Centre for Efficiency and Productivity Analysis

Working Paper Series
No. WP02/2007

Title

On The Distribution of Estimated Technical Efficiency in Stochastic Frontier Models

Authors
Wei Siang Wang & Peter Schmidt

Date: May, 2007

School of Economics
University of Queensland
St. Lucia, Qld. 4072
Australia

ISSN No. 1932 - 4398

ON THE DISTRIBUTION OF ESTIMATED TECHNICAL EFFICIENCY

IN STOCHASTIC FRONTIER MODELS

Wei Siang Wang
Michigan State University

Peter Schmidt
Michigan State University

May, 2007

1. INTRODUCTION

In this paper we consider the stochastic frontier model introduced by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977). We write the model as

(1) $$y_i = X_i\beta + \varepsilon_i \quad , \quad \varepsilon_i = v_i - u_i \quad , \quad u_i \geq 0 \ .$$

Here typically y_i is log output, X_i is a vector of input measures (e.g., log inputs in the Cobb-Douglas case), v_i is a normal error with mean zero and variance σ_v^2, and $u_i \geq 0$ represents technical inefficiency. Technical efficiency is defined as $TE_i = \exp(-u_i)$, and the point of the model is to estimate u_i or TE_i.

A specific distributional assumption on u_i is required. The papers cited above considered the case that u_i is half normal (that is, it is the absolute value of a normal with mean zero and variance σ_u^2) and also the case that it is exponential. Other distributions proposed in the literature include general truncated normal (Stevenson (1980)) and gamma (Greene (1980a, 1980b, 1990) and Stevenson (1980)). In this paper we will consider only the half normal case, but similar results would apply to the other cases. Also, our exposition is for the cross-sectional case, but we could also consider panel data as in Pitt and Lee (1981).

Define $\hat{\beta}$ to be the MLE of β, and $\hat{\varepsilon}_i = y_i - X_i\hat{\beta}$. Then the usual estimate of u_i, suggested by Jondrow et al. (1982), is $E(u_i|\varepsilon_i)$, evaluated at $\varepsilon_i = \hat{\varepsilon}_i$. We can estimate TE_i by $\widehat{TE_i} = \exp(-\hat{u}_i)$ but a preferred estimate is $\widetilde{TE_i} = E\{\exp(-u_i)|\varepsilon_i\}$ evaluated at $\varepsilon_i = \hat{\varepsilon}_i$. See Battese and Coelli (1988), who also show how to define $\hat{u}_i$ and $\widetilde{TE_i}$ in the case of panel data.

In this paper we derive the distribution of $\hat{u}_i$. (The same method of derivation would also apply to $\widetilde{TE}_i$, though we do not give the details.) It is important to realize that this is not, and should not be expected to be, the same as the distribution of u_i. In other words, if one assumes that the u_i are half normal, it is tempting to look at the $\hat{u}_i$ and see if their distribution looks half normal. It should not, unless σ_v^2 is very small. We show that the distribution of $\hat{u}_i$ becomes the same as the distribution of u_i as $\sigma_v^2 \to 0$ (with σ_u^2 fixed), and that the distribution of $\hat{u}_i$ collapses on the point $E(u)$ as $\sigma_v^2 \to \infty$. We also graph the distribution for intermediate values of σ_v^2.

One way to understand the difference between the distributions of $\hat{u}_i$ and u_i is to realize that $\hat{u}_i$ is a shrinkage of u_i toward its mean. This reflects the familiar principle that an optimal (conditional expectation) forecast is less variable than the thing being forecast. The usual breakdown of variance into explained and unexplained parts says:

(2) $$\text{var}(u_i) = \text{var}[E(u_i|\varepsilon_i)] + E[\text{var}(u_i|\varepsilon_i)]$$

so that $\text{var}(u_i)$ is greater than $\text{var}(\hat{u}_i)$ by the amount $E[\text{var}(u_i|\varepsilon_i)]$.[1] An implication of shrinkage is that on average we will overestimate u_i when it is small, and underestimate u_i when it is large. To see the exact sense in which this is true, we also derive the distribution of $\hat{u}_i$ conditional on u_i. We show that as $\sigma_v^2 \to 0$ (with σ_u^2 fixed), the distribution of $\hat{u}_i$ conditional on u_i collapses on u_i,

[1] The expectation is over the distribution of the conditioning variable, ε_i.

while as $\sigma_v^2 \to \infty$, the distribution of $\hat{u}_i$ conditional on u_i does not depend on u_i (it collapses on the point $E(u)$). Once again we graph the distribution for intermediate values of σ_v^2, for various values of u_i.

The plan of the paper is as follows. Section 2 considers the distribution of $\hat{u}_i$. Section 3 considers the distribution of $\hat{u}_i$ conditional on u_i. Section 4 gives our concluding remarks. There is also an Appendix which contains some of the derivations.

2. THE DISTRIBUTION OF $\hat{u}$

In this section we derive and discuss the distribution of $\hat{u}_i = E(u_i | \varepsilon_i)$. This is a random variable because it is a function of ε_i, which is a random variable, and its distribution follows from the distribution of ε_i.

Our discussion will ignore estimation error in β. That is, we consider $\hat{u}_i = E(u_i | \varepsilon_i)$, whereas in practice $\hat{u}_i = E(u_i | \varepsilon_i)$ *evaluated at* $\varepsilon_i = \hat{\varepsilon}_i$. The difference between ε_i and $\hat{\varepsilon}_i$ is that $\varepsilon_i = y_i - X_i\beta$ whereas $\hat{\varepsilon}_i = y_i - X_i\hat{\beta}$; that is, the difference is just the contribution of estimation error in β. The justification for ignoring this is that, in any application we can envision, the intrinsic randomness in $E(u_i | \varepsilon_i)$ due to its being a function of ε_i will dwarf the randomness due to estimation error in β. More formally, the former is $O_p(1)$ while the latter is $O_p(1/\sqrt{N})$. Also, for notational simplicity, we will henceforth omit subscript "i" from $\hat{u}, u, v$ and ε.

Since $\hat{u} = E(u | \varepsilon)$ it is a function of ε, and we can write $\hat{u} = h(\varepsilon)$. The function h was given by Jondrow et al. (1982):

(3) $$\hat{u} = h(\varepsilon) = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}[-\varepsilon + \sigma_0 \bullet \lambda(\varepsilon / \sigma_0)] \quad \text{, where } \sigma_0^2 = (\sigma_u^2 + \sigma_v^2) \bullet \sigma_v^2 / \sigma_u^2 \ ,$$

$\lambda(s) = \phi(s) / [1 - \Phi(s)]$, and where ϕ and Φ are the standard normal density and cdf, respectively.

The function *h* is a monotonic (strictly decreasing) function, so it can be inverted. That is, we can formally write

(4) $$\varepsilon = h^{-1}(\hat{u}) = g(\hat{u}) \quad .$$

We cannot express the function *g* analytically, but it is well defined and we can calculate it. For example, Figure 1 shows the function *g* for the case that $\sigma_u^2 = \sigma_v^2 = 1$.

Let f_ε and $f_{\hat{u}}$ represent the densities of ε and $\hat{u}$. Then making the simple change of variables in (4), we have

(5) $$f_{\hat{u}}(\hat{u}) = f_\varepsilon(g(\hat{u})) \bullet |g'(\hat{u})| \quad .$$

The density of ε is given by Aigner, Lovell and Schmidt:

(6) $$f_\varepsilon(\varepsilon) = (2/a) \bullet \phi(\varepsilon / a) \bullet \Phi(-\varepsilon b / a) \quad , \quad a = \sqrt{\sigma_u^2 + \sigma_v^2} \quad , \quad b = \sigma_u / \sigma_v \quad .$$ [2]

Also, we can calculate the Jacobian term $|g'(\hat{u})|$. We show in Appendix A that

(7) $$g'(\hat{u}) = \frac{a^2}{\sigma_u^2 \bullet [-1 + \lambda'(g(\hat{u}) / \sigma_0)]} \quad \text{, where } \lambda'(s) = -s\lambda(s) + \lambda^2(s) \ .$$

---

[2] This notation is slightly different from Aigner, Lovell and Schmidt. Our *a* is their σ and our *b* is their λ. But we have already used λ for the inverse Mill's ratio, and there are enough different σ's already without introducing another one.

So, substituting (6) and (7) into (5), we obtain

(8) $$f_{\hat{u}}(\hat{u}) = \frac{2a\bullet\phi(g(\hat{u})/a)\bullet\Phi(-g(\hat{u})b/a)}{\sigma_u^2\left|-1+\lambda'(g(\hat{u})/\sigma_0)\right|} .$$

Clearly this is not the same as f_u, the half normal density.

The following result shows what happens in the limit as σ_v^2 approaches zero and infinity, respectively. The proof is given in Appendix B.

THEOREM 1:

(1) As $\sigma_v^2 \to 0$, $(\hat{u}-u) \to_p 0$.

(2) As $\sigma_v^2 \to 0$, $f_{\hat{u}} \to f_u$ (pointwise).

(3) As $\sigma_v^2 \to \infty$, $\hat{u} \to_p E(u)$.

(4) As $\sigma_v^2 \to \infty$, $[\pi/(\pi-2)]\bullet(\sigma_v/\sigma_u^2)\bullet(\hat{u}-E(u)) \to_d N(0,1)$.

These results make sense if we realize that we are treating $\varepsilon = v - u$ as our observable quantity. If $\sigma_v^2 = 0$, so that $v \equiv 0$, we effectively observe *u*, and so in the limit $\hat{u} = u$ and the distribution of $\hat{u}$ equals the distribution of *u*. Conversely, when $\sigma_v^2 = \infty$, ε contains no useful information about *u*, and the best estimate of *u* is simply $\hat{u} = E(u)$. Part (4) says that, for large σ_v^2, $\hat{u}$ is approximately normally distributed around $E(u)$, with variance $[(\pi-2)/\pi]^2\bullet(\sigma_u^4/\sigma_v^2)$.

For values of σ_v^2 between zero and infinity, the density of $\hat{u}$ represents the shrinkage of u towards its mean, which is $\sqrt{(2/\pi)}\bullet\sigma_u$, or about $0.80\bullet\sigma_u$.[3] Figure 2 displays the half normal density of u, with σ_u = 1; this corresponds to the density of $\hat{u}$ when $\sigma_v^2 = 0$. Figures 3, 4, 5 and 6 give the density of $\hat{u}$ when σ_v^2 = 0.1, 1, 10 and 100, respectively. None of these densities looks much like the half normal. Comparing the densities in the different figures requires some care, since the axes are scaled differently. However, it is clearly the case that, as σ_v^2 increases, the density of $\hat{u}$ becomes more peaked and concentrated more tightly about the mean of 0.80. As σ_v^2 becomes large, the distribution of $\hat{u}$ collapses onto the point E(u), as indicated in part (3) of Theorem 1. The approximate normality of the distribution of $\hat{u}$ for large σ_v^2 is evident in Figure 6.

Finally, Figure 7 contains all of the five graphs that were in Figures 2 through 6. The use of a common set of axes makes it hard to see the detail in any one of the graphs, but seeing them all together does make clear what happens as σ_v^2 changes.

3. THE DISTRIBUTION OF $\hat{u}$ CONDITONAL ON *u*

In the previous section, we saw that the distribution of $\hat{u}$ is a shrinkage toward the mean of the distribution of *u*. Intuitively, this means that we should expect that on average we will overestimate small realizations of *u* and underestimate large ones. To see the precise sense in which this is true, in this section we derive and graph the density of $\hat{u}$ conditional on *u*.

The density of $\hat{u}$ conditional on *u* is given by the following equation.

[3] Note that, by the law of iterated expectations, the mean of $\hat{u}$ is the same as the mean of *u*.

(9) $$f(\hat{u}|u) = \frac{a^2 \bullet \exp[-(1/2\sigma_v^2)(g(\hat{u})+u)^2]}{\sqrt{(2\pi)} \bullet \sigma_u^2 \bullet \sigma_v^2 \bullet \left|-1+\lambda'(g(\hat{u})/\sigma_0)\right|}$$

The derivation is given in Appendix C.

Theorem 1 above gives some guidance as to what we should expect this density to look like. As $\sigma_v^2 \to 0$, the distribution of $\hat{u}$ conditional on u should collapse onto the point u. Conversely, as $\sigma_v^2 \to \infty$, the distribution of $\hat{u}$ conditional on u no longer depends on u; it collapses onto the point $E(u)$.

The following result shows that, approximately normalized, $\hat{u}$ conditional on u is asymptotically normal both as $\sigma_v^2 \to 0$, and as $\sigma_v^2 \to \infty$. (The normalization obviously must differ in the two cases.) The proof is given in Appendix D.

THEOREM 2:

(1) As $\sigma_v^2 \to 0$, $\frac{\hat{u}-u}{\sigma_v} \to_d N(0,1)$.

(2) As $\sigma_v^2 \to \infty$, $(\frac{\pi}{\pi-2}) \bullet (\frac{\sigma_v}{\sigma_u^2}) \bullet (\hat{u}-u) \to_d N(0,1)$.

Results (1) and (2) hold treating u as fixed. That is, they deal with the distribution of $\hat{u}$ conditional on u. Result (2) is, however, the same as the unconditional result given in result (4) of Theorem 1.

Figures 8 – 13 give the density of $\hat{u}$ conditional on u, for u = 0.1, $\sigma_u^2 = 1$, and σ_v^2 = 0.001, 0.01, 0.1, 1, 10 and 100. The value u = 0.1 is a small value (in the left tail of the distribution) and so we expect to overestimate it, on average. This does occur except perhaps for the very smallest value of σ_v^2. We do not have a strict shrinkage to the mean, in the sense that there *is* probability mass for $\hat{u}$ to the left of the true value of *u*, but except when σ_v^2 is very small the vast majority of the probability mass is to the right of *u*. For the larger values of σ_v^2 most of the probability mass is near the mean, *E*(*u*). The approximate normality of the distribution of $\hat{u}$ conditional on *u* for small σ_v^2 and for large σ_v^2 can be seen in Figures 8 and 13, respectively. For intermediate values of σ_v^2 the distribution does not look normal.

Figures 14-19 give the same results, but now for the case that $u = 2$. The value $u = 2$ is a large value (in the right tail of the distribution) and so we expect to underestimate it, on average. This does occur, and again the amount of shrinkage to the mean is small when σ_v^2 is small and large when σ_v^2 is big.

Figure 20 illustrates the point that, when σ_v^2 is large enough, the density of $\hat{u}$ conditional on *u* no longer depends on *u*. In Figure 20 we have $\sigma_u^2 = 1$ and $\sigma_v^2 = 100$, and we display the density of $\hat{u}$ conditional on *u* for $u = 0.1$, 0.5, 1 and 2. These densities are not much different. With enough noise, the data are no longer very relevant in estimating *u*, or equivalently the estimate is not very different depending on the true value of *u* that generated the data.

4. CONCLUDING REMARKS

This paper derived the distribution of the technical efficiency estimate $\hat{u} = E(u|\varepsilon)$, and also the distribution of $\hat{u}$ conditional on *u*. We used these distributions to make two main points. The first point is that the distribution of $\hat{u}$ is not, and should not be expected to be, the same as that of *u*. So, for example, if we assume a half normal distribution for *u*, and we plot the distribution of $\hat{u}$, we should not be disturbed when it does not look half normal. The second point is that $\hat{u}$ is (in a probabilistic sense) a shrinkage of *u* toward the mean. On average, we will overestimate the smaller realizations of *u* and underestimate the larger realizations.

We stress that neither of these facts means that there is anything “wrong” with $\hat{u}$. It is the optimal (rational, conditional expectation, minimum mean square error,…) forecast of *u*. What the paper illustrates is just the sense in which statistical noise is inconvenient.

APPENDIX

A. Derivation of the Jacobian in equation (7)

From equation (3), we have $\hat{u} = h(\varepsilon) = k \bullet [-\varepsilon + \sigma_0 \bullet \lambda(\varepsilon / \sigma_0)]$, where $k = \sigma_u^2 / (\sigma_u^2 + \sigma_v^2)$. So

(A1) $$\frac{d\hat{u}}{d\varepsilon} = k \bullet [-1 + \sigma_0 \bullet \lambda'(\varepsilon / \sigma_0) \bullet (1/\sigma_0)] = k \bullet [-1 + \lambda'(\varepsilon / \sigma_0)].$$ Then

(A2) $$g'(\hat{u}) = \frac{d\varepsilon}{d\hat{u}} = \left[\frac{d\hat{u}}{d\varepsilon}\right]^{-1} = \frac{1}{k \bullet [-1 + \lambda'(\varepsilon / \sigma_0)]} = \frac{\sigma_u^2 + \sigma_v^2}{\sigma_u^2 \bullet [-1 + \lambda'(g(\hat{u}) / \sigma_0)]}$$

and the Jacobian is just the absolute value of this expression.

B. Proof of Theorem 1

First we give some facts about the inverse Mill's ratio $\lambda(s) = \phi(s) / [1 - \Phi(s)]$. As $s \to -\infty$, (i) $\lambda(s) \to 0$, (ii) $s\lambda(s) \to 0$, (iii) $\lambda'(s) = -s\lambda(s) + \lambda^2(s) \to 0$. (Note that (i) and (iii) follow from (ii), and (ii) follows from the existence of the integral defining the mean of the standard normal.)

Now we start with the expression for $\hat{u}$, as given above. As $\sigma_v^2 \to 0$, $k \to 1$, $-\varepsilon \to_p u$ (since as $\sigma_v^2 \to 0, v \to_p 0$), $\sigma_0 \to 0$, and $\sigma_0 \lambda(\varepsilon / \sigma_0) \to 0 \bullet \lambda(-\infty) = 0$. Therefore $\hat{u} \to_p u$ (in the sense that the difference between $\hat{u}$ and u goes to zero). This proves part (1) of Theorem 1.

To prove part (2), consider the density of $\hat{u}$ as given in equation (8) of the text. As $\sigma_v^2 \to 0$, we have $a \to \sigma_u$, $a / \sigma_u^2 \to 1/\sigma_u$, $g(\hat{u}) \to -u$, $\phi(g(\hat{u})/a) \to \phi(-u/\sigma_u) = \phi(u/\sigma_u)$, and $\Phi(-g(\hat{u})b/a) \to \Phi(\infty) = 1$. Also the Jacobian term $\to 1$ because $\lambda'(-\infty) = 0$. Therefore $f_{\hat{u}}(\hat{u}) \to 2 \bullet (1/\sigma_u) \bullet \phi(\hat{u}/\sigma_u)$, which is the half normal density.

To prove part (3), of the Theorem, we return to the expression for $\hat{u}$ given above, which we write as $\hat{u} = -k\varepsilon + k\sigma_0\lambda(\varepsilon/\sigma_0)$. As $\sigma_v^2 \to \infty$, $k \to 0, \sigma_0^2 \to \infty, k \bullet \sigma_0 \to \sigma_u$ and $\lambda(\varepsilon/\sigma_0) \to \lambda(0) = \sqrt{(2/\pi)}$. Therefore $\hat{u} \to_p \sigma_u \bullet \sqrt{(2/\pi)} = E(u)$.

To prove part (4), we write

(A3) $$\frac{\sigma_v}{\sigma_u^2} \bullet (\hat{u} - E(u)) = -\frac{\sigma_v}{\sigma_u^2} \bullet k \bullet \varepsilon + \frac{\sigma_v}{\sigma_u^2} \bullet \left[k \bullet \sigma_0 \bullet \lambda(\frac{\varepsilon}{\sigma_0}) - \sigma_u \bullet \sqrt{\frac{2}{\pi}} \right].$$

The first term on the r.h.s. of (A3) equals

(A4) $$\frac{\sigma_v}{\sigma_u^2 + \sigma_v^2} \bullet u - \frac{\sigma_v^2}{\sigma_u^2 + \sigma_v^2} \bullet \frac{v}{\sigma_v} \simeq -\frac{v}{\sigma_v},$$

where "$\mathrm{A} \simeq \mathrm{B}$" means that $\mathrm{A} - \mathrm{B} \to 0$ with probability one as $\sigma_v^2 \to \infty$. Note that $-v/\sigma_v$ is $N(0,1)$.

The second term on the r.h.s. of (A4) is

(A5) $$\frac{\sigma_v}{\sigma_u^2} \bullet \left[k \bullet \sigma_0 \bullet \lambda(\frac{\varepsilon}{\sigma_0}) - \sigma_u \bullet \sqrt{\frac{2}{\pi}} \right]$$

$$= \frac{\sigma_v}{\sigma_u^2} \bullet \left[\frac{\sigma_u \sigma_v}{\sqrt{\sigma_u^2 + \sigma_v^2}} \bullet \lambda(\frac{\varepsilon}{\sigma_0}) - \sigma_u \bullet \sqrt{\frac{2}{\pi}} \right]$$

$$\simeq \frac{\sigma_v}{\sigma_u} \bullet \left[\lambda(\frac{\varepsilon}{\sigma_0}) - \sqrt{\frac{2}{\pi}} \right].$$

Now use the mean value theorem (delta method) to write

(A6) $$\lambda(\frac{\varepsilon}{\sigma_0}) \simeq \lambda(0) + \lambda'(0) \bullet (\frac{\varepsilon}{\sigma_0})$$

$$= \sqrt{\frac{2}{\pi}} + \frac{2}{\pi} \bullet \frac{\varepsilon}{\sigma_0}$$

and so the term in (A5) becomes

(A7) $$\frac{\sigma_v}{\sigma_u} \bullet \frac{2}{\pi} \bullet \frac{\varepsilon}{\sigma_0} \ .$$

Also

(A8) $$\frac{2}{\pi} \bullet \frac{\sigma_v}{\sigma_u} \bullet \frac{\varepsilon}{\sigma_0} = \frac{2}{\pi} \bullet \frac{\sigma_v}{\sigma_u} \bullet \left(\frac{-u}{\sigma_0} + \frac{v}{\sigma_0}\right)$$

$$\simeq \frac{2}{\pi} \bullet \frac{\sigma_v}{\sigma_u} \bullet \frac{v}{\sigma_0} = \frac{2}{\pi} \bullet \frac{\sigma_v}{\sigma_u} \bullet \frac{\sigma_u}{\sqrt{\sigma_u^2 + \sigma_v^2}} \bullet \frac{v}{\sigma_v}$$

$$\simeq \frac{2}{\pi} \bullet \frac{v}{\sigma_v} \ .$$

Combining (A8) with (A4), we have

(A9) $$\frac{\sigma_v}{\sigma_u^2} \bullet (\hat{u} - E(u)) \simeq \left(-1 + \frac{2}{\pi}\right) \bullet \frac{v}{\sigma_v^2}$$

which is distributed as $\text{N}(0,[(\pi\text{-}2)/\pi]^2)$.

C. Derivation of $f(\hat{u}|u)$ in equation (9)

We begin by noting that the joint density of (u,ε) is $f(u,\varepsilon) = f_u(u) \bullet f_v(\varepsilon + u)$. Now transform to $(u,\hat{u})$ where as before $\varepsilon = g(\hat{u})$. The Jacobian of this transformation is $|g'(\hat{u})|$ as given in equation (7) of the text. Therefore the joint density of $(u,\hat{u})$ is

(A10) $$f(u,\hat{u}) = f_u(u) \bullet f_v(g(\hat{u} + u) \bullet \text{Jacobian}$$

and the conditional density of $\hat{u}$ given u is

(A11) $$f(\hat{u}|u) = f(u,\hat{u}) / f_u(u) = f_v(g(\hat{u}) + u) \bullet \text{Jacobian} \ .$$

Substituting the normal density for f_v and the Jacobian expression in (7), we arrive at the expression in equation (9) of the text.

D. Proof of Theorem 2

To prove part (1) of the Theorem, we write

(A12) $$\frac{\hat{u}-u}{\sigma_v} = -k \bullet \frac{v}{\sigma_v} + (k-1) \bullet \frac{u}{\sigma_v} + k \bullet \frac{\sigma_0}{\sigma_v} \bullet \lambda(\frac{v-u}{\sigma_0}) \ .$$

As $\sigma_v^2 \to 0$, $k \to 1$, so the first term on the r.h.s. $\approx -v/\sigma_v$. The second term is:

(A13) $$\frac{k-1}{\sigma_v} \bullet u = -\frac{\sigma_v}{\sigma_u^2 + \sigma_v^2} \bullet u \to \frac{0}{\sigma_u^2} \bullet u = 0 \ .$$

(Remember u is fixed in this calculation.) The third term is

(A14) $$k \bullet \frac{\sigma_0}{\sigma_v} \bullet \lambda(\frac{v-u}{\sigma_0}) \simeq \lambda(-\infty) = 0$$

where we have used the facts that, as $\sigma_v^2 \to 0$, $k \to 1$ and $\sigma_0 / \sigma_v \to 1$. Therefore

$(\hat{u}-u)/\sigma_v \simeq -v/\sigma_v$ which is $N(0,1)$.

The proof of part (2) is essentially the same as the proof of part (4) of Theorem 1, and is therefore omitted.

REFERENCES

Aigner, D.J., C.A.K. Lovell, and P. Schmidt (1977), "Formulation and Estimation of Stochastic Frontier Production Function Models," *Journal of Econometrics* 6, 21-37.

Battese, G.E., and T.J. Coelli (1988), "Prediction of Firm-Level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data," *Journal of Econometrics* 38, 387-399.

Jondrow, J., C.A.K. Lovell, I.S. Materov, and P. Schmidt (1982), "On the Estimation of Technical Efficiency in the Stochastic Production Function Model," *Journal of Econometrics* 19, 233-238.

Meeusen, W., and J. Van Den Broeck (1977), "Efficient Estimation from Cobb-Douglas Production Functions with Composed Error," *International Economic Review* 18, 435-444.

Pitt, M.M., and L.F. Lee (1981), "The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry," *Journal of Development Economics* 9, 43-64.

Stevenson, R.E. (1980), "Likelihood Functions for Generalized Stochastic Frontier Estimation," *Journal of Econometrics* 13, 57-66.

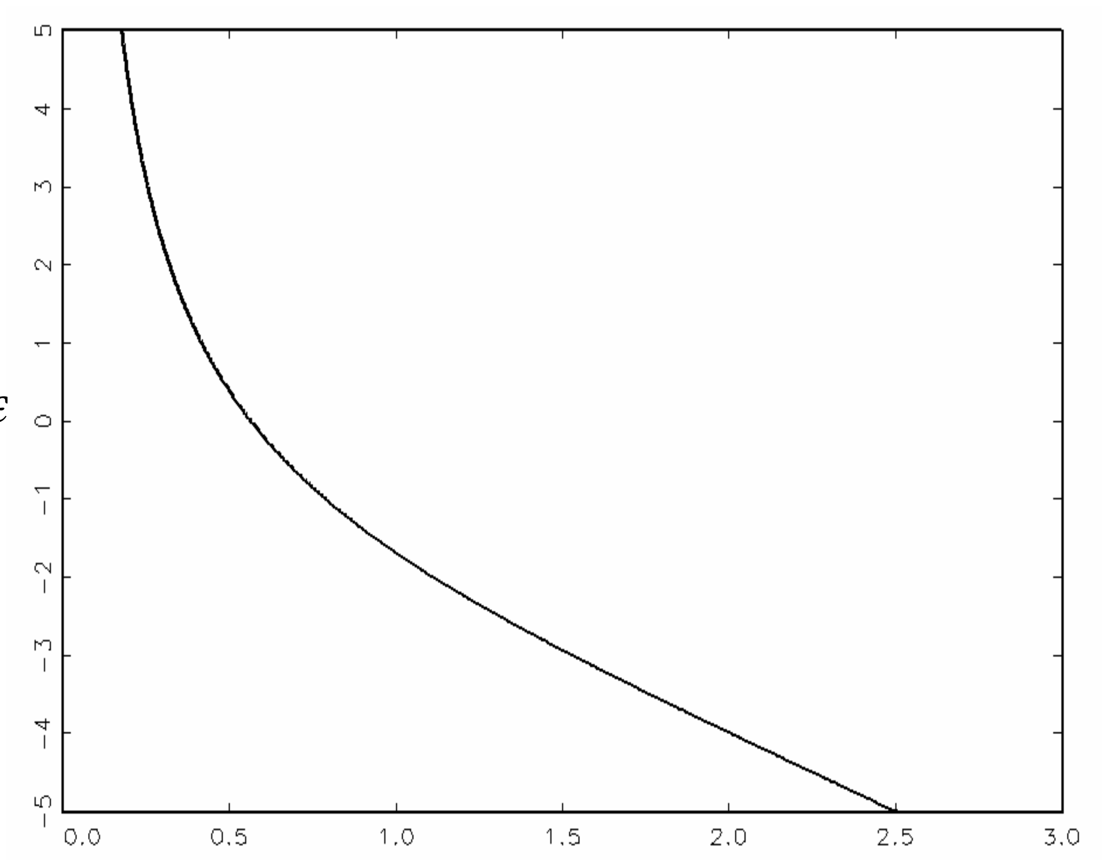
Greene, W.H. (1980a), "Maximum Likelihood Estimation of Econometric Frontier Functions," *Journal of Econometrics* 13, 27-56.

Greene, W.H. (1980b), "On the Estimation of a Flexible Frontier Production Model," *Journal of Econometrics* 13, 101-115.

Greene, W.H. (1990), "A Gamma-Distributed Stochastic Frontier Model," *Journal of Econometrics* 46, 141-164.

Figure 1

The relationship between ε and $\hat{\mathrm{u}}$ with $\sigma_{\mathrm{u}}^2 = \sigma_v^2 = 1$



ε

$\hat{u}$

Figure 2 A half normal distribution, $u \sim N(0,1)^{+}$

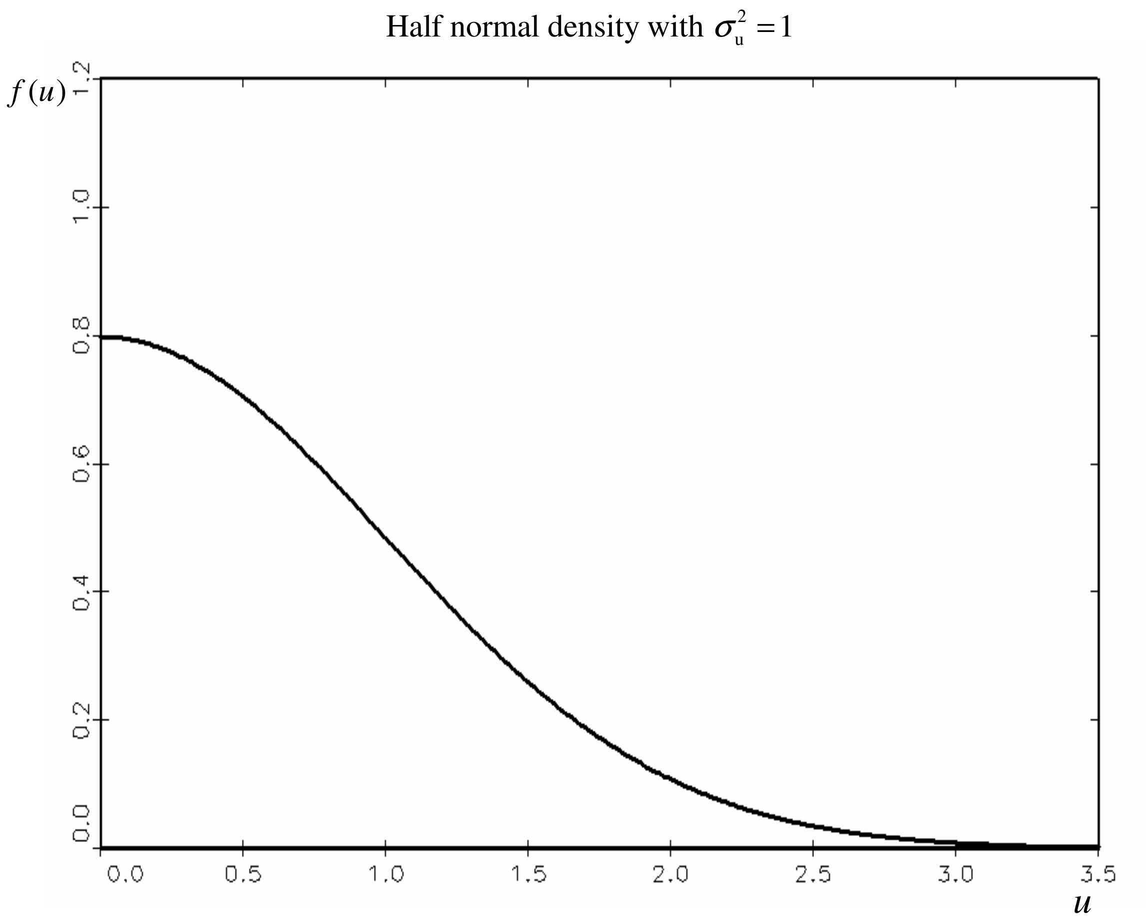


Figure 3

Density of $\hat{u}$ with $\sigma_u^2 = 1$ and $\sigma_v^2 = .1$

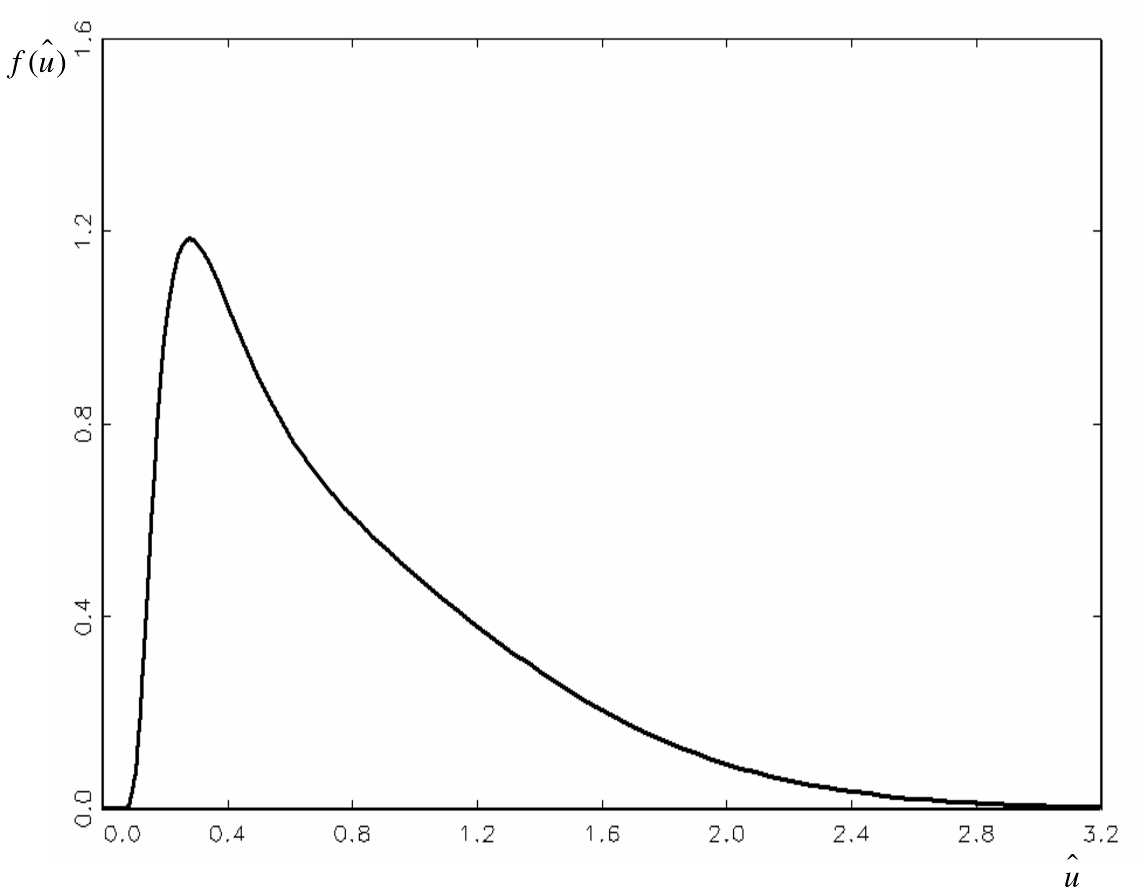


Figure 4

Density of $\hat{u}$ with $\sigma_u^2 = 1$ and $\sigma_v^2 = 1$

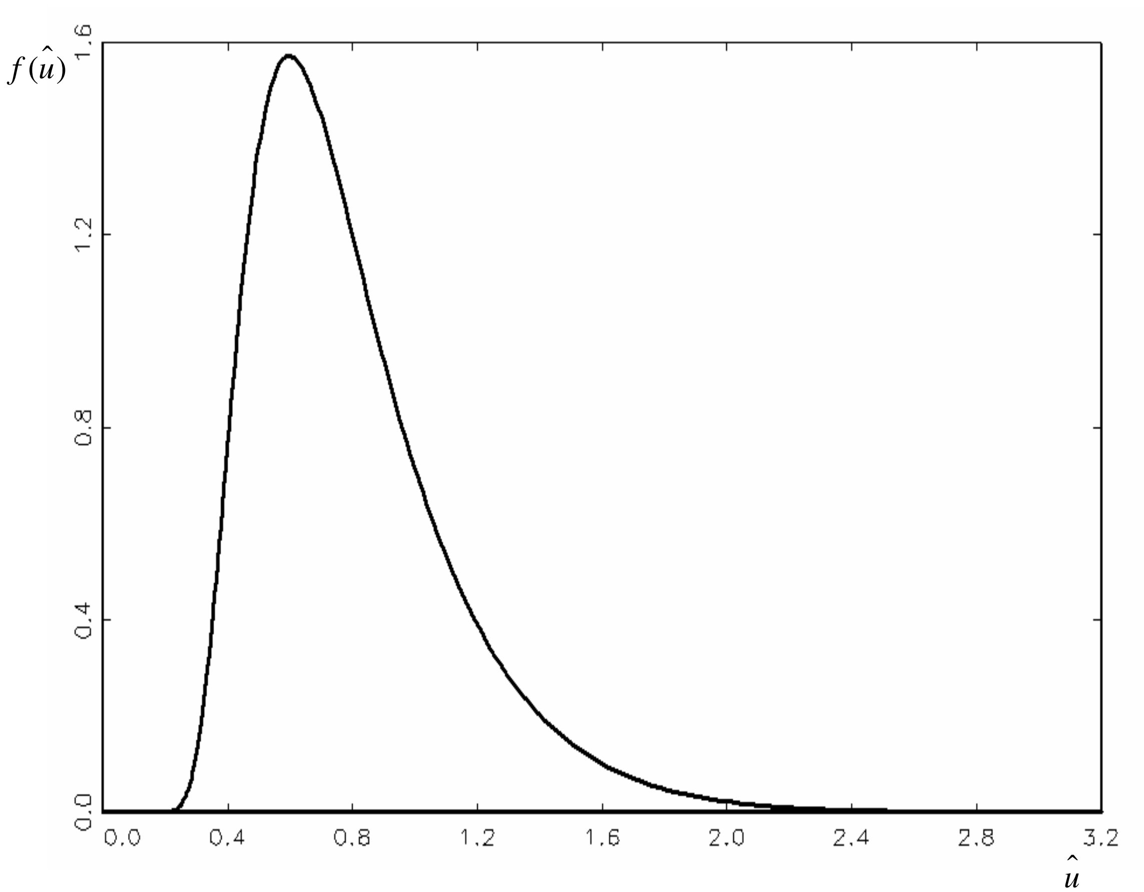


Figure 5

Density of $\hat{u}$ with $\sigma_u^2 = 1$ and $\sigma_v^2 = 10$

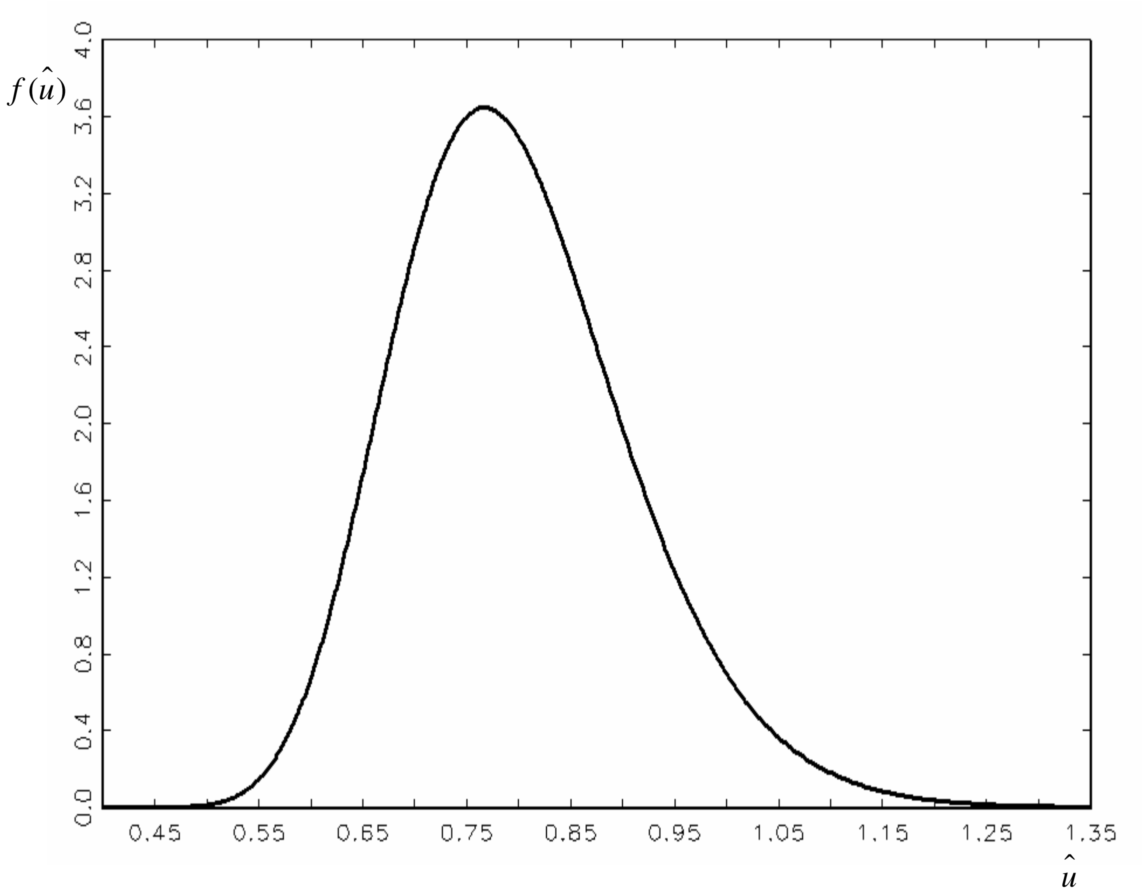


Figure 6

Density of $\hat{u}$ with $\sigma_u^2 = 1$ and $\sigma_v^2 = 100$

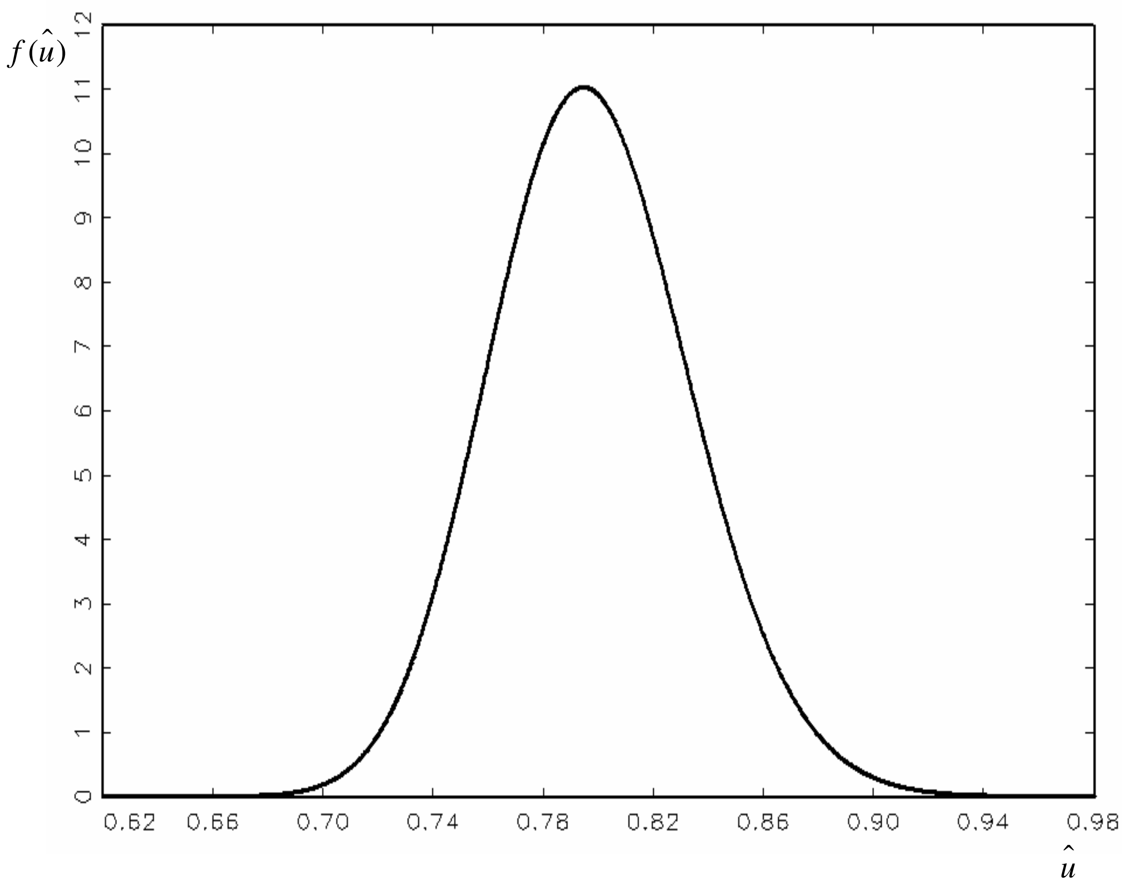


Figure 7 The combined graph for figure 2-figure 6

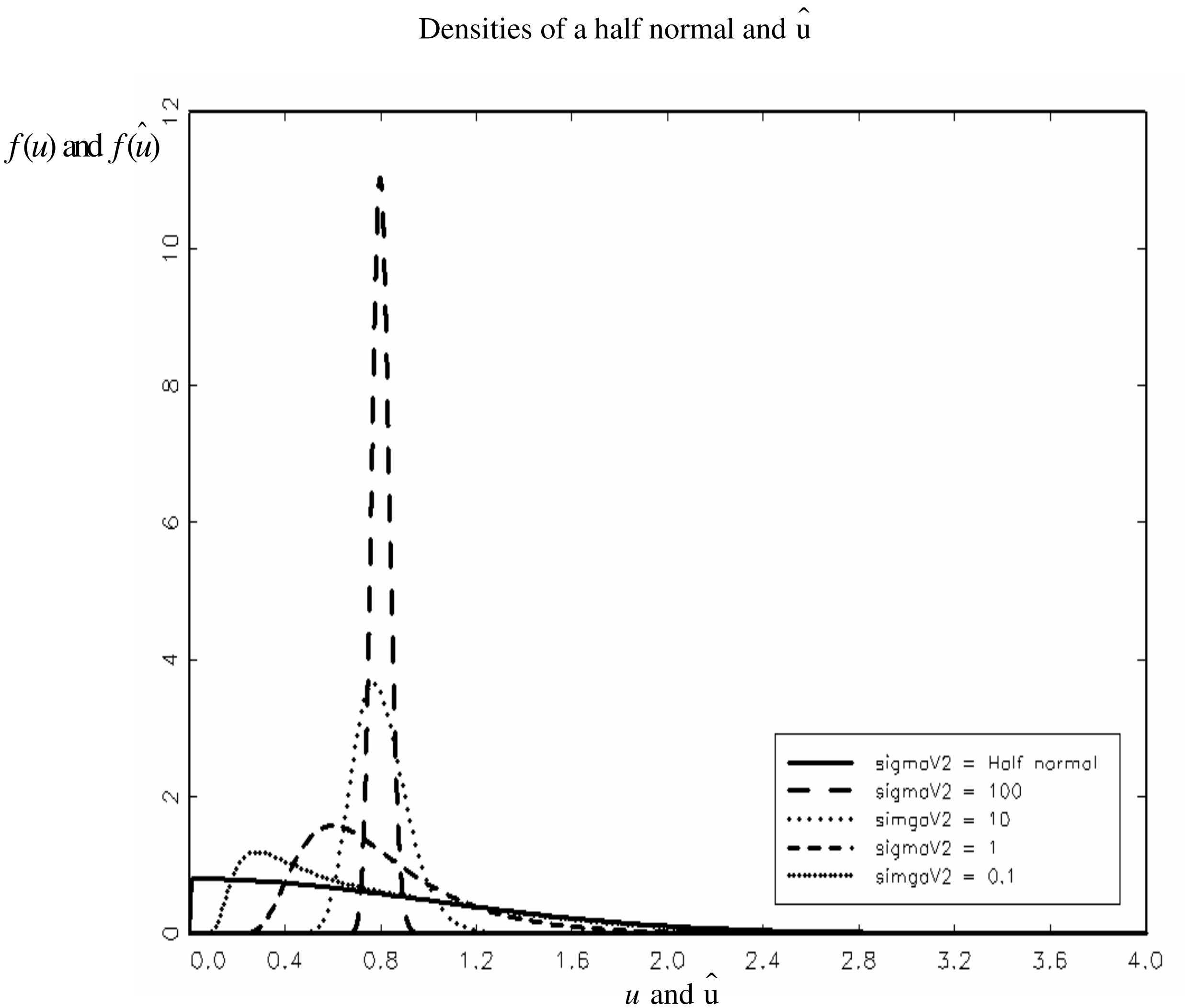


Figure 8

Density of $\hat{u} \mid u$=.1 with $\sigma_u^2 = 1$ and $\sigma_v^2 = .001$

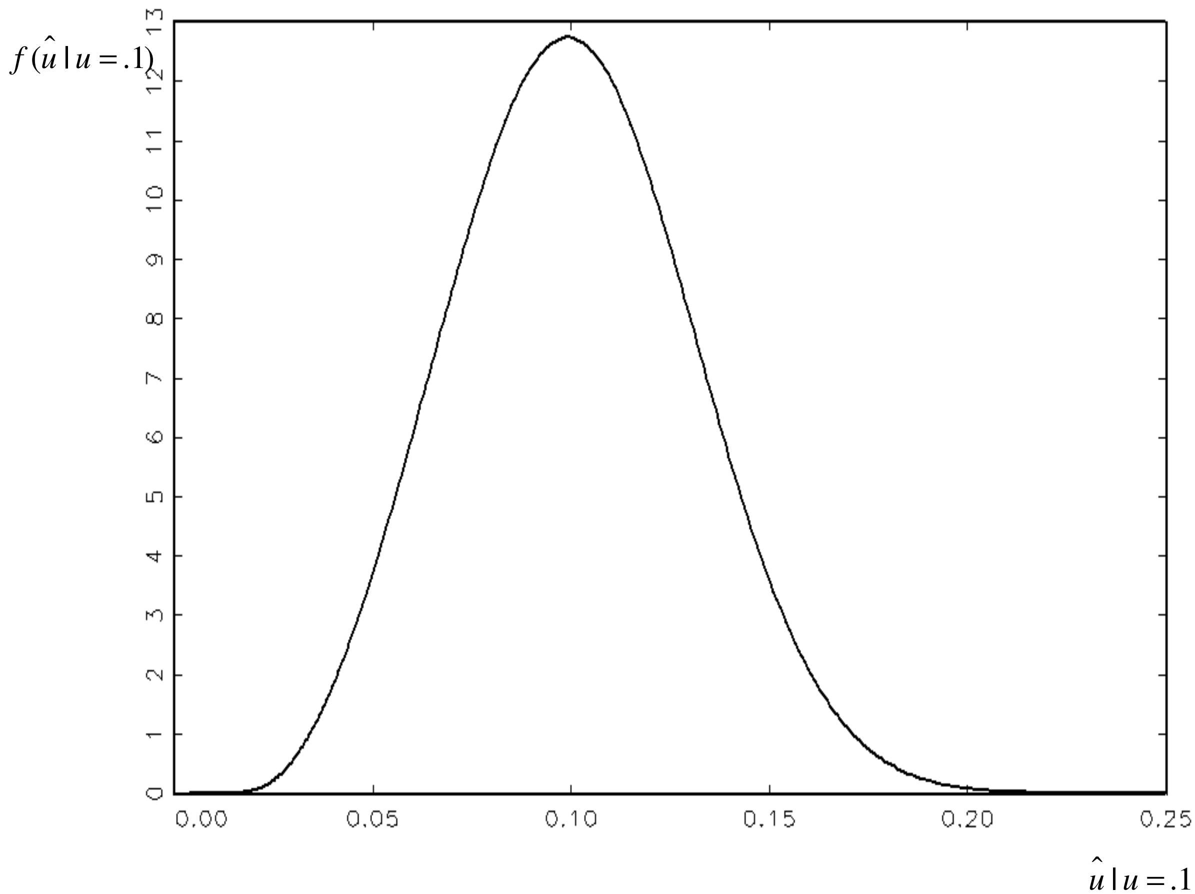


Figure 9

Density of $\hat{u} \mid u$=.1 with $\sigma_u^2 = 1$ and $\sigma_v^2 = .01$

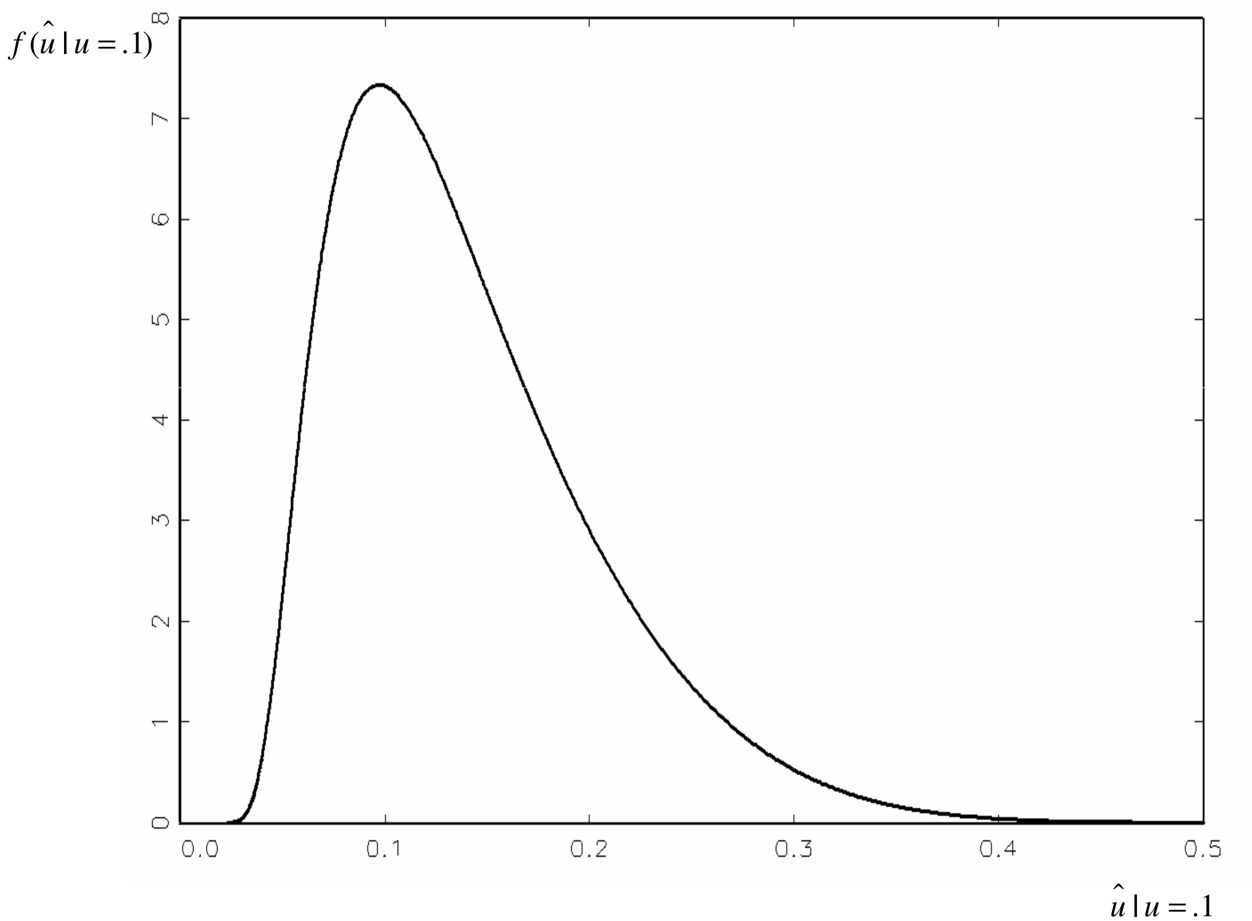


Figure 10

Density of $\hat{u}\,|\,u$=.1 with $\sigma_u^2 = 1$ and $\sigma_v^2 = .1$

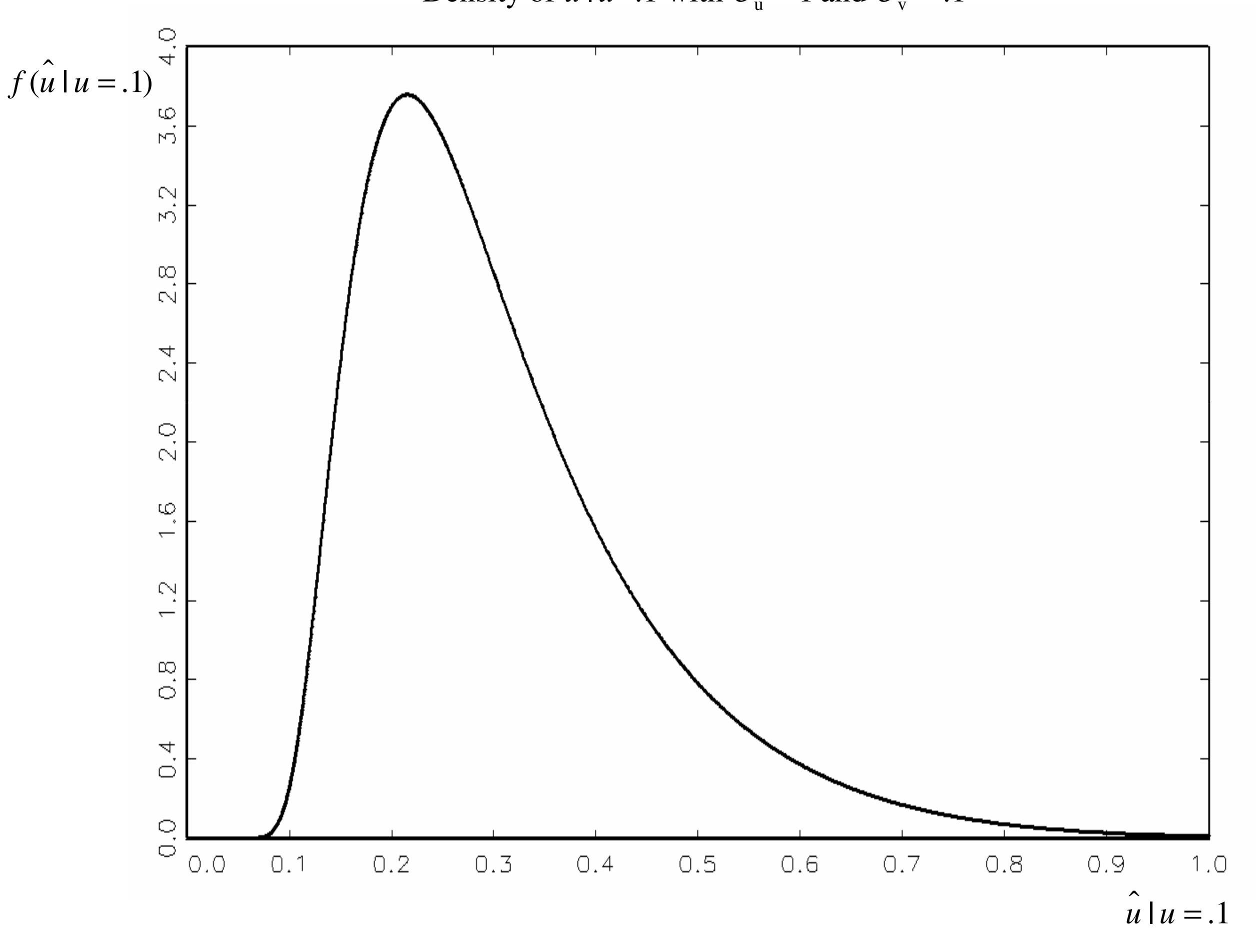


Figure 11

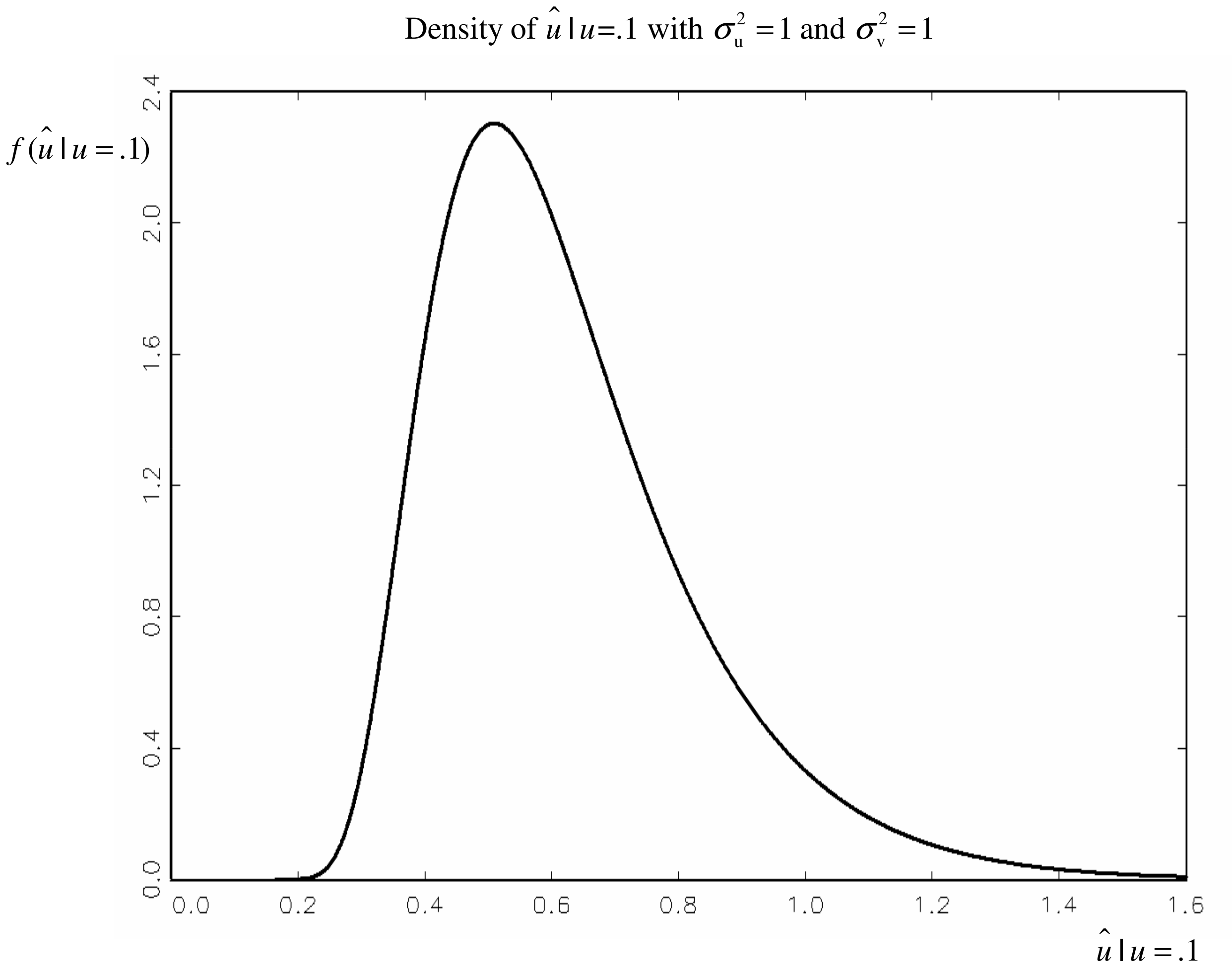

Density of $\hat{u} \mid u$=.1 with $\sigma_u^2 = 1$ and $\sigma_v^2 = 1$
$f(\hat{u} \mid u = .1)$
2.4
2.0
1.6
1.2
0.8
0.4
0.0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
$\hat{u} \mid u = .1$

Figure 12

Density of $\hat{u} \mid u$=.1 with $\sigma_u^2 = 1$ and $\sigma_v^2 = 10$

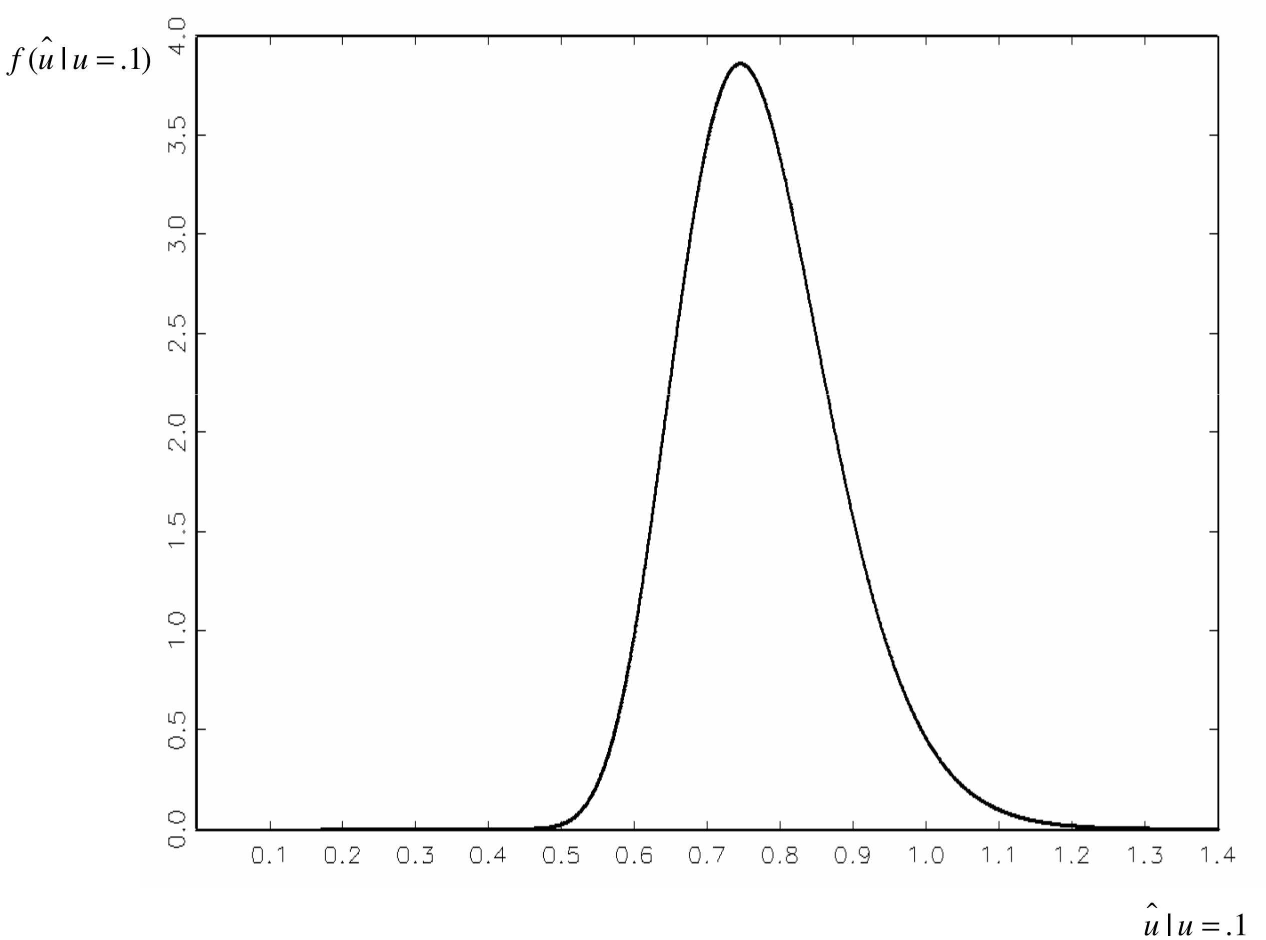


Figure 13

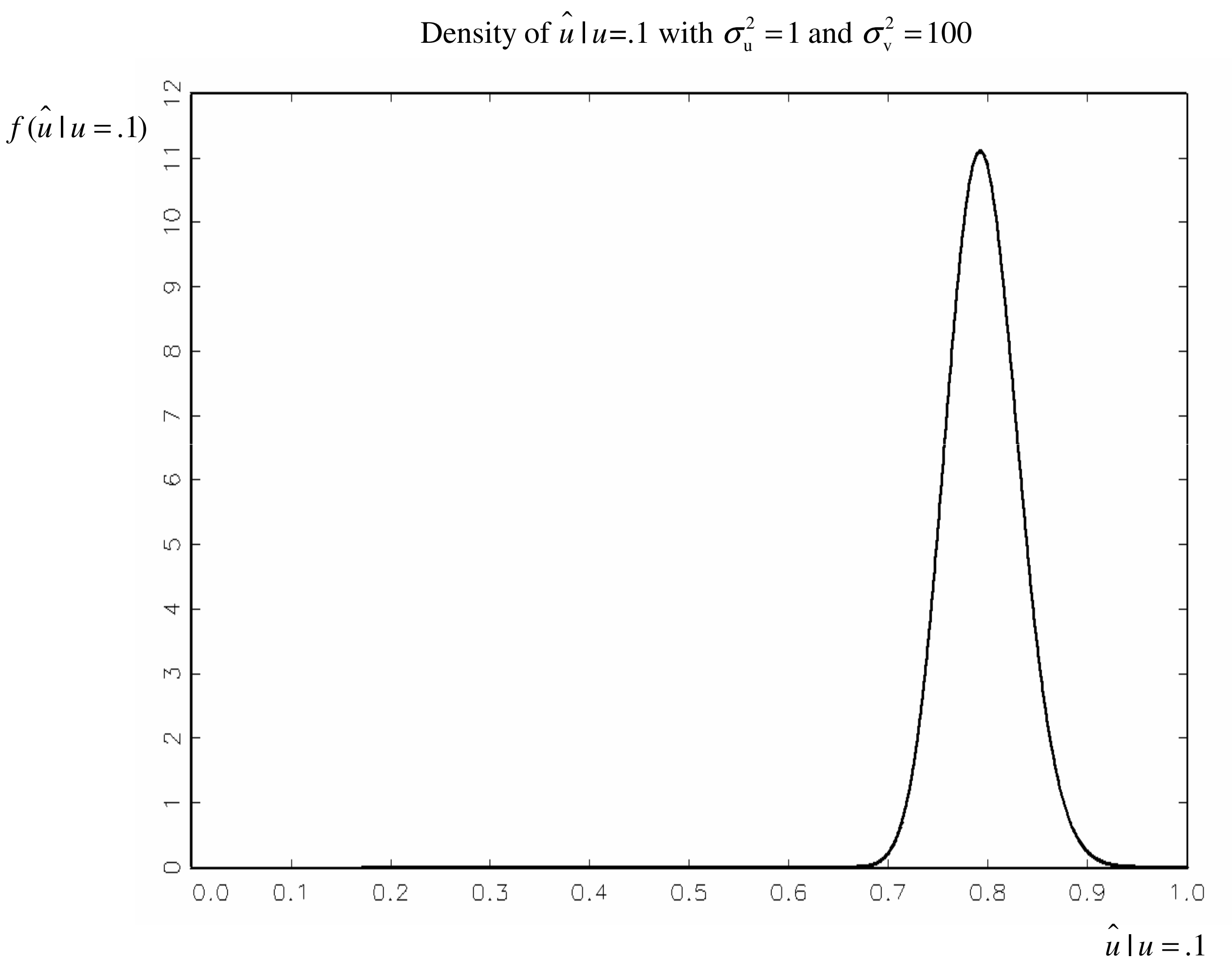

Density of $\hat{u} \mid u$=.1 with $\sigma_u^2 = 1$ and $\sigma_v^2 = 100$
$f(\hat{u} \mid u = .1)$
0
1
2
3
4
5
6
7
8
9
10
11
12
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
$\hat{u} \mid u = .1$

Figure 14

Density of $\hat{u} \mid u$=2 with $\sigma_u^2 = 1$ and $\sigma_v^2 = .001$

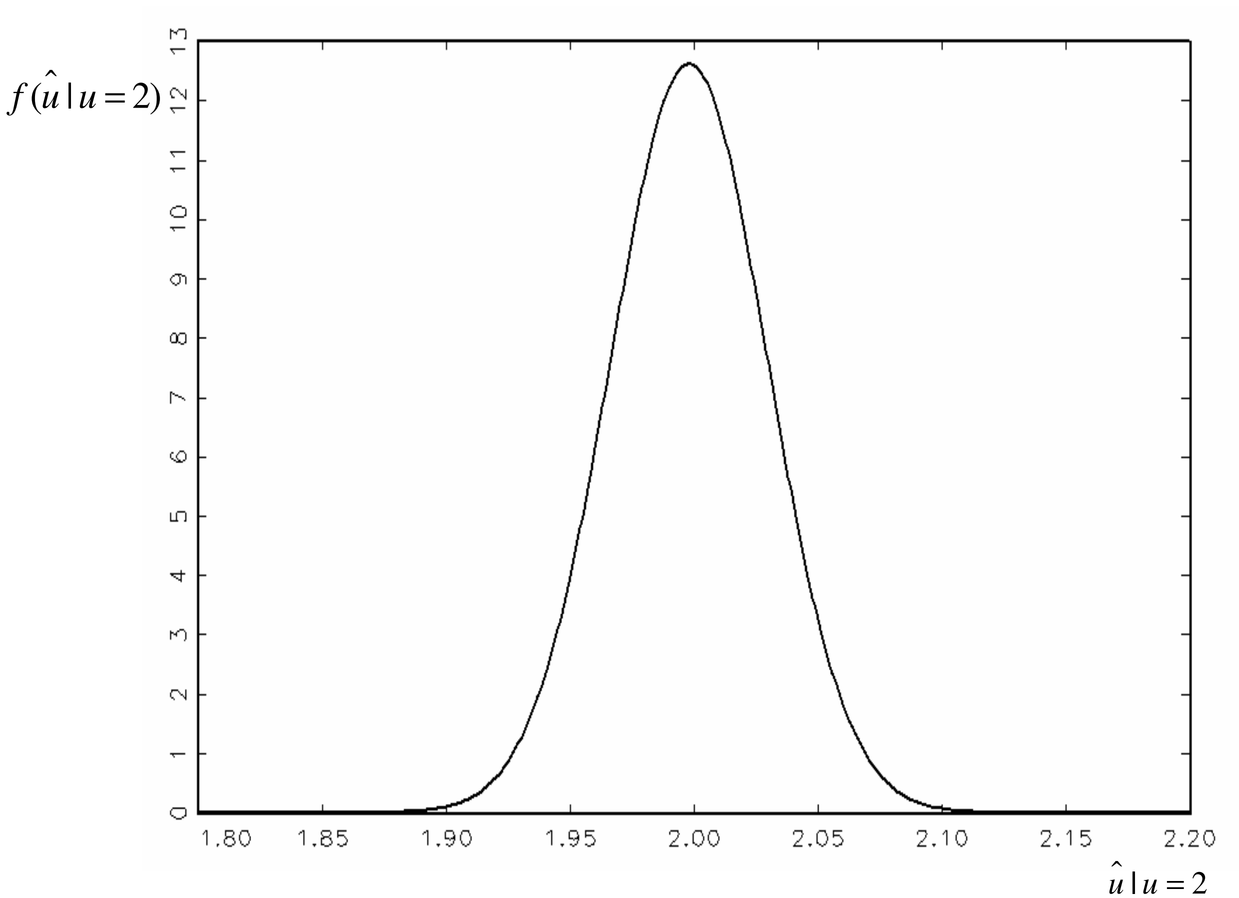


Figure 15

Density of $\hat{u} \mid u$=2 with $\sigma_u^2 = 1$ and $\sigma_v^2 = .01$

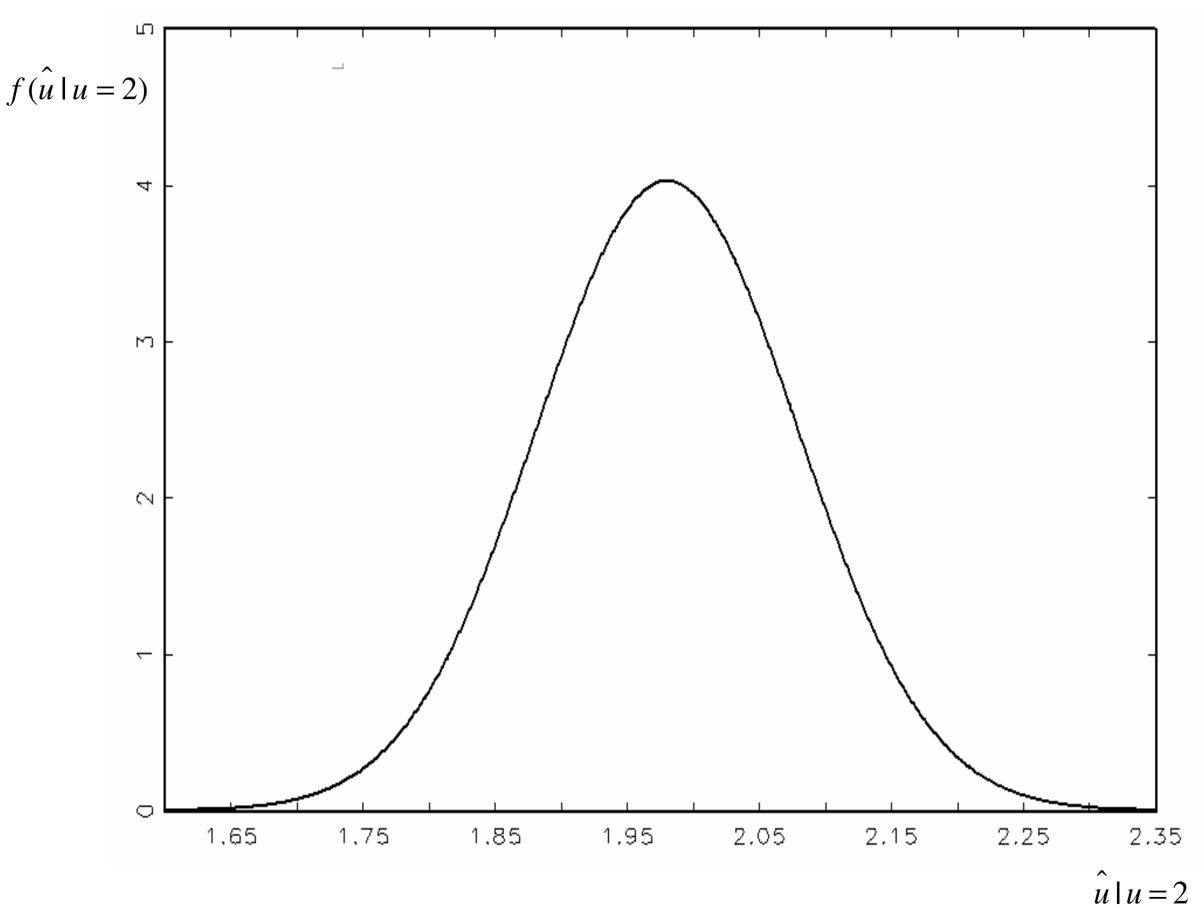


Figure 16

Density of $\hat{u} \mid u$=2 with $\sigma_u^2 = 1$ and $\sigma_v^2 = .1$

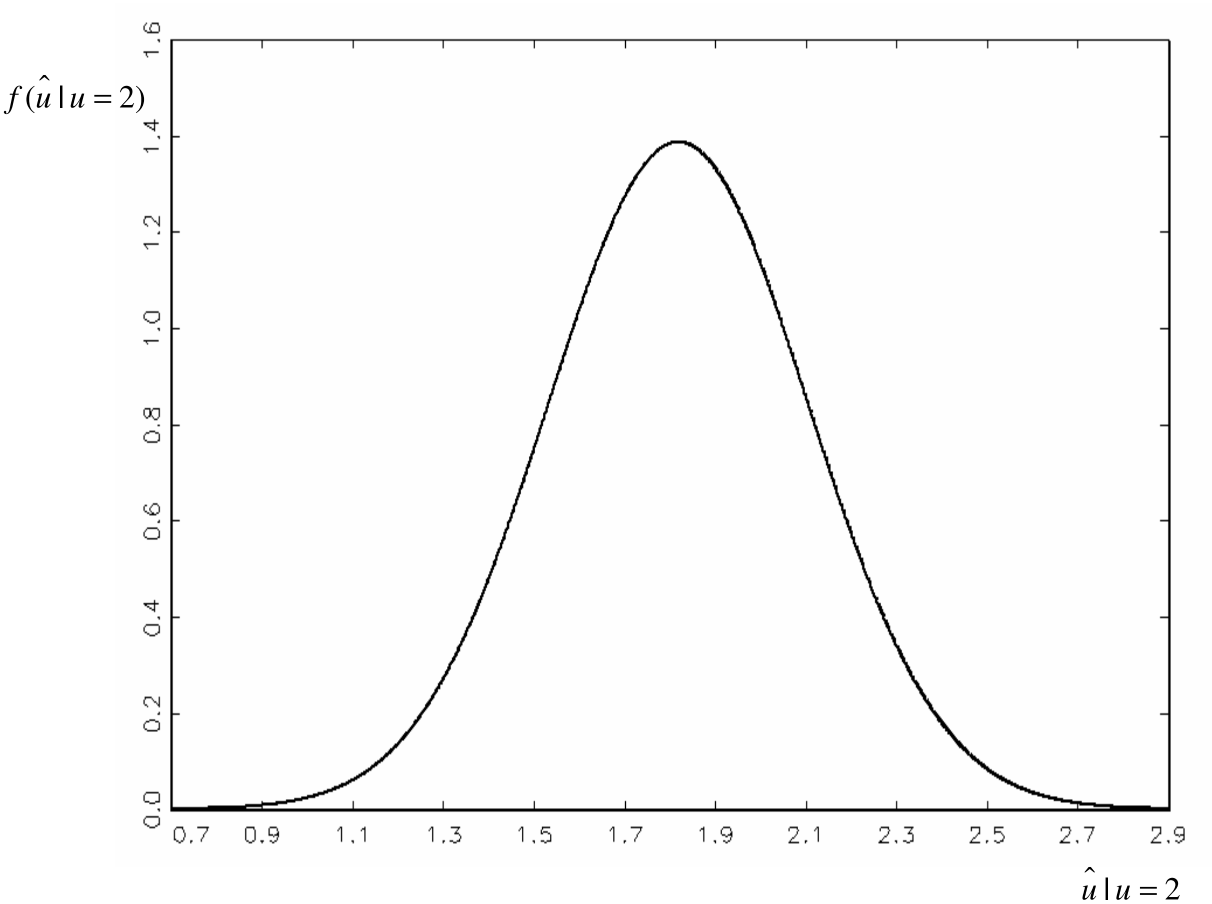


Figure 17

Density of $\hat{u} \mid u$=2 with $\sigma_u^2 = 1$ and $\sigma_v^2 = 1$

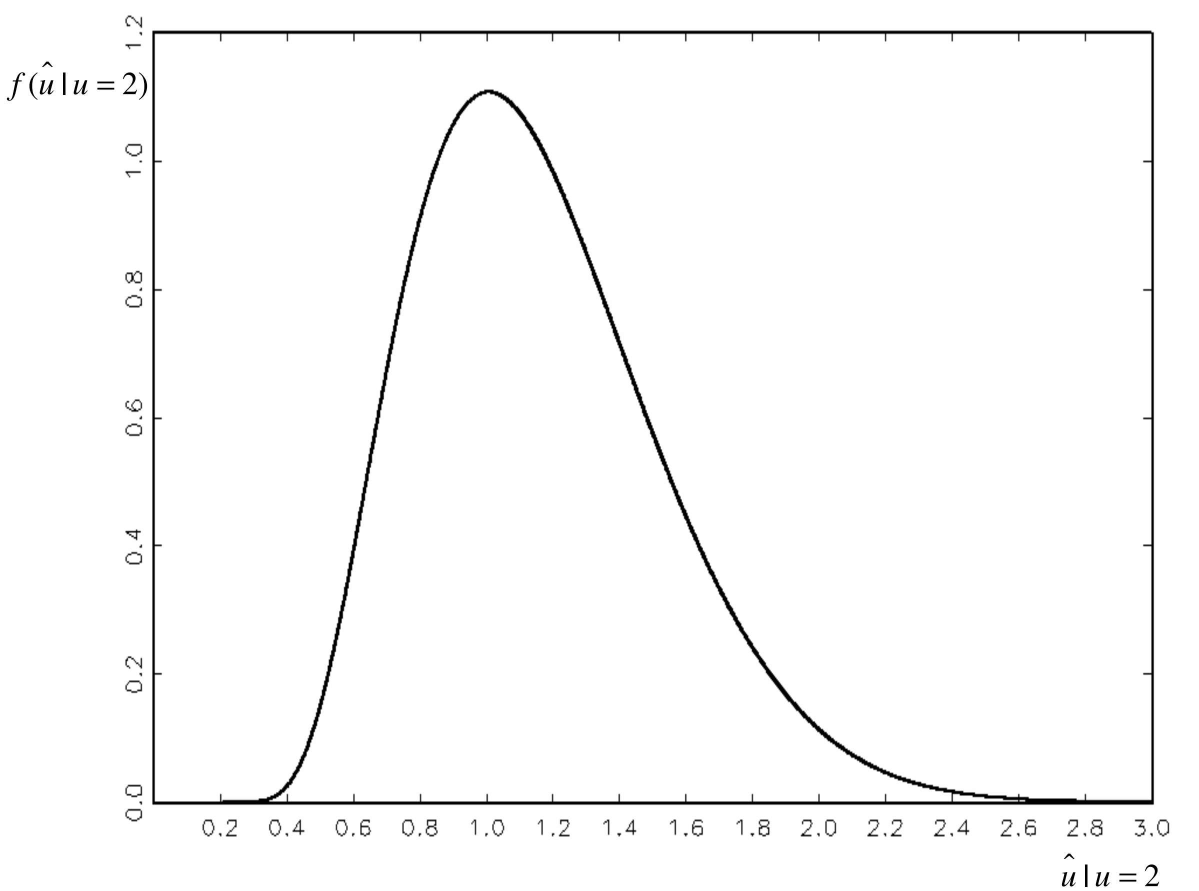


Figure 18

Density of $\hat{u} \mid u$=2 with $\sigma_u^2 = 1$ and $\sigma_v^2 = 10$

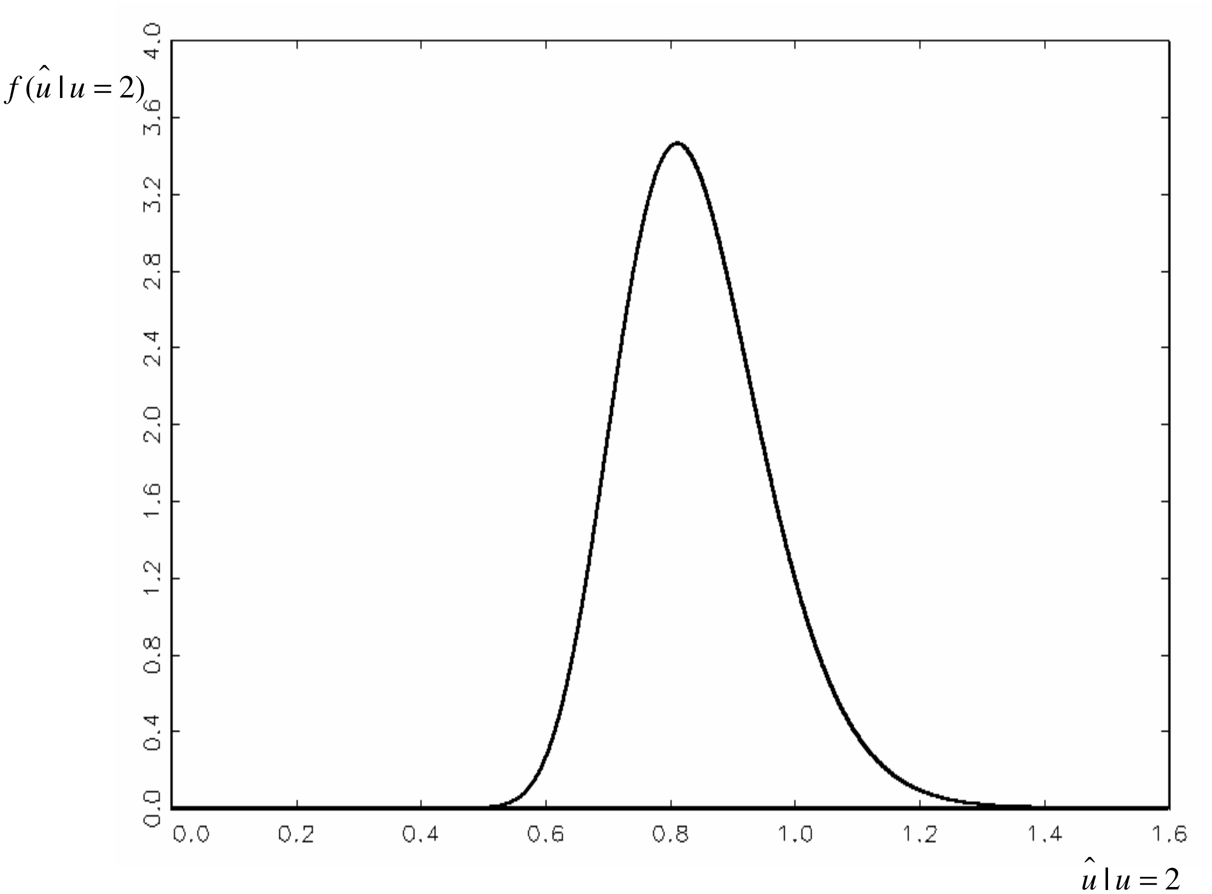


Figure 19

Density of $\hat{u} \mid u$=2 with $\sigma_u^2 = 1$ and $\sigma_v^2 = 100$

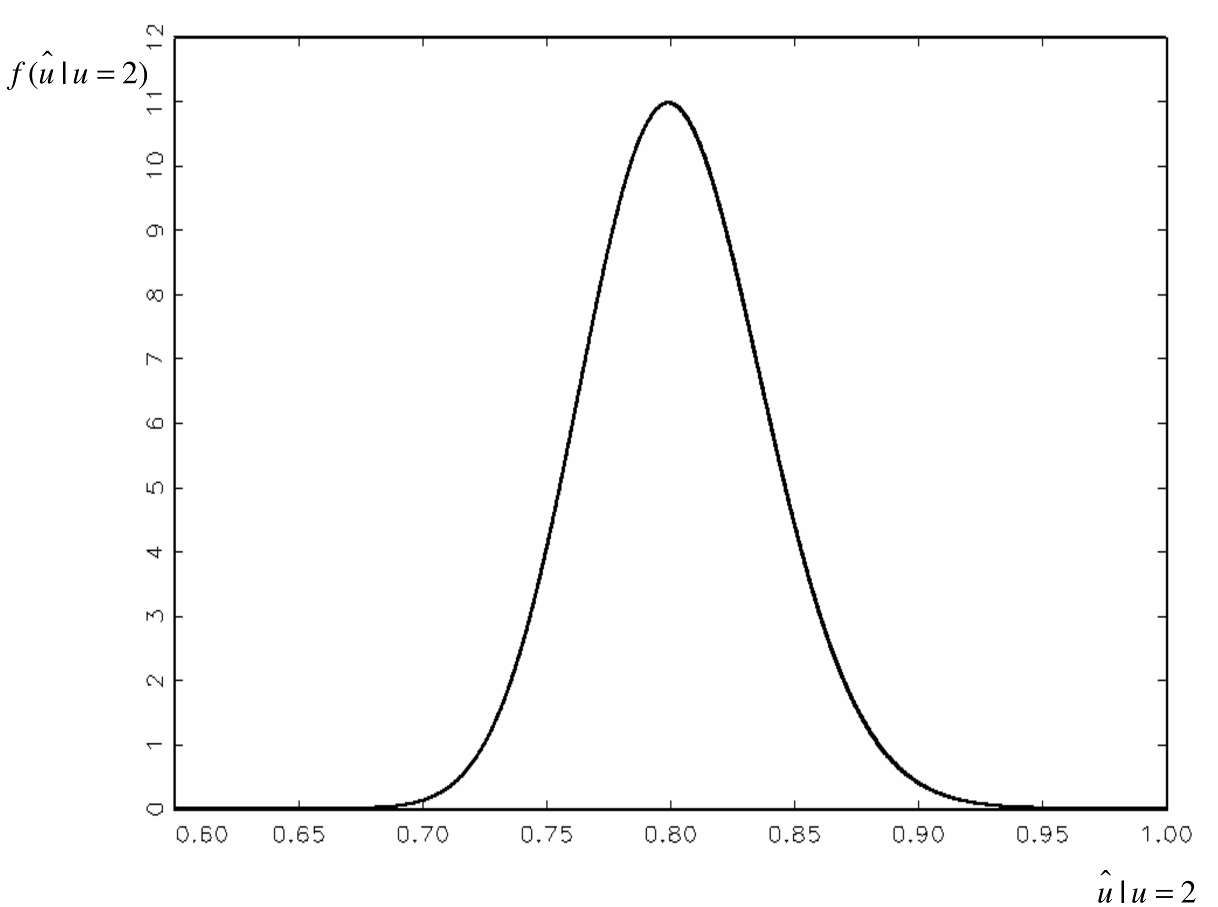


Figure 20 Density of $\hat{u} \mid u$ for u =.1, .5, 1 and 2 with $\sigma_u^2 = 1$ and $\sigma_v^2 = 100$