University of Illinois **Final Exam Answers** Department of Economics
Spring 2010 **Economics 574** Professor Roger Koenker

1. (This is pretty mechanical, but perhaps worthwhile to develop some facility with the linear algebra of 2SLS.)

a.) Premultiply y, Z, W by M_X to get $\tilde{y}, \tilde{Z}, \tilde{W}$, and now do 2SLS using $\tilde{W}$ as IV's for $\tilde{Z}$:

$$\hat{\alpha}_{\text{2SLS}} = (\tilde{Z}^\top P_{\tilde{W}} \tilde{Z})^{-1} \tilde{Z}^\top P_{\tilde{W}} \tilde{y}$$

Note that $P_{\tilde{W}} = P_{M_X W}$.

b.) Let $A = W^\top M_X W$, then $\|\hat{\gamma}(\alpha)\|_2^2 = (y - Z\alpha)^\top P_{\tilde{W}} (y - Z\alpha)$ so minimizing gives $\hat{\alpha}_{\text{2SLS}}$ by the idempotency of M_X.

c. The model asserts that W satisfies an exclusion restriction, so the intuition, or at least my intuition, for the procedure is that we want to choose α in such a way that it makes this exclusion as "true as possible."

2. This was an LLN question.

a.) Trivial

b.) Checking SOC for a.) confirms that

$$f(\theta) = p \log(1 + \theta) + (1 - p) \log(1 - \theta)$$

is increasing for $\theta < \hat{\theta}$, and decreasing for $\theta > \hat{\theta}$. A picture reveals that there are two points at which $f(\theta) = 0$, one at $\theta = 0$ where as they say, "nothing ventured, nothing gained" (or lost), another at $\approx$0.389. Let X_n be the Bernoulli random variable taking the value one with probability p and zero with probability $q = 1 - p$. Then,

$$\begin{aligned} \log W_n &= \log W_{n-1} + \log(1 - \theta) + X_n \log R(\theta) \\ &= \log W_0 + n \log(1 - \theta) + S_n \log R(\theta) \end{aligned}$$

where $S_n = \sum_{i=1}^n X_k$ and $R(\theta) = (1 + \theta)/(1 - \theta)$. Taking expectations, the crucial thing is whether the "drift" term is positive or negative. In the former case wealth increases without bound, in the latter it converges to zero by the LLN. Solving for the driftless case yields $\tilde{\theta}$.

c.) My intention in this part was that in each period you were entitled to invest proportions $\theta_k,\ k = 1, ..., K$ of your wealth on K coins having success probabilities p_k. With two coins things are simplest and we have payoffs:

$$Z = \left\{ \begin{array}{lcl} 1+\theta_1+\theta_2 & & p_1p_2 \\ 1+\theta_1-\theta_2 & & p_1(1-p_2) \\ & wp & \\ 1-\theta_1+\theta_2 & & (1-p_1)p_2 \\ 1-\theta_1-\theta_2 & & (1-p_1)(1-p_2) \end{array} \right.$$

It is advantageous to put money on any coin with $p_k > 1/2$ for diversification reasons. As one might expect, more money is allocated to higher probability coins, but you shouldn't neglect the less profitable ones entirely. At this point the problem becomes numerical, and I didn't expect anyone to go further, although some did.

3. This was a somewhat perverse MLE problem, in the sense that we were using a local maximizer when the global maximizer was known to be terrible. Surprisingly, the local maximimizer works quite well.

a.) $F_X(x) = \Phi((\log(x-\alpha)-\mu)/\sigma)$ so by the chain rule, $f_X(x) = \phi((\log(x-\alpha)-\mu)/\sigma)/(x-\alpha)$.

b.) This was intended to be a exercise in inequality manipulation:

$$\begin{aligned} \hat{\sigma}^2(\alpha) &= n^{-1}\sum(\log(x_i-\alpha)-\hat{\mu})^2 \\ &\leq n^{-1}\sum(\log(x_i-\alpha))^2 \quad \hat{\mu} \text{ is a minimizer} \\ &\leq n^{-1}\sum(\log(x_{(1)}-\alpha))^2 \quad \text{for } \alpha \text{near } x_{(1)} \\ &= (\log(x_{(1)}-\alpha))^2. \end{aligned}$$

Thus,

$$\begin{aligned} L(\alpha) &= K\hat{\sigma}^{-n}\prod_{i=1}^{n}(x_i-\alpha) \\ &\geq K\hat{\sigma}^{-n}(x_{(1)}-\alpha) \end{aligned}$$

and repeated application of l'Hôpital's rule implies that this expression tends to infinity as $\alpha \to x_{(1)}$.

c.) The local likelihood maximizer can be computed easily in R, and the local likelihood can be used to construct a co:nfidence interval. Most exams that tried to construct a confidence interval did a Wald version using the asymptotic covariance matrix – this entails evaluation of the inverse of the Fisher information matrix and is quite arduous. Using the likelihood is quite simple as the figure below illustrates.

d. Simulations are also quite easy in R. Code for my version of the simulations is included along with the code for the figure for part c.) in an appendix. One problem with the comparison of estimates of μ and

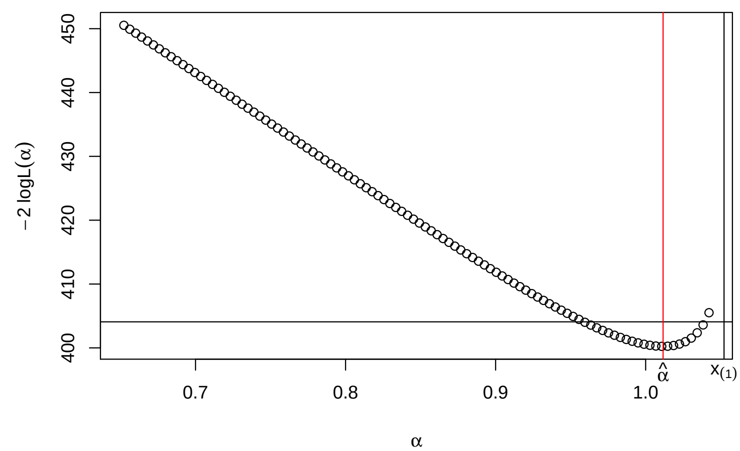

FIGURE 1. This figure plots twice the concentrated log likelihood, $-2\log L$, as a function of α for a sample of 200 observations from the three parameter lognormal distribution. A point estimate of α is indicated by $\hat{\alpha}$ and the vertical red line. The horizontal black line is drawn at $-2\log L(\hat{\alpha}) + C$ where C is the 0.95 quantile of a χ^2_1 random variable, so the 0.95 confidence interval is simply the interval for which twice the log likelihood is below this horizontal line. Note that the resulting confidence interval is asymmetric, in this situation we know "for sure" that α is less than $x_{(1)}$.

σ with and without estimating α is that there is a strong correlation between the estimates of α and the other parameters. So I thought it might be interesting to compare estimates of some functional that was invariant like the median. When you run the simulation you see that estimating α seems very costly in terms of estimating μ and σ, but if you are really interested ultimately in using these estimates to estimate the median, then estimating α is almost costless: the first component of MSEE is only slightly bigger than the second one based on knowing that $\alpha = 1$, and both of these parametric estimates are quite a bit better than the sample median. Of course knowing that we are in the three parameter lognormal model is quite a leap of faith.

Appendix A. Code for the Lognormal Figure

```
# Plot likelihood for 3 parm lognormal
loglik <- function(a,x){
```

```
        n <- length(x)
        y <- log(x-a)
        s <- sqrt(var(y))
        n * s + sum(log(x-a))
        }
x <- 1 + exp(rnorm(200))
minx <- min(x)
as <- seq(minx - .4, minx - .01,length = 100)
fs <- as
for(i in 1:length(as)){
        fs[i] <- 2 * loglik(as[i],x)
        }
pdf("lnorm.pdf", height = 5, width = 7)
plot(as,fs,xlab = expression(alpha), ylab = expression(-2~logL( alpha )))
a <- optimize(loglik,c(minx - .5, minx - .01),x = x)
ahat <- a$minimum
lhat <- a$objective
calpha <- qchisq(.95,1)
abline(h = 2*lhat + calpha)
mtext(expression(x[(1)]),1,0,at = min(x))
mtext(expression(hat(alpha)),1,0,at = ahat)
abline(v = ahat,col = "red")
abline(v = min(x),col = "black")
dev.off()
```

Appendix B. Code for the Lognormal Simulation

```
# likelihood for 3 parm lognormal
loglik <- function(a,x){
        n <- length(x)
        y <- log(x-a)
        s <- sqrt(var(y))
        n * s + sum(log(x-a))
        }
R <- 1000
n <- 200
A <- matrix(0,3,R)
B <- matrix(0,2,R)
E <- matrix(0,3,R)
for(j in 1:R){
        x <- 1 + exp(rnorm(n))
        B[1,j] <- mean(log(x - 1))
```

```
        B[2,j] <- var(log(x - 1))
        minx <- min(x)
        ahat <- optimize(loglik,c(minx - .1, minx - .0001),x = x)$minimum
        A[1,j] <- ahat
        A[2,j] <- mean(log(x - ahat))
        A[3,j] <- var(log(x - ahat))
        E[1,j] <- ahat + exp(A[2,j])
        E[2,j] <- 1 + exp(B[1,j])
        E[3,j] <- quantile(x,.5)
        }
A <- t(A - c(1,0,1))
B <- t(B - c(0,1))
E <- t(E - 2)
MSEA <- apply(A^2,2,mean) # MSE for estimated alpha
MSEB <- apply(B^2,2,mean) # MSE for known alpha
MSEE <- apply(E^2,2,mean) # MSE for median estimates
```