Estimation of a Cost-of-Living Index for a Storable Product Using A Dynamic Structural Model

Matthew Osborne*

August, 2012

Abstract

This paper develops a cost of living index (COLI) for a storable product category with periodic promotions. To construct the index, we first estimate a dynamic structural model of consumer stockpiling behavior on one year of household level scanner data on canned tuna purchases. In our model, consumers are forward-looking and expect periodic temporary product promotions: when a product goes on promotion, they store it for future consumption.

With estimates from the structural model in hand, we compute consumer utilities, which are the central input in constructing the COLI. Our COLI is a sequence of taxes that are dependent on the aggregate market state, keeping average utilities constant over time. Using a simple transformation of the COLI which recasts it as a price index, our index averages about 82% over the course of one year. We find that standard indexes are significantly higher than our index - fixed base indexes average around 90%, while chained indexes are substantially higher or lower than our index. Our average index is well approximated by a simpler COLI index proposed by Feenstra and Shapiro (2003), which predicts price falls of about 84%.

*Research Economist, Bureau of Economic Analysis, U.S. Department of Commerce, Washington, DC 20230. Email: matthew.osborne@bea.gov. The views expressed herein are my own and not necessarily those of the Bureau of Economic Analysis or the U.S. Department of Commerce.

Our work also advances structural estimation of dynamic discrete choice models. We demonstrate how to use a the recently developed Bayesian technique of Imai, Jain, and Ching (2009) to estimate a stockpiling model. Use of the Bayesian technique allows us to estimate the parameters of consumer price expectations along with utility parameters; typically these parameters are estimated in a separate step prior to estimating the dynamic model. Using estimated parameters as inputs into the estimation of the dynamic model will lead to biases in the standard error estimates for parameters that arise from the dynamic model, unless a standard two-step correction is applied. Our approach avoids this. Use of the technique also allows us to incorporate continuously distributed unobserved heterogeneity and endogenous consumption in a tractable way.

1 Introduction

The Consumer Price Index (CPI) tracks the price movements of many different household food products and is widely used in economic measurement and policy. As an example, when constructing real measures of food purchased for consumption at home, the Bureau of Economic Analysis deflates revenues in a product category by the corresponding CPI.[1] Using accurate price indexes is therefore a topic that is of great importance for accurate measurement of economic output. The increasing availability of household and store-level scanner data, as well as the development of techniques that allow for the estimation of sophisticated dynamic models of household behavior, can facilitate the improvement and accuracy of price indexes.

Our paper constructs a cost-of-living index (COLI) for a frequently-purchased product category which is storable, and where periodic promotions occur. A COLI measures how much one would need to compensate consumers for price changes over time, relative to some base period. Although standard price indexes are not constructed as COLIs, according to the Bureau of Labor Statistics (Bureau of Labor Statistics 2007) the measurement objective underlying the CPI is that of a COLI, and the CPI is often used as such in practice.[2] In order to construct a COLI, one needs to be able to compute consumer utility. Structural econometric methods are ideal for this, because they produce estimates of the parameters of a consumer's utility function. With the estimated parameters in hand, constructing a COLI is straightforward, and a simple transformation of the COLI allows us to compare it to a standard price index.

Dynamic consumer behavior can cause a "true" price index derived from a COLI to diverge from standard price indexes. In our paper, the dynamics we consider are driven by consumer stockpiling of a storable product. When the product is on promotion, consumers stockpile it for future consumption. This type of behavior is problematic

[1] See Chapter 5 of the NIPA handbook located at http://www.bea.gov/national/pdf/ch5%20PCEforposting.pdf, page 5-13. Measured in 2005 dollars, food purchased for off premises consumption totaled about 106 billion dollars for the first quarter of 2012 (source: NIPA tables at http://www.bea.gov/national/txt/dgpa.txt, Table 1.5.3.).

[2] Bureau of Labor Statistics (2007), pg 2 states "The concept of COLI provides the CPI's measurement objective."

when one tries to compute price indexes because it creates a type of substitution bias: when consumers observe a high price for a product, they wait to purchase the product until it goes on promotion again. In a standard fixed-base index, period t prices are weighted by period 0 quantities; since promotions happen infrequently, they will on average receive less weight than non-promotional prices. In reality, though, since consumers tend to wait for promotions, a more accurate price index would assign more weight to prices in promotional periods as opposed to non-promotional periods.[3]

To quantify the magnitude of the divergence between the COLI and other price indexes, we specify and estimate a dynamic structural model of consumer stockpiling behavior using one year of household level scanner data of canned tuna purchases. During the course of the year, canned tuna prices fall. The index derived from our structural model averages 82% over the course of our one-year sample, consistent with falling prices. In contrast, a fixed-base Laspeyres index averages about 90% over the course of the year.[4] The fact that the Laspeyres index is higher than our COLI confirms the intuition in the previous paragraph: relative to the base period, the Laspeyres overweights nonpromotional periods, as compared with our index. Chained indexes do an even worse job of approximating our price index: a chained Laspeyres averages 115% whereas a chained Tornqvist averages 61%. The chained Laspeyres index is upwards-biased due to it failing the "time reversal" test.[5] The chained Tornqvist index has weights that average slightly less then 1, but when this is cumulated the overall index drops sharply over time.

[3]When the BLS constructs its price indexes, it conducts random surveys of shelf prices of product in stores across the United States. BLS price indexes are therefore a measure of offered prices. Due to intertemporal substitution, offered prices and accepted prices may differ. The importance of this distinction has been discussed in earlier work by Erdem, Imai, and Keane (2003), who use a structural stockpiling model to show that offered and accepted prices differ significantly (Erdem, Imai, and Keane (2003) do not construct a COLI from their model).

[4]We construct our indexes at the weekly level.

[5]If prices in period t and period $t+2$ are $\overline{p}$, and prices are equal to $\tilde{p}$ in period $t+1$, a price index satisfies the time reversal test if a two-period chained index between periods t, $t+1$ and $t+2$, $P(\overline{p},\tilde{p})P(\tilde{p},\overline{p})$ is 1. It can be shown that this index is bigger than or equal to 1 for the chained Laspeyres, which results in an upward bias.

Our paper is not the first to consider the impact of intertemporal substitution on the construction of price indexes. Earlier work by Feenstra and Shapiro (2003) (hereby abbreviated FS) develops a COLI for canned tuna using store level scanner data. There are several key differences between their work and ours. First, the consumer model is very different. FS specify a representative consumer model where consumers make purchases over a finite planning horizon. It is assumed that consumers the path of future prices (as well as other exogenous variables) exactly, and that the expenditure function can be approximated by a translog function of prices. Under these assumptions, the "true" price index for the good takes a functional form that is similar to a fixed-base Tornqvist index. FS find that fixed-base indexes approximate their COLI index more closely than chained indexes, with the fixed-base Tornqvist index providing the best approximation.

We replicate the FS COLI using our data, finding an index value of 84%, which is quite close to our average COLI. This is interesting, because it suggests that practitioners could use the FS COLI as an approximation to the true price index when our approach is infeasible. The similarity between the estimates is likely due to weighting: The FS COLI uses current expenditure shares as part of the index, meaning that promotional periods (where most purchases occur) should receive relatively more weight than non-promotional periods. As with FS, we also find that the closest indexes to the true COLI are fixed-base indexes. In contrast to our work, FS find that both the chained Laspeyres and chained Tornqvist display significant upward biases, whereas we find that the chained Tornqvist drops over time. We suspect the difference in our findings is due to differences in the data. In FS, most quantity purchased during a promotional period occurs near the end of the promotion, and product advertising occurs near the end of the promotion as well. This imparts an upward bias to chained indexes. In contrast, in our data we find that most purchases occur at the beginning of promotional periods, and product advertising occurs then as well. This may explain why the chained Tornqvist drops so rapidly: it is likely overweighting price falls that occur at the beginning of promotions.

We note that we are not the first paper to use a dynamic structural model to impute a price index. In the context of durable goods, Gowrisankaran and Rysman

(2011) construct a cost-of-living index (COLI) using aggregate scanner data and the estimated parameters of a dynamic structural model of consumer behavior. There are some key differences between our work and that of Gowrisankaran and Rysman (2011). First, since we examine the canned tuna category our implications are likely to hold in other product categories where repeated purchases occur, and the product is storable. Gowrisankaran and Rysman (2011) examine durable goods where a product is likely only purchased once, or very infrequently. Although Gowrisankaran and Rysman (2011) find that their price index differs significantly from a BLS-style index, the dynamics that drive the difference are somewhat different than in our case. In their case, the difference arises due to a "new buyer" problem identified by Aizcorbe (2005): low value consumers enter their market near the end of their sample when prices are low, creating concavity in the price index. In contrast, in our case consumers are timing their purchases to coincide with promotions. As we described above, this behavior creates a wedge betweeen standard fixed-base price indexes, which weight by base period shares, and our COLI, which will give higher weight to promotional periods.

Before turning to a description of the data, we make some notes on our empirical methodology. Our dynamic inventory model is complicated to estimate, because we assume that consumption is endogenous, and there is continuously distributed unobserved consumer heterogeneity. To make the estimation of our model more feasible, we apply the Bayesian approach of Imai, Jain, and Ching (2009). This approach significantly reduces the computational complexity of estimation, and allows us to easily integrate over unobservables such as initial inventories and unobserved heterogeneity. One other advantage of the approach is that we can feasibly estimate parameters that govern consumer price expectations along with the other model parameters. As is standard in the literature, we assume consumer expectations are rational, so that consumer price expectations coincide with observed prices.[6] Typically, one estimates price expectations estimating the dynamic model and then uses the estimates as inputs into the model. Technically, the standard errors of the dynamic estimates should be

[6] This assumption seems reasonable for products like canned tuna that have been in existence for a long time, and where consumers are likely familiar with store pricing.

adjusted since they are functions of previously estimated parameters - as far as we are aware this is never done. The two step approach is taken because the processes that govern price expectations are typically functions of many parameters, and maximizing the model's likelihood over a high dimensional parameter space using standard methods such as the nested fixed-point method is computationally infeasible.[7] In the nested fixed point approach, one must solve for consumer value functions every time a parameter is changed. Because the Bayesian approach only involves solving for the value function once, rather than many many times, it can accommodate more parameters.

The rest of the paper proceeds as follows. Section 2 introduces and describes the data. Section 3 provides some reduced form evidence that stockpiling behavior is important in our data. Sections 4 and 5 describe our stockpiling model, and how we model consumer price expectations. Section 6 describes the estimation technique, and Section 8 describes the results.

2 The Data Set

The data set is household level Nielsen scanner data on canned tuna purchases from Sioux Falls, SD. We focus the analysis on the two most popular brands of canned tuna, Starkist and Chicken of the Sea, which comprise over 90% of all purchases. Although canned tuna is available in different package sizes, the most popular size by far is the standard 6 ounce can. Thus, for computational simplicity the analysis focuses on households who only purchase the 6 ounce can size. As we show in Table 1, the two brands have roughly equal market shares by volume. Additionally, prices are very similar, although Starkist is slightly more expensive than Chicken of the Sea. Starkist has a lower standard deviation of prices than Chicken of the Sea, indicating it goes on sale less often, and its sales are less deep than those of its competitor. In this paper, promotions are defined as dips in the observed shelf price. Coupons may also be a part of a firm's promotional strategy; however, in this market there is very little observed coupon use. Coupons are used in less than 10% of purchases. Due to the computational issues that arise due to including coupons, we do not include them in

[7] See Rust (1994) and Rust (1987) for an overview of the nested fixed point algorithm.

the analysis.[8]

Our data set spans 3 years, 1985 to 1988. The data tracks purchases of canned tuna over this period for a sample of about 1500 households at 19 different stores. When a household makes a purchase at a store, the store's identity is recorded; however, household store visits where no purchase is made are only recorded for the final 51 weeks of the data. We restrict most of the analysis to that period because in our model of household stockpiling we need to know the price of a product when a store is visited, but no purchase is made. We also remove some households from the data where the quality of the data may be suspect. First, households for whom I observe less than 25 store visits during the 51 weeks are removed. Households who were surveyed in the data were given swipe cards that they were required to bring with them on shopping trips, and it is likely that households for whom the frequency of store visits is low were forgetting to use their card, or frequently shopping at a smaller store that was not included in the Nielsen data. Additionally, to keep the model's state space tractable I limit the sample to households who purchase at most ten cans of tuna in a week. This is not a strong limitation: among household-week observations where a purchase is observed, purchases of 11 cans or more comprise about 0.2%. We also restrict the sample to households who purchase 10 or more cans of tuna over the entire year. The behavior of households who never purchase tuna or who only purchase it very infrequently is unlikely to be well explained by a stockpiling model. After these cuts, we are left with a final sample of about 600 households.[9]

In Figure 1 we plot quantity weighted average prices for our sample on a weekly basis. There is a significant amount of volatility in prices over time - the standard deviation in weekly average prices is roughly 10% of its mean. A downward trend in prices can be seen in the data - a regression of log price on a time trend produces a coefficient that implies prices drop by an average of 0.3% per week. Overall quantities, displayed in Figure 2, display a similar amount of volatility. Periodically total quantity sold spikes upwards - these spikes are at least partially driven by consumers purchasing

[8] See Osborne (2011) for discussion on including coupons in dynamic structural models.

[9] The numbers in Table 1 were computed for all households, rather than the reduced sample. Prices and shares for the reduced sample are similar to the overall sample.

the product on promotion. To make this obvious, we plot the Starkist brand's weekly prices and quantities for all households over the longer 1985-1988 sample period in Figure 3 for a typical store in the data. The price, shown in the top panel, stays relatively flat at around 60 cents for most of the time period, but occasionally it drops significantly for a short period of time. The bottom panel shows the quantity sold, measured in the number of 6 ounce cans. The quantity sold is on average about 100 cans per week, but when promotions occur it jumps significantly to over 300 cans. This behavior is consistent with consumer stockpiling behavior: stores keep the price high most of the time, but recognize that price sensitive consumers will run down their inventories. To draw these consumers into the market, a sale eventually occurs. As discussed in the previous section, accounting for this type of behavior when measuring price indexes can be important: when the BLS constructs price indexes, it randomly samples prices at stores across the United States. However, as can be seen from the figure, and as has been pointed out by Erdem, Imai, and Keane (2003), the average offer price will differ from the average accepted price since most purchases occur on promotion.

Some evidence that the two brands may be using temporary promotions as a method of competition is shown in Figure 4. This figure shows the price of Starkist in black for the same store as the previous figure, and the price of Chicken of the Sea as the red dotted line. Notice that it is rarely the case that both brands go on promotion at the same time. Often when Starkist has a sale, Chicken of the Sea has a promotion soon afterwards, and vice versa. This observation will help to motivate our specification for consumer price expectations, which will include competitor reactions.

3 Evidence of Stockpiling Behavior

The time-series patterns of price and quantity in Figure 3, while suggestive of inventory behavior, are not totally conclusive of its presence. Alternative explanations have been put forward to explain periodic sales (such as variation in store inventories). The large spikes in demand may just reflect the overall price sensitivity of consumers in the market; stockpiling is not necessary to explain them. To offer more evidence of stockpiling, I run a regression of the total quantity a household purchases during a given week on a measure of the household's inventory, the price, a dummy variable for whether a sale occurs, the interaction of that dummy variable with price, and feature and display variables.[10] In reality, a household's inventory is unobserved, so I estimate household inventories by assuming that household consumption rates are constant. One can construct an estimate of the household's inventory during a given week as the sum of quantity purchased prior to that time, minus total consumption, plus initial inventories. I estimate a household's consumption rate by dividing the total quantity consumed by the number of weeks the household is observed. Initial inventories are also unobserved, so a household fixed effect is included to control for them. Quantity is measured in ounces, while inventory is measured in ounces divided by 100. Price is measured in dollars per ounce. Table 2 shows the estimates of this regression. Inventory is negative and significant, which suggests that when a household's inventory increases, the quantity they purchase decreases. Further, when a product goes on sale, households purchase more of the product, and have a larger price coefficient, which is consistent with stockpiling. A sale is a dummy variable that is one when the price is observed to be 5% or more lower than the modal price of the product in the store. The regression also includes store and product dummy variables.

A reduced form regression like the one just presented is suitable for determining whether or not stockpiling occurs, but not for policy simulation. For example, one of the exercises I perform is to vary the frequency of discounts. Stockpiling is forward-looking behavior, so consumers will react not only to current prices, but to what they expect prices to look like in the future. Since I do not model consumer price expectations in the regression, the regression parameters will be functions of policy variables such as the parameters of the price process. To use the regression model for predictions, one would need to account for how the price process impacted the regression parameters.

[10] This type of regression analysis follows similar work in Hendel and Nevo (2006).

4 Model of Consumer Stockpiling

In each period t, J different brands of a product are available. Consumers can purchase up to N_u units every period. A period is assumed to be a one week interval. Consumer i's choice of what to purchase in period t is a J-vector of quantities, $\boldsymbol{x}_{it} = (x_{1it}, ..., x_{Jit})$, such that $0 \leq \sum_{j=1}^{J} x_{jit} \leq N_u$. Consumer i has a taste for each product j, γ_{ij}. In addition to deciding what to purchase, consumers have to decide what to consume every period. I assume that consumption is integral, that is, consumers cannot eat fractions of a unit in a period. This is a reasonable assumption for canned tuna, since the fish cannot easily be stored if only part of a can is used. Consumption is expressed as a vector $\boldsymbol{c}_{it} = (c_{1it}, ..., c_{Jit})$. Any units that are not consumed are be stored in inventory for future consumption. Inventory is an integer vector $\boldsymbol{\iota}_{it} = (\iota_{1it}, ..., \iota_{Tit})$. I assume that consumers have a maximum storage space of $2N_u$, which means that $\sum_{j=1}^{J} \iota_{jit} \leq 2N_u$. It will be convenient to denote the consumer's total inventory as $I_{it} = \sum_{j=1}^{J} \iota_{jit}$. I_{it} evolves as follows:

$$I_{it} = I_{it-1} + \sum_{j=1}^{J} x_{jit} - \sum_{j=1}^{J} c_{jit} \tag{1}$$

Each consumer observes a choice specific error, ϵ_{qit}, prior to making a purchase; q indexes each of the $N_u(N_u - 1)/2$ possible values of $\boldsymbol{x}_{it}$. The per unit price of each product[11] in period t is p_{jit}, and the vector of prices for all brands is denoted as $\boldsymbol{p}_{it}$. In each period, a consumer's flow utility from consuming $\boldsymbol{c}_{it}$ and purchasing $\boldsymbol{x}_{it}$ is

$$\begin{aligned} U(\boldsymbol{c}_{it}, \boldsymbol{x}_{it}, \boldsymbol{\iota}_{it}, \boldsymbol{p}_{it}) = & \sum_{j=1}^{J} \gamma_{ij} u(c_{jit}; \beta) - \alpha_i \sum_{j=1}^{J} x_{jit} p_{jit} - sc_0 I_{it} - sc_1 I_{it}^2 \\ & - CC\, \mathbf{1}\{\sum_{j=1}^{J} x_{jit} > 0\} + \epsilon_{qit}. \end{aligned} \tag{2}$$

In this function, the utility from consuming a given product is $\gamma_{ij} u(c_{jit}; \beta)$, where β is a parameter that impacts the shape of this subutility. We assume that flow utility for each product is quadratic, so that

[11] We have not been able to find any significant evidence of quantity discounts in our data set.

$$u(c;\beta) = c + \beta c^2. \tag{3}$$

The parameter α_i is consumer i's price sensitivity. The consumer's inventory holding costs are assumed to be quadratic, with sc_0 on the linear term and sc_1 on the quadratic. The term CC is a carrying cost, and it represents the disutility of purchasing and carrying the product.

Consumers are assumed to be forward-looking with rational expectations, and they discount the future at a discount rate $\delta > 0$. Thus, at time t, a consumer chooses her consumption and purchases in order to maximize her current flow utility plus her expected discounted future utility. There are three state variables which each consumer keeps track of every period. One is the current price vector, $\boldsymbol{p}_{it}$. Related to this is a state variable that tracks whether a promotion occurs in period t, $\boldsymbol{s}_{it}$; this is a vector of length J containing 1 in position j if product j is on promotion, and 0 otherwise. I assume that a product is on promotion if its price is observed to be below some level $\overline{p}_j$. Promotions evolve over time according to a discrete Markov process, $S(\boldsymbol{s}_{it}|\boldsymbol{p}_{it-1}, \boldsymbol{s}_{it-1})$. The probability of a promotion occurring today is a function of whether or not the product was on sale in the previous week, whether its competitors were on sale in the previous week, and what last week's prices were. Given that sales occur sporadically and are usually short, one would expect that the probability of a sale occurring given no sale last week would be low, and the probability of no sale occurring given a sale last week would be higher. Conditional on a sale occurring today, prices evolve over time according to a Markov process $P(\boldsymbol{p}_{it}|\boldsymbol{p}_{it-1}, \boldsymbol{s}_{it-1}, \boldsymbol{s}_{it})$. Although it is technically redundant to include $\boldsymbol{s}$ as an argument in P, we feel it eases the exposition to specify the price process conditional on the current promotion state. This is because if a sale occurs, the price distribution is truncated at $\overline{p}_j$. The last state variable is the consumer's inventory, $\boldsymbol{\iota}_{it}$. Inventories for individual brands evolve analogously to Equation (1):

$$\iota_{jit} = \iota_{jit-1} + x_{jit} - c_{jit} \tag{4}$$

Denote the set of state variables as $\boldsymbol{\Sigma}_{it} = (\boldsymbol{p}_{it}, \boldsymbol{s}_{it}, \boldsymbol{i}_{it})$, and denote the vector of utility parameters as $\boldsymbol{\theta} = (\gamma_{i1}, ..., \gamma_{iJ}, \beta, \alpha_i, sc_0, sc_1, CC)$. The consumer's expected discounted utility in purchase event t is

$$V(\boldsymbol{\Sigma}_{it}; \boldsymbol{\theta}) = \max_{\Pi_i} E\left[\sum_{\tau=t}^{\infty} \delta^{\tau-t} U(\boldsymbol{c}_{it}, \boldsymbol{x}_{it}, \boldsymbol{\iota}_{it}, \boldsymbol{p}_{it}) | \boldsymbol{\Sigma}_{it}, \Pi_i; \boldsymbol{\theta}\right], \tag{5}$$

where Π_i is a set of decision rules that map the state in purchase t, $\boldsymbol{\Sigma}_{it}$, into actions, which are how to purchase, $\boldsymbol{x}_{it}$, and how much to consume, $\boldsymbol{c}_{it}$. The parameter δ is a discount factor, which is assumed to equal 0.95.[12] The expectation is taken over the error term $\boldsymbol{\epsilon}$, and the evolution of future prices. The function $V(\boldsymbol{\Sigma}_{it}; \boldsymbol{\theta})$ is a value function, and is a solution to the Bellman's equation

$$V(\boldsymbol{\Sigma}_{it}; \boldsymbol{\theta}) = E_{\boldsymbol{\epsilon}}[\max_{\boldsymbol{c}_{it}, \boldsymbol{x}_{it}} \{U(\boldsymbol{c}_{it}, \boldsymbol{x}_{it}, \boldsymbol{\iota}_{it}, \boldsymbol{p}_{it}) + \delta E_{P(\boldsymbol{p}_{it+1}|\boldsymbol{p}_{it})} V(\boldsymbol{\Sigma}_{it+1}; \boldsymbol{\theta})\}]. \tag{6}$$

As a final note, we index the utility parameters γ_{ij} and α_i by i, which indicates that they are heterogeneous across the population. We assume that these parameters are lognormally distributed across the population. In our estimation we experimented with letting the other model parameters (β, SC_0, SC_1, and CC) be heterogeneous, but found it was difficult to identify the variances of those parameters.

5 Price Expectations

Consumer price expectations are an important component of a stockpiling model. We do not actually observe price expectations in the data, so we have to make an assumption about how they look. We assume that consumers are rational: their expectations about future prices correspond to the actual evolution of prices. My approach is to estimate a price process from the data, and to assume that process corresponds to how consumers expect future prices to evolve.

There are several important features of the observed price process which the estimated price process should capture. First, as can be seen in Figure 3, the price series is relatively flat for most of the time, but periodically promotions occur for a short

[12] The discount factor is usually difficult to identify in forward-looking structural models, so it is common practice to assign it a value (Magnac and Thesmar 2002).

period of time. It is easy to see from the figure when a promotion occurs, but it is more difficult to define a rule which splits prices into promotional and regular prices. We find that defining a promotion as any price below the median price works well - if one overlays the median price (which is 59 cents) on the price series graphs, the flat areas all lie above it and the promotions lie below it. The price process that we propose models the probability a promotion occurs using a discrete Markov process.

A second important feature is that, when a product is not on promotion, its price is flat for long periods of time. We model the price process during these periods using a discrete- continuous Markov process that is similar to that of Erdem, Imai, and Keane (2003). Conditional on no promotion occurring in weeks $t-1$ and t, there is a probability that the price changes. If the price does change, a truncated regression is used to predict that change.

The third feature that our price process captures is the existence of competitor reactions: when one firm's price drops, the competitor drops shortly afterwards. This can be seen in Figure 4, which shows the price paths for Starkist and Chicken of the Sea in one store. Our price process allows the probability of promotions to depend on the competitor's prices and promotional behavior.

The state variable governing promotions, s_{jt} is one when product j is on promotion in week t, and 0 otherwise. We estimate the Markov transition process on the probability of a sale occurring using a probit model:

$$\begin{aligned} &P(s_{jt}=1|s_{jt-1}=k) \\ &=1-\Phi\left(-\left[\psi^{pr}_{0jk}+\psi^{pr}_{1jk}p_{jt-1}+\psi^{pr}_{2jk}p_{-jt-1}+\psi^{pr}_{3jk}s_{-jt-1}+\psi^{pr}_{4jk}s_{jt-1}p_{-jt-1}\right]\right). \end{aligned} \tag{7}$$

The superscript *pr* stands for promotion, and the subscript $k \in \{0,1\}$. The probability of transitioning from the nonsale state into the sale state, or vice-versa, is governed by the competitor's previous price, the product's own previous price, and whether the competitor's product was on promotion.

When a product is not on sale for two consecutive weeks, its price will often stay constant for a few weeks. To account for this, I model the probability the $p_{jt}=p_{jt-1}$ using a probit process as

$$P(p_{jt} = p_{jt-1} | s_{jt} = 1, s_{jt-1} = 1) =$$
$$1 - \Phi\left(-\left[\psi^s_{0j} + \psi^s_{1j} p_{jt-1} + \psi^s_{2j} p_{-jt-1} + \psi^s_{3j} s_{-jt-1} + \psi^s_{4j} s_{jt-1} p_{-jt-1}\right]\right). \tag{8}$$

When the price of a product changes, we assume that the change is distributed according to a truncated normal distribution. If the product is transitioning into the non-sale state, then its distribution is censored at the median price, m_j; If it transitions into the sale state, we truncate from above at $m_j - 1$. I model this using a Tobit model, where I allow the parameters to depend on whether the product was previously on promotion ($s_{jt-1} = k$):

$$y_{jt} = \psi^c_{0jk} + \psi^c_{1jk} \ln(p_{jt-1}) + \psi^c_{2jk} \ln(p_{-jt-1}) + \psi^c_{3jk} s_{-jt-1} + \psi^c_{4jk} s_{jt-1} \ln(p_{-jt-1})$$
$$\ln(p_{jt}) = \begin{cases} y_{jt} & \text{if } y_{jt} > \ln(m_j) \\ \ln(m_j) & \text{if } y_{jt} \leq \ln(m_j). \end{cases} \tag{9}$$

A similar specification is run when the product transitions to a sale state. The inclusion of the competing brand's prices and promotions allow for competitor reactions. As a final note, we estimate all the parameters of consumer price expectations along with utility parameters in the model. Earlier work typically estimates the process for price expectations prior to estimating the model. Although doing this will produce consistent estimates, there could be inconsistency in the standard errors of the utility coefficients, due to them being functions of the estimated parameters of the price process. Our approach avoids this problem.

6 Estimation Technique

This section outlines the two major parts of the estimation. We use the Bayesian technique of Imai, Jain, and Ching (2009) to estimate our model. This approach has some computational and statistical advantages over classical approaches such as simulated maximum likelihood, or method of simulated moments. The primary computational difficulty that arises when estimating dynamic discrete choice models is that one must solve the Bellman equation that gives rise to the consumer value function. Typically this is done via a nested fixed point algorithm: every time one changes the parameter values when optimizing the objective function, one solves the contraction mapping that defines the value function using an iterative procedure.[13] Solving for the value function over and over again significantly increases the computational complexity of estimating dynamic discrete choice models. In contrast, in the Bayesian technique of Imai, Jain, and Ching (2009) one only solves for the value function once over the course of the estimation. In every step of the Gibbs sampler, one performs a single value function update. As the parameters of the Gibbs sampler converge to draws from the posterior distribution, the updated value functions at each Gibbs draw converge to the solution of the Bellman equation. Overall, the Gibbs chain has three main steps in it: step one draws population-varying coefficients, step two draws population-fixed coefficients, and step three draws price expectations coefficients. The following section discusses the first three steps, and the next discusses the fourth step.

6.1 Gibbs Steps for Utility Parameters

First we consider the Gibbs chain for the demand parameters. The vector of utility coefficients being estimated is $(\gamma_{i1}, ..., \gamma_{iJ}, \beta, \alpha_i, sc_0, sc_1, CC)$. Denoting the vector of population-varying parameters as $\tilde{\boldsymbol{\theta}}_i = (\gamma_{i1}, \gamma_{i2}, \alpha_i)$, we assume that $\ln(\tilde{\boldsymbol{\theta}}) \sim N(\boldsymbol{b}, \boldsymbol{W})$, where the matrix $\boldsymbol{W}$ is diagonal. We denote the vector of population-fixed parameters as $\boldsymbol{\theta} = (\beta, sc_0, sc_1, CC)$ and the vector of price expectations coefficients, which enter consumer expectations, as $\boldsymbol{\psi}$.

In many of the Gibbs steps, it is necessary to compute the probability of a consumer's sequence of observed choices conditional on a set of draws on initial inventories, $\boldsymbol{\iota}_{i0}$, observed prices, $\boldsymbol{p}_{i0}, ..., \boldsymbol{p}_{iT}$, and consumer utility coefficients $\boldsymbol{\theta}$. An observation here is a household-week, and the household's purchase decision, $\boldsymbol{x}_{it}$, is observed. We need to construct the probability of each household's sequence of purchase decisions. Household quantity decisions are unobserved, but conditional on a value of $\boldsymbol{x}_{it}$, and all the previous parameters, these can easily be calculated as

[13] See Rust (1994) and Rust (1987) for an overview of the nested fixed point algorithm.

$$\boldsymbol{c}_{it}* = arg\max_{\boldsymbol{c}_{it}} \left\{ \sum_{j=1}^{J} \gamma_{ij} u(c_{jit};\beta) - sc_0 I_{jit} - sc_1 I_{jit}^2 + EV(\boldsymbol{\Sigma}_{it}; \tilde{\boldsymbol{\theta}}_i, \boldsymbol{\theta}, \boldsymbol{\psi}) \right\}. \tag{10}$$

An approximation to $EV(\boldsymbol{\Sigma}_{it}; \tilde{\boldsymbol{\theta}}_i, \boldsymbol{\theta}, \boldsymbol{\psi})$ is computed using a nearest neighbor algorithm which is described in Section 6.3. Inventories are unobserved in period t, but they can be computed conditional on the initial inventories in period 0 (we describe how we compute period 0 inventories at the end of this section). In other words, conditional on all possible choices in period 1, one can compute period 1's consumption. Then the inventory at the beginning of period 2 is the period 0 inventory, plus period 1's observed purchase, $\boldsymbol{x}_{i1}$, minus the optimal consumption, $\boldsymbol{c}_{i1}^*$. Period 2's inventory can be constructed similarly, and so on. Denote the utility from the optimal consumption in period t as

$$\begin{aligned}\nu(\boldsymbol{x}_{it}) = &\max_{\boldsymbol{c}_{it}} \left\{ \sum_{j=1}^{J} \gamma_{ij} u(c_{jit};\beta) - sc_0 I_{jit} - sc_1 I_{jit}^2 + EV(\boldsymbol{\Sigma}_{it}; \tilde{\boldsymbol{\theta}}_i, \boldsymbol{\theta}, \boldsymbol{\psi}) \right\} \\ &- CC\, \mathbf{1}\{\sum_{j=1}^{J} x_{jit} > 0\} - \alpha_i \sum_{j=1}^{J} x_{jit} p_{jit}\end{aligned}$$

I assume that the choice specific error term, ϵ_{qit}, is logit. Denote the observed $\boldsymbol{x}_{it}$ in period t as $\boldsymbol{x}_{it}^{obs}$, and denote each possible value of $\boldsymbol{x}_{it}$, indexed by q, as $\boldsymbol{x}^q$. Then a consumer's sequence of choices can easily be computed as

$$Pr_i(\tilde{\boldsymbol{\theta}}_i, \boldsymbol{\theta}, \boldsymbol{\psi}) = \prod_{t=1}^{T} \frac{\exp(\nu(\boldsymbol{x}_{it}^{obs}))}{\sum_{q=1}^{N_u(N_u-1)/2} \exp(\nu(\boldsymbol{x}^q))}. \tag{11}$$

The sum in the denominator in equation (11) goes from $q = 1$ up to $N_u(N_u - 1)/2$ because we assume that consumers only purchase a maximum of N_u units in a single purchase occasion, and there are 2 brands, so $N_u(N_u - 1)/2$ is the total number of brand combinations that can be purchased.

To draw the utility parameters, we first draw $\tilde{\boldsymbol{\theta}}_i$ for each consumer i, conditional on $\boldsymbol{\theta}$, $\{\boldsymbol{\iota}_{i0}\}_{i=1}^{N}$ and and $\boldsymbol{\psi}$. The posterior distribution of $\tilde{\boldsymbol{\theta}}_i$ does not have a closed-form solution, so we use the Metropolis-Hastings (MH) algorithm to draw a new $\tilde{\boldsymbol{\theta}}_i$.

Denoting the previous draw on $\tilde{\boldsymbol{\theta}}_i$ as $\tilde{\boldsymbol{\theta}}_i^0$, our procedure is to draw a candidate $\ln(\tilde{\boldsymbol{\theta}})_i^1 \sim N(\ln(\tilde{\boldsymbol{\theta}})_i, \rho^2 \boldsymbol{W})$, and to accept the new draw with probability

$$\frac{Pr_i(\tilde{\boldsymbol{\theta}}_i^1, \boldsymbol{\theta}, \boldsymbol{\psi})}{Pr(\tilde{\boldsymbol{\theta}}_i^0, \boldsymbol{\theta}, \boldsymbol{\psi})}.$$

The parameter ρ is adjusted every period so that about 30% of the candidate $\tilde{\boldsymbol{\theta}}_i^1$ draws are accepted every period (see **?**)). Once we have drawn $\{\tilde{\boldsymbol{\theta}}_i\}_{i=1}^N$, we draw out $\boldsymbol{b}$ and $\boldsymbol{W}$ conditional on the logarithm of the draws. We assume that the prior on $\boldsymbol{b}$ is normal with a prior variance of $100\boldsymbol{I}$, so the prior are proper but relatively uninformative. This prior generates a normal posterior distribution for $\boldsymbol{b}$, which we can draw from using standard methods. Similarly, we put an inverse gamma prior on each of the diagonal elements of $\boldsymbol{W}$ with scale parameter 3 and shape parameter 0.5, which results in an inverse gamma posterior for the elements of $\boldsymbol{W}$.[14]

Next we draw $\boldsymbol{\theta}$ given the drawn $\{\tilde{\boldsymbol{\theta}}_i\}_{i=1}^N$, the previous draw's $\{\boldsymbol{\iota}_{i0}\}_{i=1}^N$ and $\boldsymbol{\psi}$. We also use a random walk MH for this step, with a parameter ρ_2 on the variance that periodically updates to keep the acceptance rate at 30%. Given our previous draw $\boldsymbol{\theta}^0$, we draw a candidate $\boldsymbol{\theta}^1$ where

$$\boldsymbol{\theta}^1 \sim N(\boldsymbol{\theta}^0, \rho_2^2 \boldsymbol{I}).$$

We accept the candidate draw with probability

$$\left(\prod_{i=1}^{I} \frac{Pr_i(\tilde{\boldsymbol{\theta}}_i, \boldsymbol{\theta}^1, \boldsymbol{\psi})}{Pr_i(\tilde{\boldsymbol{\theta}}_i, \boldsymbol{\theta}^0, \boldsymbol{\psi})} \right) \frac{k(\boldsymbol{\theta}^1)}{k(\boldsymbol{\theta}^0)},$$

where k is the prior distribution on $\boldsymbol{\theta}$. We assume that k is normal, with a variance matrix of $100\boldsymbol{I}$ so that priors are close to uninformative. The prior means are set to the average parameter estimates from a similar model presented in Osborne (2010).

[14]Most of the mass of this prior distribution is between 0 and 0.5, which would be too tight if we were not exponentiating the utility parameters. However, our utility parameters are lognormal, which means a relatively small variance in the underlying normal can lead to a large variance in the resulting lognormal. For example, the exponential of a normal distribution with mean 1 and variance 0.5 has variance of about 13.7.

Our choice probabilities above depend on inventories prior to period 1, $\boldsymbol{\iota}_{i0}$, which are unobserved. We integrate them out using simulation. Our approach is to simulate $\boldsymbol{\iota}_{i0}$ conditional on $\tilde{\boldsymbol{\theta}}_i$, $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$ for each consumer. The initial inventories are drawn using the incomplete data prior to the final year of the data, where consumer purchases are observed but store visits are unobserved when purchases are not made. We proceed by first drawing a sequence of store visits for each household where a visit is unobserved from the empirical density of observed store visits for that particular household.[15] We than assume that at the beginning of the data (73 weeks prior to the beginning of the sample) initial inventories are zero. We then compute optimal consumption in each period, conditional on the parameter draws, prices and quantities chosen in each period, which gives us end of period inventories. $\boldsymbol{\iota}_{i0}$ is then taken to be the level of inventories computed at the end of the 73rd week. A new draw on the initial inventories is taken in each Gibbs step.

6.2 Gibbs Step for Price Expectations

Consumer price expectations are assumed to coincide with the distribution of prices that are observed in our data. To see how we draw the parameters governing these, consider first the posterior distribution for the $\boldsymbol{\psi}^{pr}$, the parameters associated with the probability of a price change. We use standard data-augmentation techniques, with an adjustment that accounts for consumer choice probabilities. For a given consumer i in week t and a vector of observables for brand j, $\boldsymbol{w}_{j,i,t} = (1, p_{j,i,t-1}, p_{-j,i,t-1}, s_{-j,i,t-1}, s_{j,i,t-1}p_{-j,i,t-1})$, consider a latent variable formation of the probit model

[15] A slightly different procedure was presented in Osborne (2010) where the $\boldsymbol{\iota}_{i0}$'s were assigned a distribution and draws were taken on the $\boldsymbol{\iota}_{i0}$ in a separate Gibbs step to rationalize the final year of choice data. The procedure used in Osborne (2010) produces a distribution of initial inventories that is zero for more than 99% of households. Although the approach of Osborne (2010) is more nonparametric, our approach produces a distribution of initial inventories which more closely mimics the distribution of inventories we compute at the very end of our sample, which we feel is more realistic. Since the consumers in our sample are fairly regular purchasers of canned tuna, it seems unlikely that almost all of them would have no tuna in inventory at the beginning of the sample. The distribution of initial inventories may be difficult to identify nonparametrically.

$$z_{j,i,t} = \boldsymbol{w}'_{i,t}\boldsymbol{\psi}^{pr}_j + e_{j,i,t},\ e_{j,i,t} \sim N(0,1)$$
$$y_{j,i,t} = \begin{cases} 0 & \text{if } z_{j,i,t} \leq 0 \\ 1 & \text{otherwise} \end{cases}.$$

Given the prior iteration's draw on $\boldsymbol{\psi}^{pr}$, which we call $\boldsymbol{\psi}^{pr,0}$, we draw out the z's. We then draw a new $\boldsymbol{\psi}^{pr,1}$ using the MH algorithm. As a proposal density, we use the density of $\boldsymbol{\psi}^{pr,1}$ conditional on the $z_{j,i,t}$'s and the data on prices observed by all consumers, but not consumer choices.[16] Note that this density is just standard normal. Then we accept the new draw with probability $\prod_{i=1}^{I} Pr_i(\boldsymbol{\psi}^{pr,1}; \boldsymbol{\theta}, \boldsymbol{\iota}_{i0})/Pr_i(\boldsymbol{\psi}^{pr,0}; \boldsymbol{\theta}, \boldsymbol{\iota}_{i0})$. We also need to draw $\boldsymbol{\psi}^{c}$ and $\boldsymbol{\psi}^{s}$. Note that if it were not for the consumer choices, we could draw both of these parameters using standard Bayesian procedures for the Tobit and probit models. We therefore use the same type of proposal distributions with both these parameters - we draw any latent unobservables, and then draw candidate parameters conditional on the latent unobservables and the store level data. We accept or reject the vector of $\boldsymbol{\psi}^{pr}$'s, $\boldsymbol{\psi}^{c}$'s, and $\boldsymbol{\psi}^{s}$'s jointly to save computational time. The acceptance rate in this step tends to be high, as the ratio $\prod_{i=1}^{I} Pr_i(\boldsymbol{\psi}^{pr,1}; \boldsymbol{\theta}, \boldsymbol{\iota}_{i0})/Pr_i(\boldsymbol{\psi}^{pr,0}; \boldsymbol{\theta}, \boldsymbol{\iota}_{i0})$ is typically close to 1.

6.3 Approximating the Expected Value Function

One of the most computationally intensive parts of the estimation procedure is computing an estimate of the value function, which we need to draw the parameters $\tilde{\boldsymbol{\theta}}_i^1, \boldsymbol{\theta}, \boldsymbol{\psi}$,

[16] We assume that the price expectations we estimate accurately capture the price process that consumer i observes. The prices observed by the consumer will depend on both the pricing policies of the canned tuna firms, and the choice of store that the consumer makes. The equations describing price expectations outlined in Section 5 will therefore capture the composition of both these random variables, and the correct data to use will be the data on consumer prices at the level of the consumer. As discussed by Erdem, Imai, and Keane (2003), an alternative approach is to estimate the price process at the store level and to adjust for store visit probabilities afterwards. Their paper finds that both of these ways of modeling price expectations generate very similar parameter estimates. Note that this specification assumes that the choice of store is exogenous to the decision to purchase canned tuna.

to compute simulated quantities $\hat{q}$, and to update the value function (we describe the updating procedure in the next section). Here we follow the Imai, Jain, and Ching (2009) procedure of starting with a guess of the value function, and computing one update to the value function at each Gibbs iteration. In each update, it is necessary to compute an estimate of the value function at the current Gibbs draw. This is done by averaging over value functions saved in previous Gibbs draws where the parameters are "close" to the current draw. We use a nearest neighbor approach to selecting close parameter draws. Another alternative is to use a kernel-weighted average of past value functions, where the kernel weights measure the closeness of the current draw to past draws. Norets (2009) discusses advantages and disadvantages of each approach.

At step g of the Gibbs sampler, we will have $N(g)$ saved draws on $\tilde{\boldsymbol{\theta}}_i$ for every household and $N(g)$ saved draws on $\boldsymbol{\theta}, \boldsymbol{\psi}$. We will also have $N(g)$ saved value function estimates for each state space point at each of those draws.[17] Denote the draw on $\boldsymbol{\theta}$ ($\tilde{\boldsymbol{\theta}}_i$) from iteration k as $\boldsymbol{\theta}^k$ ($\boldsymbol{\theta}_i^k$), $k = g - N(g), ..., g-1$ and the observed price expectations draw as $\boldsymbol{\psi}^k$. Denote the vector composed of $\tilde{\boldsymbol{\theta}}_i^k$, $\boldsymbol{\theta}^k$ and $\boldsymbol{\psi}^k$ as $\boldsymbol{\Gamma}_i^k = (\tilde{\boldsymbol{\theta}}_i, \boldsymbol{\theta}^k, \boldsymbol{\psi}^k)$.

For each i and k we will have an associated saved value function, $V_{i,k}(\boldsymbol{\iota}, \boldsymbol{p})$. We will have a saved value function for each value of $\boldsymbol{\iota}$, since $\boldsymbol{\iota}$ is discrete, but not for $\boldsymbol{p}$, since $\boldsymbol{p}$ is continuous. Instead what we do is draw a set of N_p prices in each iteration from an importance distribution $h(\cdot)$. This means we have a set of $\{\boldsymbol{p}_{g,s}\}_{s=1,g=g-N(g)}^{N_p,g-1}$ saved prices. Suppose now that we need an estimate of the expected value function at some parameter vector $\boldsymbol{\Gamma}$, a price $\boldsymbol{p}$, and an inventory $\boldsymbol{\iota}$. We do this in two steps. First, for household i we find the closest $\boldsymbol{\Gamma}_i^k$'s to $\boldsymbol{\Gamma}$. We will find $m = 1, ..., \tilde{N}$ of them as follows, indexing them as $\tilde{l}^m$

$$\begin{aligned}
\tilde{l}^1 &= \min_{l \in \{1,...,N(g)\}} \{\|\boldsymbol{\Gamma} - \boldsymbol{\Gamma}_i^l\|\} \\
\tilde{l}^2 &= \min_{l \in \{1,...,N(g)\}/\tilde{l}^1} \{\|\boldsymbol{\Gamma} - \boldsymbol{\Gamma}_i^l\|\} \\
&\vdots \\
\tilde{l}^{\tilde{N}} &= \min_{l \in \{1,...,N(g)\}/\{\tilde{l}^1,...,\tilde{l}^{\tilde{N}-1}\}} \{\|\boldsymbol{\Gamma} - \boldsymbol{\Gamma}_i^k\|\},
\end{aligned}$$

[17] When starting the algorithm, we begin with an estimate of the value function that is zero in all states.

where the $\|\|$ function indicates the Euclidean norm and / indicates set subtraction. Our estimated value function will average over the $\tilde{l}^m$'s, as well as the saved $\boldsymbol{p}_{g,s}$'s at each Gibbs iteration associated with $\tilde{l}^m$, which we will call $g(\tilde{l}^m)$. Note that we need to average over the saved $\boldsymbol{p}_{g,s}$'s, because when we compute the expected value function at a current price $\boldsymbol{p}$, we must integrate over the distribution of possible prices tomorrow. To average over the saved $\boldsymbol{p}_{g,s}$'s, we use importance sampling. We compute the probability that $\boldsymbol{p}_{g,s}$ occurs tomorrow given today's price $\boldsymbol{p}$ using the transition density at the current $\boldsymbol{\psi}$, $P(\boldsymbol{p}_{g,s}|\boldsymbol{p},\boldsymbol{\psi})$, and use $P(\boldsymbol{p}_{g,s}|\boldsymbol{p},\boldsymbol{\psi})/h(\boldsymbol{p}_{g,s})$ as the importance weight. The estimated expected value function is then

$$\hat{EV}_i(\boldsymbol{\iota},\boldsymbol{p};\boldsymbol{\Gamma}) = \frac{1}{\tilde{N}} \sum_{m=1}^{\tilde{N}} \frac{\sum_{s=1}^{N_p} V_{i,\tilde{l}^m}\left(\boldsymbol{\iota},\boldsymbol{p}_{g(\tilde{l}^m),s}\right) \frac{P(\boldsymbol{p}_{g(\tilde{l}^m),s}|\boldsymbol{p},\boldsymbol{\psi})}{h(\boldsymbol{p}_{g(\tilde{l}^m),s})}}{\sum_{s=1}^{N_p} \frac{P(\boldsymbol{p}_{g(\tilde{l}^m),s}|\boldsymbol{p},\boldsymbol{\psi})}{h(\boldsymbol{p}_{g(\tilde{l}^m),s})}}.$$

6.4 Updating the Value Function

The last step of the algorithm is to update the value function $V_{i,k}(\boldsymbol{\iota},\boldsymbol{p}_{g,s})$ at every $\boldsymbol{\iota}$ and $\boldsymbol{p}_{g,s}$ combination. Note that every iteration, we only perform a single value function update, as outlined in Imai, Jain, and Ching (2009). Over the course of running the Gibbs sampler the value function will converge. Updating the value function is relatively straightforward. At the end of iteration g, we will have a new parameter vector $\boldsymbol{\theta}^g$, as well as a draw on $\boldsymbol{\psi}^g$. First, given some $\boldsymbol{\Gamma} = (\boldsymbol{\theta}_i^g, \boldsymbol{\theta}^g, \boldsymbol{\psi}^g)$ we compute $\hat{EV}(\boldsymbol{\iota},\boldsymbol{p}_{g,s};\boldsymbol{\Gamma})$ for all possible $\boldsymbol{\iota}$ values. Then, we compute optimal consumption at all possible $\boldsymbol{x}$ choices, which we will denote $\boldsymbol{c}^*(\boldsymbol{x}^q)$. Net of the error term, utility at this point will be

$$\hat{\nu}(\boldsymbol{x}^q) = \sum_{j=1}^{J} \gamma_{ij} u(c^*;\beta) - sc_0 I - sc_1 I^2 + \delta \hat{EV}(\boldsymbol{\iota}',\boldsymbol{p}_{g,s};\Gamma) - CC\mathbf{1}\{\sum_{j=1}^{J} x_j^q > 0\} - \alpha_i \sum_{j=1}^{J} x_j^q p_{j,g,s},$$

where $\boldsymbol{\iota}'$ denotes period $t+1$ inventory. The updated value function will then be

$$V_{i,l}(\boldsymbol{\iota},\boldsymbol{p}_{g,s}) = \log\left(\sum_{q=1}^{N_u(N_u-1)/2} \exp(\hat{\nu}(\boldsymbol{x}^q))\right).$$

7 Parameter Identification

In this section I present an informal argument for the identification of the demand and costs parameters. On the demand side, most of the argument is similar to that laid out in Erdem, Imai, and Keane (2003), which presents a qualitatively similar model. First, consider the identification of the parameters on inventory costs. An decrease in the linear parameter on inventory costs, sc_0, will cause consumers to purchase smaller quantities more often: the overall incentive to stockpile will decrease. An increase in the quadratic term will cause consumers to avoid making large increases in inventory. Thus, if a consumer just purchased a large amount of a product, and the product is observed to go on sale again, we should observe the consumer to be unlikely to stockpile in response. Next consider the identification of the taste parameters. Conditional on a given amount in inventory, increasing the quadratic term will decrease the amount of tuna consumed.[18] This parameter is analogous to a consumption rate. Changing this parameter will impact the duration dependence of the purchase hazard. The impact of changing this parameter is nonlinear: if β is high, then duration dependence increases - the time between purchases drops. If it is low, the time between purchases rises. Erdem, Imai, and Keane (2003) show that the multiplicative taste parameter (γ_j) increases overall demand, but does not affect duration dependence. Thus, market shares will drive this parameter.

Identification of the carrying cost parameter follows from the overall purchase frequency. The higher is the carrying cost, the less frequently will consumers make purchases. It is separately identified from the stockpiling costs because it does not impact the spacing between purchases - a consumer's disutility due to carrying costs from making five purchases is the same whether the five purchases occur in five consecutive weeks, or over five months, each occurring once per month. However, making five consecutive purchases will drive up stockpiling costs more than five purchases which spaced further apart.

The heterogeneity in the price coefficient will be driven by heterogeneity in con-

[18] To see this, consider the case where consumers are not forward-looking, and there are no inventory costs. A consumer's optimal consumption of c_j will be $-(1+2\beta)/2\beta$.

sumer response to promotions. In the data, we observe consumers making purchases over a year. Since different consumers shop at the same stores and observe the same prices, we will see if some consumers tend to be more likely to enter the market during promotions than other consumers. Heterogeneity in tastes will be driven by heterogeneity in consumer response to promotions for different products: we can infer that consumers who are unresponsive to promotions for Starkist, but are responsive to those for Chicken of the Sea have a preference for Chicken of the Sea, and vice versa.

8 Model Estimates

The Gibbs sampler is run for 13,000 iterations, where the first 3,000 draws are removed to reduce dependence on the starting values. We estimate the model on a 50% subsample of the data to reduce computational burden - the estimation data includes 317 households. We plot the time series of parameter draws in Figure 5, where it can be seen that a little after around 2,000 draws the parameter draws become relatively stable. To reduce autocorrelation across draws, when we present our estimates we use every 10th draw as suggested by Train (2003). We save $\tilde{N} = 30$ previous value function draws, choosing the 3 closest previous value function draws. We draw 20 new prices every iteration.

Estimates of the utility parameters and marginal costs are shown in Table 3. The first column of the table shows the average of the draws for the population mean utility parameters. For population-varying coefficients, I present the average (over draws) of the draw mean of their lognormal distribution - if the draw on the underlying normal parameters are a mean of b and variance of w, then the draw on the mean of the lognormal is $\exp(b + w/2)$. For population-fixed parameters, the population mean is simply the draw. The second column shows the posterior standard deviations of these draws. The third shows the average of the draws for the population standard deviation parameters. Again, for the population-varying parameters I present the average of draws on the standard deviation of the lognormal.[19] For population-fixed parameters

[19] For an underlying normal draw with mean parameter b and variance w, the standard deviation of the lognormal is $\sqrt{(\exp(w) - 1)\exp(2b + w)}$.

these columns are dashed out.

Turning to the estimates, the average tastes for Starkist and Chicken of the Sea are similar, which is consistent with their marketshares being similar. The curvature parameter of about -0.8 implies that, absent forward-looking behavior, a consumer would use between one and two cans of tuna per week. The small estimates of the inventory cost parameters imply negligible holding costs. The price coefficient is about 0.028, which is consistent with significant price sensitivity. Since price is measured in cents, a increase in price of ten cents decreases consumer utility by about 0.28. Flow utilities average at around -0.6, so price is a significant determinant of the purchase decision. The carrying cost is large and significant, which is consistent with purchases of canned tuna being relatively infrequent. In dollar terms, the carrying cost is measured to be a little less than one dollar for an average consumer. The estimates of the standard deviations of the population-varying parameters are consistent with a significant amount of consumer heterogeneity. Note that although the standard deviation of the price coefficient is small in absolute terms, relative to the average price coefficient it is large, indicating significant unobserved heterogeneity in price sensitivity.

Tables 4 and 6 show the estimated posterior means and standard deviations for the price process parameters. The first two sections of Table 4 show the estimates of the probit model that determines the probability a product goes on sale conditional on it being on sale in the previous period. Conditional on being on sale, a product has about a 50 percent chance of being on sale the next period; if it is not on sale, though, the product is unlikely to go on sale. This can be seen in the predicted sale probabilities in Table 5, which shows summary statistics of the predicted probability of a sale for sale or nonsale observations in the store level data.[20] If Starkist is on sale, the likelihood it stays on sale the next period is 0.5; if it is not, then the likelihood it will go on sale is only 0.16, and Chicken of the Sea looks similar to this. The transition probabilities of sale and nonsale states closely match our model predictions. This type of time-series variation in prices will help to drive stockpiling behavior: when consumers observe a promotional price, they will know it is likely to be short-lived and will stockpile in

[20]The predicted probabilities are computed for each observation at the mean of the posterior estimates.

response to it.

9 A Dynamic Price Index for Canned Tuna

This section presents the price index for canned tuna we derive from our model estimates, and compares that index to several alternative price indexes: fixed-base Laspeyres, geometric and Tornqvist indexes, chained Laspeyres and Tornqvist, and an alternative COLI index proposed by Feenstra and Shapiro (2003) (FS).[21] The first step of computing our index is to compute an estimated COLI: in a given draw of the Gibbs sampler we compute average flow utilities under optimal consumer behavior across the population for each week of the data. Following Gowrisankaran and Rysman (2011), we compute a COLI for canned tuna by constructing a sequence of taxes or subsidies that hold average flow utilities constant over time. Because the sequence of taxes/subsidies is not dependent on individual behavior, it will not affect individual consumption choices.

Price indexes such as those produced by the BLS are not typically presented as COLIs, although as we discussed in the Introduction, they are motivated by COLIs. In order to compare our sequence of taxes, which is a COLI, to a price index we need a way to map one to the other. Gowrisankaran and Rysman (2011) provide such a mapping. To convert a price index to a COLI, one first computes a share weighted average price in a base period. Then, one multiplies the index every period by the share-weighted average price, divides the result by 100, and subtracts the result from the initial average price. This index provides the change in income relative to the initial period needed to purchase the same basket of goods. Rather than taking the price indexes we will compute and turning them into COLIs, we will reverse the transformation above and turn our COLI into a price index. We will add the initial average price to the COLI each week, and divide the result by the initial average price.

As a base price, we choose the modal price of each SKU in each store in the year prior to the one in which we estimate our model. We use as quantity weights the

[21]The alternative indexes are computed using data from our full sample of 600 households, rather than the 50% subsample the estimation was done on.

average quantity sold of the SKU in the store at the modal price. Our procedure for choosing a base price follows that of FS, and will help ensure that our results are comparable to theirs. To compute the average price index in each week, we average the price index over saved Gibbs draws. Figure 6 shows the how the estimated price index evolves over time. The dotted red lines around the index show 95% confidence intervals for each period, which we can compute from the Gibbs draws. The index is quite volatile, which is consistent with the significant amount of volatility we observe in prices. Additionally, the index is usually below 1, consistent with prices falling over time which was documented earlier.

We then compare our indexes to several alternatives. The 'standard' price index for food products produced by the BLS (which is used for instance to deflate PCE food items) is a Laspeyres index:

$$P_t^L = \sum_i w_{i0} \frac{p_{it}}{p_{i0}}.$$

In the above equation, i indexes a SKU-store combination. w_{i0} is the base period expenditure share for product i, and p_{i0} is the modal price in the previous year. We also compute fixed base Geometric and Tornqvist indexes, as well as chained Laspeyres and Tornqvist indexes. The formulas for those indexes are given in FS, and we follow their procedures here. The final index we use as a comparison is a COLI index proposed by FS for frequently purchased, storable products. FS develop a representative consumer model where it is assumed that consumption and purchases differ, and consumers have perfect foresight over future prices. If one approximates the expenditure function arising from their model with a translog functional form, results due to Caves, Christensen, and Diewert (1982a) and Caves, Christensen, and Diewert (1982b) show that the resulting COLI has the same functional form as a fixed-base Tornqvist index,

$$COL^{FS} = \exp\left(\sum_{t=1}^{T}\sum_{i=1}^{N}\frac{1}{2}(s_{i0}+s_{it})\ln\left(\frac{p_{it}}{p_{i0}}\right)\right),$$

where s_{i0} is the expenditure share of SKU-store combination i in the base period, where the share is taken over all weeks during the base period. In our case we take the modal price in the previous year, times the average quantity sold at that price, and divide by total expenditures on that product-store combination over the course of the entire year. Similarly, s_{it} is the expenditure share of product-store combination i in week t of the year of analysis.

Table 7 summarizes the comparison of our index to the alternatives. The first column shows the average of each index over the course of the year. Our dynamic index, shown in the first row, is on average 82%. In comparison, all the fixed-base indexes are higher, averaging about 90%. What is likely going on here is that the fixed-base indexes put equal weight on all prices in each week relative to the base week.[22] However, our index puts more weight on sale periods relative to the base period, because consumers wait to purchase until promotions. Hence, our index will produce larger price falls relative to the base period. The chained indexes display much wider variation. The chained Laspeyres rises significantly over time, ending at a value of 1.3. As discussed in the introduction, the chained Laspeyres is known to have a significant upward bias. The chained Tornqvist, on the other hand, drops significantly over time. As we will discuss below, this is likely due to the Tornqvist overweighting price falls relative to price rises.

Finally, the COLI suggested by FS suggests an 84% fall in prices, which is similar to our dynamic index. This suggests that even though the FS COLI requires simplifying assumptions relative to our more sophisticated model, it produces a reasonably good approximation to our price index. Additionally, it is by far the closest index to ours - the next closest index, the fixed base Tornqvist, is about 7% above our index. Our results are generally in line with FS, who find that fixed base indexes approximate their COLI better than chained indexes. However, we suspect that the reasons for this differ from those found in FS. FS find that in their data, most purchases of canned tuna occur in the later weeks of a promotion. Those purchases also coincide with feature

[22]The weights in our indexes are computed using modal prices in the previous year, following FS. However, we rebase the indexes to the price we observe in the first week of the data used for estimation to retain comparability to the dynamic COLI. As a robustness check, we computed all of the indexes in FS using week 1 prices and quantities as initial weights. Our results were qualitatively similar - the FS COLI rises to 88%, and the other indexes rise by roughly 5% on average. The FS COLI is still the closest index to our dynamic index.

and display advertising. We checked our data to see whether the same phenomenon occurred. First, we categorized promotions by the number of weeks the promotion lasted.[23] We found that about 80 percent of promotions are two weeks long, and about 65 percent of the quantity purchased during a two week sale happened in the first week.[24] Additionally, for the two week promotions, the product being promoted was on advertised 99 percent of the time in the first week, but only 13 percent of the time in the second week. The reason that the FS COLI approximates our index well is likely due to the fact that the FS COLI puts more weight on weeks where promotions occur relative to non-promotion weeks. As we discussed above, neither the fixed-base nor the chained indexes correctly weight promotional weeks where more purchases occur. FS also note that the incidence of purchases near the end of a promotion tends to drive a significant amount of upward bias in chained indexes - their chained Laspeyres and chained Tornqvist indexes rise 37 to 45 percent over the course of the year. In contrast, although our chained Laspeyres also rises, our chained Tornqvist drops significantly over time. This is probably due to the fact that the chained Laspeyres uses fixed weights. The weights in the chained Tornqvist index are an average of weights in periods t and $t-1$. Because we observe purchases occurring primarily at the beginning of promotions, the chained Tornqvist will overweight initial price falls relative to price rises.

10 Discussion

We derive a price index for storable goods using the estimates of a dynamic structural model. Our results suggest that the closest standard price indexes to ours are fixed-base indexes. However, those price indexes are still significantly higher than our dynamic

[23] We denote a price drop as a promotion if the price of one of the products drops below the previous period's price and subsequently returns to a higher price. We limit the analysis to promotion spells between 1 and 3 weeks. Temporary price cuts seldom lasted longer than 3 weeks. We inspected graphs of prices for some stores and our procedure seemed to do a good job of identifying temporary price promotions.

[24] 3 week promotions occur about 10 percent of the time, and about 80 percent of purchases occur in the first or second week.

index, due to how they weight promotional periods relative to nonpromotional periods. The COLI index proposed by FS provides a good approximation to our index.

There is a significant amount of future work to be done to quantify the importance of stockpiling behavior in economic measurement. Although we have shown that stockpiling is important in purchases of canned tuna, it is likely to matter for many other products. Purchases of food at home, which is a significant part of Personal Consumption Expenditures as produced by the Bureau of Economic Analysis, are an aggregate measure of the value of all food at home products, weighted by their respective price indexes. A useful future exercise would be to investigate the impact of using alternative indexes, such as the one proposed in this paper, on measurement of output. Additionally, we found that our price index displayed more volatility from week to week than BLS type indexes - in Table 7 the standard deviation of our index is about twice that of the other three fixed base indexes. This suggests that month to month volatility in food at home purchases may be greater than what is implied by the estimates produced by the Bureau of Economic Analysis. Looking further, it is likely that other types of dynamics will have a significant impact on measures of prices. Gowrisankaran and Rysman (2011) provide an example with durable goods; dynamics such as learning or variety-seeking also will play an important role for many consumer products.

References

Aizcorbe, A. (2005). Price deflators for high technology goods and the new buyer problem. Unpublished manuscript, Board of Governors of the Federal Reserve Bank.

Bureau of Economic Analysis (2011, November). NIPA Handbook. Technical report.

Bureau of Labor Statistics (2007). BLS Handbook of Methods. Technical report.

Caves, D., L. Christensen, and W. E. Diewert (1982a). The economic theory of index numbers and the measurement of input, output, and productivity. Econometrica 50(11), 1393–1414.

Caves, D., L. Christensen, and W. E. Diewert (1982b, March). Multilateral comparisons of output, input and productivity using superlative index numbers. Economic Journal 92, 73–86.

Erdem, T., S. Imai, and M. Keane (2003). A model of consumer brand and quantity choice dynamics under price uncertainty. Quantitative Marketing and Economics 1(1), 5–64.

Feenstra, R. and M. Shapiro (2003, August). High-Frequency Substitution and the Measurement of Price Indexes. Univ. of Chicago and NBER. in Robert Feenstra and Matthew Shapiro (eds.), *Scanner Data and Price Indexes.*

Gowrisankaran, G. and M. Rysman (2011). Dynamics of consumer demand for new durable goods. Working Paper.

Hendel, I. and A. Nevo (2006). Sales and consumer inventory. The RAND Journal of Economics 37, 543–561.

Imai, S., N. Jain, and A. Ching (2009). Bayesian estimation of dynamic discrete choice models. Econometrica 77(6), 1865–1899.

Magnac, T. and D. Thesmar (2002). Identifying dynamic discrete decision processes. Econometrica 20(2), 801–816.

Norets, A. (2009). Inference in dynamic discrete choice models with serially correlated unobserved state variables. Econometrica 77, 1665–1682.

Osborne, M. (2010). Frequency versus depth: How changing the temporal process for promotions impacts demand for a storable good. Working Paper.

Osborne, M. (2011, March). Consumer learning, switching costs and heterogeneity: A structural examination. Quantitative Marketing and Economics 9(1), 25–70.

Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurchner. Econometrica 55, 993–1033.

Rust, J. (1994). Structural Estimation of Markov Decision Processes. Elsevier. Handbook of Econometrics Vol. 4 Engle R. and McFadden D (eds.).

Train, K. (2003). Discrete Choice Methods with Simulation. Cambridge University Press, New York.

Table 1: Summary of Data

	Starkist	Chicken of the Sea
Market Shares	48.2 %	51.8 %
Average Prices	$ 0.63	$ 0.61
Std Dev of Prices	$ 0.9	$ 0.11

Table 2: Test for Inventory Behavior: Household Level Regression of Quantity on Inventory

Regressor	Estimate	Std Err
Inventory	-2.60	0.119
Price	-0.985	0.046
Sale	6.48	0.525
Price*Sale	-0.583	0.053
Display	0.113	0.134
Feature	0.141	0.115

Regression includes household, store, and brand fixed effects.

Table 3: Estimates of Utility Coefficients and Costs

Coefficient	Mean	Std Dev	Variance	Std Dev
SK Taste (γ)	5.866	(0.274)	1.897	(0.18)
COS Taste (γ)	5.763	(0.253)	1.745	(0.144)
Curvature (β)	-0.797	(0.015)	-	-
Inv Cost Linear	-5e-04	(0.00014)	-	-
Inv Cost Quadratic	-1.6e-05	(4.1e-06)	-	-
Price (α_i)	0.028	(0.001)	0.005	(5e-04)
Carrying Cost	-2.346	(0.062)	-	-

Table 4: Estimates of Price Process:Probit of Transition Probabilities

	Starkist		COS	
Coefficient	Est	Std Err	Est	Std Err
Prob(sale|sale)				
Intercept	-0.878	(0.138)	2.043	(0.104)
Own Price	0.024	(0.003)	-0.027	(0.002)
Comp Price	-0.005	(0.001)	-0.008	(0.001)
Comp Sale	0.052	(0.146)	-0.721	(0.225)
Comp Sale*(Comp Price)	-0.004	(0.003)	0.017	(0.004)
Prob(sale|nonsale)				
Intercept	-0.242	(0.063)	1.068	(0.06)
Own Price	-0.015	(0.001)	-0.003	(0.001)
Comp Price	0.005	(0.001)	-0.022	(0.001)
Comp Sale	1.437	(0.138)	0.14	(0.178)
Comp Sale*(Comp Price)	-0.03	(0.003)	-0.006	(0.003)
Prob($p_t = p_{t-1}$|nonsale)				
Intercept	1.295	(0.061)	1.439	(0.069)
Own Price	-0.012	(0.001)	-0.016	(0.001)
Competitor Price	-0.003	(0.001)	0	(0.001)
Comp Sale	4.199	(0.133)	3.642	(0.19)
Comp Sale*Comp Price	-0.074	(0.002)	-0.073	(0.004)

Table 5: Predicted Transition Probabilities between Sale and Non-Sale Weeks

Starkist: Prob(sale|sale)					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.3965	0.4814	0.4849	0.4896	0.4985	0.6125
Starkist: Prob(sale|nonsale)					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.07763	0.13736	0.16496	0.15861	0.18404	0.24712
COS: Prob(sale|sale)					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.3478	0.4889	0.5643	0.5355	0.5945	0.7125
COS: Prob(sale|nonsale)					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.1224	0.1900	0.3400	0.2847	0.3500	0.3600

Table 6: Estimates of Price Process:Price Change Tobit Regressions

	Starkist		COS	
Coefficient	Est	Std Err	Est	Std Err
Nonsale-Nonsale				
Intercept	3.143	(0.246)	3.143	(0.246)
ln(ownprice)	-0.095	(0.07)	2.585	(0.255)
ln(compprice)	0.344	(0.066)	-0.327	(0.063)
Comp Sale	0.198	(0.297)	0.706	(0.076)
Comp Sale*ln(compprice)	-0.029	(0.074)	0.257	(0.306)
σ^2	0.175	(0.002)	-0.036	(0.077)
Nonsale-Sale				
Intercept	1.293	(0.255)	0.237	(0.004)
ln(ownprice)	0.36	(0.11)	1.646	(0.244)
ln(compprice)	0.273	(0.103)	0.162	(0.07)
Comp Sale	0.058	(0.29)	0.389	(0.085)
Comp Sale*ln(compprice)	0.015	(0.073)	0.069	(0.299)
σ^2	0.103	(0.008)	0.007	(0.076)
Sale-Nonsale				
Intercept	0.001	(0.292)	0.123	(0.003)
ln(ownprice)	0.664	(0.104)	-0.601	(0.274)
ln(compprice)	0.348	(0.077)	0.309	(0.076)
Comp Sale	0.156	(0.281)	0.834	(0.071)
Comp Sale*ln(compprice)	-0.032	(0.069)	0.111	(0.295)
σ^2	0.206	(0.007)	-0.008	(0.073)
Sale-Sale				
Intercept	0.471	(0.286)	0.223	(0.005)
ln(ownprice)	0.754	(0.107)	0.803	(0.24)
ln(compprice)	0.122	(0.077)	0.589	(0.07)
Comp Sale	0.066	(0.287)	0.201	(0.068)
Comp Sale*ln(compprice)	-0.013	(0.07)	0.111	(0.307)
σ^2	0.077	(0.006)	-0.024	(0.076)

Table 7: Summary of Price Indexes

Index	Mean	Min	Max	Std. Dev
Dynamic Model Index	82.34	45.81	116.92	16.35
Laspeyres	90.2	77.33	103.64	7.25
Fixed Base Geo.	89.7	76.94	103.7	7.3
Fixed Base Tornqvist	89.27	76.08	104.2	8.02
Chained Laspeyres	115.12	93.53	135.06	10.06
Chained Tornqvist	61.12	37.74	101.09	19.8
Feenstra and Shapiro (2003) Index	84.31	-	-	-

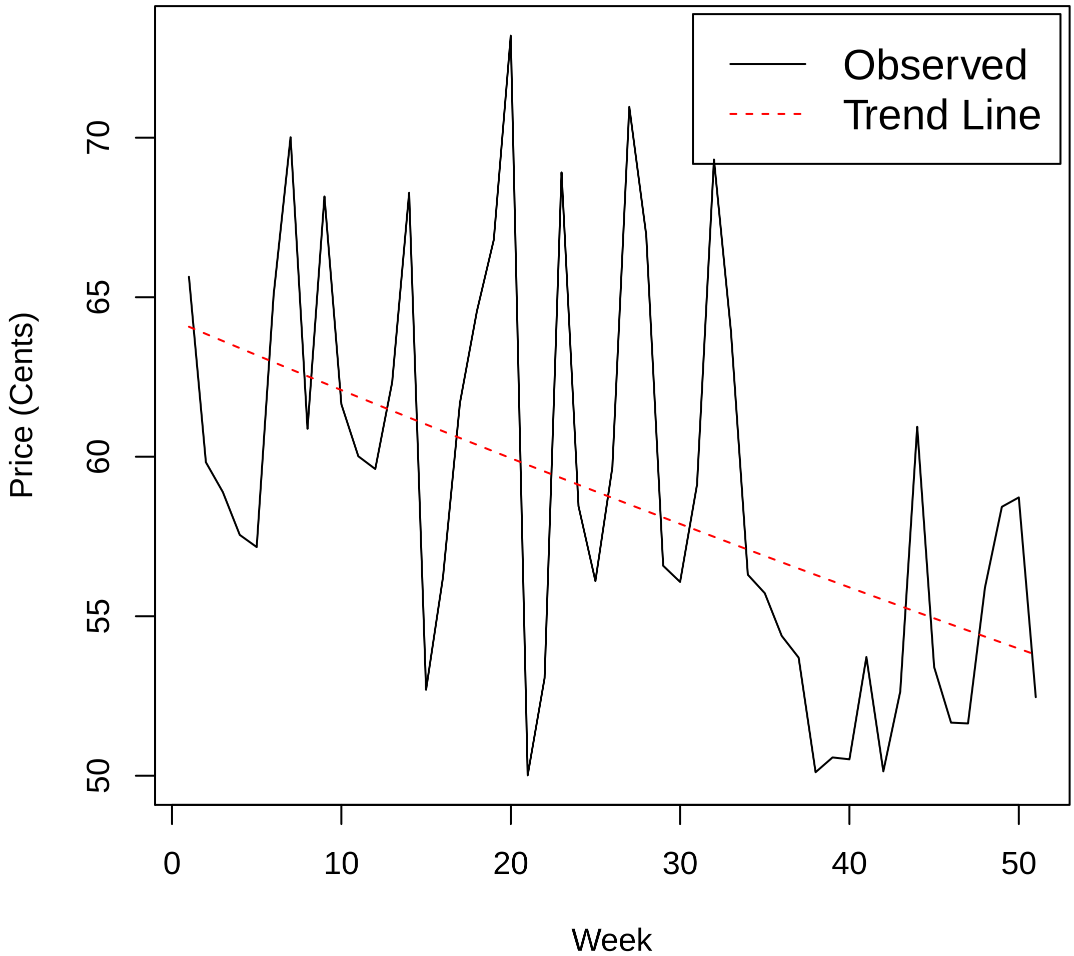

Figure 1: Average Prices Over Sample Period

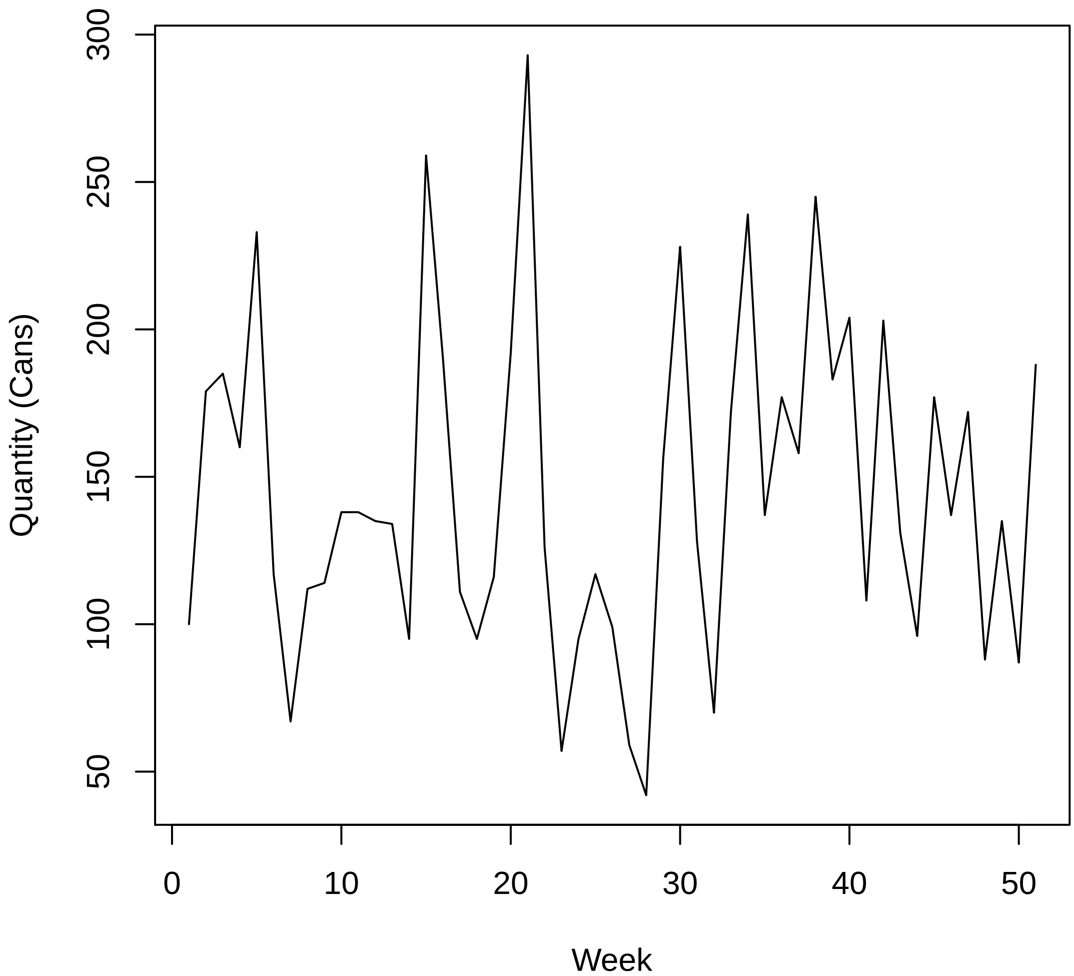

Figure 2: Quantity Sold Over Sample Period

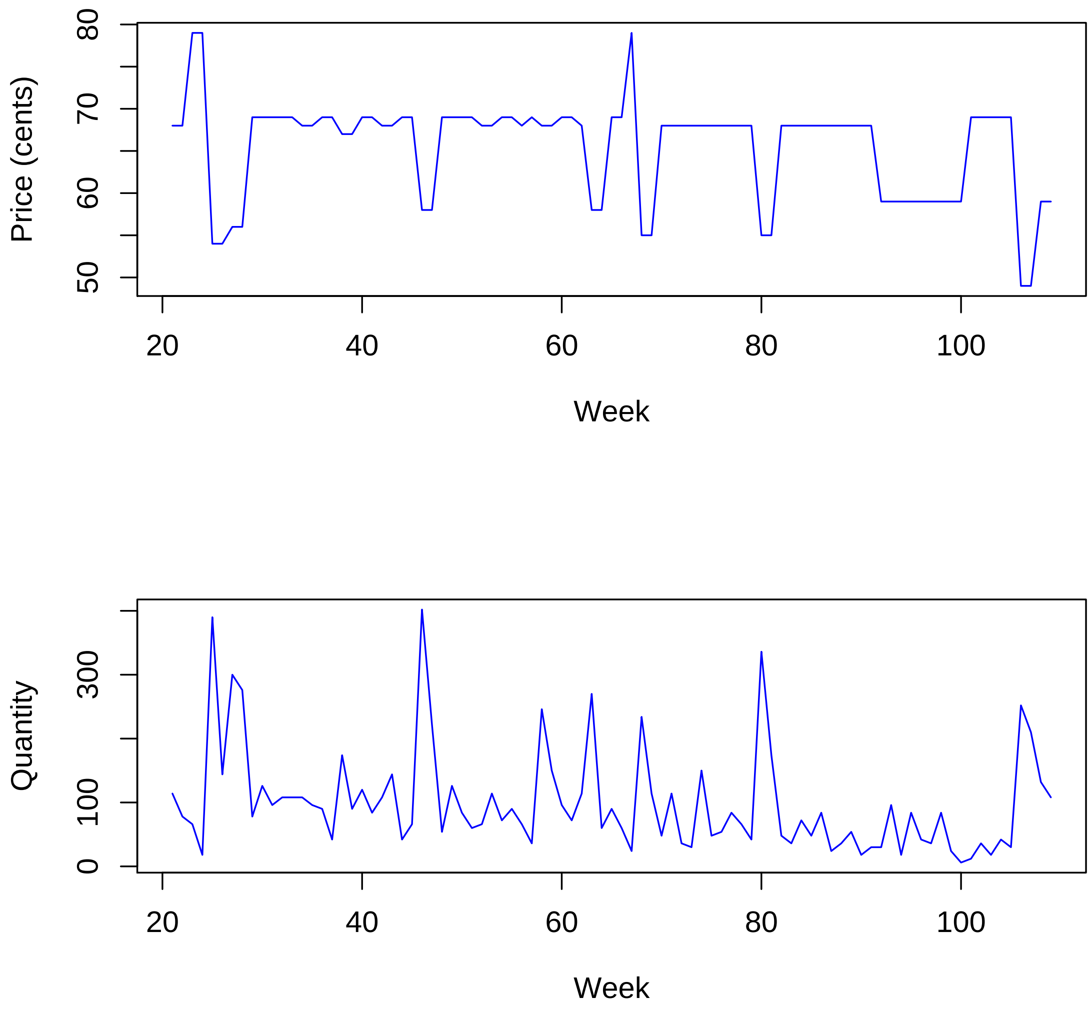

Figure 3: Prices and Quantities of Starkist for a Single Store

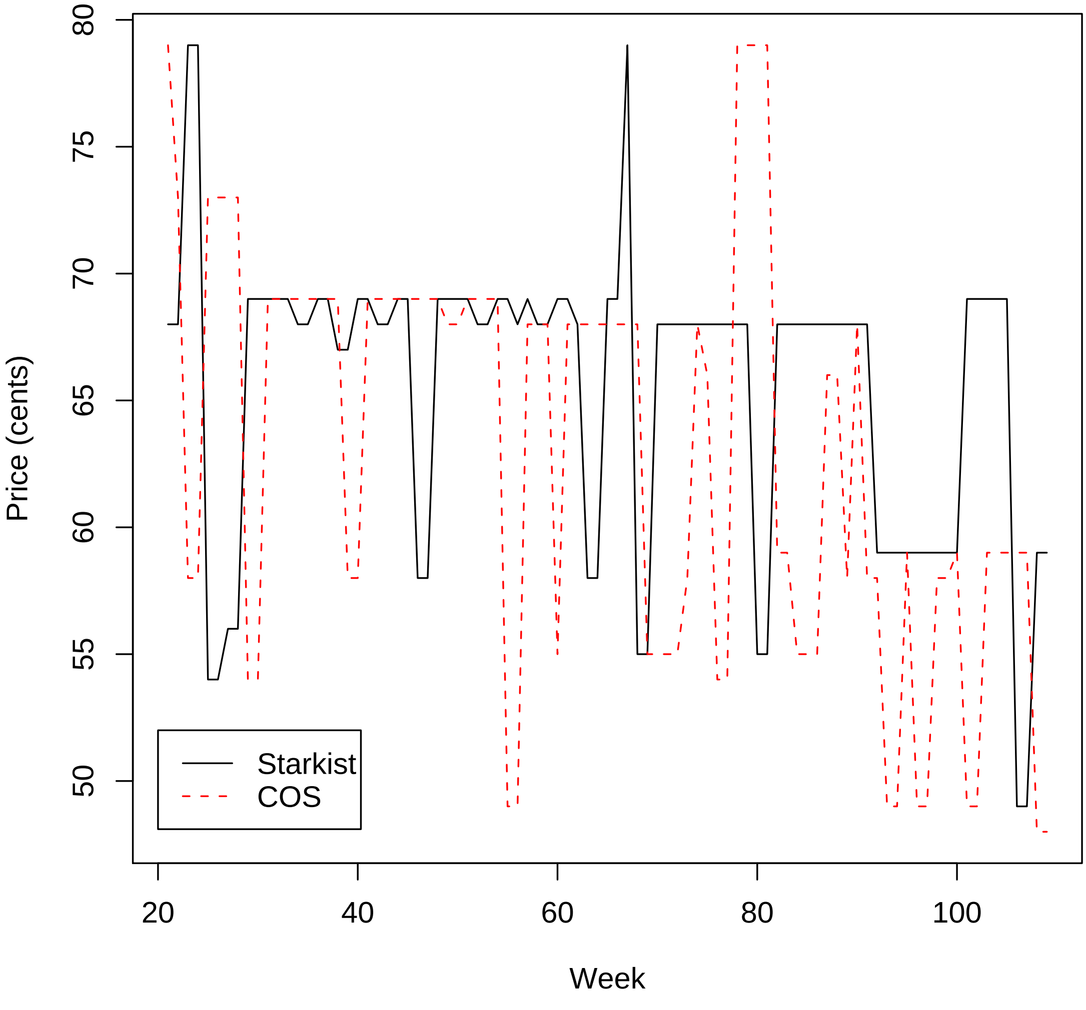

Figure 4: Prices of Starkist and Chicken of the Sea for a Single Store

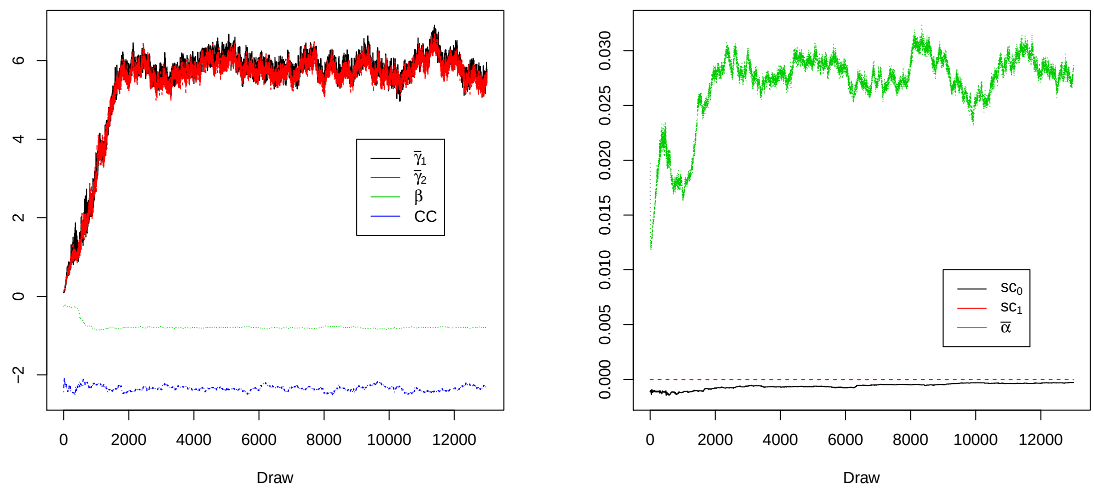

Figure 5: Plots Of MCMC Draws

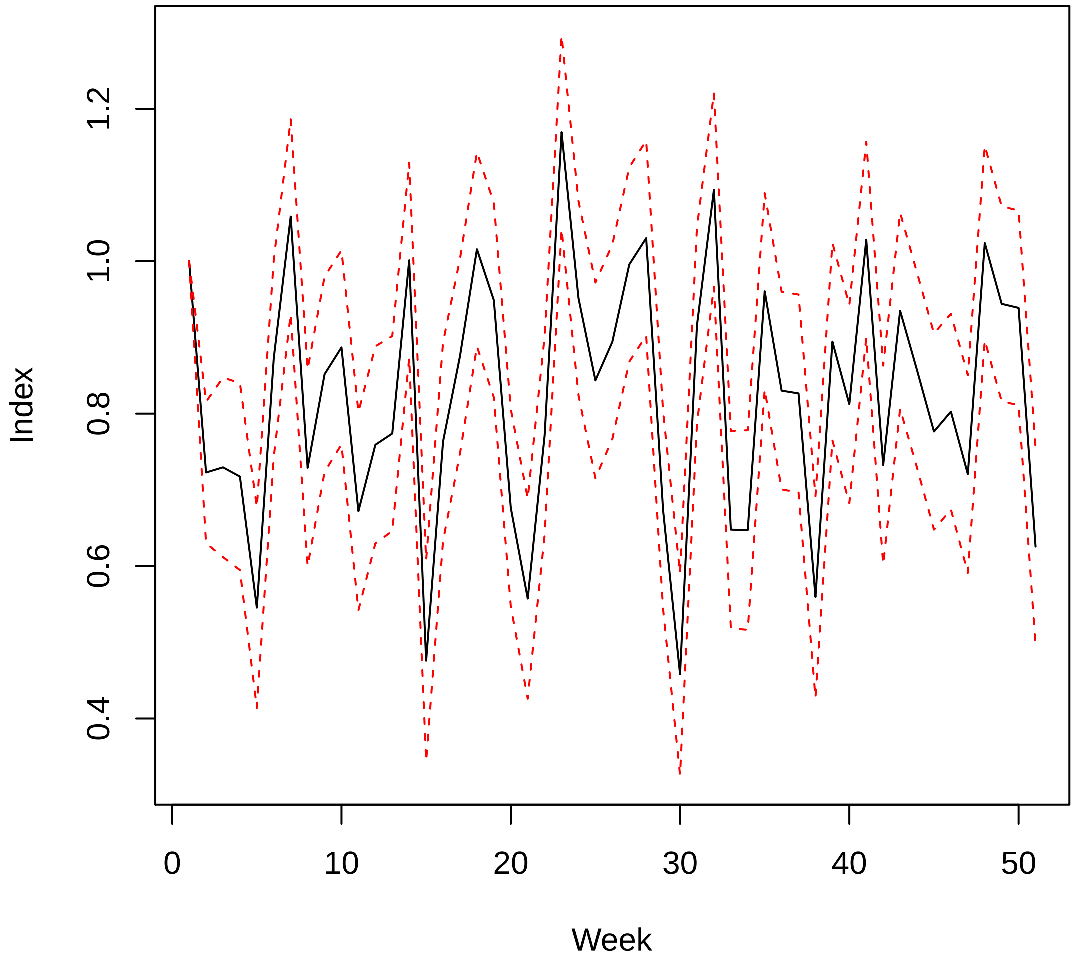

Figure 6: Dynamic Price Index for Canned Tuna