Should Consumer Subsidies Be More Flexible?

A Structural Analysis and Case Study of the Food Stamps Program*

Rudolf-Harri Oberg†

(with Andres Musalem‡)

Most recent version available at sites.duke.edu/rudolfoberg/research/

Abstract

In this paper, we study how consumers respond to subsidy designs with different spending restrictions, with implications to optimal design of consumer subsidies. In the context of SNAP (commonly known as the Food Stamps Program), we analyze the effect of food stamp benefits from the perspective of recipients (who care about overall consumer welfare) and the policymaker (who prefers that funds are used to buy food). Studying the effect of different subsidies in the context of SNAP is interesting since there are active proposals to change food stamp benefits in opposite ways: (i) to lift restrictions and provide cash with no strings attached, (ii) to add restrictions and exclude certain items (e.g., soda). To simulate consumer behavior under different subsidies, we develop a structural model of consumer demand which integrates consumer decisions for brands, categories, and stores. Our main finding is that expanding food stamp benefits to include household goods would be preferred by both benefit recipients and the policymaker. The mechanism driving this result is that flexible benefits give access to a wider selection of items which provides greater incentives to visit stores. In addition, we quantify trade-offs between different benefit designs and study the effect of banning benefit use on sweetened soda. Finally, our model of brand, category and store choice makes a technical contribution that could be interesting beyond the application considered in this paper.

*The authors thank Greg Allenby, Carl F. Mela, Bryan Bollinger, Daniel Xu and seminar participants at the 40th ISMS Marketing Science Conference and Duke University IO Lunch Group for comments and suggestions.
†Rudolf-Harri Oberg is a Ph.D. candidate in economics at Duke University (email: rудolf.harri.oberg@duke.edu).
‡Andres Musalem is an assistant professor at the Department of Industrial Engineering, University of Chile (email: amusalem@dii.uchile.cl).
1 Introduction

The Supplemental Nutrition Assistance Program (SNAP) – formerly and more familiarly known as the Food Stamps Program – provides monthly benefits for food spending. SNAP is the second-largest means-tested\(^1\) program in the United States after Medicaid, providing benefits to 40 million Americans in the fiscal year of 2018 (USDA (2018)). The program’s focus on subsidizing food expenditure, as opposed to supporting the purchase of household items or giving cash, reflects its goal to offer nutrition assistance to low-income individuals and families. At the same time, there is a growing debate over whether SNAP and other “in-kind”\(^2\) style social benefit programs should be more flexible, with some proposing replacing in-kind benefits with simple cash benefits.\(^3\) On the one hand, economic theory implies that giving cash over restrictive subsidies is always more effective from a welfare perspective since recipients are free to pick the bundle that maximizes their utility. On the other hand, the policymaker may prefer that recipients only use benefits on certain kinds of goods and as a result sets in place place different rules and restrictions on benefit spend in an attempt to achieve program goals. This difference in the perspective of recipients and the policymaker is important for evaluating SNAP and other in-kind transfer programs, as proposals to change program implementation may be derived from a singular perspective. As a result, there are proposals to change the current SNAP implementation in opposing ways, with some proposals calling for lifting restrictions on food stamp benefits, and other proposals calling for further restrictions on food stamp benefits by excluding certain items (e.g., soda).

In this paper, we analyze the effects of food stamp benefits relative to alternative (counterfactual) subsidy designs from the perspective of the policymaker and benefit recipients. We ask how key outcomes like consumer welfare or marginal propensity to consume food items (MPCF)\(^4\)

\(^1\)Means-tested benefits are available to individuals who can demonstrate that their means (income, savings, and other capital) are below a certain threshold.

\(^2\)In-kind transfer programs give benefits in the form of goods and services (e.g., SNAP, Medicare). This is opposed to in-cash transfer programs which give benefits in the form of a cash transfer (e.g., U.S. Social Security and unemployment benefits).


\(^4\)MPCF is a key indicator which reflects the main goal of the SNAP program and is also the main object of interest in many research papers. Therefore, we identify preferences of the policymaker to be in line with shifting benefit dollars toward food expenditure, i.e., maximizing MPCF out of provided benefits.
differ under (i) the usual food stamp restrictions (food only), (ii) grocery restrictions (food and household items only), and (iii) tax rebate (cash transfer). In order to simulate consumer behavior under different benefit systems, we develop a structural model of consumer demand which integrates consumer decisions for brands, categories, and stores. Using a structural model to account for direct and indirect effects of consumer behavior is crucial for understanding the true effect of different subsidies. For example, restricting consumer subsidies to food only, as opposed to including household items, may raise food expenditures conditional on store visit but may reduce the overall prospect of visiting a store, leaving it an open question which policy shifts more dollars towards food expenditure.

In addition to studying trade-offs between different subsidies, we also consider how excluding items from benefit use in one category affects purchases in other categories. The argument that some food categories – especially the category of sweetened beverages which includes sweetened soda – should be made ineligible for purchase with SNAP benefits has been gaining momentum. The idea to restrict spending on soda comes from the fact that more money is spent on soft drinks than any other item (USDA (2016)). Those who propose banning soda from SNAP argue that the ban could change purchase patterns and reduce soda consumption, whereas those who oppose the ban predict no change in consumption patterns since SNAP recipients could just buy the same amount of soda with their own cash. In addition, shopping frequency may change when subsidy use is restricted and it is an empirical question to what extent consumers substitute to other categories within a store and to what extent they substitute to the outside good (and decrease shopping frequency). We use our demand model to simulate a policy in which SNAP-eligible households face a restricted version of SNAP where spending on soda is banned. We establish that banning soda from SNAP affects the budget line in the same way as mandating a tax on soda which prompts us to compute the extent of soda taxation needed to reduce soda consumption to the level recorded under the ban.

We use household panel data from the IRI marketing data set (Bronnenberg et al. (2008)) to

---

5See for example Lane (2017), O’Connor (2017), Bittman (2012) for articles in the popular press.
estimate our structural model. The data contain information on weekly store trips and purchase decisions for a panel of over four thousand consumers, making it an ideal setting to study how consumer behavior changes in response to policy measures on all three levels (store, category, and brand choice). Our data are augmented with store locations and consumers’ pre-tax income and location.

We find evidence that expanding food stamp benefits to include grocery items (food and household goods) would be preferred by both benefit recipients and the policymaker. Our estimates imply that SNAP eligible consumers spend more money on food items using grocery benefits compared to regular food stamp benefits. Grocery benefits give access to a wider selection of items which provides greater incentives to visit stores. As a result, shopping frequency increases by 6.3% and serves as the main mechanism driving increased food spending (5.2%) relative to the regular food stamp subsidy. In addition, our simulations show that consumers would trade $100 in regular benefits for $95 in grocery benefits, which implies that an equal amount of grocery benefits would improve consumer welfare. Altogether, this finding provides a positive answer to the main question posed in the title of this paper and provides a concrete example of how benefits could be more flexible. This finding is interesting because it shows that a new design – which is not observed in practice – can improve both outcomes of interest without any compromise. Our simulations also confirm existing results about the trade-offs of regular food stamp policy and cash benefit, yielding further credibility to our main finding. Second, we find that consumers respond to a 10% soda tax and a soda ban in a similar way, reducing soda consumption by 23% and 20%, respectively. We discover that the ban policy has a differential impact on consumers’ substitution patterns and find evidence that differences in response are mainly generated by differences in preferences, as opposed to differences in travel costs.

Our results contribute to the literature on the effects of SNAP benefits on food spending. While most papers in the literature use reduced-form type of analysis to estimate key parameters like the MPCF, ours is the only paper (to the best of our knowledge) that takes a structural approach. Our structural model yields estimates that are in-line with existing estimates in the literature and has
the benefit of providing new (counterfactual) estimates for policies that are often part of the policy debate but are not observed in practice. Hoynes and Schanzenbach (2009), Castner and Mabli (2010), Beatty and Tuttle (2015) and Hastings and Shapiro (2018) all provide estimates for MPCF out of cash and SNAP and our estimates for these two quantities agree with what has been reported in these papers. Our results are also related to the literature on the effect of soda taxes (see e.g., Wang (2015), Bollinger and Sexton (2018) for overview). Most papers in this literature use event studies and/or structural approaches to predict the reduction in the consumption of sugar-sweetened beverages induced by such a tax. We approach the topic of soda taxation from the perspective of restricting food stamp benefits and calculate how the ban would be similar to a soda tax that is universally mandated or enacted only upon low income consumers. One caveat of our analysis is that our modeling framework does not account for the possibility that firms could reoptimize prices in response to new taxes or restricted benefits. However, recent research has demonstrated limited retailer price response to demand shocks and soda tax (Arcidiacono et al. (2018), Bollinger and Sexton (2018)). At the very least, our results provide a useful ballpark estimate of the effect of policy change.

Our modeling framework contributes to the literature on store choice and consumer demand. The main novelty of our model is integrating consumer demand for brands, categories and stores in a direct utility framework (see table 1). Few papers have modeled all three decisions and to our knowledge ours is the first to be based on a direct utility model of consumer behavior (Chandukala et al. (2008), Chintagunta and Nair (2011)). The primary benefit of considering all three levels of decision making in our application is that the policy of restricting or expanding food stamp benefits is often targeted at specific items and the effectiveness of the policy depends on changes in consumer demand. For example, if sugar-sweetened sodas are banned from food stamp benefits but consumers substitute to other goods that are deemed unhealthy (e.g., other sodas) then such a restriction to benefits is less effective in serving its purpose. Including store choice in the model is important for measuring how consumers substitute between the inside good (items sold at stores) and the outside good (expenditure outside of stores) in response to changes in their bud-
<table>
<thead>
<tr>
<th>Brand Category Store Direct Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bell and Lattin (1998)</td>
</tr>
<tr>
<td>Lee and Allenby (2009)</td>
</tr>
<tr>
<td>Kim (2017)</td>
</tr>
<tr>
<td>Thomassen et al. (2017)</td>
</tr>
<tr>
<td>Shriver and Bollinger (2017)</td>
</tr>
<tr>
<td>Briesch et al. (2009)</td>
</tr>
<tr>
<td>Schiraldi et al. (2012)</td>
</tr>
<tr>
<td>Pakes (2010)</td>
</tr>
<tr>
<td>Figurelli (2013)</td>
</tr>
<tr>
<td>Zhu et al. (2011)</td>
</tr>
<tr>
<td>Singh et al. (2006)</td>
</tr>
</tbody>
</table>

Table 1: Store Choice Literature

get constraint. In terms of joint consideration of how much demand is allocated across goods, categories, and stores, our work is closest to Bell et al. (1999). We build on this work by using a direct utility model where a single primitive (consumer consumption utility) drives all three decisions. Our direct utility model builds on the work of Lee and Allenby (2009) by nesting their model of brand-category choice into a model of store choice. Direct utility models have recently been applied to explain store and category choice in a variety of applications, e.g., Thomassen et al. (2017), Allcott et al. (2019), and Shriver and Bollinger (2017). We build on this research by adding more dimensions (brand choice) to the model, thereby enabling a richer set of substitution patterns, as noted above. Although not considered in the current application, the addition of brands is also useful for exploring the concept of loss leader pricing which rests on the premise that a given brand or category can drive demand for all goods within the store (Walters and MacKenzie (1988)).

To conclude, our model makes a technical contribution which is relevant beyond the application considered in this paper.

The rest of the paper is organized as follows. In section 2 we give an overview of the data. Section 3 develops the model and section 4 discusses estimation. In section 5, we perform counterfactual experiments and analyze results. We summarize our findings in section 6.
2 Data

We use household panel data from the IRI marketing data set (Bronnenberg et al. (2008)). The data consist of weekly store trips of individuals who are residents of either Eau Claire, Wisconsin, or Pittsfield, Massachusetts, and who participate in IRI’s BehaviorScan program. Our final data set follows the store, category and brand choice for 4002 consumers for 52 weeks. Weekly data include info on consumer spending, purchased quantities, store choice, distance to stores, product prices and characteristics. We focus on 10 most purchased categories with 3–6 most purchased brands in each category. We will now give details on how we process the data.

2.1 Data Set Construction

There are a number of common tasks that need to be completed to make the data usable for choice modeling. We outline our choices below. In most cases, we closely follow the methods used by Gordon et al. (2013).

The IRI data set tracks 30 product categories but we focus on the same 19 product categories that are used by Gordon et al. (2013). We rank the categories in terms of sales volume and pick the top ten categories for analysis. The remaining 9 categories are grouped into a composite category where the weight of each category is given by its sales volume. The top ten categories in terms of sales volume are carbonated beverages, frozen pizza, yogurt, frozen dinner, potato chips, toilet tissue, coffee, laundry detergent, paper towel, tortilla chips. The composite category consists of hot-dog, spaghetti sauce, margarine and butter, mayonnaise, peanut butter, shampoo, ketchup, deodorant, and mustard categories.

To make the consumer’s choice problem tractable, we aggregate UPC’s in each category into a set of brands. First, UPC’s that serve a specific market niche, have very low sales or are otherwise irrelevant to the analysis are removed. Among the UPC’s that remain, only most popular package types and sizes are used in the analysis. For example, only UPC’s of liquid detergents and their most popular package sizes are included in the category of laundry detergent since they make up...
more than 95 percent of the category sales. The same principle of UPC selection is applied in each category which on average causes a 10 percent reduction in the number of UPC’s over different categories. Second, the selected UPC’s are aggregated into brands following the decision rules discussed in Gordon et al. (2013). Since private labels cannot be identified with a specific retail chain from the data, all private labels are grouped under the same brand regardless of the chain they belong to. The resulting brands are sorted by market share and those brands that yield a cumulative market share of at least 80 percent or have an individual market share of at least 4 percent are included in the analysis. The remaining brands are grouped into a single composite brand with an average market share of 18.3 percent. Next, prices are aggregated to the brand level as an weighted average of UPC prices. The weight of an UPC is equal to its share of sales volume in the brand it belongs to where the sales volume is calculated at the store level in the given year.

In selecting the sample of households, we include all households who have visited at least two stores and for whom location and income data exists. This rule selects 4210 panelists from an initial sample of 11322. We focus on each consumer’s two stores with the highest expenditure shares. First, shopping outside the top two stores is minimal (see table 3). Second, the computational cost of accommodating additional stores are severe which is why the literature on multi-store multi-category choice has typically focused on the case of two stores.

We approximate each consumer’s weekly budget by dividing their pre-tax income by 52. If weekly spending is greater than the approximated budget, we increase the weekly spending by 5 percent and take this to be the new budget for that week. To calculate the distance to visited stores, we use latitude-longitude information provided for each consumer’s residence. Since information on consumer’s exact location is sensitive, the data has been disguised so that the true location lies within 225 meters from the provided data. We calculate distance to visited stores by executing a Python script which queries Google for driving distance. To make the sample more compact, we get rid of outliers and extreme values. For each consumer’s two most visited stores, one store is closer and the other store is further away. We drop 2.5 percent of consumer who live very close to their closer

---

[6] We approximate each consumer’s weekly budget by dividing their pre-tax income by 52. If weekly spending is greater than the approximated budget, we increase the weekly spending by 5 percent and take this to be the new budget for that week. To calculate the distance to visited stores, we use latitude-longitude information provided for each consumer’s residence. Since information on consumer’s exact location is sensitive, the data has been disguised so that the true location lies within 225 meters from the provided data. We calculate distance to visited stores by executing a Python script which queries Google for driving distance. To make the sample more compact, we get rid of outliers and extreme values. For each consumer’s two most visited stores, one store is closer and the other store is further away. We drop 2.5 percent of consumer who live very close to their closer

---

[6] With $K$ stores the state-space has $2^K$ elements, with the same number of optimization exercises for each consumer.
store and 2.5 percent of consumer who live very far from their other store. This rule ensures that each panelists distance to their two most visited stores is at least 0.4 miles and not more than 14 miles and brings the sample to 4002 panelists.

Finally, we use federal guidelines to simulate whether consumers in our sample are eligible for SNAP benefits. Using criteria from the 2018 fiscal year, we say that a consumer is SNAP eligible when his yearly reported income falls below the eligibility threshold.\(^7\) We find that around 24 percent of our sample is SNAP eligible using these guidelines. Since direct SNAP participation data is notoriously hard to get, most research relies on indirect measures of participation.\(^8\) For example, Hastings and Shapiro (2018) use method of payment to infer SNAP participation. Therefore, we believe that simulating SNAP participation in our sample using federal guidelines is sufficient for analyzing how target consumers respond to different subsidies.

2.2 Descriptive Statistics

Table 2 presents descriptive demographic and income data for the estimation sample. The high share of outside good is worthy of note as it illustrates the magnitude of economic choices relative to the reported income that we are able observe in the data. In particular, when consumers do not visit any stores in a given week, then by construction they spend all their income on the outside good. Even when consumers visit a store in a given week, observed spending is relatively small compared to unobserved spending. One concern is that we may be missing important expenditure information by only focusing on two stores. Table 3 shows that this concern is minimal and limiting our attention to two stores captures 95 percent of expenditure among all shopping outcomes. Consumer spending is in line with visit patterns and on average around two thirds of all store trips are done at store 1 (the top store in terms of total consumer expenditure). In sum, this shows that

---

\(^7\)The threshold is 130 percent of the federal poverty line. In federal fiscal year of 2018, the federal poverty line was $1702 for a household of three, making the eligibility threshold $2213 per month or about $27k per year. (USDA 2018)

\(^8\)Meyer and Goerge (2010) and Meyer and Mittag (2015) show that common surveys underreport participation by 30 percent or more. In addition, SNAP eligibility is a complex function of many types of income that cannot be measured precisely in any existing data set.
### Table 2: Summary of consumer data.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>P0.1</th>
<th>P0.25</th>
<th>P0.5</th>
<th>P0.75</th>
<th>P0.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income ($/week)</td>
<td>996</td>
<td>528</td>
<td>433</td>
<td>577</td>
<td>962</td>
<td>1346</td>
<td>1682</td>
</tr>
<tr>
<td>Distance to the closer store (mi)</td>
<td>2.7</td>
<td>1.7</td>
<td>1.1</td>
<td>1.6</td>
<td>2.3</td>
<td>3.4</td>
<td>4.7</td>
</tr>
<tr>
<td>Distance to the store further away (mi)</td>
<td>5.2</td>
<td>2.4</td>
<td>2.5</td>
<td>3.6</td>
<td>4.7</td>
<td>6.4</td>
<td>8.6</td>
</tr>
<tr>
<td>Consumer’s weekly share of outside good (cond. on visiting stores in the set {1, 2}) (%)</td>
<td>98.05</td>
<td>2.3</td>
<td>96.3</td>
<td>97.9</td>
<td>98.7</td>
<td>99.1</td>
<td>99.4</td>
</tr>
</tbody>
</table>

### Table 3: Use of stores.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>St Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prob. of store visit in any given week</td>
<td>0.58</td>
<td>0.19</td>
</tr>
<tr>
<td>Prob. of visiting store 1 (cond. on visiting stores in the set {1, 2})</td>
<td>0.68</td>
<td>0.22</td>
</tr>
<tr>
<td>Prob. of visiting store 2 (cond. on visiting stores in the set {1, 2})</td>
<td>0.21</td>
<td>0.17</td>
</tr>
<tr>
<td>Prob. of visiting store 1, 2 (cond. on visiting stores in the set {1, 2})</td>
<td>0.11</td>
<td>0.11</td>
</tr>
<tr>
<td>Expenditure share in store 1 by weekly spending (across all stores)</td>
<td>0.69</td>
<td>0.17</td>
</tr>
<tr>
<td>Expenditure share in store 2 by weekly spending (across all stores)</td>
<td>0.26</td>
<td>0.15</td>
</tr>
</tbody>
</table>

most consumers in our sample use one main store for most of their purchases and supplement their purchases by occasional visits to another store; outside of these two stores, consumer spending is minimal. Table 4 gives an overview of consumer spending across all categories that form the inside good in our analysis. All categories in the sample are actively purchased as they are top categories by construction. The most popular category is the carbonated beverages category. This observation is consistent with spending patterns reported by the USDA (USDA (2016)). Another observation consistent with the latter report is that there are only small differences in spending habits across income, and therefore we only report spending habits across all consumers. The extent of multi-brand purchasing is low, meaning that consumers tend to pick only one brand when deciding to purchase in a category – this will guide our modeling choices which we outline in the next section.

### 3 Model

In this section, we develop our demand model for brand, category, and store choice. We first outline the challenges in modeling brand-category choice and show our specification for a brand-category
model which is similar to the model specified by Lee and Allenby (2009). We then show how to nest our brand-category model into a store choice framework. Second, we develop a statistical specification of our proposed model and derive the likelihood function.

### 3.1 Direct Utility Model

There are a number of challenges in modeling brand-category demand. First, multiple discreteness, or the purchase of multiple units of multiple alternatives, is a common characteristic of market basket data. Consumers make purchases in multiple categories which means that the utility specification must be able to accommodate corner and interior solutions. Selection of multiple alternatives is also present in store choice data as consumers may visit several stores. Second, the decision to buy from a category may be related to the purchase decision in other categories which means that the model should be able to capture potential cross-category dependencies. For categories that are complements, a joint purchase gives higher utility than the combined utility from separate purchases and for categories that are substitutes the opposite holds true. Third, since we
aim to model brand-level purchase quantities rather than category-level purchase incidence, we
must adopt a specification that keeps the number of parameters at a reasonable level.

3.1.1 Model of Bundle Choice

We define utility functions that are relevant for bundle choice and set up the consumer problem.

We model consumer $i$’s utility for bundle $q = (q_{\text{in}}, q_{\text{out}})$ as

$$u(q) = \Psi^T q_{\text{in}} - H^T B H + \frac{q_{\text{out}}}{\beta_0}$$

where $q_{\text{in}}$ is the vector of inside (store) goods, $q_{\text{out}}$ is the quantity of the outside good, and $H = \left(u_c\right)_{c=1}^C$ is the vector of category utilities. Utility of category $c \in \{1, \ldots, C\}$ is denoted by $u_c$ and is computed as the sum of category-specific brand-level marginal utilities:

$$u_c = \sum_{k \in \mathcal{I}(c)} \psi_k q_{\text{in},k},$$

where $\psi_k$ is the marginal utility of brand $k$, $q_{\text{in},k}$ is the quantity of brand $k$, and $\mathcal{I}(c)$ denotes the index set of brands that belong to category $c$. The consumer problem for bundle choice is

$$\max_{q} u(q) \quad \text{s.t.} \quad p^T q_{\text{in}} + q_{\text{out}} = E,$$

where $p$ is the price vector of brands and $E$ is the consumer’s budget constraint. Notice that the price of the outside good is normalized to one.

The model specification for bundle utility in equation 1 is similar to the model in Lee and Allenby (2009) to the extent that both models feature a “linear part” which is an additive function of all the brand preferences, and a “quadratic part” which is an interaction term between category-specific sub-utilities. We specify a separate model for the outside good, whereas Lee and Allenby (2009) treat the outside good as one of the categories. Finally, our statistical specification for the model uses the same strategy but different shock distributions, see section 3.2 for details.
Key parameters in the utility function defined in equation 1 are $\Psi$ and $B$. The parameter vector $\Psi$ controls the linear part of the utility function and the matrix $B$ controls the quadratic part of the utility function. These parameters can be interpreted in greater detail by looking at the marginal utility with respect to one of the brands:

$$MU_{cb} := \frac{\partial u(q)}{\partial q_{cb}} = \Psi_{cb} \left( 1 - 2 \cdot \left( \beta_{C_c,C_k} u_{C_c} + \sum_{k \neq c}^{C} \beta_{C_c,C_k} u_{C_k} \right) \right).$$ (4)

In equation 4, we see that the marginal utility of brand $b$ in category $c$ depends on the sub-utilities of all other categories. When categories $c$ and $k$ are complements, then $\beta_{C_c,C_k} < 0$, and purchases in category $k$ increase the marginal utility, and when the categories are substitutes, then $\beta_{C_c,C_k} > 0$, and purchases in category $k$ decrease the marginal utility. The off-diagonal terms of $B$, therefore, capture cross-category relationships, while the diagonal terms capture each category’s degree of satiation. The parameter $\Psi_{cb}$ can be thought of as the baseline marginal utility for brand $b$ in category $c$ which is then either shifted up or down depending on other purchases the consumer makes.

Finally, note that the model implies choice of one brand when the consumer decides to purchase from a category. Equation 4 implies that differences between any two alternatives only depend on the baseline marginal utility parameter, the $\Psi_{cb}$ term, and so only one brand will be chosen. This choice behavior is consistent with choice behavior observed in the dataset, in section 2.2 we noted that multi-brand purchasing is fairly low across all categories.
A Simple Example

To illustrate the utility function related to bundle choice with a simple example, suppose that the number of stores, categories, and brands is two. The utility function is then

\[
\begin{align*}
    u(q_{\text{lin}}, q_0) &= \begin{pmatrix} 
    \Psi_{s_1c_1b_1} \\
    \Psi_{s_1c_1b_2} \\
    \Psi_{s_1c_2b_1} \\
    \Psi_{s_1c_2b_2} \\
    \Psi_{s_2c_1b_1} \\
    \Psi_{s_2c_1b_2} \\
    \Psi_{s_2c_2b_1} \\
    \Psi_{s_2c_2b_2}
\end{pmatrix}^T 
    \begin{pmatrix} 
    q_{s_1c_1b_1} \\
    q_{s_1c_1b_2} \\
    q_{s_1c_2b_1} \\
    q_{s_1c_2b_2} \\
    q_{s_2c_1b_1} \\
    q_{s_2c_1b_2} \\
    q_{s_2c_2b_1} \\
    q_{s_2c_2b_2}
\end{pmatrix} 
    - \begin{pmatrix} 
    u_{c_1} \\
    u_{c_2} \\
    \beta_{c_1,c_1} u_{c_1} + \beta_{c_1,c_2} u_{c_2} \\
    \beta_{c_1,c_2} u_{c_1} + \beta_{c_2,c_2} u_{c_2} \\
    \beta_{c_2,c_1} u_{c_1} + \beta_{c_2,c_2} u_{c_2} \\
    \beta_{c_2,c_2} u_{c_1} + \beta_{c_2,c_2} u_{c_2} \\
    \beta_{c_2,c_1} u_{c_1} + \beta_{c_2,c_2} u_{c_2} \\
    \beta_{c_2,c_2} u_{c_1} + \beta_{c_2,c_2} u_{c_2}
\end{pmatrix} 
    + \frac{q_0}{\beta_0}\end{align*}
\]

which shows that the marginal utility is non-constant and allows for corner and interior solutions. Figure 1 plots the indifference curves and shows that the curves are convex to the origin, allowing for an interior solution, and the curves also intersect the two axes, allowing for a corner solution.

3.1.2 Model of Store Choice

We model consumers’ store visit utility as a function of store-specific bundle utility and consumer-store distance and fixed effects. Let \( q^*_k \) denote the solution to the consumer problem defined in
Figure 1: Indifference curves for the case of two one-brand categories (two products) and one store. Parameters are $\Psi_{s_1c_1b_1} = \Psi_{s_1c_2b_1} = 0.5$, $\beta_{c_1c_1} = \beta_{c_2c_2} = 0.25$, $\beta_{c_1c_2} = -3$.

Equation 3 when bundle choice is restricted to store $s$. Utility for store $s$ is given by

$$U(s) = u(q^*_s) - \tilde{R}E_s - \rho \cdot \text{dist}_s + \eta_s,$$

where $u(q^*_s)$ is the bundle utility, $\tilde{R}E_s$ is the store fixed effect, $\rho \cdot \text{dist}_s$ is travel cost to store $s$, and $\eta_s$ is the store error term that is known to the consumer but unknown to the econometrician. The store choice problem is given by

$$\max_s U(s).$$

In our empirical specification, we use data on each consumer’s two most visited stores. We model the options of no visit and joint visit separately which means the model described in equation 7 is a four-alternative model where $s \in \{s_0, s_1, s_2, s_{12}\}$ denotes either the option of no store visit, first store visit, second store visit, or visiting both stores.
3.2 Statistical Specification

We use the following decomposition of joint probability of bundle and store choice as the basis for forming the likelihood function later on:

\[ P(q, s) = P(q|s)P(s). \]  \hspace{1cm} (9)

The decomposition in equation 9 outlines our strategy to compute the likelihood function in two parts. First, to calculate \( P(q|s) \), we introduce stochastic shocks into the marginal utility specification and use Kuhn-Tucker first order conditions for constrained utility maximization to form the likelihood for the observed bundle. Second, to calculate \( P(s) \), we define an error distribution of stochastic store shocks.

3.2.1 Computing likelihood of bundle choice, \( P(q|s) \)

The likelihood of observed bundle choice is formed based on Kuhn-Tucker or KT (1951) first order conditions for constrained utility maximization. The KT approach assumes that the utility function is random (from the analyst’s perspective), implying that the optimal consumption vector is random too. Under this approach, probabilities for corner and interior solutions can be derived by solving the constrained utility maximization problem using KT conditions.

Following a common practice in the literature, we introduce the stochastic element into the marginal utility specification:

\[ MU_{scb, \varepsilon} = MU_{scb} \exp(\varepsilon_{scb}), \quad \varepsilon_{scb} \sim iid \text{ EVT1}(0, \sigma_{\varepsilon}) \quad \forall s \in \{1, \ldots, S\}, \forall c \in \{1, \ldots, C\}, \forall b \in \{1, \ldots, B\}. \]  \hspace{1cm} (10)

The error term captures unobserved (from the analyst’s perspective) information that impacts the marginal utility for each good. Our statistical specification for product shocks relies on the extreme value distribution with unknown scale parameter, whereas Lee and Allenby (2009) use a fixed multivariate normal distribution with an identity matrix. The error term is exponentiated to guarantee...
positivity of the marginal utility. The underlying utility function changes in the following way:

\[ u(q) = (\Psi * \exp(\varepsilon))^T q_{in} - H^T B H + \frac{q_{out}^\beta_0 \exp(\varepsilon_{out})}{\beta_0}, \] 

(11)

where the operation \( * \) denotes element by element multiplication between two vectors. Equation 11 shows that the stochastic element affects the vector of baseline marginal utility, \( \Psi \); error terms do not affect the vector of category utility, \( H \), and thus there are no squared error terms. Taking derivative of equation 11 gives back the expression of marginal utility in equation 10. In addition, to guarantee that \( \Psi_{scb} > 0 \), we parameterize the non-stochastic part of baseline marginal utility as

\[ \Psi_{scb} = \exp(\psi^T x_{scb}) \] 

(12)

where \( x_{scb} \) is a set of attributes that characterize the alternative and \( \psi \) is the corresponding preference parameter which describes consumer preferences for these characteristics. In the empirical specification, brand characteristics are brand-level dummy variables and a measure of feature and display at the week level.

**Optimal Quantity Allocation**

Observed purchase quantities are associated with KT conditions that are derived from the following Lagrangian:

\[ L = (\Psi * \exp(\varepsilon))^T q_{in} - H^T B H + \frac{q_{out}^\beta_0 \exp(\varepsilon_{out})}{\beta_0} + \lambda (E - p^T q_{in} - q_{out}), \] 

(13)
where $\lambda$ is the Lagrangian multiplier associated with the budget constraint. The KT first-order conditions for optimal quantity allocation are given by

$$\frac{\partial L}{\partial q_{scb}} = \Psi_{scb} \left( \exp(\epsilon_{scb}) - 2 \cdot \left( \beta_{c_c, c_c} u_{c_c} + \sum_{k \neq c} \beta_{c_c, c_k} u_{c_k} \right) \right) - \lambda p_{scb} = 0, \text{ if } q_{scb}^* > 0 \quad (14)$$

$$\frac{\partial L}{\partial q_{scb}} = \Psi_{scb} \left( \exp(\epsilon_{scb}) - 2 \cdot \left( \beta_{c_c, c_c} u_{c_c} + \sum_{k \neq c} \beta_{c_c, c_k} u_{c_k} \right) \right) - \lambda p_{scb} < 0, \text{ if } q_{scb}^* = 0 \quad (15)$$

$$\frac{\partial L}{\partial q_{out}} = q_{out}^{\beta_0-1} \exp(\epsilon_{out}) - \lambda = 0, \text{ since } q_{out} > 0 \text{ by assumption.} \quad (16)$$

The optimal quantity allocation $q^* = (q_{in}^*, q_{out}^*)$ satisfies KT first-order conditions in equations 14–16 and the budget constraint $p^T q_{in} + q_{out} = E$. We assume that the outside good is always consumed, which makes it possible to substitute $\lambda = q_{out}^{\beta_0-1} \exp(\epsilon_{out})$ into the other first-order conditions and rewrite the KT first-order conditions as follows:

$$\epsilon_{scb} = \ln \left( 2 \cdot \left( \beta_{c_c, c_c} u_{c_c} + \sum_{k \neq c} \beta_{c_c, c_k} u_{c_k} \right) + \exp(\epsilon_{out}) \frac{p_{scb}}{\Psi_{scb}} q_{out}^{\beta_0-1} \right) =: V_{scb}, \text{ if } q_{scb}^* > 0 \quad (17)$$

$$\epsilon_{scb} < \ln \left( 2 \cdot \left( \beta_{c_c, c_c} u_{c_c} + \sum_{k \neq c} \beta_{c_c, c_k} u_{c_k} \right) + \exp(\epsilon_{out}) \frac{p_{scb}}{\Psi_{scb}} q_{out}^{\beta_0-1} \right) =: V_{scb}, \text{ if } q_{scb}^* = 0 \quad (18)$$

### Likelihood Formation

The likelihood for a vector of observed demand $q^* = (q_{in}^*, q_{out}^*)$ can be formed by using the distributional assumptions of $\epsilon$-shocks and the change-of-variables technique. In equation 10, we specified the shocks to follow extreme value type-1 distribution with location 0 and scale $\sigma_\epsilon$. Without loss of generality, assume that the consumer purchased the first $K$ of $M$ goods. The likelihood for the observed bundle is then given by

$$P(q_1, \ldots, q_K, 0, \ldots, 0|\epsilon_{out}) = \prod_{j=1}^{K} \frac{1}{\sigma_\epsilon} g \left( \frac{V_j}{\sigma_\epsilon} \right) |J| \cdot \prod_{j=K+1}^{M} G \left( \frac{V_j}{\sigma_\epsilon} \right), \quad (19)$$

where $g(\cdot)$ and $G(\cdot)$ are the pdf and cdf of the standard extreme value type-1 distribution. We can see that the likelihood is composed of a density component for all positive goods and a probability
mass component for all items that are not purchased. The Jacobian is generated from the change-
of-variables technique and is a $K \times K$ matrix with elements

$$J_{scb,s'c'b'} = \frac{\partial \epsilon_{scb}}{\partial q_{s'c'b'}} = \frac{2\beta c_{c,c}' \Psi_{s'c'b'} - \exp(\epsilon_{out}) q_{out}^{-2} \tilde{p}_{scb} (\beta_0 - 1) p_{s'c'b'}}{\exp(V_{scb})}. \quad (20)$$

In order to form the likelihood function, we also have constrain the support of the parameters. We want the utility function to be quasi-concave and all marginal utilities to be positive. The likelihood is given by

$$\mathcal{L}(\Theta | q_1, \ldots, q_K, 0, \ldots, 0) = \int P(q_1, \ldots, q_K, 0, \ldots, 0 | \epsilon_{out}) d\epsilon_{out} \quad (21)$$

$$\times 1 \left( \frac{\partial u}{\partial q_{scb}} > 0 \text{ for all } s, c, b \right)$$

$$\times 1 (u(q_{in}, q_{out}) \text{ quasiconcave on the budget constraint})$$

Quasi-concavity of the utility function ensures sufficiency of the KT first-order conditions, positivity of marginal utility is needed to ensure a proper utility function.

### 3.2.2 Computing likelihood of store choice, $P(s)$

To compute likelihood of store choice, we need to specify a distribution for store-specific error terms, $\eta_s$. We use multi-variate normal distribution for error shocks which means that our statistical specification for store choice is the probit model:

$$U(s) = u(q^*_s) - \tilde{R}E_s - \rho \text{dist}_s + \eta_s \quad (22)$$

$$\eta \sim MVN(\mathbf{0}, \tilde{\Sigma}) \quad (23)$$

where $\tilde{\Sigma}$ is the full covariance matrix of store-specific error terms. We use two stores in our empirical specification and model the options of visiting both stores and visiting no stores as separate alternatives. This means the store indicator $s \in \{s_0, s_1, s_2, s_{12}\}$ can take on four values. We will
discuss normalization for level and scale under identification in the next section.

4 Estimation

4.1 Identification

We will now discuss identification of model parameters. First, in the model for bundle choice, it is necessary to normalize preference parameters for scale. Either one of the psi parameters (describing baseline marginal utility) or one of the beta parameters (describing satiation) needs to be normalized at one for each consumer and each category. We follow the convention used by Lee and Allenby (2009) and set the diagonal terms of the satiation matrix $B$ equal to one which corresponds to normalizing the within-category satiation term for each consumer. Second, in the model for store choice, it is necessary to normalize the model for both level and scale. We use the typical normalization scheme for probit discrete choice models (e.g., Train (2009)). We normalize the four-alternative model described in equation 7 for utility levels by taking utility differences with respect to the no visit alternative, and for scale by setting the top-left element of the covariance matrix of error differences to one. Finally, we cannot separately identify individual-specific store fixed-effects and the effect of travel costs as these quantities are constant over time for any consumer-store pair. We define a generalized individual-specific fixed effect as $RE_s = \bar{R}E_s + \rho dist_s$ which is identified for $s \in \{s_1, s_2, s_{12}\}$ and let the no visit alternative serve as the baseline. To study the effect of travel costs, we regress estimated fixed effects against store distance.

4.2 Heterogeneity

We allow for heterogeneity in shopping behavior by assuming that individual parameters come from population-level distributions with hyperparameters. In the model for bundle choice, each household is characterized by their vector of baseline marginal utilities $\psi^h$, matrix of within-category and cross-category satiation parameters $B^h$, and satiation of outside good denoted pa-
parameter $\beta_0$. In the model for store choice, each household is characterized by a set of generalized store fixed effects, $RE^h_s$. To sum up, we model household specific parameters with the following population-level distributions:

$$\psi^h_i \sim \text{log-normal} \left( \bar{\psi}_i, \bar{\sigma}_{\psi_i} \right) \quad i = 1, \ldots, K$$

(24)

$$\beta^h_{C_i,C_j} \sim N \left( \tilde{\beta}_{ij}, \tilde{\sigma}_{\beta_{ij}} \right) \quad i, j = 1, \ldots, C$$

(25)

$$\frac{\beta^h_0}{1 - \beta^h_0} \sim \text{log-normal} \left( \tilde{\beta}_0, \tilde{\sigma}_{\beta_0} \right)$$

(26)

$$RE^h_s \sim N \left( \bar{RE}_s, \bar{\sigma}_{RE_s} \right) \quad s \in \{s_1, s_2, s_{12}\}$$

(27)

This set-up makes the model hierarchical with population-level hyperparameters at the top level, individual parameters at the next level, and choice data at the lowest level. Markov Chain Monte Carlo estimation is used to estimate all model parameters. We use various methods, including adaptive proposal rules, and give the details of the algorithm in the appendix.

### 4.3 Estimates and Model Fit

We will now discuss parameter estimates and model fit. Figures 2 and 3 show estimates of hyperparameters of baseline marginal utility coefficients (the $\psi$ - coefficients in equations 4 and 24). We focus on the hyperparameters of log-normally distributed individual coefficients since it is easier to spot a trend in the ranking of brands using hyperparameters, rather than individual coefficients which can be close to zero for many brands. Each category has 3–6 brands where the first brands represent top brands in the category and the last brand represents the composite brand. Figure 2 shows that the mean parameters of brand baseline marginal utility generally follow the pattern predicted by brand market shares, with top brands being associated with higher mean parameters. The ranking of the composite brand depends on the extent of market concentration in the category and may be among the lowest or highest ranking brands. Figure 3 illustrates spread among the individual coefficients of baseline marginal utility. The general pattern in figure 3 shows that higher ranking brands are associated with less spread which is an intuitive pattern since more consumers
have to value a brand relatively highly for it to be a top brand. We use the concept of Bayesian credible interval – defined as the 2.5th and 97.5th percentile of the posterior distribution – to illustrate precision of estimates. We see that most coefficients are precisely estimated, the credible interval associated with the composite category’s brand appears larger because it is the only brand in that category, in addition, the underlying coefficients of baseline marginal utility are extremely close to zero, resulting in a wide range of parameters to fit the data well. Next, figure 4 shows estimates of cross-category interactions (the $\beta$ parameters in matrix $B$ in equations 4 and 25). We focus on individual coefficients since they are directly interpretable and consumers display heterogeneity in whether two categories are treated as substitutes or complements. Figure 4 shows boxplots for each cross-category interaction (the elements below the diagonal of the matrix $B$). Figure 4 reveals that most category interactions demonstrate a slight degree of substitutability. Interactions of household categories with food categories display a greater degree of substitutability but also show that for some consumers household and food categories can be complements.

5 Results

5.1 Background on SNAP

We now use the estimated model to conduct counterfactual experiments and study the effect of various subsidy designs. Our population of interest consists of consumers who are eligible for SNAP benefits and we are interested how this population behaves in response to a policy change. We use federal guidelines to simulate whether consumers in our sample are eligible for SNAP benefits. Using criteria from the 2018 fiscal year, we say that a consumer is SNAP eligible when his yearly reported income falls below the eligibility threshold. In section 2 we discussed why observing SNAP participation is notoriously difficult. Therefore, we use federal guidelines as a proxy for consumers’ SNAP eligibility and assume no SNAP participation as the baseline status.

9 The threshold is 130 percent of the federal poverty line. In federal fiscal year of 2018, the federal poverty line was $1702 for a household of three, making the eligibility threshold $2213 per month or about $27k per year. (USDA 2018)
Consumers receive SNAP benefits via an EBT card (Electronic Benefit Transfer) which is similar to a debit card. The amount of benefits depends on household size and household’s own contribution. We use the estimated average monthly benefit for a household of three in the fiscal year of 2018 as the benefit amount we use in counterfactual experiments – that amount is $384 per month (CBPP (2018)). How do we implement the effect of receiving a food stamp subsidy? Food stamps act as an extra income which can only be spent on food. Therefore, when formulating the consumer’s problem, food stamps affect the budget constraint. As a result, the consumer problem for bundle choice in equation 3 is reformulated as follows:

\[
\max_q \quad u(q) \quad \text{s.t.} \quad p^T q_{\text{in}} + q_{\text{out}} = E + \Delta_{\text{SNAP}}, \\
\quad p^T q_{\text{food}} \leq E + \Delta_{\text{SNAP}}, \\
\quad p^T q_{\text{non-food}} + q_{\text{out}} \leq E.
\]

Equations 28—30 show how the budget constraint changes when the consumer is receiving SNAP benefits, denoted by $\Delta_{\text{SNAP}}$. In addition to the general budget constraint (equation 28), two other budget constraints need to be accounted for. First, equation 29 specifies that food spending must remain below the combined amount of the subsidy and the original budget. Second, equation 30 specifies that non-food expenditure and outside good expenditure are still constrained by the original budget. Similarly, restricting subsidy use to food and household items changes the consumer problem in the following way:

\[
\max_q \quad u(q) \quad \text{s.t.} \quad p^T q_{\text{in}} + q_{\text{out}} = E + \Delta_{\text{SNAP}}, \\
\quad p^T q_{\text{food}} + p^T q_{\text{non-food}} \leq E + \Delta_{\text{SNAP}}, \\
\quad q_{\text{out}} \leq E.
\]

Equation 32 specifies that the subsidy now applies to food and non-food (household) items and that in this case spending on the outside good is still limited by the original budget (equation 33).
For cash transfers, no additional constraints are used and the consumer problem remains standard with one budget constraint as specified in equation 3.

5.2 Marginal Propensity to Consume Food

In this section, we estimate marginal propensity to consume food (MPCF) under different subsidy designs. MPCF is a key indicator which reflects the main goal of the SNAP program and is also the object of main interest in many research papers, allowing to compare our estimates to other estimates previously reported in the literature. Our target population of interest are food stamp eligible consumers who make up around 24% of our sample. Our definition of marginal propensity to consume certain items out of a subsidy benefit relies on comparing expenditure levels on target items before (the no subsidy regime) and after (the subsidy regime) the receipt of benefits. We compare MPCF under (i) normal food stamp benefits (food only), (ii) expanded food stamp benefits (food and household items), (iii) cash subsidy (e.g., tax rebate). To simulate behavior under different policy regimes, we use \( S = 20 \) draws from the posterior distribution of model estimates. For each draw, we set up a separate optimization exercise corresponding to the subsidy design and record the outcome. We measure total food expenditure across the observation window (52 weeks) with and without a specific subsidy design and compute incremental effect of the subsidy on food expenditure.

<table>
<thead>
<tr>
<th>Mean Marginal Propensity of Spending per $1</th>
<th>SNAP Benefits</th>
<th>Grocery</th>
<th>Tax Rebate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food (MPCF)</td>
<td>0.232</td>
<td>0.244</td>
<td>0.126</td>
</tr>
<tr>
<td>Household Items</td>
<td>0.077</td>
<td>0.119</td>
<td>0.066</td>
</tr>
<tr>
<td>Outside Good</td>
<td>-0.021</td>
<td>-0.033</td>
<td>0.613</td>
</tr>
</tbody>
</table>

Table 5: Incremental impact of different subsidies.

Table 5 demonstrates how different subsidies affect consumer spending across various categories. The average marginal propensity to consume food (MPCF) out of tax rebate benefits across SNAP eligible consumers is 0.126. This estimate is in line with existing evidence that MPCF out of cash income for low-income individuals is around 0.1 (Castner and Mabli (2010), Hoynes and
Schanzenbach (2009)). We estimate that MPCF out of food stamps is 0.232 or 84% larger than MPCF out of cash which is also consistent with reported results in the literature. Our estimate that MPCF out of grocery benefits is 0.244 or 94% higher than MPCF out of cash is a novel and interesting result. The result is novel because it refers to a counterfactual scenario where SNAP spending restrictions are relaxed to allow spending on household items. In addition, the result is interesting because it is not obvious how extending spending to one set of categories affects spending in another set of categories. In our model, there are two mechanisms that tie together spending between different categories. First, the utility function in equation 1 includes cross-category interaction effects which are captured in the quadratic part of the utility function. When spending in one category increases, then spending in another category will also increase (decrease) when the two categories are complements (substitutes). Second, utility from store goods is a key part in the decision to visit a particular store or to not visit any stores (equation 7). When one category is subsidized, the overall attractiveness of visiting any store is increased. As a result, even in the absence of any cross-category effects, spending in one category may increase when another category is subsidized since consumers are likely to visit more stores and have more opportunities to make purchases from all existing categories. In the context of our counterfactuals, we find evidence that store visit decisions indirectly affect marginal propensity of spending on various categories. Both SNAP and extended SNAP benefits incentivize consumers to visit stores and increase yearly store visits by 4.8% and 5.1%, respectively. Extended SNAP benefits are more effective at driving recipients to visit stores because they can spend benefits on a wider range of items which in turn increases spending on all categories, including food. Tax rebate, on the other hand, allows recipients to use benefits without visiting any store and encourages to minimize store visits to avoid travel costs. As a consequence, yearly store visits under tax rebate benefits decrease by 4.1% which means that more purchases are concentrated on fewer trips.

\[\text{Most papers in the literature estimate an MPCF out of food stamps between 0.16–0.32 (e.g., Bruich (2014), Hoynes and Schanzenbach (2009)). Recently, Hastings and Shapiro (2018) reported higher estimate of 0.5–0.6.}\]
5.3 Consumer Welfare

In this section, we measure cost-effectiveness of different subsidy designs from the perspective of consumer welfare. We treat consumer welfare under the usual SNAP subsidy as the baseline and ask each consumer

“How many dollars of subsidy X are needed to achieve the utility under the SNAP subsidy?”.

We define this concept of equivalent variation using the following notation. Each period, consumers solve the problem of store and bundle choice. Let \((q^*_s, s^*)\) denote the generic solution to this problem where \(s^*\) solves the store choice problem

\[
    s^* = \arg\max_s U(s) = \arg\max_s u(q^*_s) - \tilde{\mathcal{R}}_s - \rho \text{dist}_s + \eta_s
\]

and \(q^*_s\) solves the bundle choice problem in store \(s^*\) subject to restrictions denoted by \(R\):

\[
    q^*_s = \arg\max_q u(q) \quad \text{s.t.} \quad (R_1) \quad p^T q_{\text{in}} + q_{\text{out}} = E.
\]

Define

\[
    \mathcal{U}\left(\{p_t\}_{t=1}^T, E, R\right) := \sum_{t=1}^T u(q^*_{s^*,t})
\]

as the “total consumption utility” over the observation window when solving the problem of bundle and store choice each period facing prices \(\{p_t\}_{t=1}^T\), budget \(E\) and spending restrictions \(R\). Note that the definition in equation 36 only involves the “bundle utility” and ignores any utility components that affect store choice. That is because we treat utility from direct consumption of goods as the natural baseline measure of consumer welfare in the context of our setting.

To compare the cost-effectiveness of different subsidy designs, we are interested in estimating the size of hypothetical benefits \(\hat{\Delta}_{\text{SNAP extended}}\) and \(\hat{\Delta}_{\text{tax break}}\) which will make a consumer indifferent between receiving the full SNAP subsidy and receiving benefits under a design where spending is
restricted to food and household items (extended SNAP) or when there are no spending restrictions (tax rebate), respectively. Using the notation above, we are interested in estimating the size of benefits $\hat{\Delta}_{\text{SNAP extended}}$ and $\hat{\Delta}_{\text{tax break}}$ such that the following holds:

$$
\mathbb{U}\left(\{p_t\}_{t=1}^T, E + \hat{\Delta}_{\text{SNAP extended}}, R_{\text{SNAP extended}}\right) = \mathbb{U}\left(\{p_t\}_{t=1}^T, E + \Delta_{\text{SNAP}}, R_{\text{SNAP}}\right)
$$

(37)

$$
\mathbb{U}\left(\{p_t\}_{t=1}^T, E + \hat{\Delta}_{\text{tax rebate}}, R_{\text{tax rebate}}\right) = \mathbb{U}\left(\{p_t\}_{t=1}^T, E + \Delta_{\text{SNAP}}, R_{\text{SNAP}}\right).
$$

(38)

The usual SNAP restrictions, denoted by $R_{\text{SNAP}}$, limit subsidy spending to food categories only. The extended SNAP subsidy restrictions, denoted by $R_{\text{SNAP extended}}$, restrict subsidy spending to food and household categories. There are no restrictions on subsidy spending when receiving a tax rebate ($R_{\text{tax rebate}}$).

We estimate the size of hypothetical benefits for all consumers who qualify for the SNAP subsidy in our sample. We find significant heterogeneity in benefit sizes which is driven by individual-level differences in consumer preferences and shopping frequency which determine the value of more flexible spending under alternative subsidy designs. We find that at mean-levels the “exchange rate” between different subsidies is the following:

$$
$100 \text{ SNAP benefits} = $93 \text{ extended SNAP benefits} = $16 \text{ tax break}.
$$

(39)

The exchange rate in equation 39 reflects the extent to which different spending restrictions affect consumer welfare. The substantial difference between tax benefits and the usual SNAP benefits reveals that consumers place a high priority on spending on the outside good. This estimate also describes the incentive to trade SNAP benefits for cash (a form of benefit fraud). Our estimate that on average $100 in SNAP benefits is equivalent to a $16 tax rebate is in line with existing evidence that SNAP benefits are traded for cash at the rate of $0.50 in cash for $1 in SNAP benefits on online bulletin boards and auction sites.\footnote{See Smith (2019) for references to articles in the popular press.} The relatively similar exchange rate between the normal SNAP subsidy and extended SNAP subsidy benefits reveals that while consumers prefer extended benefits...
which allow subsidy spending on household items, the relative impact of restricting spending to food only is low.

There are two channels in our modeling framework that drive the results above: (i) travel costs and (ii) preference parameters for store goods and outside good. Since there is heterogeneity across consumers, it would be interesting to know whether there are any systematic differences across consumers in the mechanism that drives the “exchange rate” results. First, figure 5 shows that there are systematic differences among income groups as consumers in the bottom half of the income distribution are willing to accept less cash to be indifferent between receiving the tax rebate subsidy and the food stamp subsidy. As pointed out above, there are two possible channels for this discrepancy. Figure 6 shows that travel costs are indeed an important factor in driving exchange rate results, however, there are no significant differences in travel costs across the two income groups (figure 7). As a result, the source of discrepancy in the exchange rate between the two income groups is driven by systematic differences in preferences.

We now discuss implications of these findings for tax policy, benefit design, and consumer welfare. First, government spending on in-kind programs is significant and there are active debates about the desirability of flexible benefits (Lieber and Lockwood (2019)). We quantify the relative trade-offs between normal benefits and flexible benefits in the context of SNAP. Our results show that most of the benefits from lifting restrictions on spending are accrued on the margin of the outside good and not on the inside good. For considering policy change, this result can be viewed from two angles. From the perspective of cost-effectiveness, SNAP benefits could be replaced by much smaller cash transfers in the form of tax rebates, while not affecting consumer welfare. From the perspective of consumer welfare, offering recipients a choice between SNAP benefits or a smaller cash transfer would improve consumer welfare and improve cost-effectiveness of the program. In sum, we quantify differences between various subsidy designs from the perspective of consumer welfare and point out how these results can be used to inform policy.
5.4 Restricting Soda Purchases with Soda Tax or Ban

The fact that SNAP is the largest nutrition assistance program and at the same time its benefit recipients spend more money on soft drinks than any other item\textsuperscript{12} has spurred debates on how benefit spending should be restricted\textsuperscript{13}. While several states and medical groups have urged changes to the SNAP, the Department of Agriculture (USDA) has denied every request so far. Aside from potential administrative costs of implementing additional restrictions, it is uncertain how restricting benefits would change buying habits. Since most households supplement benefits with their own income, restricting soda purchases may have no effect since households can simply use their own cash instead of the benefits to buy soda, as also pointed out in a recent expert testimony (Schanzenbach (2017)). However, there are circumstances when this prediction does not hold. Recent empirical evidence suggests that consumers reduce spending in target category in response to reduced allowances and only partially substitute reduced benefits with personal funds (Andreyeva et al. (2013)\textsuperscript{14}). First, there may be psychological reasons why consumers treat benefits and cash separately (e.g., Hastings and Shapiro (2018) find evidence of mental accounting among SNAP recipients). Second, if consumers follow traditional demand theory then restricting benefit use is equivalent to a price change: when depicting the consumer problem graphically, the budget line pivots in response to the policy change in both cases. Consistent with the latter explanation, consumers in our model respond to subsidy restriction and price increase in the same way. Since the policy maker may have particular preferences to mandate either solution, it would be interesting to know under what circumstances enacting a soda tax or banning soda from SNAP yield the same outcomes in terms of soda consumption.

We simulate consumer behavior under two scenarios: enacting a soda tax and banning soda from SNAP. We find that these two policies yield similar outcomes in terms of reducing soda con-

\textsuperscript{12}See USDA (2016) for an overview of spending habits of SNAP recipients.
\textsuperscript{13}See for example Lane (2017), O’Connor (2017), Bittman (2012) for articles in the popular press.
\textsuperscript{14}Andreyeva et al. (2013) look at recipient response to reduced allowances of sugary juices in the context of the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC). They find that after implementation of the revisions, WIC juice purchases were reduced on par with allowance changes and this reduction was only partially compensated in the form of an increase in juice purchases using personal and other non-WIC funds.
sumption. Figure 8 shows mean response under three different soda tax policies and a ban policy where soda is excluded from SNAP benefits. Consumer response is calculated with respect to the baseline policy (the full SNAP benefit). Figure 8 shows that consumers reduce soda consumption by 20 percent and 23 percent under the soda ban and 10% soda tax, respectively. How does rest of the consumer basket change? We find that consumers substitute between different categories within a store, increasing their non-soda food spending substantially (figure 9) with little changes to the aggregate share of the outside good or shopping frequency. Looking at the effect of soda ban across income groups (figure 10), we notice that the ban policy has a differential impact on two income groups. In particular, the lowest income consumers only partially substitute to other food items and instead increase their outside good spending relative to the baseline policy. In contrast, consumers whose income is above the median of benefit-eligible consumers do not change their share of outside good purchases and and instead substitute to alternative store goods that are still allowed under the SNAP subsidy with soda excluded. Similar to results in section 5.3, consumer response depends on two channels, including (i) travel costs and (ii) preferences for store goods and outside good. Since there are no significant differences in travel costs across the two income groups (figure 7), we conclude that most of the discrepancy in response is driven by differences in preferences. This contrast among different groups is interesting as it shows how a simple ban policy can have a differential impact through different substitution patterns generated by differences in underlying preferences.

6 Conclusion

Many government funded welfare programs provide its recipients in-kind transfers which are subject to different spending restrictions. Since the budgetary and administrative costs of implementing such programs are large, there is a growing debate over whether in-kind benefits should be replaced with simple cash benefits. In the context of SNAP (formerly known as the Food Stamps
Program), we analyze the effect of SNAP benefits from the perspective of the policymaker (who prefers that funds are used to buy food) and recipients (who care about overall consumer welfare). The existing literature has documented that marginal propensity to consume food (MPCF) out of food stamps is higher than out of cash, thereby rationalizing the use of food stamps from the perspective of the policymaker. We develop a structural model of consumer demand for brands, categories, and stores and study how SNAP benefits affect spending on different categories. Our structural model yields estimates that are in line with existing results in the literature and provides novel estimates for alternative subsidy designs that are not observed in practice but are often part of the policy debate. Our main finding is that expanding SNAP benefits to grocery items (food plus household goods) would yield outcomes that are preferred by both benefit recipients and the policymaker. This finding is interesting because it shows that the new design improves both measures of interest (food spending measured in MPCF and consumer welfare) without any compromise, thereby providing a positive answer to the main question posed in this paper. While expanding benefits to household goods would improve consumer welfare, we find that most benefits are accrued on the margin of the outside good and demonstrate that the exchange rate of cash subsidy to SNAP subsidy is around one fifth. Second, we study the other lever which the public has often prompted for the policymakers to pull and examine the effects of banning benefit use on certain goods. We establish that restricting benefit use has similar effects to a tax and in the context of soda category find that a ban would have similar effects on reducing soda consumption as enacting a ten percent tax on soda. We find evidence that excluding soda from SNAP benefits has a differential impact on the substitution patterns of low and high income consumers, which is mainly driven by differences in preferences.

Results and analysis in this paper are relevant for policymakers who need to consider both sides – recipients and program goals set by the policymaker – when considering policy change. Our empirical results highlight the relative trade-offs of existing and alternative policies often considered in the public debate and offer concrete mechanisms to explain the differential impact of subsidy policies.
APPENDIX

A1: Overview of Markov Chain Monte Carlo Estimation

This section gives an overview of the estimation algorithm. For individual coefficients, we use random walk Metropolis algorithm (for $\beta^h_0$) and Adaptive Metropolis (AM) algorithm of Haario et al. (2001) for vector components ($\psi^h$ and $B^h$). To update the hyperparameters, we use posterior distributions based on the updated individual coefficients. In the first case, a new individual parameter value $\beta^h_0^*$ is proposed by perturbing the current logit ratio

$$\text{logit} (\beta^h_0^*) = \text{logit} (\beta^h_0) + \sigma_{MET, \beta_0} \bar{\sigma}_{\beta_0} N (0, 1),$$

and then backing out the implied value by $\beta^h_0^* = \frac{\exp(\text{logit}(\beta^h_0^*))}{1 + \exp(\text{logit}(\beta^h_0^*))}$. The value for the current iteration is set to $\beta^h_0^*$ if

$$U < \frac{P(q | \beta^h_0^*) P (\beta^h_0 | \bar{\beta}_0, \bar{\sigma}_{\beta_0})}{P(q | \beta^h_0) P (\beta^h_0 | \bar{\beta}_0, \bar{\sigma}_{\beta_0})} \quad \text{where } U \sim \text{Uniform} (0, 1),$$

and left unchanged otherwise. The metropolis standard deviation, $\sigma_{MET, \beta_0^*}$, is individual specific and is modified after every 500 iterations to make the acceptance rate of the chain close to 44 percent, the optimal rate of acceptance in a one-dimensional context (Roberts and Rosenthal (2001)). The metropolis standard deviation is increased if the acceptance rate is greater than the target acceptance rate, and decreased if the acceptance is smaller than the target acceptance rate. In the second case, we use vector jumping and accept or reject the entire vector. For updating $\psi^h$, the proposal is given by

$$Q (\psi^h) = (1 - \beta) N \left( \psi^h, (2.38)^2 \Sigma_{\text{emp}}/d \right) + \beta N \left( \psi^h, (0.1)^2 I_d/d \right),$$

where $\Sigma_{\text{emp}}$ is the current empirical estimate of the covariance structure of target distribution based on the run so far, and where $\beta = 0.05$ is a small positive constant. The idea is that the pro-
posal $N(\cdot, (2.38)^2\Sigma/d)$ is optimal in a particular large-dimensional context (Roberts and Rosenthal (2001)), and the empirical version is an effort to approximate this. A small amount of normal noise is added to avoid the chain from getting stuck. The algorithm can be adapted during the warm-up phase by specifying $H$, the length of memory, and $c$, the coefficient in front of $\Sigma_{\text{emp}}$. We take $H = 200$ and update the empirical covariance structure after every $H$ iterations and fix the covariance structure after the warm-up period is over. For the coefficient, we start with $c = 2.38$ as shown above and modify it after every 50 iterations to make the acceptance rate close 23 percent, the optimal acceptance rate in large-dimensional contexts (Roberts and Rosenthal (2001)).

To initialize the sampler, we draw initial values for the hyperparameters and then generate individual values from their corresponding distributions. To generate initial utility shocks, we use equations:

$$
\varepsilon_{scb,t}^0 = \ln \left( 2 \cdot B_{scb, t}^0 \cdot u_t^0 + \frac{p_{scb,t}}{\psi_{scb}^0} \cdot q_{out,t}^{0\cdot-1} \exp(\varepsilon_{0,t}^0) \right) \quad q_{scb}^* > 0
$$

$$
\varepsilon_{scb,t}^0 \sim EVT(0, \sigma_{\varepsilon}^0) \quad \text{s.t} \quad \varepsilon_{scb,t}^0 < \ln \left( 2 \cdot B_{scb, t}^0 \cdot u_t^0 + \frac{p_{scb,t}}{\psi_{scb}^0} \cdot q_{out,t}^{0\cdot-1} \exp(\varepsilon_{0,t}^0) \right) \quad q_{scb}^* = 0.
$$

A2: Figures
Figure 2: Estimates of the mean parameter of log-normally distributed baseline marginal utility ($\bar{\psi}$).

Notes: This figure presents posterior mean estimates and 95% Bayesian credible intervals for population means of log-normally distributed brand preferences (the $\bar{\psi}_i$ - parameters in $\psi_i \sim \text{log-normal} (\bar{\psi}_i, \bar{\sigma}_\psi)$ $i = 1, \ldots, K$). There are 3–6 brands in each category, with the first brands representing top brands in the category and the last brand representing the composite brand. Calculations are based on $S = 100$ random samples from the thinned MCMC chain where every 300th iteration is saved.
Figure 3: Estimates of the standard deviation parameter of log-normally distributed baseline marginal utility ($\bar{\sigma}_\psi$).

Notes: This figure presents posterior mean estimates and 95% Bayesian credible intervals for population standard deviations of log-normally distributed brand preferences (the $\bar{\sigma}_\psi$ parameters in $\psi_i \sim \text{log-normal}(\bar{\psi}_i, \bar{\sigma}_\psi_i)$, $i = 1, \ldots, K$). There are 3–6 brands in each category, with the first brands representing top brands in the category and the last brand representing the composite brand. Calculations are based on $S = 100$ random samples from the thinned MCMC chain where every 300th iteration is saved.
Figure 4: Category interaction effects.

Notes: This figure presents boxplots of individual coefficients of category interaction effects. Positive coefficients indicate degree of substitutability and negative coefficients indicate degree of complementarity. Estimate of each individual’s coefficient is based on the posterior mean of $S = 100$ random samples from the thinned MCMC chain where every 300th iteration is saved.
Figure 5: Exchange Rate by Income Groups.

Notes: This figure presents boxplots of subsidy “exchange rate” across two income groups. Two income groups are generated by dividing eligible consumers (monthly income below $2203) into two groups based on the median income (monthly income $1346). The exchange rate on the y-axis is the amount of dollars in cash subsidy (no restrictions) that would make consumers indifferent between choosing the cash subsidy in the respective amount or a $100 food stamp subsidy.

Figure 6: Exchange rate by travel costs.

Notes: This figure presents a scatterplot of subsidy “exchange rate” across standardized measure of travel costs. The exchange rate on the x-axis is the amount of dollars in cash subsidy (no restrictions) that would make consumers indifferent between choosing the cash subsidy in the respective amount or a $100 food stamp subsidy. Travel costs comprise of individual-store fixed effect and a distance measure (see equation 7 for details) and are standardized (Z-score).
Figure 7: Travel Costs Across Two Income Groups.

Notes: This figure presents boxplots of standardized travel costs across two income groups. Two income groups are generated by dividing eligible consumers (monthly income below $2203) into two groups based on the median income (monthly income $1346). Travel costs comprise of individual-store fixed effect and a distance measure (see equation 7 for details) and are standardized (Z-score).

Figure 8: Reduction in Soda Consumption Across Different Policies Relative to SNAP Policy.

Notes: This figure presents a line plot of reduction in soda consumption (relative to the baseline SNAP policy) across different policies. The tax policy is computed at three points, corresponding to enacting a 5%, 10%, and 15% tax on soda while still keeping consumers subject to the usual SNAP policy. The ban policy excludes soda from SNAP benefits, while keeping benefits otherwise the same.
Figure 9: Substitution to Non-soda Food.

Substitution to Non–soda Food

Notes: This figure presents a line plot of increase in non-soda food consumption (relative to the baseline SNAP policy) across different policies. The tax policy is computed at three points, corresponding to enacting a 5%, 10%, and 15% tax on soda while still keeping consumers subject to the usual SNAP policy. The ban policy excludes soda from SNAP benefits, while keeping benefits otherwise the same.

Figure 10: Comparison of Substitution Between Non-soda and Outside Good Across Two Income Groups.

Notes: This figure presents a boxplot comparison of two variables across two income groups. The figure compares substitution patterns between non-soda and outside good across two income groups. Two income groups are generated by dividing eligible consumers (monthly income below $2203) into two groups based on the median income (monthly income $1346).
References


USDA (2016). Foods typically purchased by supplemental nutrition assistance program (snap) households.


