter estimate. In a companion paper, we use some of the analytical results presented here together with Monte Carlo studies to look at the distribution of test statistics based on instrumental variable estimation.

REFERENCES

Anderson, T.W. "Some Recent Developments on the Distributions of Single-Equation Estimators." In *Advances in Econometrics*, W. Hildebrand, ed., pp. 109–121, 1982.

Basmann, R.L., "Exact Finite Sample Distributions for Some Econometric Estimators and Test Statistics: A Survey and Appraisal," in *Frontiers of Quantitative Economics*, Vol. II, M.D. Intrilligator and D.A. Kendrick, eds., pp. 209–288.

Fieller, E.C., "The Distribution of the Index in a Normal Bivariate Population," *Biometrica* 24, 428–440, 1932.

Hinkley, D.V., "On the Ratio of Two Correlated Normal Random Variables," *Biometrika* 56, 3, pp. 635–639, 1969.

Johnson, N.L. and S. Kotz, *Distributions in Statistics: Continuous Multivariate Distributions*, John Wiley and Sons, Inc., New York, 1972.

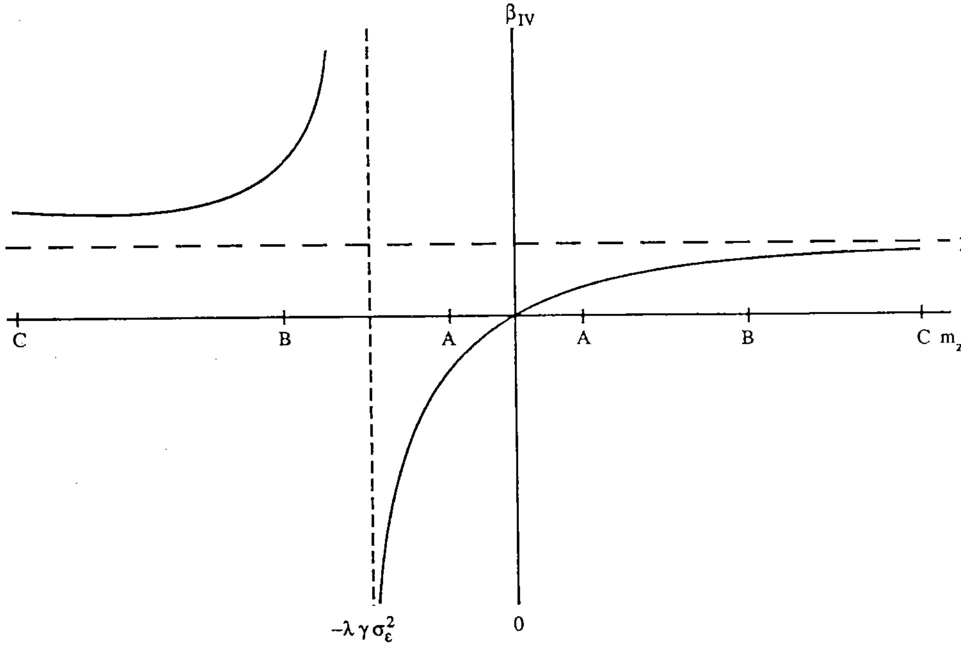
Mariano, R.S. and J.B. McDonald, "A Note on the Distribution Functions of LIML and 2SLS Structural Coefficient in the Exactly Identified Case," *Journal of the American Statistical Association*, No. 368, Vol. 74, pp. 847–848.

Marsaglia, G., "Ratios of Normal Variables and Ratios of Sums of Uniform Variables," *Journal of the American Statistical Association* 60, pp. 193–204, 1965.

Nelson, C.R. and R. Startz, "The Distribution of the Instrumental Variable Estimator and Its t–Ratio When the Instrument is a Poor One," mimeo, University of Washington, 1988.

Figure 1

β_{IV} as a Function of m_{zu}



$$\beta_{IV} = \frac{m_{zu}}{\gamma + \lambda^{-1} m_{zu}}$$

True and Asymptotic Density Functions for Instrumental Variables Lambda=1 V=4

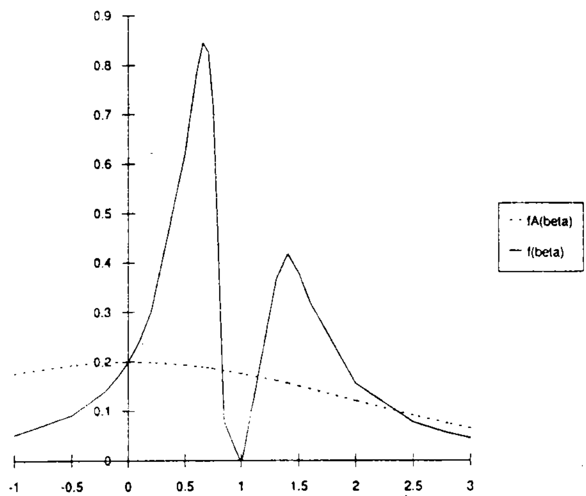


Figure 2

True and Asymptotic Density Functions for Instrumental Variables Lambda=1 V=16

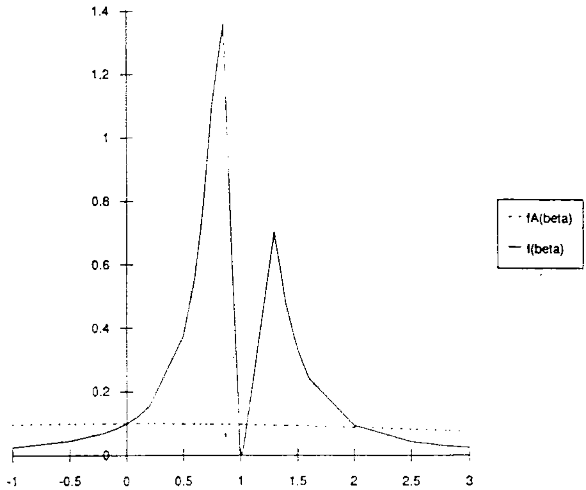


Figure 3

True and Asymptotic Density Functions for Instrumental Variables Lambda=1 V=1/2

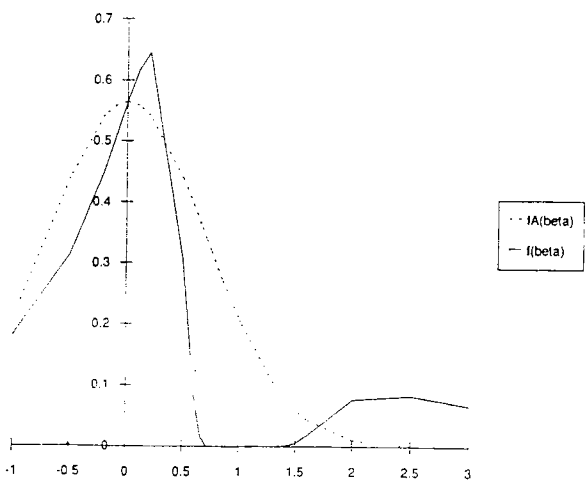


Figure 4