Does Ethnicity Influence Beer Consumption and Brand Choice?

Matthew Aaron Looney

May 08, 2018

Abstract

The beer industry in the United States represents more than $350 billion dollars of total economic impact. This makes the industry one of the largest consumer product sectors in America. As a category of consumer goods, beer is one of the most dynamically changing products in the country. The industry responds quickly to changes in consumer preferences and shifting consumer demographics. The continued ability to understand and predict shifts in consumer demand is paramount to the industries' ability to continue to innovate and selectively appeal to the heterogeneous consumer. It is well known that consumer demographics play a key role in determining consumer preference. A recent study by Lopez and Matschke (2012) examine consumer preference for beer using market level data. The authors assume a heterogeneous agent model with age and income variation across consumers. Consistent with expectations, they find a strong relationship between increased age and income and lower sensitivity to price. However, the study omits other relevant consumer demographic variables, such as ethnicity. This study seeks to understand more deeply how additional consumer characteristics influence preference for beer, in particular, how ethnicity influences beer consumption and brand choice. The methodological approach taken in this study is to estimate a differentiated demand for beer using a random coefficients discrete choice model. Information Resources, Inc. (IRI) brand level data will be used as well as consumer level demographic data from the Current Population Survey (CPS).

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

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

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

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

Table 6: Price Elasticity Matrix - LA (cont.)

	Brand	B11	B12	B13	B14	B15
11	FOSTERS LAGER	-2.092	0.008	0.001	0.0001	0.001
12	GUINNESS EXTRA STOUT	0.0002	-3.263	0.001	0.0001	0.001
13	HEINEKEN	0.0002	0.008	-2.782	0.0001	0.001
14	KEYSTONE LIGHT	0.0002	0.008	0.001	-1.239	0.001
15	MILLER GENUINE DRAFT	0.0002	0.008	0.001	0.0001	-1.858

Subset of Brands

Table 7: Price Elasticity Matrix - LA (cont.)

	Brand	B16	B17	B18	B19	B20
16	MILLER GENUINE DRAFT LIGHT 64	-1.860	0.0002	0.005	0.0001	0.00005
17	MILLER LITE	0.003	-1.854	0.005	0.0001	0.00005
18	MILWAUKEES BEST	0.003	0.0002	-1.046	0.0001	0.00005
19	MODELO ESPECIAL	0.003	0.0002	0.005	-2.291	0.00005
20	TECATE	0.003	0.0002	0.005	0.0001	-1.819

Subset of Brands

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu % Date and time: Fri, Oct 05, 2018 - 13:00:06

Table 8: Own Price - Chicago

Brand_1_25	eii_1_25	Brand_26_50	eii_26_50
BACARDI SILVER MOJITO	-1.546	MICHELOB ULTRA	-1.053
BECKS	-1.618	MICHELOB ULTRA AMBER	-0.973
BLUE MOON BELGIAN WHITE ALE	-1.704	MICHELOB ULTRA LIME CACTUS	-1.277
BUD LIGHT	-1.017	MILLER CHILL	-1.148
BUD LIGHT GOLDEN WHEAT	-1.378	MILLER GENUINE DRAFT	-1.013
BUD LIGHT LIME	-1.381	MILLER GENUINE DRAFT LIGHT 64	-0.964
BUDWEISER	-1.032	MILLER HIGH LIFE	-0.822
BUSCH	-0.708	MILLER LITE	-1.074
BUSCH LIGHT	-0.670	MILWAUKEES BEST	-0.629
COORS	-0.855	MODELO ESPECIAL	-1.517
COORS LIGHT	-1.005	MOLSON CANADIAN	-1.079
CORONA EXTRA	-1.552	NATURAL ICE	-0.685
CORONA LIGHT	-1.490	NATURAL LIGHT	-0.555
DOS EQUIS XX AMBAR LAGER	-1.459	NEGRA MODELO	-1.652
FOSTERS LAGER	-1.426	NEWCASTLE BROWN ALE	-1.788
GEORGE KILLIANS IRISH RED LAG	-1.429	ODOULS	-0.988
GUINNESS DRAUGHT	-1.792	ODOULS AMBER	-1.034
GUINNESS EXTRA STOUT	-1.841	RED STRIPE	-1.725
HEINEKEN	-1.649	SAINT PAULI GIRL	-1.395
HEINEKEN PREMIUM LIGHT LAGER	-1.647	SMIRNOFF ICE	-1.678
ICEHOUSE	-0.778	SMIRNOFF ICE GREEN APPLE BITE	-1.752
KEYSTONE LIGHT	-0.575	SMIRNOFF ICE RASPBERRY BURST	-1.757
LANDSHARK LAGER	-1.556	STELLA ARTOIS LAGER	-1.945
MICHELOB AMBER BOCK	-1.225	TECATE	-1.152
MICHELOB ULTR POMEGRANAT RSPB	-1.284	WILD BLUE	-1.657

Table 9: Own Price - Dallas

Brand_1_25	eii_1_25	Brand_26_50	eii_26_50
BACARDI SILVER MOJITO	-1.992	MICHELOB ULTRA	-1.736
BECKS	-2.049	MICHELOB ULTRA AMBER	-1.792
BLUE MOON BELGIAN WHITE ALE	-2.134	MICHELOB ULTRA LIME CACTUS	-1.892
BUD LIGHT	-1.531	MILLER CHILL	-1.684
BUD LIGHT GOLDEN WHEAT	-1.865	MILLER GENUINE DRAFT	-1.533
BUD LIGHT LIME	-1.837	MILLER GENUINE DRAFT LIGHT 64	-1.536
BUDWEISER	-1.516	MILLER HIGH LIFE	-1.322
BUSCH	-1.246	MILLER LITE	-1.564
BUSCH LIGHT	-1.238	MILWAUKEES BEST	-1.132
COORS	-1.490	MODELO ESPECIAL	-1.907
COORS LIGHT	-1.577	MOLSON CANADIAN	-1.937
CORONA EXTRA	-2.217	NATURAL ICE	-1.140
CORONA LIGHT	-2.054	NATURAL LIGHT	-1.156
DOS EQUIS XX AMBAR LAGER	-1.996	NEGRA MODELO	-2.169
FOSTERS LAGER	-1.826	NEWCASTLE BROWN ALE	-2.312
GEORGE KILLIANS IRISH RED LAG	-1.948	ODOULS	-1.621
GUINNESS DRAUGHT	-2.595	ODOULS AMBER	-1.578
GUINNESS EXTRA STOUT	-2.489	RED STRIPE	-2.201
HEINEKEN	-2.165	SAINT PAULI GIRL	-2.000
HEINEKEN PREMIUM LIGHT LAGER	-2.141	SMIRNOFF ICE	-2.289
ICEHOUSE	-1.254	SMIRNOFF ICE GREEN APPLE BITE	-2.326
KEYSTONE LIGHT	-1.278	SMIRNOFF ICE RASPBERRY BURST	-2.302
LANDSHARK LAGER	-2.230	STELLA ARTOIS LAGER	-2.603
MICHELOB AMBER BOCK	-1.828	TECATE	-1.737
MICHELOB ULTR POMEGRANAT RSPB	-1.872	WILD BLUE	-2.211

Table 10: Own Price - Houston

Brand_1_25	eii_1_25	Brand_26_50	eii_26_50
BACARDI SILVER MOJITO	-2.948	MICHELOB ULTRA	-2.560
BECKS	-3.060	MICHELOB ULTRA AMBER	-2.605
BLUE MOON BELGIAN WHITE ALE	-3.080	MICHELOB ULTRA LIME CACTUS	-2.796
BUD LIGHT	-2.125	MILLER CHILL	-2.546
BUD LIGHT GOLDEN WHEAT	-2.728	MILLER GENUINE DRAFT	-2.145
BUD LIGHT LIME	-2.683	MILLER GENUINE DRAFT LIGHT 64	-2.156
BUDWEISER	-2.107	MILLER HIGH LIFE	-1.639
BUSCH	-1.603	MILLER LITE	-2.165
BUSCH LIGHT	-1.568	MILWAUKEES BEST	-1.441
COORS	-2.132	MODELO ESPECIAL	-2.682
COORS LIGHT	-2.127	MOLSON CANADIAN	-2.742
CORONA EXTRA	-3.208	NATURAL ICE	-1.379
CORONA LIGHT	-3.038	NATURAL LIGHT	-1.497
DOS EQUIS XX AMBAR LAGER	-2.993	NEGRA MODELO	-3.063
FOSTERS LAGER	-3.086	NEWCASTLE BROWN ALE	-3.464
GEORGE KILLIANS IRISH RED LAG	-3.348	ODOULS	-2.223
GUINNESS DRAUGHT	-3.881	ODOULS AMBER	-2.212
GUINNESS EXTRA STOUT	-3.662	RED STRIPE	-3.464
HEINEKEN	-3.158	SAINT PAULI GIRL	-3.084
HEINEKEN PREMIUM LIGHT LAGER	-3.145	SMIRNOFF ICE	-3.340
ICEHOUSE	-1.782	SMIRNOFF ICE GREEN APPLE BITE	-3.396
KEYSTONE LIGHT	-1.504	SMIRNOFF ICE RASPBERRY BURST	-3.364
LANDSHARK LAGER	-3.156	STELLA ARTOIS LAGER	-3.962
MICHELOB AMBER BOCK	-2.664	TECATE	-2.347
MICHELOB ULTR POMEGRANAT RSPB	-2.769	WILD BLUE	-3.603

Table 11: Own Price - Syracuse

Brand_1_25	eii_1_25	Brand_26_50	eii_26_50
BACARDI SILVER MOJITO	-0.914	MICHELOB ULTRA	-0.574
BECKS	-0.886	MICHELOB ULTRA AMBER	-0.641
BLUE MOON BELGIAN WHITE ALE	-0.903	MICHELOB ULTRA LIME CACTUS	-0.755
BUD LIGHT	-0.530	MILLER CHILL	-0.769
BUD LIGHT GOLDEN WHEAT	-0.747	MILLER GENUINE DRAFT	-0.557
BUD LIGHT LIME	-0.729	MILLER GENUINE DRAFT LIGHT 64	-0.561
BUDWEISER	-0.551	MILLER HIGH LIFE	-0.398
BUSCH	-0.383	MILLER LITE	-0.547
BUSCH LIGHT	-0.388	MILWAUKEES BEST	-0.371
COORS	-0.504	MODELO ESPECIAL	-0.868
COORS LIGHT	-0.548	MOLSON CANADIAN	-0.543
CORONA EXTRA	-0.920	NATURAL ICE	-0.385
CORONA LIGHT	-0.858	NATURAL LIGHT	-0.386
DOS EQUIS XX AMBAR LAGER	-0.802	NEGRA MODELO	-0.888
FOSTERS LAGER	-0.728	NEWCASTLE BROWN ALE	-0.927
GEORGE KILLIANS IRISH RED LAG	-0.781	ODOULS	-0.649
GUINNESS DRAUGHT	-1.042	ODOULS AMBER	-0.659
GUINNESS EXTRA STOUT	-1.083	RED STRIPE	-0.991
HEINEKEN	-0.880	SAINT PAULI GIRL	-0.895
HEINEKEN PREMIUM LIGHT LAGER	-0.865	SMIRNOFF ICE	-0.947
ICEHOUSE	-0.361	SMIRNOFF ICE GREEN APPLE BITE	-1.012
KEYSTONE LIGHT	-0.370	SMIRNOFF ICE RASPBERRY BURST	-1.000
LANDSHARK LAGER	-0.907	STELLA ARTOIS LAGER	-1.018
MICHELOB AMBER BOCK	-0.692	TECATE	-0.705
MICHELOB ULTR POMEGRANAT RSPB	-0.759	WILD BLUE	-0.971

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

Table 13:

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
cdid	2,600	26.500	15.011	1	13.8	39.2	52
p2_Wmean	2,600	11.924	3.035	5.278	9.873	14.436	18.122
share	2,600	0.002	0.004	0.00000	0.0001	0.001	0.029
outshr	2,600	0.913	0.015	0.867	0.906	0.925	0.935
prct_PR	2,600	0.394	0.255	0.000	0.184	0.576	1.000
ABV	2,600	4.490	1.111	0	4.2	5	8
Calories_oz	2,600	11.321	3.288	5	8.5	12.5	20
Carbs_oz	2,600	0.991	0.702	0	0.5	1.2	3
USA	2,600	0.640	0.480	0	0	1	1
Mexico	2,600	0.120	0.325	0	0	0	1

1 Introduction

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] As a category of consumer goods, beer is one of the most dynamically changing products in the country. The industry responds quickly to changes in consumer preference and shifting consumer demographics. The ability of the beer industry to change quickly to appeal to a new consumer group makes the industry an interesting candidate to study with regard to changing ethnic diversity. The focus of this study is to evaluate how shifts in the consumer demographic landscape is changing how the beer industry responds to increasing populations of Hispanics and Asians. To understand this question it is important to model demand with consideration of, not only income, age, and family size, but also ethnic makeup. Ethnicity has largely been ignored in the differentiate product demand literature and this study aims to fill this

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

gap.

According to a study by the Pew Research Center, by the year 2055 the United States will no longer have a single racial or ethnic majority (Cohn and Caumont 2016). In 1965, 84% of Americans were non-Hispanic Whites, 11% Black, 4% Hispanic, less than 1% Asian and less than 1% other. By the year 2015 the ethnic distribution of the United States had shifted with non-Hispanic whites now representing 62% of the U.S. population, 12% Black, 18% Hispanic, 6% Asian and 2% other.[2] Over a 50 year period the American demographic landscape had changed with large increases in Hispanic and Asian representation. These divergent trends are expected to continue long into the future, at variable rates, depending on how immigration policy morphs with fluctuations in the political landscape (Passel and Rohal 2015).

When discussing ethnic trends it has become important to distinguish ethnicity from race and to define these terms precisely. Unfortunately this is not a trivial task. The classic definitions of race and ethnicity used by sociologists and anthropologist are given by the following two definitions. Race refers to a category of people who share certain inherited physical characteristics, such as skin color, facial features, and stature. Using physical differences as their criteria, scientists at one point identified as many as nine races (Smedley 1998).[3] Ethnicity refers to the shared social, cultural, and historical experiences, stemming from common national or regional backgrounds, that make subgroups of a population different from one another. Within the context of ethnicity, an ethnic group can be defined as a subgroup of a population with a set of shared social, cultural, and historical experiences; with relatively distinctive beliefs, values, and behaviors; and with some sense of identity of belonging to the subgroup

[2] Whites, Blacks and Asians include only single-race non-Hispanics. Asians include Pacific Islanders. Hispanics are of any race. Source: Pew research Center estimates based on adjusted census data

[3] African, American Indian or Native American, Asian, Australian Aborigine, European (more commonly called "white"), Indian, Melanesian, Micronesian, and Polynesian

(Barkan 2016).

There has been much debate and disagreement within the social science community when dealing with issues of race and ethnicity. However, the current general consensus is that biological definitions of race are antiquated and largely irrelevant. Issues of racial classification based on genetic profiling have muddied the waters even further.[4] With the current level of genetic diversity present in the world today it is not possible to classify an individual as purely one race over another. The idea of racial classification based on biology has dissolved and the idea of ethnic identity has emerged as the preferred method of individual demographic identifier. Due to the complexity of the problems related to using biological race as a demographic identifier I will instead use the concept of ethnicity as defined above. An increasing number of data sets are including questions about ethnicity. This self reporting of ethnic identity is fundamentally inline with the widely accepted definitions of ethnicity and thus I feel it more appropriate to use ethnic demographic information when studying consumer preference.

Distinct patterns of immigration from Latin America and Asia is driving much of the current change in the demographic landscape. Hispanics and Asians are estimated to represent 70% of the total immigrant population by the year 2055 with 34% and 36%, respectively (Cohn and Caumont 2016). But immigration patterns represent more than simple statistics. Immigrants are people who bring with them ethnic identity and culture. This ethnic identity is embedded in food, religion, language, and to varying degrees, permeates every aspect of an individuals identity. The United States has always been a nation of immigrants. Since it's founding, the country has undergone three distinct episodes of immigration. The most recent between

[4] With the availability of DNA testing by companies such as Ancestry.com and 23andMe many people discover the true racial diversity within their genes.

1850 to 1930 when large waves of immigrants from Germany, Italy, Ireland, Greece and many eastern European nations landed on American soil. These immigrants changed the American food scene by bringing their food preferences with them from their home countries. Given the current changes in the ethnic landscape of the United States, it is widely accepted that a fourth episode of immigration is underway. Consumer food preference last generation is looking very different from consumer food preference in the current generation. In fact, at the current pace of ethnic change, shifts in inter-generational food preference are becoming a reality that the U.S. food industry will need to manage to appeal to the heterogeneous consumer. Understanding how food preferences change with large populations from Latin America and Asia settling in to the fabric of the United States is an important research question for economist to tackle. Many studies incorporate consumer income and age variations when studying demand for differentiated goods but few explicitly model ethnicity as an added demographic variable. I argue that ethnicity is an important component to consider when studying consumer demand. This study seeks to understand more deeply how additional consumer characteristics influence product preference, and in particular, how ethnicity influences beer consumption and brand choice.

1.1 Literature Review

The theoretical framework for projecting consumer choice onto a set of product characteristics started with work by Lancaster (1966). This idea was extended further by McFadden who developed the econometric model in (1973). This projection allowed for the estimation of a large number of products using a finite number of characteristics. Prior to this line of research the problem of dimensionality plagued empirical industrial organization and differentiated product demand which rendered many interesting problems intractable.

While dimensionality was a large problem to overcome, price endogeneity is present in almost all supply and demand models and discrete choice modeling is no exception. If I fail to account for this endogeneity then the parameter estimates will be biased and depending on the exact model specification, inconsistent. Berry (1994) introduced an elegant method to overcome the endogeneity issue. He developed a method to estimate differentiated demand models in the presence of unobserved product characteristics by inverting the market share equation to obtain an estimate of mean utility. What the inversion technique provided was a method to linearize the market share equation which then allowed for the use of linear instrumental variables. Without this method the market share equation is highly non-linear and the use of standard linear instrumental variables will not work. Once the implied mean utility is obtained the estimation can proceed in a similar way as homogeneous goods modeling. While Berry's method does allow for the handling of endogenous prices through the use of standard linear instrumental variables his method does not address the well known problem of independence of irrelevant alternatives (IIA) thereby severely restricting the cross-price elasticity patterns.[5]

The incorporation of consumer heterogeneity in discrete choice modeling of differentiated products explicitly solves the IIA problem. In a paper by Berry, Levinsohn and Pakes (BLP) (1995) have shown that individual consumer heterogeneity can be applied to a discrete choice model of differentiated goods in a tractable way to uncover consistent estimations of demand elasticities. The BLP model developed a way to break the IIA property which allowed for more realistic substitution patterns across products. The BLP model has become a popular model in differentiate product

[5]For detailed discussions see Domencich and McFadden (1975), Hausman and Wise (1978), (S. Berry et al. 1995), (Nevo 2000)

demand and has become the model of choice when investigating merger simulations. However, the results obtained from the merger simulations are going to be driven by the demand model and supply side assumptions. It is therefore imperative to consider all components that may affect the demand for a product. Most studies incorporate components of income, age, and household size while studying differentiated product demand but few studies have incorporated ethnicity. I argue that ethnicity is a missing key component in determining product choice and should be included in the demand model.

It is well known that consumer demographics play a key role in determining consumer preference. A search of the literature related to discrete choice modeling of differentiated goods reveal few studies that incorporate other demographic variables beyond income and age. A study by Lopez and Matschke (2012) use a random coefficients multinomial logit model to examine consumer preference for beer using market level data. The authors assume a heterogeneous agent model with age and income variation across consumers. Consistent with expectations, they find a strong relationship between increased age and income and lower sensitivity to price. However, the authors omit other relevant consumer demographic variables, such as ethnicity. One of the few studies to incorporate an ethnic variable in a random coefficients discrete choice model was Romeo (2016). His results show that as the percentage of Hispanics in the population increase, the utility gained from consuming Corona decrease and turn negative when moving from 18 to 22 percent Hispanics. Romeo can not offer any clear explanation of these results but suggests that perhaps with increasing percentages of Hispanics in a market, competition from other Latin American beers are moving consumers away from Corona. With so few studies on how ethnicity impacts demand it is important to systematically study this missing demographic variable. This paper aims to fill this gap by studying how ethnicity impacts demand choice in different markets across the United States.

1.2 Conceptual Framework

The demand system is obtained by assuming an individual will maximize their utility by selecting a product from the choice set. The conditional indirect utility of individual i selecting product j in market t is given by $U(x_{jt}, \xi_{jt}, p_{jt}, \zeta_i | \theta)$, where x_{jt} is a K-dimensional (row) vector of observable product characteristics, ξ_{jt} is the unobserved (by the researcher) product characteristic, p_{jt} is the price of product j in market t, ζi are individual characteristics, and θ is a set of unknown parameters to be estimated.

Then indirect utility is given by the following,

$$u_{ijt} = \alpha_i f(\cdot) + \beta_i x_{jt} + \xi_{jt} + \varepsilon_{ijt} \tag{1}$$

where, $i = 1, ..., I_t,\ j = 1, ..., J,\ t = 1, ..., T,\ f(\cdot)$ is a functional form for indirect utility[6], α_i is individual i's marginal utility gained from the parameters of the utility function, β_i is a K-dimensional (column) vector of individual-specific taste coefficients, and ε_{ijt} is a mean zero stochastic error term.

The introduction of consumer heterogeneity is done through modeling individual taste preference (α_i, β_i) based on consumer characteristics, ζ_i. The individual consumer characteristics consist of two parts, observed characteristics, D_i and unobserved

[6]Depending on the research question the functional form can take on any one of a number of utility functions. For example, if wealth effects are thought to be important then a Cobb-Douglas functional may be appropriate, if not then quasi-linear may be chosen, see (Petrin 2002)

characteristics v_i. The observed characteristics can be obtained by either sampling from a joint parametric distribution (e.g., the Current Population Survey) or by using sample data to estimate the mean and standard deviation of the characteristics of interest from a population. For example, assume a log-normal distribution on income and take random draws using the mean and standard deviation from census data. The unobserved characteristics, v_i, can be generated by random sampling from a parametric distribution.

Following nomenclature used by Nevo gives the following,

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Pi D_i + \Sigma v_i,\ v_i \sim P_v^*(v),\ \mathrm{D}_i \sim \hat{P}_D^*(D) \tag{2}$$

where, D_i is observed characteristic, v_i is unobserved characteristics, $P_v^*(\cdot)$ is a parametric distribution, $\hat{P}_D^*(\cdot)$ is either a parametric or non-parametric distribution. Π is a matrix of parameter estimates that measure the variation in taste characteristics across demographics and Σ is a matrix of parameter estimates for the unobserved demographic characteristics.

For completion of the demand system I need to define the outside good. Without the outside good construction, all individuals in all markets would be purchasing all products all of the time. For the econometrician to see a change in quantities demanded with a homogeneous price increase across all products (relative to other markets), I need people moving to the margins which means I need to define the option where people can make a zero purchase. I define the mean utility gained from the outside option as U_{i0t}. It is important to note that the outside option is not identified in the current context.[7]

[7]It is possible to identify the outside option with further restrictions and normalization, see Nevo

Combining equations (1) and (2) I obtain the expression for the conditional indirect utility,

$$
\begin{aligned}
u_{ijt} &= \delta_{jt}(x_{jt}, p_{jt}, \xi_{jt}|\theta_1) + \mu_{ijt}(x_{jt}, p_{jt}, v_i, D_i|\theta_2) + \varepsilon_{ijt}, \\
\delta_{jt} &= \beta x_{jt} - \alpha p_{jt} + \xi_{jt}, \\
\mu_{ijt} &= [-p_{jt}, x_{jt}](\Pi D_i + \Sigma v_i)
\end{aligned}
\tag{3}
$$

where, δ_{jt} is the mean utility, and μ_{ijt} and ε_{ijt} are the mean zero deviations from the mean utility which represent the random coefficient interactions between the observed and unobserved parts of the product characteristics and the demographics. The vectors θ_1 and θ_2 contain the linear and nonlinear parameters of the model with $\theta_1 = (\alpha, \beta)$ and $\theta_2 = (\Pi, \Sigma)$.

The model parameters (θ_1 and θ_2) can be estimated using Hansen's generalized method of moments estimation technique (Hansen 1982) and following the procedures detailed in Berry, et. al. (1995) and Nevo (2000).

1.3 Data, Methods and Procedures

The methodological approach taken in this study is to estimate a differentiated demand for beer using the random coefficients multinomial logit model developed by Berry, et. al. (1995) and extended by Nevo (2001), Chidmi (2006), Chidmi and Lopez (2007), and Chidmi and Murova (2011). The data required to consistently estimate the model outlined in the previous section consist of the following variables: market shares, prices, brand characteristics, and information on the distribution of consumer demographics. Unlike other papers that have studied differentiated product demand

(2000) for a more detailed discussion.

this paper will incorporate a more rich set of demographic variables to elucidate how ethnic characteristics influence brand choice. In this paper a market is defined as beer sales during 2010 in Los Angeles, California over 52 weeks at the chain/brand level. The market level sales data are obtained from Information Resource Inc. (IRI) and consist of all dollar by volume sales for a brand of beer at a specific chain during a one week period. Market shares are calculated by adding up the total gallons of beer sold, by brand, in a given market and dividing by the total potential market size. The potential market size is assumed to be the per capita consumption of beer in Los Angeles, California.[8] The product characteristics used in this study are as follows: alcohol by volume (ABV), international bitterness units (IBU), a measure of beer color, referred to as the standard reference method (SRM), calories per once, and carbohydrates per once. Indicator variables are included for the brand country of origin[9] and the style of beer.[10] Consumer demographics are obtained by pooling data from the 2009, 2010, and 2011 Current Population Survey, Annual Social and Economic (ASEC) Supplement.[11] The demographic variables used are total family income, individual age, individual educational attainment, and a series of survey questions that ask the following: Are you Spanish, Hispanic, or Latino?, if yes then from where[12], In what country were you born?, In what country was your mother born?, and In what country was your father born? All demographic variables are taken

[8]For example, the per capita consumption of beer for consumers over the age of 21 in the state of California is given as 25.5 gallons, according to 2012 data by the Beer Institute. The 2010 census estimates the population of individuals over the age of 21 in Los Angeles as 2.6 million. For weekly data, the total potential market size is 1,297,558 gallons of beer for consumers over 21.

[9]Was the brewer located in the United States, Mexico, Germany, the Netherlands, etc.

[10]For example, Pale Lagers & Pilsners, German Style Bocks, India Pale Ale (IPA), Wheat Beer, etc.

[11]Very little variation in demographics occur over any three year period. The pooling of data allows me to obtain a larger sample of individuals to minimize the error propagation during the numerical integration step in the estimation phase.

[12]The respondent has a choice of five countries/regions to chose from; Mexico, Puerto Rico, Cuba, Central/South America, or other Spanish.

at the individual level with the exception of total family income.[13] All demographic variables are for individuals over the age of 21, the age at which alcohol is legally purchased and consumed in the United States.

Price data and unobserved product characteristics will be correlated with econometric error term and cause problems with endogeneity. These problem can be overcome by specifying a set of instruments that are correlated with prices and the unobserved product characteristics but uncorrelated with the error terms. (need to define and describe the IV's used... not done yet...)

The estimation proceeds using GMM methods to form a population moment condition which is a product of the instrumental variables and a structural error term to form a nonlinear GMM estimator. Let $Z = [z_1, ...z_m]$ be a set of instruments where $E[Z' \cdot \omega(\theta^*)] = 0$, and ω (the error term) is a function of the model parameters and θ^* denotes the true values of the model parameters. The GMM estimate is

$$\hat{\theta} = \arg\min_{\theta}[\omega(\theta)' Z A^{-1} Z' \omega(\theta)] \tag{4}$$

where, A is a consistent estimate of $E[Z'\omega\omega'Z]$. Following the procedures detailed in Berry (1994), the unobserved product characteristics error term is given by $\xi_j + \Delta\xi_{jt}$. The unobserved product can then be computed as a function od the data and the parameters by solving for the mean utility levels, $\delta_{\cdot t}$, that solves the implicit equations,

$$s_{\cdot t}(x, p_{\cdot t}, \delta_{\cdot t}|\theta_2) = S_{\cdot t} \tag{5}$$

where, $s_{\cdot t}(\cdot)$ is the market share function defined by equation (define market share

[13] I assume that families pool their resources and adult individuals have equal access to total family income

function in Conceptual Framework...), and $s_{\cdot t}$ are the observed market shares. The inversion can be done either numerically (for the full model) or by using the McFadden result for the logit model which gives, $\delta_{jt}(x, p_{\cdot t}, S_{\cdot t}|\theta_2) = ln(S_{jt}) - ln(S_{0t})$. The error term is now defined as $\omega_{jt} = \delta_{jt}(x, p_{\cdot t}, S_{\cdot t}|\theta_2) - (x_j\beta + \alpha p_{jt})$. As Nevo (2001) notes, the reason is now clear for separating out the θ_1 and θ_2 terms, θ_1 enters the error term and the GMM objective function linearly while θ_2 enters nonlinearly.

The estimation algorithm computer code was developed by closely following the Matlab computer code developed by Nevo (2000) and Chidmi and Murova (2011).

References

Barkan, S. E. (2016). *Sociology : understanding and changing the social world.* Minneapolis: Minneapolis, MN : University of Minnesota Libraries Publishing.

Berry, S. T. (1994). Estimating Discrete-Choice Models of Product Differentiation. *The RAND Journal of Economics*, *25*(2), 242–262.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, *63*(4), 841.

Chidmi, B. (2006). *Brand -supermarket level demand for breakfast cereals and multi-dimensional competition* (PhD thesis). ProQuest Dissertations Publishing.

Chidmi, B., & Lopez, R. A. (2007). Brand-supermarket demand for breakfast cereals and retail competition.(Author abstract). *American Journal of Agricultural Economics*, *89*(2), 324.

Chidmi, B., & Murova, O. (2011). Measuring market power in the supermarket industry: the case of the SeattleTacoma fluid milk market. *Agribusiness*, *27*(4), 435–449.

Cohn, D., & Caumont, A. (2016). 10 demographic trends that are shaping the US and the world. *Fact Tank, Pew Research Center, March*, *31*.

Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, *50*(4), 1029–1054.

Lancaster, K. J. (1966). A New Approach to Consumer Theory. *Journal of political Economy*, *74*(2), 132–157.

Lopez, R. A., & Matschke, X. (2012). Home Bias in US Beer Consumption. *Pacific Economic Review*, *17*(4), 525–534.

McFadden, D., & others. (1973). Conditional logit analysis of qualitative choice behavior.

Nevo, A. (2000). A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand. *Journal of Economics & Management Strategy*, *9*(4), 513–548.

Nevo, A. (2001). Measuring Market Power in the Ready-to-Eat Cereal Industry. *Econometrica*, *69*(2), 307–342.

Passel, J., & Rohal, M. (2015). *Modern Immigration Wave Brings 59 Million to US, Driving Population Growth and Change Through 2065.*

Petrin, A. (2002). Quantifying the Benefits of New Products: The Case of the Minivan.

Journal of political Economy, *110*(4), 705–729.

Romeo, C. J. (2016). Incorporating Prior Information into A GMM Objective For Mixed Logit Demand Systems. *Journal of Industrial Economics*, *64*(2), 336–363.

Smedley, A. (1998). “Race” and the Construction of Human Identity. *American Anthropologist*, *100*(3), 690–702.