Econ 203c, Spring 2003, Final Exam

Question 1

$y_i = x_i'\beta + \gamma z_i^* + u_i$
$z_{1i} = z_i^* + v_{1i}$
$z_{2i} = z_i^* + v_{2i}$
Assume: $E[u_i|x_i, z_i^*] = cov(u_i, v_{1i}) = cov(u_i, v_{2i}) = 0$ and $cov(v_{1i}, v_{2i}) \neq 0$

Note: There was a typo that was not corrected during the exam: $cov(v_{1i}, v_{2i}) = 0$. I will be very lenient in grading this question. The rest of the answers for this question assume $cov(v_{1i}, v_{2i}) = 0$.

1) Show LS for γ not consistent
Using z_{1i} and substituting

$y_i = x_i'\beta + \gamma(z_{1i} - v_{1i}) + u_i$
$= x_i'\beta + \gamma z_{1i} + (u_i - \gamma v_{1i})$
$= x_i'\beta + \gamma z_{1i} + \varepsilon_i$

Prove consistency:
Let $W = [x\ z_1]$

$$\begin{bmatrix} \widehat{\beta} \\ \widehat{\gamma} \end{bmatrix} = \begin{bmatrix} \beta \\ \gamma \end{bmatrix} + (W'W)^{-1}W'\varepsilon$$

$$p\lim\left\{\begin{bmatrix} \widehat{\beta} \\ \widehat{\gamma} \end{bmatrix} - \begin{bmatrix} \beta \\ \gamma \end{bmatrix}\right\} = p\lim(1/nW'W)^{-1}p\lim 1/nW'\varepsilon$$

Assume $p\lim(1/nW'W)^{-1}$ is finite.

$$p\lim 1/nW'\varepsilon = p\lim 1/n\sum W_i\varepsilon_i = \lim_{n\to\infty} 1/n\sum E[W_i\varepsilon_i] = \lim_{n\to\infty}\begin{bmatrix} 1/n\sum E[x_i\varepsilon_i] \\ 1/n\sum E[z_{1i}\varepsilon_i] \end{bmatrix}$$

$E[x_i\varepsilon_i] = E[x_i(u_i - \gamma v_{1i})] = E[x_iu_i] - \gamma E[x_iv_{1i}]$

$E[x_iu_i] = 0$ by assumption.
Although it is not stated in the problem, we could make a further assumption that

$E[x_iv_{1i}] = 0$.
With these two assumptions, $E[x_i\varepsilon_i] = 0$.
$E[z_{1i}\varepsilon_i] = E[(z_i^* + v_{1i})(u_i - \gamma v_{1i})]$
$= E[z_i^*u_i] + E[v_{1i}u_i] - \gamma E[v_{1i}v_{1i}] - \gamma E[z_i^*v_{1i}]$

$= -\gamma E[v_{1i}v_{1i}] - \gamma E[z_i^* v_{1i}]$

This does not in general equal zero.

Therefore, $p\lim 1/nW'\varepsilon \neq 0$ and $p\lim\begin{bmatrix} \widehat{\beta} \\ \widehat{\gamma} \end{bmatrix} \neq \begin{bmatrix} \beta \\ \gamma \end{bmatrix}$.

2) A consistent estimator for γ could be found by using z_{2i} to instrument for z_{1i}.

Let $W = [x\ z_1]$ and $G = [x\ z_2]$.

Following the same steps as above

$$\begin{bmatrix} \widehat{\beta} \\ \widehat{\gamma} \end{bmatrix} = \begin{bmatrix} \beta \\ \gamma \end{bmatrix} + (G'W)^{-1}G'\varepsilon$$

$$p\lim\{\begin{bmatrix} \widehat{\beta} \\ \widehat{\gamma} \end{bmatrix} - \begin{bmatrix} \beta \\ \gamma \end{bmatrix}\} = p\lim(1/nG'W)^{-1}p\lim 1/nW'\varepsilon$$

Assume $p\lim(1/nG'W)^{-1}$ is finite.

$$p\lim 1/nG'\varepsilon = p\lim 1/n\sum G_i\varepsilon_i = \lim_{n\to\infty} 1/n\sum E[G_i\varepsilon_i] = \lim_{n\to\infty}\begin{bmatrix} 1/n\sum E[x_i\varepsilon_i] \\ 1/n\sum E[z_{2i}\varepsilon_i] \end{bmatrix}$$

$$E[x_i\varepsilon_i] = E[x_i(u_i - \gamma v_{1i})] = E[x_iu_i] - \gamma E[x_iv_{1i}]$$

Here we need the assumption that $E[x_iv_{1i}] = 0$.

Therefore, $E[x_i\varepsilon_i] = 0$.

$$E[z_{2i}\varepsilon_i] = E[(z_i^* + v_{2i})(u_i - \gamma v_{1i})]$$

$$= E[z_i^*u_i] + E[v_{2i}u_i] - \gamma E[v_{2i}v_{1i}] - \gamma E[z_i^*v_{1i}]$$

For this to be zero, we need to assume each of these four components is zero.

With these assumptions, $p\lim 1/nG'\varepsilon = 0$ and $p\lim\begin{bmatrix} \widehat{\beta} \\ \widehat{\gamma} \end{bmatrix} = \begin{bmatrix} \beta \\ \gamma \end{bmatrix}$.

3) If $cov(v_{2i}, u_i) \neq 0$, then the estimator in 2) will not be in general consistent.

4) If $E[u_i|x_i] \neq 0$, the estimator in 2) will not be consistent. Any endogeneity in a model will contaminate all of the coefficient estimates.

5) Given that x_i is now assumed to endogenous, we would need an instrument for it. We could use one of the z_is as an instrument but there are not enough of them to both instrument for x_i and for the other z_i. Hence without more data, there is no

consistent estimator for γ.

6) With the assumption that $E[u_i|z_i^*] \neq 0$, we would need an instrument for z_i^*. Since both z_{i1} and z_{i2} would be endogenous, neither are legitimate instruments. Again without more data (more instruments), there is no consistent estimator for γ.

Question 2

$y_i^* = z_i'\gamma_0 + v_i$

1) This is an ordered multinomial choice model.
$L(\gamma_0, \mu_1, \mu_2) = \Pi_{i=1}^n [\Pi_{j=1}^3 pr(y_i = j|z_i)^{z_{ji}}]$,
where
$z_{ji} = 1$ if i chooses choice j, 0 otherwise

and
$pr(y_i = 1|z_i) = pr(y_i^* \leq \mu_1|z_i)$
$= pr(z_i'\gamma_0 + v_i \leq \mu_1|z_i)$
$= pr(v_i \leq \mu_1 - z_i'\gamma_0|z_i)$,

$pr(y_i = 2|z_i) = pr(\mu_1 < y_i^* \leq \mu_2|z_i)$
$= pr(\mu_1 < z_i'\gamma_0 + v_i < \mu_2)$
$= pr(v_i < \mu_2 - z_i'\gamma_0|z_i) - pr(v_i \leq \mu_1 - z_i'\gamma_0|z_i)$,
and
$pr(y_i = 3|z_i) = pr(y_i^* > \mu_2|z_i)$
$= pr(z_i'\gamma_0 + v_i > \mu_2|z_i)$
$= pr(v_i > \mu_2 - z_i'\gamma_0|z_i)$

2) Assume $v_i|z_i \sim i.i.d. N(0, \sigma_v^2)$

First normalize $\sigma_v = 1$.
The likelihood function is
$L(\gamma_0, \mu_1, \mu_2) = \Pi_{i=1}^n [\Pi_{j=1}^3 pr(y_i = j|z_i)^{z_{ji}}]$,

where $pr(y_i = 1|z_i) = \Phi(\mu_1 - z_i'\gamma_0)$, $pr(y_i = 2|z_i) = \Phi(\mu_2 - z_i'\gamma_0) - \Phi(\mu_1 - z_i'\gamma_0)$, $pr(y_i = 3|z_i) = 1 - \Phi(\mu_2 - z_i'\gamma_0)$, and $\Phi(\bullet)$ is the standard normal CDF.

Completely written out, the likelihood function is
$L(\gamma_0,\mu_1,\mu_2) = \Pi_{i=1}^{n}\{\Phi(\mu_1 - z_i'\gamma_0)^{z_{1i}}[\Phi(\mu_2 - z_i'\gamma_0) - \Phi(\mu_1 - z_i'\gamma_0)]^{z_{2i}}[1 - \Phi(\mu_2 - z_i'\gamma_0)]^{z_{3i}}\}$

In this model we are estimating γ_0/σ_v, i.e. γ_0 up to some unknown scale factor.

3) $\widehat{\gamma}_n = \arg\max L(\gamma_0,\mu_1,\mu_2)$

Show $\widehat{\gamma}_n$ is a consistent estimator for γ_0. This proof is a direct application of the consistency of MLE proof from Lecture Note 8. A very brief sketch of the proof is all that is needed.

4) Testing procedure

$$H_0 : \sum_{k=1}^{K}(\gamma_{0k})^2 - K = 0$$
$$H_1 : \sum_{k=1}^{K}(\gamma_{0k})^2 - K \neq 0$$

Testing procedure
First, construct a test statistic
Wald test

$$W_n = \frac{n\left[\sum_{k=1}^{K}(\widehat{\gamma_{0k}})^2 - K\right]^2}{est.cov\left[\sum_{k=1}^{K}(\widehat{\gamma_{0k}})^2 - K\right]},$$

where $est.cov\left[\sum_{k=1}^{K}(\widehat{\gamma_{0k}})^2 - K\right] = \Delta cov(\widehat{\gamma_0})\Delta'$, $\Delta = \left[\begin{array}{ccc} 2\widehat{\gamma_{01}} & \ldots & 2\widehat{\gamma_{0K}} \end{array}\right]$, and

$$cov(\widehat{\gamma_0}) = \left[-1/n\sum_{i=1}^{n}\frac{\partial^2 \ln L_i(\widehat{\gamma_0},\widehat{\mu_1},\widehat{\mu_2})}{\partial\gamma_0^2}\right]^{-1}$$

Next, derive the asymptotic distribution of the test statistic under the null hypothesis
$W_n \sim \chi^2(1)$

Finally, choose a confidence level and test the hypothesis.
At a confidence level of 95 percent,
Reject null hypothesis if $W_n > 3.84$
Fail to reject null hypothesis if $W_n \leq 3.84$.

Question 3
$y_i = \beta_1 x_i^{\beta_2} + u_i$, $E[u_i|x_i] = 0$

1) Two moment functions: There are many possibilities.
The most obvious are based on the FOC from the Non-Linear Least Squares criteria.
$\varphi_{1i}(y,x,\beta) = (y_i - \beta_1 x_i^{\beta_2})x_i^{\beta_2}$

$\varphi_{2i}(y,x,\beta) = (y_i - \beta_1 x_i^{\beta_2})\beta_1 x_i^{\beta_2}\ln x_i$
{The last result is from the following:

$\frac{\partial}{\partial\beta}a^{\beta} = \frac{\partial}{\partial\beta}\exp(\beta\ln a) = \exp(\beta\ln a)\ln a = a^{\beta}\ln a$.}

The MM estimator is $\widehat{\beta}_n = \arg\min \varphi(y,x,\beta)'\varphi(y,x,\beta)$,
where $\varphi(y,x,\beta) = 1/n\sum_{i=1}^{n}\begin{bmatrix} (y_i - \beta_1 x_i^{\beta_2})x_i^{\beta_2} \\ (y_i - \beta_1 x_i^{\beta_2})\beta_1 x_i^{\beta_2}\ln x_i \end{bmatrix}$

If this system is full rank, we can estimate a unique $\widehat{\beta}_n$.

2) Asymptotic distribution
$\sqrt{n}\,(\widehat{\beta}_n - \beta) \xrightarrow{d} N(0, \Lambda(\beta_0))$,

where $\Lambda(\beta_0) = (A(\beta_0)A(\beta_0)')^{-1}A(\beta_0)W(\beta_0)A(\beta_0)'(A(\beta_0)A(\beta_0)')^{-1}$,
$A(\beta_0) = E[\frac{\partial}{\partial\beta}\varphi(y,x,\beta_0)]$, and $W(\beta_0) = E[\varphi(y,x,\beta_0)\varphi(y,x,\beta_0)']$

3) Assume $E[u_i|x_i] \neq 0$, $E[u_i|z_i] = 0$, and $E[z_i x_i] \neq 0$.
New MM estimator
$\widehat{\beta}_n = \arg\min \varphi(y,x,z,\beta)'V_n^{-1}\varphi(y,x,z,\beta)$,

where $\varphi(y,x,z,\beta) = 1/n \sum_{i=1}^{n} \begin{bmatrix} (y_i - \beta_1 x_i^{\beta_2}) z_i^{\beta_2} \\ (y_i - \beta_1 x_i^{\beta_2}) \beta_2 \beta_1 z_i^{\beta_2 - 1} \end{bmatrix}$, and V_n^{-1} is some weight matrix.

Another possible set of moments is simply
$\varphi_i(y,x,z,\beta) = (y_i - \beta_1 x_i^{\beta_2}) z_i$

4) $\widehat{x}_i = z_i' \widehat{\alpha}_{ols}$, where $\widehat{\alpha}_{ols}$ is the OLS estimate of x_i on z_i.
$\widehat{\alpha}_{ols} = (z'z)^{-1} z'x$

Using this predicted $\widehat{x}_i$, the moment condition is
$\varphi_i(y,x,z) = (y_i - \beta_1 \widehat{x}_i^{\beta_2}) z_i$
$= (y_i - \beta_1 x_i^{\beta_2} + \beta_1 x_i^{\beta_2} - \beta_1 \widehat{x}_i^{\beta_2}) z_i$
$= (y_i - \beta_1 x_i^{\beta_2}) z_i + (\beta_1 x_i^{\beta_2} - \beta_1 \widehat{x}_i^{\beta_2}) z_i$
$= u_i z_i + (\beta_1 x_i^{\beta_2} - \beta_1 \widehat{x}_i^{\beta_2}) z_i$

$E[\varphi_i(y,x,z)] \;= E[u_i z_i] + E[(\beta_1 x_i^{\beta_2} - \beta_1 \widehat{x}_i^{\beta_2}) z_i]$
$E[u_i z_i] = 0$ by assumption.
$E[(\beta_1 x_i^{\beta_2} - \beta_1 \widehat{x}_i^{\beta_2}) z_i] = \beta_1 E[(x_i^{\beta_2} - (z_i' \widehat{\alpha}_{ols})^{\beta_2}) z_i]$

By Law of Iterative Expecations,
$E[(x_i^{\beta_2} - (z_i' \widehat{\alpha}_{ols})^{\beta_2}) z_i] = E_z\{E(x_i^{\beta_2}|z) - E((z_i' \widehat{\alpha}_{ols})^{\beta_2}|z)] z_i\}$
$E(x_i^{\beta_2}|z) \neq E((z_i' \widehat{\alpha}_{ols})^{\beta_2}|z)$ in general
Thus, $E[(\beta_1 x_i^{\beta_2} - \beta_1 \widehat{x}_i^{\beta_2}) z_i] \neq 0$ in general.
This implies that this moment condition is not satisfied for this non-linear model.

To see why this non-linear model breaks down, consider the following linear case where $y_i = \beta_1 x_i + u_i$
$E[\varphi_i(y,x,z)] \;= E[u_i z_i] + E[(\beta_1 x_i - \beta_1 \widehat{x}_i) z_i]$
$E[(\beta_1 x_i - \beta_1 \widehat{x}_i) z_i] = \beta_1 E[(x_i - (z_i' \widehat{\alpha}_{ols})) z_i]$

Note that $(x_i - (z_i' \widehat{\alpha}_{ols}))$ is the residual for observation i.
By Law of Iterated Expectations,
$E[(x_i - (z_i' \widehat{\alpha}_{ols})) z_i] = E_z\{E[(x_i - z_i' \widehat{\alpha}_{ols})|z] z_i\}$.
$E[(x_i - (z_i' \widehat{\alpha}_{ols})|z] = E[x_i|z] - z_i' E[\widehat{\alpha}_{ols}|z]$.
A basic OLS result is that $E[\widehat{\alpha}_{ols}|z] = \alpha$ (i.e. OLS estimates are unbiased.)
[proof: $E[\widehat{\alpha}_{ols}|z] = \alpha + (z'z)^{-1} z' E[\varepsilon|z]$, $E[\varepsilon|z] = 0$ by assumption.]
And since the OLS assumption is $E[x_i|z] = z_i' \alpha$,
this implies $E[x_i|z] - z_i' E[\widehat{\alpha}_{ols}|z] = 0$.

Combining all the results from above,
$E[\varphi(y,x,z,\beta)] \; = 0$ in the linear case.

5) Testing procedure
$H_0 \,:\, \beta_1 - \exp(\beta_2) = 0$
$H_1 \,:\, \beta_1 - \exp(\beta_2) \neq 0$

Wald test
$$W_n = \frac{n[\beta_1 - \exp(\beta_2)]^2}{est.cov(\beta_1 - \exp(\beta_2))},$$

where $est.cov(\beta_1 - \exp(\beta_2)) \; = \Delta cov(\widehat{\beta}_n)\Delta'$, $\Delta = [1 \; -\exp(\beta_2)]$,
and $cov(\widehat{\beta}_n) = (A(\widehat{\beta}_n)V(\widehat{\beta}_n)A(\widehat{\beta}_n)')^{-1}$ (the finite sample estimator of the asymptotic covariance matrix defined above)

Under null hypothesis, $W_n \sim \chi^2(1)$
Choose a confidence level and test the hypothesis.

Confidence level 95 percent.
Reject null hypothesis if $W_n > 3.84$

Fail to reject null hypothesis if $W_n \leq 3.84$.

Question 4

$y_i^* = x_i'\beta + \varepsilon_i$, where $E[\varepsilon_i|x_i] = 0$
Let $y_i = y_i^*$ if $y_i < y^0$ and $y_i = y^0$ if $y_i^* \geq y^0$ (Notice: typo in original)
1) Compute $E[y_i|x_i, y_i^* < y^0]$

This calculation is the same as for the Tobit model with the censoring or cutoff point moved from zero to y^0.

$E[y_i|x_i, y_i^* < y^0] = E[x_i'\beta + \varepsilon_i | x_i'\beta + \varepsilon_i < y^0, x_i]$
$= x_i'\beta + \sigma_\varepsilon E[\frac{\varepsilon_i}{\sigma_\varepsilon} | \frac{\varepsilon_i}{\sigma_\varepsilon} < \frac{y^0 - x_i'\beta}{\sigma_\varepsilon}, x_i]$, where σ_ε is a constant and $\sigma_\varepsilon > 0$

$$= x_i'\beta + \sigma_\varepsilon \frac{\int_{-\infty}^{\frac{y^0 - x_i'\beta}{\sigma_\varepsilon}} \frac{\varepsilon_i}{\sigma_\varepsilon} f(\frac{\varepsilon_i}{\sigma_\varepsilon}) d(\frac{\varepsilon_i}{\sigma_\varepsilon})}{F(\frac{y^0 - x_i'\beta}{\sigma_\varepsilon})}, \qquad (4.1)$$

where $f(\bullet)$ is the PDF and $F(\bullet)$ is the CDF for $\frac{\varepsilon_i}{\sigma_\varepsilon}|x_i$.

At this point, an assumption about the distribution of $\frac{\varepsilon_i}{\sigma_\varepsilon}|x_i$ is needed.

2) Two step estimator

First assume $\frac{\varepsilon_i}{\sigma_\varepsilon}|x_i \sim N(0,1)$.

With this assumption we can complete the computation of $E[y_i|x_i, y_i^* < y^0]$.

The numerator from (4.1) is

$$\int_{-\infty}^{\frac{y^0 - x_i'\beta}{\sigma_\varepsilon}} \frac{\varepsilon_i}{\sigma_\varepsilon} f(\frac{\varepsilon_i}{\sigma_\varepsilon}) d(\frac{\varepsilon_i}{\sigma_\varepsilon}) = \int_{-\infty}^{\frac{y^0 - x_i'\beta}{\sigma_\varepsilon}} z \frac{1}{(2\pi)^{1/2}} e^{-z^2/2} dz.$$

Using $\frac{d}{dz} e^{-z^2/2} = -ze^{-z^2/2}$, yields

$$= -\frac{1}{(2\pi)^{1/2}} e^{-z^2/2}\Big|_{-\infty}^{\frac{y^0 - x_i'\beta}{\sigma_\varepsilon}} = -\frac{1}{(2\pi)^{1/2}} e^{(\frac{y^0 - x_i'\beta}{\sigma_\varepsilon})^2/2}$$

$= -\phi(\frac{y^0 - x_i'\beta}{\sigma_\varepsilon})$, where $\phi(\bullet)$ is the standard normal PDF

This is the negative of the case in the lecture notes.

Combining these results

$E[y_i|x_i, y_i^* < y^0] = x_i'\beta - \sigma_\varepsilon \lambda(\frac{y^0 - x_i'\beta}{\sigma_\varepsilon})$, where $\lambda(a) = \phi(a)/\Phi(a)$ and $\phi(\bullet)$ is the standard normal PDF and $\Phi(\bullet)$ is the standard normal CDF.

Estimation

Step 1: Compute a probit model for the probability that a firm's expenditures are top coded.

The likelihood for the probit model is

$L(y, x, z, \beta, \sigma_\varepsilon) = \Pi_{i=1}^n \Phi(\frac{y^0 - x_i'\beta}{\sigma_\varepsilon})^{z_i}[1 - \Phi(\frac{y^0 - x_i'\beta}{\sigma_\varepsilon})]^{1-z_i}$,

where $z_i = 1$ if if $y_i < y^0$, 0 otherwise.

Note that not all of the parameters of this model are identified.

We can re-write $\frac{y^0 - x_i'\beta}{\sigma_\varepsilon}$.

$$\frac{y^0 - x_i'\beta}{\sigma_\varepsilon} = \frac{y^0}{\sigma_\varepsilon} - \frac{\beta_1}{\sigma_\varepsilon} - \frac{\beta_2}{\sigma_\varepsilon} x_{2i} - \ldots - \frac{\beta_K}{\sigma_\varepsilon} x_{Ki}$$

$= \gamma_1 + \gamma_2 x_{2i} + \ldots + \gamma_K x_{Ki}$, where $\gamma_1 = \frac{y^0}{\sigma_\varepsilon} - \frac{\beta_1}{\sigma_\varepsilon}$, $\gamma_2 = -\frac{\beta_2}{\sigma_\varepsilon}, \ldots, \gamma_K = -\frac{\beta_K}{\sigma_\varepsilon}$

$= x_i'\gamma$

With this substitution, the probit likelihood is
$L(y,x,z,\gamma) = \Pi_{i=1}^{n}\Phi(x_i'\gamma)^{z_i}[1-\Phi(x_i'\gamma)]^{1-z_i}$
$\widehat{\gamma} = \arg\max L(y,x,z,\gamma)$

Step 2: Use the estimate from Step 1 to compute $\widehat{\lambda}_i(x_i'\widehat{\gamma})$ and estimate an OLS regression of observed y_i on x_i and $\widehat{\lambda}_i(x_i'\widehat{\gamma})$. [Note: I put an i subscript on the inverse Mill's ratio simply to indicate that it varies with the xs.]

With this second step we obtain an estimate for $\widehat{\sigma_\varepsilon}$ as the coefficient estimate on $\widehat{\lambda}_i(x_i'\widehat{\gamma})$. Therefore the estimate of β is
$\widehat{\beta}_k = -\widehat{\gamma}_k\widehat{\sigma_\varepsilon}$ for $k = 2,\ldots,K$ and $\widehat{\beta}_1 = y^0 - \widehat{\gamma}_1\widehat{\sigma_\varepsilon}$

3) If each firm had a different, but known, topcode, this estimation method would still apply. This is only a trivial difference. Note that if the topcode was unknown, we could not do this.

4) Two step estimator as GMM estimator
The moments for this two step estimator consist of the score function for the first

step probit and the OLS moments from the second step.

First step probit
$l(\gamma,x,z) = \sum z_i\ln\Phi(x_i'\gamma) + (1-z_i)\ln(1-\Phi(x_i'\gamma))$
$$\varphi_1(\gamma,x,z) = \frac{\partial l(\gamma,x,z)}{\partial\gamma} = 1/n\sum\{\frac{z_i}{\Phi(x_i'\gamma)}\phi(x_i'\gamma)x_i - \frac{1-z_i}{1-\Phi(x_i'\gamma)}\phi(x_i'\gamma)x_i\}$$

Second step OLS
$$\varphi_2(\sigma_\varepsilon,\beta,x,y) = 1/n\sum[y_i - (x_i'\beta - \sigma_\varepsilon\lambda_i(x_i'\gamma))]\begin{bmatrix} x_i \\ \lambda_i(x_i'\gamma) \end{bmatrix}$$

GMM estimator

$$(\widehat{\gamma},\widehat{\beta},\widehat{\sigma_\varepsilon}) = \arg\min \begin{bmatrix} \varphi_1(\gamma,x,z) \\ \varphi_2(\sigma_\varepsilon,\beta,x,y) \end{bmatrix}' \begin{bmatrix} \varphi_1(\gamma,x,z) \\ \varphi_2(\sigma_\varepsilon,\beta,x,y) \end{bmatrix}$$

5) The two step GMM estimator in 4) is just-identified. We could find a better (more efficient) estimator by adding additional moments. For example,

$$(\widehat{\gamma},\widehat{\beta},\widehat{\sigma_\varepsilon}) = \arg\min \begin{bmatrix} \varphi_1(\gamma,x,z) \\ \varphi_2(\sigma_\varepsilon,\beta,x,y) \\ \varphi_3(\sigma_\varepsilon,\beta,\gamma,x,y) \end{bmatrix}' V_n^{-1} \begin{bmatrix} \varphi_1(\gamma,x,z) \\ \varphi_2(\sigma_\varepsilon,\beta,x,y) \\ \varphi_3(\sigma_\varepsilon,\beta,\gamma,x,y) \end{bmatrix}',$$

where $\varphi_3(\sigma_\varepsilon,\beta,\gamma,x,y)$ is an additional set of moments and V_n^{-1} is some weight matrix. Note that now that the system is over-identified, we must include a weight matrix to obtain a unique estimate.

Question 5

1) $L(p|x_1,\ldots,x_n) = \Pi_{i=1}^n p^{x_i}(1-p)^{1-x_i}$

2) $\ln L(p|x_1,\ldots,x_n) = \ln p \sum x_i + \ln(1-p)\sum(1-x_i)$

FOC

$$\frac{\partial \ln L(p|x_1,\ldots,x_n)}{\partial p} = 0$$

$$\Rightarrow \frac{\sum x_i}{p} = \frac{1}{1-p}\sum(1-x_i)$$

$$\Rightarrow \sum x_i - p\sum x_i = pn - p\sum x_i$$

$$\Rightarrow \widehat{p} = 1/n\sum x_i$$

3) Asymptotic distribution for p

Using the MLE results

$$\sqrt{n}\,(\widehat{p}-p) \xrightarrow{d} N(0,I(p)^{-1}),$$

where $I(p) = -E[\frac{\partial^2 \ln L(p|x)}{\partial p^2}]$

$$\frac{\partial^2 \ln L(p|x)}{\partial p^2} = -\frac{x}{p^2} - \frac{1-x}{(1-p)^2}$$

$$-E[\frac{\partial^2 \ln L(p|x)}{\partial p^2}] = \frac{E[x]}{p^2} + \frac{1-E[x]}{(1-p)^2}$$

$$E[x] = p$$

$\Rightarrow I(p) = 1/p + \dfrac{1}{(1-p)}$

$\Rightarrow I(p) = \dfrac{1}{p(1-p)}$

Thus, $\sqrt{n}\,(\widehat{p} - p) \overset{d}{\to} N(0, p(1-p))$

4) MLE for $p(1-p)$

Let $\alpha = p(1-p)$

By Invariance Property of MLE, $\widehat{\alpha} = \widehat{p}(1-\widehat{p})$

Asymptotic distribution

$\sqrt{n}\,(\widehat{\alpha} - \alpha) \overset{d}{\to} N(0, \Lambda(p))$,

where $\Lambda(p) = \Delta I(p)^{-1}\Delta'$ by Delta Method and $\Delta = \dfrac{\partial p(1-p)}{\partial p} = 1 - 2p$.

Thus, $\Lambda(p) = (1-2p)^2 p(1-p)$

5) Consistent estimator for $\Lambda(p)$

$\widehat{\Lambda(p)} = (1-2\widehat{p})^2\widehat{p}(1-\widehat{p})$

An application of the Slutsky theorems is sufficient to show $p\lim\widehat{\Lambda(p)} = \Lambda(p)$.