Learning in Credence Good Markets: An Example of Vitamins

Iryna Demko, PhD student, Agricultural Economics, Sociology, and Education Department

Pennsylvania State University, iryna.demko@gmail.com

Edward C. Jaenicke, Associate Professor of Agricultural Economics

Pennsylvania State University, ecj3@psu.edu

DRAFT: May 25, 2014

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2014 AAEA Annual Meeting, Minneapolis, MN, July 27-29, 2014.

We wish to thank the Data Center at The University of Chicago Booth School of Business for providing data. Information on availability and access to the data is available at http://research.chicagobooth.edu/nielsen.

Abstract

Unlike many studies of learning and pharmaceuticals, this paper considers credence goods such as vitamins and the role of consumer experience in resolving uncertainty when the user cannot observe the effects of the goods after consumption. The Homescan data justifies variations in the purchases: 45% of households choose different Universal Product Code (UPC) items during subsequent shopping trips than the ones they bought originally. My findings suggest that the probability of choosing Brand 1 increases after a positive experience with Brand 1 and declines after a positive experience with Brand 2. This is based on the assumption that the consumer has had a positive experience about the product if she bought it with a current purchase and three periods back. In a structural model I intend to relax this assumption and compare the endogenous speed of learning about vitamins with the speed of learning about drugs.

Keywords: credence good, vitamin, learning, speed of learning, spillover effect

JEL Classification: D83, L15, I1

1 Introduction

In all honesty, how do I really know the effectiveness of this supplement? The health media highly recommends taking a multi...I choose Naturemade because of the often 2 for 1 sales. Yet I do not REALLY know if they are as wonderful as I hope they are. What difference would feel if I quit taking them? I trust in Naturemade...that's the best I can say and hope for.

from a consumer review, http://www.naturemade.com/products/multivitamins/multi-daily

People take vitamins to remedy gaps in their diets and to be protected against illnesses and disease. Americans are now more health conscious than ever before. According to the Mintel Group Report, sales of vitamins in the US increased by 38% from 2008-2013, and are expected to climb another 30% by 2018 (Mintel Group, 2013).

Uncertainty about brand quality influences individuals when they select vitamins. Consumers may not know whether a vitamin is going to be effective in improving their health. This is further complicated since the impact of vitamins may differ for people with different habits, diets, different levels of stress, and different baseline measures of health, etc.

Furthermore, treatment outcomes of vitamins are difficult to study[1]. With prescription pharmaceuticals, the gold standard for research is the randomized clinical trial in which some patients take a drug and others take a placebo. Since people ingest vitamins as essential nutrients in their daily diets, however, there is no way to withhold vitamins altogether from research subjects, thus ensuring a completely accurate study.

Also, the form of supplement available for retail purchase may not be the same as the form used for research. Since the vitamins, minerals, and supplements category is not regulated by the Food and Drug Administration, products might not always be what producers claim (Mintel Group, 2013).

There is no observed symptom relief after a vitamin is taken, which makes it difficult or impossible to measure utility gains or losses. Thus, I consider over-the-

[1]Parker-Pope T. 'Vitamin Pills: A False Hope?' *New York Times*, February 16, 2009.

counter vitamins as an excellent example of *credence* goods. In this environment, buyers can only ascertain the quality of products through experimentation. Since consumer experience leads to consumers' updating of the match values, vitamins, though credence goods, share some features with experience goods as well.

This study aims to understand how a consumer resolves uncertainty in the vitamin market.

The environment itself stimulates learning since the cost of experimenting is not high; in fact, the average cost of a vitamin may be as low as 3 cents per pill (Amazon Web-site). Furthermore, vitamins are easy to buy without a prescription, and as long as the shoppers do not feel worse after taking the vitamins, nothing really stops them from trying a different brand or form of vitamins.

Using a reduced form model that assumes uncertainty is resolved in three periods, I found that people tend to buy a cheaper brand of vitamins if they do not have any experience taking them. However, after consumers have tried several brands, it is not clear which brand they will choose. Individuals do change which vitamins they buy, but rather than switching brands, they switch the type of vitamin they purchase. In other words, the brands become clusters. Subsequently, the model can be estimated in a Dickstein (2004)'s fashion using the index rule articulated Pandey et al. (2007). First, however, I need to address the problem of left-censored data, since the history of vitamin consumption prior to 2004 is not observable.

While there are many studies of pharmaceuticals, to the best of my knowledge none of them consider credence goods such as vitamins and the role of consumer experience in resolving uncertainty regarding their quality. Crawford and Shum (2005), for example, study learning by employing a unique data set of anti-ulcer prescriptions in Italy. Unlike vitamins, prescription drugs are experience goods that have clear effects: symptom relief after consumption. This kind of effect is not observable in the retail vitamin market. Moreover, Crawford and Shum (2005) use a patient-level dataset with both observed prescriptions and observed changes. This type of data set would not be appropriate or available for the market because vitamins can be purchased over-the-counter. Finally, unlike in the retail market for vitamins, patient prices for anti-ulcer drugs do not vary, because all patients' medical costs are covered by national health system.

In the following sections I give a profile of the typical vitamin consumer and use data to explain the evidence of consumer learning. I subsequently use a reduced form model to show that learning is happening, discuss possible spillover effects, and suggest potential ways to address the problem of left-censored data.

2 Data

I use Homescan Consumer Panel Data from the Nielsen Company (US), LLC provided by the Marketing Data Center at the University of Chicago Booth School of Business.

Consumers and Products

The Nielsen Company recruited 39,577 households in 2004, with 17,755 of them continuing to be tracked in 2011.

In order to identify the *regular* consumers of vitamins, I draw a sample from 24 of major markets in the US, all of which are available for the years 2004-2011.

The initial filters lead to a balanced panel of 336 unique households that, taken together, made 7,158 shopping trips to buy vitamins. I eliminate from the sample the households that made multiple purchases of vitamins per trip. This does not change the pattern of switches between vitamins. As a result, the panel is reduced to 202 unique households and 3,623 shopping trips.

A typical household in the panel has 2.2 family members and buys multivitamins 2.2 times per year. Since an average pack of vitamins contains 163 tablets and one tablet is usually taken per day, there may be just one regular consumer of vitamin per household or there may be seasonal spikes in the consumption of vitamins.

On average, the household whose members consume vitamins has a total annual income of approximately $40,000-44,999. This annual income is tracked two years in advance of the panel year.

Older consumers are the core consumers of vitamins and the drivers of growth in this purchasing category. For the male head of the vitamin-consuming household, the mean age is 45-49 years, whereas the female head is about 50-54 years old. 32% of the male heads are retired or unemployed according to the Nielsen Company occupation classification. 21% - graduated high school, 21% have some college education, and 19% graduated from college. In comparison, 34% of the female heads graduated high school, 23% went to college, 22% graduated college, and 46% are retired or unemployed.

171 households, or 85% of the sample, have no children under 18; this corresponds to the age profile of the households. 18 households have three kids, 5 households have one child, and 4 households have two. 64% of the households consist of the married couples, 18% are females living alone, and 8% are males living alone. For other details on the household composition, please see Appendix A.

I combine 325 of the available Universal Product Codes (UPCs) for *multivitamins* into five brands. A Private Label Brand is created from all private label brands and accounts for 78% of all purchases. Brand 2 was chosen during 326 shopping trips thereby representing 9% of the purchases. Brand 3 constitutes 5.3% of the vitamin purchases, and Brand 4 – 1.66%. Brand 5 consists of all the remaining brands and accounts for 6.21% of the purchases during shopping trips.

The Private Label Brand or Brand 1 is significantly cheaper than the other brands and has the most pills per bottle, as shown in Appendix B.

3 Identification

Identification of learning model parameters requires a sufficiently rich set of observed switches among vitamins. In 1,541 out of 3,623 shopping trips (45%), households chose different UPC items than the ones they bought during the previous shopping trips. The same proportion of switches occurred with the Private Label Brand. The households buying Brand 2 switched 70 times, or during 25% of all shopping trips for this brand of multivitamin. Within Brand 3, the households changed their preferences 53 times or during 33% of the purchases; within Brand 4, they shifted 14 times, or during 32% of the purchases. The households selecting Brand 5 had only 23 switches, for a total of 12%. This may be because Brand 5 consisted of small distinct brands that attracted loyal consumers.

In noticing these switches, it is important to remember that the households with more than one member might be buying vitamins for other family members during the different shopping trips, i.e. a wife buys a special female-targeted vitamin for herself one day and a special male-targeted vitamin for her husband another day. The wife would not have received any new information about her vitamin, but the data would identify her as having practiced learning because of the observed switch. However, in my data set, the same proportion of switches (45%) holds even when conditioned on the households with one member only. Please check Appendix C for brand-specific variations.

The average household makes eight switches over the course of seven years while nine times it chooses the same UPC item as during the previous shopping occasion. This variation in the purchases enables identification of the learning coefficients.

In Figures 1-5 of Appendix D illustrate the variation in prices for each of the five brands. The variation is reported for the most popular size of the bottle within each brand category and can support the identification of model parameters.

4 Reduced Form Model

To test for learning, I estimate a linear probability model for two brands. Brand 1 combines all private label brands, and Brand 2 combines all other brands of multi-vitamins. As shown in the Data Section, Brand 1 is generally cheaper, so shoppers may trade down to it in order to save money.

$$P_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon_i \quad (1)$$

where P_i is a dummy variable indicating whether or not the consumer i bought Brand 1. β_0 is a constant and ϵ_i is an independently and identically distributed error term. The parameters of interest are β_1 and β_2 . They show the effects of past experiences on the probability of buying Brand 1.

I assume that consumers had positive experiences with the product if they bought this product in period $t-3$ and then again in period t. They could have bought something else during period $t-2$ and period $t-1$, but in period $t-3$, any uncertainty has been resolved. In this way, the speed of learning is considered to be three periods. I introduce an explanatory variable called *A Positive Experience from Buying Brand* based on this and such that:

- x_1 is a dummy variable indicating whether a consumer had a positive experience with Brand 1. x_1 equals zero if the consumer bought Brand 1 in period $t-3$ but not in period t; it also equals zero if the consumer did not buy Brand 1 in period $t-3$. I expect the estimates of β_1 to be positive, because a positive past experience with Brand 1 would increase the probability of choosing it again.

- x_2 is a dummy variable indicating whether a consumer had a positive experience with Brand 2. A positive experience with Brand 2 decreases the likelihood that the consumer will choose Brand 1, and I expect the estimates of $\beta_2$to have a positive sign.

Since the purchasing of one brand may also reveal information about the other brand, an interaction term of the positive experiences with Brand 1 and Brand 2 is introduced to Model (1):

$$P_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \epsilon_i \tag{2}$$

In my understanding, having positive experiences with taking any brand of vitamins increases the probability of buying Brand 1. Consequently, the estimates of β_3 from Model (2) should be positive.

5 Results

The estimation of the linear probability model produces similar results with the logit and probit models, as shown in Appendix E. All coefficients are statistically significant and show that the probability of choosing Brand 1 increases alongside consumers' positive experiences of Brand 1 and decreases alongside consumers' positive experiences of Brand 2.

Appendix F shows the predicted probability from Models (1) and (2).

The first thing to notice is that the predictions of the linear probability model are in line with the logit and probit models.

When consumers have no experience with Brand 1 or Brand 2, they will choose Brand 1 with a probability between 68-73%. If consumers have had positive experience with Brand 1 but do not know anything about Brand 2, they will choose Brand 1 almost surely, with a probability between 90-93%. Having positive experiences with Brand 2 and no experience with Brand 1 dramatically decreases the probability of buying Brand 1, to about 10-18%. Finally, if the consumers have tried both brands, Model (2) predicts that they will buy Brand 1 with a probability of 64%. An outcome is not that clear with Model (1) since the predicted probability of choosing Brand 1 ranges from 44 to 51%.

To conclude, since Brand 1 is generally cheaper, users may trade down to it in order to save money. If the shoppers have tried both brands, it is not clear which brand they will choose to purchase.

6 Further Steps

In the previous sections, I demonstrated the presence of learning. As I develop this study further, I will extend it into a structural model with the endogenous speed of learning. Then, I plan to also compare the speed of learning about vitamins with the speed of learning about drugs that have the same frequency of usage as vitamins.

Dickstein (2014) presents spillovers in learning for the antidepressant market. In the antidepressant market, there is a natural clustering of the available treatments according to the way the drugs function in the brain. Consumers may learn about the quality of vitamins in a correlated fashion, too. After a poor outcome from a vitamin sold by Brand 1, for example, the consumer may avoid Brand 2 and Brand 3 if they share Brand 1's characteristics, i.e. size of pill. Because the purchasing of one brand reveals information not only about that specific brand but also about vitamins as a group, signals about the quality of the brands become correlated. From the number of switches in purchases, it is clear that people often try different brands. On the other hand, Appendix G reveals that these switches mostly happen within the brands. Consequently, it seems that within in a consumer's mind, brands exist as clusters, and the model can be estimated following Dickstein (2004) along with the index rule by Pandey et al. (2007).

In the case of vitamins, I have a left-truncation problem since the history of vitamins' consumption is not observable prior to 2004. Ackerberg (2003) does not face this problem because Yoplait 150 is a brand that is new to the market and the consumers do not have a history of previous purchases. Hendel and Nevo (2006) have a left-truncation problem with their dynamic storable goods model because they do not know how much detergent the consumer has in inventory when the data set starts. They use the distribution of inventories over the entire population after three periods as the distribution of possible inventories for each consumer at $t = 1$. Crawford and Shum (2005) solve the problem by including only patients who are first observed after the sixth month of the sample. To address the problem of the left-censoring in my data, I plan to exclude from the data households that do not buy vitamins in 2004 but start buying them in 2005 and continue purchasing them through 2011. Since there are only 96 of such households, I will construct an unbalanced panel with an outside option/brand of not buying vitamins. In this way, excluding any households that did not buy vitamins in 2004 will not reduce the data that dramatically.

References

Ackerberg, Daniel, (2003). 'Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination', *International Economic Review*, 44(3): 1007-1040.

Crawford G. and M. Shum. (2005). 'Uncertainty and learning in pharmaceutical demand', *Econometrica*, 73(4): 1137-1173.

Dickstein, Michael, (2014). 'Efficient Provision of Experience Goods: Evidence from Antidepressant Choice', working paper.

Hendel I. and A. Nevo (2006). 'Measuring the Implications of Sales and Consumer Inventory Behavior', *Econometrica*, 74(6): 1637-1673.

Pandey, S., D. Chakrabarti, and D. Agarwal. (2007) 'Multi-Armed Bandit Problems with Dependent Arms', *Proceedings of the 24th International Conference on Machine Learning*, ACM, 721-728.

Vitamins, Minerals and Supplements. Issues and Insights, *The Mintel Group Report*, December 2013: 2.

Appendix A

Table 1. Household Composition

	Freq.	Percent	Cum.
Married	129	63.86	63.86
Female Head Living wiht Others Related	13	6.44	70.30
Male Head Living with Others Related	2	0.99	71.29
Female Living Alone	36	17.82	89.11
Female Living with Non-Related	1	0.50	89.60
Male Living Alone	16	7.92	97.72
Male Living with Non-Related	5	2.48	100.00
Total	202	100.00	

Appendix B

Table 1. Summary Statistics of Prices Per Pill

	Obs	Mean	Std. Dev.	Min	Max
Brand 1	2820	0.038	0.017	0	0.449
Brand 2	326	0.107	0.051	0	0.299
Brand 3	192	0.061	0.043	0	0.217
Brand 4	60	0.117	0.037	0.03	0.187
Brand 5	225	0.142	0.103	0.00	0.358

Table 2. Summary Statistics of Bottle Sizes

	Obs	Mean	Std. Dev.	Min	Max
Brand 1	2820	205.56	115.93	10	700
Brand 2	326	113.60	88.51	50	250
Brand 3	192	123.30	68.59	50	300
Brand 4	60	110	77.66	45	325
Brand 5	225	103.02	33.04	30	200

Appendix C

Table 1. Frequency of Brands Purchases for Households of Size One

	Freq.	Percent	Cum.
Brand 1	674	75.65	75.65
Brand 2	41	4.60	80.25
Brand 3	89	9.99	90.24
Brand 4	16	1.80	92.03
Brand 5	71	7.97	100.00
Total	202	100.00	

Appendix D

Figure 1. Prive Variation Over 2004-2011, Brand 1 for the most popular bottle size - 100 pills

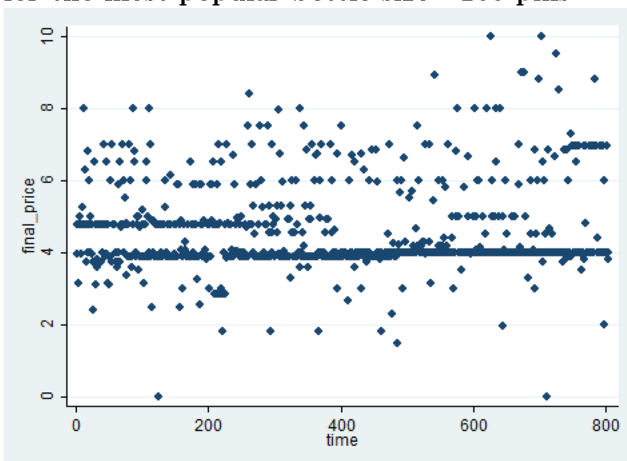


Figure 2. Prive Variation Over 2004-2011, Brand 2 for the most popular bottle size - 50 pills

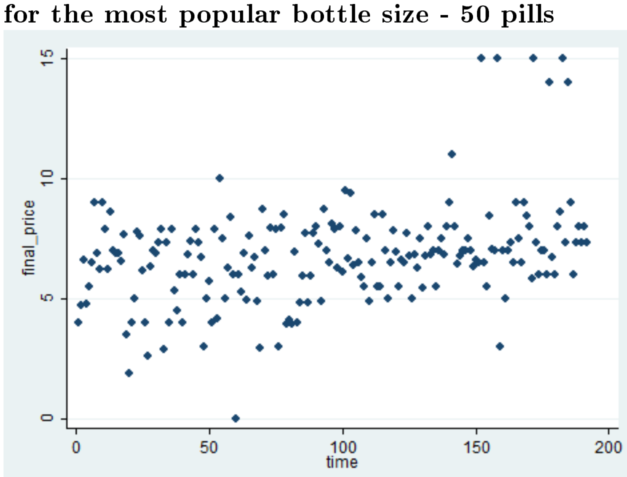


Figure 3. Prive Variation Over 2004-2011, Brand 3 for the most popular bottle size - 90 pills

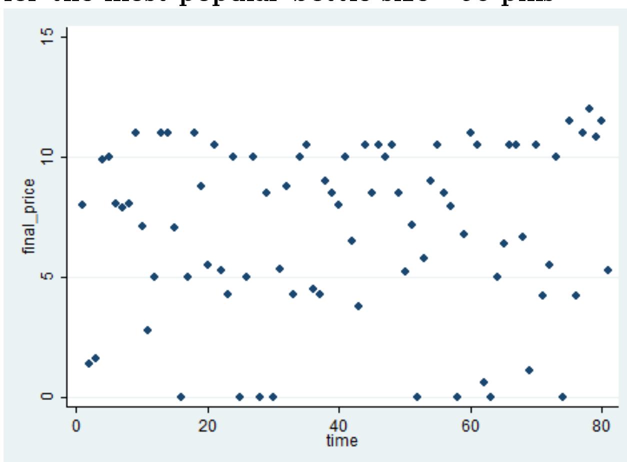


Figure 4. Prive Variation Over 2004-2011, Brand 4 for the most popular bottle size - 100 pills

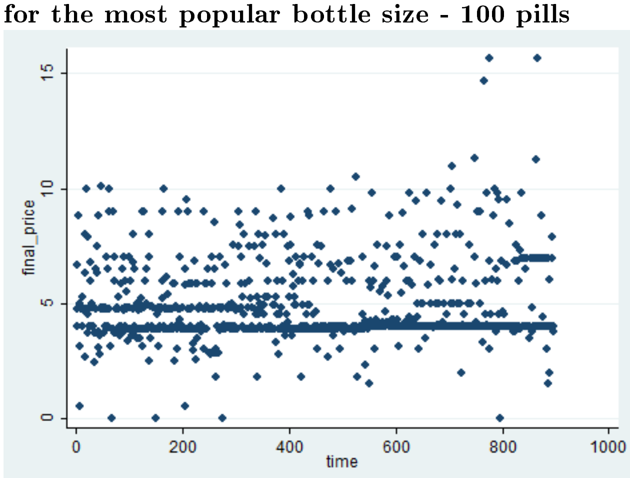
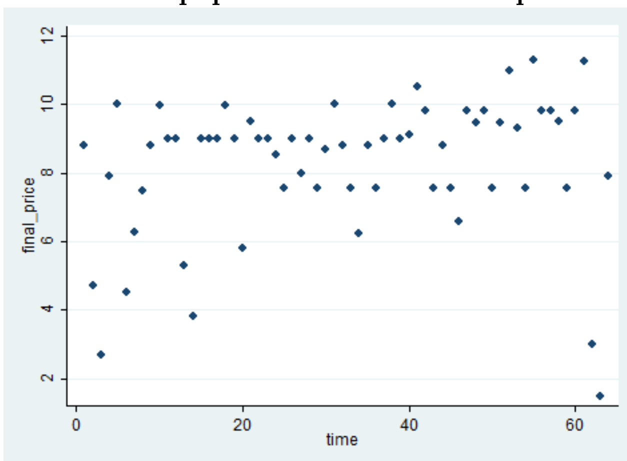


Figure 5. Prive Variation Over 2004-2011, Brand 5 for the most popular bottle size - 100 pills



Appendix E

Table 1. Three Period Learning: Linear Probability Model Estimates

	Single Brand	Two Brands	w/ Interaction
Constant, $\hat{\beta}_0$	0.500 (0.011)	0.680 (0.011)	0.730 (0.012)
Positive Experience Brand 1, $\hat{\beta}_1$	0.385 (0.014)	0.250 (0.013)	0.179 (0.014)
Positive Experience Brand 2, $\hat{\beta}_2$	-	-0.490 (0.016)	-0.626 (0.019)
Interaction, $\hat{\beta}_3$	-	-	0.362 (0.032)
Number of Observations	3,623	3,623	3,623
R^2	0.180	0.355	0.377

Note: Dependent variable: dummy on whether the consumer bought Brand 1. Standard errors in parentheses

Table 2. Three Period Learning: Logit Model Estimates

	Single Brand	Two Brands	w/ Interaction
Constant, $\hat{\beta}_0$	0.000 (0.054)	0.828 (0.069)	0.994 (0.077)
Positive Experience Brand 1, $\hat{\beta}_1$	2.035 (0.085)	1.662 (0.095)	1.309 (0.109)
Positive Experience Brand 2, $\hat{\beta}_2$	-	-2.469 (0.106)	-3.150 (0.165)
Interaction, $\hat{\beta}_3$	-	-	1.442 (0.232)
Number of Observations	3,623	3,623	3,623
Log Likelihood	-1754.0373	-1455.9451	-1431.1295

Note: Dependent variable: dummy on whether the consumer bought Brand 1. Standard errors in parentheses

Table 3. Three Period Learning: Probit Model Estimates

	Single Brand	Two Brands	w/ Interaction
Constant, $\hat{\beta}_0$	0.000 (0.048)	0.499 (0.042)	0.613 (0.046)
Positive Experience Brand 1, $\hat{\beta}_1$	0.198 (0.048)	0.923 (0.053)	0.723 (0.060)
Positive Experience Brand 2, $\hat{\beta}_2$	-	-1.472 (0.062)	-1.873 (0.088)
Interaction, $\hat{\beta}_3$	-	-	0.908 (0.131)
Number of Observations	3,623	3,623	3,623
Log Likelihood	-1754.0373	-1455.9451	-1431.1295

Note: Dependent variable: dummy on whether the consumer bought Brand 1. Standard errors in parentheses

Appendix F

Table 1. Predicted Probability of Choosing Brand 1, $\hat{P}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$

	Linear Probability	Logit	Probit
Brand 1 (-), Brand 2 (-)	0.68	0.70	0.69
Brand 1 (+), Brand 2 (-)	0.93	0.92	0.92
Brand 1 (-), Brand 2 (+)	0.18	0.16	0.17
Brand 1 (+), Brand 2 (+)	0.44	0.51	0.48

Note: (+) if experience is positive, (-) if no experience

Table 2. Predicted Probability of Choosing Brand 1, $\hat{P}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_1 x_2$

	Linear Probability	Logit	Probit
Brand 1 (-), Brand 2 (-)	0.73	0.73	0.73
Brand 1 (+), Brand 2 (-)	0.90	0.91	0.91
Brand 1 (-), Brand 2 (+)	0.10	0.10	0.10
Brand 1 (+), Brand 2 (+)	0.64	0.64	0.64

Note: (+) if experience is positive, (-) if no experience