Logit Demand Profit Maximization

Matthew Guzman

March 2021

Abstract

This article presents recent research on profit maximization under logit demand. Logit demand models represent a popular tool in discrete choice literature, and a computationally tractable methodology to solve the share equations that result was developed in the early 2000s. This methodology makes use of the LambertW equation, and once it is implemented, the profit maximizing price can be found using popular statistical tools such as R Studio, Python, and MATLAB. Current research in this field is investigating a generalized methodology to solve for the profit maximizing price under nested logit models as well.[1]

[1]The author would like to thank Joseph Podwol, Jeff Qiu, and Robert Draba for their helpful comments and constant support for this article. All errors are my own.

Contents

1 Introduction

Since the late 1970s, discrete choice models have become increasingly popular tools to estimate consumer demand. These models represent flexible, easily implementable mathematical representations of systems, and since 2000, multiple extensions have been developed to better model realistic consumer behavior.

1.1 Literature Review

A "discrete choice" model is a mathematical tool that simulates the likelihood that an individual will choose one of multiple options. These options can be for truly anything, and researchers have used these models to study such varied products as residential housing choices[2], ground coffee[3], automobile manufacturing[4], and breakfast cereals[5]. [**Need to add detail on the discete choice literature.**]

1.2 The Logit Demand Discrete Choice Model

A key characteristic of a discrete choice model is the inclusion of two product-specific parameters. In this paper, we will express these parameters as a and b. Here, a represents the "utility" for the consumer, and b represents the "price coefficient." In this way, the value a consumer, say Jane, receives from the product or the "utility" can be modeled in the following way:

$$U_{Jane} = a_{Jane} + b_{Jane}p$$

Here, we are adding Jane's utility a to Jane's price coefficient b multiplied by the price p. Usually, the price coefficient is negative, but this doesn't need to be the case. Aggregating these utility functions over many consumers, we can fit a model to all consumers' utilities that takes the following form:

$$U = a + bp + \epsilon$$

This is exactly the same equation, but we have now taken the average of all consumers' utilities and their price coefficients. To that end, we refer to a as the "mean utility", and we add an error term ϵ to this equation to account for idiosyncrasies among particular consumers.

These parameters allow for semi-realistic simulations of consumer choice. Consumers choose products based on their personal valuations, and this valuation is captured in the mean utility variable a. Furthermore, the price has some impact on each consumer's decision, and the importance of the price is captured by the price coefficient b. Using datasets that indicate how many individuals purchased each product at different prices, we can estimate the values of a and b, and this information can inform us as analysts and policy makers on what consumers actually want.

Notably, these parameters circumvent the need to use survey data to understand consumers' preferences. Discrete choice models rely on "revealed preference," which is to say that they model consumers' values based on what they are choose, not what they claim that they would choose.

We can mathematically derive the share equations for each option available to consumers in a general discrete choice problem using this setup. These derivations are outside the scope of this paper, and we will instead focus on the "logit" model in this paper. The logit discrete choice model is a shorthand way of saying that the error term in the equation is distributed according to a multivariate extreme value. When this is the case, the share equations have closed form solutions of the following form:

$$s_i = \frac{\exp(a_i + b_i p_i)}{\exp(a_i + b_i p_i) + \exp(a_j + b_j p_j)}$$

In this case, a consumer has two choices i and j. The share of option i, or the probability that a consumer will choose option i, can be given be the equation above which is the exponential of the mean utility plus

[2]See McFadden 1978
[3]See Guadagni and Little, 1983
[4]See Berry, Levinson, and Pakes 1995
[5]See Nevo, 2001

the price coefficient multiplied by the price of option i, divided by this value plus the exponential of mean utility plus the price coefficient multiplied by the price of option j.

A key feature of the logit demand discrete choice model is the ease with which it can be scaled. Discrete choice models implicitly model multiple options, and thus they lend themselves to systems in which the total options range from 5 options to 500 options. In the logit model, this can be completed straightforwardly by adding the exponential values to the denominator of each share function.

With this background in mind, this article describes the logit demand model and demonstrates a methodology to find the profit maximizing price under its assumptions. A closed form solution to this problem was discovered in the early 2000s, and this represents a particularly useful application for businesses.

The rest of this article is organized as follows. Section 2 provides a methodology for solving for the Bertrand-Nash equilibrium and profit maximizing price under logit demand with two firms and no outside good. Section three generalizes this model to solve for the equilibrium with N firms. Section four generalizes this further to solve for the equilibrium with N firms under a nested logit demand model. Section five provides a summary of these results. A discussion of the convergence methodology and additional derivations are provided in the Appendices, and code to solve for the profit maximizing price under logit demand is provided on github.

2 Bertrand-Nash Duopoly Equilibrium under Logit Demand

2.1 Problem Setup

Let's set up the Bertrand-Nash equilibrium problem with two firms under a logit demand discrete choice model. Suppose that there are two coffee shops in downtown Washington, D.C.: Starbucks and Dunkin Donuts. Let's go ahead and assume that these companies do not engage in price discrimination and there is no cost associated with switching to a different coffee shop. Additionally, all government employees are addicted to coffee, so consumers cannot choose to not drink coffee. Thus, government employees in this area have only the following two coffee options:

1. Starbucks coffee
2. Dunkin coffee

In addition to the imaginary scenario we have set up, let's assume that no one is buying coffee from both Starbucks and Dunkin Donuts. Also, Starbucks and Dunkin are not colluding, and each company wants to maximize its profits in downtown D.C.

2.2 Logit Demand Setup

We now consider the pricing problem faced by these two competing coffee shops, and we treat their products as imperfect substitutes in the minds of the government employees in the Washington D.C. area.

Let s_S denote the fraction of government workers N that chooses Starbucks and let s_D denote the fraction that chooses Dunkin Donuts. These shares can be interpreted as Starbucks and Dunkin Donuts' "market shares". A logit representation of these shares is:

$$s_S = \frac{\exp(a_S + b_S p_S)}{\exp(a_S + b_S p_S) + \exp(a_D + b_D p_D)}$$

$$s_D = \frac{\exp(a_D + b_D p_D)}{\exp(a_S + b_S p_S) + \exp(a_D + b_D p_D)}$$

Here, a represents the hypothetical "mean utility" that a consumer receives from buying coffee, and b represents the hypothetical "price coefficient," or a quantitative measurement of how much the price matters to the consumer. Note that the price coefficient is always negative by assumption in this model.

In addition to these market share equations, we know that each company in this model has a profit equation of the following form:

$$\pi = (p - c)Ns$$

Here, the profit denoted by π is equal to the price charged per unit of coffee minus the cost of the coffee, multiplied by the firm's share and the total size of the market.

In a typical economic or antitrust analysis, the main intent of this model would be to estimate the consumer parameters a_S, b_S, a_D, and b_D. Usually, data would provide the market shares, prices, and costs. Given these values, we could solve the equations above for the consumer parameters. This type of analysis could inform our understanding of how much consumers value coffee through the mean utility parameter a and how important the price is when they make their decisions as indicated by the value of the price coefficient b.

On the other hand, we can also use this model to calculate the market equilibrium. Instead of using observed data to estimate the shares to solve these equations, we could estimate the consumers' utilities and price sensitivities to solve these equations for the market shares. In this way, we could decide what the hypothetical consumer values a_S, b_S, a_D, and b_D should be and solve for the market shares using those values. Then, if we further choose hypothetical costs to the firms, we can even calculate what the profit-maximizing prices, shares, and profits would be in equilibrium given all of these assumptions.

This second option represents an extremely useful tool for firms. Companies such as coffee shops may need to regularly adjust prices based on their own costs and limited information about their competitors prices. The logit model provides an ideal tool to predict optimal pricing levels given this limited information. Furthermore, the mean utility parameters and the price coefficients are very flexible, so analysts can attempt different combinations of these variables in order to tweak consumer demand in realistic ways.

With that, the rest of this section outlines how to solve for the profit-maximizing prices under logit demand given consumer values and costs to firms. You can go ahead and assume the following values:

1. $a_S = 6$, $b_S = -0.8$, $c_S = 1.50$
2. $a_D = 5$, $b_D = -1$, $c_D = 1.25$
3. N = 1000

Here, we are assuming that the average consumer receives 6 utility from a Starbucks coffee, Starbucks's price coefficient is -0.8, and it costs Starbucks \$1.50 to make a coffee. Furthermore, the average consumer only receives 5 utility from a Dunkin Donuts coffee, Dunkin's price coefficient is -1, and it costs Dunkin \$1.25 to make a coffee. Finally, we are assuming that there are 1000 government employees buying coffee.[6]

So, what do we need to do in order to find the equilibrium prices and profits? We can start by writing each firm's profit function:

$$\pi_S = (p_S - c_S)Ns_S$$

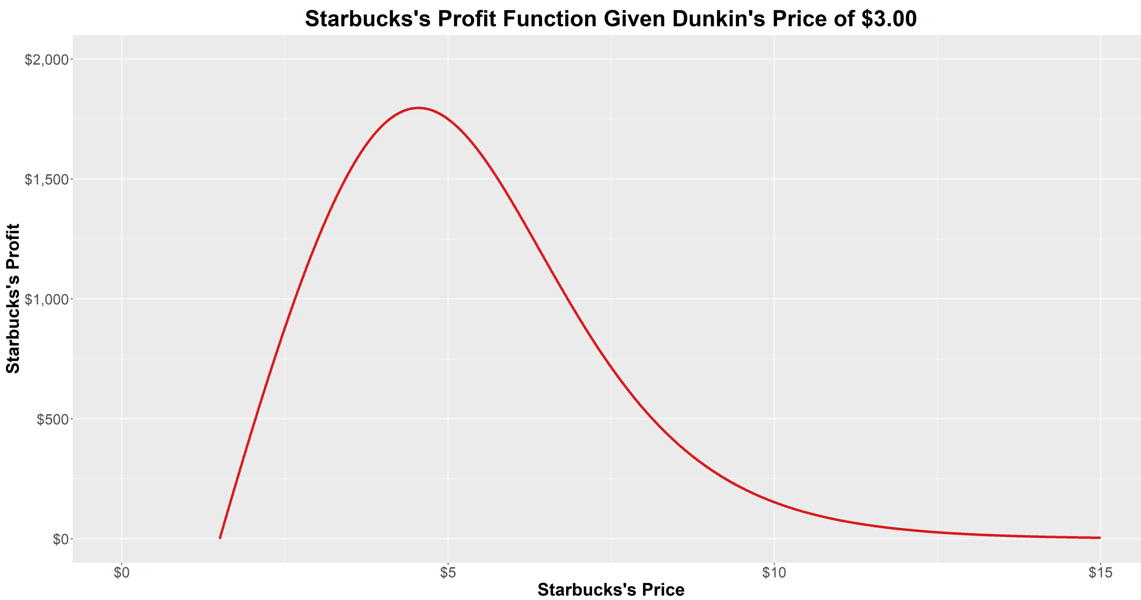
$$\pi_D = (p_D - c_D)Ns_D$$

Breaking these equations down, each firm's profit denoted by π is equivalent to its price minus its cost multiplied by the number of consumers in the market multiplied by the firm's share. This can be more easily interpreted as the amount of money each firm makes on each coffee, multiplied by the number of coffees it sells.

2.3 Geometric Setup

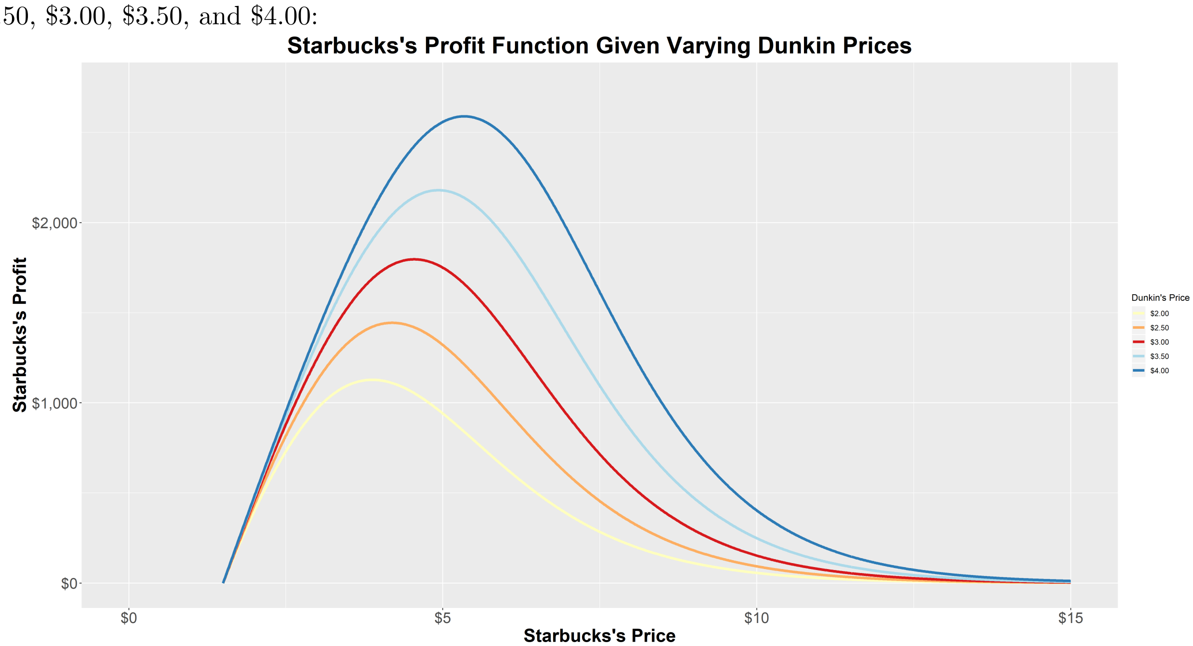
Let's develop some intuition for this problem by looking at it geometrically. Let's assume that Dunkin's price is set to \$3.00. Plugging \$3.00 into Starbucks's profit equation above, we can graph Starbucks's profit as Starbucks changes its price:

[6] Note that the size of the market N is really just a scaling factor. We can think of this more realistically as 10,000 government employees in the downtown D.C. area, but this parameter does not actually impact the equilibrium prices.



Here, Starbucks's profit maximizing price occurs near the \$4.55 mark, and Starbucks's market share is 58.9%.[7] Therefore, if Starbucks were to face a price of \$3.00 from Dunkin (and if Starbucks assumed the same mean utility and price coefficient parameters that we set above), then Starbucks would know that its best strategy would be to charge \$4.55 for each cup of coffee. This would lead to a profit of $(\$4.55 - \$1.50) * 1000 * .589 = \$1,796.45$ for Starbucks and a profit of $(\$3.00 - \$1.25) * 1000 * .411 = \$719.25$ for Dunkin Donuts.[8]

This example illustrates how we want to find the profit maximizing price, but how can we achieve this when both firms can change their prices in response to the other firm? In the slice above, Starbucks knew that Dunkin would price at \$3.00 and it could respond accordingly. Let's consider a few different options. In the figure below, we see five curves that illustrate Starbucks's profit curve when Dunkin prices at \$2.00, \$2.50, \$3.00, \$3.50, and \$4.00:



From these curves, we can see that Starbucks should price lower when Dunkin prices lower, and Starbucks

[7] We calculated Starbucks's market share by plugging these two prices into Starbucks's share equation above.

[8] Notice that if Starbucks were to match Dunkin's price of \$3.00, then Starbucks and Dunkin would have an equivalent profit of $(\$3.00 - \$1.50) * 1000 * .50 = \$750.00$. Therefore, Starbucks can actually do better if it chooses to price higher than Dunkin.

should price higher when Dunkin prices higher. The question is, how should Starbucks choose its price when it doesn't know how Dunkin will price? At best, Starbucks has an approximate guess of where Dunkin is going to price, but it has no idea which of these prices Dunkin will choose.[9]

To make things more complex, Dunkin is actually trying to do the same thing with Starbucks's price, and this implies that the two coffee shops are actually changing their prices at the same time. Dunkin might choose \$3.00, then realize that Starbucks is undercutting them with a price of \$2.75 and then change their price to \$2.50, and Starbucks might react after that as well. The key point here is that both firms will try to maximize their profits depending on the other firm's price, and we want to know where the "equilibrium" is. That is, eventually the two firms will stop undercutting each other and choose a set of prices at which neither can lower their price without being worse off. The question is, what is that set of prices?

To answer this question, we want to *maximize* the profit functions for each firm. Firms want to make the most money they can given the fact that other firms will try to do the same. Mathematically, we can write this as:

$$p_S^* = \underset{p_S}{\text{argmax}}(p_S - c_S)Ns_S$$

$$p_D^* = \underset{p_D}{\text{argmax}}(p_D - c_D)Ns_D$$

Here, we just rewrote the equations above with the "argmax" operator with respect to p_S and p_D. This is to say that we want to know the values of p_S and p_D that return the highest outputs of the profit functions. We know that at the equilibrium, the values of p_S^* and p_D^* are exactly the profit maximizing prices for Starbucks and Dunkin. Note that here, the $*$ indicates that these values for p_S and p_D are the highest profit-yielding choices for the two coffee shops.

2.4 Algebraic Setup

So, this begs the question of how to find the maximum outputs of these two profit functions in two variables. We will find where these profit functions change from increasing to decreasing, and then we will find the intersection of profit maximizing choices for Starbucks and Dunkin. As you may have guessed, this implies that we will take the First Order Conditions of both profit functions, and then we will solve these two equations for the profit maximizing prices. That is, we will find where the derivatives of these functions are equal to 0, and then we will solve the system of equations that results.

Let's look at the derivative of each of the coffee shops' profit functions. Notably, these are functions of two variables. Returning to your multi-variable calculus days, you may recall that the **partial** of an equation of two variables $f(x, y) = x^2 + y^2$ is:

$$f_x(x, y) = f_x = \frac{\partial f}{\partial x} = \frac{\partial}{\partial x} f(x, y) = 2x$$

Recognizing that our profit functions are also equations of two variables ($\pi = f(p_S, p_D)$ if you were keeping track), we have the following **partial derivatives**:[10]

$$\frac{\partial \pi_S}{\partial p_S} = Ns_S + N\frac{\partial s_S}{\partial p_S}(p_S - MC_S)$$

$$\frac{\partial \pi_D}{\partial p_D} = Ns_D + N\frac{\partial s_D}{\partial p_D}(p_D - MC_D)$$

Setting these equal to 0, we have:

$$0 = Ns_S + N\frac{\partial s_S}{\partial p_S}(p_S - MC_S)$$

$$0 = Ns_D + N\frac{\partial s_D}{\partial p_D}(p_D - MC_S)$$

[9] See Appendix 1 for Starbucks's three-dimensional profit surface.

[10] Note that we are not taking the **total derivatives** of the Starbucks and Dunkin's profit functions. While the derivative of Starbucks's profit with respect to Dunkin's price does exist, we don't particularly care what this value is. This is a real-world example, and in the real world, Starbucks and Dunkin each only have control over their own prices. Since Starbucks can't change Dunkin's price, it doesn't make sense to model Starbucks's profit with respect to Dunkin's changing price.

We just need to solve these two equations to get what we need. We have two equations, are our two unknowns are p_S and p_D. Let's go ahead and solve these equations for p_S and p_D:

$$p_S^* = MC_S - \left(\frac{\partial s_S}{\partial p_S}\right)^{-1} \cdot s_S$$

$$p_D^* = MC_D - \left(\frac{\partial s_D}{\partial p_D}\right)^{-1} \cdot s_D$$

One problem though - you will note that these equations are still in terms of s_S and s_D, and these terms are in fact functions of p_S and p_D. Specifically, the equations for the shares s_S and s_D are exponential equations of p_S and p_D. While these derivatives are easy enough to calculate and put into their respective equations, you will find that it is in fact impossible to algebraically solve for p_S and p_D not in terms of s_S and s_D. The technical term for this is that there is no "closed form" solution for these equations. In principle, we have two equations and two unknowns, but we can't solve for the profit maximizing prices in a straightforward manner.

Instead of solving these equations directly, we will approach them using the "guess and check" method. The idea of this methodology is to try prices essentially at random until we find a "fixed point" of the function, or the point at which the function stops changing. We will conduct this in an iterative fashion - that is, we will try a set of prices (p_{S_0}, p_{D_0}) on the right hand side of the equation and we will calculate the resulting (p_S^*, p_D^*). Then, we will calculate the distance between (p_{S_0}, p_{D_0}) and (p_S^*, p_D^*). Once this distance is less than a very small number (say 0.00001), we will be able to safely assume that we have reached a "fixed point" and that the resulting (p_S^*, p_D^*) are the profit-maximizing prices.[11]

2.5 Solve for the Profit Maximizing Prices

Now that the methodology is set, let's calculate the derivatives $\left(\frac{\partial s_S}{\partial p_S}\right)^{-1}$, and $\left(\frac{\partial s_D}{\partial p_D}\right)^{-1}$:

$$\left(\frac{\partial s_S}{\partial p_S}\right)^{-1} = \left(\frac{b_S \cdot \exp(a_S + b_S p_S) \cdot \exp(a_D + b_D p_D)}{(\exp(a_S + b_S p_S) + \exp(a_D + b_D p_D))^2}\right)^{-1} = \frac{(\exp(a_S + b_S p_S) + \exp(a_D + b_D p_D))^2}{b_S \cdot \exp(a_S + b_S p_S) \cdot \exp(a_D + b_D p_D)}$$

$$\left(\frac{\partial s_D}{\partial p_D}\right)^{-1} = \left(\frac{b_D \cdot \exp(a_D + b_D p_D) \cdot \exp(a_S + b_S p_S)}{(\exp(a_S + b_S p_S) + \exp(a_D + b_D p_D))^2}\right)^{-1} = \frac{(\exp(a_S + b_S p_S) + \exp(a_D + b_D p_D))^2}{b_D \cdot \exp(a_D + b_D p_D) \cdot \exp(a_S + b_S p_S)}$$

With these, the profit maximizing equations to solve are:

$$p_S^* = MC_S - \left(\frac{b_S \cdot \exp(a_S + b_S p_D) \cdot \exp(a_D + b_D p_D)}{(\exp(a_S + b_S p_S) + \exp(a_D + b_D p_D))^2}\right)^{-1} \cdot s_S$$

$$p_D^* = MC_D - \left(\frac{b_D \cdot \exp(a_D + b_D p_D) \cdot \exp(a_S + b_S p_S)}{(\exp(a_S + b_S p_S) + \exp(a_D + b_D p_D))^2}\right)^{-1} \cdot s_D$$

At this point, we could enter these equations into a computer program and use the "guess and check" methodology described above to solve the equations. However, this is an extremely finicky procedure in practice. Solving systems of nonlinear equations is always a challenge, and we can in fact use a mathematical trick to make this an easier problem to solve.

[11] See Appendix 2 for a more robust description of this procedure. This represents a computationally tractable way to approach this problem in a common programming language such as R or Python, however, there are numerous other approaches to solving systems of non-linear equations that could be useful here as well. For example, the "nleqslv" package in R could be used, or the "optim" command in base R would be another viable option to maximize these equations.

Let's put the profit maximizing price for Starbucks into a more workable form. Starting from the equation above, we have:

$$
\begin{aligned}
p_S^* &= MC_S - \left(\frac{b_S \cdot \exp(a_S + b_S p_S) \cdot \exp(a_D + b_D p_D)}{(\exp(a_S + b_S p_S) + \exp(a_D + b_D p_D))^2}\right)^{-1} \cdot s_S \\
&= MC_S - \left(\frac{b_S \cdot \exp(a_S + b_S p_S) \cdot \exp(a_D + b_D p_D)}{(\exp(a_S + b_S p_S) + \exp(a_D + b_D p_D))^2}\right)^{-1} \cdot \frac{\exp(a_S + b_S p_S)}{\exp(a_S + b_S p_S) + \exp(a_D + b_D p_D)} \\
&= MC_S - \left(\frac{(\exp(a_S + b_S p_S) + \exp(a_D + b_D p_D))^2}{b_S \cdot \exp(a_S + b_S p_S) \cdot \exp(a_D + b_D p_D)} \cdot \frac{\exp(a_S + b_S p_S)}{\exp(a_S + b_S p_S) + \exp(a_D + b_D p_D)}\right) \\
&= MC_S - \left(\frac{(\exp(a_S + b_S p_S) + \exp(a_D + b_D p_D))^{\cancel{2}}}{b_S \cdot \cancel{\exp(a_S + b_S p_S)} \cdot \exp(a_D + b_D p_D)} \cdot \frac{\cancel{\exp(a_S + b_S p_S)}}{\cancel{\exp(a_S + b_S p_S) + \exp(a_D + b_D p_D)}}\right) \\
&= MC_S - \left(\frac{\exp(a_S + b_S p_S) + \exp(a_D + b_D p_D)}{b_S \cdot \exp(a_D + b_D p_D)}\right) \\
&= MC_S - \frac{1}{b_S \cdot s_D} \\
&= MC_S - \frac{1}{b_S} \cdot \frac{1}{s_D}
\end{aligned}
$$

In the final expression, we know that the shares s_S and s_D sum to 1. With that in mind, we can make the following substitution:

$$
\begin{aligned}
p_S^* &= MC_S - \frac{1}{b_S} \cdot \frac{1}{s_D} \\
&= MC_S - \frac{1}{b_S} \cdot \frac{1}{1 - s_S} \\
&= MC_S - \frac{1}{b_S \cdot (1 - s_S)}
\end{aligned}
$$

This hasn't solved our problem yet, but we have now manipulated p_S^* to appear in terms of Starbucks's own market share. Now, we will use the "lambertW" function to find a closed form solution to this problem. The lambertW function, also called the omega function, was developed by Johann Heinrich Lambert in 1758, and it takes the following form:

$$W(x)e^{W(x)} = x$$

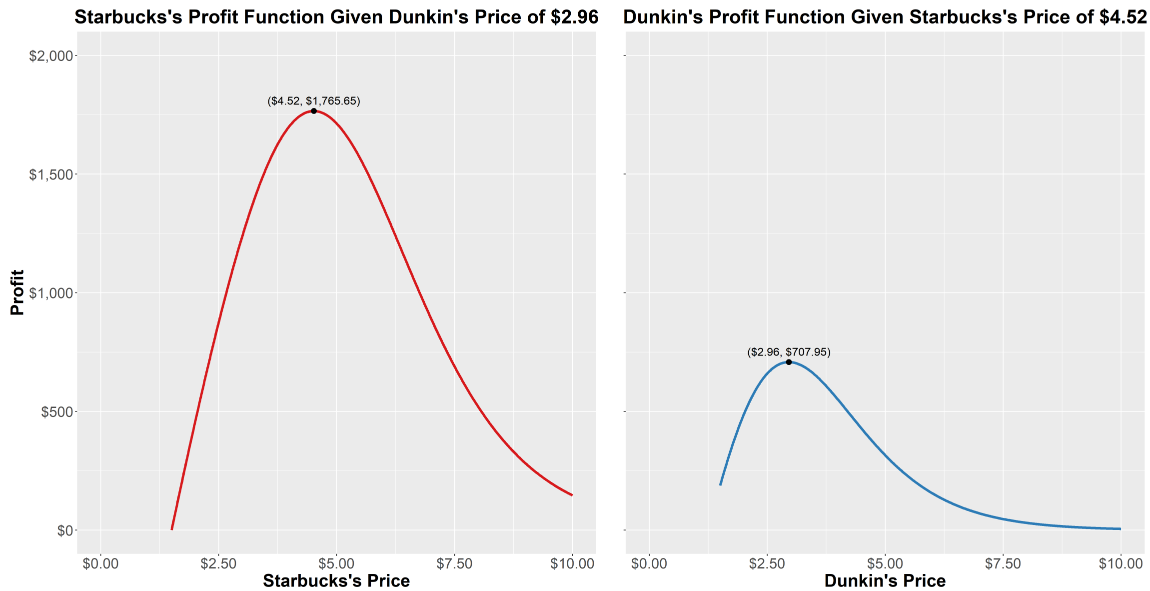
Using algebraic manipulation and the lambertW function, we can rewrite the optimal price equations above in closed forms as:

$$p_S^* = MC_S - \frac{1 + W\left(\frac{\exp(a_S - 1 + MC_S \cdot b_S)}{\exp(a_D + b_D \cdot p_D)}\right)}{b_S}$$

$$p_D^* = MC_D - \frac{1 + W\left(\frac{\exp(a_D - 1 + MC_D \cdot b_D)}{\exp(a_S + b_S \cdot p_S)}\right)}{b_D}$$

A proof of this equality is provided in Appendix 3, and we will solve these equations using the guess and check methodology described above. The lambertW function can be approximated very accurately, so this is a much more easily solved system of non-linear equations.

Solving these equations for p_S^* and p_D^*, we find that the profit maximizing prices for Starbucks and Dunkin are \$4.52 and \$2.96, respectively. We can graphically confirm this by looking at the two coffee shops' prices graphically below:



When Starbucks faces a price of \$2.96 from Dunkin, Starbucks will maximize its profit when it prices at \$4.52. Conversely, when Dunkin faces a price of \$4.52 from Starbucks, Dunkin will maximize its profit when it prices at \$2.96. Both firms would see a lower profit from this point, so we have found the Bertrand-Nash equilibrium for these firms with these consumer parameters.

3 Bertrand-Nash Equilibrium with N Firms under Logit Demand

3.1 Setup

Now that we have a strategy in place to solve for the profit maximizing price, let's expand this methodology to solve for the profit maximizing price with N firms. Rather than just considering Starbucks and Dunkin Donuts as coffee options, we now consider there to be any number of firms N from which consumers can choose to buy coffee in downtown D.C. In this generalized case, the share of firm i can be calculated in the following equation:

$$s_i = \frac{\exp(a_i + b_i p_i)}{\sum_{j=1}^{N} \exp(a_j + b_j p_j)}$$

Here, we are dividing the exponential of firm i's value by the sum of the exponentials of the values of all firms in the market.

Note that if there is an "outside good" assumed to be in the market with a value of 0, then the mean utility for the outside good will typically be $a = 0$ and the price will be 0. Therefore, the valuation of the outside good will be $\exp(0 - b_0 \cdot 0) = e^0 = 1$. An example with an outside good included will often be written with this assumption with the following form:

$$s_i = \frac{\exp(a_i + b_i p_i)}{1 + \sum_{j=1}^{N} \exp(a_j + b_j p_j)}$$

This would be different if we were to assume that the outside option had some value as well, but assuming that the outside option has a value of 0 is a very common assumption.

3.2 Algebraic Setup

We can set up the profit function for firm i in a very similar manner to the duopoly example. The profit function for coffee shop i is:

$$\pi_i = (p_i - c_i) N s_i$$

As in the two firm case, we need to take the first order condition of the profit function for each firm in order to solve for the profit maximizing prices. Taking the derivative of the profit function for firm i with respect to p_i, we have:

$$\frac{\partial \pi_i}{\partial p_i} = N s_i + N \frac{\partial s_i}{\partial p_i}(p_i - MC_i)$$

Setting this equal to 0, we have:

$$0 = N s_i + N \frac{\partial s_i}{\partial p_i}(p_i - MC_i)$$

Solving for firm i's profit maximizing price, we have:

$$p_i^* = MC_i - \left(\frac{\partial s_i}{\partial p_i}\right)^{-1} \cdot s_i$$

3.3 Solving for the Profit Maximizing Price

With this we need to solve the system of nonlinear equations with N equations where each equation is of the following form:

$$p_i^* = MC_i - \left(b_i \cdot \frac{\exp(a_i + b_D p_D)}{\sum_{i=1}^{N}(\exp(a_i + b_i p_i))^2}\right)^{-1} \cdot s_i$$

Applying the lambertW function from above, the closed form solution for the profit maximizing price under logit demand for firm i is given by:[12]

$$p_i^* = MC_i - \frac{1 + W\left(\frac{\exp(a_i - 1 + MC_i \cdot b_i)}{\sum_{i \neq j}^{j} \exp(a_j + b_j \cdot p_j)}\right)}{b_i}$$

Notably, as the number of firms N approaches infinity, the second term in the equation above approaches 0. At that point, the optimal price of firm i would be equal to its marginal cost.[13]

4 Introduction to Nested Logit Models

4.1 Introduction

In addition to finding the profit maximizing price with any number of firms, we can also find the profit maximizing price under a nested logit model. A "nested" logit model is very similar to the standard logit model described above, but in addition to comparing different products, the nested logit model allows products to be grouped into different "nests" based on their similarity.

Let's build some intuition for this problem by tweaking the duopoly example from before. In the original case, we looked at two products: a coffee from Starbucks and a coffee from Dunkin. In the real world though, there are many more options than standard coffee at each of these stores. Customers could pick several types of coffee drinks including mochas, cappuccinos, lattes, and standard coffees among others. This could quickly spiral into a very long list if we were to include every permutation of every coffee drink these stores sell, so for this example, we are just going to stop there. Suppose that Starbucks sells these four types of coffees, and and Dunkin sells three of these four options: mochas, lattes, and standard coffees.

Normally, we think of market shares with respect to firms. That is to say that when we analyze these two coffee shops, we would typically be interested in Starbucks's market share and Dunkin's market share. How can we calculate market shares when each of these stores sells multiple types of coffee? In fact, each store's market share is comprised of the shares of each of the products it sells. When we assume that Starbucks and Dunkin Donuts each only sell standard coffee, then each company only sells one product and each company's

[12] See Appendix 3 for derivation.

[13] Also of note: in the monopoly case, the denominator of the lambertW function is $e^0 = 1$.

market share is equal to the share of its own coffee. When we assume that each firm sells two types of coffee, then each company sells two products and each company's market share is equal to the sum of the shares of the two types of coffee it sells.

Because each product actually has its own market share, we actually need to set up as many share equations as the number of products. In this example, we have two coffee shops with four types of coffee at Starbucks and three types of coffee at Dunkin. Therefore, we need seven share equations in total.

Recognizing that we need seven market share equations in total, let's begin by just setting up the share equation for a Starbucks mocha in this world. Here, we need to look at the share of a Starbucks mocha within the mocha "nest":

$$s_{SMocha} = s_{S|Mocha} \cdot s_{Mocha}$$

In this equation, the share of a Starbucks mocha is equal to share of the Starbucks mocha in the mocha "nest" multiplied by the share of mochas in the entire world of coffee. For example, if 20% of all coffee sales were mochas and consumers chose the Starbucks mocha 50% of the time over the Dunkin mocha, then this share equation would be:

$$\begin{aligned} s_{SMocha} &= s_{S|Mocha} \cdot s_{Mocha} \\ &= 0.50 \cdot 0.20 \\ &= 0.10 \end{aligned}$$

In this way, the share of a Starbucks mocha would be 10% over the entire coffee market. Similarly, Starbucks lattes, cappuccinos, and standard coffees would all have their own shares in the entire market, and we could sum all of these to find Starbucks's total market share and total profit.

4.2 Share Equation Breakdown

Let's break down the share equation from above. In that formula, we set up two terms: the share of a Starbucks mocha within the mocha "nest" $s_{S|Mocha}$ and the share of mochas in the overall market S_{Mocha}. We can model the first term explicitly in the equation below:

$$s_{S|Mocha} = \frac{\exp(\frac{a_S + b_S \cdot p_S}{\sigma_{Mocha}})}{\exp(\frac{a_S + b_S \cdot p_S}{\sigma_{Mocha}}) + \exp(\frac{a_D + b_D \cdot p_D}{\sigma_{Mocha}})}$$

This is very similar to the share equation for one product that we considered before. Here, we are dividing the exponential of the sum of the consumer parameters and dividing by the sum of the exponentials of consumer parameters of all products in this nest. In this case there are exactly two products in the mocha "nest", and they are a Starbucks mocha and a Dunkin mocha.

Notably, there is one additional parameter σ_{Mocha} that was not present in the original logit demand example. This is called a "nesting parameter," and we assign each nest its own nesting parameter between 0 and 1. Intuitively, we are dividing each consumer valuation by a fraction which is increasing the valuation by some amount. Smaller values of σ_{Mocha} lead to higher overall valuations, than larger values of σ_{Mocha}, and because the valuation of each product in the mocha "nest" is divided by σ_{Mocha}, each valuation for the nest is "scaled" by the same amount.

This "nesting parameter" approach provides a mathematical way to account for the fact that some products are better substitutes than others. Consumers' valuations of a mocha from Starbucks and a mocha from Dunkin will be "scaled" by the same amount and will be slightly closer in terms of their values than consumers' valuations of lattes. If the price of a Starbucks mocha increases, then a consumer might switch to a Dunkin mocha as that might be the closest substitute. Alternatively however, consumers might choose to switch to a latte from Starbucks instead because lattes are less expensive. This represents a very flexible model that can incorporate any number of products, and products could be grouped in to nests in numerous ways.[14]

[14] Here, we are grouping types of coffees into nests. This assumes that a mocha from Starbucks is a better substitute for a mocha from Dunkin than a latte is. Alternatively, we could group all Starbucks products into a nest. This would imply that a mocha from Starbucks is a better substitute for a latte from Starbucks than a mocha from Dunkin. This might be a more realistic model if consumers are unlikely to go to leave Starbucks to go to Dunkin Donuts instead.

Now that we've walked through the explicit formula for $s_{S|Mocha}$, let's turn to the explicit formula for s_{Mocha}:

$$I_{Mocha} = \log\left(\exp\left(\frac{a_{SMocha} + b_{SMocha}}{\sigma_{Mocha}}\right) + \exp\left(\frac{a_{DMocha} + b_{DMocha}}{\sigma_{Mocha}}\right)\right)$$

$$s_{Mocha} = \frac{\exp\left(\sigma_{Mocha} \cdot I_{Mocha}\right)}{\exp(\sigma_{Mocha} \cdot I_{Mocha}) + \exp(\sigma_{Latte} \cdot I_{Latte}) + \exp(\sigma_{Cap} \cdot I_{Cap}) + \exp(\sigma_{Coffee} \cdot I_{Coffee})}$$

These equations might look big and scary, but with a few rules about e and logarithms, they are actually intuitively on par with what we've looked at before. I_{Mocha} is equivalent to the natural log of the sum of the valuations of the mocha "nest". This is also the denominator of the share equation that we just set up for $s_{S|Mocha}$. Therefore, the numerator of the equation for s_{Mocha} is equivalent to the exponential of the nesting parameter σ_{Mocha} multiplied by I_{Mocha}. One useful trick with logarithms is provided below:

$$\begin{aligned} e^{c \cdot log(e^{a/c} + e^{b/c})} &= (e^{a/c} + e^{b/c})^c \\ &= e^a + e^b \end{aligned}$$

Interestingly, σ_{Mocha} cancels out, so we could rewrite the equation for s_{Mocha}:

$$s_{Mocha} = \frac{\exp(a_{SMocha} + b_{SMocha} \cdot p_{SMocha}) + \exp(a_{DMocha} + b_{DMocha} \cdot p_{DMocha})}{\exp(\sigma_{Mocha} \cdot I_{Mocha}) + \exp(\sigma_{Latte} \cdot I_{Latte}) + \exp(\sigma_{Cap} \cdot I_{Cap}) + \exp(\sigma_{Coffee} \cdot I_{Coffee})}$$

This should intuitively make sense. Once we've grouped products into their respective nests, the valuations of each nest do not require any σ nesting parameters.

4.3 Solving for Profit Maximizing Prices Under a Nested Logit Model

We have arrived at the frontier of research in terms of solving for the profit maximizing price under a nested logit model. This model is extremely attractive as it approaches a real-world model design. The thing is, this is again not a straightforward problem to solve. We have seven equations and seven unknowns, and this time, we can't use the lambertW function to ease the computational difficulty. Let's set up the profit equations for Starbucks and Dunkin to see why:

$$\begin{aligned} \pi_S = (p_{SMocha} - c_{SMocha}) \cdot N \cdot s_{SMocha} + (p_{SCap} - c_{SCap}) \cdot N \cdot s_{SCap} + \\ (p_{SLatte} - c_{SLatte}) \cdot N \cdot s_{SLatte} + (p_{SCoffee} - c_{SCoffee}) \cdot N \cdot s_{SCoffee} \end{aligned}$$

$$\begin{aligned} \pi_D = (p_{DMocha} - c_{DMocha}) \cdot N \cdot s_{DMocha} + (p_{DLatte} - c_{DLatte}) \cdot N \cdot s_{DLatte} + \\ (p_{DCoffee} - c_{DCoffee}) \cdot N \cdot s_{DCoffee} \end{aligned}$$

Starbucks and Dunkin want to maximize these profit equations with respect to the prices that they are able to control. Starbucks has four prices that it can change, and Dunkin has three prices that it can change. As such, we can mathematically write the profit maximization problems of these firms as:

$$\boldsymbol{p_S^*} = (p^*_{SMocha}, p^*_{SCap}, p^*_{SLatte}, p^*_{SCoffee})$$

$$\begin{aligned} \boldsymbol{p_S^*} = \underset{\boldsymbol{p_S^*}}{\text{argmax}}[(p_{SMocha} - c_{SMocha}) \cdot N \cdot s_{SMocha} + (p_{SCap} - c_{SCap}) \cdot N \cdot s_{SCap} + \\ (p_{SLatte} - c_{SLatte}) \cdot N \cdot s_{SLatte} + (p_{SCoffee} - c_{SCoffee}) \cdot N \cdot s_{SCoffee}] \end{aligned}$$

$$\boldsymbol{p_D^*} = (p^*_{DMocha}, p^*_{DLatte}, p^*_{DCoffee})$$

$$\begin{aligned} \boldsymbol{p_D^*} = \underset{\boldsymbol{p_D^*}}{\text{argmax}}[(p_{DMocha} - c_{DMocha}) \cdot N \cdot s_{DMocha} + (p_{DCap} - c_{DCap}) \cdot N \cdot s_{DCap} + \\ (p_{DLatte} - c_{DLatte}) \cdot N \cdot s_{DLatte} + (p_{DCoffee} - c_{DCoffee}) \cdot N \cdot s_{DCoffee}] \end{aligned}$$

Notably, Starbucks and Dunkin no longer care about maximizing any individual price. For example, if Starbucks mochas are priced very high at say \$6.50 per cup and Starbucks lattes are priced very low at say \$3.50 per cup, then more customers will be driven to purchase lattes. It is possible that so many individuals would switch to lattes that this could actually increase Starbucks's profit. In addition to Starbucks now facing this multidimensional profit maximization problem, it is trying to maximize its revenue given Dunkin's multidimensional profit maximization choice as well.

This is an extremely difficult problem to solve. The model must compare every possible pricing option for the two firms. From Starbucks pricing mochas at \$10.00 and all other drinks at \$1.00 while Dunkin charges a flat rate of \$2.50 for all drinks, to Starbucks and Dunkin both choosing to price all drinks at \$4.50, every permutation of prices must be considered. These profit equations are considered "high dimensional" as they can have many different variables. This simple example is already seven-dimensional, and we haven't even considered additional types of coffee drinks or additional coffee shops in the market.

Research has shown that these high dimensional profit functions are not "quasi-concave" in $\boldsymbol{p^*}$. That is to say that under varying values of $(p^*_{SMocha}, p^*_{SCap}, p^*_{SLatte}, p^*_{SCoffee})$, there are multiple maxima and minima of Starbucks's profit function. Even if we were able to solve Starbucks's profit equation in terms of these four prices, we would end with a set of multiple possible profit-maximizing choices for the firms. Instead of approaching this by analyzing price combinations, other research has approached this problem by solving for the profit maximizing prices in terms of $\boldsymbol{s^*}$, the vector of shares of each product. Still more research has shown that these equations can be solved in terms of the "markup" of each nest. The markup is defined as the price minus the cost of each product. This recent research shows optimistic possibilities for solving these equations, but there is no simple way to solve them at this point.

4.4 Generalized forms of Nested Logit Equations

The consumer valuation for product i in nest h in a nested logit example takes on the following form:

$$V_i = \frac{a_i + b_i \cdot p_i}{\sigma_h}$$

This is very similar to the previous valuation expression, but we have now included the σ parameter. Here, each valuation for a product is divided by a σ parameter that identifies their nest. σ is always a value between 0 and 1, and lower values correspond to higher overall valuations of the nest.

With this in mind, the share of product i in nest h is given by:

$$s_i = s_{i|h} \cdot s_h$$

$$s_{i|h} = \frac{\exp(\frac{V_i}{\sigma_h})}{\sum_{k=1}^{H} \exp(\frac{V_k}{\sigma_h})}$$

$$s_h = \frac{\exp(\sigma_h \cdot I_h)}{\sum_{l=1}^{H} \exp(\sigma_l \cdot I_l)}$$

5 The Other Way: Estimating a Logit Model with Simulated Data

5.1 Simulating Data

One final manipulation is that we can estimate a logit model using the optimal prices we found to recover the consumer parameters with which we began.

A word of caution: The process we have outlined here takes advantage of a mathematical trick to solve a system of nonlinear equations. Solving a system of nonlinear equations is always a challenge, and estimating the values of the consumer parameters given the prices, shares, and margins is an imperfect science. That being said, we will explore the "antitrust" package in R as one possible avenue to estimate this logit model.

Given the prices and shares, we could theoretically assign a proportion of all consumers each product each product. Given a large enough sample size, we could take this one step further so that each consumer values a cup of coffee at some amount that is taken from a uniform distribution between 0 and 1. To make this even more realistic, we could add a normal random variable to each customer's valuation so that it is somewhat random how much the consumer wants the coffee.

Given this setup, we could simulate each customer's valuation of coffee and which coffee they chose to buy. Given more information, we could take this one step further and create a set of fixed effect indicators for each customer to estimate precise price sensitivities for different groups of customers.

While it is unlikely that simulating data in this way will ever be truly useful, it is possible, and once you've simulated this data, you could then use a logit demand solver to calculate the consumer parameters with which we began. One example of this type of solver is contained in the Antitrust R package. Here, the "logit" function can calculate the average price coefficient and mean utility for any number of competitors with a high degree of accuracy.

6 Conclusion

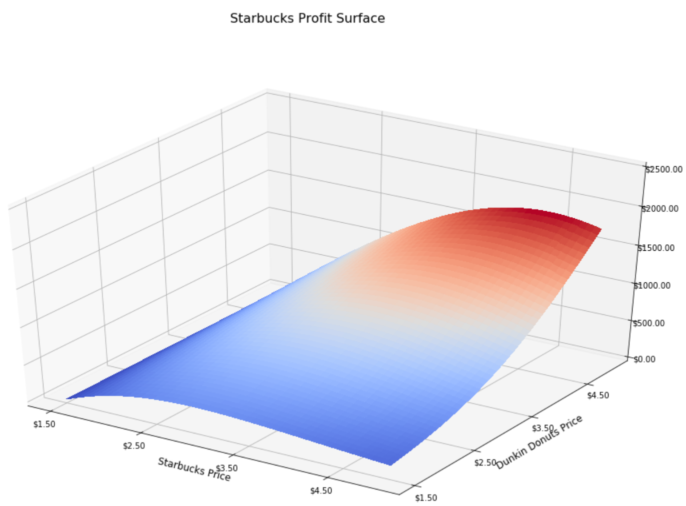
This article outlines how to find the profit maximizing price under logit demand. We can do this easily using a mathematical trick in the system of nonlinear equations, and this greatly simplifies the calculations.

While discrete choice models are primarily used for market analysis, it could be easily be applied to many other fields. For example, school choice, political choice, and employment choice could all be modelled as discrete choice problems. Applications such as these could have interesting results in the future.

7 Appendix 1: Profit Surface in Duopoly Setting

From the duopoly example of Starbucks and Dunkin Donuts in Section 2, we can go one step further to understand this problem geometrically. Below is the same plot from Figure 2 with Starbucks's profit curve at different settings of Dunkin's price:

This indicates that when Dunkin prices low, Starbucks should also price low, and when Dunkin prices high, Starbucks should also price high. Rather than looking at discrete values of p_D, we can also model Starbucks's profit function in two variables p_S and p_D. In the figure below, Starbucks's and Dunkin's prices are plotted along the x and y axes, and Starbucks's profit is plotted on the z-axis. Here, Starbucks's profit increases as Dunkin and Starbucks increase their prices. When either of the two competitors' prices are low, the both competitors' profits are low as indicated by the blue regions of the surface. On the other hand, if both firms price high, then both will see higher profits as indicated by the red regions of the surface.



There are two important points on this surface. First is the point at the maximum of the red region where Starbucks receives the most profit. At this point, Starbucks and Dunkin and both pricing as high as possible, and this in fact represents the monopolist's price. On the other hand, the lowest point on this surface at the bottom left indicates the perfect competition price at which Starbucks and Dunkin both price at cost. With any discrete number of firms $N > 1$, Starbucks and Dunkin will both price somewhere on this surface between these two points.

8 Appendix 2: Guess and Check Convergence Methodology

Let's make this methodology slightly more tractable with a simple example. Say you had two linear equations of two variables:

$$x = \frac{y-9}{-2}$$
$$y = 5 - x$$

These is a simple enough system of equations that you could solve it algebraically, but say that you wanted to solve it by the guess and check method as we do in our logit problem. If you guessed values of $(x, y) = (3, 5)$, then you would have the following result:

$$3 = \frac{5-9}{-2} = 2$$
$$5 = 5 - 3 = 2$$

Checking the distance between these left hand side and the right hand side results, we have:

$$distance = \sqrt{(y_2 - y_1)^2 + (x_2 - x_1)^2} = \sqrt{(2-5)^2 + (2-3)^2} = 3.162$$

Now, if we tried new values of $(x, y) = (3.99, 1.01)$ we could solve the equations again:

$$3.99 = \frac{1.01-9}{-2} = 3.995$$
$$1.01 = 5 - 3.99 = 1.01$$

Checking the distance between these left hand side and the right hand side results, we have:

$$distance = \sqrt{(3.995 - 3.99)^2 + (1.01 - 1.01)^2} = .005$$

The distance between the second set of values we tried is small, and we could in fact try sets of values that would allow us to get progressively closer to where the distance is 0. In this simple example, we can algebraically solve the system of equations to find that $(x, y) = (4, 1)$, but this illustrates how we're going to approach the logit problem when we can't directly solve the system of equations.

One additional point is that we won't exactly try values at random. In our model, the first values we try will be random. After that though, we will always use the combination (x_{n-1}, y_{n-1}) as our next combination. The table below shows the results of the iterative process for the linear example. Here, we iterate through the formula 15 times, each time using the resulting value from the previous run:

iteration	(x_{right}, y_{right})	(x_{left}, y_{left})	distance
1	$(3, 5)$	$(2, 2)$	3.162
2	$(2, 2)$	$(3.5, 3)$	1.803
3	$(3.5, 3)$	$(3, 1.5)$	1.581
4	$(3, 1.5)$	$(3.75, 2)$	0.901
5	$(3.75, 2)$	$(3.5, 1.25)$	0.791
6	$(3.5, 1.25)$	$(3.875, 1.5)$	0.451
7	$(3.875, 1.5)$	$(3.75, 1.125)$	0.395
8	$(3.75, 1.125)$	$(3.938, 1.25)$	0.225
9	$(3.938, 1.25)$	$(3.875, 1.063)$	0.198
10	$(3.875, 1.063)$	$(3.969, 1.125)$	0.113
11	$(3.969, 1.125)$	$(3.938, 1.033)$	0.099
12	$(3.938, 1.033)$	$(3.984, 1.063)$	0.056
13	$(3.984, 1.063)$	$(3.969, 1.016)$	0.049
14	$(3.969, 1.016)$	$(3.992, 1.031)$	0.028
15	$(3.992, 1.031)$	$(3.984, 1.007)$	0.025

With each iteration, the distance becomes smaller and (x_{right}, y_{right}) and (x_{left}, y_{left}) become closer. After we iterate enough times, the distance between our trial values and the returned values approaches 0, and we can stop whenever we get past a very small threshold as we can safely assume that we have approximated the solution.

This is precisely the methodology that we will used to find the profit maximizing prices in our logit question. We will solve the profit maximizing equations as much as we can algebraically, and then we will use a computer to iterate through possible price combinations until the price converges.

9 Appendix 3: Additional Derivations

9.1 Profit Maximizing Price with Two Firms

This derivation was liberally adapted from Ashwin Aravindakshan and Brian Rathford's 2011 paper in the *Review of Marketing Science* entitled "Solving Share Equations in Logit Models Using the LambertW Function." All credit is assigned to them for this incredibly useful result. From Section 2, let's return to the equation for Starbucks's profit maximizing price in terms of Dunkin's share:

$$p_S^* = MC_S - \frac{1}{b_S \cdot s_D}$$

From here, we can make the following substitutions to implement the lambertW function in this problem:

Proof.

$$\begin{aligned}
p_S^* &= MC_S - \frac{1}{b_S \cdot s_D} \\
&= MC_S - \frac{1}{b_S \cdot (\frac{\exp(a_D + b_S \cdot p_D)}{\exp(a_S + b_S \cdot p_S) + \exp(a_D + b_D \cdot p_D)})} && \text{substitute the value of } s_D \\
&= MC_S - \frac{\exp(a_S + b_S \cdot p_S) + \exp(a_D + b_D \cdot p_D)}{b_S \cdot \exp(a_D + b_D \cdot p_D)} \\
&= MC_S - \left(\frac{1}{b_S} + \frac{\exp(a_S + b_S \cdot p_S)}{b_S \cdot \exp(a_D + b_D \cdot p_D)}\right)
\end{aligned}$$

Multiplying both sides by b_S and also subtracting a_S from both sides, we have:

$$\begin{aligned}
b_S \cdot p_S^* - a_S &= b_S \cdot \left(MC_S - \left(\frac{1}{b_S} + \frac{\exp(a_S + b_S \cdot p_S)}{b_S \cdot \exp(a_D + b_D \cdot p_D)}\right)\right) - a_S \\
&= b_S \cdot MC_S - b_S \cdot \frac{1}{b_S} - b_S \cdot \frac{\exp(a_S + b_S \cdot p_S)}{b_S \cdot \exp(a_D + b_D \cdot p_D)} - a_S \\
&= b_S \cdot MC_S - 1 - \frac{\exp(a_S + b_S \cdot p_S)}{\exp(a_D + b_D \cdot p_D)} - a_S
\end{aligned}$$

Rewriting the equation above, we have:

$$\frac{\exp(a_S + b_S \cdot p_S)}{\exp(a_D + b_D \cdot p_D)} + b_S \cdot p_S^* + a_S = -1 + b_S \cdot MC_S + a_S$$

Taking exponentials on both sides, we then have:

$$e^{\frac{\exp(a_S + b_S \cdot p_S)}{\exp(a_D + b_D \cdot p_D)}} \cdot e^{b_S \cdot p_S^* + a_S} = e^{-1 + b_S \cdot MC_S + a_S}$$

Multiplying both sides by $\frac{1}{\exp(a_D + b_D \cdot p_D)}$, we then have:

$$e^{\frac{\exp(a_S + b_S \cdot p_S)}{\exp(a_D + b_D \cdot p_D)}} \cdot \frac{e^{b_S \cdot p_S^* + a_S}}{\exp(a_D + b_D \cdot p_D)} = \frac{e^{-1 + b_S \cdot MC_S + a_S}}{\exp(a_D + b_D \cdot p_D)}$$

Now, make the following substitution for W:

$$W = \frac{e^{b_S \cdot p_S^* + a_S}}{\exp(a_D + b_D \cdot p_D)}$$

This implies that We^W is:

$$We^W = \frac{e^{b_S \cdot p_S^* + a_S}}{\exp(a_D + b_D \cdot p_D)} \cdot e^{\frac{e^{b_S \cdot p_S^* + a_S}}{\exp(a_D + b_D \cdot p_D)}}$$

This is exactly equal the equation above.[15] Therefore, we can make the following substitution:

$$We^W = \frac{e^{-1+b_S \cdot MC_S + a_S}}{\exp(a_D + b_D \cdot p_D)}$$

We will use two properties of the lambertW function to complete the proof. First, any lambertW equation of the form $We^W = x$ has a solution at $W = W(x)$.[16] Therefore, we know that a solution to the equation above is given by:

$$W = W\left(\frac{e^{-1+b_S \cdot MC_S + a_S}}{\exp(a_D + b_D \cdot p_D)}\right)$$

Substituting for W, we then have:

$$\frac{e^{b_S \cdot p_S^* + a_S}}{\exp(a_D + b_D \cdot p_D)} = W\left(\frac{e^{-1+b_S \cdot MC_S + a_S}}{\exp(a_D + b_D \cdot p_D)}\right)$$

The second property of the lambertW function that we will use is the logarithmic property of the lambertW function. This tells us that $\ln(W(x)) = \ln(x) - W(x)$. Therefore, taking the natural logs of both sides of this equation, we have:

$$b_S \cdot p_S^* + a_S - \ln(a_D + b_D \cdot p_D) = -1 + b_S \cdot MC_S + a_S - \ln(a_D + b_D \cdot p_D) - W\left(\frac{e^{-1+b_S \cdot MC_S + a_S}}{\exp(a_D + b_D \cdot p_D)}\right)$$

Finally, solving this equation for p_S^*, the final result is:

$$p_S^* = MC_S - \frac{1 + W\left(\frac{e^{a_S - 1 + b_S \cdot MC_S}}{\exp(a_D + b_D \cdot p_D)}\right)}{b_S}$$

□

9.2 Market Share with Two Firms

This derivation was also adapted from Ashwin Aravindakshan and Brian Rathford's 2011 paper in the *Review of Marketing Science* entitled "Solving Share Equations in Logit Models Using the LambertW Function." All credit is assigned to them for this incredibly useful result. From Section 2, let's return to the equation for Starbucks's market share:

$$s_S^* = \frac{\exp(a_S + b_S \cdot p_S)}{\exp(a_S + b_S \cdot p_S) + \exp(a_D + b_D \cdot p_D)}$$

Splitting the numerator we have:

$$s_S^* = \frac{\exp(a_S) \cdot \exp(b_S \cdot p_S)}{\exp(a_S + b_S \cdot p_S) + \exp(a_D + b_D \cdot p_D)}$$

[15] Look very carefully at the the equation above. Notice that the numerator is just $exp(a_S + b_S \cdot p_S)$. This don't look the same because of the way you wrote the es!

[16] Would like a footnote to explain this.

Furthermore, we can bring the second half of the numerator to the denominator using exponent rules:

$$
\begin{aligned}
s_S^* &= \frac{\exp(a_S)}{(\exp(a_S + b_S \cdot p_S) + \exp(a_D + b_D \cdot p_D)) \cdot \exp(-b_S \cdot p_S)} \\
&= \frac{\exp(a_S)}{(\exp(a_S) + \exp(a_D + b_D \cdot p_D) \cdot \exp(-b_S \cdot p_S))}
\end{aligned}
$$

Now, we can actually use the equation for the profit maximizing price to solve for the closed form of the market share at the profit maximizing price:

Proof.

$$
\begin{aligned}
s_S^* &= \frac{\exp(a_S)}{\exp(a_S) + \exp(a_D + b_D \cdot p_D) \cdot \exp(-b_S \cdot p_S^*)} \\
&= \frac{\exp(a_S)}{\exp(a_S) + \exp(a_D + b_D \cdot p_D) \cdot \exp\left(-b_S \cdot \left(MC_S - \frac{1 + W\left(\frac{\exp(a_S - 1 + b_S \cdot MC_S)}{\exp(a_D + b_D \cdot p_D)}\right)}{b_S}\right)\right)} \\
&= \frac{\exp(a_S)}{\exp(a_S) + \exp(a_D + b_D \cdot p_D) \cdot \exp\left(-b_S \cdot MC_S + 1 + W\left(\frac{\exp(a_S - 1 + b_S \cdot MC_S)}{\exp(a_D + b_D \cdot p_D)}\right)\right)} \\
&= \frac{\exp(a_S)}{\exp(a_S) + \exp(a_D + b_D \cdot p_D) \cdot \exp(1 - b_S \cdot MC_S) \cdot \exp\left(W\left(\frac{\exp(a_S - 1 + b_S \cdot MC_S)}{\exp(a_D + b_D \cdot p_D)}\right)\right)}
\end{aligned}
$$

Using the lambertW function formulation $We^W = x$, we can rewrite the W term in this equation as:

$$
s_S^* = \frac{\exp(a_S)}{\exp(a_S) + \exp(a_D + b_D \cdot p_D) \cdot \exp(1 - b_S \cdot MC_S) \cdot \left(\frac{\exp(a_S - 1 + b_S \cdot MC_S)}{\exp(a_D + b_D \cdot p_D) \cdot W\left(\frac{\exp(a_S - 1 + b_S \cdot MC_S)}{\exp(a_D + b_D \cdot p_D)}\right)}\right)}
$$

Let's split up some terms and do some algebra to take this to a more workable form:

$$
\begin{aligned}
s_S^* &= \frac{\exp(a_S)}{\exp(a_S) + \exp(a_D + b_D \cdot p_D) \cdot \exp(1) \cdot \exp(-b_S \cdot MC_S) \cdot \left(\frac{\exp(a_S) \cdot \exp(-1) \cdot \exp(b_S \cdot MC_S)}{\exp(a_D + b_D \cdot p_D) \cdot W\left(\frac{\exp(a_S - 1 + b_S \cdot MC_S)}{\exp(a_D + b_D \cdot p_D)}\right)}\right)} \\
&= \frac{\exp(a_S)}{\exp(a_S) + \exp(a_D + b_D \cdot p_D) \cdot \left(\frac{\exp(a_S) \cdot \exp(1-1) \cdot \exp((-b_S \cdot MC_S) + (b_S \cdot MC_S))}{\exp(a_D + b_D \cdot p_D) \cdot W\left(\frac{\exp(a_S - 1 + b_S \cdot MC_S)}{\exp(a_D + b_D \cdot p_D)}\right)}\right)} \\
&= \frac{\exp(a_S)}{\exp(a_S) + \cancel{\exp(a_D + b_D \cdot p_D)} \cdot \left(\frac{\exp(a_S) \cdot \cancel{\exp(1-1)} \cdot \exp((-b_S \cdot MC_S) + (b_S \cdot MC_S))}{\cancel{\exp(a_D + b_D \cdot p_D)} \cdot W\left(\frac{\exp(a_S - 1 + b_S \cdot MC_S)}{\exp(a_D + b_D \cdot p_D)}\right)}\right)} \\
&= \frac{\exp(a_S)}{\exp(a_S) + \left(\frac{\exp(a_S) \cdot \exp(\cancel{(-b_S \cdot MC_S) + (b_S \cdot MC_S)})}{W\left(\frac{\exp(a_S - 1 + b_S \cdot MC_S)}{\exp(a_D + b_D \cdot p_D)}\right)}\right)} \\
&= \frac{\exp(a_S)}{\exp(a_S) + \left(\frac{\exp(a_S)}{W\left(\frac{\exp(a_S - 1 + b_S \cdot MC_S)}{\exp(a_D + b_D \cdot p_D)}\right)}\right)}
\end{aligned}
$$

Dividing all the numerator by $\exp(a_S)$, we have:

$$
s_S^* = \frac{1}{1 + \left(\frac{1}{W\left(\frac{\exp(a_S - 1 + b_S \cdot MC_S)}{\exp(a_D + b_D \cdot p_D)}\right)}\right)}
$$

Rewriting the 1 in the denominator as a fraction we have:

$$s_S^* = \frac{1}{\frac{W\left(\frac{\exp(a_S-1+b_S\cdot MC_S)}{\exp(a_D+b_D\cdot p_D)}\right)}{W\left(\frac{\exp(a_S-1+b_S\cdot MC_S)}{\exp(a_D+b_D\cdot p_D)}\right)} + \left(\frac{1}{W\left(\frac{\exp(a_S-1+b_S\cdot MC_S)}{\exp(a_D+b_D\cdot p_D)}\right)}\right)}$$
$$= \frac{1}{\frac{W\left(\frac{\exp(a_S-1+b_S\cdot MC_S)}{\exp(a_D+b_D\cdot p_D)}\right)+1}{W\left(\frac{\exp(a_S-1+b_S\cdot MC_S)}{\exp(a_D+b_D\cdot p_D)}\right)}}$$
$$= \frac{W\left(\frac{\exp(a_S-1+b_S\cdot MC_S)}{\exp(a_D+b_D\cdot p_D)}\right)}{1+W\left(\frac{\exp(a_S-1+b_S\cdot MC_S)}{\exp(a_D+b_D\cdot p_D)}\right)}$$

□

9.3 Profit Maximizing Price with N Firms

From Section 3, let's return to the equation for the profit maximizing price p_i^* with N firms:

$$p_i^* = MC_i - \left(b_i \cdot \frac{\exp(a_i + b_D p_D)}{\sum_{i=1}^N (\exp(a_i + b_i p_i))^2}\right)^{-1} \cdot s_i$$

So we have:

$$p_i^* = MC_i - \frac{1}{b_S \cdot \sum_{j=1}^N s_j}$$

From here, we can make the following substitutions to implement the lambertW function in this problem:

Proof.

$$p_i^* = MC_i - \frac{1}{b_S \cdot \sum_{j=1}^N s_j}$$
$$= MC_i - \frac{1}{b_S \cdot \left(\sum_{j=1}^N \frac{\exp(a_j+b_j\cdot p_j)}{\sum_{i=1}^N \exp(a_i+b_i\cdot p_i)}\right)} \qquad \text{substitute the value of } s_j$$
$$= MC_S - \frac{\sum_{i=1}^N \exp(a_i + b_i \cdot p_i)}{b_i \cdot \sum_{j=1}^N \exp(a_j + b_j \cdot p_j)}$$
$$= MC_S - \left(\frac{1}{b_i} \cdot \frac{\sum_{i=1}^N \exp(a_i + b_i \cdot p_i)}{\sum_{j=1}^N \exp(a_j + b_j \cdot p_j)}\right)$$

Multiplying both sides by b_i and also subtracting a_i from both sides, we have:

$$b_i \cdot p_i^* - a_i = b_i \cdot \left(MC_i - \left(\frac{1}{b_i} \cdot \frac{\sum_{i=1}^N \exp(a_i + b_i \cdot p_i)}{\sum_{j=1}^N \exp(a_j + b_j \cdot p_j)}\right)\right) - a_i$$
$$= b_i \cdot MC_i - b_i \cdot \frac{1}{b_i} - b_i \cdot \frac{\exp(a_i + b_i \cdot p_i)}{\sum_{j=1}^N \exp(a_j + b_j \cdot p_j)} - a_i$$
$$= b_i \cdot MC_i - 1 - \frac{\exp(a_i + b_i \cdot p_i)}{\sum_{j=1}^N \exp(a_j + b_j \cdot p_j)} - a_i$$

Rewriting the equation above, we have:

$$\frac{\exp(a_i + b_i \cdot p_i)}{\sum_{j=1}^N \exp(a_j + b_j \cdot p_j)} + b_i \cdot p_i^* + a_i = -1 + b_i \cdot MC_i + a_i$$

Taking exponentials on both sides, we then have:

$$e^{\frac{\exp(a_i+b_i\cdot p_i)}{\sum_{j=1}^{N}\exp(a_j+b_j\cdot p_j)}}\cdot e^{b_i\cdot p_i^*+a_i}=e^{-1+b_i\cdot MC_i+a_i}$$

Multiplying both sides by $\frac{1}{\sum_{j=1}^{N}\exp(a_j+b_j\cdot p_j)}$, we then have:

$$e^{\frac{\exp(a_i+b_i\cdot p_i)}{\sum_{j=1}^{N}\exp(a_j+b_j\cdot p_j)}}\cdot\frac{e^{b_i\cdot p_i^*+a_i}}{\sum_{j=1}^{N}\exp(a_j+b_j\cdot p_j)}=\frac{e^{-1+b_i\cdot MC_i+a_i}}{\sum_{j=1}^{N}\exp(a_j+b_j\cdot p_j)}$$

Now, make the following substitution for W:

$$W=\frac{e^{b_i\cdot p_i^*+a_i}}{\sum_{j=1}^{N}\exp(a_j+b_j\cdot p_j)}$$

This implies that We^W is:

$$We^W=\frac{e^{b_i\cdot p_i^*+a_i}}{\sum_{j=1}^{N}\exp(a_j+b_j\cdot p_j)}\cdot e^{\frac{e^{b_i\cdot p_i^*+a_i}}{\sum_{j=1}^{N}\exp(a_j+b_j\cdot p_j)}}$$

This is exactly equal the equation above.[17] Therefore, we can make the following substitution:

$$We^W=\frac{e^{-1+b_i\cdot MC_i+a_i}}{\sum_{j=1}^{N}\exp(a_j+b_j\cdot p_j)}$$

We will use two properties of the lambertW function to complete the proof. First, any lambertW equation of the form $We^W=x$ has a solution at $W=W(x)$.[18] Therefore, we know that a solution to the equation above is given by:

$$W=W\left(\frac{e^{-1+b_i\cdot MC_i+a_i}}{\sum_{j=1}^{N}\exp(a_j+b_j\cdot p_j)}\right)$$

Substituting for W, we then have:

$$\frac{e^{b_i\cdot p_i^*+a_i}}{\sum_{j=1}^{N}\exp(a_j+b_j\cdot p_j)}=W\left(\frac{e^{-1+b_i\cdot MC_i+a_i}}{\sum_{j=1}^{N}\exp(a_j+b_j\cdot p_j)}\right)$$

The second property of the lambertW function that we will use is the logarithmic property of the lambertW function. This tells us that $\ln(W(x))=\ln(x)-W(x)$. Therefore, taking the natural logs of both sides of this equation, we have:

$$b_i\cdot p_i^*+a_i-\ln(\sum_{j=1}^{N}(a_j+b_j\cdot p_j)=-1+b_i\cdot MC_i+a_i-\ln(\sum_{j=1}^{N}(a_j+b_j\cdot p_j)-W\left(\frac{e^{-1+b_i\cdot MC_i+a_i}}{\sum_{j=1}^{N}\exp(a_j+b_j\cdot p_j)}\right)$$

Finally, solving this equation for p_S^*, the final result is:

$$p_i^*=MC_i-\frac{1+W\left(\frac{e^{a_i-1+b_S\cdot MC_i}}{\sum_{j=1}^{N}\exp(a_j+b_j\cdot p_j)}\right)}{b_i}$$

□

[17] Look very carefully at the the equation above. Notice that the numerator is just $exp(a_S+b_S\cdot p_S)$. This don't look the same because of the way you wrote the es!

[18] Would like a footnote to explain this.

9.4 Market Share with N Firms

This derivation was also adapted from Ashwin Aravindakshan and Brian Rathford's 2011 paper in the *Review of Marketing Science* entitled "Solving Share Equations in Logit Models Using the LambertW Function." All credit is assigned to them for this incredibly useful result. From Section 3, let's return to the equation for a firm i's market share:

$$s_i^* = \frac{\exp(a_i + b_i \cdot p_i)}{\sum_{i=1}^{N}(\exp(a_i + b_i \cdot p_i)}$$

Splitting the numerator we have:

$$s_i^* = \frac{\exp(a_i) \cdot \exp(b_i \cdot p_i)}{\sum_{i=1}^{N}(\exp(a_i + b_i \cdot p_i)}$$

Furthermore, we can bring the second half of the numerator to the denominator using exponent rules:

$$\begin{aligned} s_i^* &= \frac{\exp(a_i)}{\sum_{i=1}^{N}(\exp(a_i + b_i \cdot p_i)) \cdot \exp(-b_i \cdot p_i)} \\ &= \frac{\exp(a_i)}{\exp(a_i) + \sum_{j=1}^{N}\exp(a_j + b_j \cdot p_j) \cdot \exp(-b_i \cdot p_i))} \end{aligned}$$

Now, we can actually use the equation for the profit maximizing price to solve for the closed form of the market share at the profit maximizing price:

Proof.

$$\begin{aligned} s_i^* &= \frac{\exp(a_i)}{\exp(a_i) + \sum_{j=1}^{N}\exp(a_j + b_j \cdot p_j) \cdot \exp(-b_i \cdot p_i^*))} \\ &= \frac{\exp(a_i)}{\exp(a_i) + \sum_{j=1}^{N}\exp(a_j + b_j \cdot p_j) \cdot \exp\left(-b_i \cdot \left(MC_i - \frac{1+W\left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N}\exp(a_j + b_j \cdot p_j)}\right)}{b_i}\right)\right)} \\ &= \frac{\exp(a_i)}{\exp(a_i) + \sum_{j=1}^{N}\exp(a_j + b_j \cdot p_j) \cdot \exp\left(-b_i \cdot MC_i + 1 + W\left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N}\exp(a_j + b_j \cdot p_j)}\right)\right)} \\ &= \frac{\exp(a_i)}{\exp(a_i) + \sum_{j=1}^{N}\exp(a_j + b_j \cdot p_j) \cdot \exp(1 - b_i \cdot MC_i) \cdot \exp\left(W\left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N}\exp(a_j + b_j \cdot p_j)}\right)\right)} \end{aligned}$$

Using the lambertW function formulation $We^W = x$, we can rewrite the W term in this equation as:

$$s_i^* = \frac{\exp(a_i)}{\exp(a_i) + \sum_{j=1}^{N}\exp(a_j + b_j \cdot p_j) \cdot \exp(1 - b_i \cdot MC_i) \cdot \left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N}\exp(a_j + b_j \cdot p_j) \cdot W\left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N}\exp(a_j + b_j \cdot p_j)}\right)}\right)}$$

Let's split up some terms and do some algebra to take this to a more workable form:

$$
\begin{aligned}
s_i^* &= \frac{\exp(a_i)}{\exp(a_i) + \sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j) \cdot \exp(1) \cdot \exp(-b_i \cdot MC_i) \cdot \left(\frac{\exp(a_i) \cdot \exp(-1) \cdot \exp(b_i \cdot MC_i)}{\sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j) \cdot W\left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j)} \right)} \right)} \\
&= \frac{\exp(a_i)}{\exp(a_i) + \sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j) \cdot \left(\frac{\exp(a_i) \cdot \exp(1-1) \cdot \exp((-b_i \cdot MC_i) + (b_i \cdot MC_i))}{\sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j) \cdot W\left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j)} \right)} \right)} \\
&= \frac{\exp(a_i)}{\exp(a_i) + \cancel{\sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j)} \cdot \left(\frac{\exp(a_i) \cdot \cancel{\exp(1-1)} \cdot \exp((-b_i \cdot MC_i) + (b_i \cdot MC_i))}{\cancel{\sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j)} \cdot W\left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j)} \right)} \right)} \\
&= \frac{\exp(a_i)}{\exp(a_i) + \left(\frac{\exp(a_i) \cdot \exp(\cancel{(-b_i \cdot MC_i) + (b_i \cdot MC_i)})}{W\left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j)} \right)} \right)} \\
&= \frac{\exp(a_i)}{\exp(a_i) + \left(\frac{\exp(a_i)}{W\left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j)} \right)} \right)}
\end{aligned}
$$

Dividing the numerator by $\exp(a_S)$, we have:

$$
s_i^* = \frac{1}{1 + \left(\frac{1}{W\left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j)} \right)} \right)}
$$

Rewriting the 1 in the denominator as a fraction we have:

$$
\begin{aligned}
s_i^* &= \frac{1}{\frac{W\left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j)} \right)}{W\left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j)} \right)} + \left(\frac{1}{W\left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j)} \right)} \right)} \\
&= \frac{1}{\frac{W\left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j)} \right) + 1}{W\left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j)} \right)}} \\
&= \frac{W\left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j)} \right)}{1 + W\left(\frac{\exp(a_i - 1 + b_i \cdot MC_i)}{\sum_{j=1}^{N} \exp(a_j + b_j \cdot p_j)} \right)}
\end{aligned}
$$

$\square$

REFERENCES

Aravindakshan, Ashwin, and Brian Ratchford. "Solving Share Equations in Logit Models Using the LambertW Function." *Review of Marketing Science*, vol. 9, 2011, ashwin.faculty.ucdavis.edu/wp-content/uploads/sites/169/2014/12/Publication-5.pdf. Accessed 28 Feb. 2021.

Berry, Steven, et al. "Automobile Prices in Market Equilibrium." Econometrica, vol. 63, no. 4, 1995, pp. 841–890. JSTOR, www.jstor.org/stable/2171802. Accessed 1 Mar. 2021.

Hanson, Ward, and Kipp Martin. "Optimizing Multinomial Logit Profit Functions." Management Science, vol. 42, no. 7, 1996, pp. 992–1003. JSTOR, www.jstor.org/stable/2634363. Accessed 28 Feb. 2021.

Sands, Emily Glassberg, and Duncan Gilchrist. "How to Get the Price Right: Tools for Optimal Price Setting." *Medium*, 30 Oct. 2016, medium.com/teconomics-blog/how-to-get-the-price-right-9fda84a33fe5. Accessed 28 Feb. 2021.

Shum, Matthew. "Lecture Notes: Discrete-Choice Models of Demand." Economics, California Institute of Technology. *Caltech.edu*, California Institute of Technology, www.its.caltech.edu/ mshum/ec106/discrete.pdf. Accessed 28 Feb. 2021. Working paper.