Answer key for the 2003 final exam (econ 220b).

1. (a) Since we assume that $C_i^* = \beta Y_i^*$, we can substitute for C^* and Y^* to get

$$C_i - a_{ci} = \beta(Y_i - a_{yi})$$

and then $C_i = Y_i + (a_{ci} - \beta a_{yi})$. So

$$\epsilon_i = a_{ci} - \beta a_{yi}.$$

(b) The plim of an estimator b_T is the constant b (if it exists) such that, for all $\delta > 0$,

$$\text{Prob}[|b_T - b| > \delta] \to 0 \qquad \text{as } T \to \infty.$$

For the OLS estimator,

$$\begin{aligned} b_T &= \left(\sum Y_i^2\right)^{-1} \sum Y_i C_i \\ &= 1/T \sum (Y_i^* + a_{yi})(C_i^* + a_{ci}) \Big/ 1/T \sum (Y_i^* + a_{yi})^2 \\ &\xrightarrow{p} E\big[(Y_i^* + a_{yi})(C_i^* + a_{ci})\big] \big/ E\big[(Y_i^* + a_{yi})^2\big] \\ &= \beta E Y_i^{*2} \big/ \big[E Y_i^{*2} + E a_{yi}^2\big]. \end{aligned}$$

Since the expected value of the OLS estimator is the product of β and a positive fraction less than one, the OLS estimator has a downward bias.

(c) It is clear from part (a) that the error term is correlated with the regressor Y_i.

(d) i. $EY_i > 0$ (it's income).

ii. $E(\epsilon_i) = E(a_{ci} - \beta a_{yi}) = 0$.

(e) Regressing C_i and Y_i on a vector of ones gives $\bar{C}$ and $\bar{Y}$ respectively. So the two-stage least squares estimator is $(\bar{Y})^{-2}\bar{C}\bar{Y}$ which simplifies to $\bar{C}/\bar{Y}$.

(f)

$$\begin{aligned}\hat{\beta}_{IV} &\xrightarrow{p} EC_i/EY_i \\ &= E(C_i^* + a_{ci})/E(Y_i^* + a_{yi}) \\ &= \beta EY_i/EY_i\end{aligned}$$

2. (a) $h(\theta, w_t) = x_t\epsilon_t = x_t(y_t - x_t'\beta)$ and so $g(\theta, Y_T) = 1/T(X'Y - X'X\beta)$. The GMM estimator is given by

$$\hat{\beta} = (X'X)^{-1}X'Y.$$

(b) The OLS, 2SLS and GMM estimators are all the same. The GLS estimator will have a different formula reflecting the serial correlation in the sequence $\{x_t\epsilon_t\}$.

(c) We know that

$$S = \sigma^2 \lim T^{-1} \sum_{t=1}^{T} \sum_{v=-\infty}^{\infty} \rho^v E\, x_t x_{t-v}' \qquad \hat{D}' = -X'X/T$$

and so

$$\hat{V} = T\hat{\sigma}^2(X'X)^{-1}\Big[\sum_t \sum_v \rho^v x_t x_{t-v}'\Big](X'X)^{-1}$$

(d) The appropriate test statistic is

$$T(R\hat{\beta} - r)'\big(R\hat{V}R'\big)^{-1})(R\hat{\beta} - r)$$

and it has a chi-square distribution with degrees of freedom equal to the dimension of r under the null hypothesis. This is a scaled version of the usual F-test, but with a different weighting matrix that reflects the serial correlation in the model's innovations.

(e) The Wald statistic is similar—we use the fact that the function λ is approximately linear at the limit of our estimator. Let $R(\cdot)$ represent the $m \times k$ matrix of derivatives of $\lambda(\cdot)$. Then our test statistic has the same distribution under the null as in the previous question, and is given by the formula

$$T\lambda(\hat{\beta})'\big[R(\hat{\beta})\hat{V}R(\hat{\beta})\big]^{-1}\lambda(\hat{\beta}).$$

3. (a) $\mathcal{L}(\lambda, y_t) = T\log\lambda - \lambda\sum y_t$.

   (b) The first order conditions from maximizing the log likelihood give us

   $$0 = T/\lambda - \sum y_t$$

   and so $\hat{\lambda} = 1/\bar{y}$.

   (c) $h(\lambda, y_t) = 1/\lambda - y_t$.

   (d) Straightforward substitution gives

   $$S = Eh(\lambda, y_t)^2 \qquad \hat{D}' = -\hat{\lambda}^{-2}.$$
   $$= \lambda^{-2}.$$

   (e) Using the plig-in estimator of S, we find $\hat{V} = \hat{\lambda}^2$, implying that

   $$\sqrt{T}(\hat{\lambda} - \lambda_0) \xrightarrow{L} N(0, \lambda_0^2).$$