Econ 311: Labor Supply and the Two-Step Estimator
Revised Version

One period models

\[
\max U(c, l) = \left( \frac{c^\alpha - 1}{\alpha} \right) + b \left( \frac{l^\phi - 1}{\phi} \right)
\]

st. \( c + wL \leq wT + A \).

The Euler equation is \( w = \frac{bL^{\phi-1}}{c^{\alpha-1}} \) and the reservation wage is given by
\[
wr = \left[ \frac{bl^{\phi-1}}{c^{\alpha-1}} \right]_{l=1,c=A} = \frac{b}{A^{\alpha-1}} \rightarrow \ln wr
\]

\[= \ln b + (1 - \alpha) \ln A\]

Assume \(\ln b = x\beta + e, \ e \perp \perp (x,A,w), \ e \sim N(0, \sigma_e^2)\). Assume wages observed for everyone
\text{Pr (Person works}|X, A)\\
= \text{Pr} (\ln w_r \leq \ln w|x, A)\\
= \text{Pr} \left( \frac{e}{\sigma_e} \leq \frac{\ln w - x\beta - (1 - \alpha) \ln A}{\sigma_e} \right)\\
= \Phi(c)\\
\text{where } c = \frac{\ln w - x\beta - (1 - \alpha) \ln A}{\sigma_e}.

\textbf{Grouped data estimator}

Each cell has common values of \\
\text{ }\\
w_i, x_i, A_i. \text{ For each cell, obtain } \hat{P}_i(D_i = \\
1|w_i, x_i, A_i) = \text{ cell proportion working}
= \Phi(\hat{c}_i). Calculate \( \hat{c}_i = \Phi^{-1}(\hat{P}_i) \). (Instead of standard normal (\( \Phi \)) could use standard logistic \( \Lambda(c) = \frac{e^c}{1 + e^c} \)) or linear probability model

\( F(c) = \frac{c}{c_u - c_l}, c_l \leq \frac{e}{\sigma_e} \leq c_u \).)

Regress \( \hat{c}_i \) on \( \frac{\ln w - x\beta - (1 - \alpha) \ln A}{\sigma_e} \)

and obtain estimates of \( \sigma_e, \beta, \alpha \).
Justifying the grouped data estimator. Suppose

\[ D_i = 1 \] if agent i works, 0 if not. For each cell,

get \( \hat{p}(D = 1|w, A). \) By WLLN and Slutsky,

\[
\text{plim}\Phi^{-1}(\hat{p}(D = 1|w, A)) = \Phi^{-1}(\text{plim}\hat{p}(D = 1|w, A))
\]

\[ = \Phi^{-1}(p(D = 1|w, A)) \]

\[ = \frac{\ln w - x\beta - (1 - \alpha) \ln A}{\sigma_e}. \]

Set up the regression function
\[ \Phi^{-1}(\hat{p}) = \Phi^{-1}(p) + [\Phi^{-1}(\hat{p}) - \Phi^{-1}(p)] \]
\[ = \frac{\ln w - x\beta - (1 - \alpha) \ln A}{\sigma_e} + \nu \]

We need to characterize \( V = \Phi^{-1}(\hat{P}) - \Phi^{-1}(P) \). By the delta method, assuming \( g \) is continuously differentiable,
\[ \sqrt{N}(g(\hat{p}) - g(p)) = \left. \frac{\partial g}{\partial p} \right|_{p^*} \sqrt{N}(\hat{p} - p) \]

where \( \min(\hat{p}, p) \leq p^* \leq \max(\hat{p}, p) \)
\[ \sqrt{N}(\hat{p} - p) \sim N(0, \sigma^2_p). \]
Applying this to our regression function, we get

\[ \sqrt{N_i} (\Phi^{-1}(\hat{p}_i) - \Phi^{-1}(p_i)) = \frac{1}{\phi(p_i^*)} \sqrt{N_i} (\hat{p}_i - p_i) = v_i \]

\[ \forall \text{ cells } i. \]

Suppose \( \{N_i\} \to \infty \) and errors are uncorrelated asymptotically. Then

\[ v_i \xrightarrow{d} N \left(0, \left[ \frac{1}{\phi(p_i^*)} \right]^2 \frac{p_i(1 - p_i)}{N_i} \right) \]

where \( \frac{p_i(1 - p_i)}{N_i} \) is the variance of the binary
random variable $p_i$. Obtain the feasible GLS estimator by regressing

\[
\frac{\Phi^{-1}(\hat{p}_i)}{\phi(p^*_i)} \text{ on } \frac{\sigma_e}{\sqrt{p_i(1-p_i)}} \frac{\ln w - x\beta - (1-\alpha)\ln A}{\sqrt{p_i(1-p_i) N_i}}
\]

Can show that the estimates are asymptotically efficient and satisfy the orthogonality condition:
\[
\sum_{i=1}^{I} \text{plim} \frac{1}{\phi(p_i^*)} \left( \frac{\ln w - x\beta - (1 - \alpha) \ln A}{\sigma_e} \right)
\]

\[(\hat{p}_i - p_i) = 0.\]
Microdata analogue

\[ L = \prod_{d_i=1} \Phi(c_i) \prod_{d_i=0} \Phi(-c_i) = \prod_i \Phi(c_i[2D_i - 1]) \]

MLE gives consistent and asymptotically normal estimates of \( \sigma_e, \beta, \alpha \). Assume you don’t know wages for anyone, but do know

\[ \ln w = z\gamma + u \text{ where } u \sim N(0, \sigma_u^2), \]

\( (e - u) \perp \perp x,A,z \), and

\( (e - u) \sim N(0, \sigma_e^2 + \sigma_u^2 - 2\sigma_{eu}) = N(0, \sigma^*^2) \)

\[ (\sigma^*)^2 = (\sigma_u^2 + \sigma_e^2 - 2\sigma_{ue})^{1/2}. \]
Then

\[ \Pr(i \text{ works}) = \Pr \left( \frac{e - u}{\sigma^*} \leq \frac{z\gamma - x\beta - (1 - \alpha) \ln A}{\sigma^*} \right). \]

Note: if \((e, u)\) are Weibull, then \((e - u)\) is logistic.

**Identification (2 cases)**

(i) If \(z, x\) distinct, then can estimate

\[ \left( \frac{\gamma}{\sigma^*}, \frac{\beta}{\sigma^*}, \frac{1 - \alpha}{\sigma^*} \right), \]

but can’t estimate \(\sigma^*\).

(ii) If \(z, x\) have elements in common \((x_c = z_c, x_d, z_d)\), then can estimate only

\[ \left( \frac{\beta_d}{\sigma^*}, \frac{\gamma_d}{\sigma^*}, \frac{\gamma_c - \beta_c}{\sigma^*}, \frac{(1 - \alpha)}{\sigma^*} \right) \]

and again not \(\sigma^*\). Assume you observe
wages for workers only
\[ \ln w = z\gamma + u \]

\[ \ln w_r = x\beta + (1 - \alpha) \ln A + e \]

\((e, u) \perp (x, z, A)\)

\[ y_i = \ln w - \ln w_r = z\gamma - x\beta - (1 - \alpha) \ln A + u - e. \]

Agent \(i\) works if \(y_i > 0\). This is the Roy model with 2 sectors: market \((u)\) and nonmarket \((e)\).

\[ E(\ln w | \ln w > \ln w_r, x, z, A) \]
\[ z\gamma + \frac{\sigma_{uu} - \sigma_{ue}}{\sigma^*} \lambda \left( \frac{x\beta - z\gamma + (1 - \alpha) \ln A}{\sigma^*} \right) \]

\[ = z\gamma + \frac{\sigma_{uu} - \sigma_{ue}}{\sigma^*} \lambda(-c) \]

where \( c = \frac{z\gamma - x\beta - (1 - \alpha) \ln A}{\sigma^*} \)

2-step Estimator

(i) Run probit on LFP decision

\[ \left( \frac{\hat{\gamma}}{\sigma^*}, \frac{\hat{\beta}}{\sigma^*}, \frac{\hat{\alpha}}{\sigma^*} \right) = \arg\max \sum_i \ln \Phi[c_i(2D_i - 1)] \]

Form

\[ \lambda(-\hat{c}_i) = \frac{\phi(-\hat{c}_i)}{1 - \Phi(-\hat{c}_i)}, \]

\[ \hat{c}_i = \frac{z_i\hat{\gamma} - x_i\hat{\beta} - (1 - \hat{\alpha}) \ln A_i}{\hat{\sigma}^*} \]
(ii) Estimate \((\hat{\gamma}, \frac{\sigma_{uu} - \sigma_{ue}}{\sigma^*})\) via OLS on

\[ \ln w_i = z_i \gamma + \frac{\sigma_{uu} - \sigma_{ue}}{\sigma^*} \lambda(-\hat{c}_i) + \eta_i \]

using sample of workers only.

**Identification**

With one exclusion restriction (1 variable, call it \(z_1\), in \(z\) not in \(x\) or \(\ln A\) or \(\ln A\)). We can now identify everything

\( (\gamma, \beta, \alpha, \sigma^*, \sigma_{uu}, \sigma_{ee}, \sigma_{ue}) \):
(i) Step 1 of 2-step gives $\frac{\gamma}{\sigma^*}$ as well as 
\[ \left( \frac{\gamma - \beta}{\sigma^*}, \frac{1 - \alpha}{\sigma^*} \right). \]
Step 2 gives $\gamma$ as well as 
\[ \left( \gamma, \frac{\sigma_{uu} - \sigma_{ue}}{\sigma^*} \right). \]
Solve for $\sigma^*$. Use $\sigma^*$ to solve for $(\sigma_{uu} - \sigma_{ue}, \beta, \alpha)$.

(ii) Look at residuals from step 2:
\[
\ln w_i = z_i \gamma + \frac{\sigma_{uu} - \sigma_{ue}}{\sigma^*} \lambda(-\hat{c}_i) + v_i
\]
\[
E(v_i^2) = \sigma_{uu} \left( \rho^2 \left[ 1 - \lambda^2(-c_i) - \lambda(-c_i)c_i \right] \right) + [1 - \rho^2]
\]
\[
= \sigma_{uu} - \sigma_{uu} \rho^2 \left[ \lambda_i^2 + \lambda_i c_i \right].
\]
Regression of $\hat{v}_i^2$ on $\lambda_i^2 + \lambda_i \hat{c}_i$ gives

consistent estimates of $(\sigma_{uu}, \rho)$. Solve for
\[ \sigma_{ue}. \] Use \( \sigma^* = \sigma_{uu} + \sigma_{ee} - 2\sigma_{ue} \) to solve for \( \sigma_{ee} \).

Without any exclusion restrictions on \( z \).

We can only identify \((\gamma, \sigma_{uu})\): We cannot uniquely identify \( \sigma_{ee} \) or \( \sigma_{ue} \).

(i) Step 2 of 2-step gives \( \left( \gamma, \frac{\sigma_{uu} - \sigma_{ue}}{\sigma^*} \right) \).

(ii) Obtain \((\sigma_{uu}, \rho)\) from residual regression as above.

(iii) To solve for \((\sigma_{ee}, \sigma_{ue})\) either normalize \( \sigma_{ee} = 1 \) or \( \sigma_{ue} = 0 \). If we assume \( \sigma_{ee} = 1 \), then from step 2 we obtain

\[
\frac{\sigma_{uu} - \sigma_{ue}}{1 + \sigma_{uu} - 2\sigma_{ue}},
\]

from which we can ob-
tain $\sigma_{ue}$. If we assume $\sigma_{ue} = 0$, then from step 2 we get \( \frac{\sigma_{uu}}{\sigma_{uu} + \sigma_{ee}} \) and can solve for $\sigma_{ee}$.


$$\ln w = z \gamma + \sigma \lambda(x \hat{\beta}) + \sigma \left[ \lambda(x \beta) - \lambda(x \hat{\beta}) \right] + v.$$ 

Define $-c = \frac{x \beta - z \gamma + (1 - \alpha) \ln A}{\sigma^*}$.

OLS gives consistent estimates of $\gamma, \sigma$

but the variance of the OLS estimates $\gamma, \sigma$
equal the usual OLS variance matrix plus an additional term due to the \( \sigma(\lambda - \hat{\lambda}) \) term, so we have heteroskedasticity and extra variability.

\[
\lambda(x\hat{\beta}) = \lambda(x\beta) + \frac{\partial \lambda}{\partial c} x(\hat{\beta} - \beta) + o(\cdot)
\]

\[
\sqrt{N}(\hat{\lambda} - \lambda) = \frac{\partial \lambda}{\partial c} \sqrt{N} x(\hat{\beta} - \beta) + o(\cdot) \xrightarrow{d} N \left( 0, \frac{\partial \lambda}{\partial c} x \sum_{\beta} x' \frac{\partial \lambda'}{\partial c} \right)
\]

\[
\sum_{\beta} = \text{asy.var} \left( \sqrt{N} x(\hat{\beta} - \beta) \right).
\]

Sampling distribution of the OLS coefficient
is:

\[
\begin{pmatrix}
\hat{\gamma} \\
\hat{\sigma}
\end{pmatrix} = \begin{pmatrix}
\gamma \\
\sigma
\end{pmatrix} + \begin{pmatrix}
z'z & z'\hat{\lambda} \\
\hat{\lambda}'z & \hat{\lambda}'\hat{\lambda}
\end{pmatrix}^{-1}
\begin{pmatrix}
\hat{z} \\
\hat{\lambda}
\end{pmatrix} \left(\sigma(\hat{\lambda} - \lambda) + \nu\right)
\]

\[
Z'Z = \sum Z_iZ_i' \quad Z'\hat{\lambda} = \sum_{i=1}^{1} Z_i\lambda_i \text{ etc.}
\]

Rearranging we get
\[
\sqrt{N} \left[ \left( \hat{\gamma} \right) - \left( \gamma \right) \right] = \left( \frac{z^' z}{N} \frac{z^' \hat{\lambda}}{N} \right)^{-1} \left( \frac{z}{\sqrt{N}} \frac{\sqrt{N}}{\hat{\lambda}} \right) \\
\left( \sigma(\hat{\lambda} - \lambda) + v \right).
\]

Taylor expanding around the true \(\beta\) and taking probability limits of each element on the rhs gives
\[
\frac{z' \hat{\lambda}}{N} = \frac{z' [\lambda + \frac{\partial \lambda}{\partial c} x(\hat{\beta} - \beta)]}{N} = \frac{z' \lambda}{N} \\
+ \frac{z' \frac{\partial \lambda}{\partial c} x(\hat{\beta} - \beta)}{N} \overset{p}{\rightarrow} \frac{z' \lambda}{N}
\]

\[
\frac{\hat{\lambda}' \lambda}{N} = \frac{\left[ \lambda + \frac{\partial \lambda}{\partial c} x(\hat{\beta} - \beta) \right]' \left[ \lambda + \frac{\partial \lambda}{\partial c} x(\hat{\beta} - \beta) \right]}{N}
\]

\[
= \frac{\chi' \lambda}{N} + 2 \frac{\chi' \frac{\partial \lambda}{\partial c} x}{N \sqrt{N}} \sqrt{N} (\hat{\beta} - \beta)
\]

\[
+ \left[ \frac{\partial \lambda}{\partial c} x \frac{1}{N \sqrt{N}} \sqrt{N} (\hat{\beta} - \beta) \right]' \cdot
\]

\[
\left[ \frac{\partial \lambda}{\partial c} x \frac{1}{N \sqrt{N}} \sqrt{N} (\hat{\beta} - \beta) \right] \overset{p}{\rightarrow} \frac{\chi' \lambda}{N}
\]
\[
\frac{z'\sigma(\hat{\lambda} - \lambda)}{\sqrt{N}} = -z'\sigma\left(\frac{\partial\lambda}{\partial c} x(\hat{\beta} - \beta)\right) = \frac{z'\left(\frac{\partial\lambda}{\partial c}\right)_x}{\sqrt{N}} \sqrt{N}(\hat{\beta} - \beta) \xrightarrow{p} -\sigma \sum_1 N(0, \Sigma_{\beta}) = N(0, \sigma^2 \sum_1 \Sigma_{\beta} \sum_1').
\]

Where \( \sum_1 = \text{plim} \left[ \frac{z'\partial\lambda}{\partial c} x \right] \)
\[ \hat{\lambda} \left[ \sigma (\lambda - \hat{\lambda}) + v \right] = -\sigma \left[ \lambda + \frac{\partial \lambda}{\partial c} x (\hat{\beta} - \beta) \right]' \left[ \frac{\partial \lambda}{\partial c} x (\hat{\beta} - \beta) \right] \]

\[ + \frac{\lambda + \frac{\partial \lambda}{\partial c} x (\hat{\beta} - \beta)}{\sqrt{N}} v \left[ \frac{\partial \lambda}{\partial c} x (\hat{\beta} - \beta) \right]' \left[ \frac{\partial \lambda}{\partial c} x (\hat{\beta} - \beta) \right] \]

\[ - \sigma \left[ \lambda + \frac{\partial \lambda}{\partial c} x (\hat{\beta} - \beta) \right]' \left[ \frac{\partial \lambda}{\partial c} x (\hat{\beta} - \beta) \right] \]

\[ \chi' v + \frac{\chi' v}{\sqrt{N}} + \frac{\chi' v}{\sqrt{N}} \left[ \frac{\partial \lambda}{\partial c} x (\hat{\beta} - \beta) \right]' \left[ \frac{\partial \lambda}{\partial c} x (\hat{\beta} - \beta) \right] \]

\[ = -\sigma \left[ \frac{\partial \lambda}{\partial c} x \right]' \left[ \frac{\partial \lambda}{\partial c} x \right] \frac{1}{\sqrt{N}} \sqrt{N} (\hat{\beta} - \beta) \]

\[ - \sigma \left[ \frac{1}{N} \left[ \frac{\partial \lambda}{\partial c} x \right]' \left[ \frac{\partial \lambda}{\partial c} x \right] \sqrt{N} (\hat{\beta} - \beta) \right]' \]

24
\[
\begin{align*}
\left[ \sqrt{N}(\hat{\beta} - \beta) \right] \\
= \sqrt{N}(\hat{\beta} - \beta) + \frac{\lambda' v}{\sqrt{N}} + \frac{\left[ \frac{\partial \lambda}{\partial c} x \right]' v}{\sqrt{N}} + \frac{\chi'}{N} \frac{\partial \lambda}{\partial c} x \sqrt{N}(\hat{\beta} - \beta) \xrightarrow{P} P - \sigma \text{plim } 25
\end{align*}
\]

\[
\begin{align*}
\left[ \sqrt{N}(\hat{\beta} - \beta) \right] + \frac{\lambda' v}{\sqrt{N}} \xrightarrow{P} -\sigma \sum_2 N(0, \sum_B) \\
= N(0, \sigma^2 \sum_2 \sum_\beta \sum'_2)
\end{align*}
\]

(assuming \( v \perp \frac{\partial \lambda}{\partial c} x \) so that plim \( 25 \))
\[
\left[ \frac{\partial \lambda}{\partial c} \right]^N_v = 0
\]

where \( \sum_2 = \text{plim} \left[ \chi' \left( \frac{\partial \lambda}{\partial c} \right) \right] \). Then let

\[
Q_0 = E \left( \begin{pmatrix} z'z & z'\lambda \\ \lambda'z & \lambda'\lambda \end{pmatrix} \right)
\]

\[
Q_1 = \left( \begin{pmatrix} \sum_1^1 \\ \sum_2 \end{pmatrix} \right)
\]

Putting this all together and assuming that the random components of the first and second step are independent (\(i.e.\) sequence of estimates \(\hat{\beta}\) is independent of \(v\))
\[
\sqrt{N} \left[ \left( \hat{\gamma} \right) - \left( \gamma \right) \right] \xrightarrow{d} N[0, V]
\]

\[
V = \sigma^2 Q_0^{-1} + Q_0^{-1} Q_1 \sum_{\beta} Q_1 Q_0^{-1}.
\]

Note that since \( \hat{\beta} \) is estimated using MLE,

\[
\sum_{\beta} = -E \left[ \frac{\partial^2 f(x; \beta)}{\partial \beta \partial \beta'} \right]^{-1}
\]

In the non-independence case we have show that

\[
\sqrt{N} \left[ \left( \hat{\gamma} \right) - \left( \gamma \right) \right] \xrightarrow{d} N[0, \Sigma]
\]

where

27
\[ \sum = \sigma_v^2 Q_0^{-1} + Q_0^{-1}[Q, R_1^{-1}(\beta)Q_1' - Q_1 R_1^{-1}(\beta)Q_2' - Q_2 R_1^{-1}(\beta)Q_1]Q_0^{-1} \]

where \( W = [z\lambda] \)

\[ Q_0 = \text{plim} \frac{1}{n}[W'W] \]

\[ Q_1 = \text{plim} \frac{1}{n} \left[ W' \left[ \frac{\partial \lambda}{\partial c} \right] \right] \]
\[ R_1(\beta) = -E \left[ \frac{\partial^2 f(x; \beta)}{\partial \beta \partial \beta'} \right]^{-1} \]

\[ Q_2 = \text{plim} \frac{1}{n} \sum_{i=1}^{N} W_i' \frac{\partial f}{\partial \beta}(x; \beta) \]