Econometrics

The answer keys below are suggested solutions, and students may propose alternative answers. As long as they provide valid arguments and well-founded equations, (partial) credits may be awarded at the grader's own judgement.

Question 10 (Econometrics, 60 points). There has been recent work using supermarket scanner data to form longitudinal data sets to study consumer behavior. For example, the following paper was presented in March at the Harvard-MIT Industrial Organization Workshop: "State Dependence and Alternative Explanations for Consumer Inertia." From the abstract: "For many consumer packaged goods products, researchers have documented a form of state dependence whereby consumers become "loyal" to products they have consumed in the past. ... However, it has not been established that this form of state dependence can be identified in the presence of consumer heterogeneity of an unknown form. Most importantly, before this inertia can be given a structural interpretation and used in policy experiments such as counterfactual pricing exercises, alternative explanations which might give rise to similar consumer behavior must be ruled out."

Here is a simple version of the problem. We observe $Y_{it} = 1$ if household i purchases a particular product at date t, and $Y_{it} = 0$ otherwise ($i = 1, \ldots, N$; $t = 1, \ldots, T$). There are $N = 500$ households and $T = 9$ observations per household.

a. (**20 points**) Suppose that the conditional expectation function

$$E(Y_{it} \,|\, Y_{i,t-1})$$

does not depend upon t. Provide an explicit estimator for this function.

Answer:

Because $Y_{i,t-1}$ is binary, the most general functional form is

$$E(Y_{it} \,|\, Y_{i,t-1}) = \alpha + \beta Y_{i,t-1}.$$

Estimate α from the fraction of observations with $Y = 0$ that are followed by $Y = 1$; estimate $\alpha + \beta$ by the fraction of observations with $Y = 1$ that are followed by $Y = 1$.

$$\hat{\alpha} = \sum_{i=1}^{N} \sum_{t=2}^{T} 1(Y_{it} = 1, Y_{i,t-1} = 0) \Big/ \sum_{i=1}^{N} \sum_{t=2}^{T} 1(Y_{i,t-1} = 0),$$

$$\hat{\beta} = \sum_{i=1}^{N} \sum_{t=2}^{T} 1(Y_{it} = 1, Y_{i,t-1} = 1) \Bigg/ \sum_{i=1}^{N} \sum_{t=2}^{T} 1(Y_{i,t-1} = 1) - \hat{\alpha}.$$

Points:
5 points for noting the general functional form $E(Y_{it} \,|\, Y_{i,t-1}) = \alpha + \beta Y_{i,t-1}$.
5 points for getting $\hat{\alpha}$ correct.
5 points for getting $\hat{\beta}$ correct.
5 points for a well-illustrated answer.
Partial credit at the grader's own judgement.

b. (**20 points**) Now bring in the heterogeneity issue. Suppose that there is a household-specific probability p_i of purchasing the product, and that, conditional on p_i, the random variables $(Y_{i1}, \ldots, Y_{iT})$ are mutually independent, so that

$$\text{Prob}(Y_{it} = 1 \,|\, Y_{i,t-1}, p_i) = p_i.$$

Then there is no "loyalty" (state dependence) once we condition on p_i. Explain why the conditional expectation function in (a) will nevertheless show a dependence on $Y_{i,t-1}$.

Answer:
The slope coefficient in the linear predictor is

$$\beta = \text{Cov}(Y_{it}, Y_{i,t-1})/\text{Var}(Y_{i,t-1}).$$

Use iterated expectations:

$$\begin{aligned} \text{Cov}(Y_{it}, Y_{i,t-1}) &= E(Y_{it}Y_{i,t-1}) - E(Y_{it})E(Y_{i,t-1}) \\ &= E[E(Y_{it}Y_{i,t-1} \,|\, p_i)] - E(p_i)E(p_i) \\ &= E(p_i^2) - [E(p_i)]^2 = \text{Var}(p_i) > 0, \end{aligned}$$

unless there is no heterogeneity.

Points:
20 points for a well-illustrated answer.
Partial credit at the grader's own judgement.

c. (**20 points**) We do not have data on p_i but want to know whether some state dependence would remain if we could condition on p_i. We do have data on an indicator variable X with $X_{it} = 1$ if household i faced a discount price or special promotion for the product at date t, and $X_{it} = 0$ otherwise. Explain how this data could be used to determine whether or not there is state dependence, allowing for heterogeneity.

Answer:

If there is state dependence, then a sale or promotion in the past should continue to affect current behavior, controlling for whether or not there is currently a sale or promotion. So calculate the magnitudes of

$$\sum_{i=1}^{N}\sum_{t=2}^{T} 1(Y_{it}=1, X_{it}=1, X_{i,t-1}=1) \Big/ \sum_{i=1}^{N}\sum_{t=2}^{T} 1(X_{it}=1, X_{i,t-1}=1)$$
$$-\sum_{i=1}^{N}\sum_{t=2}^{T} 1(Y_{it}=1, X_{it}=1, X_{i,t-1}=0) \Big/ \sum_{i=1}^{N}\sum_{t=2}^{T} 1(X_{it}=1, X_{i,t-1}=0)$$

and of

$$\sum_{i=1}^{N}\sum_{t=2}^{T} 1(Y_{it}=1, X_{it}=0, X_{i,t-1}=1) \Big/ \sum_{i=1}^{N}\sum_{t=2}^{T} 1(X_{it}=0, X_{i,t-1}=1)$$
$$-\sum_{i=1}^{N}\sum_{t=2}^{T} 1(Y_{it}=1, X_{it}=0, X_{i,t-1}=0) \Big/ \sum_{i=1}^{N}\sum_{t=2}^{T} 1(X_{it}=0, X_{i,t-1}=0).$$

Economically significant magnitudes would be consistent with state dependence. (Question does not ask for details on providing standard errors.)

Points:

20 points for a well-illustrated answer.
Partial credit at the grader's own judgement.