

# AAEC 6312 Homework 2

## Due November 6, 2017

October 28, 2017

1. **Problem 1: Consumer's Behavior:** The accompanying data describes the weekly buying behavior of consumers in Pittsfield, Massachusetts and Eau Claire, Wisconsin for coffee across different supermarket chains in each city from June 18, 2001 to July 30, 2006. Your task is to analyze the effect of the factors that affect consumer's brand and chain choice. To start your analysis, you decide to pick only one city and one week (pick a week that gives you at least 200 observations). The data (Homescan<sup>1</sup> data from Information Resource infoscan, IRI) consist of panelist id, panelist income (category), race, and household size; the week, the number of units purchased by the panelist, the dollar amount spent (this, is not the price. The price can be obtained using the dollar amount, the units, and the volume equivalent), universal product code (COLUPC), the city, the chain, the volume equivalent,<sup>2</sup> and the brand. In what follows, we assume that brand  $j$  purchased in chain  $k$  is different than brand  $j$  purchased in chain  $l$ , i.e., there is differentiation at the brand and chain level. Perform the following analysis:
  - (a) Estimate a multinomial logit model to explain the brand/chain choice using demographic variables. Choose a brand and estimate the effect of moving from income category .
  - (b) Estimate a conditional logit model to explain the brand/chain choice using the price and the brand and chain dummy variables as product characteristics. Choose a brand and estimate the effect of decreasing the price by \$1.25.
  - (c) Estimate a mixed logit model to explain the brand/chain choice using demographic variables and the price.
  - (d) Estimate a nested logit model with the upper level being the chain (use chain dummy variables as explanatory variables) and the lower

---

<sup>1</sup>Each consumer in the sample is given a scanner and all the items bought are weekly scanned for 18 months

<sup>2</sup>The volume equivalent allows us to go from units to weights. For example, a consumer buying two units of coffee with volume equivalent of 0.75 implies a  $2 \times 0.75 = 1.50$  pounds of coffee.

level the brands (use the price as the explanatory variable). Test the validity of IIA.

2. **Problem 2. Censoring/Truncation.** Greene (2007) analyzed the default behavior and monthly behavior of a large sample of credit card users (13,444). The Data consist of

- Cardhldr = Dummy variable, 1 if application for credit card accepted, 0 if **not**
- Default = 1 if defaulted 0 if not (observed when Cardhldr = 1, 10,499 observations),
- Age = Age in years plus twelfths of a year,
- Adepct = 1 + number of dependents,
- Acadmos = months living at current address,
- Majordrg = Number of major derogatory reports,
- Minordrg = Number of minor derogatory reports,
- Ownrent = 1 if owns their home, 0 if rent
- Income = Monthly income (divided by 10,000),
- Selfempl = 1 if self employed, 0 if not,
- Inc\_per = Income divided by number of dependents,
- Exp.Inc = Ratio of monthly credit card expenditure to yearly income,
- Spending = Average monthly credit card expenditure (for Cardhldr = 1),
- Logspend = Log of spending.

Using this data, do the following:

1. Estimate the following model

$$\logspend = \beta_1 + \beta_2 \ln income + \beta_3 Age + \beta_4 Adepct + \beta_5 ownrent + \epsilon$$

- (a) Using OLS. What is the effect of 10% increase in income on credit card expenditure?
- (b) Using Censored regression. What is the effect of 10% increase in income on credit card expenditure?
- (c) Using Heckman Two-Step Estimator. What the is effect of 10% increase in income on credit card expenditure?

2. Create a subsample where only credit cardholders appear and do the following

- (a) Estimate the above model using OLS. What is the difference in credit card spending between home owner and renter?
  - (b) Estimate the above model using truncated regression. What is the difference in credit card spending between home owner and renter?
3. Now we are interested in explaining the number of major derogatory reports as function of log income, age, the number of dependents, home ownership status and ratio of monthly credit card expenditure to yearly income.
- (a) Estimate this model using Poisson regression for credit cardholders only. What is the effect of 10% increase in income on the expected value (mean) of the number of major derogatory reports? Is Poisson regression a good specification for the data at hand?
  - (b) Estimate this model using negative binomial regression for credit cardholders only. What is the effect of 10% increase in income on the expected value (mean) of the number of major derogatory reports?
  - (c) Estimate the two models taking into account the truncation. What is the effect of 10% increase in income on the expected value (mean) of the number of major derogatory reports?