

APPENDIX TO:

**BAYES ESTIMATES OF DISTANCE-TO-MARKET:
 TRANSACTIONS COSTS, COOPERATIVES AND MILK-MARKET
 DEVELOPMENT IN THE ETHIOPIAN HIGHLANDS**

**DISTRIBUTIONS OF MODEL PARAMETERS AND DISTANCE ESTIMATES
 and
 MATLAB™ CODE**

With reference to the distributions of the model parameters, given $\pi(\theta|y, z) \propto f^{MN}(v|qz_p + x\beta, \Sigma) f^{TN}(z_p|w\delta, \omega^2 I_N) \pi(\theta)$, and noting that $\pi(\theta) \propto \omega^{-1} |\Sigma|^{-(m+1)/2}$, $\pi(\Sigma|\beta, z_s, \tau, z_p, \delta, \omega, y) \propto f^{MN}(v|qz_p + x\beta, \Sigma) |\Sigma|^{-(m+1)/2} \propto |\Sigma|^{-(S+m+1)/2} \exp\{-\frac{1}{2} \text{trace } \mathbf{A} \Sigma^{-1}\} \propto f^{iW}(\Sigma|\mathbf{A}, S, m)$, where $\mathbf{A} = \begin{pmatrix} \mathbf{a}_{ss} & \mathbf{a}_{sp} \\ \mathbf{a}_{ps} & \mathbf{a}_{pp} \end{pmatrix}$, $\mathbf{a}_{ss} = (\mathbf{v}_s - \mu_s)' (\mathbf{v}_s - \mu_s)$, $\mathbf{a}_{sp} = (\mathbf{v}_s - \mu_s)' (\mathbf{v}_p - \mu_p)$, $\mathbf{a}_{ps} = (\mathbf{v}_p - \mu_p)' (\mathbf{v}_s - \mu_s)$, $\mathbf{a}_{pp} = (\mathbf{v}_p - \mu_p)' (\mathbf{v}_p - \mu_p)$, $\mu_s = \mathbf{x}_s \beta_s$, $\mu_p = \mathbf{h} z_p + \mathbf{x}_p \beta_p$, $S = \sum_i T_i$ and $m = 2$.

Second, $\pi(\beta|z_s, \tau, z_p, \delta, \omega, \Sigma, y) \propto f^{MN}(v|qz_p + x\beta, \Sigma) \propto \exp\{-\frac{1}{2} (\mathbf{v} - \mathbf{q}z_p - \mathbf{x}\beta)' (\Sigma^{-1} \otimes \mathbf{I}_S) (\mathbf{v} - \mathbf{q}z_p - \mathbf{x}\beta)\} \propto \exp\{-\frac{1}{2} (\mathbf{A} - \mathbf{B}\beta)' \mathbf{C} (\mathbf{A} - \mathbf{B}\beta)\} \propto f^{MN}(\beta | (\mathbf{B}' \mathbf{C} \mathbf{B})^{-1} (\mathbf{B}' \mathbf{C} \mathbf{A}), (\mathbf{B}' \mathbf{C} \mathbf{B})^{-1})$, $\mathbf{A} = \mathbf{v}_p - \mathbf{h} z_p$, $\mathbf{B} = \mathbf{h}$, $\mathbf{C} = (\Sigma^{-1} \otimes \mathbf{I}_S)$. Third, noting that the censoring implies truncation at the threshold, τ , we have, for each $t \in C_i$, $\pi(z_{sit} | \tau, z_p, \delta, \omega, \Sigma, \beta, y) \propto \exp\{-\frac{1}{2} (\mathbf{z}_{it} - \mu_i)' \Sigma^{-1} (\mathbf{z}_{it} - \mu_i)\} [I(y_{sit} = 0) I(z_{sit} < \tau)]$, where I denotes the indicator function that equals one if the random variable is contained within the interval, and equals zero otherwise. Hence, $\pi(z_{sit} | \tau, z_p, \delta, \omega, \Sigma, \beta, y) \propto \exp\{-\frac{1}{2} (z_{si} - \mu_{s|pit})' \sigma_{s|p}^{-1} (z_{si} - \mu_{s|pit})\} [I(y_{sit} = 0) I(z_{sit} < \tau)] \propto f^{TN}(z_{sit} | \mu_{s|pit}, \sigma_{s|p})$, $\mu_{s|pit} = \mu_{sit} + \sigma_{sp} \sigma_{ss}^{-1} [y_{pit} - \mu_{pit}]$, $\mu_{sit} = \mathbf{x}_{sit} \beta_s$, $\sigma_{s|p} \equiv \sigma_{ss} - \sigma_{sp} \sigma_{pp}^{-1} \sigma_{ps}$. Fourth $\pi(\tau | z_p, \delta, \omega, \Sigma, \beta, z_s, y) \propto \prod_{i=1}^N \prod_{t \in C_i} [I(y_{sit} = z_{sit}) I(\tau < z_{sit})]$

$\prod_{t \in C_i} [I(y_{sit}=0)I(z_{sit}<\tau)]$ and τ has the uniform distribution on the interval $[a, b]$ where $a = \max$

$\{z_{sit}:y_{sit}=0\}$ and $b = \min \{z_{sit}:y_{sit}=z_{sit}\}$. That is, $\pi(\tau|z_p, \delta, \omega, \Sigma, \beta, z_s, y) = f^U(\tau|a, b)$. Fifth, $\pi(z_p|\delta, \omega, \Sigma, \beta, z_s, \tau, y) \propto f^{MN}(v|qz_p+x\beta, \Sigma) f^{TN}(z_p|w\delta, \omega^2 I_N) \propto \exp\{-\frac{1}{2}(v-qz_p-x\beta)'(\Sigma^{-1} \otimes I_S)(v-qz_p-x\beta)\} \exp\{-\frac{1}{2}(z_p-w\delta)'(\omega^{-2} I_N)(z_p-w\delta)\} \propto \exp\{-\frac{1}{2}(A-Bz_p)'C(A-Bz_p)\} \exp\{-\frac{1}{2}(z_p-D)'E(z_p-D)\} \propto f^{MN}(\beta|(B'CB+E)^{-1}(B'CA+ED), (B'CB+E)^{-1}), A = v-x\beta, B = h, C = \Sigma^{-1} \otimes I_S, D = w\delta$ and $E = \delta, \omega^2 I_N$. Sixth, $\pi(\delta|\omega, \Sigma, \beta, z_s, \tau, y) \propto f^{TN}(z_p|w\delta, \omega^2 I_N) \propto \exp\{-\frac{1}{2}(z_p-w\delta)'(\omega^{-2} I_N)(z_p-w\delta)\} \propto f^{MN}(\delta|(w'w)^{-1}w'z_p, \omega^{-2}(w'w)^{-1})$. Finally, $\pi(\omega|\Sigma, \beta, z_s, \tau, \delta, y) \propto \omega^{-(N+1)} \exp\{-\frac{1}{2}(z_p-w\delta)'(\omega^{-2} I_N)(z_p-w\delta)\} \propto f^{iG}(\omega|N, s^2), Ns^2 = (z-w\delta)'(z_p-w\delta)$.

With reference to the distributions of the distance estimates, we are interested in the existence of distributions and moments of the quantities in text-equation (13). Henceforth, (with a slight abuse of notation compared to that appearing in the text), consider the quotient

$z = \frac{x}{y}$, which is the first term on the right-side of the equation above and note that the joint

distribution of x and y is bivariate normal $\pi(x, y) \propto$

$\exp\left(-\frac{1}{2}\begin{pmatrix} x-\theta_x & y-\theta_y \end{pmatrix} \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix} \begin{pmatrix} x-\theta_x \\ y-\theta_y \end{pmatrix}\right)$, where $\begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix} \equiv \Sigma^{-1}$ is the inverse of

the variance-covariance matrix for (x, y) . We are interested in the marginal distribution of z

$= x/y$. This distribution is presented in Hinkley (1969, equations (1) and (2), p. 636), who

attributes the result to Fieller (1932). The density (in our notation) is

$$\pi(z) = \frac{b(z)d(z)}{\sqrt{2\pi}\sigma_x\sigma_y a(z)^3} \left[\Phi\left\{\frac{b(z)}{(1-\rho^2)a(z)}\right\} - \Phi\left\{-\frac{b(z)}{(1-\rho^2)a(z)}\right\} \right] + \frac{\sqrt{(1-\rho^2)}}{\pi\sigma_x\sigma_y a(z)^2}$$

$$\times \exp\left\{\frac{c}{2(1-\rho^2)}\right\}, \text{ where } a(z) \equiv \left(\frac{z^2}{\sigma_x^2} - \frac{2\rho z}{\sigma_x\sigma_y} + \frac{1}{\sigma_y^2}\right)^{1/2}, b(z) \equiv \frac{\theta_x z}{\sigma_x^2} - \frac{\rho(\theta_x + \theta_y z)}{\sigma_x\sigma_y} + \frac{\theta_y}{\sigma_y^2}, c \equiv$$

$$\frac{\theta_x^2}{\sigma_x^2} - \frac{2\rho\theta_x\theta_y}{\sigma_x\sigma_y} + \frac{\theta_y^2}{\sigma_y^2}, d(z) \equiv \exp\left\{\frac{b(z)^2 - ca(z)^2}{2(1-\rho^2)a(z)^2}\right\}, \Phi\{r\} \equiv \int_{-\infty}^r \phi(u)du, \phi(u) \equiv \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}u^2},$$

and $\rho \equiv \frac{\sigma_{xy}}{\sigma_x \sigma_y}$. Due to its complexity, obtaining modal estimates of this density using, say,

the EM algorithm is infeasible. Fieller (1932) studies the existence of moments and offers relevant comment (p. 432, paragraph at the bottom of the page; p. 434, below equation (35); and p. 435, entire discussion) about the use of the normal approximation when the ratio θ_y/σ_y is ‘large’—a result he attributes to Geary. Hinkley (1969) studies the approximation and concludes that it is very accurate for moderate sample sizes under the restriction $0 < \sigma_y \leq \theta_y$. The conclusion that the density of z is proper is stated below equation (8). Other relevant works on the topic are Merrill (1928) and Marsaglia (1965). Thus, we have the results that the marginal density for z is proper; that, under certain conditions, its moments may not be finite; but that, when those conditions are met, the normal distribution provides a useful approximation to the true distribution. Turning, now, to the question of conditional expectations of z , consider the change of variables from (x, y) to (z, w) , where $z = g(x, y) \equiv x/y$ and $w = h(x, y) \equiv y$. We have, $y = h^{-1}(z, w) \equiv w$ and $x = g^{-1}(z, w) \equiv zw$, so the transformation is one-to-one. The Jacobian of the transformation is $|J| =$

$$\left| \begin{array}{cc} \frac{\partial g^{-1}(z, w)}{\partial z} & \frac{\partial g^{-1}(z, w)}{\partial w} \\ \frac{\partial h^{-1}(z, w)}{\partial z} & \frac{\partial h^{-1}(z, w)}{\partial w} \end{array} \right| = |w|, \text{ which is non-zero for } w \neq 0. \text{ Thus, the joint distribution of}$$

(z, w) under this latter condition is

$$\pi(z, w) \propto |w| \exp \left(-\frac{1}{2} (zw - \theta_x \quad w - \theta_y) \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix} \begin{pmatrix} zw - \theta_x \\ w - \theta_y \end{pmatrix} \right),$$

where “ \propto ” denotes “is proportional to.” The conditional distribution, $\pi(z | w)$, follows from multiplying out the quadratic form in the kernel

$$\begin{aligned} \pi(z | w) &\propto \exp \left(-1/2 \{ (zw - \theta_x)^2 \sigma_x^2 + 2(zw - \theta_x) \sigma_{xy} (w - \theta_y) + (w - \theta_y)^2 \sigma_y^2 \} \right) \\ &\propto \exp \left(-1/2 \{ (z^2 w^2 - 2zw\theta_x) \sigma_x^2 + 2zw\sigma_{xy} (w - \theta_y) \} \right) \end{aligned}$$

$$\propto \exp (-1/2 \{ (z^2 w^2 \sigma_x^2 + z (2 w (\sigma_{xy} (w - \theta_y) - \sigma_x^2 \theta_x)) \})$$

$$\propto \exp (-1/2 \{ (z^2 A + z B) \})$$

where $A \equiv w^2 \sigma_x^2$ and $B \equiv 2 w (\sigma_{xy} (w - \theta_y) - \sigma_x^2 \theta_x)$. Now, completing the square in z

$$\pi(z | w) \propto \exp (-A/2 \{ (z^2 + z B/A) \}),$$

$$\propto \exp (-A/2 \{ (z - C)^2 \}),$$

and it follows from properties of the normal distribution that the conditional distribution of z is normal with variance A and mean $C = B/(-2A)$ so that the conditional expectation of (the ratios of normals) $z = x/y$ is

$$E(z | w) = 2 w (\sigma_{xy} (w - \theta_y) - \sigma_x^2 \theta_x) / -2 w^2 \sigma_x^2$$

$$= w^{-1} (\theta_x - \sigma_x^{-2} \sigma_{xy} (w - \theta_y)).$$

It is this measure that corresponds to the expectation of the quotient in the right-side of text-equation (13).

REFERENCES

Fieller E C. 1932. The Distribution of the Index in a Normal Bivariate Population.

Biometrika **24**: 428-40.

Geary R C. 1930. The Frequency Distribution of the Quotient of Two Normal Variates.

Journal of the Royal Statistical Society **93**: 442-46.

Hinkley D V. 1969. On the Ratio of Two Correlated Normal Random Variables. Biometrika

56: 635-39.

Marsaglia G. 1965. Ratios of Normal Variables and Ratios of Sums of Uniform Variables.

Journal of the American Statistical Association **60**: 193-204.

Merrill A S. 1928. Frequency Distribution of an Index When Both the Components Follow

the Normal Law. Biometrika **20A**: 53-63.

MATLAB™ code:

%Notes to readers and practitioners:

%The MATLAB code that follows implements the estimation algorithm described
%in the Journal of Applied Econometrics paper entitled:
%"Bayes Estimates of Distance to Market: Transactions Costs, Cooperatives
%and Milk Market Development in the Ethiopian Highlands,"
%which is co-authored with Simeon Ehui and Amare Teklu.

%The code was written November 30, 2004.
%Annotations, designed to facilitate second-use by readers were added today
%September 12, 2006.

%Readers interested in the paper, the data, the primary-data collection
%survey instruments and, possibly, the broader objectives of the research
%project within which the data were obtained, are invited to contact me at:

%Garth Holloway
%Department of Agricultural and Food Economics,
%University of Reading,
%4 Earley Gate,
%Whiteknights Road,
%PO Box 237,
%Reading
%Berkshire
%RG66AR
%United Kingdom
%phone: (44) (118) 378 6775
%facsimile (44) (118) 975 6567
%email g.holloway@cgiar.org;

%Finally, the data used in the article are a subset of quantities sampled
%from the households, including demographic and spatial characteristics,
%farm and off-farm employment, animal and human consumption and animal
%health status of dairy and non-dairy animals. These data have been made
%available by the International Livestock Research Institute, Addis Ababa,
%and are available from the author upon request. Readers wishing to use
%the data in their own research are kindly requested to acknowledge the
%International Livestock Reserarch Institute. Thank you.

%The reminder of this program describes and implements the estimation
%algorithm layed out in the body of the journal article and described in
%detail in the accompanying appendix. Readers having difficulty
%interpreting code, or its representation in the text of the article, are
%encouraged to contact me.

%Garth Holloway
%September 12, 2006.

%Setup
clear
format bank
format compact
%pause

```

%Time
tic
%pause

```

```

%Fetch
data=load('distancetomarketdataascii');
production=data(:,2);
sales=data(:,12);
crossbreed=data(:,86);
localbreed=data(:,87);
extension=data(:,90);
distance=data(:,20);
education=data(:,24);
ilukura=data(:,32);
mirti=data(:,33);
%pause

```

```

%Define
a=[sales production];
b=[crossbreed localbreed extension distance education ilukura mirti];
n=68;
for i=1:n
    t(i,:)=3;
    cst(i,:)=sum(t);
    h(cst(i,:)-t(i,:)+1:cst(i,:),:)=ones(t(i,:),1)*i;
    w(cst(i,:)-t(i,:)+1:cst(i,:),i)=ones(t(i,:),1);
    y(cst(i,:)-t(i,:)+1:cst(i,:),:)=a([68 2*68 3*68]-68+i,[1 2]);
    c(cst(i,:)-t(i,:)+1:cst(i,:),:)=b([68 2*68 3*68]-68+i,:);
    x(cst(i,:)-t(i,:)+1:cst(i,:),:)=c(cst(i,:)-t(i,:)+1:cst(i,:),[1 2 3 4 5
6 7]);
    r(i,:)=c(3*(i-1)+1,[6 7]);
    i
end
[s k]=size(x);
f=c(:,[1 2 3 4 5]);
%pause

```

```

%Size
[s k]=size(x);
[s n]=size(w);
[s m]=size(y);
[s l]=size(f);
i=find(y(:,1)==0);
j=find(y(:,1)~=0);
ni=length(i);

```

```
nj=length(j);  
%pause
```

```
%Design  
v=reshape(y,s*m,1);  
p=[zeros(s,n);w];  
q=[x zeros(s,1);zeros(s,k) f];  
%pause
```

```
%Schedule  
burnin=5000;  
gibbs=10000;  
%pause
```

```
%Start  
z=y;  
bs=0;  
g=0;  
beta=inv(q'*q)*q'*v;  
gamma=inv(p'*p)*p'*v;  
delta=inv(r'*r)*r'*gamma;  
omega=sqrt((gamma-r*delta)'*(gamma-r*delta)/n);  
mu=reshape(p*gamma+q*beta,s,m);  
tau=0;  
%pause
```

```
%Simulate  
for h=1:burnin+gibbs
```

```
    %Inverted-Wishart  
    u=mvnrnd(zeros(m,1),inv((z-mu)'*(z-mu)),s)';  
    sigma=inv(u*u');  
    %pause
```

```
    %Multivariate-Normal  
    a=v-p*gamma;  
    b=q;
```

```

c=kron(inv(sigma),eye(s));
betacov=inv(b'*c*b);
betahat=betacov*(b'*c*a);
beta=mvnrnd(betahat,betacov,1)';
%pause

%Truncated-Normal
mu=reshape(p*gamma+q*beta,s,m);
mz=mu(i,1)+(z(i,2)-mu(i,2))*inv(sigma(2,2))*sigma(2,1);
sz=sqrt(sigma(1,1)-sigma(1,2)*inv(sigma(2,2))*sigma(2,1));
z(i,1)=norminv(unifrnd(0,1,ni,1).*normcdf(ones(ni,1)*tau,mz,sz),mz,sz);
%pause

%Uniform
tau=unifrnd(max(z(i,1)),min(z(j,1)));
%pause

%Truncated-Normal
a=v-q*beta;
b=p;
c=kron(inv(sigma),eye(s));
d=r*delta;
e=omega^-2*eye(n);
gammacov=inv(b'*c*b+e);
gammahat=gammacov*(b'*c*a+e*d);
mg=gammahat;
sg=sqrt(unique(diag(gammacov)));
gamma=norminv(unifrnd(0,1,n,1).*normcdf(0,mg,sg),mg,sg);
%pause

%Multivariate Normal
delta=mvnrnd(inv(r'*r)*r'*gamma,omega^2*inv(r'*r),1)';
%pause

%Scaled-Inverse Chi-Squared
omega=sqrt((gamma-r*delta)'*(gamma-r*delta)/sum(normrnd(0,1,n,1).^2));
%pause

```



```

%Collect
bs=bs+1;
if h>burnin
    bs=burnin;
    g=g+1;

%Store
sigmas(g,:)=reshape(sigma,m*m,1)';
betas(g,:)=beta';
distances(g,:)=z(i,1)';
taus(g,:)=tau';
gammas(g,:)=gamma';
omegas(g,:)=omega';
deltas(g,:)=delta';

%Likelihood
yj=y(j,:);
mu=reshape(p*gamma+q*beta,s,m);
muj=muj(j,:);
logunc=sum(-.5*log(det(sigma))- .5*m*log(2*pi)-.5*diag((yj-
muj)*inv(sigma)*(yj-muj)'));
m12i=mu(i,1)+(y(i,2)-mu(i,2))*inv(sigma(2,2))*sigma(2,1);
s12i=sqrt(sigma(1,1)-sigma(1,2)*inv(sigma(2,2))*sigma(2,1));
y2i=y(i,2);
m2i=mu(i,2);
s2i=sqrt(sigma(2,2));
logcens=sum(log(normcdf(tau,m12i,s12i))+log(normpdf(y2i,m2i,s2i)));
loglikes(g,:)=logunc+logcens;
if g==gibbs
    [maxlogl gstar]=max(loglikes);
    sigmatar=reshape(sigmas(gstar,:),m,m);
    betatar=betas(gstar)';
    taugar=taus(gstar);
    gammatar=gammas(gstar)';
    omegatar=omegas(gstar)';
    deltar=deltas(gstar)';
end
end

%Track
where=[bs g]
end
%pause

%Plot

```

```

figure(1)
plot(sigmas)
title('sigmas')
axis tight
figure(2)
plot(betas)
title('betas')
axis tight
figure(3)
plot(distances)
title('distances')
axis tight
figure(4)
plot(taus)
title('censoring point')
axis tight
figure(5)
plot(gammas)
title('gammas')
axis tight
figure(6)
plot(omegas)
title('omegas')
axis tight
figure(7)
plot(deltas)
title('deltas')
axis tight
%pause

%Look
sigmahat=[prctile(sigmas,5)' mean(sigmas)' prctile(sigmas,95)']
betahat=[prctile(betas,5)' mean(betas)' prctile(betas,95)']
tauhat=[prctile(taus,5)' mean(taus)' prctile(taus,95)']
gammahat=[prctile(gammas,5)' mean(gammas)' prctile(gammas,95)']
omegahat=[prctile(omegas,5)' mean(omegas)' prctile(omegas,95)']
deltahat=[prctile(deltas,5)' mean(deltas)' prctile(deltas,95)']
muhat=p*gammahat(:,2)+q*betahat(:,2);
yhat=reshape(muhat,s,m);
corrrun=corrcoef([y(j,:) yhat(j,:)]);
rsquaredun=[corrrun(1,3) corrrun(2,4)].^2
pospredun=[length(find(yhat(j,1)>mean(taus))) length(find(yhat(j,2)>0))]
negpredun=[length(find(yhat(j,1)<=mean(taus))) length(find(yhat(j,2)<=0))]
corrccen=corrcoef([mean(distances)' y(i,2) yhat(i,:)]);
rsquaredcen=[corrccen(1,3) corrccen(2,4)].^2
pospredcen=[length(find(yhat(i,1)>mean(taus))) length(find(yhat(i,2)>0))]
negpredcen=[length(find(yhat(i,1)<=mean(taus))) length(find(yhat(i,2)<=0))]
maxlogl
%pause

%Look
tsigmahat=[ mean(sigmas)' mean(sigmas)'./std(sigmas)']
tbetahat=[mean(betas)' mean(betas)'./std(betas)']

```

```
ttauhat=[mean(taus) mean(taus)'./std(taus)']
tgammahat=[mean(gammas) ' mean(gammas)'./std(gammas)'];
tomegahat=[mean(omegas) ' mean(omegas)'./std(omegas)']
tdeltaht=[mean(deltas) ' mean(deltas)'./std(deltas)']
maxlogl
cond(x)
cond(r)
%pause

%Time
time=toc
timeinhours=toc/60/60
projectedl0=timeinhours*10
%pause

%Store
save bayesdistancetomarketdata
```