This is to certify that the
thesis entitled

THREE ESSAYS ON ECONOMETRICS

presented by

WEI SIANG WANG

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in ECONOMICS

Peter Schmidt
Major Professor's Signature

JULY 24, 2009

Date

THREE ESSAYS ON ECONOMETRICS

BY

WEI SIANG WANG

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

ECONOMICS

2009

ABSTRACT

THREE ESSAYS ON ECONOMETRICS

BY

WEI SIANG WANG

The first essay, "On the Distribution of Estimated Technical Efficiency in Stochastic Frontier Models," considers a stochastic frontier model with error $\varepsilon = v - u$, where v is normal and u is half normal. We derive the distribution of the usual estimate of u, $E(u \mid \varepsilon)$. We show that as the variance of v approaches zero, $E(u \mid \varepsilon) - u$ converges to zero, while as the variance of v approaches infinity, $E(u \mid \varepsilon)$ converges to $E(u)$. We graph the density of $E(u \mid \varepsilon)$ for intermediate cases. To show that $E(u \mid \varepsilon)$ is a shrinkage of u towards its mean, we derive and graph the distribution of $E(u \mid \varepsilon)$ conditional on u. We also consider the distribution of estimated inefficiency in the fixed-effects panel data setting.

The second essay, "Goodness of Fit Tests in Stochastic Frontier Models," discusses goodness of fit tests for the distribution of technical inefficiency in stochastic frontier models. If we maintain the hypothesis that the assumed normal distribution for statistical noise is correct, the assumed distribution for technical inefficiency is testable. We show that a goodness of fit test can be based on the distribution of estimated technical efficiency, or equivalently on the distribution of the composed error term. We consider both the Pearson chi-squared test and the Kolmogorov-Smirnov test. The bootstrap can be used to account for the effects of parameter estimation. Alternatively, for the Pearson test, we use existing results in the literature to account for the fact that estimated

parameters are used to construct the actual and/or the expected cell counts. Finally, we provide simulation results to show the extent to which the tests are reliable in finite samples.

The third essay, "Testing Equality of Distribution for Two Correlated Variables," discusses how to test the null hypothesis that $y_1, y_2, \ldots, y_n$ and $x_1, x_2, \ldots, x_n$ from a correlated paired sample of size n: (y_i, x_i), $i = 1, 2, 3, \ldots, n$, have the same distribution. We implement the Pearson chi-squared test, based on differences of frequencies in non-overlapping intervals (cells) that span the support of the variables, in a GMM setting. This procedure makes no assumption about the correlation between the two variables. We also suggest a novel bootstrapping procedure that enables us to generate asymptotically valid critical values for the Kolmogorov-Smirnov and Baumgartner-Weiss-Schindler tests.

To my parents

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the help of so many people in so many ways. First, I express my most heartfelt thanks to Professor Peter Schmidt. He is an exemplary mentor through his continued patience, understanding, and guidance. He went beyond normal means in helping me and giving his support for my endeavors. His guidance is priceless, and for this I will be forever grateful. I wish to express my sincere gratitude to Professor Christine Amsler for her contribution on my dissertation. In addition, she is always enthusiastic and willing to share her wealth of knowledge and life experience with me since the first day I met her. She has broadened my outlook and mind. Her encouragement has made East Lansing a more welcome place.

In addition, I would like to thank Professor Emma Iglesias for listening and giving me sage comments as I struggled to solve problems. I wanted to thank Professor Scott Swinton for his suggestions on my dissertation and research career. I also owe a debt of gratitude to Jennifer and Stephanie for their help whenever I needed it.

I wish to thank my fellow graduate students. My special thanks go to Pei-Lin Simon, Kampon, Gabriel, Brian, Dooyeon, Do Won, Sung-Guan, Meng-Chi, and Tum.

Finally and most importantly, I want to thank my parents, Hong Kiew and Swee Yong, and my Uncle Jerry and Aunt Janice for their unconditional love, advice, and support over the years.

TABLE OF CONTENTS

TABLE OF CONTENTS..........vi

LIST OF TABLES..........vii

LIST OF FIGURES..........ix

1. ON THE DISTRIBUTION OF ESTIMATED TECHNICAL EFFICIENCY IN STOCHASTIC FRONTIER MODELS **1**
1.1 Introduction.......... 2
1.2 The Distribution of $\hat{u}$ 5
1.3 The Distribution of $\hat{u}$ Conditional on u.......... 8
1.4 Panel Data..........11
1.5 Concluding Remarks..........14
Appendix 1.......... 16

2. GOODNESS OF FIT TESTS IN STOCHASTIC FRONTIER MODELS **31**
2.1 Introduction..........31
2.2 Tests Based on the Distribution of ε..........33
2.3 Simple Hypotheses..........35
2.4 Composite Hypotheses..........37
2.5 An Introductory Example: Testing Normality..........44
2.6 The Stochastic Frontier Model..........47
    2.6.1 Size of the Test..........48
    2.6.2 Power of the Test.......... 51
2.7 Concluding Remarks..........54
Appendix 2.......... 57

3. TESTING EQUALITY OF DISTRIBUTION FOR TWO CORRELATED VARIABLES **81**
3.1 Introduction..........81
3.2 Tests of Equality of Two Distributions..........82
    3.2.1 Chi-squared Test.......... 82
    3.2.2 Kolmogorov-Smirnov Test.......... 84
    3.2.3 The Baumgartner-Weiss-Schindler Test..........85
    3.2.4 The Bootstrap Resampling Method.......... 86
    3.2.5 Pairwise T-Test..........87
3.3 Monte Carlo Simulations..........88
    3.3.1 Size of the Test..........88
    3.3.2 Power of the Test.......... 90
3.4 Concluding Remarks..........91

BIBLIOGRAPHY..........**111**

LIST OF TABLES

Table 2.1 Size of the test of the hypothesis that the data are N(0, 1). 62

Table 2.2 Size of the test of the hypothesis that the data are normal.......................63

Table 2.3 Quantiles of the distribution of the normal / half normal composed error......64

Table 2.4 Size of test of the hypothesis that the data are normal / half-normal ($\lambda = 0.1$)... 67

Table 2.5 Size of test of the hypothesis that the data are normal / half-normal ($\lambda = 0.5$) 68

Table 2.6 Size of test of the hypothesis that the data are normal / half-normal ($\lambda = 1$) 69

Table 2.7 Size of test of the hypothesis that the data are normal / half-normal ($\lambda = 2$) 70

Table 2.8 Size of test of the hypothesis that the data are normal / half-normal ($\lambda = 10$) 71

Table 2.9 Power of the test of the hypothesis that the data are normal / half-normal
Alternative: the data are normal / exponential(θ)
Number of cells: $k = 3$ 72

Table 2.10 Power of the test of the hypothesis that the data are normal / half-normal
Alternative: the data are normal / exponential(θ)
Number of cells: $k = 5$...73

Table 2.11 Power of the test of the hypothesis that the data are normal / half-normal
Alternative: the data are normal / exponential(θ)
Number of cells: $k = 10$...74

Table 2.12 Power of the test of the hypothesis that the data are normal / half-normal
Alternative: the data are normal / gamma (u is c times gamma(m))
$m = 0.1$...75

Table 2.13 Power of the test of the hypothesis that the data are normal / half-normal
Alternative: the data are normal / gamma (u is c times gamma(m))
$m = 0.5$...76

Table 2.14 Power of the test of the hypothesis that the data are normal / half-normal Alternative: the data are normal / gamma (u is c times gamma(m)) $m = 2$ 77

Table 2.15 Power of the test of the hypothesis that the data are normal / half-normal Alternative: the data are normal / gamma (u is c times gamma(m)) $m = 10$ 78

Table 2.16 Size of the test of the hypothesis that the data are normal / half normal with $\sigma_u^2 = \sigma_v^2 = 1$ (simple hypothesis) 79

Table 2.17 Size of the test of the hypothesis that the data are normal / half normal Results conditional on negative skew (Tauchen version of Pearson test)...... 80

Table 3.1 Size of tests of the hypothesis that the data are bivariate- normal........ 93

Table 3.2 Power of the tests when $\rho = 0.00$........ 96

Table 3.3 Power of the tests when $\rho = 0.20$........ 97

Table 3.4 Power of the tests when $\rho = 0.40$........ 98

Table 3.5 Power of the tests when $\rho = 0.60$........ 99

Table 3.6 Power of the tests when $\rho = 0.80$........ 100

Table 3.7 Power of the tests when $\rho = 0.90$........ 101

Table 3.8 Power of the tests when $\rho = 0.95$........ 102

Table 3.9 Power of the tests when $\rho = -0.20$........ 103

Table 3.10 Power of the tests when $\rho = -0.40$........ 104

Table 3.11 Power of the tests when $\rho = -0.60$........ 105

Table 3.12 Power of the tests when $\rho = -0.80$........ 106

Table 3.13 Power of the tests when $\rho = -0.90$........ 107

Table 3.14 Power of the tests when $\rho = -0.95$........ 108

LIST OF FIGURES

FIGURE 1.1 The relationship between ε and $\hat{u}$ with $\sigma_u^2 = \sigma_v^2 = 1$ 20

FIGURE 1.2 Density of $\hat{u}$ with $\sigma_u^2 = 1$ and $\sigma_v^2 = .1, 1, 10, 100$ 21

FIGURE 1.3 The combined graph for Figure 1.2.. 22

FIGURE 1.4 Density of $\hat{u} \mid u = .1$ with $\sigma_u^2 = 1$ and $\sigma_v^2 = .001, .01, .1, 1, 10, 100$ 23

FIGURE 1.5 Density of $\hat{u} \mid u = 2$ with $\sigma_u^2 = 1$ and $\sigma_v^2 = .001, .01, .1, 1, 10, 100$ 25

FIGURE 1.6 Density of $\hat{u} \mid u$ for u =.1, .5, 1 and 2 with $\sigma_u^2 = 1$ and $\sigma_v^2 = 100$ 27

FIGURE 1.7 Density of $\hat{u}$ and $\tilde{u}$ with $\sigma_u^2 = 1, \sigma_v^2/T = 1$ and N = 10, 100, 1000
Density of $\hat{u}$ and $\tilde{u}$ with $\sigma_u^2 = 1, \sigma_v^2/T = 0.1$ and N = 10, 100, 1000
Density of $\hat{u}$ and $\tilde{u}$ with $\sigma_u^2 = 1, \sigma_v^2/T = 0.01$ and N = 10, 100, 1000......28

FIGURE 1.8 Density of $\hat{u} \mid u = 0.1$ and $\tilde{u} \mid u = 0.1$ with $\sigma_u^2 = 1, \sigma_v^2/T = 1$ and N = 10, 100, 1000
Density of $\hat{u} \mid u = 0.1$ and $\tilde{u} \mid u = 0.1$ with $\sigma_u^2 = 1, \sigma_v^2/T = 0.1$ and N = 10, 100, 1000
Density of $\hat{u} \mid u = 0.1$ and $\tilde{u} \mid u = 0.1$ with $\sigma_u^2 = 1, \sigma_v^2/T = 0.01$ and N = 10, 100, 1000...29

FIGURE 1.9 Density of $\hat{u} \mid u = 2$ and $\tilde{u} \mid u = 2$ with $\sigma_u^2 = 1, \sigma_v^2/T = 1$ and N = 10, 100, 1000
Density of $\hat{u} \mid u = 2$ and $\tilde{u} \mid u = 2$ with $\sigma_u^2 = 1, \sigma_v^2/T = 0.1$ and N = 10, 100, 1000
Density of $\hat{u} \mid u = 2$ and $\tilde{u} \mid u = 2$ with $\sigma_u^2 = 1, \sigma_v^2/T = 0.01$ and N = 10, 100, 1000...30

Essay 1

Wang, W.S. and P. Schmidt (2009), "On the Distribution of Estimated Technical Efficiency in Stochastic Frontier Models," *Journal of Econometrics* 148, 36-45.

Essay 1

ON THE DISTRIBUTION OF ESTIMATED TECHNICAL EFFICIENCY IN STOCHASTIC FRONTIER MODELS

1.1 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

$$(1.1) \qquad y_i = X_i\beta + \varepsilon_i \quad , \quad \varepsilon_i = v_i - u_i \quad , \quad u_i \geq 0 .$$

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.

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:

$$\text{var}(u_i) = \text{var}[E(u_i|\varepsilon_i)] + E[\text{var}(u_i|\varepsilon_i)] \tag{1.2}$$

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, 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 relevance of these results is not for inference about u_i. To construct confidence intervals for u_i, we simply need the distribution of u_i conditional on ε_i, as in Horrace and Schmidt (1996). Rather, the practical usefulness of our results is for testing the adequacy (goodness of fit) of the assumed distribution of u_i. Partly our message is negative: as noted above, it is not legitimate to test the model's distributional assumptions by comparing the observed distribution of $\hat{u}_i$ to the assumed distribution of u_i. However, there is also a positive message: it *is* legitimate to test the model's distributional assumptions by comparing the observed distribution of $\hat{u}_i$ to the distribution that it should have if the model's distributional assumptions are correct. That distribution is what is given in this paper. The mechanics of such a test are the subject of a subsequent paper.

Although the exposition so far is for the cross-sectional case, our analysis also applies to the case of panel data, as in Pitt and Lee (1981) and Battese and Coelli (1988). However, in the case of panel data, one can also consider the alternative of a fixed-effects

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

treatment as in Schmidt and Sickles (1984). We analyze the distribution of the fixed-effects estimate of u_i via simulations, and compare it to the distribution of the random-effects estimate that assumes a distribution for u_i. The fixed-effect estimates show serious bias, as expected, unless σ_v^2 is quite small and/or the time-series sample size is quite large. However, when σ_v^2 / T is small and the number of firms is not too large, the fixed-effect estimates are a reasonable alternative to $\hat{u}_i$.

The plan of the paper is as follows. Section 1.2 considers the distribution of $\hat{u}_i$. Section 1.3 considers the distribution of $\hat{u}_i$ conditional on u_i. Section 1.4 discusses the case of panel data. Section 1.5 gives our concluding remarks. There is also an Appendix which contains some of the derivations.

1.2 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):

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

$\lambda(s) = \phi(s)/[1 - \Phi(s)]$, 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

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

We cannot express the function g analytically, but it is well defined and we can calculate it. For example, Figure 1.1 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 (1.4), we have

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

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

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

---

[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.

Also, we can calculate the Jacobian term $|g'(\hat{u})|$. We show in Appendix 1-A that

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

So, substituting (1.6) and (1.7) into (1.5), we obtain

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

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 1-B.

THEOREM 1.2.1:

(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 $[(\pi-2)/\pi]^2 \cdot (\sigma_u^4 / \sigma_v^2)$.

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 $\sqrt{(2/\pi)} \cdot \sigma_u$, or about $0.80 \cdot \sigma_u$.[3] Figure 1.2 gives the density of $\hat{u}$ for σ_v^2 = 0.1, 1, 10 and 100. 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 1.2.1. The approximate normality of the distribution of $\hat{u}$ for large σ_v^2 is evident in the last panel (corresponding to σ_v^2 = 100).

Figure 1.3 contains the four graphs that were in Figure 1.2, plus the half-normal density, on a common set of axes. 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.

1.3 THE DISTRIBUTION OF $\hat{u}$ CONDITONAL ON u

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

$$(1.9) \qquad f(\hat{u}|u) = \frac{a^2 \bullet \exp[-(1/2\sigma_v^2)(g(\hat{u})+u)^2]}{\sqrt{(2\pi)} \bullet \sigma_u^2 \bullet \sigma_v^2 \bullet |-1+\lambda'(g(\hat{u})/\sigma_0)|}.$$

The derivation is given in Appendix 1-C.

Theorem 1.2.1 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 1-D.

THEOREM 1.3.1:

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

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

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 1.2.1.

Figure 1.4 gives the density of $\hat{u}$ conditional on u, for $u = 0.1$, $\sigma_u^2 = 1$, and $\sigma_v^2 =$ 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 the first and last panels of Figure 1.4, respectively. For intermediate values of σ_v^2 the distribution does not look normal.

Figure 1.5 gives 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 1.6 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 1.6 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.

We emphasize that the fact that the conditional expectations estimate $\hat{u}$ underestimates large realizations of u and overestimates small realizations does not mean that there is anything "wrong" with this estimator. It is, after all, the minimum mean square error estimate of u, and it is unbiased in the unconditional sense [$E(\hat{u}-u)=0$] even though it is not unbiased in the conditional sense [$E(\hat{u}|u)=u$]. Waldman (1984) considers two alternatives: (i) the "best linear predictor" $\breve{u}=a+b\varepsilon$, where $b=-\text{var}(u)/[\text{var}(u)+\text{var}(v)]$ and $a=(1+\beta)E(u)$; and (ii) the "linear unbiased estimator" $\ddot{u}=-\varepsilon$. The best linear predictor is also a shrinkage estimator and so it also underestimates large u and overestimates small u. The linear unbiased estimator has the conditional unbiasedness property and is the only estimator that does, so far as we are aware. However, we do not find it very appealing, because it makes no attempt to remove noise, and indeed Waldman's calculations show that it performs very poorly, in terms of mean square error or in terms of correlation with u, if there is substantial noise in the model.

1.4 PANEL DATA

Although the exposition so far is for the cross-sectional case, our analysis also applies to the case of panel data, as in Pitt and Lee (1981) and Battese and Coelli (1988). Now the model is

(1.10) $$y_{it}=X_{it}\beta+\varepsilon_{it} \quad , \quad \varepsilon_{it}=v_{it}-u_i \quad , \quad i=1,\ldots,N \ , \ t=1,\ldots,T.$$

Note that the u_i are time-invariant. For the moment we assume that the v_{it} are i.i.d. normal and the u_i are i.i.d. half normal. We will call this the random effects case.

Ignoring the effect of parameter estimation and suppressing the subscript i, as above, we observe $\varepsilon_1,...,\varepsilon_T$ and the estimate of u is $\hat{u} = E(u|\varepsilon_1,...,\varepsilon_T)$, as suggested by Battese and Coelli (1988). However, for the case that the v's are normal, this is the same as $\hat{u} = E(u|\bar{\varepsilon})$ where obviously $\bar{\varepsilon} = \bar{v} - u$ and $\bar{v}$ is normal with mean zero and variance σ_v^2 / T. Therefore the results of Sections 1.2 and 1.3 apply also to the random effects panel data case, if we simply reinterpret σ_v^2 as σ_v^2 / T.

We now consider the fixed effects case, as in Schmidt and Sickles (1984). Here we would obtain an estimated intercept, say $\hat{\alpha}_i$, for each firm, and then $\tilde{u}_i = (\max_j \hat{\alpha}_j) - \hat{\alpha}_i$. Ignoring estimation error in β, this is the same as $\tilde{u}_i = (\max_j \bar{\varepsilon}_j) - \bar{\varepsilon}_i$. For small (fixed) N, this is best regarded as an estimate of $u_i^* = u_i - (\min_j u_j) < u_i$. As $N \to \infty$, $(\min_{j=1}^{N} u_j) \to 0$ and the distinction between u_i and u_i^* disappears. The relevance of this distinction is that $\hat{u}_i$ of the previous section is an estimate of u_i. While $\tilde{u}_i$ is biased upward as an estimate of u_i^*, because of the "max" operation, it may or may not be biased upward as an estimate of u_i, unless N is large.

The $\bar{\varepsilon}_i$ are the difference between a variable distributed as $N(0, \sigma_v^2 / T)$ and a variable distributed as $N(0, \sigma_u^2)^+$. So the distribution of $\tilde{u}_i$ depends on σ_v^2 / T, σ_u^2 and N.

We cannot find an analytical (closed form) expression for the density of $\widetilde{u}$,[4] so we resort to simulation to generate it. Our simulations are based on 100,000 replications. We can then compare the density of $\widetilde{u}$ (for various values of N) to the density of $\hat{u}$ (which does not depend on N). For $\widetilde{u}$ we will consider $N = 10$, 100 and 1000. We also did simulations for $N = 2000$ but they were not very different from those for $N = 1000$, and in any case such large values of N do not seem empirically relevant for stochastic frontier models.

Figure 1.7 compares the density of $\hat{u}$, and the densities of $\widetilde{u}$ for $N = 10$, 100 and 1000, for $\sigma_u^2 = 1$ and for $\sigma_v^2 / T = 1, 0.1$ and 0.01. When $\sigma_v^2 / T = 1$, it is easy to see the upward bias of $\widetilde{u}$, in the sense that the distributions are centered well to the right of $E(u) = 0.8$. As expected, this is so especially for the larger values of N. This bias diminishes as σ_v^2 / T decreases. When $\sigma_v^2 / T = 0.01$, there does not appear to be much bias and the density of $\widetilde{u}$ for $N = 10$ bears a fairly close resemblance to half normal.

Perhaps a more interesting issue is the behavior of the densities conditional on u. Here we know that $\hat{u}$ is not conditionally unbiased because it is a shrinkage toward the mean, and we expect that $\tilde{u}$ will also fail to be conditionally unbiased because it is biased upward due to the max operation. Figure 1.8 gives the results for the distribution conditional on $u = 0.1$, a very small value of u. In the first panel, with lots of noise ($\sigma_v^2 / T = 1$), all of the point estimates are biased upward, and $\tilde{u}$ is worse than $\hat{u}$. As σ_v^2 / T decreases all of the estimates improve, but $\tilde{u}$ is still pretty bad for $N = 100$ and 1000.

---

4 We can derive the joint distribution of $(\max_j \bar{\varepsilon}_j)$ and $\bar{\varepsilon}_i$, but the density of the difference between these two quantities requires an integral that we cannot calculate.

Figure 1.9 gives the same kinds of results but for the distribution conditional on u = 2, a very large value of u. When there is lots of noise ($\sigma_v^2 / T = 1$), the downward bias of $\hat{u}$ and the upward bias of $\tilde{u}$ are apparent. As the amount of noise decreases, all of the estimates improve, but once again $\widetilde{u}$ is still pretty bad for N = 100 and 1000. For N = 10, $\tilde{u}$ is nearly conditionally unbiased for the two smaller values of σ_v^2 / T.

To interpret this last result, one should remember that for small N the difference between u_i and $u_i^* = u_i - (\min_j u_j)$ becomes relevant. With $\sigma_u^2 = 1$, the expected value of $(\min_j u_j)$ is about 0.13 for N = 10. This explains why, in the last panel of Figure 1.9, the density of $\tilde{u}$ is centered to the *left* of the true value of u = 2, and also why its bias performance in the panel corresponding to $\sigma_v^2 / T = 0.1$ is as favorable as it is. However, having said that, it remains the case that when N is small and there is not much noise, $\tilde{u}$ is a reasonably good estimator.

1.5 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. A goodness of fit test, whether formal or informal, should compare the distribution of $\hat{u}$ to the distribution it should have when u is half normal, which is what this paper provides. 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. The amount of shrinkage depends on the amount of noise in the model; it is large when there is lots of noise (σ_v^2 is large) and it is small when there is little noise.

We also consider the distribution of the estimate $\tilde{u}$ from the fixed effects panel data model. This suffers from a well-known upward bias due to the maximum involved in the estimation of the efficient frontier. On average we overestimate both small realizations and large realizations of u. This bias can be severe, but it is not large when N is relatively small and when there is not much noise.

Our summary of these results is straightforward. If we have the distributional assumptions correct, it is hard to argue with $\hat{u}$, which after all is the optimal (rational, minimum mean square error, ...) forecast of u. However, if we have panel data where N is not too large and where σ_v^2 / T is small relative to σ_u^2, the estimate $\tilde{u}$ based on the fixed effects model is a plausible alternative, and it has the advantage of not depending on a distributional assumption.

APPENDIX 1

Appendix 1-A Derivation of the Jacobian in equation (1.7)

From equation (1.3), we have $\hat{u} \doteq h(\varepsilon) = k \bullet [-\varepsilon + \sigma_0 \bullet \lambda(\varepsilon / \sigma_0)]$, where $k = \sigma_u^2 / (\sigma_u^2 + \sigma_v^2)$.

So

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

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

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

Appendix 1-B Proof of Theorem 1.2.1

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 1.2.1.

To prove part (2), consider the density of $\hat{u}$ as given in equation (1.8) 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

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

The first term on the r.h.s. of (1.13) equals

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

where "$A \simeq B$" means that $A - B \to 0$ 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 (1.14) is

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

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

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

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

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

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

and so the term in (1.15) becomes

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

Also

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

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

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

Combining (1.18) with (1.14), we have

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

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

Appendix 1-C Derivation of $f(\hat{u}|u)$ in equation (1.9)

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 (1.7) of the text. Therefore the joint density of $(u,\hat{u})$ is

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

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

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

Substituting the normal density for f_v and the Jacobian expression in (1.7), we arrive at the expression in equation (1.9) of the text.

Appendix 1-D Proof of Theorem 1.3.1

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

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

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:

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

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

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

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 1.2.1, and is therefore omitted.

Figure 1.1

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

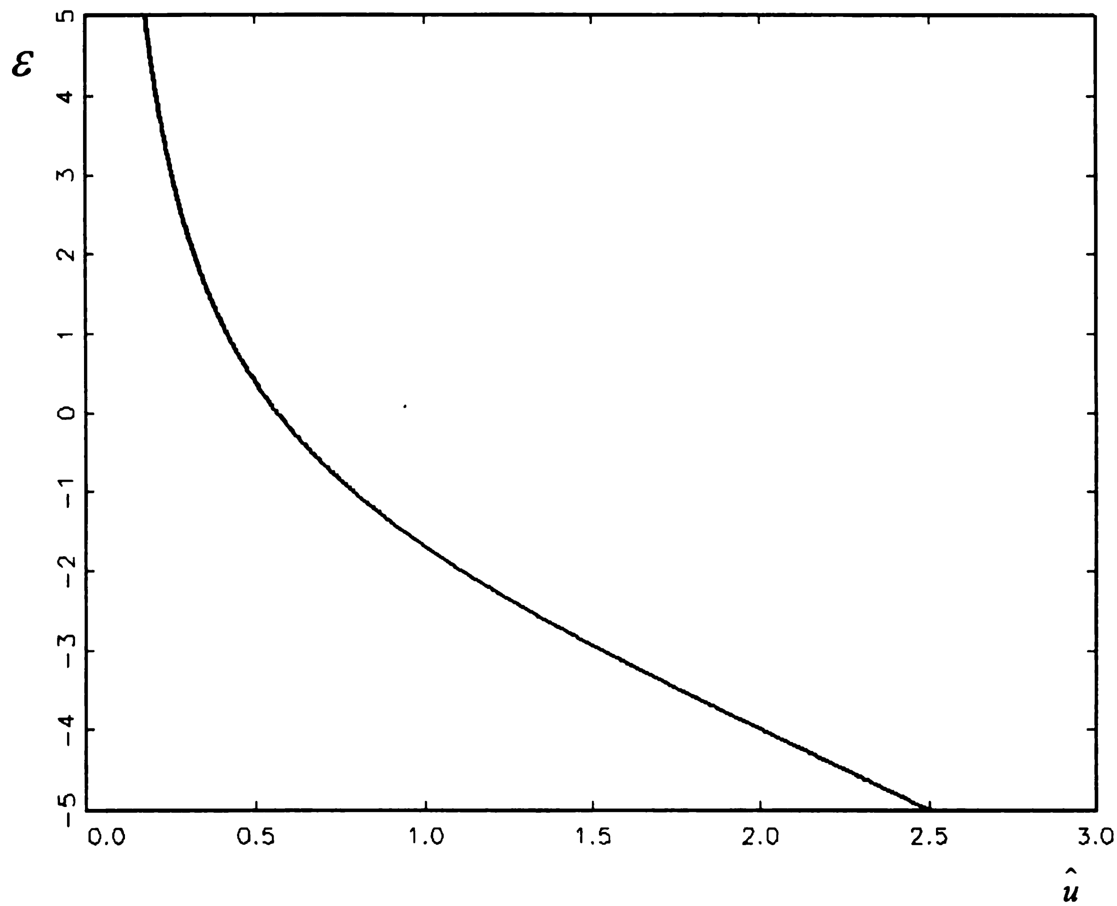


Figure 1.2
Density of $\hat{u}$ with $\sigma_u^2 = 1$ and $\sigma_v^2 = .1,\ 1,\ 10,\ 100$

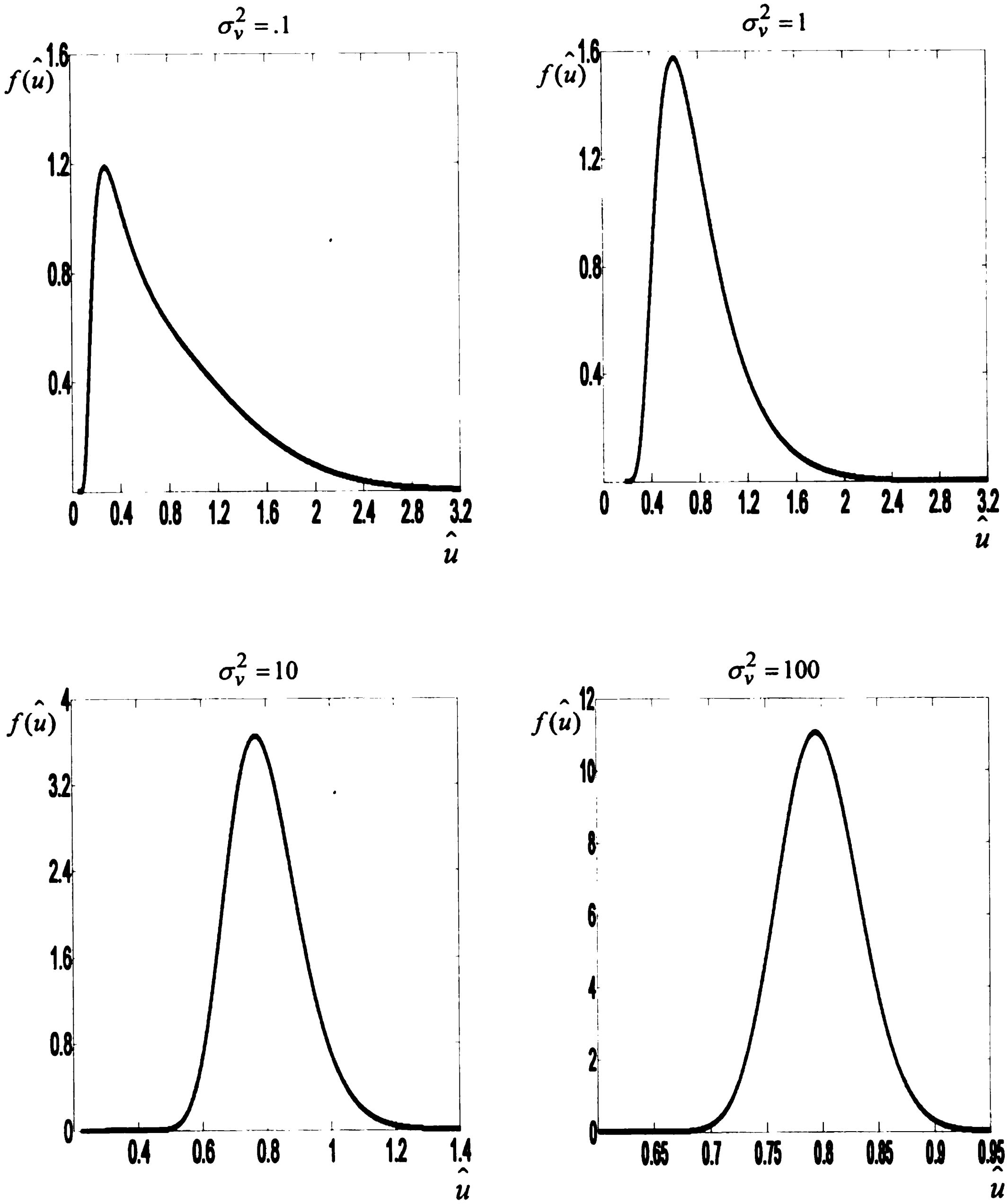


Figure 1.3 The combined graph for Figure 1.2

Density of a half normal and $\hat{u}$

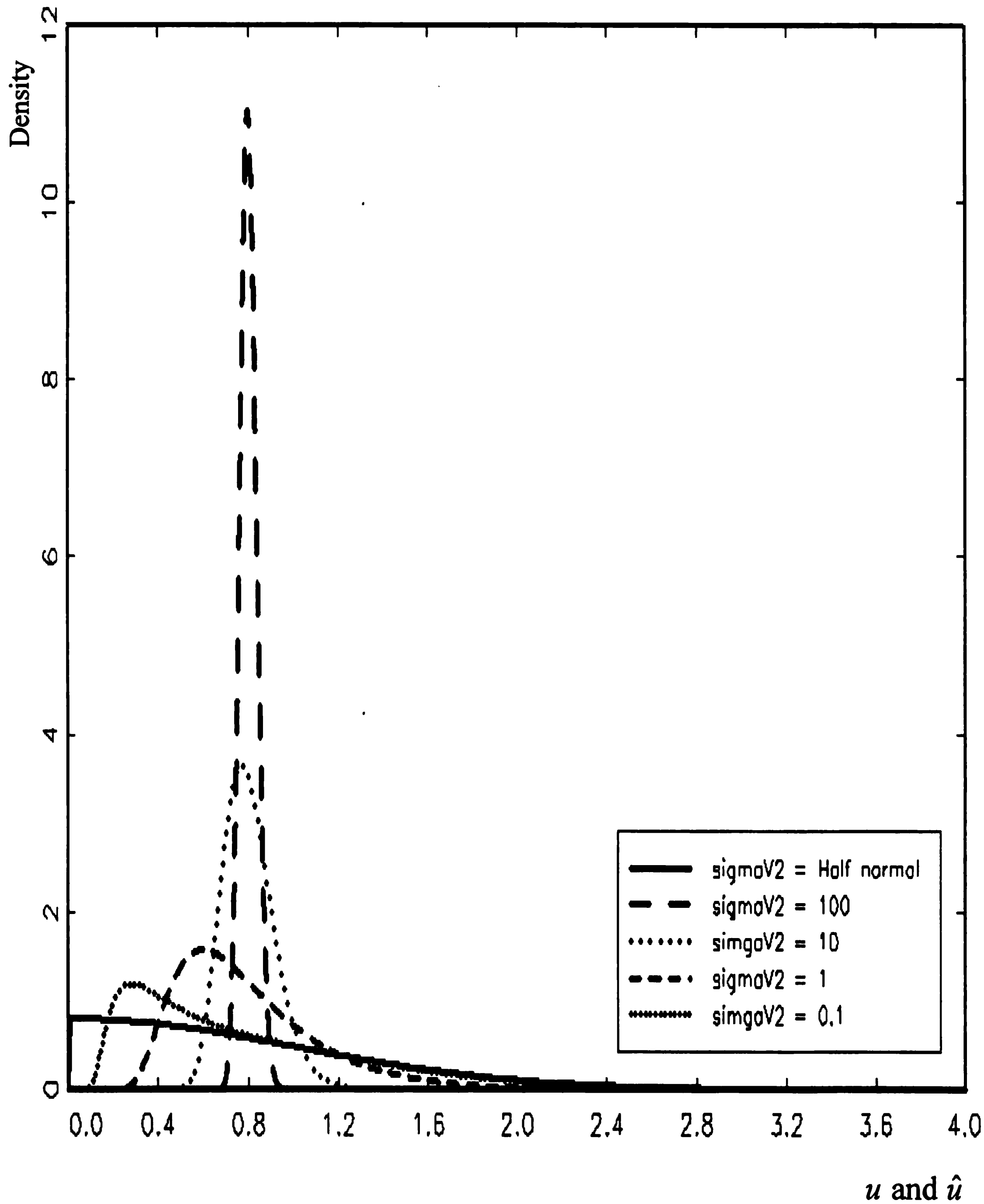


Figure 1.4

Density of $\hat{u} \mid u = .1$ with $\sigma_u^2 = 1$ and $\sigma_v^2 = .001, .01, .1, 1, 10, 100$

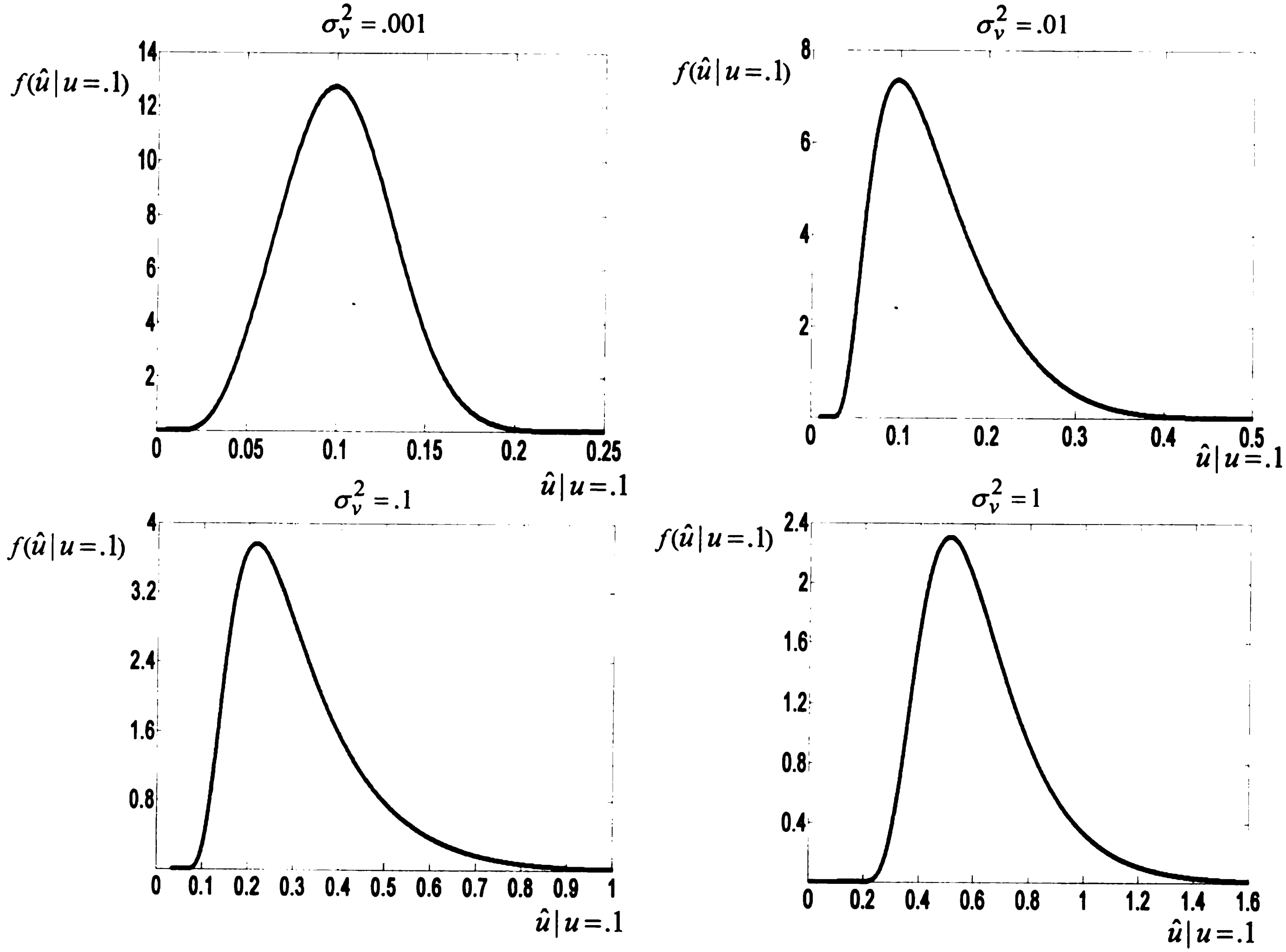


Figure 1.4 (cont'd)

Density of $\hat{u} \mid u = .1$ with $\sigma_u^2 = 1$ and $\sigma_v^2 = .001, .01, .1, 1, 10, 100$

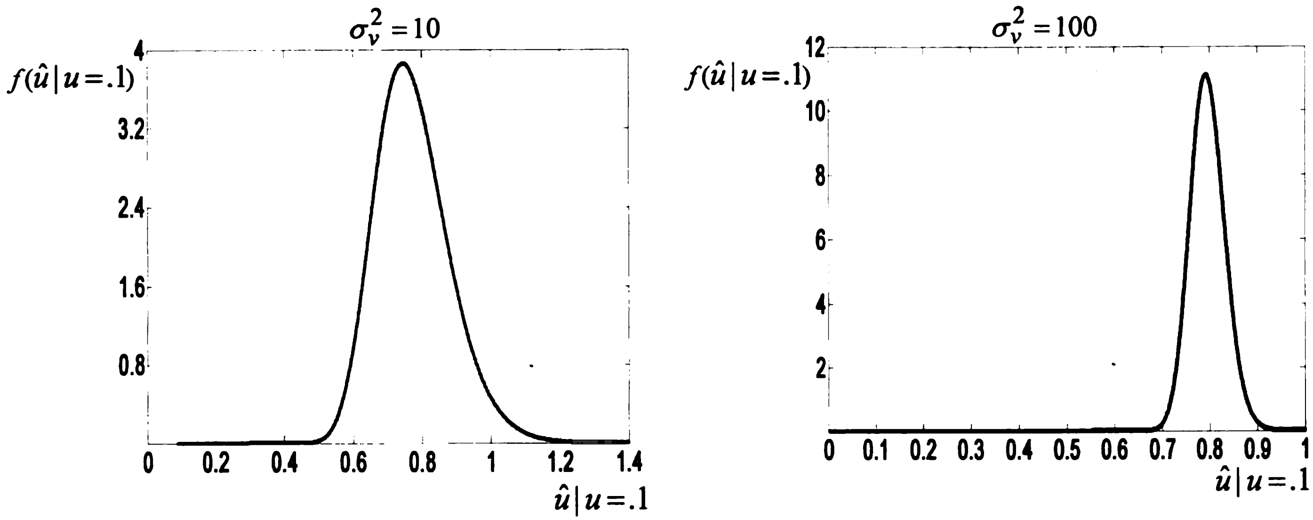


Figure 1.5

Density of $\hat{u} \mid u=2$ with $\sigma_u^2 = 1$ and $\sigma_v^2 = .001, .01, .1, 1, 10, 100$

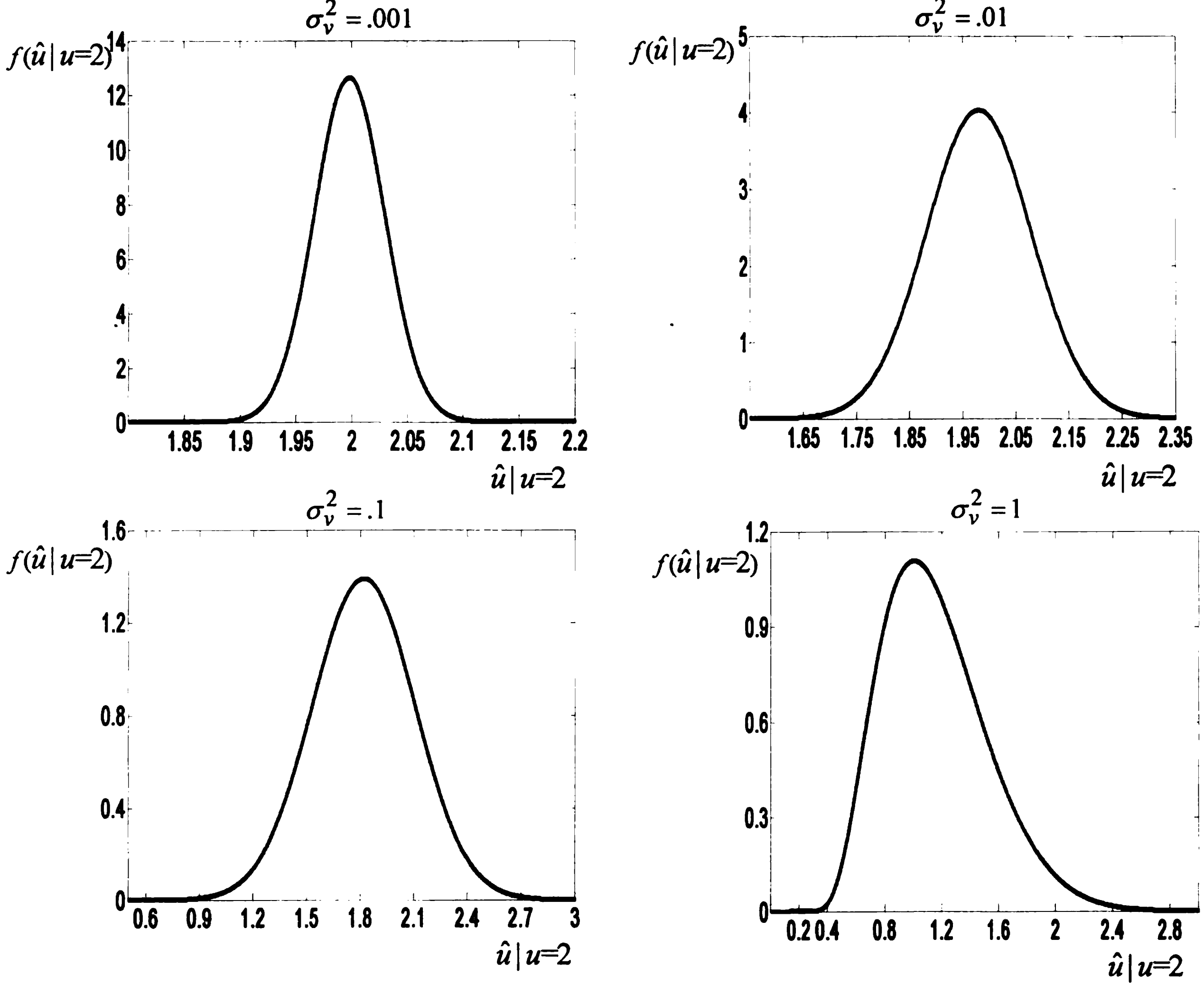


Figure 1.5 (cont'd)

Density of $\hat{u} \mid u=2$ with $\sigma_u^2 = 1$ and $\sigma_v^2 = .001, .01, .1, 1, 10, 100$

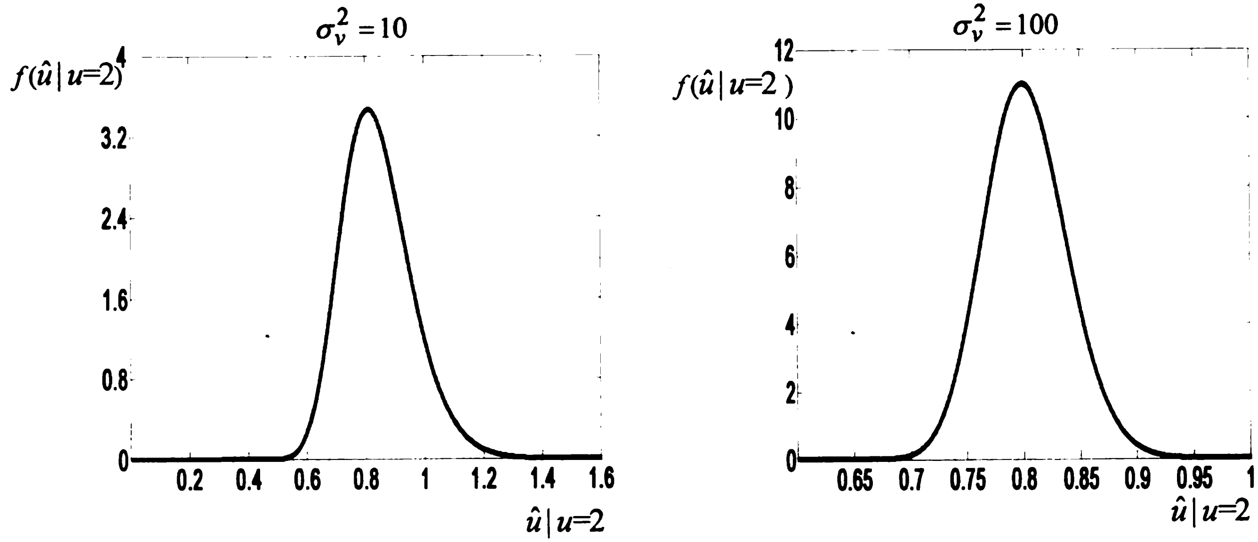


Figure 1.6

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

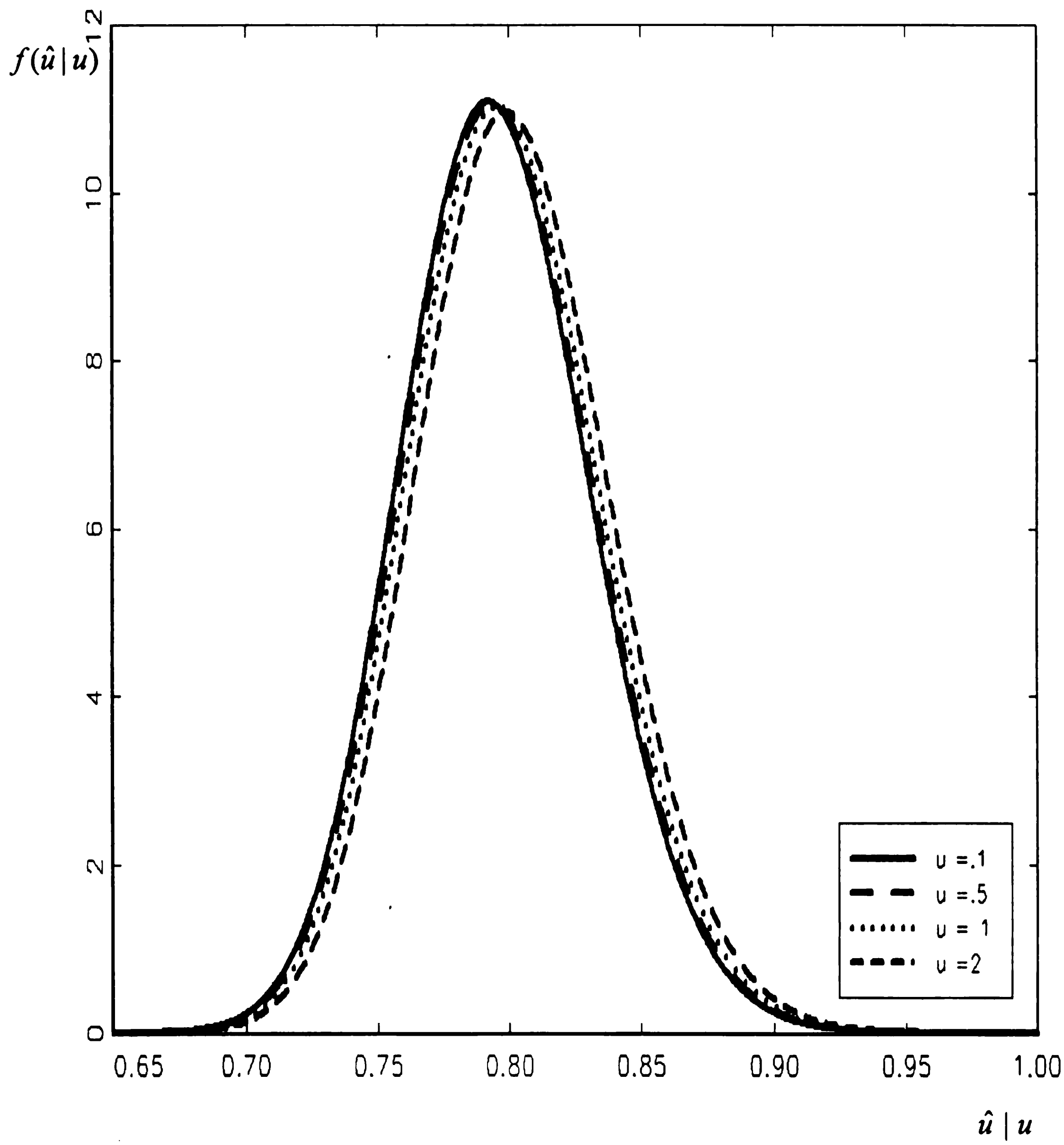
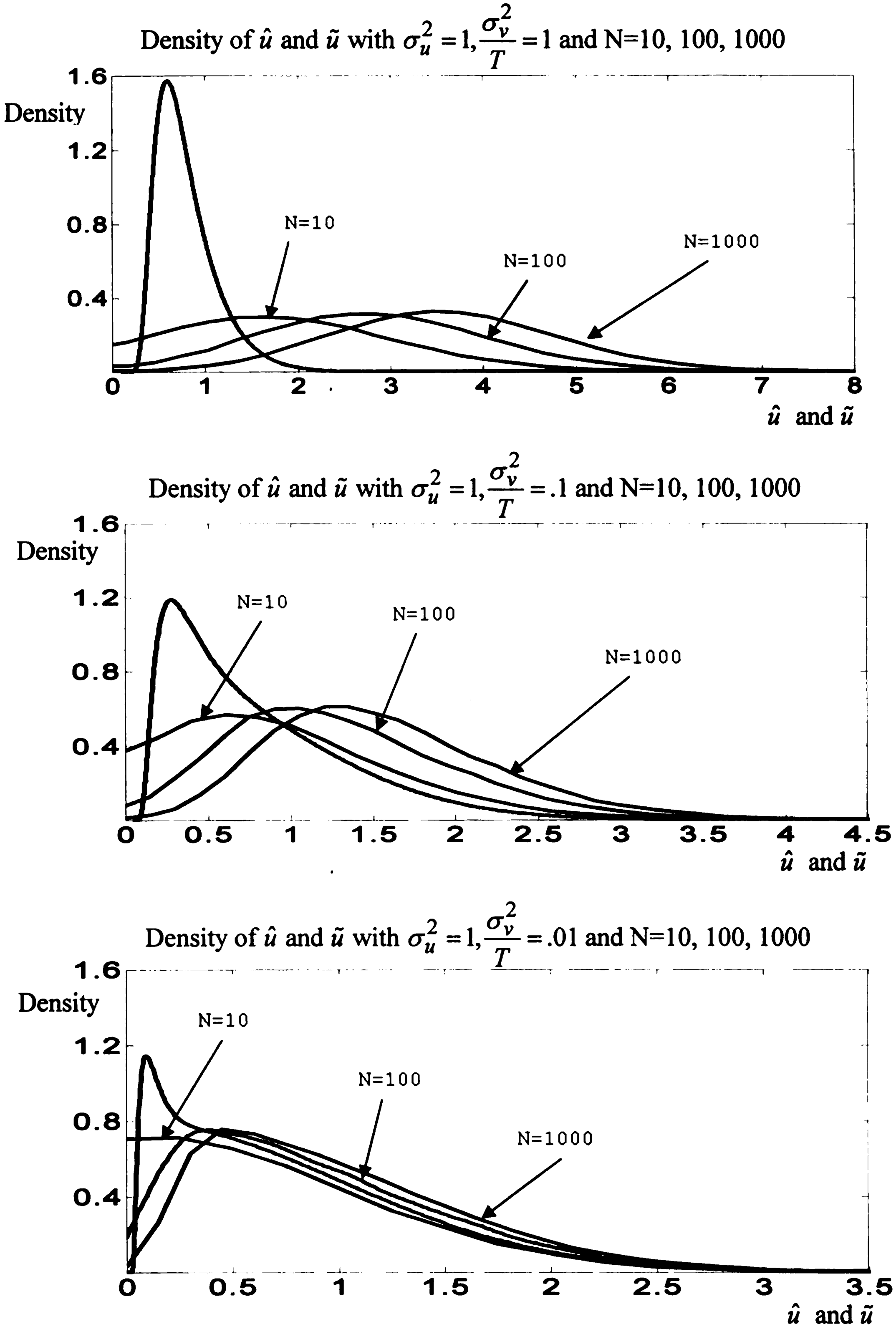


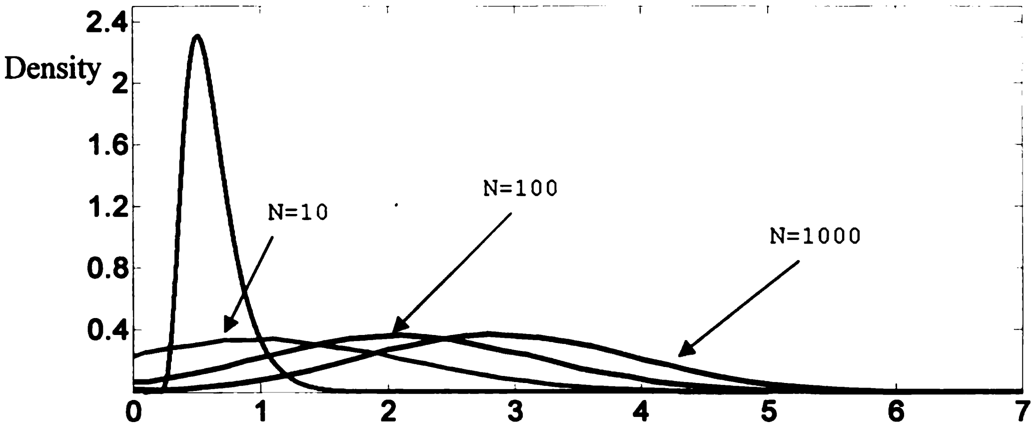
Figure 1.7



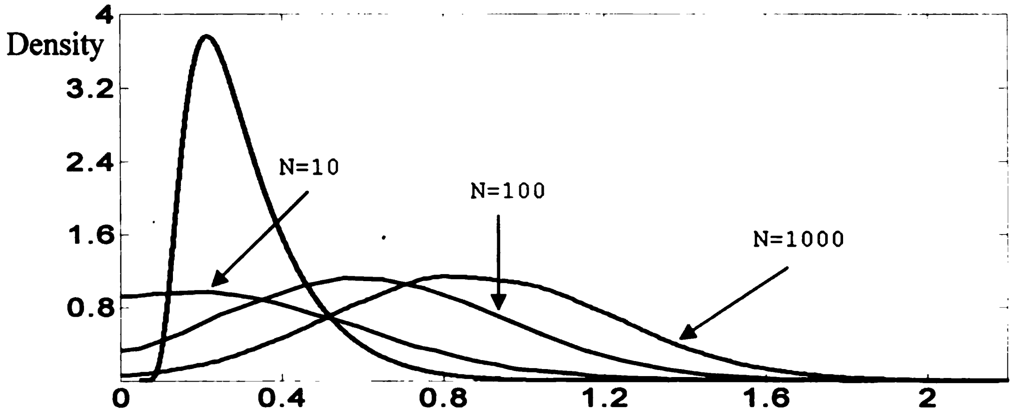
Density of $\hat{u}$ and $\tilde{u}$ with $\sigma_u^2 = 1, \frac{\sigma_v^2}{T} = 1$ and N=10, 100, 1000
Density
1.6
1.2
0.8
0.4
N=10
N=100
N=1000
0
1
2
3
4
5
6
7
8
$\hat{u}$ and $\tilde{u}$
Density of $\hat{u}$ and $\tilde{u}$ with $\sigma_u^2 = 1, \frac{\sigma_v^2}{T} = .1$ and N=10, 100, 1000
Density
1.6
1.2
0.8
0.4
N=10
N=100
N=1000
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
$\hat{u}$ and $\tilde{u}$
Density of $\hat{u}$ and $\tilde{u}$ with $\sigma_u^2 = 1, \frac{\sigma_v^2}{T} = .01$ and N=10, 100, 1000
Density
1.6
1.2
0.8
0.4
N=10
N=100
N=1000
0
0.5
1
1.5
2
2.5
3
3.5
$\hat{u}$ and $\tilde{u}$

Figure 1.8

Density of $\hat{u} \mid u = .1$ and $\tilde{u} \mid u = .1$ with $\sigma_u^2 = 1, \frac{\sigma_v^2}{T} = 1$ and N=10, 100, 1000



Density of $\hat{u} \mid u = .1$ and $\tilde{u} \mid u = .1$ with $\sigma_u^2 = 1, \frac{\sigma_v^2}{T} = .1$ and N=10, 100, 1000



Density of $\hat{u} \mid u = .1$ and $\tilde{u} \mid u = .1$ with $\sigma_u^2 = 1, \frac{\sigma_v^2}{T} = .01$ and N=10, 100, 1000

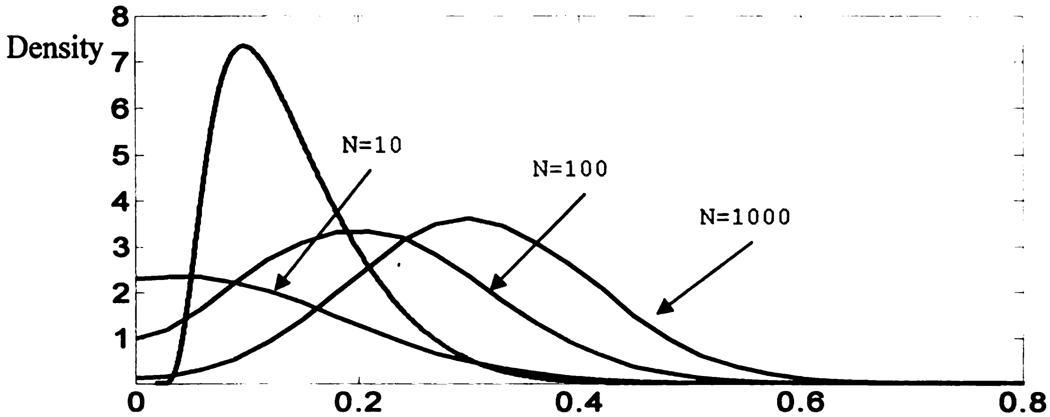
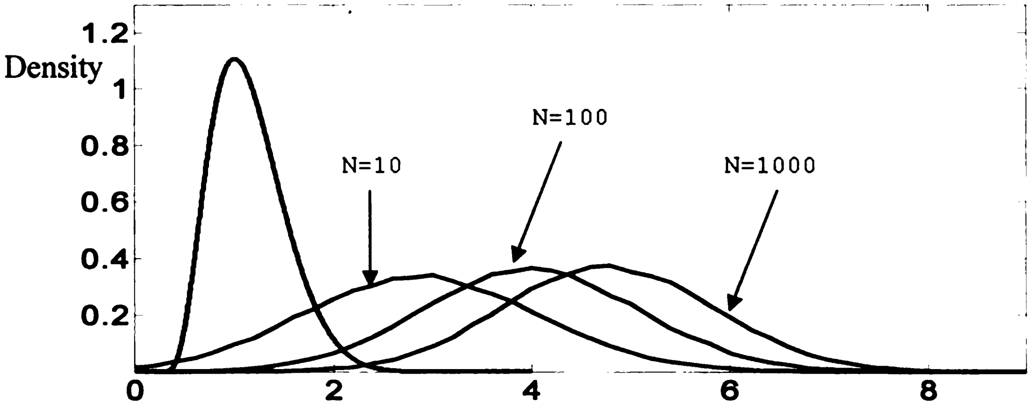
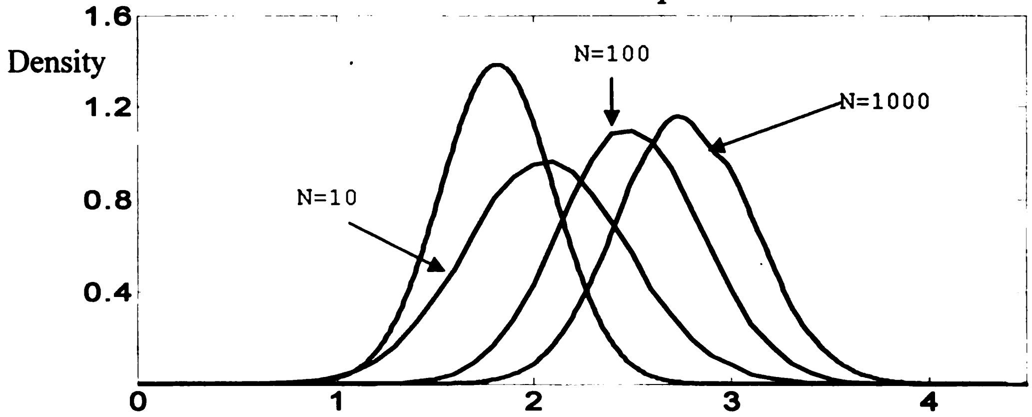


Figure 1.9

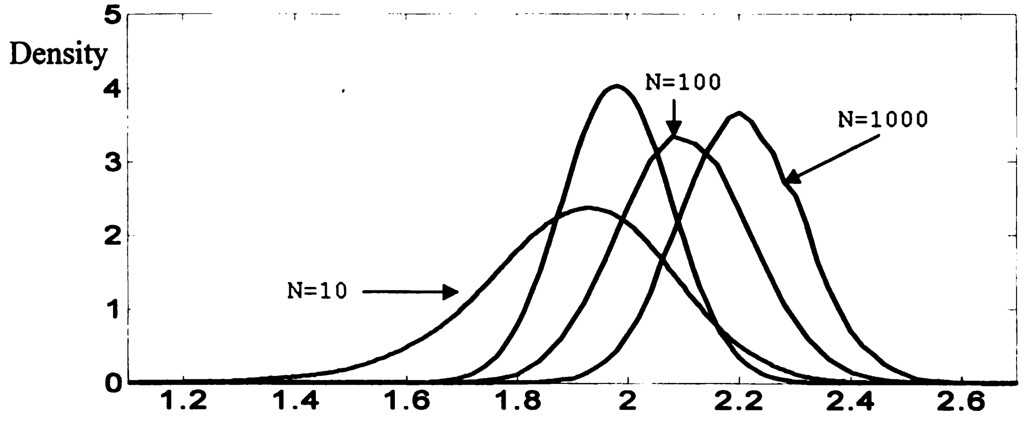
Density of $\hat{u} \mid u = 2$ and $\tilde{u} \mid u = 2$ with $\sigma_u^2 = 1, \frac{\sigma_v^2}{T} = 1$ and N=10, 100, 1000



Density of $\hat{u} \mid u = 2$ and $\tilde{u} \mid u = 2$ with $\sigma_u^2 = 1, \frac{\sigma_v^2}{T} = .1$ and N=10, 100, 1000



Density of $\hat{u} \mid u = 2$ and $\tilde{u} \mid u = 2$ with $\sigma_u^2 = 1, \frac{\sigma_v^2}{T} = .01$ and N=10, 100, 1000



Essay 2

GOODNESS OF FIT TESTS IN STOCHASTIC FRONTIER MODELS

2.1 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

(2.1) $$y_i = X_i\beta + \varepsilon_i \quad , \quad \varepsilon_i = v_i - u_i \quad , \quad u_i \geq 0 \quad , \; \mathrm{i} = 1,\ldots n.$$

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)). Our exposition is for the cross-sectional case, but we could also consider panel data as in Pitt and Lee (1981).

Our interest is in testing the distributional assumption on u_i. We will do this while maintaining the other assumptions that underlie the model, such as the functional form of the regression, the exogeneity of the X_i, and the normality of v_i. This viewpoint is motivated by the fact that in this literature the specification of the distribution of u_i is often regarded as being subject to the most doubt.

The problem then arises that u_i is not observable, and in fact cannot be consistently estimated. To be more precise, 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 $\hat{u}_i = E(u_i|\varepsilon_i)$, evaluated at $\varepsilon_i = \hat{\varepsilon}_i$. The distribution of $\hat{u}_i$ has been derived by Wang and Schmidt (2009). It is not the same as the distribution of u_i, even for large *n*. Therefore it is not legitimate to test goodness of fit by comparing the observed distribution of $\hat{u}$ to the assumed distribution of *u*. It is legitimate to test goodness of fit by comparing the observed distribution of $\hat{u}$ to the distribution derived by Wang and Schmidt. However, it is easier to base the tests instead on the distribution of ε_i that is implied by normality of v_i and the assumed distribution of u_i. This is reasonable because, given that normality of v_i is maintained, a rejection of the implied distribution of ε_i is a rejection of the assumed distribution of u_i.

We consider the usual χ^2 goodness of fit test based on expected and actual numbers of observations in cells, and also the Kolmogorov-Smirnov test based on the maximal difference between the empirical and theoretical cdf. For these tests the only technical problem of note is how to handle the issue of parameter estimation. This is relevant because both the "observations" $\hat{\varepsilon}_i = y_i - X_i\hat{\beta}$ and the expected numbers of observations in various cells depend on estimated parameters. For the chi-squared test, the relevant asymptotic theory was developed by Heckman (1984), Tauchen (1985) and Newey (1985), and we explain how this theory allows asymptotically valid tests in the stochastic frontier setting. For the Kolmogorov-Smirnov test, the comparable asymptotic theory does not exist. However, the bootstrap can be used to construct asymptotically valid tests (either for the chi-squared test or for the Kolmogorov-Smirnov test).

The plan of the paper is as follows. In Section 2.2 we discuss further the basics of goodness of fit testing in the stochastic frontier model. Sections 2.3 and 2.4 contain a general exposition of goodness of fit tests for simple and composite hypotheses, respectively. Section 2.5 gives a brief discussion of a prototypical problem, testing for normality, and presents some simulations. In Section 2.6 we discuss the problem of main interest, testing the error distribution in the stochastic frontier model, and we present detailed simulation evidence on the accuracy (size) and the power of various tests. Finally, Section 2.7 gives our concluding remarks.

2.2 TESTS BASED ON THE DISTRIBUTION OF ε

As noted above, the usual estimate of u_i is $\hat{u}_i = E(u_i|\varepsilon_i)$. (This is evaluated at $\varepsilon_i = \hat{\varepsilon}_i$, a point that we ignore in the rest of this section but address subsequently, when we discuss the relevance of allowing for the effects of parameter estimation.) The distribution of $\hat{u}_i$ is given by Wang and Schmidt (2009). It depends on the assumed distributions for both v_i and u_i, and it is not the same as the distribution of u_i. Therefore it is not legitimate to test goodness of fit by comparing the observed distribution of $\hat{u}$ to the assumed distribution of u. So, for example, if u is assumed to be half-normal, this does not imply that $\hat{u}$ should be half-normal, and it is not correct to test the half-normal assumption by seeing whether the distribution of $\hat{u}$ appears to be half-normal.

This does not mean that the observed distribution of $\hat{u}$ is uninformative. It is perfectly legitimate to test goodness of fit by comparing the observed distribution of $\hat{u}$ to the distribution that it should have under the distributional assumptions being made, as derived by Wang and Schmidt. Because this distribution depends on the distribution of both v and u, we have to maintain the correctness of the assumed (normal) distribution of v to test the correctness of the distributional assumption on u. This issue is inevitable in this context.

Such a comparison is complicated because the distribution of $\hat{u}$ is complicated. It is much easier to base a goodness of fit test on the distribution of ε. The distribution of ε also follows from the assumed distributions of v and u, and so if we maintain the correctness of the assumed distribution for v, we can test the correctness of the assumed distribution for u a goodness of fit test based on the distribution of ε. This is computationally easier than a test based on the distribution of $\hat{u}$. The following simple point is therefore relevant: $\hat{u}$ is a monotonic function of ε. This implies that most goodness of fit tests based on the distribution of $\hat{u}$ will be equivalent to the same goodness of fit tests based on the distribution of ε. For example, the Kolmogorov-Smirmov test will be exactly the same whether it is based on the distribution of $\hat{u}$ or the distribution of ε. For the Pearson χ^2 tests based on the observed versus actual numbers of observations in cells, again the test is exactly the same whether it is based on the distribution of $\hat{u}$ or the distribution of ε, provided that the cells are defined conformably.

Therefore, for reasons of computational simplicity, we will consider tests based on the distribution of ε that is implied by the assumed distributions for v and u. We maintain the correctness of the assumed (normal) distribution of v, and therefore interpret the tests as tests of the correctness of the assumed distribution of u.

2.3 SIMPLE HYPOTHESES

Suppose that we have a random sample $y_1, y_2, \ldots, y_n$ and we wish to test the hypothesis that the population distribution is characterized by the pdf $f(y, \theta_0)$. The subscript "zero" on θ indicates the true value of the parameter θ, which we assume to be the same as the value specified by the hypothesis being tested. That is, in this section we take θ_0 as given. Thus, for example, we could be testing the simple hypothesis that y is distributed as N(0,1), as opposed to the composite hypothesis that y is normal with μ and σ^2 unspecified.

To define the Kolmogorov-Smirnov statistic, let $F(y, \theta_0)$ be the cdf corresponding to the pdf $f(y, \theta_0)$. Also let $F_n(y)$ be the empirical cdf of the sample: $F_n(y) =$ (number of $y_i \le y$)/n. Then the Kolmogorov-Smirnov statistic is

(2.2) $$\text{KS} = \sup_y \left| F(y, \theta_0) - F_n(y) \right| .$$

The asymptotic distribution of KS is known and widely tabulated. It does not depend on the form of the distribution (f, or F).

Now consider the Pearson χ^2 statistic. Let the possible range of y be split into k "cells" (intervals) $A_1, \ldots, A_k$, such that any value of y is in one and only one cell. Let

$1(y \in A_j)$ be the "indicator function" that equals one if y is in cell A_j, and equals zero otherwise. Let $p_j = p_j(\theta_0) = P(y \in A_j) = E[1(y \in A_j)]$. With n observations as above, we define the observed (O) and expected (E) numbers of observations in each cell:

(2.3) $$O_j = \sum_{i=1}^{n} 1(y_i \in A_j) \quad , \quad E_j = np_j \quad , \quad j = 1,\ldots,k .$$

Then the Pearson χ^2 statistic is given by:

(2.4) $$\chi^2 = \sum_{j=1}^{k} (O_j - E_j)^2 / E_j$$

Asymptotically (as $n \to \infty$) its distribution is chi-squared with (k-1) degrees of freedom.

It is interesting (and later it is useful) to put these results into a generalized method of moments (GMM) framework. We begin with the set of moment conditions

(2.5) $$E[g(y,\theta_0)] = 0$$

where $g(y,\theta)$ is a vector of dimension (k-1) whose j^{th} element equals $[1(y \in A_j) - p_j(\theta)]$. The subscript "zero" on θ reinforces the point that the expectation in (2.5) equals zero only at θ_0, the true value of θ. Also, note that we have omitted one cell so as to avoid a subsequent singularity. We have omitted cell A_k but the choice of which cell to omit does not matter. Now define

(2.6) $$\bar{g}(\theta) = \frac{1}{n} \sum_{i=1}^{n} g(y_i,\theta)$$

and note that the j^{th} element of $\bar{g}(\theta)$ is equal to $\frac{1}{n}(O_j - E_j(\theta))$. We also need to define the variance matrix of the moment conditions $g(y,\theta_0)$. This variance matrix is the matrix $V(\theta_0)$, of dimension (k-1) by (k-1), whose j^{th} diagonal element equals ($p_j - p_j^2$), and whose i^{th}, j^{th} off diagonal element ($i \neq j$) equals ($-p_i p_j$), with all probabilities evaluated at θ_0.

A central limit theorem implies that the asymptotic distribution of $\sqrt{n}\bar{g}(\theta_0)$ is N(0, $V(\theta_0)$). From this fact it follows that

$$(2.7) \qquad n\bar{g}(\theta_0)'V(\theta_0)^{-1}\bar{g}(\theta_0) \to_d \chi^2_{k-1} .$$

To link this to the distributional result given above for the test of the simple hypothesis that y has density $f(y,\theta_0)$, we simply observe that

$$(2.8) \qquad n\bar{g}(\theta_0)'V(\theta_0)^{-1}\bar{g}(\theta_0) = \sum_{j=1}^{k}(O_j - E_j)^2 / E_j,$$

the Pearson χ^2 statistic. The equality in (2.8) is proved in Appendix 2-A. So this establishes the distributional result given in the sentence following equation (2.4).

2.4 COMPOSITE HYPOTHESES

Now suppose that we wish to test the composite hypothesis that the population distribution is characterized by the pdf $f(y,\theta)$ for some (unspecified and unknown) value of θ. This is the empirically relevant case.

For the Kolmogorov-Smirnov test, we can estimate θ by MLE. Denote this estimate by $\hat{\theta}$. Now we can use $\hat{\theta}$ in place of θ_0 in equation (2.2) to calculate the statistic. The problem is that the distribution of the statistic is changed, even asymptotically, and furthermore there is no general asymptotic theory to show how to alter the asymptotic distribution to reflect the effects of parameter estimation. For some cases (e.g. the case that the distribution being tested is exponential), it is known that the distribution of the KS statistic using $\hat{\theta}$ does not depend on the value of θ_0, so that critical values can be calculated by simulation. However, there is no such result that would apply to the stochastic frontier model.

An asymptotically valid Kolmogorov-Smirnov test for a composite hypothesis can be constructed using bootstrapping. Let $f(y,\hat{\theta})$ be the hypothesized density, evaluated at the estimate $\hat{\theta}$. Now we use a "parametric bootstrap": for b = 1, 2, ..., B, where B (the number of bootstrap draws) is large, draw $y_1^{(b)}, y_2^{(b)}, \ldots, y_n^{(b)}$ from $f(y,\hat{\theta})$. Based on this data, calculate the estimate $\hat{\theta}^{(b)}$ and the KS statistic in (2.2) based on $\hat{\theta}^{(b)}$. Then use the critical values derived from the appropriate quantiles of the empirical distribution of these B values of the statistic. The asymptotic validity of this procedure has been established by Giné and Zinn (1990) and Stute, Gonzáles and Presedo (1993).

Next we will consider the Pearson χ^2 test. As for the Kolmogorov-Smirnov test, it is not legitimate to ignore parameter estimation. Also as for the Kolmogorov-Smirnov test, an asymptotically valid test can be obtained using critical values from the parametric bootstrap. However, for the χ^2 test the necessary asymptotic theory to correct for parameter estimation is known, and tests based on this theory are an alternative to tests using the bootstrap.

To discuss this asymptotic theory, recall that the number of cells was k, and let the dimension of θ be m, with $m \leq k-1$. We have $p_j = p_j(\theta)$ and $E_j(\theta) = np_j(\theta)$; that is, the expected numbers of observations in the cells depend on θ. Thus the value of the statistic in (2.4) or (2.8), say $\chi^2(\theta)$, depends on θ. As above, let $\hat{\theta}$ be the MLE of θ, and let $\tilde{\theta}$ be the value of θ that minimizes $\chi^2(\theta)$. A famous result is that $\chi^2(\tilde{\theta})$ is asymptotically distributed as chi-squared with (k-1-m) degrees of freedom. That is, we still have a chi-squared distribution but the number of degrees of freedom is reduced by one for every estimated parameter. This is a nice result but it is not altogether satisfying, since we have reduced the number of degrees of freedom, and because $\tilde{\theta}$ is in general an inefficient estimator. It is much more natural to consider the statistic $\chi^2(\hat{\theta})$ which uses the MLE. However, this is not asymptotically distributed as chi-squared, and using the chi-squared distribution with (k-1-m) degrees of freedom results in a test which is conservative, and therefore presumably less powerful than is possible. (For this result, and the result referred to above as famous, see, e.g., Tallis (1983, p. 457).)

To understand these results, and the way in which parameter estimation by MLE is successfully accommodated, we return to the GMM interpretation of the χ^2 test given at the end of Section 2.3. The value of θ is unknown, but we can estimate θ by GMM based on the moment conditions (2.5). In this case we will minimize the GMM criterion function

(2.9) $$n\bar{g}(\theta)'\hat{V}^{-1}\bar{g}(\theta) ,$$

where $\hat{V}$ is either $V(\theta)$, in the case of the "continuous updating" GMM estimator, or is any consistent estimate of $V(\theta_0)$, in the case of the "two step" GMM estimator. The first possibility corresponds to the minimization of $\chi^2(\theta)$ with respect to θ, and yields the estimator $\tilde{\theta}$ discussed in the previous paragraph. In either the continuous updating case or the two step case, standard GMM results indicate that the minimized value of the criterion function (2.9) is asymptotically distributed as chi-squared with degrees of freedom equal to the number of moment conditions minus the number of parameters estimated, that is, $(k\text{-}1\text{-}m)$ degrees of freedom. (This is generally referred to in the GMM literature as the "test of overidentifying restrictions.") This argument establishes the "famous result" referred to above.

The estimator $\tilde{\theta}$ is not generally efficient, and so we ought to be able to do better than this. As above, let $\hat{\theta}$ be the MLE, which is (asymptotically) efficient. Unfortunately $\chi^2(\hat{\theta})$ does not generally have a chi-squared distribution. This raises the question of whether we can construct a goodness of fit statistic based on $\hat{\theta}$ that does have a chi-squared distribution. The answer is yes, as was shown by Heckman (1984), Tauchen (1985) and Newey (1985). Our discussion will follow Tauchen. We wish to test the composite hypothesis that the density of y is $f(y,\theta)$. Define the "score function"

$$(2.10) \qquad s(y,\theta) = \frac{\partial \ln f(y,\theta)}{\partial \theta} .$$

The MLE satisfies the first order condition $\sum_{i=1}^{n} s(y_i,\hat{\theta}) = 0$ and is the GMM estimator based on the (exactly identified) set of moment conditions: $Es(y,\theta_0) = 0$. Now the technical trick that leads to the test is to combine these moment conditions based on the score function with the moment conditions that we want to test, based on numbers of observations falling into various cells. Formally we write the full set of moment conditions as $Eh(y,\theta_0) = 0$, where

$$(2.11) \qquad h(y,\theta) = \begin{bmatrix} h_1(y,\theta) \\ h_2(y,\theta) \end{bmatrix} = \begin{bmatrix} s(y,\theta) \\ g(y,\theta) \end{bmatrix} .$$

Here $h_1 = s$ is the score function and $h_2 = g$ is the vector of $(k\text{-}1)$ functions given in equation (2.5) above. We wish to maintain the correctness of h_1 (to obtain $\hat{\theta}$) and test the correctness of h_2.

The test statistic is of the form

$$\text{(2.12)} \qquad \text{CMT} = n\bar{g}(\hat{\theta})'\hat{C}^{22}\bar{g}(\hat{\theta}) \ ,$$

where $\hat{C}^{22}$ will be defined in the next paragraph. The relevant distributional result is that CMT is asymptotically distributed as chi-squared with $(k\text{-}1)$ degrees of freedom. That is, we do obtain a chi-squared limiting distribution *and* there is no loss in degrees of freedom due to estimation of θ.

The difference between this statistic and $\chi^2(\hat{\theta})$ is that the conditional moment test (CMT) uses $\hat{C}^{22}$ where $\chi^2(\hat{\theta})$ uses $V(\hat{\theta})^{-1}$. The matrix $\hat{C}^{22}$ is defined as follows. Let C be the variance matrix of the vector $h(y,\theta)$. Its dimension is $(m+k\text{-}1)$ by $(m+k\text{-}1)$. Let C^{-1} be its inverse. We partition C and C^{-1} correspondingly to the partitioning of h into h_1 and h_2:

$$\text{(2.13)} \qquad C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} , \quad C^{-1} = \begin{bmatrix} C^{11} & C^{12} \\ C^{21} & C^{22} \end{bmatrix} .$$

So C^{22} is the lower right submatrix, of dimension $(k\text{-}1)$ by $(k\text{-}1)$, of C^{-1}. Then $\hat{C}^{22}$ is any consistent estimate of C^{22}. (A specific estimate will be discussed below.) We can note that $C_{22} = V(\theta_0)$ is the variance matrix of $g(y,\theta)$ and so basically $\chi^2(\hat{\theta})$ uses (an estimate of) C_{22}^{-1} whereas CMT uses (an estimate of) C^{22}. A standard matrix equality says that

(2.14) $$C^{22} = (C_{22} - C_{21}C_{11}^{-1}C_{12})^{-1}$$

which is bigger than C_{22}^{-1}. That is the sense in which the CMT adjusts for the fact that $\chi^2(\hat{\theta})$ is too conservative.

From equation (2.14), an estimate of C^{22} requires an estimate of all of C. The most commonly used estimate is the "OPG" (for "outer product of the gradient") estimate:

(2.15) $$\hat{C} = \hat{C}(\hat{\theta}) = \frac{1}{n}\sum_{i=1}^{n} h(y_i,\hat{\theta})h(y_i,\hat{\theta})'.$$

The CMT using this estimate can be calculated as the uncentered R^2 in a regression of a constant (one) on $h(y_i,\hat{\theta})$. See Newey (1985, p. 1052) for this expression.

Alternatively, for some problems we may be able to obtain an analytical expression for C, and then evaluate it at $\hat{\theta}$. The submatrix C_{11} is the information matrix for the estimation of θ by MLE, and may be evaluated using expectations of second derivatives or cross products of first derivatives. The submatrix C_{22} can be evaluated as $V(\hat{\theta})$ where the matrix V is defined in the discussion following equation (2.16). This leaves $C_{12} = Eh_1h_2'$ for which there may be an analytical expression in certain simple cases (e.g. testing for normality), but not in general.

An alternative to the CMT test is to estimate θ by GMM using the full set of moment conditions h, and then perform the usual GMM overidentification test. Let $\breve{\theta}$ be this estimate. It is different from the MLE $\hat{\theta}$, but it has the same asymptotic distribution since the second set of moment conditions (g) is statistically redundant for estimation given the score (s). The GMM overidentification test statistic is $n\bar{h}(\breve{\theta})'\hat{C}^{-1}\bar{h}(\breve{\theta})$. This can be compared to the CMT statistic $n\bar{g}(\hat{\theta})'\hat{C}^{22}\bar{g}(\hat{\theta}) = n\bar{h}(\hat{\theta})'\hat{C}^{-1}\bar{h}(\hat{\theta})$, where the last equality follows from the fact that $\bar{s}(\hat{\theta}) = 0$. So the difference between the CMT and the overidentification test is just due to the difference between $\hat{\theta}$ and $\breve{\theta}$, which is asymptotically negligible. The reason the CMT is preferable is a matter of simplicity – if one has estimated the model by MLE, which would typically be the case, then there is no need to reestimate the model.

There is a further theoretical point worth making. The discussion above takes the cells (A_j) as given, so that the probabilities of observations falling into the cells are what depend on θ. This fits well into the GMM setting because the usual asymptotic distribution theory for GMM depends on the moment conditions being differentiable with respect to θ. In practice, however, the cell definitions will naturally depend on θ. For example, if we test normality based on five equi-probable cells, the first cell will be $(-\infty, \mu - 0.84\sigma]$. So then the probabilities of being in the various cells are given, but the observed numbers in the cells depend on μ and σ. Now the moment conditions depend on the indicators $1(y \in A_j(\theta))$ which are not differentiable with respect to θ. However, Tauchen shows that the distributional theory for the CMT still holds in this case.

An interesting theoretical detail that we have not found in the literature is the following. Instead of comparing observed and expected numbers of observations in cells, we could compare the sample and population quantiles. For example, in the normal example mentioned above, in the first case we fix the cell boundary at $(\hat{\mu} - 0.84\hat{\sigma})$ and see how close the number of observations less than this is to 0.2. Alternatively, we could calculate the 20^{th} percentile in the sample and see how close it is to $(\hat{\mu} - 0.84\hat{\sigma})$. Intuition would suggest that the difference between these two tests ought to be asymptotically negligible. The sense in which that is true is discussed in Appendix 2-B.

A final point is that it is also possible to apply the bootstrap to the asymptotically valid (Tauchen) form of the χ^2 statistic. It is well known that the bootstrap can provide higher-order refinements to the asymptotically valid distributions of estimators or test statistics. That is, using asymptotic theory plus the bootstrap may give better (more accurate in finite samples) critical values than using asymptotic theory alone or the bootstrap alone. We will investigate that issue in our simulations.

2.5 AN INTRODUCTORY EXAMPLE: TESTING NORMALITY

Our interest in this paper is testing distributional assumptions in stochastic frontier models. However, first we will present a few simulations for a simpler problem, testing normality. The point is to see how the tests work in a very simple setting, and in particular in one where we can do some things analytically that we cannot do in the more complicated model.

All of our tests will be based on standard normal data. The number of replications in the simulations is 10,000, except that for the bootstrap tests we use 1000 replications and 999 bootstrap samples per replication. We will consider sample sizes n = 50, 100, 250 and 500. We use tests with nominal size of 0.05. (We also calculated results for sizes of 0.01 and 0.10 but they would not lead to different conclusions than the results for size of 0.05.)

For the χ^2 tests we will consider three different numbers of cells: k = 3, 5 and 10. In all cases we use equiprobable cells.

We first consider the test of the simple hypothesis that the data are N(0,1). That is, the values $\mu_0 = 0$ and $\sigma_0 = 1$ are specified by the null hypothesis (i.e. assumed known). The χ^2 statistic is $n\bar{g}(\theta_0)'V(\theta_0)^{-1}\bar{g}(\theta_0)$ as given in equation (2.8) above, while the KS statistic is as given in equation (2.2). The results for these tests are given in Table 2.1, using the asymptotic critical values and the bootstrapped critical values. They are very easy to summarize. All of the tests are quite accurate, in the sense that actual size is quite close to nominal size. The KS test is very slightly undersized for the smaller sample sizes (n = 50 and 100) but this is not a serious discrepancy. Bootstrapping fixes this problem.

Next we consider the test of the composite hypothesis that the data are N(μ,σ^2) for unknown μ and σ^2. In Table 2.2 we give the empirical size of the four asymptotically valid tests discussed above. The first (under the heading "Pearson (Tauchen)") is the Tauchen version of the χ^2 statistic, equal to $n\bar{g}(\hat{\theta})'\hat{C}^{22}\bar{g}(\hat{\theta})$, as given in equation (2.12). The MLE of θ is $\hat{\theta} = (\hat{\mu},\hat{\sigma}^2)$ where $\hat{\mu} = \bar{y}$ and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2$. Then $\hat{C}^{22}$ is obtained from the inverse of $\hat{C}$, the OPG estimate as in equation (2.15), with $h(y,\theta)$ evaluated at the MLE $\hat{\theta}$. This expression requires the score function for the normal density, which is given in Appendix 2-C. The second test ("Bootstrap Pearson (Tauchen)") uses the same test statistic but uses bootstrapped critical values. The third test ("Bootstrap Pearson") is the usual Pearson χ^2 test given in (2.8), using $\hat{\theta}$ in place of θ_0, with bootstrapped critical values. The fourth test ("Bootstrap KS") is the Kolmogorov-Smirnov test given in (2.2), but using $\hat{\theta}$ in place of θ_0, with bootstrapped critical values.

The results for these four tests in Table 2.2 are easy to summarize. All of the tests that use critical values from the bootstrap are quite accurate, in the sense that empirical size is close to nominal size. The results for the Tauchen version of the Pearson test, which relies on asymptotic theory instead of the bootstrap, are less favorable. For this test there are moderate to large size distortions in small samples, especially when a large number of cells is used. For example, with $k = 3$ the actual size of the nominal 0.05 level test is 0.067 for $n = 50$ and 0.055 for $n = 100$, which is not too bad. For $k = 5$, we have actual sizes of 0.086, 0.070 and 0.057 for $n = 50$, 100 and 250, respectively, so the size distortions disappear more slowly. For $k = 10$, we have actual sizes of 0.185, 0.117, 0.079 and 0.060 for $n = 50$, 100, 250 and 500, respectively, so that a rather large sample size is required to have an accurate test. An obvious implication is not to use too many cells unless the sample size is large (or, to use the bootstrap).

It is interesting to ask why it is that this test is less accurate than the Pearson test for the simple hypothesis (Table 2.1). The current test differs from the Pearson test of the simple hypothesis in a number of ways. First, it uses cells defined on the basis of $\hat{\theta}$ rather than θ_0. Second, it obtains the weighting matrix as a submatrix of the variance matrix of the moment conditions after they have been augmented with the score. Third, it evaluates the variance matrix of the (augmented) moment conditions using the OPG estimate, as opposed to using an analytical expression.

The question is which of these differences is the one that matters. To provide evidence on this question, we consider two other tests. One ("Simple Hypothesis OPG" in

Table 2.2) is based on the test statistic for the simple hypothesis (equation (2.8) above) except that we replace the variance matrix $V(\theta_0)$ with the OPG estimate $\hat{C}_{22}(\theta_0) = \frac{1}{n}\sum_{i=1}^{n} g(y_i,\theta_0)g(y_i,\theta_0)'$. So we are using the OPG estimate unnecessarily, but we do not have estimation error in θ. The second test ("Composite Hypothesis C Known" in Table 2.2) is based on the test stastistic $n\bar{g}(\hat{\theta})'C^{22}(\theta_0)\bar{g}(\hat{\theta})$, where C^{22} is the relevant submatrix of the true (not estimated) variance matrix C of the augmented moment conditions. This is possible because for this problem we can calculate C analytically. This expression is given in Appendix 2-C. This test statistic uses the cell definitions based on $\hat{\theta}$ but uses the analytical expression for C and evaluates it at θ_0. This would be infeasible in actual practice but it is feasible in the Monte Carlo setting because we know the value of θ_0.

The results for "Composite Hypothesis C Known" are quite good. The results for "Simple Hypothesis OPG" are much less accurate. The size distortions follow the same pattern as for "Pearson (Tauchen)" though they are a little smaller. Thus it appears that the finite sample inaccuracy of the non-bootstrapped Tauchen version of the Pearson test is due primarily to the use of the OPG estimate of C. The true value of C accounts properly for the effects of parameter estimation but the OPG estimated $\hat{C}$ does not.

2.6 THE STOCHASTIC FRONTIER MODEL

We now return to the stochastic frontier model. The model is as given in equation (2.1) above. We will consider the case in which v is normal and u is half-normal. The parameters of the model are β, σ_v^2 and σ_u^2. We follow the usual convention that σ_u^2 is the variance of the normal random variable of which u is the absolute value, so that $\text{var}(u) = \frac{\pi - 2}{\pi}\sigma_u^2$. We also adopt the standard notation $\lambda = \sigma_u / \sigma_v$ and $\sigma^2 = \sigma_u^2 + \sigma_v^2$.

We will refer to $\varepsilon = v - u$ as the "composed error". Its density is known (e.g. Aigner, Lovell and Schmidt (1977, p. 26)) but its cdf does not have any known closed-form expression. We need the cdf, or a tabulation of it, to calculate the cell probabilities or the cell boundaries. We have therefore tabulated the cdf via a simulation, with 10,000,000 replications for each simulation. In Table 3 we present the quantiles 0.1, 0.2, ..., 0.9 for values of λ between zero and 10,000, and for $\sigma^2 = 1$. For a given value of λ and for a different value of σ^2, one just needs to multiply the quantile by σ. (For example, for $\lambda = 1$ and $\sigma^2 = 1$, the 0.40 quantile is -0.754. So for $\lambda = 1$ and $\sigma^2 = 2$, the 0.40 quantile is (-0.754)($\sqrt{2}$) = -1.066.)

2.6.1 Size of the Test

Primarily to check these tabulations, we first briefly consider the case in which the composed error ε is observed (equivalently, β is known or specified by the null hypothesis), and the null hypothesis specifies the value of σ_v^2 and σ_u^2. Thus we are testing a simple null hypothesis under which the distribution of ε is completely specified. We note that, apart from the randomness of the simulation, the results should be exactly the same as the results in Table 2.1, where we were testing the null that the data are standard normal. (If the null completely and correctly specifies the distribution of the data, the test is the same as if we compared the percentiles corresponding to the observations to the uniform distribution, and the nature of the parent distribution is irrelevant.) We also give results for the composed error with $\sigma_u^2 = \sigma_v^2 = 1$ in (Supplemental) Table 2.16. This table corresponds to Table 2.1 of this paper for the standard normal case.

Now we turn to the case of main interest, in which we wish to test that the composed error has the normal / half-normal distribution with unspecified (unknown) values of the parameters. In this case our model for the simulations will be:

(2.16) $\qquad y_i = \alpha + \varepsilon_i \quad , \quad \varepsilon_i = v_i - u_i \quad ,$

and the unknown parameters are α, σ_u^2 and σ_v^2. There are no regressors other than intercept, and no slope coefficients, because estimation error in these is not likely to be important. However, the presence of intercept is important. It is obviously empirically relevant, and it affects the results because, by demeaning the data, it prevents the level of the series from containing information about σ_u^2. In performing our experiments, we parameterize in terms of $\alpha, \lambda = \sigma_u / \sigma_v$ and σ_v^2. The results depend only on λ (for a given value of λ, the chi-squared statistics are not affected by changes in α or σ_v^2) but in estimation we estimate three parameters.

A technical complication worth mentioning is the "wrong skew" problem pointed out by Waldman (1982). The distribution of ε has a negative skew (negative third central moment). However, in the data we may encounter a positive skew of the residuals. In our case, this would correspond to $m_3 > 0$ where $m_3 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^3$, an occurrence of the "wrong skew." When we have the wrong skew, the MLE's are as follows:

$$(2.17) \qquad \hat{\alpha} = \bar{y}, \hat{\sigma}_u^2 = 0, \hat{\sigma}_v^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2 .$$

This happens with a positive probability that depends on λ and n. For example, when λ is near zero and n is small, the wrong skew problem occurs nearly half of the time. It is widely argued (e.g. Simar and Wilson (2009)) that the wrong skew problem causes considerable difficulties in inference in stochastic frontier models. One of the points of our experiments will be to see whether this is true for goodness of fit testing.

Our results for the cases where the null is true are given in Tables 2.4-2.8, which correspond to $\lambda = 0.1, 0.5, 1, 2$ and 10, respectively. These tables have essentially the same format as Table 2.2 (minus its last two columns), except that they also report the frequency of occurrence of the wrong skew problem.

One striking result in these tables is that the frequency of rejection (size of the test) does not depend very strongly on λ. That is, for a given value of n and for a given test, the size of the test is approximately the same in all five tables. In fact, the results in these tables are very similar to the results in Table 2.2, which was for the case of testing normality with unknown mean and variance. The parameter estimation problem is much simpler in the normal case than in the normal /·half-normal case, so we might expect larger size distortions in Tables 2.4-2.8 than in Table 2.2. However, we don't actually find that; any differences are very slight. Correspondingly, the main conclusions are the same as in Section 2.4. All of the tests that use bootstrapped critical values are quite accurate (size close to nominal size). The Tauchen version of the Pearson test, which relies on asymptotic theory instead of the bootstrap, is less reliable. There are noticeable size distortions unless the sample size is very large or the number of cells used is small. Based on these results we would recommend using critical values from the bootstrap. The choice of which test to use logically would depend on considerations of power, which we will discuss in the next subsection.

The frequencies of occurrence of the wrong skew problem are in line with previous evaluations, such as in Simar and Wilson (2009). The fact that the frequency of occurrence of the wrong skew problem varies strongly with λ, but the size of the test does not, would seem to imply that any size distortions we encounter are not primarily a reflection of this problem. As a matter of curiosity, we also calculated the frequency of rejection for those samples where the wrong skew problem did and did not occur. We did this for the Tauchen version of the Pearson test only, since that was the only test with significant size distortions. These results are given in (Supplemental) Table 2.17. The frequencies of rejection are different but not too different for the samples with the wrong skew than they are for the samples with the correct skew. For example, with $n = 50$ and $\lambda = 1$, we have 6399 replications with the correct (negative) skew and 3601 with the wrong (positive) skew. For the Tauchen χ^2 test with $k = 3$, we have rejection frequencies of 0.094 conditional on correct skew and 0.055 conditional on wrong skew; for $k = 5$ we have 0.104 and 0.074. These numbers are clearly different, but it is not the case that the rejections are coming overwhelmingly from one case or the other.

2.6.2 Power of the Test

Now we turn to the question of the power of the test. This requires specification of the alternative hypothesis. The null is exactly as in the previous section: the model is as given in equation (2.16), and the null is that the composed error $\varepsilon = v - u$ has the distribution implied by v being normal and u being half-normal. The alternatives that we consider will be based on the same model, except that u will follow some other one-sided distribution. Specifically, we will consider exponential and gamma distributions for u.

For the simulations in this subsection, we still use 10,000 replications for the Tauchen version of the Pearson test, and 1000 replications with 999 bootstrap samples for the tests based on the bootstrap, except that for the bootstrapped KS test, we use 1000 replications with 399 bootstrap samples.

Tables 2.9, 2.10 and 2.11 give the power of the test when v is N(0,1) and u is exponential with mean equal to θ (and, correspondingly, variance equal to θ^2). We consider $\theta = 0.1, 0.5, 1, 2, 5$ and 10. Varying θ changes the relative importance of noise and one-sided error. Since the results of the tests are invariant to linear transformation of the data, we could equally have kept θ fixed and changed the variance of v. (For example, the results with $\theta = 5$ and $\sigma_v^2 = 1$ are the same as with $\theta = 1$ and $\sigma_v^2 = 1/25$.) Larger values of θ correspond to less noise relative to one-sided error, and presumably should lead to higher power, since it is easier to distinguish half-normal and exponential data if they are contaminated with less noise. As a result, as we move down in each section of these tables, power should increase as θ increases. However, it is not the case that the power goes to one as $\theta \to \infty$. Rather, as $\theta \to \infty$, power should approach the power that we would have if there were no noise and we were testing the null that the data are half-normal against the alternative that they are exponential.

The KS test using bootstrapped critical values is generally the most powerful test. It clearly dominates the other two tests that use bootstrapped critical values. Its comparison to the non-bootstrapped Tauchen version of the Pearson test is somewhat ambiguous, because the Tauchen test sometimes appears to be more powerful, but this occurs in cases (small n and/or large k) in which the Tauchen test had non-trivial size distortions. Even in those cases the bootstrapped KS test is more powerful if θ is large enough that power is non-trivial. Basically whenever power is over 0.2, the bootstrapped KS test is best.

Comparing results for the various Pearson tests across the three tables, we see that power is generally higher when less cells are used. That is, power is higher with three cells than with five, and higher with five cells than with ten. (There are a few exceptions for the non-bootstrapped Tauchen test when power is low and n and k are such that size distortions were found under the null.) Since size distortions are smaller and power is higher when a small number of cells is used, it is obvious to recommend using a small number of cells. Precisely how small is a question that could be investigated further.

Unfortunately, we can also see that power is rather low unless the sample size is quite large and/or the variance of u is much larger than the variance of v. For example, when $\theta = 1$, which corresponds to equal variance for v and u, power for the bootstrapped KS test is only 0.054 for $n = 50$, 0.089 for $n = 100$, and 0.150 for $n = 250$. When θ is bigger the situation is more favorable. For example, when $\theta = 5$, power is 0.530 for $n = 100$ and 0.930 for $n = 250$. However, $\theta = 5$ corresponds to var(u) = 25var(v), which is generally not common in empirical applications.

Another way to summarize these results is that we can expect to distinguish exponential data from half-normal, but that this becomes difficult if the data are contaminated by normal noise.

In Tables 2.12-2.15 we consider the case that the one-sided error has a gamma distribution. Now $u = cu^*$ where u^* follows the standard gamma distribution with density

$$f(u^*) = \frac{(u^*)^{m-1} e^{-u^*}}{\Gamma(m)} \quad (2.18)$$

The parameter *m* governs the shape of u^*. When $m = 1$ we have the exponential distribution which we have just considered. Values of *m* less than one lead to densities with a mode at zero and very steep decline as u^* increases. Values of *m* larger than one lead to a positive mode, and the distribution approaches normality as $m \to \infty$. We consider $m = 0.1, 0.5, 2$ and 10. The mean and variance of the standard gamma distribution in (2.16) both equal *m*, so the mean of $u = cu^*$ equals *cm* and the variance equals $c^2 m$. Thus for a given value of *m*, we expect power to increase when *c* increases.

In Tables 2.12-2.15, the results for the Pearson tests are for $k = 3$ only.

The general pattern of results is similar to what was found for the exponential case. Power increases as *n* increases and as *c* increases. The Kolmogorov-Smirnov test is generally the most powerful. And, again, the power of the tests is low over the part of the parameter space that would seem to be empirically most common.

An interesting feature of these results is that the power is quite low for the values of *m* greater than one, even for large values of *c*. This is so despite the fact that the gamma distribution with *m* greater than one does not at all resemble the half-normal distribution. The reason for this low power is presumably that the gamma distribution with large *m* resembles the normal distribution, and therefore is mistaken for part of the noise.

2.7 CONCLUDING REMARKS

In this paper we have considered goodness of fit tests for the stochastic frontier model. We are interested in testing the distributional assumption for the one-sided error (inefficiency term). The essential difficulty is that we can only observe the composed error, which is the sum of the one-sided error and normal random noise. So in the end we test the hypothesis that the composed error has the distribution that is implied by normality of the noise and the assumed distribution for the one-sided error.

We considered Pearson χ^2 goodness of fit tests based on expected and actual numbers of observations in cells defined by values of the composed error, and also the Kolmogorov-Smirnov test. We discussed the asymptotic theory that corrects the Pearson test for the effects of parameter estimation. We also noted that asymptotically correct critical values can be found by boostrapping, for either the Pearson test or the Kolmogorov-Smirnov test.

We performed simulations to investigate the size and power properties of the tests. In terms of size, bootstrapping works better than asymptotic theory. In terms of power, the Kolmogorov-Smirnov test dominates the Pearson tests, so that the best test overall appears to be the Kolmogorov-Smirnov test using critical values from the bootstrap.

The remaining problem is that the power of these tests against plausible alternative distributions is disappointingly low. Reasonable power seems to require sample sizes and/or signal to noise ratios that are not commonly found empirical applications. Making the same point somewhat differently, it is easy to distinguish an exponential distribution from a half-normal. However, it is hard to distinguish the sum of a normal and an exponential from the sum of a normal and a half-normal, unless the variance of the normal component is very small or the sample size is very large.

Further research is needed to understand the empirical significance of these findings. Philosophically, it does not matter if different models yield different results if we can distinguish statistically between the models; conversely, it does not matter if we cannot distinguish statistically between models, if the models give more or less the same results. It is only a problem if we cannot distinguish statistically between models *and* the models give substantively different results. Intuitively, it seems reasonable to conjecture that data sets for which it is hard to distinguish between different distributions of inefficiency are also data sets for which the different distributions lead to similar empirical results. (Presumably these are cases in which different distributions for inefficiency lead to essentially the same distribution of the composed error ε.) Therefore the relationship between robustness of results and the power of goodness of fit tests (or, more generally, the ability of any model selection method to distinguish between different distributions of inefficiency) is obviously an important issue to investigate.

APPENDIX 2

Appendix 2-A

In this Appendix we establish equation (8) of the text. We write $\bar{g}(\theta_0) = P - \hat{P}$, where P is the (k-1)-dimensional vector with j^{th} element $p_j = p_j(\theta_0)$ and $\hat{P}$ is the (k-1)-dimensional vector with j^{th} element $\hat{p}_j = O_j / n$. Also we write $V(\theta_0) = \Pi - PP'$ where Π is the diagonal matrix with j^{th} diagonal element equal to p_j. Now we use the fact (e.g. Abadir and Magnus (2005), p. 87) that

$$[\Pi - PP']^{-1} = \Pi^{-1} + \frac{1}{1 - P'\Pi^{-1}P}\Pi^{-1}PP'\Pi^{-1} \tag{2.19}$$

Therefore

$$\begin{aligned} n\bar{g}(\theta_0)'V(\theta_0)^{-1}\bar{g}(\theta_0) &= n(\hat{P} - P)'\Pi^{-1}(\hat{P} - P) \\ &+ \frac{n}{1 - P'\Pi^{-1}P}(\hat{P} - P)'\Pi^{-1}PP'\Pi^{-1}(\hat{P} - P) \end{aligned} \tag{2.20}$$

The first term on the right hand side of (2.20) equals $n\sum_{j=1}^{k-1}(\hat{p}_j - p_j)^2 / p_j =$ $\sum_{j=1}^{k-1}(O_j - E_j)^2 / E_j$. For the second term, note that $1 - P'\Pi^{-1}P = 1 - \sum_{j=1}^{k-1} p_j = p_k$ and that $(\hat{P} - P)'\Pi^{-1}P = (\hat{P} - P)'e_{k-1}$ (where e_{k-1} is a vector of dimension (k-1) with each element equal to one) $= [(1 - \hat{p}_k) - (1 - p_k)] = (p_k - \hat{p}_k)$. Therefore

$n\bar{g}(\theta_0)'V(\theta_0)^{-1}\bar{g}(\theta_0) = \sum_{j=1}^{k-1}(O_j - E_j)^2 / E_j + n(p_k - \hat{p}_k)^2 / p_k =$

$\sum_{j=1}^{k}(O_j - E_j)^2 / E_j$.

Appendix 2-B

In this Appendix we discuss the goodness of fit test based on quantiles and its relationship to the Pearson test based on actual and expected cell counts. Suppose that we pick (k-1) probabilities $0 < p_1 < p_2 \cdots < p_{k-1} < 1$. Let the corresponding population quantiles be $m_1(\theta) < m_2(\theta) \cdots < m_{k-1}(\theta)$, so that $P(y \leq m_j(\theta)) = p_j$, and let the sample quantiles be $\hat{m}_1 \leq \hat{m}_2 \cdots \leq \hat{m}_{k-1}$. So now the test will depend on ($\hat{m} - m$), the vector whose j^{th} element equals ($\hat{m}_j - m_j(\theta)$), and the test statistic equals $n(\hat{m} - m(\hat{\theta}))'W(\hat{m} - m(\hat{\theta}))$ with an appropriate choice of W.

To see how this compares to the CMT test, we note that $\sqrt{n}(\hat{m}_j - m_j(\theta))$ is asymptotically normal, and so it must be expressable as an average (plus an asymptotically negligible term). This is the "influence function representation," which is given by:

(2.21) $$\sqrt{n}(\hat{m}_j - m_j(\theta)) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} r_{ij}(\theta) + o_p(1),$$

where $o_p(1)$ is an asymptotically negligible term (i.e., it converges in probability to zero), and where

(2.22) $$r_{ij}(\theta) = \frac{1}{f(m_j(\theta))}[p_j - 1(y_i \leq m_j(\theta))],$$

where f is the pdf of y. See, for example, Ruppert and Carroll (1980), p. 832. Therefore the test based on ($\hat{m} - m$) is equivalent in large samples to the CMT test based on the moment conditions $E[1(y \leq m_j(\theta)) - p_j]$, $j = 1, 2, \ldots, k$-1. This is an overlapping set of cells. However, it is also equivalent to consider the non-overlapping cells:

$A_1 = \{y | y \le m_1(\theta)\}$, $A_2 = \{y | m_1(\theta) < y \le m_2(\theta)\}$, etc. The resulting test is the CMT test based on observed versus actual cell counts, as discussed in the text.

Appendix 2-C

In this Appendix we derive analytically the variance matrix C used in the conditional moment test, for the case of a normal distribution. We wish to evaluate

(2.23) $$C_{11} = E(ss') \;,\; C_{12} = E(sg') \;,\; C_{22} = E(gg').$$

Here $s = s(y,\theta)$ is the score function for the normal distribution, given by

(2.24) $$s(y,\theta) = \begin{bmatrix} \frac{1}{\sigma^2}(y-\mu) \\ \frac{-1}{2\sigma^2} + \frac{1}{2\sigma^4}(y-\mu)^2 \end{bmatrix}$$

and $g = g(y,\theta)$ is the vector whose j^{th} element equals $[1(y \in A_j) - p_j]$.

It is well known that C_{11} is the information matrix for the normal distribution, given by

(2.25) $$\begin{bmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{bmatrix}.$$

Also C_{22} equals the matrix $V(\theta)$ as defined in the discussion following equation (2.4) of the text.

This leaves the submatrix C_{12}. It is of dimension 2 by (k-1). We will evaluate in turn the $(1, j)$ and $(2, j)$ elements of this matrix. To do so we make the reasonable assumption that the cells are intervals, so that $A_j = (a,b]$, where for notational simplicity we do not express the subscript "j" that should appear on a and b.

Then element $(1, j)$ of C_{12} equals

$$\frac{1}{\sigma^2} E(y-\mu)[1(y \in A_j) - p_j] = \frac{1}{\sigma^2} Ey[1(y \in A_j) - p_j]$$

$$= \frac{1}{\sigma^2} Ey1(y \in A_j) - \frac{1}{\sigma^2} p_j \mu$$

$$= \frac{p_j}{\sigma^2}[E(y|a < y \le b) - \mu]$$

$$= \frac{1}{\sigma^2}\left[\varphi\left(\frac{a-\mu}{\sigma}\right) - \varphi\left(\frac{b-\mu}{\sigma}\right)\right],$$

where "ϕ" is the standard normal density function. Here we have evaluated the conditional expectation $E(y|a < y \le b) = \mu + \frac{1}{p_j}\left[\varphi\left(\frac{a-\mu}{\sigma}\right) - \varphi\left(\frac{b-\mu}{\sigma}\right)\right]$ as in Johnson and Kotz (1970), equation (79), p. 81.

Similarly element $(2, j)$ of C_{12} equals

$$E\left[\frac{-1}{2\sigma^2} + \frac{1}{2\sigma^4}(y-\mu)^2\right][1(y \in A_j) - p_j]$$

$$= E\left[\frac{1}{2\sigma^4}(y-\mu)^2\right][1(y \in A_j) - p_j]$$

$$= \frac{1}{2\sigma^4} E(y-\mu)^2 1(a < y \le b) - \frac{p_j}{2\sigma^4}$$

$$= \frac{1}{2\sigma^4} Ey^2 1(y \in A_j) + \frac{1}{2\sigma^4}(-2\mu)Ey1(y \in A_j) + \frac{1}{2\sigma^4}\mu^2 p_j - \frac{p_j}{2\sigma^4}$$

$$= \frac{1}{2\sigma^4} Ey^2 1(y \in A_j) - \frac{\mu}{\sigma^4} p_j E(y|y \in A_j) + \frac{p_j \mu^2}{2\sigma^4} - \frac{p_j}{2\sigma^4},$$

where $Ey^2 1(y \in A_j) = p_j \operatorname{var}(y | y \in A_j) + p_j \left[E(y | y \in A_j) \right]^2$.

Furthermore,

$$Ey^2 1(y \in A_j) = p_j \sigma^2 \left\{ 1 - \frac{b\phi(b) - a\phi(a)}{\Phi(b) - \Phi(a)} - \left[\frac{\phi(b) - \phi(a)}{\Phi(b) - \Phi(a)} \right]^2 \right\} + p_j \left[\mu - \sigma \frac{\phi(b) - \phi(a)}{\Phi(b) - \Phi(a)} \right]^2 .$$

Table 2.1

Size of the test of the hypothesis that the data are N(0,1)

Nominal size = 0.05

k	*n*	Pearson	Bootstrap Pearson	KS	Bootstrap KS
3	50	0.049	0.053	0.040	0.045
	100	0.054	0.047	0.040	0.048
	250	0.053	0.039	0.046	0.048
	500	0.047	0.045	0.050	0.056
5	50	0.043	0.050	*	*
	100	0.049	0.044	*	*
	250	0.049	0.041	*	*
	500	0.051	0.058	*	*
10	50	0.048	0.052	*	*
	100	0.049	0.043	*	*
	250	0.052	0.056	*	*
	500	0.051	0.058	*	*

* The number of cells (k) is not relevant for the Kolmogorov-Smirnov test.

Table 2.2

Size of the test of the hypothesis that the data are normal

Nominal size = 0.05

k	n	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS	Simple Hypothesis OPG	Composite Hypothesis C Known
3	50	0.067	0.056	0.051	0.057	0.050	0.049
	100	0.055	0.052	0.053	0.054	0.050	0.050
	250	0.052	0.050	0.049	0.047	0.054	0.045
	500	0.053	0.058	0.053	0.048	0.048	0.051
5	50	0.086	0.055	0.044	*	0.067	0.049
	100	0.070	0.042	0.055	*	0.059	0.050
	250	0.057	0.043	0.060	*	0.053	0.050
	500	0.061	0.054	0.062	*	0.053	0.048
10	50	0.185	0.054	0.060	*	0.139	0.044
	100	0.117	0.053	0.050	*	0.101	0.049
	250	0.079	0.052	0.042	*	0.071	0.051
	500	0.060	0.044	0.058	*	0.059	0.047

* The number of cells (k) is not relevant for the Kolmogorov-Smirnov test.

Table 2.3

Quantiles of the distribution of the normal / half normal composed error

$$\sigma^2 = \sigma_u^2 + \sigma_v^2 = 1\text{, various } \lambda$$

λ	Quantile								
	.10	.20	.30	.40	.50	.60	.70	.80	.90
0.0	-1.281	-0.841	-0.524	-0.253	0.000	0.253	0.524	0.841	1.281
0.1	-1.357	-0.918	-0.602	-0.332	-0.080	0.173	0.444	0.759	1.198
0.2	-1.423	-0.987	-0.675	-0.407	-0.156	0.094	0.362	0.675	1.109
0.3	-1.477	-1.048	-0.739	-0.475	-0.228	0.018	0.282	0.590	1.017
0.4	-1.522	-1.099	-0.796	-0.537	-0.294	-0.053	0.206	0.508	0.926
0.5	-1.556	-1.141	-0.843	-0.589	-0.353	-0.117	0.135	0.430	0.837
0.6	-1.582	-1.176	-0.884	-0.635	-0.405	-0.174	0.071	0.358	0.754
0.7	-1.602	-1.202	-0.917	-0.674	-0.449	-0.224	0.014	0.292	0.675
0.8	-1.616	-1.223	-0.943	-0.706	-0.486	-0.269	-0.037	0.233	0.603
0.9	-1.626	-1.238	-0.964	-0.733	-0.518	-0.306	-0.081	0.180	0.537
1.0	-1.632	-1.250	-0.981	-0.754	-0.545	-0.339	-0.120	0.133	0.478
1.1	-1.637	-1.259	-0.994	-0.772	-0.567	-0.366	-0.153	0.091	0.425
1.2	-1.640	-1.266	-1.004	-0.786	-0.586	-0.389	-0.183	0.054	0.377
1.3	-1.642	-1.271	-1.012	-0.797	-0.601	-0.409	-0.209	0.022	0.334
1.4	-1.643	-1.274	-1.018	-0.806	-0.614	-0.426	-0.230	-0.006	0.295
1.5	-1.644	-1.276	-1.023	-0.814	-0.625	-0.440	-0.249	-0.032	0.261
1.6	-1.644	-1.278	-1.026	-0.819	-0.633	-0.453	-0.266	-0.054	0.230
1.7	-1.644	-1.279	-1.029	-0.824	-0.640	-0.463	-0.280	-0.074	0.202
1.8	-1.645	-1.280	-1.031	-0.828	-0.646	-0.472	-0.293	-0.091	0.177
1.9	-1.645	-1.280	-1.033	-0.831	-0.651	-0.480	-0.304	-0.107	0.154
2.0	-1.645	-1.281	-1.034	-0.833	-0.656	-0.486	-0.313	-0.120	0.134
2.1	-1.645	-1.281	-1.034	-0.835	-0.659	-0.491	-0.322	-0.132	0.115
2.2	-1.645	-1.281	-1.035	-0.837	-0.662	-0.496	-0.329	-0.144	0.098
2.3	-1.645	-1.282	-1.035	-0.838	-0.664	-0.500	-0.335	-0.154	0.083
2.4	-1.645	-1.282	-1.036	-0.838	-0.666	-0.504	-0.341	-0.162	0.069
2.5	-1.645	-1.282	-1.036	-0.839	-0.667	-0.507	-0.346	-0.170	0.056
2.6	-1.645	-1.282	-1.036	-0.840	-0.669	-0.509	-0.350	-0.177	0.044
2.7	-1.645	-1.282	-1.036	-0.840	-0.670	-0.512	-0.354	-0.184	0.034
2.8	-1.645	-1.282	-1.036	-0.841	-0.671	-0.513	-0.358	-0.190	0.024
2.9	-1.645	-1.282	-1.036	-0.841	-0.672	-0.515	-0.361	-0.195	0.014
3.0	-1.645	-1.282	-1.036	-0.841	-0.672	-0.516	-0.364	-0.200	0.006

Table 2.3 (cont'd.)

λ	Quantile								
	.10	.20	.30	.40	.50	.60	.70	.80	.90
3.1	-1.645	-1.282	-1.036	-0.841	-0.673	-0.518	-0.366	-0.204	-0.002
3.2	-1.645	-1.282	-1.036	-0.841	-0.673	-0.519	-0.368	-0.208	-0.009
3.3	-1.645	-1.282	-1.036	-0.841	-0.673	-0.519	-0.370	-0.212	-0.016
3.4	-1.645	-1.282	-1.036	-0.841	-0.674	-0.520	-0.372	-0.215	-0.022
3.5	-1.645	-1.282	-1.036	-0.841	-0.674	-0.521	-0.373	-0.218	-0.027
3.6	-1.645	-1.282	-1.036	-0.841	-0.674	-0.521	-0.374	-0.221	-0.033
3.7	-1.645	-1.282	-1.036	-0.841	-0.674	-0.522	-0.376	-0.224	-0.038
3.8	-1.645	-1.282	-1.036	-0.841	-0.674	-0.522	-0.377	-0.226	-0.043
3.9	-1.645	-1.282	-1.036	-0.841	-0.674	-0.523	-0.378	-0.228	-0.047
4.0	-1.645	-1.282	-1.036	-0.842	-0.674	-0.523	-0.379	-0.230	-0.051
4.1	-1.645	-1.282	-1.036	-0.842	-0.674	-0.523	-0.379	-0.232	-0.055
4.2	-1.645	-1.282	-1.036	-0.842	-0.674	-0.523	-0.380	-0.233	-0.059
4.3	-1.645	-1.282	-1.036	-0.842	-0.674	-0.523	-0.381	-0.235	-0.062
4.4	-1.645	-1.282	-1.036	-0.842	-0.674	-0.524	-0.381	-0.236	-0.065
4.5	-1.645	-1.282	-1.036	-0.842	-0.674	-0.524	-0.382	-0.238	-0.068
4.6	-1.645	-1.282	-1.036	-0.842	-0.674	-0.524	-0.382	-0.239	-0.071
4.7	-1.645	-1.282	-1.036	-0.842	-0.674	-0.524	-0.383	-0.240	-0.073
4.8	-1.645	-1.282	-1.036	-0.842	-0.674	-0.524	-0.383	-0.241	-0.076
4.9	-1.645	-1.282	-1.036	-0.842	-0.674	-0.524	-0.383	-0.242	-0.078
5.0	-1.645	-1.282	-1.036	-0.842	-0.674	-0.524	-0.383	-0.243	-0.081
5.1	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.383	-0.244	-0.083
5.2	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.384	-0.244	-0.085
5.3	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.384	-0.245	-0.086
5.4	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.384	-0.246	-0.088
5.5	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.384	-0.246	-0.090
5.6	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.384	-0.247	-0.092
5.7	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.384	-0.247	-0.093
5.8	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.248	-0.094
5.9	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.248	-0.096
6.0	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.248	-0.097
6.1	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.249	-0.098
6.2	-1.645	-1.282	-1.036	-0.842	-0.675	-0.525	-0.385	-0.249	-0.100
6.3	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.250	-0.101
6.4	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.250	-0.102
6.5	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.250	-0.103
6.6	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.250	-0.104
6.7	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.251	-0.105
6.8	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.251	-0.105
6.9	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.251	-0.106
7.0	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.251	-0.107

Table 2.3 (cont'd.)

λ	Quantile								
	.10	.20	.30	.40	.50	.60	.70	.80	.90
7.1	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.251	-0.108
7.2	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.251	-0.109
7.3	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.109
7.4	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.110
7.5	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.111
7.6	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.111
7.7	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.112
7.8	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.112
7.9	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.113
8.0	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.113
8.1	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.252	-0.114
8.2	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.114
8.3	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.115
8.4	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.115
8.5	-1.645	-1.282	-1.036	-0.842	-0.675	-0.525	-0.385	-0.253	-0.116
8.6	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.116
8.7	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.116
8.8	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.117
8.9	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.117
9.0	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.117
9.1	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.118
9.2	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.118
9.3	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.118
9.4	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.119
9.5	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.119
9.6	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.119
9.7	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.119
9.8	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.120
9.9	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.120
10	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.120
11	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.122
12	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.123
13	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.124
14	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.124
15	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.125
20	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.125
50	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.125
100	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.125
1000	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.126
10000	-1.645	-1.282	-1.036	-0.842	-0.675	-0.524	-0.385	-0.253	-0.126

Table 2.4

Size of test of the hypothesis that the data are normal / half-normal

$\lambda = 0.1$

Nominal size = 0.05

k	*n*	Wrong Skew (%)	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
3	50	50.2	0.077	0.051	0.057	0.048
	100	50.9	0.062	0.053	0.055	0.062
	250	50.2	0.056	0.049	0.040	0.041
5	50	50.2	0.092	0.043	0.043	*
	100	50.9	0.068	0.046	0.046	*
	250	50.2	0.060	0.049	0.049	*
10	50	50.2	0.200	0.045	0.047	*
	100	50.9	0.119	0.049	0.054	*
	250	50.2	0.076	0.051	0.044	*

*The number of cells (k) is not relevant for the Kolmogorov-Smirnov test.

Table 2.5

Size of the test of the hypothesis that the data are normal / half-normal

$\lambda = 0.5$

Nominal size = 0.05

k	*n*	Wrong Skew (%)	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
3	50	47.8	0.079	0.048	0.031	0.047
	100	47.4	0.055	0.037	0.041	0.050
	250	44.3	0.051	0.050	0.046	0.045
5	50	47.8	0.089	0.044	0.044	*
	100	47.4	0.065	0.056	0.055	*
	250	44.3	0.055	0.044	0.057	*
10	50	47.8	0.191	0.053	0.053	*
	100	47.4	0.113	0.056	0.055	*
	250	44.3	0.073	0.055	0.049	*

* The number of cells (k) is not relevant for the Kolmogorov-Smirnov test.

Table 2.6

Size of the test of the hypothesis that the data are normal / half-normal

$\lambda = 1$

Nominal size = 0.05

k	*n*	Wrong Skew (%)	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
3	50	36.0	0.080	0.051	0.041	0.042
	100	30.8	0.059	0.053	0.037	0.041
	250	19.5	0.049	0.048	0.053	0.040
5	50	36.0	0.093	0.042	0.039	*
	100	30.8	0.065	0.056	0.042	*
	250	19.5	0.053	0.043	0.043	*
10	50	36.0	0.190	0.039	0.050	*
	100	30.8	0.111	0.041	0.044	*
	250	19.5	0.068	0.061	0.055	*

* The number of cells (k) is not relevant for the Kolmogorov-Smirnov test.

Table 2.7

Size of the test of the hypothesis that the data are normal / half-normal

$\lambda = 2$

Nominal size = 0.05

k	*n*	Wrong Skew (%)	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
3	50	11.5	0.074	0.053	0.034	0.037
	100	4.1	0.067	0.046	0.039	0.042
	250	0.2	0.050	0.056	0.044	0.045
5	50	11.5	0.107	0.043	0.042	*
	100	4.1	0.072	0.064	0.048	*
	250	0.2	0.055	0.060	0.053	*
10	50	11.5	0.233	0.040	0.043	*
	100	4.1	0.122	0.048	0.046	*
	250	0.2	0.069	0.058	0.052	*

* The number of cells (k) is not relevant for the Kolmogorov-Smirnov test.

Table 2.8

Size of the test of the hypothesis that the data are normal / half-normal

$\lambda = 10$

Nominal size = 0.05

k	*n*	Wrong Skew (%)	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
3	50	0.1	0.060	0.049	0.050	0.044
	100	0	0.053	0.057	0.051	0.038
	250	0	0.051	0.048	0.048	0.045
5	50	0.1	0.090	0.041	0.051	*
	100	0	0.064	0.049	0.048	*
	250	0	0.059	0.053	0.048	*
10	50	0.1	0.238	0.038	0.045	*
	100	0	0.116	0.054	0.057	*
	250	0	0.073	0.052	0.043	*

* The number of cells (k) is not relevant for the Kolmogorov-Smirnov test.

Table 2.9

Power of the test of the hypothesis that the data are normal / half-normal

Alternative: the data are normal / exponential(θ)

Nominal size = 0.05

Number of cells: $k = 3$

n	θ	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
50	0.1	0.069	0.059	0.059	0.046
	0.5	0.071	0.055	0.053	0.041
	1	0.079	0.056	0.059	0.054
	2	0.142	0.110	0.098	0.147
	5	0.269	0.217	0.195	0.340
	10	0.367	0.285	0.294	0.494
100	0.1	0.064	0.049	0.054	0.056
	0.5	0.060	0.055	0.050	0.048
	1	0.080	0.084	0.069	0.089
	2	0.196	0.188	0.207	0.239
	5	0.468	0.440	0.384	0.530
	10	0.633	0.563	0.537	0.700
250	0.1	0.064	0.049	0.051	0.057
	0.5	0.052	0.051	0.041	0.048
	1	0.101	0.084	0.108	0.150
	2	0.416	0.407	0.476	0.507
	5	0.879	0.842	0.789	0.930
	10	0.959	0.945	0.922	0.966

Table 2.10

Power of the test of the hypothesis that the data are normal / half-normal

Alternative: the data are normal / exponential(θ)

Nominal size = 0.05

Number of cells: $k = 5$

n	θ	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
50	0.1	0.086	0.058	0.052	0.046
	0.5	0.087	0.042	0.057	0.041
	1	0.097	0.065	0.064	0.054
	2	0.150	0.097	0.079	0.147
	5	0.265	0.137	0.123	0.340
	10	0.343	0.167	0.185	0.494
100	0.1	0.062	0.054	0.055	0.056
	0.5	0.087	0.045	0.044	0.048
	1	0.107	0.076	0.071	0.089
	2	0.202	0.167	0.135	0.239
	5	0.399	0.354	0.282	0.530
	10	0.537	0.435	0.383	0.700
250	0.1	0.061	0.056	0.051	0.057
	0.5	0.059	0.061	0.056	0.048
	1	0.110	0.101	0.091	0.150
	2	0.383	0.382	0.302	0.507
	5	0.851	0.799	0.695	0.930
	10	0.944	0.917	0.833	0.966

Table 2.11

Power of the test of the hypothesis that the data are normal / half-normal

Alternative: the data are normal / exponential(θ)

Nominal size = 0.05

Number of cells: k = 10

n	θ	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
50	0.1	0.191	0.044	0.047	0.046
	0.5	0.167	0.049	0.045	0.041
	1	0.203	0.050	0.037	0.054
	2	0.245	0.065	0.063	0.147
	5	0.370	0.079	0.102	0.340
	10	0.485	0.093	0.110	0.494
100	0.1	0.119	0.050	0.042	0.056
	0.5	0.104	0.049	0.047	0.048
	1	0.140	0.061	0.061	0.089
	2	0.225	0.108	0.101	0.239
	5	0.364	0.209	0.210	0.530
	10	0.499	0.295	0.260	0.700
250	0.1	0.082	0.052	0.045	0.057
	0.5	0.077	0.045	0.052	0.048
	1	0.124	0.075	0.059	0.150
	2	0.338	0.267	0.217	0.507
	5	0.716	0.671	0.494	0.930
	10	0.879	0.815	0.789	0.966

Table 2.12

Power of the test of the hypothesis that the data are normal / half-normal

Alternative: the data are normal / gamma (u is c times gamma(m))

Nominal size = 5%

$m = 0.1$

n	c	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
50	0.1	0.069	0.051	0.046	0.042
	0.5	0.060	0.061	0.061	0.051
	1	0.073	0.065	0.066	0.058
	2	0.081	0.135	0.120	0.117
	5	0.407	0.365	0.403	0.467
	10	0.785	0.695	0.730	0.814
100	0.1	0.059	0.055	0.046	0.055
	0.5	0.061	0.052	0.049	0.068
	1	0.075	0.046	0.046	0.056
	2	0.099	0.095	0.127	0.177
	5	0.614	0.602	0.607	0.978
	10	0.962	0.960	0.965	0.984
250	0.1	0.062	0.058	0.060	0.060
	0.5	0.052	0.055	0.051	0.052
	1	0.063	0.060	0.062	0.059
	2	0129	0.134	0.175	0.308
	5	0.888	0.921	0.954	1.000
	10	1.000	1.000	1.000	1.000

Table 2.13

Power of the test of the hypothesis that the data are normal / half-normal

Alternative: the data are normal / gamma (u is c times gamma(m))

Nominal size = 5%

$m = 0.5$

n	c	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
50	0.1	0.068	0.054	0.049	0.061
	0.5	0.049	0.046	0.046	0.046
	1	0.073	0.066	0.066	0.056
	2	0.153	0.132	0.132	0.154
	5	0.511	0.416	0.416	0.621
	10	0.791	0.740	0.740	0.887
100	0.1	0.074	0.061	0.061	0.053
	0.5	0.055	0.061	0.061	0.050
	1	0.061	0.078	0.078	0.078
	2	0.260	0.271	0.271	0.328
	5	0.784	0.732	0.732	0.869
	10	0.974	0.948	0.948	0.945
250	0.1	0.067	0.053	0.053	0.060
	0.5	0.061	0.069	0.069	0.067
	1	0.080	0.082	0.082	0.120
	2	0.516	0.583	0.583	0.685
	5	0.995	0.973	0.973	1.000
	10	1.000	1.000	1.000	1.000

Table 2.14

Power of the test of the hypothesis that the data are normal / half-normal

Alternative: the data are normal / gamma (u is c times gamma(m))

Nominal size = 5%

$m = 2$

n	c	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
50	0.1	0.058	0.054	0.061	0.043
	0.5	0.050	0.038	0.043	0.029
	1	0.073	0.057	0.053	0.050
	2	0.072	0.075	0.067	0.049
	5	0.095	0.048	0.064	0.073
	10	0.110	0.076	0.070	0.082
100	0.1	0.071	0.053	0.050	0.060
	0.5	0.044	0.063	0.049	0.046
	1	0.066	0.062	0.053	0.068
	2	0.091	0.075	0.092	0.107
	5	0.106	0.096	0.094	0.127
	10	0.125	0.110	0.107	0.132
250	0.1	0.040	0.061	0.057	0.057
	0.5	0.049	0.052	0.054	0.063
	1	0.080	0.088	0.079	0.100
	2	0.164	0.141	0.158	0.198
	5	0.210	0.196	0.188	0.220
	10	0.219	0.210	0.217	0.233

Table 2.15

Power of the test of the hypothesis that the data are normal / half-normal

Alternative: the data are normal / gamma (u is c times gamma(m))

Nominal size = 5%

$m = 10$

n	c	Pearson (Tauchen)	Bootstrap Pearson (Tauchen)	Bootstrap Pearson	Bootstrap KS
50	0.1	0.061	0.050	0.052	0.060
	0.5	0.062	0.054	0.043	0.055
	1	0.065	0.042	0.052	0.072
	2	0.065	0.054	0.044	0.068
	5	0.095	0.074	0.076	0.098
	10	0.110	0.098	0.099	0.115
100	0.1	0.062	0.051	0.049	0.055
	0.5	0.064	0.060	0.055	0.053
	1	0.065	0.048	0.052	0.068
	2	0.068	0.062	0.063	0.075
	5	0.093	0.070	0.065	0.099
	10	0.100	0.082	0.085	0.121
250	0.1	0.055	0.055	0.054	0.039
	0.5	0.059	0.063	0.059	0.055
	1	0.055	0.066	0.066	0.043
	2	0.071	0.069	0.072	0.058
	5	0.082	0.079	0.090	0.087
	10	0.104	0.095	0.088	0.119

(Supplemental) Table 2.16

Size of the test of the hypothesis that the data are normal / half normal

with $\sigma_u^2 = \sigma_v^2 = 1$ (simple hypothesis)

Nominal size = 0.05

K	*N*	Pearson	Bootstrap Pearson	KS	Bootstrap KS
3	50	0.050	0.044	0.042	0.048
	100	0.049	0.047	0.040	0.053
	250	0.050	0.051	0.047	0.048
5	50	0.043	0.045	*	*
	100	0.049	0.044	*	*
	250	0.047	0.050	*	*
10	50	0.045	0.054	*	*
	100	0.048	0.046	*	*
	250	0.040	0.052	*	*

* The number of cells (k) is not relevant for the Kolmogorov-Smirnov test.

(Supplemental) Table 2.17

Size of the test of the hypothesis that the data are normal / half normal

Results conditional on negative skew

Tauchen version of Pearson test

Nominal size = 0.05

k	*n*	λ = 0.1	λ = 0.5	λ = 1	λ = 2	λ = 10
3	50	0.089	0.092	0.094	0.124	0.060
	100	0.066	0.055	0.065	0.068	0.053
	250	0.058	0.051	0.052	0.050	0.052
	500	0.057	0.058	0.052	0.052	0.052
5	50	0.097	0.092	0.104	0.111	0.090
	100	0.068	0.065	0.067	0.072	0.064
	250	0.064	0.059	0.057	0.056	0.059
	500	0.057	0.060	0.055	0.053	0.054
10	50	0.209	0.207	0.206	0.241	0.238
	100	0.119	0.111	0.111	0.122	0.116
	250	0.080	0.078	0.079	0.070	0.073
	500	0.069	0.072	0.065	0.061	0.064

Essay 3

TESTING EQUALITY OF DISTRIBUTION FOR TWO CORRELATED VARIABLES

3.1 INTRODUCTION

In this paper, we consider the problem of testing equality of distribution for two variables. One is often confronted with the question whether two samples have the same distribution, for instance, whether a government policy program changes the earnings distribution, or whether students from two high schools in the same county have the same SAT score distribution, or whether the temperature distribution changes before and after the Kyoto Protocol.

Many nonparametric and parametric tests are available for testing equality of two distribution functions. Examples include the Kolmogorov-Smirnov (KS) test, the Cramer-von Mises (CM) test, and the Baumgartner-Weiss-Schindler (BWS) test of Baumgartner et al. (1998). These tests assume the two data series are independent. However, this assumption is restrictive. As an example, if we have a cross-section of cities, and for each we have an average temperature before and after the Kyoto Protocol, obviously these two observations are correlated for each city. In this paper we wish to allow pairwise correlation across the two data sets, to accommodate cases like this.

Li et al. (2009) is the only paper of which we are aware that allows this type of pairwise correlation. Their test is somewhat complicated because they consider the case that each observation is a vector random variable, and they allow mixed continuous and categorical variables. Our tests are much simpler, since we simply ask how we can modify the simple tests of the previous paragraph to allow for pairwise correlation.

In this paper we show how to implement a Pearson χ^2 test, based on differences of frequencies in non-overlapping intervals (cells) that span the support of the variables, in a GMM setting. This procedure makes no assumption about the correlation between the two variables. We also suggest a novel bootstrapping procedure that enables us to generate asymptotically valid critical values for the KS and BWS tests.

The plan of the paper is as follows. In section 3.2 we present the tests used to test equality of distribution. Section 3.3 gives simulation results to assess the size and power of the different relevant tests. Section 3.4 gives our concluding remarks.

3.2 TESTS OF EQUALITY OF TWO DISTRIBUTIONS

Suppose that we have a paired sample of size n: (y_i, x_i), $i = 1, 2, 3, \ldots, n$. We have independence over i, but for a given i, y_i and x_i may be correlated. We wish to test the null hypothesis that $y_1, y_2, \ldots, y_n$ and $x_1, x_2, \ldots, x_n$ are from the same distribution. That is, our interest is in testing the hypothesis that the dependent random variables Y and X have the same distribution.

3.2.1 Chi-squared Test

We follow Wang, Amsler and Schmidt (2009) in considering a Pearson χ^2 statistic using the GMM framework. They showed how to test whether a single sample is from a hypothesized distribution. Here we extend their methods to the case of testing for equality of distribution when the two samples are dependent.

Consider the GMM structure for the Pearson χ^2 statistic, as follows. Let the possible range of the random variables Y and X be split into k cells (intervals) $A_1 .., A_k$, such that any value of y_i and x_i, respectively, are in one and only one cell. Furthermore, let $1(y \in A_j)$ (or $1(x \in A_j)$) be the indicator function that equals one if y_i (or x_i) is in cell A_j, and equals zero otherwise. Under the null, the events $1(y \in A_j)$ and $1(x \in A_j)$ have the same probability, and thus we have the set of moment conditions:

$$E[g(y,x)] = 0 \qquad (3.1)$$

where $g(y,x)$ is a vector of dimension (k-1) whose j^{th} element equals $[1(y \in A_j) - 1(x \in A_j)]$. We have omitted one cell so as to avoid a subsequent singularity. We have omitted cell A_k but the choice of which cell to omit does not matter. If the dependent sample y and x are not from the same distribution, the observed frequencies of y would not be similar to the corresponding observed frequencies of x for some or perhaps all cells. As a result, the sample analog of equation (3.1) would not be close to zero.

Now define

$$\bar{g} = \frac{1}{n} \sum_{i=1}^{n} g(y_i, x_i) \qquad (3.2)$$

and note that the j^{th} element of $\bar{g}$ is equal to $\frac{1}{n}\left[\sum_{i=1}^{n} 1(y_i \in A_j) - \sum_{i=1}^{n} 1(x_i \in A_j)\right]$.

We also need to define the variance matrix of the moment conditions $g(y,x)$. If the two variables are independent, we can use the variance matrix of dimension (k-1) by (k-1), whose j^{th} diagonal element equals $2*(p_j - p_j^2)$, and whose i^{th}, j^{th} off diagonal element ($i \neq j$) equals $-2*p_i p_j$, with $p_i = p_j$ equal to the probability of each cell. But this does not apply here as it does not take into account the correlation structure that we allow.

However, we can use the outer product of the gradient (OPG) estimate of the variance matrix of the moment conditions, which is:

(3.3) $$\hat{V} = \frac{1}{n}\sum_{i=1}^{n} g(y_i, x_i) g(y_i, x_i)'.$$

This is a consistent estimate of V, the variance matrix of the moment conditions; that is, $V = Eg(y_i, x_i)g(y_i, x_i)'$.

A central limit theorem then implies that the asymptotic distribution of $\sqrt{n}\overline{g}$ is $N(0, V)$. From this fact it follows that

(3.4) $$n\overline{g}'\hat{V}^{-1}\overline{g} \rightarrow_d \chi_{k-1}^2.$$

To implement the test, for a given value of k, we use equiprobable cells, where "equiprobable" is defined using the sample quantiles of $x_1, x_2, ..., x_n$. We could alternatively use the sample quantiles of $y_1, y_2, ..., y_n$ or of the combined sample but we do not pursue these alternatives here. Tauchen (1985) showed that using sample quantiles to define the cells does not invalidate the distributional result (3.4).

3.2.2 Kolmogorov-Smirnov Test

The nonparametric KS two sample test is widely applied to test whether two independent samples are from the same distribution. The test is sensitive to any kind of distributional differences, for example, in the mean, variance, and kurtosis; however, the test assumes independent samples.

The KS test compares the vertical difference of the empirical cdf between the two random samples $y_1, y_2, \ldots, y_{n_1}$ and $x_1, x_2, \ldots, x_{n_2}$ and is defined by the largest value of this difference. If the two random samples are from the same distribution, then both empirical cdfs should be expected to be very similar to one another. The statistic is

$$KS = \sup \left| \frac{1}{n_1} \sum_i^{n_1} 1(y_i \le x) - \frac{1}{n_2} \sum_i^{n_2} 1(x_i \le x) \right| \quad . \tag{3.5}$$

If $KS \bullet \sqrt{n_1 \bullet n_2 / (n_1 + n_2)} > K_\alpha$ where K_α is the critical value[1] at the α significance level, then we reject the null hypothesis that two random variables have the same distribution at the α level. In our case, we assume paired data and so $n_1 = n_2 = n$.

3.2.3 The Baumgartner-Weiss-Schindler Test

Baumgartner et al. (1998) introduced a new nonparametric rank test---the Baumgartner-Weiss-Schindler (BWS) test. It is also used to test the null hypothesis that two independently samples have the same distribution. The test only requires that the underlying distribution function is continuous.

They suggested that the test should be based on the square of the difference between the empirical cdfs of two samples and be weighted by its variance. The statistic is $B = (n_1 \bullet n_2)/(n_1 + n_2) \bullet \int_0^1 (\hat{F}_{n_1}(z) - \hat{F}_{n_2}(z)) / (z \bullet (1-z)) dz$. To calculate this statistic, suppose that we have the two samples $y_1, y_2, \ldots, y_{n_1}$ and $x_1, x_2, \ldots, x_{n_2}$. Then we mix the

[1] Pearson et al. (1972) tabulated the critical points for finite samples.

two samples together and define H_j and G_i as the rank numbers of y_j and x_i in the pooled sample. We can calculate the BWS test statistic B, as follows:

$$\hat{B} = \frac{(\hat{B}_Y + \hat{B}_X)}{2}, \text{ where} \tag{3.6}$$

$$\hat{B}_Y = \frac{1}{n_1}\sum_{j=1}^{n_1} \frac{(H_j - \frac{n_1+n_2}{n_1} \bullet j)^2}{\frac{j}{n_1+1} \bullet (1-\frac{j}{n_1+1}) \bullet \frac{n_2(n_1+n_2)}{n_1}},$$

$$\hat{B}_X = \frac{1}{n_2}\sum_{i=1}^{n_2} \frac{(G_i - \frac{n_1+n_2}{n_2} \bullet i)^2}{\frac{i}{n_2+1} \bullet (1-\frac{i}{n_2+1}) \bullet \frac{n_1(n_1+n_2)}{n_2}}.$$

The critical values are tabulated in their paper.

Baumgartner et al. showed that the power of the test is at least as good as previously proposed nonparametric tests such as the KS, CM, and Wilcoxon tests. In addition, they showed that the BWS test tends to be a more powerful test for samples from a distribution with heavy tails because it takes into consideration the weighting variance, $z \bullet (1-z)$.

3.2.4 The Bootstrap Resampling Method

Although the regular KS and BWS tests are only valid for independent samples, we can use the percentile bootstrap approach to find the critical values of the KS and BWS tests for dependent samples. The bootstrap can be used to construct asymptotically valid tests provided the resampling method is correct.

A proper bootstrapping procedure to obtain critical values must ensure that (i) the bootstrap samples satisfy the null, and (ii) the correlation structure of the assumed DGP is maintained. Li et al. (2009) propose a bootstrapping procedure that satisfies (i) but not (ii). They merged the two data sets X and Y into one combined series Z, and then they drew the bootstrap samples of X (or Y) for the data set Z. So half of the bootstrap X observations are original X observations and half are original Y observations, and similarly for the bootstrap Y observations. Obviously, the bootstrap X and bootstrap Y series are from the same (Z) distribution, so they satisfy the null, but the pairwise correlation pattern has been destroyed.

Our bootstrap procedure is as follows. We construct a bootstrap sample of pairs (y_i, x_i) from the original sample. Because we draw pairs, the pairwise correlation structure is maintained. Then with probability 0.5, we accept the pair as it is, but with probability 0.5, we rename y_i as x_i and x_i as y_i. So now, as in Li et al. (2009), half of the bootstrap X observations are original X observations and half are original Y observations, and similarly for the bootstrap Y observations. So the null is true (the bootstrap X and bootstrap Y series are from the same distribution) and the correlation structure is maintained.

3.2.5 Pairwise T-Test

The pairwise t-test is used to test the null hypothesis that two samples $y_1, y_2, ..., y_n$ and $x_1, x_2, ..., x_n$ from a normal distribution have the same mean. It is simply a test of whether the differences $D_i = y_i - x_i$ have zero mean. First, we sum the differences across n pairs and then calculate the sample variance of the differences. Second, we use the classical t-test statistic

(3.7) $$t = 1/\sqrt{n}\sum_{i=1}^{n} D_i \Big/ \sqrt{\sum_{i=1}^{n}(D_i - \bar{D}_i)^2 / (n-1)}.$$

If the statistic value is larger than $t_{n-1,\alpha}$, we reject the null that both samples have the equal mean. Therefore, two distributions are not equal.

This test is asymptotically (large n) valid without normality. Our interest in it is that its power should serve as a standard of comparison for our two-sample tests if the two series are normal but with different means, a case that we will consider in our simulations.

3.3 MONTE CARLO SIMULATIONS

Our interest is to test equality of distribution for correlated variables. We will consider a paired sample $(y_i, x_i)_{i=1}^{n}$ that is drawn from a bivariate normal distribution with varying degree of correlation between Y and X. Therefore we simulate the samples based on the DGP

$$(3.8) \qquad y_i = \delta + \rho \bullet x_i + \sqrt{1-\rho^2} \bullet z_i \quad , \quad |\rho| < 1$$

where ρ is the correlation coefficient between y_i and x_i, and x_i and z_i are independent and identically distributed as $N(0,1)$. So $X \sim N(0,1)$ while $Y \sim N(\delta,1)$. The null hypothesis is $\delta = 0$ and the alternative is $\delta \neq 0$.

3.3.1 Size of the Test

In this section, we show the sizes of the various tests at the 0.05 nominal level. We will consider sample sizes of 50, 100, 250, 500, and 1000 with the correlation coefficient ρ varying from -0.99 to 0.95. In the χ^2 test, we will consider three different numbers of non-overlapping equiprobable cells: k = 4, 7, and 10.

In Table 3.1, columns 3, 4 and 5 show the χ^2 GMM test results based on 20,000 replications. The test is relatively accurate in all cases when the sample size is greater or equal to 250, in the sense that the actual size level is close to the nominal level. The test is slightly undersized for the smaller sample sizes ($n = 50$ and 100), but this is not a serious discrepancy for the 4 cell case.

For the smaller sample size cases, the simulations show that the χ^2 test with fewer cells has smaller size distortions. However, the size differences between tests with larger and smaller numbers of cells used in the test are not statistically different from zero when the sample size is large enough, e.g., 500 or 1000.

The size of the χ^2 test does not depend much on ρ. This agrees with the theoretical result that the test is asymptotically valid even if X and Y are correlated.

In Table 3.1, column 6 shows the size for the KS test. When $\rho = 0$, so that X and Y are independent, the size of the test is more or less correct so long as the sample size is not too small. However, the KS test is undersized when ρ is positive and oversized when ρ is negative. These size distortions are fairly severe when $|\rho|$ is large. For the BWS test, column 8 of Table 3.1 shows a similar pattern of size distortions, but the degree of oversize for negative ρ is larger than for the KS test.

Next we consider the bootstrapped version of the KS and BWS tests. We use 1000 replications in the simulations, and 399 bootstrap replications to calculate the critical values used in each replication of the simulations. We use the "novel" bootstrap method described in section 3.2.4.

In Table 3.1, columns 7 and 8 give these results. They are easy to summarize. The bootstrap-based tests are quite accurate (size is close to the nominal significance level.) except when ρ is positive and large, and n is small. In those cases, the bootstrap KS test is a little better (in terms of size) than the bootstrap BWS test.

As expected, the pairwise t-test (column 10) has correct size.

3.3.2 Power of the Test

Now we turn to the question of the power of the test. The alternative is $\delta \neq 0$ in (3.8) above. Tables 3.2-3.8 give the results for different non negative value of ρ. Also Tables 3.9-3.14 give similar results, which we will display but not discuss, for negative values of ρ.

Columns 3, 4, and 5 in Tables 3.2-3.8 show the power of the χ^2 test for the cases of k = 4, 7, and 10 cells. As expected, the power of the tests increases as the sample size is larger and as δ is larger. The power increases also when ρ is larger (and, from Table 3.1, this is *not* a reflection of size distortions). Finally, the power is larger when we use less cells. From Table 3.1, this last result may be a reflection of larger size distortion under the null when more cells are used.

When $\rho = 0$, results not reported here show that the KS and BWS tests (using the usual tabulated critical values) are more powerful than the χ^2 test. However, we do not recommend these tests because we have seen that they have large size distortion when $\rho \neq 0$.

Columns 6 and 7 in Tables 3.2-3.8 show the powers of the bootstrap KS and the bootstrap BWS tests. The results show both bootstrap tests are much more powerful than the χ^2 test. Moreover, the bootstrap BWS test is superior to the bootstrap KS test. The possible reason that the bootstrap BWS is better is that the weighting emphasizes the tails of the distribution functions, which increases the power of the test. (Baumgartner et al. (1998)). As for the χ^2 test, the power of the bootstrapped KS and BWS tests increases as n increases, as δ increases, and as ρ increases.

Finally, the last column of Tables 3.2-3.8 gives the power of the t-test. We expect this test to be more powerful than the other tests because the DGP satisfies exactly the assumption that underlies the test: normal distributions with equal variance but unequal means. And, indeed, the power of the t-test is greater than the power of the other tests (except when $\rho = 0$). However, the difference in power between the t-test and the bootstrapped BWS test is not very large. We interpret this as evidence favorable to the good power properties of the bootstrapped BWS test.

3.4 CONCLUDING REMARKS

In this paper, we have considered tests for equality of distributions. More specifically, we are interested in testing whether two correlated variables have the same distribution. This is different from many classical tests, e.g., the KS, the CM, and the Wilcoxon tests, which are widely used to test whether two *independent* samples belong to the same distribution.

We first apply the Pearson χ^2 GMM test based on equiprobable cells over the support to perform the test. This test compares the difference of the samples' observation counts in each cell. Cell boundaries are based on sample quantiles. Simulations show that the size of the tests is acceptably accurate in finite samples, provided not too many cells are used.

Next we consider the KS and BWS tests. These tests assume independence and have substantial size distortions when X and Y are correlated. We therefore propose a novel bootstrapping resampling scheme to obtain valid critical values. The bootstrapped KS and BWS tests have more or less correct size, and they are more powerful than the χ^2 tests. The bootstrapped BWS test appears to be the most powerful. In fact, its power is almost as large as that of the t-test for equality means, despite the fact that the DGP for the simulations favors the t-test.

In further research we plan to consider different types of alternatives. The results here were based on an alternative which is different from the null only by a constant mean. We plan to use copulas to simulate more flexible underlying distributions, including correlated distributions with non-normal marginals. In such cases, we can see whether the BWS test still outperforms the χ^2 test.

Table 3.1

Size of tests of the hypothesis that the data are bivariate- normal

Nominal size = 0.05

ρ	n	$\chi^2(3)$	$\chi^2(6)$	$\chi^2(9)$	KS	Bt.strp KS	BWS	Bt.strp BWS	T-test
-0.99	50	0.041	0.021	0.012	0.108	0.050	0.154	0.058	0.051
	100	0.043	0.040	0.028	0.131	0.044	0.154	0.057	0.053
	250	0.049	0.045	0.040	0.113	0.049	0.150	0.041	0.049
	500	0.047	0.047	0.045	0.125	0.042	0.157	0.046	0.050
	1000	0.050	0.046	0.047	0.124	0.045	0.151	0.049	0.051
-0.95	50	0.044	0.029	0.016	0.105	0.050	0.150	0.056	0.050
	100	0.044	0.043	0.034	0.131	0.054	0.149	0.049	0.053
	250	0.047	0.046	0.043	0.115	0.064	0.143	0.042	0.047
	500	0.051	0.048	0.049	0.124	0.055	0.151	0.044	0.050
	1000	0.051	0.050	0.049	0.126	0.044	0.146	0.054	0.051
-0.90	50	0.044	0.030	0.017	0.103	0.049	0.146	0.052	0.050
	100	0.044	0.046	0.033	0.131	0.054	0.146	0.054	0.053
	250	0.045	0.046	0.045	0.109	0.051	0.138	0.048	0.048
	500	0.049	0.048	0.047	0.124	0.053	0.146	0.052	0.051
	1000	0.051	0.049	0.049	0.123	0.049	0.141	0.051	0.051
-0.80	50	0.044	0.033	0.018	0.097	0.046	0.138	0.056	0.050
	100	0.047	0.046	0.035	0.124	0.048	0.135	0.056	0.053
	250	0.049	0.046	0.042	0.102	0.049	0.128	0.043	0.048
	500	0.047	0.047	0.046	0.116	0.049	0.135	0.048	0.051
	1000	0.050	0.050	0.052	0.114	0.051	0.134	0.053	0.051
-0.60	50	0.045	0.031	0.017	0.082	0.046	0.115	0.055	0.050
	100	0.046	0.045	0.034	0.106	0.056	0.116	0.059	0.053
	250	0.050	0.046	0.044	0.091	0.061	0.109	0.046	0.047
	500	0.046	0.050	0.046	0.098	0.048	0.115	0.054	0.050
	1000	0.051	0.049	0.048	0.101	0.050	0.116	0.051	0.051
-0.40	50	0.046	0.034	0.018	0.067	0.040	0.095	0.048	0.050
	100	0.045	0.042	0.032	0.089	0.056	0.096	0.053	0.053
	250	0.049	0.043	0.044	0.072	0.054	0.900	0.041	0.048
	500	0.048	0.047	0.044	0.082	0.051	0.093	0.057	0.050
	1000	0.051	0.049	0.050	0.087	0.052	0.096	0.048	0.051

Table 3.1 (cont'd.)

ρ	n	$\chi^2(3)$	$\chi^2(6)$	$\chi^2(9)$	KS	Bt.strp KS	BWS	Bt.strp BWS	T-test
-0.20	50	0.046	0.031	0.017	0.050	0.041	0.072	0.043	0.051
	100	0.045	0.043	0.034	0.070	0.052	0.076	0.055	0.053
	250	0.048	0.046	0.044	0.057	0.051	0.069	0.046	0.049
	500	0.048	0.047	0.043	0.064	0.061	0.072	0.046	0.050
	1000	0.052	0.049	0.050	0.072	0.054	0.076	0.050	0.052
0.00	50	0.043	0.033	0.017	0.037	0.048	0.052	0.036	0.050
	100	0.045	0.041	0.034	0.054	0.053	0.054	0.053	0.053
	250	0.050	0.045	0.043	0.042	0.048	0.052	0.045	0.049
	500	0.049	0.051	0.043	0.049	0.066	0.052	0.049	0.050
	1000	0.053	0.049	0.052	0.053	0.049	0.054	0.051	0.052
0.20	50	0.042	0.032	0.016	0.024	0.041	0.030	0.036	0.049
	100	0.044	0.042	0.035	0.036	0.045	0.033	0.046	0.054
	250	0.048	0.048	0.046	0.029	0.051	0.028	0.044	0.048
	500	0.051	0.048	0.044	0.034	0.039	0.029	0.050	0.051
	1000	0.053	0.050	0.051	0.035	0.047	0.033	0.050	0.052
0.40	50	0.044	0.030	0.015	0.012	0.044	0.015	0.033	0.049
	100	0.044	0.044	0.034	0.018	0.050	0.013	0.051	0.052
	250	0.049	0.045	0.045	0.015	0.045	0.013	0.044	0.049
	500	0.048	0.049	0.049	0.018	0.041	0.012	0.050	0.051
	1000	0.054	0.050	0.050	0.030	0.046	0.016	0.052	0.053
0.60	50	0.043	0.029	0.015	0.004	0.053	0.003	0.0300	0.050
	100	0.046	0.041	0.033	0.006	0.050	0.030	0.045	0.052
	250	0.048	0.047	0.045	0.006	0.045	0.002	0.041	0.048
	500	0.048	0.048	0.048	0.006	0.040	0.002	0.051	0.052
	1000	0.052	0.051	0.049	0.007	0.045	0.003	0.049	0.053
0.80	50	0.039	0.026	0.013	0.000	0.019	0.000	0.022	0.050
	100	0.045	0.041	0.033	0.000	0.039	0.000	0.038	0.051
	250	0.048	0.048	0.043	0.000	0.045	0.000	0.034	0.050
	500	0.048	0.048	0.045	0.000	0.042	0.000	0.053	0.052
	1000	0.049	0.049	0.050	0.000	0.048	0.000	0.048	0.051

Table 3.1 (cont'd.)

ρ	n	$\chi^2(3)$	$\chi^2(6)$	$\chi^2(9)$	KS	Bt.strp KS	BWS	Bt.strp BWS	T-test
0.90	50	0.034	0.023	0.010	0.000	0.025	0.000	0.007	0.051
	100	0.043	0.040	0.031	0.000	0.036	0.000	0.025	0.051
	250	0.046	0.044	0.044	0.000	0.037	0.000	0.030	0.050
	500	0.047	0.044	0.044	0.000	0.054	0.000	0.051	0.053
	1000	0.049	0.049	0.049	0.000	0.051	0.000	0.050	0.050
0.95	50	0.027	0.014	0.007	0.000	0.012	0.000	0.004	0.050
	100	0.034	0.033	0.023	0.000	0.024	0.000	0.020	0.051
	250	0.046	0.042	0.040	0.000	0.053	0.000	0.028	0.049
	500	0.047	0.047	0.044	0.000	0.055	0.000	0.050	0.053
	1000	0.049	0.050	0.047	0.000	0.049	0.000	0.053	0.051

Table 3.2

Power of the tests when $\rho = 0.00$

Nominal size = 0.05

Null: the data are bivariate-normal with the equal means and equal variances

Alternative: the data are from bivariate-normal with the mean difference δ

δ	n	$\chi^2(3)$	$\chi^2(6)$	$\chi^2(9)$	Bt.strp KS	Bt.strp BWS	T-test
0.10	50	0.061	0.036	0.020	0.066	0.060	0.056
	100	0.068	0.057	0.045	0.066	0.098	0.095
	250	0.122	0.087	0.078	0.150	0.175	0.141
	500	0.198	0.161	0.133	0.280	0.328	0.269
	1000	0.388	0.305	0.270	0.505	0.585	0.483
0.20	50	0.105	0.060	0.033	0.143	0.147	0.113
	100	0.156	0.116	0.088	0.211	0.271	0.235
	250	0.392	0.292	0.244	0.500	0.553	0.462
	500	0.681	0.594	0.423	0.774	0.863	0.777
	1000	0.951	0.926	0.890	0.979	0.990	0.972
0.30	50	0.182	0.103	0.054	0.238	0.273	0.216
	100	0.318	0.244	0.181	0.459	0.519	0.455
	250	0.753	0.649	0.577	0.883	0.905	0.810
	500	0.969	0.949	0.913	0.984	0.994	0.985
	1000	0.999	0.999	0.999	1.000	1.000	1.000
0.40	50	0.300	0.175	0.103	0.374	0.445	0.356
	100	0.545	0.440	0.351	0.684	0.759	0.698
	250	0.954	0.908	0.8693	0.973	0.993	0.970
	500	0.996	0.991	0.998	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
0.50	50	0.447	0.283	0.171	0.509	0.627	0.528
	100	0.763	0.669	0.563	0.856	0.915	0.865
	250	0.997	0.990	0.982	0.998	1.000	0.999
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
1.00	50	0.975	0.925	0.827	1.000	1.000	0.998
	100	1.000	1.000	0.993	1.000	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.3
Power of the tests when $\rho = 0.20$
Nominal size = 0.05
Null: the data are bivariate-normal with the equal means and equal variances
Alternative: the data are from bivariate-normal with the mean difference δ

δ	n	$\chi^2(3)$	$\chi^2(6)$	$\chi^2(9)$	Bt.strp KS	Bt.strp BWS	T-test
0.10	50	0.064	0.038	0.020	0.059	0.068	0.084
	100	0.076	0.061	0.049	0.089	0.107	0.120
	250	0.145	0.103	0.092	0.182	0.205	0.237
	500	0.232	0.186	0.158	0.310	0.377	0.420
	1000	0.459	0.379	0.329	0.578	0.630	0.711
0.20	50	0.115	0.063	0.034	0.134	0.159	0.192
	100	0.178	0.141	0.103	0.242	0.319	0.345
	250	0.468	0.364	0.307	0.564	0.651	0.706
	500	0.773	0.699	0.623	0.829	0.918	0.939
	1000	0.981	0.969	0.953	0.993	1.000	0.999
0.30	50	0.211	0.120	0.062	0.262	0.314	0.376
	100	0.380	0.298	0.224	0.489	0.606	0.645
	250	0.835	0.758	0.689	0.878	0.953	0.963
	500	0.990	0.982	0.967	0.995	1.000	1.000
	1000	1.000	1.000	0.999	1.000	1.000	1.000
0.40	50	0.356	0.217	0.125	0.447	0.507	0.591
	100	0.637	0.536	0.439	0.739	0.842	0.876
	250	0.981	0.964	0.942	0.990	1.000	0.991
	500	1.000	1.000	0.999	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
0.50	50	0.523	0.357	0.221	0.635	0.714	0.784
	100	0.843	0.771	0.676	0.904	0.961	0.974
	250	0.999	0.999	0.996	0.999	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
1.00	50	0.991	0.972	0.918	0.997	0.999	0.999
	100	1.000	1.000	1.000	1.000	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.4
Power of the tests when $\rho = 0.40$
Nominal size = 0.05
Null: the data are bivariate-normal with the equal means and equal variances
Alternative: the data are from bivariate-normal with the mean difference δ

δ	n	$\chi^2(3)$	$\chi^2(6)$	$\chi^2(9)$	Bt.strp KS	Bt.strp BWS	T-test
0.10	50	0.067	0.041	0.021	0.070	0.065	0.096
	100	0.079	0.069	0.052	0.105	0.188	0.146
	250	0.172	0.131	0.109	0.197	0.247	0.303
	500	0.290	0.243	0.251	0.361	0.457	0.526
	1000	0.564	0.494	0.433	0.648	0.745	0.828
0.20	50	0.136	0.075	0.038	0.165	0.181	0.245
	100	0.221	0.169	0.131	0.297	0.384	0.436
	250	0.576	0.475	0.409	0.640	0.763	0.826
	500	0.871	0.831	0.776	0.904	0.969	0.984
	1000	0.995	0.999	0.990	1.000	1.000	0.999
0.30	50	0.262	0.150	0.088	0.295	0.381	0.473
	100	0.477	0.385	0.310	0.558	0.702	0.768
	250	0.918	0.872	0.831	0.936	0.985	0.990
	500	0.999	0.998	0.959	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
0.40	50	0.439	0.283	0177	0.506	0.610	0.717
	100	0.752	0.670	0.580	0.847	0.915	0.949
	250	0.996	0.672	0.985	1.000	1.000	1.000
	500	1.000	0.999	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
0.50	50	0.522	0.466	0.316	0.696	0.836	0.887
	100	0.843	0.884	0.825	0.951	0.987	0.994
	250	0.999	1.000	0.999	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
1.00	50	0.991	0.994	0.980	1.000	1.000	1.000
	100	1.000	1.000	1.000	1.000	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.5
Power of the tests when $\rho = 0.60$
Null: the data are bivariate-normal with the equal means and equal variances
Alternative: the data are from bivariate-normal with the mean difference δ

δ	n	$\chi^2(3)$	$\chi^2(6)$	$\chi^2(9)$	Bt.strp KS	Bt.strp BWS	T-test
0.10	50	0.075	0.041	0.023	0.056	0.079	0.121
	100	0.094	0.083	0.055	0.112	0.139	0.197
	250	0.224	0.174	0.149	0.270	0.334	0.423
	500	0.394	0.349	0.302	0.430	0.602	0.702
	1000	0.721	0.683	0.626	0.758	0.899	0.946
0.20	50	0.172	0.097	0.052	0.189	0.236	0.350
	100	0.295	0.246	0.184	0.346	0.482	0.595
	250	0.731	0.659	0.599	0.770	0.891	0.944
	500	0.959	0.949	0.932	0.969	0.996	0.999
	1000	0.999	1.000	1.000	1.000	1.000	1.000
0.30	50	0.353	0.222	0.132	0.352	0.484	0.642
	100	0.625	0.558	0.473	0.676	0.837	0.909
	250	0.981	0.970	0.957	0.978	1.000	0.997
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
0.40	50	0.438	0.420	0.279	0.604	0.752	0.875
	100	0.752	0.851	0.782	0.906	0.976	0.993
	250	0.996	0.999	0.999	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
0.50	50	0.634	0.649	0.496	0.696	0.946	0.973
	100	0.925	0.974	0.956	0.951	1.000	1.000
	250	0.999	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
1.00	50	1.000	0.999	0.999	1.000	1.000	1.000
	100	1.000	1.000	1.000	1.000	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.6
Power of the tests when $\rho = 0.80$
Null: the data are bivariate-normal with the equal means and equal variances
Alternative: the data are from bivariate-normal with the mean difference δ

δ	n	$\chi^2(3)$	$\chi^2(6)$	$\chi^2(9)$	Bt.strp KS	Bt.strp BWS	T-test
0.10	50	0.095	0.046	0.025	0.083	0.081	0.195
	100	0.138	0.114	0.086	0.158	0.205	0.342
	250	0.369	0.309	0.274	0.372	0.524	0.710
	500	0.633	0.629	0.577	0.646	0.856	0.941
	1000	0.933	0.941	0.931	0.945	0.993	0.998
0.20	50	0.269	0.159	0.092	0.256	0.340	0.592
	100	0.492	0.454	0.376	0.481	0.710	0.876
	250	0.934	0.928	0.901	0.926	0.999	0.999
	500	0.999	1.000	0.999	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
0.30	50	0.570	0.420	0.287	0.537	0.724	0.908
	100	0.874	0.864	0.815	0.881	0.974	0.996
	250	0.999	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
0.40	50	0.832	0.733	0.592	0.789	0.955	0.992
	100	0.989	0.991	0.983	0.881	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
0.50	50	0.963	0.930	0.861	0.931	0.997	1.000
	100	0.999	1.000	1.000	0.994	0.999	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
1.00	50	1.000	1.000	1.000	1.000	1.000	1.000
	100	1.000	1.000	1.000	1.000	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.7

Power of the tests when $\rho = 0.90$

Nominal size = 0.05

Null: the data are bivariate-normal with the equal means and equal variances

Alternative: the data are from bivariate-normal with the mean difference δ

δ	n	$\chi^2(3)$	$\chi^2(6)$	$\chi^2(9)$	Bt.strp KS	Bt.strp BWS	T-test
0.10	50	0.121	0.064	0.030	0.086	0.099	0.340
	100	0.202	0.184	0.147	0.207	0.324	0.592
	250	0.574	0.562	0.529	0.531	0.797	0.941
	500	0.865	0.901	0.890	0.877	0.988	0.999
	1000	0.995	0.999	0.998	0.999	1.000	1.000
0.20	50	0.426	0.293	0.188	0.308	0.515	0.875
	100	0.730	0.752	0.701	0.674	0.926	0.993
	250	0.996	0.998	0.998	0.995	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
0.30	50	0.799	0.719	0.600	0.653	0.932	0.996
	100	0.984	0.990	0.990	0.977	0.999	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
0.40	50	0.971	0.958	0.922	0.919	0.998	1.000
	100	1.000	1.000	1.000	1.000	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
0.50	50	0.998	0.998	1.000	0.980	1.000	1.000
	100	1.000	1.000	1.000	1.000	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
1.00	50	1.000	1.000	1.000	1.000	1.000	1.000
	100	1.000	1.000	1.000	1.000	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.8

Power of the tests when $\rho = 0.95$

. Nominal size = 0.05

Null: the data are bivariate-normal with the equal means and equal variances

Alternative: the data are from bivariate-normal with the mean difference δ

δ	n	$\chi^2(3)$	$\chi^2(6)$	$\chi^2(9)$	Bt.strp KS	Bt.strp BWS	T-test
0.10	50	0.160	0.084	0.042	0.102	0.112	0.595
	100	0.300	0.313	0.279	0.270	0.484	0.876
	250	0.780	0.834	0.836	0.736	0.966	0.999
	500	0.974	0.995	0.997	0.975	0.999	1.000
	1000	0.999	1.000	1.000	1.000	1.000	1.000
0.20	50	0.612	0.528	0.412	0.431	0.758	0.993
	100	0.960	0.958	0.951	0.853	0.995	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
0.30	50	0.940	0.943	0.915	0.803	1.000	1.000
	100	0.999	1.000	1.000	1.000	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
0.40	50	0.998	0.999	1.000	0.973	1.000	1.000
	100	1.000	1.000	1.000	1.000	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
0.50	50	1.000	1.000	1.000	1.000	1.000	1.000
	100	1.000	1.000	1.000	1.000	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
1.00	50	1.000	1.000	1.000	1.000	1.000	1.000
	100	1.000	1.000	1.000	1.000	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.9
Power of the tests when ρ = - 0.20
Nominal size = 0.05
Null: the data are bivariate-normal with the equal means and equal variances
Alternative: the data are from bivariate-normal with the mean difference δ

δ	n	$\chi^2(3)$	$\chi^2(6)$	$\chi^2(9)$	Bt.strp KS	Bt.strp BWS	T-test
0.10	50	0.055	0.045	0.024	0.056	0.060	0.063
	100	0.055	0.054	0.047	0.099	0.095	0.094
	250	0.110	0.088	0.078	0.132	0.151	0.167
	500	0.187	0.255	0.102	0.253	0.279	0.298
	1000	0.317	0.255	0.218	0.432	0.482	0.508
0.20	50	0.090	0.046	0.033	0.121	0.137	0.141
	100	0.141	0.108	0.008	0.192	0.225	0.242
	250	0.319	0.240	0.209	0.415	0.495	0.524
	500	0.609	0.514	0.417	0.722	0.804	0.818
	1000	0.914	0.854	0.808	0.958	0.980	0.985
0.30	50	0.163	0.082	0.043	0.215	0.241	0.265
	100	0.286	0.216	0.158	0.398	0.473	0.496
	250	0.661	0.562	0.467	0.763	0.846	0.863
	500	0.930	0.903	0.864	0.967	0.981	0.987
	1000	0.999	0.998	0.994	1.000	1.000	1.000
0.40	50	0.245	0.137	0.074	0.314	0.394	0.425
	100	0.480	0.376	0.286	0.590	0.695	0.719
	250	0.921	0.855	0.798	0.954	0.983	0.981
	500	0.998	0.994	0.988	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
0.50	50	0.369	0.219	0.136	0.490	0.570	0.611
	100	0.674	0.573	0.464	0.789	0.864	0.878
	250	0.992	0.976	0.962	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
1.00	50	0.954	0.881	0.714	0.975	0.990	0.993
	100	0.999	0.999	0.996	1.000	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.10
Power of the tests when ρ = - 0.40
Nominal size = 0.05
Null: the data are bivariate-normal with the equal means and equal variances
Alternative: the data are from bivariate-normal with the mean difference δ

δ	n	$\chi^2(3)$	$\chi^2(6)$	$\chi^2(9)$	Bt.strp KS	Bt.strp BWS	T-test
0.10	50	0.060	0.032	0.028	0.061	0.067	0.065
	100	0.050	0.048	0.050	0.090	0.090	0.090
	250	0.097	0.077	0.068	0.119	0.143	0.153
	500	0.171	0.128	0.100	0.238	0.253	0.262
	1000	0.278	0.209	0.197	0.394	0.432	0.439
0.20	50	0.077	0.059	0.027	0.110	0.126	0.117
	100	0.129	0.081	0.066	0.163	0.224	0.220
	250	0.289	0.206	0.192	0.405	0.434	0.466
	500	0.539	0.438	0.379	0.651	0.763	0.758
	1000	0.865	0.787	0.732	0.926	0.970	0.974
0.30	50	0.126	0.077	0.042	0.203	0.225	0.233
	100	0.256	0.191	0.135	0.359	0.426	0.432
	250	0.598	0.481	0.412	0.729	0.795	0.812
	500	0.896	0.845	0.777	0.949	0.973	0.975
	1000	0.997	0.993	0.988	1.000	1.000	1.000
0.40	50	0.224	0.117	0.061	0.317	0.351	0.372
	100	0.439	0.315	0.251	0.576	0.633	0.657
	250	0.884	0.772	0.701	0.919	0.960	0.968
	500	0.993	0.982	0.975	1.000	1.000	1.000
	1000	1.000	1.00	1.000	1.000	1.000	1.000
0.50	50	0.339	0.210	0.177	0.437	0.512	0.544
	100	0.622	0.506	0.380	0.722	0.823	0.840
	250	0.979	0.950	0.927	0.986	0.990	0.999
	500	1.000	1.000	0.999	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
1.00	50	0.927	0.806	0.636	0.958	0.979	0.989
	100	0.999	0.994	0.989	1.000	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.11

Power of the tests when ρ = - 0.60

Nominal size = 0.05

Null: the data are bivariate-normal with the equal means and equal variances

Alternative: the data are from bivariate-normal with the mean difference δ

δ	n	$\chi^2(3)$	$\chi^2(6)$	$\chi^2(9)$	Bt.strp KS	Bt.strp BWS	T-test
0.10	50	0.064	0.035	0.021	0.059	0.070	0.065
	100	0.077	0.061	0.048	0.080	0.084	0.086
	250	0.096	0.078	0.077	0.118	0.126	0.142
	500	0.141	0.120	0.109	0.220	0.232	0.232
	1000	0.242	0.175	0.177	0.349	0.389	0.394
0.20	50	0.086	0.044	0.018	0.111	0.115	0.116
	100	0.120	0.105	0.066	0.152	0.205	0.194
	250	0.267	0181	0.160	0.350	0.410	0.419
	500	0.481	0.400	0.332	0.613	0.673	0.683
	1000	0.804	0.718	0.656	0.896	0.947	0.951
0.30	50	0.132	0.070	0.035	0.184	0.200	0.205
	100	0.217	0.166	0.118	0.333	0.386	0.387
	250	0.528	0.424	0.366	0.671	0.744	0.754
	500	0.862	0.776	0.712	0.915	0.955	0.960
	1000	0.991	0.984	0.974	1.000	1.000	1.000
0.40	50	0.195	0.115	0.070	0.279	0.321	0.332
	100	0.379	0.281	0.209	0.509	0.574	0.582
	250	0.828	0.716	0.632	0.885	0.930	1.000
	500	0.982	0.965	0.947	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
0.50	50	0.295	0.185	0.108	0.394	0.464	0.478
	100	0.554	0.460	0.365	0.699	0.788	0.787
	250	0.959	0.915	0.881	0.988	0.997	0.996
	500	1.000	1.000	0.993	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
1.00	50	0.877	0.737	0.574	0.935	0.964	0.971
	100	0.997	0.985	0.965	1.000	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.12
Power of the tests when $\rho = -0.80$
Nominal size = 0.05
Null: the data are bivariate-normal with the equal means and equal variances
Alternative: the data are from bivariate-normal with the mean difference δ

δ	n	$\chi^2(3)$	$\chi^2(6)$	$\chi^2(9)$	Bt.strp KS	Bt.strp BWS	T-test
0.10	50	0.051	0.036	0.023	0.061	0.073	0.066
	100	0.057	0.055	0.054	0.082	0.083	0.088
	250	0.083	0.076	0.067	0.101	0.121	0.132
	500	0.134	0.105	0.108	0.205	0.205	0.208
	1000	0.225	0.166	0.157	0.312	0.357	0.363
0.20	50	0.082	0.043	0.024	0.102	0.110	0.106
	100	0.102	0.076	0.068	0.153	0.187	0.173
	250	0.226	0.165	0.150	0.317	0.364	0.379
	500	0.435	0.356	0.301	0.553	0.625	0.632
	1000	0.747	0.661	0.603	0.855	0.918	0.919
0.30	50	0.129	0.065	0.036	0.1650	0.178	0.188
	100	0.199	0.145	0.100	0.286	0.336	0.338
	250	0.477	0.347	0.309	0.608	0.681	0.705
	500	0.800	0.718	0.663	0.886	0.921	0.927
	1000	0.989	0.973	0.956	0.993	0.995	0.999
0.40	50	0.197	0.095	0.057	0.271	0.300	0.298
	100	0.333	0.259	0.183	0.464	0.522	0.535
	250	0.764	0.636	0.566	0.885	0.904	0.907
	500	0.964	0.940	0.918	0.987	0.990	0.992
	1000	1.000	1.000	0.998	1.000	1.000	1.000
0.50	50	0.280	0.149	0.084	0.362	0.419	0.445
	100	0.504	0.402	0.323	0.643	0.725	0.731
	250	0.930	0.862	0.806	0.965	0.988	0.988
	500	0.999	0.999	0.990	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
1.00	50	0.843	0.675	0.512	0.905	0.942	0.950
	100	0.992	0.969	0.942	0.907	0.989	0.999
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.13
Power of the tests when $\rho = -0.90$
Nominal size = 0.05
Null: the data are bivariate-normal with the equal means and equal variances
Alternative: the data are from bivariate-normal with the mean difference δ

δ	n	$\chi^2(3)$	$\chi^2(6)$	$\chi^2(9)$	Bt.strp KS	Bt.strp BWS	T-test
0.10	50	0.054	0.034	0.020	0.065	0.079	0.071
	100	0.051	0.047	0.044	0.081	0.090	0.093
	250	0.083	0.062	0.058	0.104	0.123	0.128
	500	0.126	0.119	0.093	0.184	0.193	0.197
	1000	0.207	0.159	0.126	0.297	0.338	0.345
0.20	50	0.069	0.033	0.022	0.105	0.109	0.105
	100	0.109	0.085	0.061	0.160	0.179	0.172
	250	0.215	0.147	0.132	0.293	0.346	0.360
	500	0.397	0.328	0.282	0.515	0.588	0.613
	1000	0.712	0.621	0.568	0.821	0.895	0.904
0.30	50	0.116	0.072	0.042	0.159	0.171	0.177
	100	0.189	0.138	0.107	0.263	0.317	0.326
	250	0.451	0.346	0.298	0.578	0.655	0.680
	500	0.771	0.682	0.644	0.860	0.903	0.916
	1000	0.986	0.963	0.940	0.995	0.999	0.999
0.40	50	0.180	0.093	0.050	0.224	0.276	0.284
	100	0.307	0.236	0.168	0.435	0.506	0.518
	250	0.746	0.623	0.522	0.824	0.890	0.889
	500	0.952	0.937	0.911	0.980	0.990	0.990
	1000	1.000	0.999	0.998	1.000	1.000	1.000
0.50	50	0.268	0.142	0.079	0.351	0.414	0.418
	100	0.477	0.384	0.288	0.605	0.703	0.708
	250	0.920	0.854	0.792	0.953	0.979	0.983
	500	0.998	0.994	0.986	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
1.00	50	0.830	0.646	0.478	0.881	0.929	0.940
	100	0.991	0.968	0.936	0.991	0.995	0.990
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.14

Power of the tests when $\rho = -0.95$

. Nominal size = 0.05

Null: the data are bivariate-normal with the equal means and equal variances

Alternative: the data are from bivariate-normal with the mean difference δ

δ

δ	n	$\chi^2(3)$	$\chi^2(6)$	$\chi^2(9)$	Bt.strp KS	Bt.strp BWS	T-test
0.10	50	0.056	0.035	0.023	0.063	0.078	0.072
	100	0.048	0.053	0.042	0.085	0.088	0.095
	250	0.085	0.660	0.059	0.097	0.116	0.126
	500	0.132	0.108	0.088	0.191	0.188	0.190
	1000	0.210	0.162	0.131	0.287	0.298	0.338
0.20	50	0.073	0.038	0.019	0.091	0.105	0.110
	100	0.090	0.072	0.069	0.144	0.174	0.172
	250	0.211	0.155	0.135	0.294	0.342	0.352
	500	0.388	0.323	0.260	0.503	0.503	0.597
	1000	0.699	0.617	0.562	0.819	0.884	0.897
0.30	50	0.106	0.057	0.040	0.167	0.172	0.172
	100	0.190	0.137	0.101	0.243	0.308	0.317
	250	0.438	0.342	0.295	0.592	0.655	0.667
	500	0.754	0.668	0.602	0.855	0.909	0.910
	1000	0.985	0.953	0.934	1.000	1.000	1.000
0.40	50	0.170	0.089	0.050	0.239	0.284	0.281
	100	0.304	0.230	0.166	0.411	0.493	0.495
	250	0.727	0.617	0.520	0.811	0.874	0.882
	500	0.994	0.925	0.892	1.000	1.000	1.000
	1000	1.000	0.999	0.999	1.000	1.000	1.000
0.50	50	0.259	0.153	0.083	0.344	0.404	0.416
	100	0.452	0.368	0.277	0.616	0.679	0.697
	250	0.913	0.837	0.781	0.943	0.981	0.978
	500	0.996	0.990	0.982	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000
1.00	50	0.814	0.639	0.460	0.879	0.924	0.936
	100	0.986	0.962	0.914	0.993	0.990	0.990
	250	1.000	1.000	1.000	1.000	1.000	1.000
	500	1.000	1.000	1.000	1.000	1.000	1.000
	1000	1.000	1.000	1.000	1.000	1.000	1.000

BIBLIOGRAPHY

Abadir, K.M. and J.R. Magnus (2005), *Matrix Algebra* (Econometric Exercises, Volume 1), Cambridge, Cambridge University Press.

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.

Baumgartner, W., P. Weiss and H. Schindler (1998), "A Nonparametric Test for the General Two-Sample Problem," *Biometrics* 54, 1129-1135.

Davidson, R. and J.G. MacKinnon (2000), "Bootstrap Tests: How Many Bootstraps?" *Econometric Reviews* 19, 55-68.

Giné, E. and J. Zinn (1990), "Bootstrapping General Empirical Measures," *Annals of Probability* 18, 851-869.

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.

Hansen, L.P. (1982)," Large Sample Properties of Generalized Method Moment Estimators," *Econometrica* 50, 1029-1054.

Heckman, J. (1984), "The χ^2 Goodness of Fit for Models Estimated from Microdata," *Econometrica* 52, 1543-1548.

Horrace, W.C. and P. Schmidt (1996), "Confidence Statements for Efficiency Estimates from Stochastic Frontier Models," *Journal of Productivity Analysis* 7, 257-282.

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.

Kumbhakar, S.C., B.U. Park, L. Simar and E.G. Tsionas (2007), "Nonparametric Stochastic Frontiers: A Local Maximum Likelihood Approach," *Journal of Econometrics* 137, 1-27.

Li, Q., E. Maasoumi and J. Racine (2009), "A Nonparametric Test for Equality of Distributions with Mixed Categorical and Continuous Data," *Journal of Econometrics* 148, 186-200.

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

Newey, W.K. (1985), "Maximum Likelihood Specification Testing and Conditional Moment Tests," *Econometrica* 53, 1047-1070.

Pearson E.S. and H.O. Hartely (1972), "Biometrika Tables for Statisticians Vol. 2," *Cambridge University Press*, Cambridge.

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.

Ruppert, D. and R.J. Carroll (1980), "Trimmed Least Squares Estimation in the Linear Model," *Journal of the American Statistical Association* 75, 828-838.

Schmidt, P. and R. Sickles (1984), "Production Frontiers and Panel Data," *Journal of Business and Economic Statistics* 2, 367-374.

Sheskin, D.J. (2003), "Handbook of Parametric and Nonparametric Statistical Procedures," *Chapman and Hall*, 3rd edition.

Simar, L. and P.W. Wilson (2009), "Inferences from Cross-Sectional Stochastic Frontier Models," *Econometric Reviews*, forthcoming.

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

Stuart, A. and J.K. Ord (1987), "Advanced Theory of Statistics Vol. 2," *Oxford University Press*, 1159-1196.

Stute, W., W. Gonzáles Manteiga and M. Presedo Quindimil (1993), "Bootstrap Based Goodness of Fit Tests," *Metrika* 40, 243-256.

Tauchen, G. (1985), "Diagnostic Testing and Evaluation of Maximum Likelihood Models," *Journal of Econometrics* 30, 415-444.

Tallis, G.M. (1983), "Goodness of Fit," pp. 451-461 in S. Kotz and N.L. Johnson, *Encyclopedia of Statistical Sciences,* Vol. 3, New York, Wiley.

Waldman, D.M. (1984), "Properties of Technical Efficiency Estimators in the Stochastic Frontier Model," *Journal of Econometrics* 25, 353-364.

Wang, W.S., C. Amsler and P. Schmidt (2009), "Goodness of Fit Tests in Stochastic Frontier Models," *Working Paper.*

Wang, W.S. and P. Schmidt (2009), "On the Distribution of Estimated Technical Efficiency in Stochastic Frontier Models," *Journal of Econometrics* 148, 36-45.