Does Ethnicity Influence Beer Consumption and Brand Choice

A Multinomial Logit Approach

Matthew Aaron Looney

Department of Agricultural and Applied Economics
Texas Tech University

October 16, 2018

Outline

- Introduction and Background
  - The Beer Industry
  - Changing Demographics
  - Data
- The Empirical Model
  - Multinomial Logit
  - Instrumental Variables
  - The IIA Property
    - Implications on Elasticities
- Empirical Results
- Conclusion and Future Work

Introduction and Background

Goals of this Research

Use a Discrete Choice Behavior Model to:

- Explore how the Demographic Composition of a Market (City) may Influence Beer Consumption and Brand Choice.
  - Do markets with high Hispanic populations prefer to consume beer produced by Mexican firms over beer produced by U.S. firms?
  - Do markets with high non-Hispanic populations prefer to consume beer produced by U.S. firms over beer produced by Mexican firms?

The Beer Industry

One of the Largest Consumer Product Sectors in America is the Beer Industry.

- As a whole, this industry represents more than $350 billion dollars of total economic impact.[1]

[1] Based on data from 2016 in a study by John Dunham & Associates.

The Beer Industry

In 2017, the U.S. beer industry shipped (sold) 207.4 million barrels of beer – equivalent to more than 2.9 billion cases of 24-12 ounce servings.[2]

- Putting this into perspective:
  - In 2017 65 billion gallons of beer sold
  - Same year, 5.45 billion gallons of milk sold in the US[3]

[2] U.S. Alcohol and Tobacco Tax and Trade Bureau (TTB) and U.S. Commerce Department, 2018

[3] Source: Statista

The Beer Industry

In 2017, 83 percent of all beer was domestically produced, and 17 percent was imported from more than 100 different countries around the world.[4]

[4] TTB and U.S. Commerce, 2018

The Beer Industry

Based on beer shipment data and U.S. Census population statistics, U.S. consumers 21 years and older consumed, on average, 26.9 gallons of beer and cider per person during 2017.[5]

[5] National Beer Wholesalers Association (NBWA) Industry Affairs, 2018

The Beer Industry

The industry structure has changed significantly over the past 30 years. In 1983, there were 49 breweries. At the end of 2016 there were 5,096 reporting brewers, and by the end of CY 2017 the most recent data show there were 8,863 TTB permitted breweries in the United States.

A quick glance at this trend would suggest the beer industry has moved from a highly concentrated to a highly diverse industry.

The Beer Industry

Market Share of Brewers from 2007 to 2017

The CR_4 has moved from 87.2% in 2007 to 78.6% in 2017.[6]

By most accounts the decrease in the CR_4 suggests the beer industry is moving toward a more competitive environment. With the introduction of many small craft brewers the market share shift from the largest firms has been small but noticeable.

[6]The Concentration Ratios CR_n is a measure of market concentration and market power. The CR_n has become a standard tools of competition economists and competition authorities to assess the competitive health of a market industry

The Beer Industry

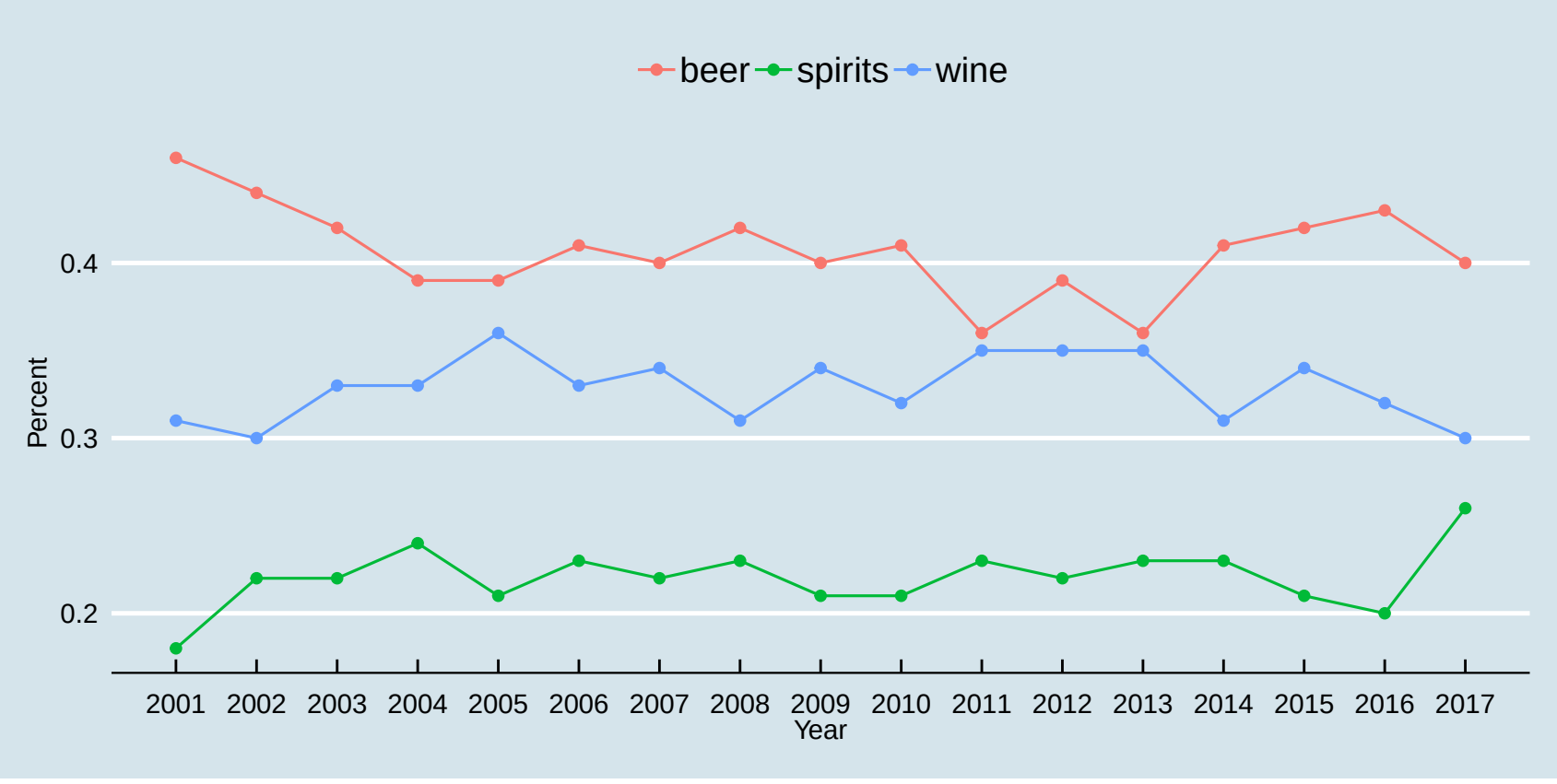
Market Share of Brewers from 2007 to 2017

Brewer/Importer	2007 Share	2017 Share
Anheuser-Busch Inbev	48.3%	41.6%
MillerCoors, LLC	29.4%	24.3%
Constellation	5.4%	8.9%
Heineken USA	4.1%	3.8%
Pabst Brewing	2.8%	2.3%
Other Domestic/Imports	10%	19%
Total	100%	100%

Source: Beer Marketer's Insights, 2018

The Beer Industry

Consumer Preference (Source: Gallup Poll, 2017)



Changing Demographics

By the year 2055 the United States will no longer have a single racial or ethnic majority.[7]

[7] Pew Research Center, (Cohn and Caumont 2016)

Changing Demographics

Growth in the Hispanic and Asian populations is predicted to almost triple over the next 40 years. By 2055, the breakdown is estimated to be 47% White, 29% Hispanic, 9% Asian, and 13% Black.[8]

[8] World Population Review, 2018

Changing Demographics

U.S. Population

	1960 (%)	2005 (%)	2055 (%)
White	85	67	47
Hispanic	3.5	14	29
Black	11	13	13
Asian	0.6	5	9

Source: Pew research Center Projections

Changing Demographics

City	population > 21	Hispanic (%)	Mexican (%)
Los Angeles, CA	2,798,161	48.5	31.9
Chicago, IL	1,996,235	28.9	21.4
Dallas, TX	849,924	42.4	36.7
Houston, TX	1,497,533	43.8	32.1
Syracuse, NY	103,268	8.3	0.7
Spokane, WA	155,143	5.0	3.4

Source: 2010 Census Data

Data

- Information Resources, Inc. (IRI) brand level data
- 50 Brands of beer representing ~85% gallons sold in a city
- 5 Firms
  - AB Inbev, MillerCoors, Constellation, Heineken, Diageo
- Weekly data (52 weeks)
- 6 Cities
  - Los Angeles, Chicago, Dallas, Houston, Syracuse, Spokane

Data

- Brands are aggregated by gallons sold in a 1 week period
- Construction of shares:
  - sum over gallons of beer sold by brand across weeks divided by market size
  - market size = City pop>21 X per capita beer consumption per week
- Beer characteristics:
  - price = US dollars per gallon
  - prct_PR = percent brand on price reduction (>5%) ($I_{[0,1]}$)
  - ABV = alcohol by volume
  - Calories_oz = calories per ounce
  - Carbs_oz = carbohydrates per ounce
  - USA = US produced beer ($I_{[0,1]}$)
  - Mexico = Mexican produced beer ($I_{[0,1]}$)

Data

Table 4: Summary Statistics - Los Angeles

Statistic	N	Mean	St. Dev.	Min	Max
price	2,600	11.924	3.035	5.278	18.122
share	2,600	0.002	0.004	0.00000	0.029
outshr	2,600	0.913	0.015	0.867	0.935
prct_PR	2,600	0.394	0.255	0.000	1.000
ABV	2,600	4.490	1.111	0	8
Calories_oz	2,600	11.321	3.288	5	20
Carbs_oz	2,600	0.991	0.702	0	3
USA	2,600	0.640	0.480	0	1
Mexico	2,600	0.120	0.325	0	1

price = weighted mean average price
share = share of brand
outshr = share of the outside good
prct PR = percent brand on price reduction
ABV = alcohol by volume

Data

Table 5: Summary Statistics - Spokane

Statistic	N	Mean	St. Dev.	Min	Max
price	2,600	11.776	3.029	5.513	20.064
share	2,600	0.002	0.004	0.00001	0.035
outshr	2,600	0.889	0.021	0.826	0.917
prct_PR	2,600	0.328	0.285	0.000	1.000
ABV	2,600	4.490	1.111	0	8
Calories_oz	2,600	11.321	3.288	5	20
Carbs_oz	2,600	0.991	0.702	0	3
USA	2,600	0.640	0.480	0	1
Mexico	2,600	0.120	0.325	0	1

price = weighted mean average price
share = share of brand
outshr = share of the outside good
prct PR = percent brand on price reduction
ABV = alcohol by volume

The Empirical Model

A Riduculously Short History

- Discrete choice behavior modeling can be traced back to the American psychologist L. L. Thurstone who developed the concept of the Law of Comparative Judgment (1927).
- Marshak (1960) re-framed Thurstone contributions in terms of utility and developed what is now called the Random Utility Model (RUM).
- Lancaster (1966) developed the theoretical framework for projecting consumer choice onto a set of product characteristics.

A Riduculously Short History

- McFadden (1973) built on the RUM by showing that the utility of an individual i making decision j, from a finite set of choices, will equal U_{ij} iff $U_{ij} > U_{ik}$, $\forall\ k \neq j$.
- However, U_{ij} can only be known up to an econometrically observable point. In other words, $U_{ij} = V_{ij} + \varepsilon_{ij}$, where V_{ij} is the deterministic part of the utility and ε_{ij} is an unobservable econometric error term.

Multinomial Logit

$$s_j = \frac{\exp(\delta_j)}{\sum_{k=0}^{J} \exp(\delta_k)} \tag{1}$$

where, $\delta_j = x_j\beta - \alpha p_j + \varepsilon_{ij}$

Multinomial Logit

$$\log \frac{s_j}{s_0} = x_j\beta - \alpha p_j + \varepsilon_{ij} \tag{2}$$

where,

s_j is the share of brand j
s_0 is the share of the outside good(s).
x_j is a vector of brand j characteristics
β and α are the parameters to be estimated

Instrumental Variables

- By normalizing the outside good (k=0 in eq. 1) and assuming the relationship between observed and predicted shares are invertible (Berry, 1994) we obtain the following linear-in-log difference form:

$$\log(s_j) - \log(s_0) = x_j\beta - \alpha p_j + \varepsilon_{ij} \tag{3}$$

- We can now employ a standard linear instrumental variables technique to deal with any endogeneity issues.

The IIA Property

- If we assume our error terms are distributed iid Extreme Value Type 1, as we have, then the ratio of the probabilities of any two alternatives are independent of all other alternatives in the choice set, a rather awkward result!

$$\frac{s_j}{s_m} = \frac{\exp(\delta_j)/\sum_{k=0}^{J}\exp(\delta_k)}{\exp(\delta_m)/\sum_{k=0}^{J}\exp(\delta_k)} = \frac{\exp(\delta_j)}{\exp(\delta_m)} = \exp(\delta_j - \delta_m) \qquad (4)$$

The IIA Property

- The implication of the IIA property can be seen most clearly when we consider the cross-price elasticities.

$$\begin{aligned}\underset{\substack{j\neq m\\ m\in k}}{\varepsilon_{jm}} &= \frac{\partial s_j}{\partial p_m}\frac{p_m}{s_j} = \left[\frac{0}{1+\Sigma\exp(\delta_k)} + \frac{-\exp(\delta_j)}{1+\Sigma\exp(\delta_k)}\left(-\alpha\exp(\delta_k)\right)\right]\frac{p_m}{s_j} \\ &= \alpha s_j s_m \frac{p_m}{s_j} \\ &= \alpha s_m p_m \end{aligned} \tag{5}$$

Empirical Results

Regression Results

Table 1: Multinomial Logit Regression

	Dependent variable:					
	log(share) - log(outshr)					
	LA	Chicago	Dallas	Houston	Syracuse	Spokane
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	−6.745***	−8.486***	−6.479***	−6.070***	−3.796***	−5.530***
	(0.276)	(0.220)	(0.256)	(0.245)	(0.221)	(0.207)
p2_Wmean	−0.175***	−0.041***	−0.151***	−0.274***	−0.133***	−0.220***
	(0.016)	(0.014)	(0.016)	(0.015)	(0.012)	(0.012)
ABV	−0.255***	−0.472***	−0.385***	−0.287***	0.032	−0.132***
	(0.049)	(0.042)	(0.041)	(0.042)	(0.041)	(0.039)
prct_PR	2.138***	3.772***	1.654***	2.699***	1.408***	0.470***
	(0.136)	(0.128)	(0.146)	(0.148)	(0.140)	(0.098)
Calories_oz	0.219***	0.270***	0.212***	0.178***	−0.011	0.207***
	(0.033)	(0.028)	(0.027)	(0.028)	(0.027)	(0.026)
Carbs_oz	−1.208***	−1.728***	−1.594***	−1.092***	−0.691***	−1.395***
	(0.140)	(0.121)	(0.117)	(0.119)	(0.119)	(0.107)
USA	−0.297***	0.249***	0.552***	0.177*	0.102	0.455***
	(0.103)	(0.092)	(0.090)	(0.096)	(0.088)	(0.080)
Mexico	1.350***	0.726***	1.009***	0.901***	−1.388***	0.260***
	(0.124)	(0.104)	(0.104)	(0.110)	(0.102)	(0.097)
Observations	2,600	2,600	2,600	2,600	2,600	2,600
R^2	0.240	0.413	0.323	0.408	0.337	0.358
Adjusted R^2	0.237	0.412	0.321	0.407	0.335	0.356
Residual Std. Error (df = 2592)	1.716	1.448	1.435	1.433	1.422	1.348
F Statistic (df = 7; 2592)	116.620***	261.049***	176.928***	255.383***	188.118***	206.251***
Note:						*p<0.1; **p<0.05; ***p<0.01

Regression Results

Table 2: Instrumental Variable - Multinomial Logit Regression

	Dependent variable:					
	log(share) - log(outshr)					
	LA	Chicago	Dallas	Houston	Syracuse	Spokane
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	−6.596***	−7.634***	−6.460***	−6.645***	−4.686***	−5.244***
	(0.284)	(0.229)	(0.275)	(0.258)	(0.228)	(0.212)
p2_Wmean	−0.188***	−0.122***	−0.152***	−0.229***	−0.063***	−0.247***
	(0.017)	(0.015)	(0.018)	(0.016)	(0.013)	(0.013)
prct_PR	2.157***	3.771***	1.656***	2.644***	1.378***	0.495***
	(0.136)	(0.129)	(0.147)	(0.148)	(0.141)	(0.098)
ABV	−0.250***	−0.438***	−0.385***	−0.314***	0.001	−0.121***
	(0.049)	(0.042)	(0.041)	(0.042)	(0.041)	(0.039)
Calories_oz	0.217***	0.255***	0.211***	0.193***	0.013	0.205***
	(0.033)	(0.028)	(0.027)	(0.028)	(0.028)	(0.026)
Carbs_oz	−1.170***	−1.487***	−1.591***	−1.233***	−0.958***	−1.338***
	(0.141)	(0.123)	(0.119)	(0.121)	(0.121)	(0.108)
USA	−0.343***	−0.085	0.547***	0.379***	0.407***	0.359***
	(0.105)	(0.096)	(0.094)	(0.100)	(0.091)	(0.082)
Mexico	1.325***	0.608***	1.007***	1.017***	−1.298***	0.225**
	(0.124)	(0.105)	(0.105)	(0.111)	(0.103)	(0.097)
Observations	2,600	2,600	2,600	2,600	2,600	2,600
R^2	0.239	0.406	0.323	0.406	0.328	0.357
Adjusted R^2	0.237	0.404	0.321	0.404	0.326	0.355
Residual Std. Error (df = 2592)	1.716	1.457	1.435	1.435	1.432	1.350

Note: *p<0.1; **p<0.05; ***p<0.01

Regression Results

Table 3: Own Price Elasticities- LA

Brand_1_25	eii_1_25	Brand_26_50	eii_26_50
BACARDI SILVER MOJITO	-2.963	MICHELOB ULTRA	-2.009
BECKS	-2.425	MICHELOB ULTRA AMBER	-2.019
BLUE MOON BELGIAN WHITE ALE	-2.980	MICHELOB ULTRA LIME CACTUS	-2.349
BUD LIGHT	-1.851	MILLER CHILL	-1.885
BUD LIGHT GOLDEN WHEAT	-2.245	MILLER GENUINE DRAFT	-1.858
BUD LIGHT LIME	-2.276	MILLER GENUINE DRAFT LIGHT 64	-1.860
BUDWEISER	-1.852	MILLER HIGH LIFE	-1.533
BUSCH	-1.544	MILLER LITE	-1.854
BUSCH LIGHT	-1.275	MILWAUKEES BEST	-1.046
COORS	-1.664	MODELO ESPECIAL	-2.291
COORS LIGHT	-1.911	MOLSON CANADIAN	-2.166
CORONA EXTRA	-2.731	NATURAL ICE	-1.264
CORONA LIGHT	-2.500	NATURAL LIGHT	-1.258
DOS EQUIS XX AMBAR LAGER	-2.360	NEGRA MODELO	-2.564
FOSTERS LAGER	-2.092	NEWCASTLE BROWN ALE	-3.174
GEORGE KILLIANS IRISH RED LAG	-2.534	ODOULS	-1.996
GUINNESS DRAUGHT	-3.142	ODOULS AMBER	-2.107
GUINNESS EXTRA STOUT	-3.263	RED STRIPE	-2.817
HEINEKEN	-2.782	SAINT PAULI GIRL	-2.282
HEINEKEN PREMIUM LIGHT LAGER	-2.576	SMIRNOFF ICE	-2.772
ICEHOUSE	-1.683	SMIRNOFF ICE GREEN APPLE BITE	-2.844
KEYSTONE LIGHT	-1.239	SMIRNOFF ICE RASPBERRY BURST	-2.849
LANDSHARK LAGER	-2.731	STELLA ARTOIS LAGER	-3.244
MICHELOB AMBER BOCK	-2.187	TECATE	-1.819
MICHELOB ULTR POMEGRANAT RSPB	-2.248	WILD BLUE	-2.910

Regression Results

Table 4: Price Elasticity Matrix - LA

	Brand	B1	B2	B3	B4	B5
1	BLUE MOON BELGIAN WHITE ALE	-2.980	0.0004	0.013	0.0002	0.0001
2	BUD LIGHT	0.039	-1.851	0.013	0.0002	0.0001
3	BUD LIGHT LIME	0.039	0.0004	-2.276	0.0002	0.0001
4	BUDWEISER	0.039	0.0004	0.013	-1.852	0.0001
5	BUSCH	0.039	0.0004	0.013	0.0002	-1.544

Subset of Brands

Table 5: Price Elasticity Matrix - LA (cont.)

	Brand	B6	B7	B8	B9	B10
6	COORS	-1.664	0.019	0.003	0.0005	0.001
7	COORS LIGHT	0.024	-1.911	0.003	0.0005	0.001
8	CORONA EXTRA	0.024	0.019	-2.731	0.0005	0.001
9	CORONA LIGHT	0.024	0.019	0.003	-2.500	0.001
10	DOS EQUIS XX AMBAR LAGER	0.024	0.019	0.003	0.0005	-2.360

Subset of Brands

Regression Results

Table 12: Own Price - Spokane

Brand_1_25	eii_1_25	Brand_26_50	eii_26_50
BACARDI SILVER MOJITO	-3.503	MICHELOB ULTRA	-2.711
BECKS	-3.312	MICHELOB ULTRA AMBER	-2.569
BLUE MOON BELGIAN WHITE ALE	-3.614	MICHELOB ULTRA LIME CACTUS	-2.896
BUD LIGHT	-2.382	MILLER CHILL	-2.777
BUD LIGHT GOLDEN WHEAT	-2.996	MILLER GENUINE DRAFT	-2.372
BUD LIGHT LIME	-2.939	MILLER GENUINE DRAFT LIGHT 64	-2.351
BUDWEISER	-2.404	MILLER HIGH LIFE	-2.039
BUSCH	-1.868	MILLER LITE	-2.395
BUSCH LIGHT	-1.855	MILWAUKEES BEST	-1.665
COORS	-2.272	MODELO ESPECIAL	-3.093
COORS LIGHT	-2.458	MOLSON CANADIAN	-2.490
CORONA EXTRA	-3.661	NATURAL ICE	-1.626
CORONA LIGHT	-3.402	NATURAL LIGHT	-1.640
DOS EQUIS XX AMBAR LAGER	-3.245	NEGRA MODELO	-3.504
FOSTERS LAGER	-2.940	NEWCASTLE BROWN ALE	-4.205
GEORGE KILLIANS IRISH RED LAG	-2.626	ODOULS	-2.169
GUINNESS DRAUGHT	-4.108	ODOULS AMBER	-2.248
GUINNESS EXTRA STOUT	-4.238	RED STRIPE	-3.636
HEINEKEN	-3.706	SAINT PAULI GIRL	-2.976
HEINEKEN PREMIUM LIGHT LAGER	-3.348	SMIRNOFF ICE	-3.529
ICEHOUSE	-2.002	SMIRNOFF ICE GREEN APPLE BITE	-3.638
KEYSTONE LIGHT	-2.001	SMIRNOFF ICE RASPBERRY BURST	-3.584
LANDSHARK LAGER	-3.621	STELLA ARTOIS LAGER	-4.306
MICHELOB AMBER BOCK	-3.216	TECATE	-2.661
MICHELOB ULTR POMEGRANAT RSPB	-2.849	WILD BLUE	-3.543

Conclusion and Future Work

Conclusion

- The goal of this research was to explore how the demographic composition of a market (City) may influence beer consumption and brand choice.
- The results indicate that markets with high Hispanic populations have a higher utility for Mexican produced beer over domestically produced beer.
- Further, the results suggest that markets with low Hispanic populations have a higher utility for domestically produced beer.
- The estimated price coefficient and own price elasticities are in agreement with other studies that have looked at beer.
- Understanding how consumer demands change based on ethnic composition have important implications for how firms respond to the changing demographic landscape.

Future Work

- The price elasticity matrix is severely restricted using the multinomial logit.
- Incorporate consumer heterogeneity into the modeling framework to uncover higher fidelity estimates. This will provide better insight into how different types of consumers are making different choices.
- Expand the study to include more cities across time.
- Utilize a random coefficients logit model to help deal with the lack of consumer heterogeneity and severely restricted elasticity matrix.