Beginning Doctoral Program Gerzensee

Lectures held by Mark Watson and Bo Honoré

Econometrics Finals Solutions

List of Contributors:

Tom Rodriguez (University of Fribourg)

May 14, 2024

1 Econometrics Final 2015 / 16

Watson

Exercise 1

(a) No. Since $|\theta| > 1$, this process is not invertible:

$$
\begin{aligned}
X_t &= (1-2L)\varepsilon_t \\
X_t \frac{1}{1-2L} &= \varepsilon_t
\end{aligned}
$$

We cannot turn $\frac{1}{1-2L}$ into an infinite sum. Or, one con show that:

$$
\begin{aligned}
\varepsilon_t =& X_t + \theta\varepsilon_{t-1} \\
=& X_t + \theta\left(X_{t-1} + \theta\varepsilon_{t-2}\right) \\
&\vdots \\
=& \sum_{j=0}^{t} \theta^j X_{t-j} + \theta^t \varepsilon_0
\end{aligned}
$$

If $|\theta| > 1, \varepsilon_0$ gets more important as t grows. Thus, ε_t cannot be recovered from past values of X_t.

(b) (i) By observational equivalence, one could estimate:

$$X_t = \tilde{\varepsilon}_t - \tilde{\theta}\tilde{\varepsilon}_{t-1} \text{ where } \tilde{\theta} = 1/2; \text{Var}\left(\tilde{\varepsilon}_t\right) = 1$$

Then we also find

$$X_t \frac{1}{1-\tilde{\theta}L} = \tilde{\varepsilon}_t \Rightarrow \tilde{\varepsilon}_T = \sum_{j=0}^{T-1} \tilde{\theta}^j X_{T-j}$$

$$\mathbb{E}_T\left(X_{T+1}\right) = -\tilde{\theta} \sum_{j=0}^{T-1} \tilde{\theta}^j X_{T-j}$$

(ii) The first line directly follows from the result above. The jump to the second line is legal because $X_{1:T}$ are known, therefore they have no variance.

$$\begin{aligned}\mathrm{Var}\left(X_{T+1}\right) &= \mathrm{Var}\left(\tilde{\varepsilon}_{T+1} - \sum_{j=0}^{T-1} \tilde{\theta}^j X_{T-j}\right) \\ &= \mathrm{Var}\left(\varepsilon_{T+1}\right) = 1\end{aligned}$$

Exercise 2

Did not cover this in lectures

Exercise 3

Correlated errors. Use the Kalman Filter derived in exercise session 2:

$$\begin{bmatrix} w_t \\ v_t \end{bmatrix} \overset{\text{iid}}{\sim} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} R & G \\ G' & Q \end{bmatrix}\right) \qquad \begin{aligned} y_t &= A'X_t + H'\xi_t + w_t \\ \xi_t &= F\xi_{t-1} + v_t \end{aligned}$$

Here: $R = 1; Q = 2; G = 1; A'X_t = 0; F = 1; H = 1$
$\xi_{t-1|t-1} = 3; P_{t-1|t-1} = 0.5; y_t = 2$

#	Variable	Formula	Value
1	$\xi_{t\|t-1}$	$F\xi_{t-1\|t-1}$	3
2	$Y_{t\|t-1}$	$A'X_t + H'\xi_{t\|t-1}$	3
3	$P_{t\|t-1}$	$FP_{t-1\|t-1}F' + Q$	2.5
4	h_t	$H'P_{t\|t-1}H + R + H'G' + GH$	5.5
5	K_t	$\left(P_{t\|t-1}H + G\right) \cdot h_t^{-1}$	0.64
6	η_t	$Y_t - Y_{t\|t-1}$	-1
7	$\xi_{t\|t}$	$\xi_{t\|t-1} + K_t\eta_t$	2.36
8	$P_{t\|t}$	$P_{t\|t-1} - K_t\left(P_{t\|t-1}H + G\right)$	1.08

Therefore:

$$\xi_{t|t} = 2.36$$
$$P_{t|t} = 1.08$$

Exercise 4

(a) OLS:

$$\sqrt{T}(\hat{\beta} - \beta) = \left(\frac{1}{T}\sum_{t=1}^{T} x_t^2\right)^{-1} \left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} x_t u_t\right)$$

$$\mathbb{E}\left(x_t u_t\right) = \mathbb{E}\left(\left(\varepsilon_{t+1} + \varepsilon_{t+2}\right) u_t\right) = 0$$

Unfortunately, $(x_t u_t\}$ is not a martingale difference sequence:

$$\begin{aligned}\mathbb{E}\left(x_t u_t \mid \Omega_{t-1}\right) &= \mathbb{E}\left(\left(\varepsilon_{t+1} + \varepsilon_{t+2}\right)\left(\phi u_{t-1} + \varepsilon_t\right) \mid \Omega_{t-1}\right) \\ &= \mathbb{E}\left(\phi u_{t-1}\varepsilon_{t+1} + \phi u_{t-1}\varepsilon_{t+2} + \varepsilon_t\varepsilon_{t+1} + \varepsilon_t\varepsilon_{t+2} \mid \Omega_{t-1}\right)\end{aligned}$$

What is Ω_{t-1} in this case? $\{u_j\}_{j=0}^{t=1}$ and $\{x_j\}_{j=0}^{t-1}$.

Then, we can find:

$$\left.\begin{aligned} x_{t-1} &= \varepsilon_t + \varepsilon_{t+1} \\ x_{t-2} &= \varepsilon_{t-1} + \varepsilon_t \\ &\vdots \\ x_0 &= \varepsilon_1 + \varepsilon_2 \end{aligned}\right| \Rightarrow \text{ recover } \varepsilon_{t+1} \pm \varepsilon_1$$

Using $\{u_j\}_{j=0}^{t-1}: \quad \varepsilon_1 = u_1 - \phi u_0$

Since we are able to recover $\varepsilon_{t+1} \pm \varepsilon_1$, and ε_1, we can also recover all $\{\varepsilon_j\}_{j=1}^{t+1}$. Therefore, $\varepsilon_{t+1} \in \Omega_{t-1}$ and $\varepsilon_t \in \Omega_{t-1}$. Thus, we conclude that

$$\mathbb{E}\left(x_t u_t \mid \Omega_{t-1}\right) \neq 0$$

Thus, we have that:

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} x_t u_t \xrightarrow{d} N\left(0, \sum_{j=-\infty}^{\infty} \lambda_j\right)$$

Now, we must find all λ_j:

$$\begin{aligned}
\lambda_0 &= \mathbb{E}\left(x^2 u_t^2\right) = \mathbb{E}\left(x_t^2\right) \mathbb{E}\left(u_t^2\right) = 2\sigma^2 \cdot \sigma^2 \left(1-\phi^2\right)^{-1} \\
\lambda_1 &= \mathbb{E}\left(x_t x_{t-1}\right) \mathbb{E}\left(u_t u_{t-1}\right) = \sigma^2 \cdot \phi\sigma^2 \left(1-\phi^2\right)^{-1} \\
\lambda_j &= 0 \quad \forall j > 2 \text{ since } \mathbb{E}\left(x_t x_{t-j}\right) = 0 \quad \forall j > 2
\end{aligned}$$

We can then find the sum we were looking for:

$$\sum_{j=-\infty}^{\infty} \lambda_j = 2\sigma^2 \cdot \sigma^2 \left(1-\phi^2\right)^{-1} (1+\phi) = \frac{2\sigma^4}{1-\phi}$$

Also:

$$\left(\frac{1}{T} \sum x_t^2\right)^{-1} \xrightarrow{p} \mathbb{E}\left(x_t^2\right)^{-1} = \left(2\sigma^2\right)^{-1}$$

Conclusion:

$$\sqrt{T}(\hat{\beta} - \beta) \xrightarrow{d} N\left(0, \frac{1}{2(1-\phi)}\right)$$

(b) We know that

$$\hat{\beta} \overset{a}{\sim} N\left(\beta, \frac{1}{T} \frac{1}{2(1-\phi)}\right)$$

Thus, we construct the following CI :

$$
\begin{aligned}
CI_{95} &= \hat{\beta} \pm 1.96\sqrt{\frac{1}{T}\left(\sum_{j=-1}^{1}\hat{\lambda}_j\right)\left(\frac{1}{T}\sum x_t^2\right)^{-2}} \\
&= 0.08 \pm 1.96\sqrt{\frac{1}{400}(2.43+6.15+2.43)(1.95)^{-2}} \\
&= [-0.087; 0.247]
\end{aligned}
$$

Since $0 \in CI_{95}$, we cannot reject the null hypothesis.

Exercise 5

$$
\left| \begin{array}{l} Y_t = Y_{t-1} + X_t = \sum_{u=1}^{t} X_u \\ X_k = X_{k-1} + \varepsilon_k = \sum_{j=1}^{u} \varepsilon_j \end{array} \right| \Rightarrow Y_t = \sum_{u=1}^{t}\sum_{j=1}^{u}\varepsilon_j
$$

$$
\begin{aligned}
\sum_{t=1}^{T} Y_t &= \sum_{t=1}^{T}\sum_{u=1}^{t}\sum_{j=1}^{u}\varepsilon_j \\
T^{-5/2}\sum_{t=1}^{T} Y_t &= \frac{1}{T}\sum_{t=1}^{T}\boxed{\frac{1}{T}\sum_{u=1}^{t}\frac{1}{\sqrt{T}}\sum_{j=1}^{u}\varepsilon_j}
\end{aligned}
$$

The blue block is analogue to what we saw in class. Let $\xi(t/T) = T^{-1/2}\sum_{i=1}^{t}\varepsilon_i(t/T)$, and $\xi_T(s)$ the function that linearly interpolates between these points, then $\xi_T(s) \xrightarrow{d} W$ as $\sigma^2 = 1$ in this case. W is a Wiener process. As we are summing over n observations, we have

$$
\boxed{\frac{1}{T}\sum_{u=1}^{t}\frac{1}{\sqrt{T}}\sum_{j=1}^{u}\varepsilon_j} \xrightarrow{d} \int_0^u W(s)ds
$$

And now, sum this over all t, then we get

$$
T^{-5/2}\sum_{t=1}^{T} Y_t = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\sum_{u=1}^{t}\frac{1}{\sqrt{T}}\sum_{j=1}^{u}\varepsilon_j \xrightarrow{d} \int_0^1\int_0^u W(s)dsdu
$$

Honoré

Exercise 1

(a) Solve the maximization problem:

$$\max_{m>0} \mathbb{E}\left(-\frac{Y}{m}\right) - \ln(m)$$
$$\text{FOC: } \mathbb{E}(Y)\frac{1}{m^2} - \frac{1}{m} = 0$$
$$\Rightarrow m = \mathbb{E}(Y) = \mu$$

(b) Solve the maximization problem:

$$\max_b \sum_{i=1}^n \frac{y_i}{f(x_i, b)} - \ln(f(x_i, b))$$
$$\text{FOC: } \sum_{i=1}^n -y_i f(x_i, b)^{-2} f'(x_i, b) - f(x_i, b)^{-1} f'(x_i b) = 0$$
$$\sum_{i=1}^n \left(y_i f(x_i, b)^{-1} + 1\right) f(x_i, b)^{-1} f'(x_{i,b} b) = 0$$

No clue how to continue

Exercise 2

No clue. I don't think we looked at ordered logit in lectures.

Exercise 3

$$P(D=1) = \alpha \quad \mathbb{E}(Y_1 \mid D=1) = 10 \quad \mathbb{E}(Y_0 \mid D=0) = 5$$
$$0 \leq Y_0 \leq Y_1 \qquad 0 \leq Y_1 \leq 15$$

$$\left.\begin{array}{l}\alpha 10+(1-\alpha)0 \leq \mathbb{E}\left(Y_1\right) \\ \mathbb{E}\left(Y_1\right) \leq \alpha 10+(1-\alpha)15\end{array}\right| \Rightarrow 10\alpha \leq \mathbb{E}\left(Y_1\right) \leq 15-5\alpha$$

$$\left.\begin{array}{l}(1-\alpha)5+\alpha 0 \leq \mathbb{E}\left(Y_0\right) \\ \mathbb{E}\left(Y_0\right) \leq (1-\alpha)5+\alpha 15\end{array}\right| \Rightarrow 5-5\alpha \leq \mathbb{E}\left(Y_0\right) \leq 5+10\alpha$$

By these conditions, we conclude that:

$$\mathbb{E}\left(Y_1-Y_0\right) \in [15\alpha-5, 10-15\alpha]$$

Size of interval:

$$15\alpha-5-10+15\alpha=30\alpha-15$$

Smallest if size is zero: $\alpha=1/2$

Exercise 4

(a) Instead of regressing the mean, one wants to find a quantile. Let's say one is interested in the median (50% quantile) effect that smoking has on the risk of lung cancer. Then, one would run a quantile regression, and would obtain the constant (risk of cancer for non-smokers), and $\hat{\beta}$ (the median increase in risk for smokers).

(b) Say, one has data that does not fit the linear model very well. Instead of going non-linear one could run multiple linear regressions on subsets of the data.

Exercise 5

Difference the model:

$$\Delta\varepsilon_{i2}=\Delta y_{i2}-\Delta x_{1i2}'\beta_1-\Delta x_{2i2}'\beta_2 \quad i=1,\ldots,n$$

Use the information on $\mathbb{E}(\varepsilon \mid x)$.

$$
\begin{aligned}
\mathbb{E}\left(\Delta\varepsilon_{i2} \mid x_{1is}\right) = 0 \quad & \text{for } s = 1, 2 \\
\mathbb{E}\left(\Delta\varepsilon_{i2} \mid x_{2is}\right) = 0 \quad & \text{for } s = 1
\end{aligned}
$$

Thus, we found 3 moment conditions. The model is over-identified if $\dim\left(x_{1it}\right) + \dim\left(x_{2it}\right) < 3$. I.e. only if x_{1it} & x_{2it} are scalars.

2 Econometrics Final 2016 / 17

Watson

Exercise 1

Kalman Filter equations:

$$F = 0.9; \quad H = R = Q = 1$$
$$\text{Var}\left(x_t\right) = \frac{1}{1-0.81} \cong 5.263$$

#	Variable		Formula	Value
1	$x_{t\|t-1}$	(μ_1)	$F \cdot \mathbb{E}\left(x_{t-1} \mid y_{1:t-1}\right)$	0
2	$y_{t\|t-1}$	(μ_2)	$H \cdot \mu_1$	0
3	$P_{t\|t-1}$	(Σ_{11})	$F^2 \cdot V\left(x_t\right) + Q$	5.263
4	h_t	(Σ_{22})	$H^2 P_{t\|t-1} + R$	6.263
5	K_t	$\left(\Sigma_{12}\Sigma_{22}^{-1}\right)$	$P_{t\|t-1} \cdot H \cdot h_t^{-1}$	0.840
6	η_t	$(z_2 - \mu_2)$	$y_t - y_{t\|t-1}$	1
7	$x_{t\|t}$	$(\mathbb{E}\left(z_1 \| z_2\right))$	$x_{t\|t-1} + K_t \eta_t$	0.840
8	$P_{t\|t}$	$(V\left(z_1 \| z_2\right))$	$P_{t\|t-1} - K_t H P_{t\|t-1}$	0.841

(a) Since everything is normally distributed, we find:

$$\begin{aligned} f\left(y_t \mid x_{t-1} = 0, y_{t-1} = 2\right) &= \frac{1}{\sqrt{2\pi h_t}} \exp\left[-\frac{1}{2}\frac{\eta_t^2}{h_t}\right] \\ &\cong \frac{1}{\sqrt{2\pi 6.263}} \exp\left[-\frac{1}{2}\frac{y_t^2}{6.263}\right] \end{aligned}$$

Remember, we don't know y_t in (a).

(b) Also a normal distribution. Thus:

$$
\begin{aligned}
f\left(x_t \mid y_t=1, x_{t-1}=0, y_{t-1}=2\right) &= \frac{1}{\sqrt{2\pi P_{t|t}}} \exp\left[-\frac{1}{2} \frac{\left(x_t - x_{t|t}\right)^2}{P_{t|t}}\right] \\
&\cong \frac{1}{\sqrt{2\pi 0.841}} \exp\left[-\frac{1}{2} \frac{\left(x_t - 0.84\right)^2}{0.841}\right]
\end{aligned}
$$

Exercise 2

Could not solve this, here's what I did:

$$
y_t \left|x_{t-1} = x_t\right| x_{t-1} + v_t
$$

$$
\begin{aligned}
P\left(x_t = 1 \mid x_{t-1} = 0\right) &= 0.2 \\
P\left(x_t = 0 \mid x_{t-1} = 0\right) &= 0.8 \\
\Rightarrow f\left(x_t \mid x_{t-1} = 0\right) &= 0.2^{x_t} \cdot 0.8^{1-x_t}
\end{aligned}
$$

$$
\begin{aligned}
f\left(y_t \mid y_{1:t-1}\right) =& f\left(y_t \mid x_t = 1\right) P\left(x_t = 1 \mid y_{1:t-1}\right) \\
& + f\left(y_t \mid x_t = 0\right) P\left(x_t = 0 \mid y_{1:t-1}\right) \\
=& f\left(1 + v_t\right) \cdot 0.2 + f\left(v_t\right) \cdot 0.8 \\
=& \frac{0.2}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(y_t - 1\right)^2\right] + \frac{0.8}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} y_t^2\right]
\end{aligned}
$$

Exercise 3

(a)

$$
\begin{aligned}
\operatorname{Var}\left(x_t\right) = 1 + 4 = 5 \text{ from } \varepsilon_t \overset{\text{iid}}{\sim} N(0,1) \\
\operatorname{Var}\left(x_t\right) = \left(1 + \theta^2\right) \sigma_\eta^2 \text{ from } MA(1)
\end{aligned}
$$

Combine the two:

$$
5 = \left(1 + \theta^2\right) \sigma_\eta^2 \tag{1}
$$

Also get auto-covariance:

$$\begin{aligned}
\operatorname{Cov}(x_t, x_{t+1}) &= \operatorname{Cov}(\varepsilon_t + 2\varepsilon_{t-1}, \varepsilon_{t+1} + 2\varepsilon_t) \\
&= 2\operatorname{Cov}(\varepsilon_t, \varepsilon_t) = 2 \\
\operatorname{Cov}(x_t, x_{t+1}) &= \operatorname{Cov}(\eta_t + \theta\eta_{t-1}, \eta_{t+1} + \theta\eta_t) \\
&= \theta \operatorname{Cov}(\eta_t, \eta_t) = \theta\sigma_\eta^2
\end{aligned}$$

Combine the two:

$$2 = \theta\sigma_\eta^2 \tag{2}$$

Plug (2) into (1):

$$\begin{aligned}
& 5 = \left(1 + \theta^2\right)\frac{2}{\theta} \\
\Leftrightarrow \quad & 0 = 2\theta^2 - 5\theta + 2 \\
\Leftrightarrow \quad & 0 = \theta^2 - 2.5\theta + 1 \\
\Leftrightarrow \theta_{1/2} &= \frac{2.5 \pm \sqrt{2.25}}{2} = \frac{2.5 \pm 1.5}{2} = \{1/2; 2\}
\end{aligned}$$

$$\begin{aligned}
\theta_1 &= \frac{1}{2} \quad \text{By invertibility} \\
\sigma_\eta^2 &= 4 \quad \text{By (2)}
\end{aligned}$$

(b)

$$\begin{aligned}
&\eta_t + \theta\eta_{t-1} = (1 + \theta L)\eta_t = \varepsilon_t + 2\varepsilon_{t-1} = (1 + 2L)\varepsilon_t \\
&\eta_t = \frac{1 + 2L}{1 + \theta L}\varepsilon_t = (1 + 2L)\left(1 - \theta L + \theta^2 L^2 - \theta^3 L^3 + \ldots\right)\varepsilon_t \\
&\quad = \left(1 + 2L - \theta L - 2\theta L^2 + \theta^2 L^2 + 2\theta^2 L^3 - \theta^3 L^3 - 2\theta^3 L^4 + \ldots\right)\varepsilon_t \\
&\quad = \left(1 + (2 - \theta)L + \left(-2\theta + \theta^2\right)L^2 + \left(2\theta^2 - \theta^3\right)L^3 + \ldots\right)\varepsilon_t \\
&\quad = \left(1 + (2 - \theta)L + (2 - \theta)(-\theta)L^2 + (2 - \theta)(-\theta)^2 L^3 + \ldots\right)\varepsilon_t \\
&\eta_t = \varepsilon_t + (2 - \theta)\sum_{i=0}^{t}(-\theta)^i\varepsilon_{t-1-i}
\end{aligned}$$

Exercise 4

Did not cover this in lectures.

Exercise 5

(a)

(i)

$$\sqrt{T}(\hat{\alpha}-\alpha) = \left(\frac{1}{T}\sum x_{t-1}^2\right)^{-1}\left(\frac{1}{\sqrt{T}}\sum x_{t-1}\varepsilon_t\right)$$

$$\left(\frac{1}{T}\sum x_{t-1}^2\right)^{-1} \xrightarrow{p} \mathbb{E}\left(x_{t-1}^2\right)^{-1} = \frac{1-\phi^2}{2}$$

$$\left(\frac{1}{\sqrt{T}}\sum x_{t-1}\varepsilon_t\right) \xrightarrow{d} N\left(0, \mathbb{E}\left(x_{t-1}^2\varepsilon_t^2\right)\right)$$

Use the fact that $x_t \perp \varepsilon_t \forall t$

$$\left(\frac{1}{\sqrt{T}}\sum x_{t-1}\varepsilon_t\right) \xrightarrow{d} N\left(0, \mathbb{E}\left(x_{t-1}^2\right)\mathbb{E}\left(\varepsilon_t^2\right)\right)$$

Combine the two: (Slutsky)

$$\sqrt{T}(\hat{\alpha}-\alpha) \xrightarrow{d} N\left(0, \frac{1-\phi^2}{2}\right)$$

(ii)

$$\sqrt{T}(\hat{\phi}-\phi) = \left(\frac{1}{T}\sum x_{t-1}^2\right)^{-1}\left(\frac{1}{\sqrt{T}}\sum x_{t-1}v_t\right)$$

$$\left(\frac{1}{T}\sum x_{t-1}^2\right)^{-1} \xrightarrow{p} \mathbb{E}\left(x_{t-1}^2\right)^{-1} = \frac{1-\phi^2}{2}$$

$$\left(\frac{1}{\sqrt{T}}\sum x_{t-1}v_t\right) \xrightarrow{d} N\left(0, \mathbb{E}\left(x_{t-1}^2 v_t^2\right)\right)$$

Use the fact that $x_t \perp v_t \forall t$

$$\left(\frac{1}{\sqrt{T}}\sum x_{t-1}v_t\right) \xrightarrow{d} N\left(0, \mathbb{E}\left(x_{t-1}^2\right)\mathbb{E}\left(v_t^2\right)\right)$$

Combine the two: (Slutsky)

$$\sqrt{T}(\hat{\phi}-\phi) \xrightarrow{d} N\left(0, \frac{1-\phi^2}{2}\right)$$

(iii) Write in matrix notation:

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \underbrace{\begin{bmatrix} x_{t-1} & 0 \\ 0 & x_{t-1} \end{bmatrix}}_{X} \begin{bmatrix} \alpha \\ \phi \end{bmatrix} + \underbrace{\begin{bmatrix} \varepsilon_t \\ v_t \end{bmatrix}}_{\eta_t}$$

Apply GMM:

$$\sqrt{T}\begin{pmatrix} \hat{\alpha}-\alpha \\ \hat{\phi}-\phi \end{pmatrix} \xrightarrow{\alpha} N(0,\Omega)$$

And $\Omega = \mathbb{E}\left(XX'\right)^{-1}\mathbb{E}\left(\left(\eta' X\right)'\left(\eta X\right)'\right)\mathbb{E}\left(XX'\right)^{-1}$

$$\begin{aligned}
\mathbb{E}(XX')^{-1} &= \begin{bmatrix} \mathbb{E}\left(x_{t-1}^2\right)^{-1} & 0 \\ 0 & \mathbb{E}\left(x_{t-1}^2\right)^{-1} \end{bmatrix} = \frac{1-\phi^2}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
\eta' X &= \begin{bmatrix} \varepsilon_t \\ v_t \end{bmatrix}' \begin{bmatrix} x_{t-1} & 0 \\ 0 & x_{t-1} \end{bmatrix} = \begin{bmatrix} \varepsilon_t x_{t-1} & v_t x_{t-1} \end{bmatrix} \\
\mathbb{E}\left(\left(\eta' X\right)'\left(\eta X\right)'\right) &= \begin{bmatrix} \mathbb{E}\left(\varepsilon_t^2 x_{t-1}^2\right) & \mathbb{E}\left(\varepsilon_t v_t x_{t-1}^2\right) \\ \mathbb{E}\left(\varepsilon_t v_t x_{t-1}^2\right) & \mathbb{E}\left(v_t^2 x_{t-1}^2\right) \end{bmatrix} \\
&= \begin{bmatrix} \mathbb{E}\left(\varepsilon_t^2\right)\mathbb{E}\left(x_{t-1}^2\right) & \mathbb{E}\left(\varepsilon_t\right)\mathbb{E}\left(v_t\right)\mathbb{E}\left(x_{t-1}^2\right) \\ \mathbb{E}\left(\varepsilon_t\right)\mathbb{E}\left(v_t\right)\mathbb{E}\left(x_{t-1}^2\right) & \mathbb{E}\left(v_t^2\right)\mathbb{E}\left(x_{t-1}^2\right) \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \cdot \frac{2}{1-\phi^2} \\
\Omega &= \frac{1-\phi^2}{2}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}
\end{aligned}$$

(b)

$$CI = \left[\hat{\hat{\alpha}} \pm 1.96 \frac{1}{\sqrt{T}} \sqrt{\frac{1-\hat{\phi}^2}{2}} \right] = [1.203; 1.396]$$

Honoré

Exercise 1

(a) The issue is that the asymptotics for OLS only hold for $n \to \infty$, with a fixed number of parameters. But as $n \to \infty$, the number of the α_i also goes to infinity.

Then, we cannot say anything about the distributions of (β, γ, δ).

(b) We should use first differences (α_i drop out):

$$\Delta y_{it} = \Delta x_{it}'\beta + \Delta x_{it-1}'\gamma + \Delta x_{it-2}'\delta + \Delta\varepsilon_{it}$$

We need to start at $T = 4$, otherwise the explanatory variables are not well defined. Also note, that by assumption:

$$\mathbb{E}\left(\Delta\varepsilon_{it} \mid x_{it}, x_{it-1}, x_{it-2}, \ldots\right) = 0$$

Thus, the errors are uncorrelated, and OLS should recover the coefficients.

Not super sure if GMM would be better with moment conditions:

$$\mathbb{E}\left(\Delta\varepsilon_{it} x_{it}\right) = \mathbb{E}\left(\left(\Delta y_{it} - \Delta x_{it}'\beta - \Delta x_{it-1}'\gamma - \Delta x_{it-2}'\delta\right) x_{it}\right) = 0$$

Exercise 2

(a)

$$\begin{aligned}
&\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N\left(0, A^{-1}BA^{-1}\right) \\
&A = \mathbb{E}\left(x_i^2 \exp\left(x_i\beta\right)^2\right) = \mathbb{E}\left(x_i^2 \exp\left(2x_i\beta\right)\right) \\
&B = \mathbb{E}\left(\varepsilon_i^2 x_i^2 \exp\left(2x_i\beta\right)\right) = \mathbb{E}\left(\mathbb{E}\left(\varepsilon_i^2 \mid x_i\right) x_i^2 \exp\left(2\beta x_i\right)\right) \\
&\quad = \mathbb{E}\left(x_i^2 \exp\left(3\beta x_i\right)\right) \\
&\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N\left(0, \frac{\mathbb{E}\left(x_i^2 \exp\left(3\beta x_i\right)\right)}{\left[\mathbb{E}\left(x_i^2 \exp\left(2x_i\beta\right)\right)\right]^2}\right)
\end{aligned}$$

(b)

$$
\begin{aligned}
&\sqrt{n}(\hat{\beta}-\beta) \xrightarrow{d} N\left(0, \Gamma^{-1} S \Gamma^{-1}\right) \\
&\Gamma = \mathbb{E}\left(-x_i \exp\left(x_i \beta\right)\right) \\
&S = V\left(f\left(x_i, \beta\right)\right) = V\left(y_i - \exp\left(x_i \beta\right)\right) = V\left(\varepsilon_i\right) = \mathbb{E}\left(\exp\left(x_i \beta\right)\right) \\
&\sqrt{n}(\hat{\beta}-\beta) \xrightarrow{d} N\left(0, \frac{\mathbb{E}\left(\exp\left(x_i \beta\right)\right)}{\left[\mathbb{E}\left(x_i \exp\left(x_i \beta\right)\right)\right]^2}\right)
\end{aligned}
$$

(c)

$$
\begin{aligned}
&\sqrt{n}(\hat{\beta}-\beta) \xrightarrow{d} N\left(0, \left(G' S^{-1} G\right)^{-1}\right) \\
&G = \mathbb{E}\begin{bmatrix} -x_i \exp\left(x_i \beta\right) \\ -x_i^2 \exp\left(x_i \beta + 2\right) \end{bmatrix} \\
&S^{-1} = V\begin{bmatrix} \varepsilon_i \\ \left(y_i - \exp\left(x_i \beta\right) \underbrace{\exp(2)}_{\text{WTF?}} x_i\right) \end{bmatrix}
\end{aligned}
$$

Exercise 3

Approximately: (Nonparametrics, slide 11)

$$
\begin{aligned}
\text{Bias}(\hat{f}(x)) &\cong \frac{1}{2} h^2 f''(x) \int v^2 K(v) dv \\
&= \frac{1}{2} h^2 f''(x) \int v^2 \frac{1}{2} dv \\
&= \frac{1}{2} h^2 \frac{1}{4} \exp\left(-\frac{x}{2}\right)\left(\frac{x^2}{2} - 1\right) \frac{1}{2}\left[\frac{1}{3} v^3\right]_{-1}^{1} \\
&= \frac{1}{2} h^2 \frac{1}{4} \exp\left(-\frac{x}{2}\right)\left(\frac{x^2}{2} - 1\right) \frac{1}{3} \\
\text{Bias}(\hat{f}(1)) &\cong \frac{1}{n^{1/4}} \exp\left(-\frac{1}{2}\right)\left(-\frac{1}{16}\right)
\end{aligned}
$$

$$
\begin{aligned}
V(\hat{f}(x)) &\cong \frac{1}{nh} f(x) \int K(v)^2 dv \\
&= \frac{1}{3n^{3/4}} \frac{1}{2} \exp\left(-\frac{x}{2}\right) \int_{-1}^{1} \frac{1}{4} dv \\
&= \frac{1}{3n^{3/4}} \frac{1}{2} \exp\left(-\frac{x}{2}\right) \frac{1}{2} \\
V(\hat{f}(1)) &= \frac{1}{12} \frac{1}{n^{3/4}} \exp\left(-\frac{1}{2}\right) \\
\mathrm{MSE}(\hat{f}(1)) &= \left[\frac{1}{n^{1/4}} \exp\left(-\frac{1}{2}\right)\left(-\frac{1}{16}\right)\right]^2 + \frac{1}{12} \frac{1}{n^{3/4}} \exp\left(-\frac{1}{2}\right) \\
&= n^{-1/2} \exp(-1) 2^{-8} + n^{-3/4} \frac{1}{12} \exp\left(-\frac{1}{2}\right) \\
&= \mathrm{const}_1\, n^{-1/2} + \mathrm{const}_2\, n^{-3/4}
\end{aligned}
$$

Exercise 4

(a) It is given by $\phi\left(x_0'\beta\right)\beta_l$, where $\phi(\cdot)$ is the pdf of a standard normal, and β_l the coefficient on the explanatory variable.

(b) We can estimate the marginal effect by:

$$
g(\hat{\beta}) = \phi\left(x_0'\hat{\beta}\right)\hat{\beta}_l
$$

Also recall from the lecture that

$$
\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \Sigma)
$$

Now, we can apply the delta-method as $g(\cdot)$ is a non-linear function of β.

$$
\sqrt{n}(g(\hat{\beta}) - g(\beta)) \xrightarrow{d} N\left(0, \left(\frac{\partial g(\beta)}{\partial \beta}\right)' \Sigma \frac{\partial g(\beta)}{\partial \beta}\right)
$$

(c) Very long answer...

3 Econometrics Final 2017 / 18

Watson

Exercise 1

(a)

$$\begin{aligned}
\mathbb{E}_t\left(Y_{t+1} \mid \varepsilon_t + \varepsilon_{t-1} = 2\right) &= \mathbb{E}_t\left(\varepsilon_{t+1} + 0.8\varepsilon_t \mid \varepsilon_t = 2 - \varepsilon_{t-1}\right) \\
&= \mathbb{E}_t\left(\varepsilon_{t+1} + 0.8\left(2 - \varepsilon_{t-1}\right)\right) \\
&= \mathbb{E}_t\left(\varepsilon_{t+1}\right) + 1.6 - 0.8\mathbb{E}_t\left(\varepsilon_{t-1}\right) \\
&= 1.6
\end{aligned}$$

The optimal forecast is the conditional expectation of Y_{t+1} given the information that $\varepsilon_t + \varepsilon_{t-1} = 2$.

(b)

$$\begin{aligned}
\hat{\phi} &= \frac{\widehat{\operatorname{Cov}}\left(Y_t, Y_{t-1}\right)}{\widehat{\operatorname{Var}}\left(Y_{t-1}\right)} \\
\xrightarrow{p} \frac{\operatorname{Cov}\left(Y_t, Y_{t-1}\right)}{\operatorname{Var}\left(Y_{t-1}\right)} &= \frac{\operatorname{Cov}\left(\varepsilon_t + 0.8\varepsilon_{t-1}, \varepsilon_{t-1} + 0.8\varepsilon_{t-2}\right)}{\operatorname{Var}\left(\varepsilon_{t-1} + 0.8\varepsilon_{t-2}\right)} \\
&= \frac{\operatorname{Var}\left(\varepsilon_{t-1}\right) \cdot 0.8}{\operatorname{Var}\left(\varepsilon_{t-1}\right) + 0.64\operatorname{Var}\left(\varepsilon_{t-2}\right)} \\
&= \frac{0.8}{1.64} = 0.488
\end{aligned}$$

Exercise 2

(a) It is not invertible. Let $X_t = Y_t - \beta$, then

$$\begin{aligned}
X_t &= \varepsilon_t - \theta\varepsilon_{t-1} \\
\Leftrightarrow \varepsilon_t &= X_t + \theta\varepsilon_{t-1} \\
&= X_t + \theta\left(X_{t-1} + \theta\varepsilon_{t-2}\right) \\
&\vdots \\
&= \theta^t\varepsilon_0 + \sum_{i=0}^{t-1}\theta^i X_{t-i}
\end{aligned}$$

If $|\theta| \geq 1$, then $\theta^t \varepsilon_0$ does not converge towards zero as t gets large. Therefore, X_t cannot be expressed only by its lagged values plus the period t error.

(b)

$$
\begin{aligned}
\sqrt{T}(\bar{Y} - \beta) = \sqrt{T}\frac{1}{T}\sum_{t=1}^{T}(Y_t - \beta) &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\varepsilon_t - \theta\varepsilon_{t-1}) \\
&= \frac{1}{\sqrt{T}}\left(\sum_{t=1}^{T}\varepsilon_t - \theta\sum_{t=0}^{T-1}\varepsilon_t\right) \\
&= \frac{1}{\sqrt{T}}\left(\sum_{t=1}^{T}\varepsilon_t - \theta\sum_{t=1}^{T}\varepsilon_t - \theta\varepsilon_0 + \theta\varepsilon_T\right) \\
&= \underbrace{(1-\theta)\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t}_{\xrightarrow{d} N(0,\sigma^2(1-\theta)^2)} + \underbrace{\frac{\theta}{\sqrt{T}}(\varepsilon_T - \varepsilon_0)}_{\xrightarrow{p} 0}
\end{aligned}
$$

Apply Slutsky:

$$
\begin{aligned}
&\sqrt{T}(\bar{Y} - \beta) \xrightarrow{d} N\left(0, \sigma^2(1-\theta)^2\right) \\
&\sqrt{T}(\bar{Y} - \beta) \xrightarrow{d} N(0, 1)
\end{aligned}
$$

(c) Delta method: $g(x) = x^2; (g'(x))^2 = 4x^2$

$$
\sqrt{T}\left(\bar{Y}^2 - \beta^2\right) \xrightarrow{d} N\left(0, 1 \cdot g'(\beta)^2\right)
$$

Use $g'(\beta)^2 = 4\beta^2$ and $\beta = 5$. Then:

$$
\sqrt{T}\left(\bar{Y}^2 - 25\right) \xrightarrow{d} N(0, 100)
$$

(d) If $\theta = 1$, this would lead to $V = 0$ which is clearly incorrect. There, I would do the following:

$$Y_t - \beta = u_t = \varepsilon_t - \varepsilon_{t-1}$$

$$\frac{1}{\sqrt{T}} \sum Y_t - \beta = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_t$$

Now $u_t \sim N(0, 2)$ but not iid anymore.

How further??

Exercise 3

(a)

$$\bar{X}_t = \frac{1}{n} \sum_{i=1}^{n} \xi_t + \varepsilon_{it} = \xi_t + \frac{1}{n} \sum_{i=1}^{n} \varepsilon_{it}$$

$$\xi_t \xrightarrow{p} \xi_t \text{ as } \xi_t = \xi_t$$

$$\frac{1}{n} \sum_{i=1}^{n} \varepsilon_t \xrightarrow{p} \mathbb{E}\left(\varepsilon_{it}\right) = 0 \text{ by LLN}$$

I conclude that

$$\bar{X}_t \longrightarrow \xi_t$$

(b)

$$\begin{aligned}
MSE &= \mathbb{E}\left(\left(\bar{X}_t - \xi_t\right)^2\right) \\
&= \mathbb{E}\left(\left(\frac{1}{n} \sum_{i=1}^{n} \varepsilon_{it}\right)^2\right) = \left(\frac{1}{n}\right)^2 \mathbb{E}\left(\sum_{i=1}^{n} \varepsilon_{it}^2 + \sum_{j \neq i} \varepsilon_{it}\varepsilon_{jt}\right) \\
&= \frac{1}{n^2} \sum_{i=1}^{n} \underbrace{\mathbb{E}\left(\varepsilon_{it}^2\right)}_{\sigma^2 = 1} + \sum_{j \neq i} \underbrace{\mathbb{E}\left(\varepsilon_{it}\varepsilon_{jt}\right)}_{=0 \text{ by iid } N(0,1)} \\
&= \frac{1}{n}
\end{aligned}$$

(c) Maybe MLE-estimator?

Honoré

Exercise 1

(a) $CI = [\hat{\beta} \pm 1.96 \cdot \hat{SE}(\hat{\beta})] \cong [0.227; 0.962]$

(b) Reject the hypothesis:

$$t = \frac{\hat{\beta} - \beta_0}{\hat{SE}(\hat{\beta})} = \frac{0.092 + 0.2}{0.115} = 2.545 > 1.96$$

(c)

$$\begin{aligned} x_i'\beta &= 1.1 \cdot 0.98 + 0.8 \cdot 0.152 + 1 \cdot (-0.223) - 1.872 \\ &= -0.896 \\ P\left(y_i = 1 | x_i\right) &= \frac{\exp\left(x_i'\beta\right)}{1 + \exp\left(x_i'\beta\right)} \cong 28.993\% \end{aligned}$$

(d)

$$\begin{aligned} \frac{\partial P\left(y_i = 1 \mid x_i\right)}{\partial bloodp} &= P\left(y_i = 1 \mid x_i\right) P\left(y_i = 0 \mid x_i\right) \cdot \beta_{\text{bloodp}} \\ &\cong 0.202 \end{aligned}$$

Exercise 2

(1)

$$\begin{aligned} \sqrt{n}(\hat{\beta} - \beta) &\xrightarrow{d} N\left(0, A^{-1}BA^{-1}\right) \\ A &= \mathbb{E}\left((1 + 2\beta x_i)^2\right) \\ B &= \sigma^2 \mathbb{E}\left((1 + 2\beta x_i)^2\right) \end{aligned}$$

Thus, we can simplify:

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N\left(0, \sigma^2 \mathbb{E}\left((1 + 2\beta x_i)^2\right)^{-1}\right)$$

(2)

$$\sqrt{n}(\hat{\beta}-\beta) \xrightarrow{d} N\left(0, G^{-1}SG^{-1}\right)$$

$$\begin{aligned} G &= \mathbb{E}\left(-x_i\left(1+2\beta x_i\right)\right) = -\left(\mathbb{E}\left(x_i\right) + 2\beta\mathbb{E}\left(x_i^2\right)\right) \\ S &= V\left(\left(y_i - \left(\beta+\beta^2 x_i\right)\right)x_i\right) = V\left(\varepsilon_i x_i\right) \\ &= \mathbb{E}\left(\varepsilon_i^2 x_i^2\right) - \mathbb{E}\left(\varepsilon_i x_i\right)^2 = \mathbb{E}\left(\mathbb{E}\left(\varepsilon_i^2 \mid x_i\right) x_i^2\right) \\ &= \sigma^2\mathbb{E}\left(x_i^2\right) \end{aligned}$$

(3)

$$\sqrt{n}(\hat{\beta}-\beta) \xrightarrow{d} N\left(0, \left(G'S^{-1}G\right)^{-1}\right)$$

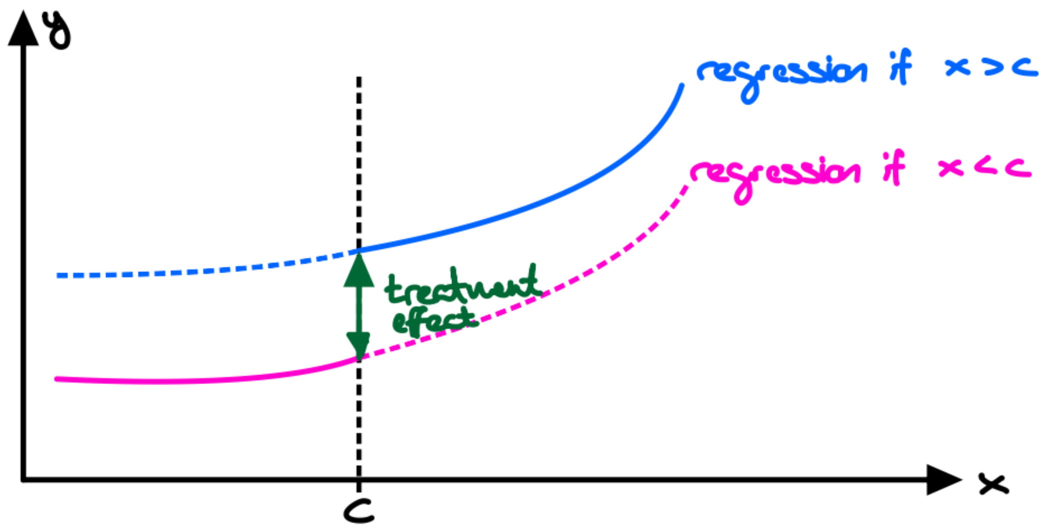
$$G = \mathbb{E}\begin{bmatrix} -\left(1+2\beta x_i\right) \\ -x_i\left(1+2\beta x_i\right) \end{bmatrix}$$

$$S = V\begin{bmatrix} \varepsilon_i \\ \varepsilon_i x_i \end{bmatrix} = \mathbb{E}\begin{bmatrix} \varepsilon_i^2 & \varepsilon_i^2 x_i \\ \varepsilon_i^2 x_i & \varepsilon_i^2 x_i^2 \end{bmatrix} = \sigma^2\begin{bmatrix} 1 & \mathbb{E}\left(x_i\right) \\ \mathbb{E}\left(x_i\right) & \mathbb{E}\left(x_i^2\right) \end{bmatrix}$$

Exercise 3

- When we think that the treatment is that some variable x is greater then some threshold c. Examples would be:
  - let x be time, c be the year 1989, and y is GDP growth in eastern Germany. Since before c, eastern Germany was under communist rule, one could interpret 1989 as the threshold after which the treatment "capitalism" was implemented.
  - let x be school grades, c be the cutoff to get into med-school, and y be earnings. We can use this cut-off as a treatment.
- It assumes, that the regressions in the counterfactual would continue continuously. Also, that the treatment at c actually causes a jump in y.

The following graph helps to drive the idea home:



Exercise 4

First, we get rid of the α_i by taking first differences:

$$\Delta y_{it} = \Delta x_{it}\beta_1 + \Delta x_{2it}\beta_2 + \Delta\varepsilon_{it} \quad t = 2, 3$$

Note, that we can use the following to find the moment conditions:

$$\mathbb{E}\left(\Delta\varepsilon_{it} \mid x_{1is}\right) = 0 \quad \forall t, s$$
$$\mathbb{E}\left(\Delta\varepsilon_{it} \mid x_{2is}\right) = 0 \quad \forall t \geq s$$

$$\begin{aligned}
\text{for } x_{1it}: \quad & \mathbb{E}\left(\Delta\varepsilon_{it} x_{1is}\right) = 0 \quad \forall s, t \\
& \mathbb{E}\left(\Delta y_{it} - \Delta x_{1it}\beta_1 + \Delta x_{2it}\beta_2\right) x_{1is} = 0 \quad \forall s, t \\
& \longrightarrow 6 \text{ moment conditions} \\
\text{for } x_{2it}: \quad & \mathbb{E}\left(\Delta\varepsilon_{it} x_{2is}\right) = 0 \quad \forall (s,t) \in \{(1,2), (1,3), (2,3)\} \\
& \mathbb{E}\left(\left(\Delta y_{it} - \Delta x_{1it}\beta_1 + \Delta x_{2it}\beta_2\right) x_{2is}\right) = 0 \quad \forall (s,t) \in \{(1,2), (1,3), (2,3)\} \\
& \longrightarrow 3 \text{ moment conditions}
\end{aligned}$$

Therefore, we can use 9 moment conditions in total, and GMM will work to estimate (β_1, β_2).

4 Econometrics Final 2018 / 19

Watson

Exercise 1

$$\sqrt{T}(\hat{\mu} - \mu) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_t$$

$\hat{\mu}$ is the OLS estimator of the mean (regressing on a vector of ones).

(a)

(i) u_t is not an mds. Thus, we must use the ACGF:

$$\frac{1}{\sqrt{T}} \sum_{t=0}^{T} u_t \xrightarrow{d} N\left(0, \sum_{j=-\infty}^{\infty} \lambda_j\right)$$

$MA(\infty)$ representation:

$$u_t = (1 - \phi L)^{-1} \varepsilon_t \Rightarrow \sum_{j=-\infty}^{\infty} \lambda_j = \frac{\sigma^2}{(1-\phi)^2}$$

Putting everything together:

$$\sqrt{T}(\bar{Y} - \mu) \xrightarrow{d} N\left(0, \frac{\sigma^2}{(1-\phi)^2}\right)$$

(ii) Since we don't have $\hat{\sigma}^2$ and $\hat{\phi}$, we will use $V = \sum_{i=-\infty}^{\infty} \lambda_i$, and $\sum_{i=-2}^{2} \lambda_i$ is a consistent estimator for V. As Y is a scalar: $\lambda_i = \lambda_{-i}$

$$CI = \left[\hat{\mu} \pm 1.96 \sqrt{\frac{1}{T} \sum_{i=-2}^{2} \hat{\lambda}_i}\right] = [11.493, 12.707]$$

(b)

(i) Again, u_t is not a mds. Use ACGF:

$$u_t = (1+\theta L)\varepsilon_t \Rightarrow \sum_{j=-\infty}^{\infty} \lambda_j = \sigma^2(1+\theta)^2$$
$$\sqrt{T}(\bar{Y}-\mu) \xrightarrow{d} N\left(0, \sigma^2(1+\theta)^2\right)$$

(ii)

$$\lambda_0 = \operatorname{Var}(Y_t) = \operatorname{Var}(u_t) = \sigma^2 + \theta^2\sigma^2$$
$$\lambda_1 = \operatorname{Cov}(Y_t, Y_{t+1}) = \operatorname{Cov}(u_t, u_{t+1}) = \theta\sigma^2$$
$$\operatorname{Cov}(Y_t, Y_{t+k}) = 0 \quad \forall k > 1$$
$$CI = \left[\hat{\mu} \pm 1.96 \frac{1}{\sqrt{T}} \left(\hat{\lambda}_0 + 2\hat{\lambda}_1\right)^{1/2}\right] = [11.560; 12.640]$$

Exercise 2

(a) First. I will rewrite the model as

$$\bar{Y}_t = \xi_t + \underbrace{\frac{1}{3}(\varepsilon_{1t} + \varepsilon_{2t} + \varepsilon_{3t})}_{\equiv u_t} = \xi_t + u_t$$
$$\xi_t = F\xi_{t-1} + e_t \quad ; \quad F = 0.62$$
$$\begin{bmatrix} u_t \\ e_t \end{bmatrix} \overset{\text{iid}}{\sim} N\left(0, \begin{bmatrix} R & 0 \\ 0 & Q \end{bmatrix}\right) \quad \Bigg| \quad \begin{matrix} R = 1/3 \\ Q = 1 \end{matrix}$$

From here, we can apply the Kalman Filter. The eight equations are given by (ignore the blue text):

(1) $\xi_{t|t-1} = F\xi_{t-1|t-1} = 0.8 \cdot 0 = 0$

(2) $\bar{Y}_{t|t-1} = \xi_{t|t-1} = 0$

(3) $P_{t|t-1} = F^2 P_{t-1|t-1} + Q = 2.778$

(4) $h_t = P_{t|t-1} + R = 3.111$

(5) $K_t = P_{t|t-1} h_t^{-1} = 0.893$

(6) $\eta_t = \bar{Y}_t - \bar{Y}_{t|t-1} = 2$

(7) $\xi_{t|t} = \xi_{t|t-1} + K_t \eta_t = 1.786$

(8) $P_{t|t} = P_{t|t-1} - K_t P_{t|t-1}$

We need one more thing to start the recursion, which is the initial values. As it is AR(1) & stationary, use

$$\xi_{0|0} = \mathbb{E}(\xi_0) = 0$$
$$P_{0|0} = \operatorname{Var}(\xi_0) = \frac{\operatorname{Var}(e_t)}{1-F^2} = \frac{1}{0.36} \cong 2.778$$

Now, we can go through the equations and plug in numerical values (in blue) to find $\hat{\xi}_{t|t} = 1.786$

(b) Use the eight equations to iterate through t periods until we find $\xi_{t|t}$ as the best guess for ξ_t.

Exercise 3

(a) I don't think so. My explanation would be that we have 2 innovations for only 1 variable. Thus, we cannot recover ε_{1t} ad ε_{2t} from past values of Y_t. Therefore, not invertible.

(b)

$$\begin{aligned}
Y_t =& \varepsilon_{1t} + \Bigg(\theta_{11}\varepsilon_{1t-1} + \theta_{21}\varepsilon_{1t-2} + \theta_{31}\boxed{\left[Y_{t-3} - \sum_{h=1}^{\infty}\theta_{h1}\varepsilon_{1,t-h-3} - \sum_{h=0}^{\infty}\theta_{h,2}\varepsilon_{2,t-h-3,3}\right]} \\
&+\theta_{41}\varepsilon_{1t-4} + \ldots) + \sum_{h=0}^{\infty}\theta_{h,2}\varepsilon_{2t-h} \\
=& \theta_{31}\boxed{\left[Y_{t-3} - \sum_{h=1}^{\infty}\theta_{h1}\varepsilon_{1,t-h-3} - \sum_{h=0}^{\infty}\theta_{h,2}\varepsilon_{2,t-h-3,3}\right]} \\
&+ \varepsilon_{1t} + \sum_{h=1}^{\infty}\theta_{1,h}\varepsilon_{t-h} + \sum_{h=0}^{\infty}\theta_{h,2}\varepsilon_{2,t-h} - \theta_{31}\varepsilon_{1t-3}
\end{aligned}$$

We see that the "error" is serially correlated. The OLS regressor is inconsistent.

Cannot really prove why though. Maybe be cause we don't know is $|\theta_j| < 1$, which would make it stationary.

(c)

$$\varepsilon_{1t} = \frac{1}{12}\left(Z_t - e_t\right) = \frac{1}{12}Z_t - \tilde{e}_t$$

WTF?

Honoré

Exercise 1

(1) $CI = [\hat{\beta} \pm 1.96 \cdot \hat{SE}(\hat{\beta})] = [-0.132; 0.765]$

(2)

$$x'\beta = 3.134$$
$$P(y = 1 \mid x) = \frac{\exp(x'\beta)}{1 + \exp(x'\beta)} \cong 0.93209 = 93.209\%$$

We see that the estimated probability is higher than the average at ca. 90%. This is the case because ...?

(3) Logit-model:

$$\frac{\partial P(y_i = 1 \mid x_i)}{\partial x_{il}} = \frac{\exp(x_i'\beta)}{(1 + \exp(x_i'\beta))^2}\beta_l \cong -0.035$$

Linear model:

$$\frac{\partial P(y_i = 1 \mid x_i)}{\partial x_{il}} = \beta_l \cong -0.046$$

Probiy model:

$$\frac{\partial P(y_i = 1 \mid x_i)}{\partial x_{il}} \cong -0.037$$

We see that (in absolute terms), the linear model gives age the highest marginal effect, followed by the probit model. The logit gives age the lowest marginal effect.

Exercise 2

(1) Let $f(x_i, \beta) = \exp(\beta_1 + x_i\beta_2)$

$$\Rightarrow \nabla f(x_i,\beta) = \begin{bmatrix} \exp(\beta_1 + x_i\beta_2) \\ \exp(\beta_1 + x_i\beta_2)\, x_i \end{bmatrix} = \begin{bmatrix} f(x_i,\beta) \\ f(x_i,\beta)\, x_i \end{bmatrix}$$

$$\Rightarrow \nabla f(\cdot)\cdot(\nabla f(\cdot))' = \begin{bmatrix} f(x_i,\beta)^2 & f(x_i,\beta)^2 x_i \\ f(x_i,\beta)^2 x_i & f(x_i,\beta)^2 {x_i}^2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & x_i \\ x_i & x_i^2 \end{bmatrix} f(x_i,\beta)^2$$

From lecture & by heteroskedasticity:

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N\left(0, \sigma^2 \mathbb{E}\left\{ \begin{bmatrix} 1 & x_i \\ x_i & x_i^2 \end{bmatrix} f(x_i,\beta)^2 \right\}^{-1}\right)$$

(2) Assume efficient MoM, then we have

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N\left(0, \mathbb{E}\left(\frac{\partial f(x_i,\beta)}{\partial \beta}\right)^{-1} S \mathbb{E}\left(\frac{\partial f(x_i,\beta)}{\partial \beta}\right)^{-1}\right)$$

$$f(x_i,\beta) = \begin{bmatrix} y_i - \exp(\beta_1 + x_i\beta_2) \\ y_i - \exp(\beta_1 + x_i\beta_2)\, x_i \end{bmatrix} \equiv \begin{bmatrix} f_1(\cdot) \\ f_2(\cdot) \end{bmatrix}$$

$$\frac{\partial f(x_i,\beta)}{\partial \beta} = \begin{bmatrix} \partial f_1(\cdot)/\partial\beta_1 & \partial f_1(\cdot)/\partial\beta_2 \\ \partial f_2(\cdot)/\partial\beta_1 & \partial f_2(\cdot)/\partial\beta_2 \end{bmatrix}$$

$$= \begin{bmatrix} -\exp(\beta_1 + x_i\beta_2) & -\exp(\beta_1 + x_i\beta_2)\, x_i \\ -\exp(\beta_1 + x_i\beta_2)\, x_i & -\exp(\beta_1 + x_i\beta_2)\, x_i^2 \end{bmatrix}$$

$$\mathbb{E}\left[\frac{\partial f\left(x_i, \beta\right)}{\partial \beta}\right]^{-1} = \mathbb{E}\left[\begin{array}{cc} -\exp\left(\beta_1 + x_i\beta_2\right) & -\exp\left(\beta_1 + x_i\beta_2\right) x_i \\ -\exp\left(\beta_1 + x_i\beta_2\right) x_i & -\exp\left(\beta_1 + x_i\beta_2\right) x_i^2 \end{array}\right]$$

$$S = V\left[\ f\left(x_i, \beta\right)\ \right] = V\left[\begin{array}{c} y_i - \exp\left(\beta_1 + x_i\beta_2\right) \\ y_i - \exp\left(\beta_1 + x_i\beta_2\right) x_i \end{array}\right]$$

$$= V\left[\begin{array}{c} \varepsilon_i \\ \varepsilon_i x_i \end{array}\right] = \mathbb{E}\left[\begin{array}{cc} \varepsilon_i^2 & x_i\varepsilon_i^2 \\ x_i\varepsilon_i^2 & x_i^2\varepsilon_i^2 \end{array}\right] = \sigma^2\mathbb{E}\left[\begin{array}{cc} 1 & x_i \\ x_i & x_i^2 \end{array}\right]$$

Exercise 3

(1) Synthetic controls are used to see if a treatment had an effect on some aggregate outcome (e.g. on city level). Since no two cities (or places) are identical, one might struggle to find a perfect control. Therefore, one might construct a synthetic (i.e. artificial) control by averaging other cities' characteristics (using weights if desired).

(2) This is the average treatment effect for the complying subjects in a randomized experiment. One can estimate it using 2SLS where the treatment group is the instrument, and the treatment is the variable of interest. It is used if one is concerned with heterogenous & unobservable treatment effects as well as if there is a reason to believe that there may be noncompliers.

Exercise 4

(1) We assume that there are matches across treatment groups: $0 < \Pr(D = 1 \mid X) < 1$. Also, we must assume that conditional on X (which is age here), the treatment outcomes (Y_1, Y_0) are independent of D (which is assignment of treatment group).

(2)

$$\text{ATET} = 3.5$$

Treated		Untreated		Differences
Age	Y	Age	Y	ΔY
25	100	25	80	20
30	50	30	60	-10
35	40	35	40	0
40	40	40	32.5	7.5
45	25	45	25	0

5 Econometrics Final 2019 / 20

Watson

Exercise 1

(a)

$$\begin{aligned}
\varepsilon_t &= \frac{1-\phi L}{1-\theta L} y_t \\
&= (1-\phi L)\left(1+\theta L+(\theta L)^2+(\theta L)^3+\ldots\right) y_t \\
&= \left(1-\phi L+\theta L-\phi\theta L^2+(\theta L)^2-\phi\theta^2 L^3+(\theta L)^3-\phi\theta^3 L^4+\ldots\right) y_t \\
&= \left(1+(\theta-\phi)L+\theta(\theta-\phi)L^2+\theta^2(\theta-\phi)L^3+\ldots\right) y_t \\
&= y_t+(\theta-\phi)\left[y_{t-1}+\theta y_{t-2}+\theta^2 y_{t-3}+\ldots\right]
\end{aligned}$$

Thus, if we know ϕ and θ and all $y_{t-i} \forall i \geqslant 0$, were able to reconstruct ε_t.

(b)

$$\begin{aligned}
&\varepsilon_t \sim N(0,1) \\
&y_t \left|\varepsilon_t = \phi_{y_{t-1}}+\varepsilon_t\right| \varepsilon_t \sim N\left(\varepsilon_t, \frac{\phi^2}{1-\phi^2}\right) \\
&\Rightarrow \begin{pmatrix} y_t \\ \varepsilon_t \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{1-\phi^2} & 1 \\ 1 & 1 \end{bmatrix}\right) \\
&\mathbb{E}\left(\varepsilon_t \mid y_t=2\right)=0+2\left(1-\phi^2\right)=2(1-0.64)=0.72
\end{aligned}$$

Exercise 2

(a)

$$\begin{aligned} g_t &= (\rho x_{t-1} + e_t)(\phi u_{t-1} + \varepsilon_t) \\ &= \rho\phi(x_{t-1}u_{t-1}) + \phi u_{t-1}e_t + \rho x_{t-1}\varepsilon_t + e_t\varepsilon_t \\ &= \gamma g_{t-1} + a_t \end{aligned}$$

where $\gamma = \rho\phi; a_t = \phi u_{t-1}e_t + \rho x_{t-1}\varepsilon_t + e_t\varepsilon_t$

Now, also $a_{t-1} = \phi u_{t-2}e_{t-1} + \rho x_{t-2}\varepsilon_{t-1} + e_{t-1}\varepsilon_{t-1}$

Note, that $\mathbb{E}(a_t a_{t-1}) = 0$, as either e_t or ε_t will enter every part of the sum, and both are independent of the rest. This also holds for $\mathbb{E}(a_t a_{t-j})$, where $j \geqslant 1$. Therefore, a_t is serially uncorrelated.

(b)
- $\{y_t, x_t\}$ is ergodic & stationary by $|\rho| < 1$ & $|\phi| < 1$
- $\mathbb{E}(g_t) = 0$ since $\mathbb{E}(x_t) = \mathbb{E}(u_t) = 0$
- $\{g_t\}$ is not mds

(1) Construct estimator

$$\begin{aligned} \hat{\beta}_{OLS} &= \left(\frac{1}{T}\sum x_t^2\right)^{-1}\left(\frac{1}{T}\sum x_t y_t\right) = \left(\frac{1}{T}\sum x_t^2\right)^{-1}\left(\frac{1}{T}\sum \beta x_t^2 + g_t\right) \\ &= \beta + \left(\frac{1}{T}\sum x_t^2\right)^{-1}\left(\frac{1}{T}\sum g_t\right) \end{aligned}$$

(2) Show convergences

$$\left(\frac{1}{T}\sum x_t^2\right)^{-1} \xrightarrow{p} \mathbb{E}\left(x_t^2\right)^{-1} = 1 - \rho^2 \quad \text{by LLN \& CMT}$$

$$\frac{1}{\sqrt{T}}\sum g_t \xrightarrow{d} N(0, \Omega) \quad \text{by CLT}$$

(3) Combine (1) \$ (2):

$$\sqrt{T}(\hat{\beta} - \beta) \xrightarrow{d} N\left(0, \mathbb{E}\left(x_t^2\right)^{-1}\Omega\mathbb{E}\left(x_t^2\right)^{-1}\right)$$

(4) Find Ω :

First, the $MA(\infty)$ representation of g_t is

$$g_t = (1-\gamma L)^{-1} a_t$$

Therefore, the ACGF gives us

$$\Omega = \sum_{j=-\infty}^{\infty} \lambda_j = \left(\frac{1}{1-\gamma}\right)^2 \operatorname{Var}(a_t)$$

Now, we find $\operatorname{Var}(a_t)$:

$$\begin{aligned}
\operatorname{Var}(a_t) &= \mathbb{E}\left[(\phi u_{t-1} e_t + \rho x_{t-1} \varepsilon_t + e_t \varepsilon_t)^2\right] \\
&= \mathbb{E}\left[\phi^2 u_{t-1}^2 e_t^2 + \rho^2 x_{t-1}^2 \varepsilon_t^2 + e_t^2 \varepsilon_t^2\right] \text{ by } e_t \perp \varepsilon_t \\
&= \phi^2 \mathbb{E}\left[u_{t-1}^2\right] + \rho^2 \mathbb{E}\left[x_{t-1}^2\right] + 1 \\
&= \frac{\phi^2}{1-\phi^2} + \frac{\rho^2}{1-\rho^2} + 1
\end{aligned}$$

(5) Express V:

$$\begin{aligned}
V = \Omega \mathbb{E}\left(x_t^2\right)^{-2} &= \left(\frac{1}{1-\gamma}\right)^2 \left[\frac{\phi^2}{1-\phi^2} + \frac{\rho^2}{1-\rho^2} + 1\right] \left(1-\rho^2\right)^2 \\
&= \frac{\phi^2\left(1-\rho^2\right) + \rho^2\left(1-\phi^2\right) + \left(1-\rho^2\right)\left(1-\phi^2\right)}{\left(1-\phi^2\right)\left(1-\rho^2\right)(1-\gamma)^2} \left(1-\rho^2\right)^2 \\
&= \frac{1-\rho^2\phi^2}{\left(1-\phi^2\right)(1-\gamma)^2} \left(1-\rho^2\right) = \frac{(1-\gamma)(1+\gamma)}{\left(1-\phi^2\right)(1-\gamma)^2} \left(1-\rho^2\right) \\
&= \frac{1+\gamma}{1-\gamma} \frac{1-\rho^2}{1-\phi^2}
\end{aligned}$$

Exercise 3

(a)

$$\min_b \sum_{t=1}^{T} (y_t - x_t b)^2 + \lambda b^2$$

$$\min_b \sum_{t=1}^{T} y_t^2 - b \cdot 2 \sum_{t=1}^{T} y_t x_t + b^2 \sum_{t=1}^{T} x_t^2 + b^2 \lambda$$

$$FOC: \quad -2 \sum_{t=1}^{\top} y_t x_t + 2b \sum_{t=1}^{T} x_t^2 + 2b\lambda = 0$$

$$\Leftrightarrow b = \left(\lambda + \sum_{t=1}^{T} x_t^2\right)^{-1} \left(\sum_{t=1}^{T} y_t x_t\right) \equiv \tilde{\beta}$$

(b)

$$\begin{aligned}
\mathbb{E}\left(\tilde{\beta} \mid \{x_t\}_{t=1}^{T}\right) &= \mathbb{E}\left(\left(\lambda + \sum_{t=1}^{T} x_t^2\right)^{-1} \left(\sum_{t=1}^{T} x_t \left(x_t \beta + \varepsilon_t\right)\right) \mid \{x_t\}_{t=1}^{T}\right) \\
&= \mathbb{E}\left(\left(\lambda + \sum_{t=1}^{T} x_t^2\right)^{-1} \left(\sum_{t=1}^{T} x_t^2 \beta + x_t \varepsilon_t\right) \mid \{x_t\}_{t=1}^{T}\right) \\
&= \left(\lambda + \sum_{t=1}^{T} x_t^2\right)^{-1} \left(\sum_{t=1}^{T} x_t^2 \beta\right) \neq \beta \quad \text{if } \lambda \neq 0
\end{aligned}$$

Yes, the estimator is biased. λ "punishes" large values (in absolute terms) of $\tilde{\beta}$.

(c)

$$
\begin{aligned}
V\left(\tilde{\beta} \mid \{x_t\}_{t=1}^T\right) &= \mathbb{E}\left(\left(\tilde{\beta} - \mathbb{E}\left(\tilde{\beta} \mid \{x_t\}_{t=1}^T\right)\right)^2 \mid \{x_t\}_{t=1}^T\right) \\
&= \mathbb{E}\left(\left[\left(\lambda + \sum_{t=1}^T x_t^2\right)^{-1} \left(\sum_{t=1}^T x_t \varepsilon_t\right)\right]^2 \mid \{x_t\}_{t=1}^T\right) \\
&= \left(\lambda + \sum_{t=1}^T x_t^2\right)^{-2} \mathbb{E}\left(\left(\sum_{t=1}^T x_t \varepsilon_t\right)^2 \mid \{x_t\}_{t=1}^T\right) \\
&= \left(\lambda + \sum_{t=1}^T x_t^2\right)^{-2} \left(\sum_{t=1}^T x_t^2\right)
\end{aligned}
$$

For the OLS, we would have $V_{\text{OLS}} = \left(\sum_{t=1}^T x_t^2\right)^{-1}$. Thus, the ridge estimator has a lower variance!

Exercise 4

(a)

$$
\frac{\partial \Delta y_{t+h}}{\partial \varepsilon_t} = \begin{cases} 1 & \text{if } h = 0 \\ 0.4 & \text{if } h = 1 \\ 0.2 & \text{if } h = 2 \\ 0 & \text{else} \end{cases}
$$

(b) Take this with a grain of salt, I am not sure if I did this correctly.

$$
\begin{aligned}
y_t &= y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \\
&= y_{t-2} + \varepsilon_{t-1} + \theta_1 \varepsilon_{t-2} + \theta_2 \varepsilon_{t-3} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \\
&= y_{t-3} + \varepsilon_{t-2} + \theta_1 \varepsilon_{t-3} + \theta_2 \varepsilon_{t-4} + \varepsilon_{t-1} + \theta_1 \varepsilon_{t-2} + \theta_2 \varepsilon_{t-3} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \\
&= \dots \\
&= y_0 + \sum_{i=0}^{t-1} \varepsilon_{t-i} + \theta_1 \sum_{i=0}^{t-1} \varepsilon_{t-1-i} + \theta_2 \sum_{i=0}^{t-2} \varepsilon_{t-2-i}
\end{aligned}
$$

$$
\frac{\partial y_{t+h}}{\partial \varepsilon_t} = \frac{\partial y_t}{\partial \varepsilon_{t-h}} = \begin{cases} 1 & \text{if } h = 0 \\ 1 + \theta_1 & \text{it } h = 1 \\ 1 + \theta_1 + \theta_2 & \text{if } h > 1 \end{cases}
$$

Honoré

Exercise 1

(a) see table

treated		untreated		ATET	
age	outcome	age	outcome	Diff	$P(X \mid D = 1)$
25	80	25	100	-20	1/8
30	60	30	50	10	2/8
35	40	35	40	0	2/8
40	35	40	40	-5	2/8
45	25	45	25	0	1/8

$$\text{ATET} \;= \frac{1}{8}(-20 + 2 \cdot 10 - 2 \cdot 5) = -10/8 = -1.25$$

(b) That the outcome given age is independent of the treatment. I.e. there is no self-selection. Also, $\Pr(D \mid \text{age }) \in (0, 1)$, i.e. for all ages I can find observations in either group.

Exercise 2

(a) No. This is the reference category & its effect is included in the intercept. Including it would introduce perfect multicollinearity, breaking the model.

(b) $CI = [\hat{\beta} \pm 1.96 \cdot \hat{SE}(\hat{\beta})] = [0.118; 0.309]$

(c) Linear model:

$$
\begin{aligned}
P\left(Y_i=1 \mid x_i\right)=x_i^{\prime} \beta & \\
=40 \cdot 0.0402242 & \quad \text{age} \\
+40^2 \cdot(-0.0005327) & \quad \text{age}^2 \\
+1 \cdot 0.0260638 & \quad \text{white} \\
+1 \cdot 0.3038465 & \quad \text{college} \\
+(-0.1445535) & \quad \text{intercept} \\
\cong 94.200\% &
\end{aligned}
$$

Logit model:

$$
\begin{aligned}
P\left(y_i=1 \mid x_i\right) &= \frac{\exp \left(x_i^{\prime} \beta\right)}{1+\exp \left(x_i \beta\right)} \\
x_i^{\prime} \beta &= 40 \cdot 0.2134573 && \text{age} \\
&+40^2 \cdot(-0.002834) && \text{age}^2 \\
&+1 \cdot 0.1350566 && \text{white} \\
&+1 \cdot 1.499036 && \text{college} \\
&+(-3.335203) && \text{intercept} \\
P\left(y_i=1 \mid x_i\right) &\cong 90.911\%
\end{aligned}
$$

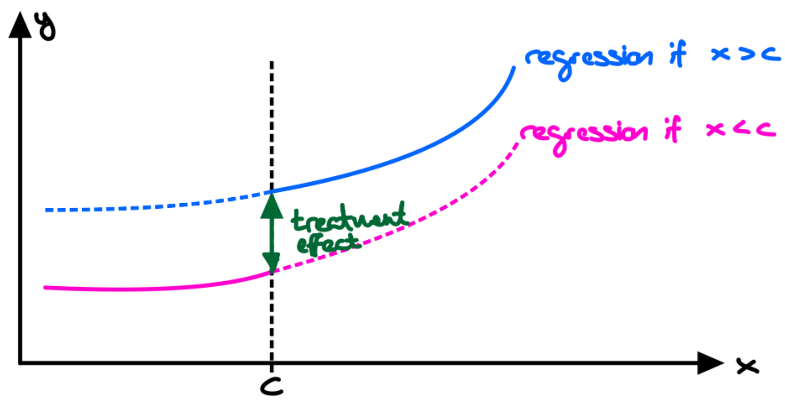
Probit model:

$$
\begin{aligned}
x_i^{\prime} \beta &= 40 \cdot 0.1267975 && \text{age} \\
&+40^2 \cdot(-0.0016807) && \text{age}^2 \\
&+1 \cdot 0.0979536 && \text{white} \\
&+1 \cdot 0.9029787 && \text{college} \\
&+(-2.010371) && \text{intercept} \\
P\left(y_i=1 \mid x_i\right) &= \Phi\left(x_i \beta\right) \cong \Phi(1.373)
\end{aligned}
$$

(d) Since the logic uses a non-liner function, one should use the delta method to find the estimators' distribution. Alternatively, one could obtain the CI on the coefficients $x_i'\beta$ and then apply $g\left(x_i'\beta\right) = \exp\left(x_i'\right) / \left(1 + \exp\left(x_i'\beta\right)\right)$ to the bounds, since $g(\cdot)$ is strictly monotone & its output is one-dimensional.

Exercise 3

The regression discontinuity makes sense, if a treatment is considered as x being greater than some cutoff c. The idea is that the relationship between x & y is different when $x \leqslant c$, than when $x > c$. One assumes that the two regressions continue smoothly in the counterfactual areas. The following graph helps to show the idea:



Example: Let x be time, c is 1989, and y the GDP growth in eastern Germany. Since the Berlin wall fell in 1989, it makes sense to model this as a regression discontinuity (ignoring the time series properties for the moment).

Exercise 4

(a)

$$f\left(x_i, b\right) = y_i - \exp\left(x_i b\right)$$
$$f'\left(x_i, b\right) = -x_i \cdot \exp\left(x_i b\right)$$
$$\sqrt{n}(b - \beta) \xrightarrow{d} N\left(0, \mathbb{E}\left(-x_i \exp\left(x_i \beta\right)\right)^{-2} \mathbb{E}\left(\exp\left(2 x_i \beta\right)\right)\right)$$

(b)

$$\frac{\partial f\left(x_i b\right)}{\partial b} = \begin{bmatrix} -\exp\left(x_i b\right) x_i \\ -\exp\left(x_i b\right) x_i^2 \end{bmatrix}$$

$$\sqrt{n}(b-\beta) \xrightarrow{d} N(0,\Sigma)$$

$$\Sigma = A^{-1} B' I_2 S I_2 B A^{-1}$$

$$A = \mathbb{E}\left(\frac{\partial f(x_i b)}{\partial b}\right)' I_2 \mathbb{E}\left(\frac{\partial f(x_i b)}{\partial b}\right)$$

$$= \mathbb{E}\left(\exp\left(x_i b\right) x_i\right)^2 + \mathbb{E}\left(\exp\left(x_i b\right) x_i^2\right)^2$$

$$B = \mathbb{E}\left(\frac{\partial f\left(x_i b\right)}{\partial b}\right) = \mathbb{E}\begin{bmatrix} -\exp\left(x_i b\right) x_i \\ -\exp\left(x_i b\right) x_i^2 \end{bmatrix}$$

$$S = \mathrm{Var}\left(f\left(x_i, \beta\right)\right) = \mathrm{Var}\begin{bmatrix} \varepsilon_i \\ \varepsilon_i x_i \end{bmatrix} = \mathbb{E}\begin{bmatrix} \varepsilon_i^2 & x_i\varepsilon_i^2 \\ x_i\varepsilon_i^2 & x_i^2\varepsilon_i^2 \end{bmatrix}$$

$$= \mathbb{E}\left[\mathbb{E}\begin{bmatrix} \varepsilon_i^2 & x_i\varepsilon_i^2 \\ x_i\varepsilon_i^2 & x_i^2\varepsilon_i^2 \end{bmatrix} \middle| x_i\right]$$

$$= \mathbb{E}\begin{bmatrix} \exp(2x_i\beta) & \exp(2x_i\beta)x_i \\ \exp(2x_i\beta)x_i & \exp(2x_i\beta)x_i^2 \end{bmatrix}$$

Done?

6 Econometrics Final 2020 / 21

Watson

Exercise 1

(a)

$$x_t = y_t + y_{t-2}$$

(1)

$$\mathbb{E}(x_t) = \mathbb{E}(y_t) + \mathbb{E}(y_{t-2}) = \mu_y + \mu_y = 2\mu_y \quad \forall t$$

(2)

$$\begin{aligned}
\text{Cov}(x_t, x_{t+k}) =& \text{Cov}(y_t + y_{t-2}, y_{t+k} + y_{t+k-2}) \\
=& \text{Cov}(y_t + y_{t-2}, y_{t+k}) + \text{Cov}(y_t + y_{t-2}, y_{t+k-2}) \\
=& \text{Cov}(y_t, y_{t+k}) + \text{Cov}(y_{t-2}, y_{t+k}) + \text{Cov}(y_t, y_{t+k-2}) + \text{Cov}(y_{t-2}, y_{t+k-2}) \\
=& \lambda_k + \lambda_{k+2} + \lambda_{k-2} + \lambda_k \\
=& 2\lambda_k + \lambda_{k+2} + \lambda_{k-2} \quad \forall t
\end{aligned}$$

Where $\lambda_k = \text{Cov}(y_t, y_{t+k})$ does not depend on t by stationarity of y_t. This concludes the proof.

(b) I do not believe that x_t is strictly stationary, since it is made up of stationary series: $x_t = y_t + y_{t-2}$.

Exercise 2

(a)

$$\begin{aligned}\mathbb{E}\left((\hat{x}_t - x_t)^2\right) &= \mathbb{E}\left[\left(\frac{1}{2}(2x_t + e_{1t} + e_{2t}) - x_t\right)^2\right] \\ &= \mathbb{E}\left[\left(\frac{1}{2}(e_{1t} + e_{2t})\right)^2\right] \\ &= \frac{1}{4}\left(\mathbb{E}\left(e_{1t}^2\right) + \mathbb{E}\left(e_{2t}^2\right) + 2\mathbb{E}\left(e_{1t}e_{2t}\right)\right) \\ &= \frac{1}{4}(1+4+0) = \frac{5}{4}\end{aligned}$$

(b)

$$\begin{aligned}&\mathbb{E}\left((\lambda_1 y_{1t} + \lambda_2 y_{2t} - x_t)^2\right) \\ =&\mathbb{E}\left((\lambda_1(x_t + e_{1t}) + \lambda_2(x_t + e_{2t}) - x_t)^2\right) \\ =&\mathbb{E}\left(((\lambda_1 + \lambda_2 - 1)x_t + \lambda_1 e_{1t} + \lambda_2 e_{2t})^2\right)\end{aligned}$$

Let $\lambda_1 + \lambda_2 = 1$, then:

$$\begin{aligned}\text{MSE} &= \mathbb{E}\left((\lambda_1 e_1 + \lambda_2 e_{2t})^2\right) \\ &= \lambda_1^2 + 4(1-\lambda_1)^2\end{aligned}$$

FOC :

$$\begin{aligned}&2\lambda_1 + 8(1-\lambda_1)(-1) = 0 \\ &\lambda_1 = 4/5 \longrightarrow \lambda_2 = 1/5\end{aligned}$$

Exercise 3

(a)

$$
\begin{aligned}
\hat{\beta} &= \left(\frac{1}{T}\sum_{t=1}^{T} x_t^2\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} x_t\left(\beta x_t + e_t\right)\right)\\
&= \underbrace{\left(\frac{1}{T}\sum_{t=1}^{T} x_t^2\right)^{-1}}_{\xrightarrow{p}\mathbb{E}\left(x_t^2\right)^{-1}}\underbrace{\left(\frac{1}{T}\sum_{t=1}^{T} x_t e_t\right)}_{\xrightarrow{p}\mathbb{E}(x_t e_t)} + \beta
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}\left(x_t^2\right)^{-1} &= 8/3\\
\mathbb{E}\left(x_t e_t\right) &= \mathbb{E}\left(0.5x_{t-1} + \varepsilon_t + \eta_t\right)\left(0.8e_{t-1} + \eta_t\right)\\
&= 0.4\mathbb{E}\left(x_{t-1}e_{t-1}\right) + \mathbb{E}\left(\eta_t^2\right) \neq 0
\end{aligned}
$$

Since x_t and e_t are correlated, $\hat{\beta}$ is inconsistent!

(b)

$$
\begin{aligned}
\tilde{\beta} &= \left(\frac{1}{T}\sum_{t=1}^{T} z_t x_t\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} z_t y_t\right)\\
&= \left(\frac{1}{T}\sum_{t=1}^{T} z_t x_t\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} z_t\left(\beta x_t + e_t\right)\right)\\
&= \beta + \left(\frac{1}{T}\sum_{t=1}^{T} z_t x_t\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} z_t e_t\right)\\
\sqrt{T}(\tilde{\beta} - \beta) &= \left(\frac{1}{T}\sum_{t=1}^{T} z_t x_t\right)^{-1}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_t e_t\right)
\end{aligned}
$$

$$
\left(\frac{1}{T}\sum_{t=1}^{T} z_t x_t\right)^{-1} \xrightarrow{p} \mathbb{E}\left(z_t x_t\right)^{-1} = \mathbb{E}\left(\left(\varepsilon_t + v_t\right)\left(0.5x_{t-1} + \varepsilon_t + \eta_t\right)\right)^{-1} = \mathbb{E}\left(\varepsilon_t^2\right)^{-1} = 1
$$

$$
\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_t e_t\right) \xrightarrow{d} N\left(0, \mathbb{E}\left(z_t^2\right)\mathbb{E}\left(e_t^2\right)\right) = N\left(0, 2\frac{1}{1-0.64}\right)
$$

In combination, this tells us (using Slutsky):

$$\sqrt{T}(\tilde{\beta}-\beta) \xrightarrow{d} N\left(0, \frac{2}{(1-0.64)}\right) = N(0, 5.6)$$

Exercise 4

(a)

$$P\left(\bar{y}^2 > \frac{\ln(T)}{T}\middle|\, y_t = \mu + \varepsilon_t\right) = P\left(\left(\frac{1}{T}\sum_{t=1}^{T} y_t\right)^2 > \frac{\ln(T)}{T}\middle|\, y_t = \mu + \varepsilon_t\right)$$
$$= P\left(\left(\frac{1}{T}\sum_{t=1}^{T} \mu + \varepsilon_t\right)^2 > \frac{\ln(T)}{T}\right) = P\left(\left(\mu + \frac{1}{T}\sum_{t=1}^{T} \varepsilon_t\right)^2 > \frac{\ln(T)}{T}\right)$$
$$= P\left(\left(\frac{\ln(T)}{T}\right)^{1/2} < \mu + \frac{1}{T}\sum_{t=1}^{T} \varepsilon_t\right) + P\left(\mu + \frac{1}{T}\sum_{t=1}^{T} \varepsilon_t < -\left(\frac{\ln(T)}{T}\right)^{1/2}\right)$$
$$= P\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \varepsilon_t < \sqrt{T}\mu - \ln(T)^{1/2}\right) + P\left(\frac{1}{T}\sum_{t=1}^{T} \varepsilon_t < -\ln(T)^{1/2} - \sqrt{T}\mu\right)$$

Since $\frac{1}{\sqrt{T}}\sum \varepsilon_t \Rightarrow N(0,1)$:

$$= \Phi\left(\sqrt{T}\mu - \ln(T)^{T/2}\right) + \Phi\left(-\ln(T)^{1/2} - \sqrt{T}\mu\right) \xrightarrow{T\to\infty} \Phi(\infty) + \Phi(-\infty) = 1$$

(b)

$$P\left(\bar{y}^2 < \frac{\ln(T)}{T}\middle|\, y_t = \varepsilon_t\right) = P\left(\left(\frac{1}{T}\sum_{t=1}^{T} y_t\right)^2 < \frac{\ln(T)}{T}\middle|\, y_t = \varepsilon_t\right)$$
$$= P\left(\left(\frac{1}{T}\sum_{t=1}^{1} \varepsilon_t\right)^2 < \frac{\ln(T)}{T}\right) = P\left(\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \varepsilon_t\right)^2 < \ln(T)\right)$$
$$\cong P\left(\chi_1^2 < \ln(T)\right) \xrightarrow{T\to\infty} 1$$

Honoré

Exercise 1

(1) MLE:

$$
\begin{aligned}
L &= \prod_{i=1}^{n} P\left(y = y_i \mid x_i\right) \\
&= \prod_{i=1}^{n} P\left(y = 1 \mid x_i\right)^{y_i} \left(1 - P\left(y = 1 \mid x_i\right)\right)^{1-y_i} \\
&= \prod_{i=1}^{n} \left[\frac{\exp\left(x_i'\beta\right)}{1+\exp\left(x_i'\beta\right)}\right]^{y_i} \left[\frac{1}{1+\exp\left(x_i'\beta\right)}\right]^{1-y_i} \\
l &= \sum_{i=1}^{n} y_i \ln\left(\frac{\exp\left(x_i'\beta\right)}{1+\exp\left(x_i'\beta\right)}\right) + \left(1-y_i\right)\ln\left(\frac{1}{1+\exp\left(x_i'\beta\right)}\right) \\
&= \sum_{i=1}^{n} y_i x_i'\beta - \ln\left(1+\exp\left(x_i'\beta\right)\right) \\
\frac{\partial l}{\partial b} &= \sum_{i=1}^{n} y_i x_i' - \frac{\exp\left(x_i'b\right)}{1+\exp\left(x_i'b\right)} x_i' \overset{!}{=} 0 \\
\frac{\partial^2 l}{\partial b^2} &= -\sum_{i=1}^{n} \left[\frac{\exp\left(x_i'b\right) x_i \left(1+\exp\left(x_i'b\right)\right) - \exp\left(x_i'b\right) x_i \exp\left(x_i'b\right)}{\left(1+\exp\left(x_i'b\right)\right)^2} x_i'\right] \\
&= -\sum_{i=1}^{n} \frac{\exp\left(x_i'b\right) x_i}{1+\exp\left(x_i'b\right)} \left(1 - \frac{\exp\left(x_i'b\right)}{1+\exp\left(x_i'b\right)}\right) x_i' \\
&= -\sum_{i=1}^{n} \frac{\exp\left(x_i'b\right)}{1+\exp\left(x_i'b\right)} \frac{1}{1+\exp\left(x_i'b\right)} x_i x_i' \\
-\mathbb{E}\left(\frac{\partial^2 e}{\partial b^2}\right)^{-1} &= n \cdot \mathbb{E}\left[\frac{\exp\left(x_i'b\right)}{1+\exp\left(x_i'b\right)} \frac{1}{1+\exp\left(x_i'b\right)} x_i x_i'\right]^{-1}
\end{aligned}
$$

Thus:

$$
\sqrt{\sqrt{n}(b-\beta) \xrightarrow{d} N\left(0, \mathbb{E}\left[\frac{\exp\left(x_i'b\right)}{1+\exp\left(x_i^ib\right)} \frac{1}{1+\exp\left(x_i'b\right)} x_i x_i'\right]^{-1}\right)}
$$

(2)

$$\begin{aligned}
\sqrt{n}(b-\beta) &\xrightarrow{d} N\left(O, A^{-1}BA^{-1}\right) \\
\frac{\partial f\left(x_i,\beta\right)}{\partial\beta} &= \frac{\exp\left(x_i'\beta\right)}{1+\exp\left(x_i'\beta\right)}\frac{1}{1+\exp\left(x_i'\beta\right)}x_i \\
A &= \mathbb{E}\left[\left(\frac{\partial f\left(x_i,\beta\right)}{\partial\beta}\right)\left(\frac{\partial f\left(x_i,\beta\right)}{\partial\beta}\right)'\right] = \mathbb{E}\left[\frac{\exp\left(2x_i'\beta\right)}{\left(1+\exp\left(x_i'\beta\right)\right)^4}x_ix_i'\right] \\
B &= \mathbb{E}\left[\mathbb{E}\left(\varepsilon_i^2 \mid x_i\right)\left(\frac{\partial f\left(x_i,\beta\right)}{\partial\beta}\right)\left(\frac{\partial f\left(x_i,\beta\right)}{\partial\beta}\right)'\right] \\
&= \mathbb{E}\left[\frac{\exp\left(x_i'\beta\right)}{\left(1+\exp\left(x_i'\beta\right)\right)^2}\frac{\exp\left(2x_i'\beta\right)}{\left(1+\exp\left(x_i'\beta\right)\right)^4}x_i'x_i'\right]
\end{aligned}$$

(3) $\sqrt{n}(b-\beta) \xrightarrow{d} N(0,V)$

Let it be efficient GMM : $V = (G'S^{-1}G)^{-1}$

$$\begin{aligned}
G &= \mathbb{E}\left(\frac{\partial f\left(x_i,\beta\right)}{\partial\beta}\right) = \mathbb{E}\left[\frac{\exp\left(x_i'\beta\right)}{\left(1+\exp\left(x_i'\beta\right)\right)^2}x_ix_i'\right] \\
S &= \mathbb{E}\left[\mathbb{E}\left(\varepsilon_i^2 \mid x_i\right)x_ix_i'\right] = \mathbb{E}\left[\frac{\exp\left(x_i'\beta\right)}{\left(1+\exp\left(x_i'\beta\right)\right)^2}x_ix_i'\right] \\
\sqrt{n}(b-\beta) &\xrightarrow{d} N\left(0, \mathbb{E}\left[\frac{\exp\left(x_i'\beta\right)}{\left(1+\exp\left(x_i'\beta\right)\right)^2}x_ix_i'\right]^{-1}\right)
\end{aligned}$$

(same as in (1))

(4) $\hat{\beta} \xrightarrow{p} \underset{b}{\text{argmax}}\mathbb{E}\left(\ln\left(f\left(x_i'b\right)\right)\right) \equiv \tilde{\beta} \neq \beta$

$$\sqrt{n}(\hat{\beta}-\tilde{\beta}) \xrightarrow{d} N\left(0, \mathbb{E}\left(\frac{\partial^2 \ln\left(f\left(x_i'\beta\right)\right)}{\partial\beta\partial\beta'}\right)^{-1} V\left(\tilde{\varepsilon}_i \mid x_i\right)\mathbb{E}\left(\frac{\partial^2 \ln\left(f\left(x_i'\beta\right)\right)}{\partial\beta\partial\beta'}\right)^{-1}\right)$$

It will be inconsistent but it will choose the "best" estimator in the class of $f(\cdot)$.

Exercise 2

(1) This will lead to issues since we are actually regressing y on its lagged values. This means we have an endogenous error term & the estimator is not consistent.

(2) We should take first differences:

t	regression equation
5	$\Delta y_{i5} = \Delta x'_{i5}\beta_1 + \Delta y_{i4}\beta_2 + \Delta\varepsilon_{i5}$
4	$\Delta y_{i4} = \Delta x'_{i4}\beta_1 + \Delta y_{i3}\beta_2 + \Delta\varepsilon_{i4}$
3	$\Delta y_{i3} = \Delta x'_{i3}\beta_1 + \Delta y_{i2}\beta_2 + \Delta\varepsilon_{i3}$

We can then use the following instruments:

t	instruments	
5	y_{i3}, y_{i2}, y_{i1}	$\{x_{is}\}_{s=1}^{5}$
4	y_{i2}, y_{i1}	$\{x_{is}\}_{s=1}^{5}$
3	y_{i1}	$\{x_{is}\}_{s=1}^{5}$

Note: cannot use forward looking instrument due to exogeneity constraint.

We must assume that the instruments are valid. The model is over-identified if there are more instruments than regressors.

Exercise 3

(1) Conditional on age, the assignment is independent of the outcomes that a person would have.

(2)

$$\begin{aligned}
ATE &= \mathbb{E}\left(Y_1 - Y_0\right) = \mathbb{E}\left(\mathbb{E}\left(Y_1 \mid X_1, D=1\right) - \mathbb{E}\left(Y_0 \mid X_1, D=0\right)\right) \\
&= \frac{3}{17}\left(\frac{100+80}{2} - 80\right) \\
&+ \frac{4}{17}\left(\frac{55+50}{2} - \frac{55+65}{2}\right) \\
&+ \frac{3}{17}\left(40 - \frac{50+30}{2}\right) \\
&+ \frac{4}{17}\left(\frac{40+35}{2} - \frac{45+20}{2}\right) \\
&+ \frac{3}{17}\left(\frac{20+25}{2} - 25\right) \\
&\cong 0.735
\end{aligned}$$

7 Econometrics Final 2021 / 22

Watson

Exercise 1

First, notice that by $y_0 = 0$ we know that

$$\begin{aligned} f(y_1) &= f(\varepsilon_1) \Rightarrow y_1 \sim N(0,1) \\ f(y_t \mid y_{t-1}) &= f(\varnothing y_{t-1} + \varepsilon_t \mid y_{t-1}) \Rightarrow y_t \mid y_{t-1} \sim N(\phi y_{t-1}, 1) \end{aligned}$$

$$\begin{aligned} f(Y_{1:50}, Y_{52:100}) =& f(Y_{52:100} \mid Y_{1:50}) f(Y_{1:50}) \\ =& f(Y_{52:100} \mid Y_{1:50}) \prod_{t=2}^{50} f(Y_t \mid Y_{t-1}) f(Y_1) \\ =& f(Y_{52:100} \mid Y_{1:50}) f(Y_1) \prod_{t=2}^{50} f(Y_t \mid Y_{t-1}) \\ =& f(Y_{52:100} \mid Y_{1:50}) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} Y_1^2\right) \prod_{t=2}^{50} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(Y_t - \phi Y_{t-1})^2\right) \\ =& \prod_{t=53}^{100} f(Y_t \mid Y_{t-1}, Y_{1:50}) f(Y_{52} \mid Y_{1:50}) \\ & \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} Y_1^2\right) \prod_{t=2}^{50} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(Y_t - \phi Y_{t-1})^2\right) \\ =& f(Y_{52} \mid Y_{1:s0}) \prod_{t=53}^{100} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(Y_t - \phi Y_{t-1})^2\right) \\ & \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} Y_1^2\right) \prod_{t=2}^{50} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(Y_t - \phi Y_{t-1})^2\right) \end{aligned}$$

At this point, we should find out how Y_{52} is distributed:

$$\begin{aligned} & Y_{52} = \phi Y_{51} + \varepsilon_{52} = \phi^2 Y_{50} + \phi \varepsilon_{51} + \varepsilon_{52} \\ \Rightarrow\ & Y_{52} \sim N\left(\phi^2 Y_{50}, 1 + \phi^2\right) \end{aligned}$$

Use this information in the expression above to find:

$$
\begin{aligned}
f\left(Y_{1:50}, Y_{52:100}\right) =& \frac{1}{\sqrt{2\pi\left(1+\phi^2\right)}} \exp\left(-\frac{1}{2\left(1+\phi^2\right)}\left(Y_{52}-\phi^2 Y_{50}\right)^2\right) \\
& \cdot \prod_{t=53}^{100} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(Y_t-\phi Y_{t-1}\right)^2\right) \\
& \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} Y_1^2\right) \prod_{t=2}^{50} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(Y_t-\phi Y_{t-1}\right)^2\right) \\
=& \left(\frac{1}{\sqrt{2\pi}}\right)^{99} \exp\left[-\frac{1}{2}\left(Y_1^2+\frac{\left(Y_{52}-\phi^2 Y_{50}\right)^2}{1+\phi^2}\right)\right] \\
& \cdot \prod_{t=2}^{50} \exp\left[-\frac{1}{2}\left(Y_t-\phi Y_{t-1}\right)^2\right] \cdot \prod_{t=53}^{100} \exp\left[-\frac{1}{2}\left(Y_t-\phi Y_{t-1}\right)^2\right]
\end{aligned}
$$

Exercise 2

$$
\begin{aligned}
y_t - \mu &= u_t \\
\frac{1}{T}\sum_{t=1}^{T} y_t - \mu &= \frac{1}{T}\sum_{t=1}^{T} u_t = \frac{1}{T}\left(\sum_{t=1}^{T/2} u_t + \sum_{t=T/2+1}^{T} u_t\right) \\
&= \frac{1}{T}\left(\sum_{t=1}^{T/2} u_t + \sum_{t=T/2+1}^{T} \varepsilon_t + \sum_{t=T/2+1}^{T} \varepsilon_{t-1}\right) \\
&= \frac{1}{T}\left(\sum_{t=1}^{T/2} u_t + \sum_{t=T/2+1}^{T} \varepsilon_t + \sum_{t=T/2+1}^{T} \varepsilon_t + \varepsilon_{T/2} - \varepsilon_T\right) \\
&= \frac{1}{T}\sum_{t=1}^{T/2} u_t + \frac{2}{T}\sum_{t=T/2+1}^{T} \varepsilon_t + \frac{\varepsilon_{T/2} - \varepsilon_T}{T} \\
\sqrt{T}(\bar{y} - \mu) &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T/2} u_t + \frac{2}{\sqrt{T}}\sum_{t=T/2+1}^{T} \varepsilon_t + \frac{\varepsilon_{T/2} - \varepsilon_T}{\sqrt{T}} \\
&= \frac{1}{\sqrt{T}}\frac{\sqrt{2}}{\sqrt{2}}\sum_{t=1}^{T/2} u_t + \frac{2}{\sqrt{T}}\frac{\sqrt{2}}{\sqrt{2}}\sum_{t=T/2+1}^{T} \varepsilon_t + \frac{\varepsilon_{T/2} - \varepsilon_T}{\sqrt{T}} \\
&= \frac{1}{\sqrt{2}}\underbrace{\frac{1}{(T/2)^{1/2}}\sum_{t=1}^{T/2} u_t}_{\xrightarrow{d} N(0,1)} + \sqrt{2}\underbrace{\frac{1}{(T/2)^{1/2}}\sum_{t=T/2+1}^{T} \varepsilon_t}_{\xrightarrow{d} N(0,1)} + \underbrace{\frac{\varepsilon_{T/2} - \varepsilon_T}{\sqrt{T}}}_{\xrightarrow{p} 0} \\
\sqrt{T}(\bar{y} - \mu) &\xrightarrow{d} N(0, \frac{1}{2} + 2) = N(0, \frac{5}{2})
\end{aligned}
$$

Exercise 3

(a)

$$
\begin{aligned}
y_t &= \varepsilon_{t+1}\beta + u_t \\
\hat{\beta} &= \left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{t+1}^2\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{t+1}\left(\varepsilon_{t+1}\beta + u_t\right)\right) \\
&= \beta + \left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{t+1}^2\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{t+1}v_t\right) \\
\hat{\beta} - \beta &= \left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{t+1}^2\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{t+1}u_t\right) \qquad (1)
\end{aligned}
$$

We know

$$
\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{t+1}^2\right)^{-1} \xrightarrow{P} \mathbb{E}\left(\varepsilon_{t+1}^2\right)^{-1} = 1 \qquad (2)
$$

Now, look at the other term. Note the following:

$$
\begin{aligned}
u_t &= u_{t-1} + \varepsilon_t \\
\sum u_t^2 &= \sum u_{t-1}^2 + 2\sum u_{t-1}\varepsilon_t + \sum \varepsilon_t^2 \\
\frac{1}{T}\sum u_{t-1}\varepsilon_t &= \frac{1}{2}\left[\frac{1}{T}\sum u_t^2 - \frac{1}{T}\sum u_{t-1}^2 - \frac{1}{T}\sum \varepsilon_t^2\right] \\
&= \frac{1}{2}\left[\underbrace{\frac{1}{T}u_T^2}_{\frac{1}{T}\left(\sum_{t=1}^{T}\varepsilon_t\right)^2} - \frac{1}{T}\sum \varepsilon_t^2\right] \xrightarrow{d} \frac{1}{2}\left(\chi_1^2 - 1\right) \qquad (3)
\end{aligned}
$$

Apply Slutsky in (1) using (2) & (3):

$$\hat{\beta}-\beta=\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{t+1}^{2}\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{t+1}u_{t}\right)\xrightarrow{d}\frac{1}{2}\left(\chi_{1}^{2}-1\right)$$

(b) First, note that (2) would change:

$$\left(\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{t+1}^{2}\right)^{-1}\xrightarrow{P}\mathbb{E}\left(\varepsilon_{t+1}^{2}\right)^{-1}=\frac{1}{5}$$

Second, (3) would change:

$$\frac{1}{T}\sum u_{t-1}\varepsilon_{t}\xrightarrow{d}\frac{1}{2}\left(s^{2}-5\right)\text{ where s}\sim\text{N}(0,5)$$

Thus:

$$\hat{\beta}-\beta\xrightarrow{d}\frac{1}{10}\left(s^{2}-5\right)=\frac{1}{2}\left(\chi_{1}^{2}-1\right)$$

Exercise 4

(a)

$$\begin{aligned}
y_t = x_t\beta + u_t \implies \quad & \sigma_t = x_t u_t \\
\mathbb{E}\left(\sigma_t \mid \Omega_{t-1}\right) &= \mathbb{E}\left(x_t \mid \Omega_{t-1}\right)\mathbb{E}\left(u_t \mid \Omega_{t-1}\right) \\
&= \mathbb{E}\left(e_t + \gamma e_{t-1} \mid \Omega_{t-1}\right)\mathbb{E}\left(\varepsilon_t + \theta\varepsilon_{t-1} \mid \Omega_{t-1}\right) \\
&= \gamma e_{t-1}\theta\varepsilon_{t-1} \neq 0
\end{aligned}$$

Thus, σ_t is not a MDS.

$$\sqrt{T}(\beta-\beta)=\left[\frac{1}{T}\sum x_{t}^{2}\right]^{-1}\left[\frac{1}{\sqrt{T}}\sum\sigma_{t}\right]$$

(1) $\left[\frac{1}{T}\sum x_t^2\right]^{-1}\xrightarrow{p}\Sigma_{xx}^{-1}=\mathbb{E}\left(x_t^2\right)^{-1}=\left(\sigma_e^2\left(1+\gamma^2\right)\right)^{-1}$

(2) $\left[\frac{1}{\sqrt{T}}\sum\sigma_t\right]\xrightarrow{d}N\left(0,\sum_{j=-\infty}^{\infty}\lambda_j\right)$ where λ_j is the j-th auto-covariance of σ_t.

$$
\begin{aligned}
\lambda_0 &= \mathbb{E}\left(\sigma_t^2\right) = \mathbb{E}\left(x_t^2 u_t^2\right) = \mathbb{E}\left(x_t^2\right)\mathbb{E}\left(v_t^2\right) = \sigma_e^2\left(1+\gamma^2\right)\sigma_\varepsilon^2\left(1+\theta^2\right) \\
\lambda_1 &= \lambda_{-1} = \mathbb{E}\left(x_t x_{t-1} u_t u_{t-1}\right) = \mathbb{E}\left(x_t x_{t-1}\right)\mathbb{E}\left(u_t u_{t-1}\right) \\
&= \gamma\mathbb{E}\left(e_{t-1}^2\right)\theta\mathbb{E}\left(\varepsilon_{t-1}^2\right) = \gamma\sigma_e^2\theta\sigma_\varepsilon^2 \\
\lambda_j &= \lambda_{-j} = 0 \quad \forall j \geqslant 2
\end{aligned}
$$

(3) by Slutsky:

$$
\begin{aligned}
&\left[\frac{1}{T}\sum x_t^2\right]^{-1}\left[\frac{1}{\sqrt{T}}\sum \sigma_t\right] \xrightarrow{d} N(0,V) \\
V &= \left(\sigma_e^2\left(1+\gamma^2\right)\right)^{-2}\left[2\gamma\sigma_e^2\theta\sigma_\varepsilon^2 + \sigma_e^2\left(1+\gamma^2\right)\sigma_\varepsilon^2\left(1+\theta^2\right)\right] \\
&= \frac{2\gamma\sigma_e^2\theta\sigma_\varepsilon^2 + \sigma_e^2\left(1+\gamma^2\right)\sigma_\varepsilon^2\left(1+\theta^2\right)}{\left(\sigma_e^2\left(1+\gamma^2\right)\right)^2}
\end{aligned}
$$

(b) We saw, that $\lambda_2 = \lambda_{-2} = 0$. Therefore, we ignore it.

$$
\begin{aligned}
CI_{gS} &= [\hat{\beta} \pm 1.96\sqrt{\hat{V}/T}] \\
&= \left[\hat{\beta} \pm 1.96\frac{1}{\sqrt{T}}\sqrt{\left(\frac{1}{T}\sum x_t^2\right)^{-2}\left(\hat{\lambda}_0 + 2\hat{\lambda}_1\right)}\right] \\
&= \left[\hat{\beta} \pm 1.96\frac{1}{\sqrt{T}}\sqrt{\left(\frac{1}{T}\sum x_t^2\right)^{-2}\left(\hat{\lambda}_0^x\hat{\lambda}_0^u + 2\hat{\lambda}_1^x\hat{\lambda}_1^u\right)}\right] \\
&= \left[2.1 \pm 1.96\frac{1}{10}\sqrt{5^{-2}(5\cdot 4 + 2\cdot 1\cdot 1.4)}\right] \\
&= \left[2.1 \pm 0.196\left(\frac{20+2.8}{25}\right)^{1/2}\right] \\
&= [1.913; 2.287]
\end{aligned}
$$

Honoré

Exercise 1

(1)

(a) $CI = [\hat{\beta} \pm 1.96 \cdot S\hat{E}(\hat{\beta})] = [-0.652; -0.214]$

(b) Since 0.2 is outside the CI, reject.

(2)

$$x_i = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \longrightarrow x_i'\beta \cong -2.189$$

$$P(y_i = 1 \mid x_i) = \frac{\exp(x_i'\beta)}{1 + \exp(x_i'\beta)} \cong 10.078\%$$

(3) Linear: $\frac{\partial P(y_x=1|x_1)}{\partial x_1} = \hat{\beta}_1 \cong 0.307$

Logit: $\frac{\partial P(y_x=1|x_1)}{\partial x_1} = \hat{\beta}_1 P(y_i = 1 \mid x_i) P(y_i = 0 \mid x_i) \cong 0.199$

Exercise 2

Use first differences. Also note that the exogeneity holds for forward looking instruments.

Differences:

$$\Delta y_{i4} = \gamma \cdot \Delta y_{i3} + \beta \cdot \Delta x_{i4} + \Delta\varepsilon_{i4}$$
$$\Delta y_{i3} = \gamma \cdot \Delta y_{i2} + \beta \cdot \Delta x_{i3} + \Delta\varepsilon_{i3}$$

Moment condition: (since $\mathbb{E}(x_{is}\Delta\varepsilon_{it}) = 0 \quad \forall t, s$)

$$\mathbb{E}(x_{is}(\Delta y_{i4} - \gamma \cdot \Delta y_{i3} - \beta \cdot \Delta x_{i4})) = 0; \quad s = 1, 2, 3, 4$$
$$\mathbb{E}(x_{is}(\Delta y_{i3} - \gamma \cdot \Delta y_{i2} - \beta \cdot \Delta x_{i3})) = 0; \quad s = 1, 2, 3, 4$$

Thus we have 8 moment conditions.

Exercise 3

This is a sequential estimator.

Let $f(X_i, \mu, \psi) = \begin{pmatrix} \mu - X_i \\ \psi_k - (X_i - \mu)^k \end{pmatrix}$

Define

$$R_1 = \mathbb{E}\left[\frac{\partial\left(\psi_k - (X_i - \mu)^k\right)}{\partial \mu}\right] = \mathbb{E}\left[k (X_i - \mu)^{k-1}\right] = k\mathbb{E}\left[(X_i - \mu)^{k-1}\right]$$

As we saw in the lecture, $R_1 = 0$ would be sufficient for the limiting distributions to be the same. If $\{X_i\}_{i=1}^n$ follow a symmetric distribution, then $R_1 = 0$ for all even k. For odd k, we would need to correct, i.e. the distributions won't be identical!

Exercise 4

(a) $\sqrt{n}(\hat{\rho} - \beta) \xrightarrow{d} N(0, A^{-1}BA^{-1})$

where $A = \mathbb{E}\left[(\exp(x_i\beta) x_i)^2\right] = \mathbb{E}\left[\exp(2x_i\beta) x_i^2\right]$

$$\begin{aligned} B &= \mathbb{E}\left[\varepsilon_i^2 \exp(2x_i\beta) x_i^2\right] \\ &= \mathbb{E}\left[\mathbb{E}\left(\varepsilon_i^2 \mid x_i\right) \exp(2x_i\beta) x_i^2\right] \\ &= \mathbb{E}\left[\exp(x_i\beta) \cdot \exp(2x_i\beta) x_i^2\right] \\ &= \mathbb{E}\left[\exp(3x_i\beta) x_i^2\right] \end{aligned}$$

(b) $\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, G^{-1}SG^{-1})$

where

$$
\begin{aligned}
G &= \mathbb{E}\left[-y_i x_i \exp\left(-x_i\beta\right)\right] \\
&= \mathbb{E}\left[-\mathbb{E}\left(y_i \mid x_i\right) x_i \exp\left(-x_i\beta\right)\right] \\
&= \mathbb{E}\left[-x_i\right] \\
S &= V\left(f\left(x_i,\beta\right)\right) = V\left(y_i \exp\left(-x_i\beta\right)\right) \\
&= \mathbb{E}\left[y_i^2 \exp\left(-2x_i\beta\right)\right] - \mathbb{E}\left[y_i \exp\left(-x_i\beta\right)\right]^2 \\
&= \mathbb{E}\left[\mathbb{E}\left(y_i^2 \mid x_i\right) \exp\left(-2x_i\beta\right)\right] - \mathbb{E}\left[\mathbb{E}\left(y_i \mid x_i\right) \exp\left(-x_i\beta\right)\right]^2 \\
&= \mathbb{E}\left[\left(\exp\left(x_i\beta\right) + \exp\left(2x_i\beta\right)\right) \exp\left(-2x_i\beta\right)\right] - \mathbb{E}[1]^2 \\
&= \mathbb{E}\left[\exp\left(-x_i\beta\right) + 1\right] - 1 = \mathbb{E}\left[\exp\left(-x_i\beta\right)\right]
\end{aligned}
$$

Exercise 5

When (Y_1, Y_0), i.e. the outcome, is independent of D conditional on X, then it is also independent of D conditional an $P(X)$. Therefore, one can also match based on $P(x)$ instead of x.

This does not rely on a functional form or parametric assumptions for identification.