Assignment 2 in Methods:

Using parts of old assignments:

Question 1

a

eststo: reg lnc lnp lny, cluster(state)
Results can be seen in Table 1

b

quietly tab year, gen(period)
eststo: xtreg lnc lnp lny period*,fe cluster(state)
eststo: reg d.lnc d.lnp d.lny d.period*, cluster(state)
xtserial lnc lnp lny

$$lnC_{it} = \alpha + \beta_1 lnP_{it} + \beta_2 lnY_{it} + \gamma_2 D_{2i} + ... + \gamma_{46} D_{46i} + \delta_1 T1_t + ... + \delta_{29} T29_t + u_{it}$$

In order to account for time effects dummy variables (T1-T29) for each but one year are included. To include state effects there are several options. Firstly dummy variables (D2-D46) for each state can be added to the regression. Equivalently, a fixed effect within transformation can be applied to get rid of unobserved state-specific effects. If the state-specific effects can be argued to be uncorrelated with the included regressors, a random effects model can be estimated. In fact, the Hausmann test does no reject the RE model. Since the RE model yields very similar estimates we will stick with the FE model for comparative purposes.

Given that smoking is an addiction it is plausible that there might be serial correlation, which is confirmed by a test for serial correlation in the residuals obtained from the FD regression. Given substantial serial correlation, the FD regression should generally be preferred.

e

eststo: xtreg lnc lnp lny period* gln*, cluster(state)
eststo: reg d.lnc d.lnp d.lny d.period* d.gln*, cluster(state)
As argued in the paper by Baltagi & Li (1999) when modelling the demand of cigarettes it is appropriate to take into account spatial correlation across states due to border effect purchases. Further, interaction between states can result

in correlated outcome. Over this suggests that the error terms are correlated across states. To account for this correlation amongst states, interaction variables that reflect endogenous and exogenous interaction effects can be added to the regression.

Endogenous effects suggest that the demand of a state is influenced by its neighbours' outcomes. We are trying to capture this effect by including *glnc*, which, for each individual state, represents the average cigarette demand of those other states it is related to as determined by the matrix D. This effect is likely to be endogenous as there is a feedback process taking place between neighbours, since links between states should be reciprocal given the spatial weight matrix. However, only the final outcome of the feedback process will be measured and it is impossible to differentiate the causal affect of a particular state j on i and vice versa. In order to make out this causal impact instrumental variables are a solution. Availability of instrumental variables is determined by the existence of intransitive triads in the network structure. As each individual state has its own group that it interacts with, it should be possible to find a state k that is related to state j but not i such that $G_{ik} = 0$ and $G_{ij} \cdot G_{jk} > 0$. Exogenous characteristics of this state k can then be used as an instrument for j and thus the causal effect of j on i can be isolated.

With instrumental variables, it is also possible to differentiate exogenous from endogenous effects and thus to break the reflection problem. Exogenous interaction effects refer to correlated outcomes as j's socioeconomic characteristics impact on i's outcome, vice versa.

Question 2

a

$$lnC_{it} = \alpha + \beta_1 lnP_{it} + \beta_2 lnY_{it} + \gamma_2 D_{2i} + ... + \gamma_{46} D_{46i} + \delta_1 T1_t + ... + \delta_{29} T29_t + u_{it}$$

$$
\begin{aligned}
E[y_i|a_i^l = 0, s_i = 0] &= \alpha - \gamma\frac{\pi_0}{\pi_2} \\
E[y_i|a_i^l = 0, s_i = 1] &= \alpha - \gamma\frac{\pi_0}{\pi_2} + \beta - \frac{1}{\pi_2}(\gamma\pi_1 + \tau(\pi_0 + \pi_1)) \\
E[y_i|a_i^l = 1, s_i = 0] &= \alpha - \gamma\frac{\pi_0}{\pi_2} + \frac{\gamma}{\pi_2} \\
E[y_i|a_i^l = 1, s_i = 1] &= \alpha + \beta - \frac{1}{\pi_2}(\gamma(\pi_0 + (\pi_1 - 1) + \tau(\pi_0 + \pi_1 - 1))
\end{aligned}
$$

b

$$E[y_i|a_i^l=0,s_i=1]-E[y_i|a_i^1=0,s_i=0]=\beta-\gamma\frac{\pi_1}{\pi_2}-\frac{\tau}{\pi_2}(\pi_0+\pi_1)$$

If $\pi_0=\pi_1=0$,

$$E[y_i|a_i^l=0,s_i=1]-E[y_i|a_i^1=0,s_i=0]=\beta-\gamma\frac{0}{\pi_2}-\frac{\tau}{\pi_2}(0+0)=\beta$$

Proxy controls partially account for OVB, which is the main reason for including them in a regression. Yet, they are simultaneously determined by the regressor of interest. This is also the case in our example where late ability is causally related to the omitted variable ability, which is meant to represent, yet it is equally a function of the regressor high school education. Including proxy controls will thus lead to biased estimates as highlighted by the equations in part a (though this bias might be an improvement over the bias present when not at all accounting for OVB). The restriction suggests that the proxy variable late ability is neither causally affected by high school education, so it is not itself an intermediate outcome, nor is it subject to a general shift which would be reflected in the constant. Instead, this assumption implies that the proxy variable late ability is purely a function of the latent variable, ability a_i, that we are trying to capture with it. Given this assumption it is possible to identify the direct effect of completed high school education on future earnings.

$$\begin{aligned}
\text{If } -\frac{\tau}{\gamma}=\frac{\pi_1}{\pi_0+\pi_1}&\Rightarrow -\tau=\frac{\gamma\pi_1}{\pi_0+\pi_1}\\
E(y_i|\alpha_i^l=0,s_i=1)-E(y_i|\alpha_i^l=0,s_i=0)&=\beta-\gamma\frac{\pi_1}{\pi_2}+\frac{\gamma\pi_1}{\pi_0+\pi_1}\cdot\frac{\pi_0+\pi_1}{\pi_2}\\
&=\beta-\gamma\frac{\pi_1}{\pi_2}+\gamma\frac{\pi_1}{\pi_2}\\
&=\beta
\end{aligned}$$

Question 3 - extra

Let's see what nice more math and pictures I can find:

$max\ u_1(x)$ s.t. $p_1x_1 + p_2x_2 \leq w_1$

$L(x_1, x_2, p_1, p_2, w_1) = x_1^{3/4}x_2^{1/4} - \lambda(p_1x_1 + p_2x_2 - w_1)$

FOC

$\dfrac{\partial L}{\partial x_1} = \dfrac{3}{4}x_1^{-1/4}x_2^{1/4} - \lambda p_1 = 0$

$\dfrac{\partial L}{\partial x_2} = \dfrac{1}{4}x_1^{3/4}x_2^{-3/4} - \lambda p_2 = 0$

$\dfrac{\partial L}{\partial x_1} / \dfrac{\partial L}{\partial x_2} := \dfrac{\frac{3}{4}x_1^{-1/4}x_2^{1/4}}{\frac{1}{4}x_1^{3/4}x_2^{-3/4}} = \dfrac{\lambda p_1}{\lambda p_2}$

$\iff 3x_1^{-1}x_2 = \dfrac{p_1}{p_2}$

$\iff x_2 = \dfrac{x_1p_1}{3p_2}, \quad x_1 = \dfrac{x_2p_23}{p_1}$

$p_1x_1 + p_2x_2 = w_1 \iff p_1x_1 + p_2\dfrac{x_1p_1}{3p_2} = w_1$

$\iff p_1x_1 + \dfrac{x_1p_1}{3} = w_1$

$\iff x_1(1 + \dfrac{1}{3}) = \dfrac{w_1}{p_1}$

$\iff x_1 = \dfrac{3w_1}{4p_1} = x_{1,1}\ ,\ (x_{good,consumer})$

$$p_1x_1 + p_2x_2 = w_1 \iff p_1\frac{x_2p_23}{p_1} + p_2x_2 = w_1$$

$$\iff x_2 = \frac{w_1}{4p_2} = x_{2,1}\ ,\ (x_{good,consumer})$$

$max\ u_1(x)$ s.t. $p_1x_1 + p_2x_2 \leq w_1$

$L(x_1, x_2, p_1, p_2, w_1) = x_1^{3/4}x_2^{1/4} - \lambda(p_1x_1 + p_2x_2 - w_1)$

FOC

$$\frac{\partial L}{\partial x_1} = \frac{3}{4}x_1^{-1/4}x_2^{1/4} - \lambda p_1 = 0$$

$$\frac{\partial L}{\partial x_2} = \frac{1}{4}x_1^{3/4}x_2^{-3/4} - \lambda p_2 = 0$$

$$\frac{\partial L}{\partial x_1} / \frac{\partial L}{\partial x_2} := \frac{\frac{3}{4}x_1^{-1/4}x_2^{1/4}}{\frac{1}{4}x_1^{3/4}x_2^{-3/4}} = \frac{\lambda p_1}{\lambda p_2}$$

$$\iff 3x_1^{-1}x_2 = \frac{p_1}{p_2}$$

$$\iff x_2 = \frac{x_1p_1}{3p_2}, \quad x_1 = \frac{x_2p_23}{p_1}$$

$$p_1x_1 + p_2x_2 = w_1 \iff p_1x_1 + p_2\frac{x_1p_1}{3p_2} = w_1$$

$$\iff p_1x_1 + \frac{x_1p_1}{3} = w_1$$

$$\iff x_1(1 + \frac{1}{3}) = \frac{w_1}{p_1}$$

$$\iff x_1 = \frac{3w_1}{4p_1} = x_{1,1}\ ,\ (x_{good,consumer})$$

$$p_1x_1 + p_2x_2 = w_1 \iff p_1\frac{x_2p_23}{p_1} + p_2x_2 = w_1$$

$$\iff x_2 = \frac{w_1}{4p_2} = x_{2,1}\ ,\ (x_{good,consumer})$$

(i)
$D_px_i(p, w_i)$ is negative semidefinite if $x_i(p, w_i)$ satisfies ULD.
Hence, if $(p' - p)[x_i(p', w_i) - x_i(p, w_i)] \leq 0$, $x_i(p, w_i)$ satisfies ULD.
Let $(p' - p) = [0, ..., 0, p'_i - p_i, 0, ..., 0] = dp > 0$ and $\exists\epsilon,\ dp < \epsilon\ \forall\ \epsilon$.
Then, we have $(p' - p)[x_i(p', w_i) - x_i(p, w_i)] = dp' \frac{\partial x_i(p,w_i)}{\partial p_i} dp \leq 0$.
Hence $D_px_i(p, w_i)$ is negative semidefinite $\forall p$
(ii)
Show that $x_i(p, w_i)$ satisfies the uncompensated law of demand if $D_px_i(p, w_i)$ is negative definite $\forall p$.
Let $p(\lambda) = (1 - \lambda)p + \lambda p' > p,\ \lambda \in (0, 1)$

The question ask to transfer initial income, here denoted by the numeraire m, so that the product of the utilities is maximized.

$$\max \Pi_{i=1}^{4} u_i \quad s.t. \sum_{i=1}^{4} u_i = 130$$

$$L = (m_1 + 2 \cdot 1\sqrt{x_1})(m_2 + 2 \cdot 2\sqrt{x_2})(m_3 + 2 \cdot 3\sqrt{x_3})(m_4 + 2 \cdot 4\sqrt{x_4})$$
$$- \lambda(m_1 + m_2 + m_3 + m_4 - 100)$$

FOC (everything with equality because we know that everybody consumes) :

$$\frac{\partial L}{\partial m_1} := (m_2 + 2 \cdot 2\sqrt{x_2})(m_3 + 2 \cdot 3\sqrt{x_3})(m_4 + 2 \cdot 4\sqrt{x_4}) - \lambda = 0$$
$$\frac{\partial L}{\partial m_2} := (m_1 + 2 \cdot 1\sqrt{x_1})(m_3 + 2 \cdot 3\sqrt{x_3})(m_4 + 2 \cdot 4\sqrt{x_4}) - \lambda = 0$$
$$\frac{\partial L}{\partial m_3} := (m_2 + 2 \cdot 2\sqrt{x_2})(m_1 + 2 \cdot 1\sqrt{x_1})(m_4 + 2 \cdot 4\sqrt{x_4}) - \lambda = 0$$
$$\frac{\partial L}{\partial m_4} := (m_2 + 2 \cdot 2\sqrt{x_2})(m_3 + 2 \cdot 3\sqrt{x_3})(m_1 + 2 \cdot 1\sqrt{x_1}) - \lambda = 0$$
$$\frac{\partial L}{\partial \lambda} := m_1 + m_2 + m_3 + m_4 - 100 = 0$$
$$\frac{\partial L/\partial m_1}{\partial L/\partial m_2} = \frac{m_2 + 2 \cdot 2\sqrt{x_2}}{m_1 + 2 \cdot 1\sqrt{x_1}} = 1 \iff m_2 + 2 \cdot 2\sqrt{x_2} = m_1 + 2 \cdot 1\sqrt{x_1}$$

We further get:

$$m_1 + 2 \cdot 1\sqrt{x_1} = m_2 + 2 \cdot 2\sqrt{x_2} = m_3 + 2 \cdot 3\sqrt{x_3} = m_4 + 2 \cdot 4\sqrt{x_4}$$

The calculation of the individual i's choice of x does not depend on the numeraire in the first subsection, so I use this results here.

$$m_1 + 2 = m_2 + 8 = m_3 + 18 = m_4 + 32$$

Initially, the distribution of the numeraire was given by:

$$m_1 = 10,\ m_2 = 20,\ m_3 = 30,\ m_4 = 40$$

Now endowments have to be changed so that the condition in equation (1) holds:

$$m_1^n = 38,\ m_2^n = 32,\ m_3^n = 22,\ m_4^n = 8$$

Then, the amount of taxes is just the difference between the initial m_i and the nash bargaining m_i^n:

$$tax_1 = m_1 - m_1^n = -28,\ tax_2 = m_2 - m_2^n = -12,$$
$$tax_3 = m_3 - m_3^n = 8$$
$$,\ tax_4 = m_4 - m_4^n = -32$$

Negative taxes mean that the individual obtains money and positives taxes mean that the individual has to pay taxes.
We than get

$$W^n(u_i^n) = \Pi_{i=1}^{4}(u_i^n)^{\alpha_i} = (m_1^n + 2)^{\alpha_1}(m_2^n + 8)^{\alpha_2}(m_3^n + 18)^{\alpha_3}(m_4^n + 32)^{\alpha_4} = 40 > 37.44$$

Partial derivatives of the expenditure function with respect to prices

$$\begin{aligned}
h_1(p,u) &= \frac{\partial e(p,w)}{\partial p_1} = \frac{u}{2} \\
h_2(p,u) &= \frac{\partial e(p,w)}{\partial p_2} = u \\
\rightarrow h(p,u) &= \begin{pmatrix} \frac{u}{2} \\ u \end{pmatrix}
\end{aligned}$$

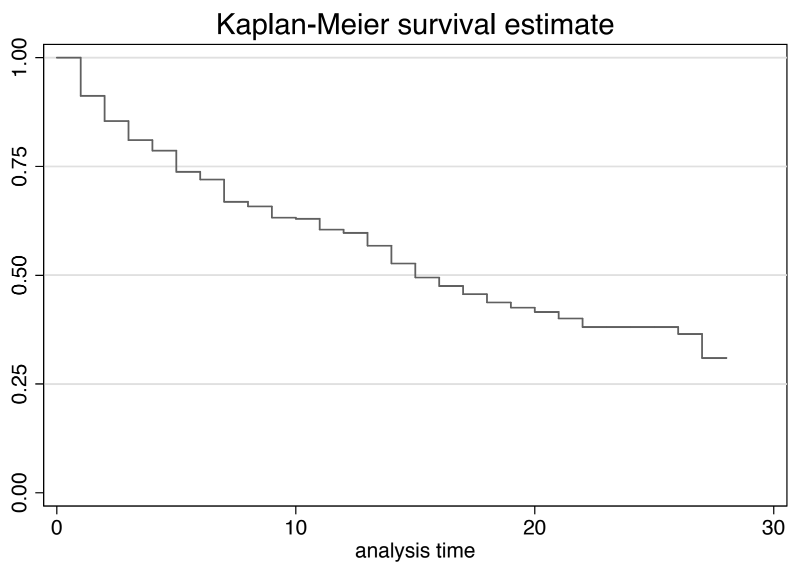

Figure 1: *Notes:* The survivor function does not go to zero by the end of the time frame. This is due to censoring at the right. Some people still did not find a new job.

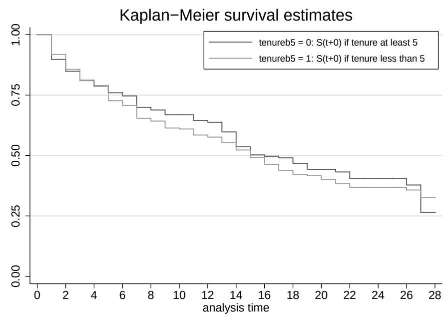

Figure 2: *Notes:* And since it makes fun, I will add one more, with a different width

Figure 3: