Essays on the Estimation of Demand for Complementary Goods

Ludovic Stourm
University of Pennsylvania, stourm@alumni.upenn.edu

Follow this and additional works at: http://repository.upenn.edu/edissertations
Part of the Advertising and Promotion Management Commons, and the Marketing Commons

Recommended Citation
http://repository.upenn.edu/edissertations/2042

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/edissertations/2042
For more information, please contact libraryrepository@pobox.upenn.edu.
Essays on the Estimation of Demand for Complementary Goods

Abstract
This dissertation investigates the problem of demand estimation across complementary goods at the level of individual purchase decisions made by consumers. At this micro level, package sizes may restrict the amount of goods that can be purchased, and temporary price discounts may induce consumers to stockpile goods in anticipation of their future consumption needs, leading to dynamics in purchasing behavior over time. We address these issues in two essays by proposing new structural models of demand. In the first essay, we develop a new approach to modeling consumer preferences for complements which is based on household production theory, and we study the importance of accounting for package size restrictions in modeling cross-category price effects. In the second essay, we embed this approach in a model of consumer forward-looking behavior to study the consequences of stockpiling behavior on the spillover effect of prices across complementary products that are storable.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Marketing

First Advisor
Raghuram Iyengar

Second Advisor
Eric T. Bradlow

Subject Categories
Advertising and Promotion Management | Marketing

This dissertation is available at ScholarlyCommons: http://repository.upenn.edu/edissertations/2042
ESSAYS ON THE ESTIMATION OF DEMAND FOR COMPLEMENTARY GOODS

Ludovic Stourn

A DISSERTATION

in

Marketing

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

Raghuram Iyengar,
Associate Professor of Marketing

Co-Supervisor of Dissertation

Eric T. Bradlow,
K. P. Chao Professor of Marketing,
Statistics and Education

Graduate Group Chairperson

Eric T. Bradlow, K. P. Chao Professor of Marketing, Statistics and Education

Dissertation Committee

David R. Bell, Xinmei Zhang and Yongge Dai Professor of Marketing
Jean-François Houde, Assistant Professor of Business Economics and Public Policy
Dedicated to Valeria, Cecilia and Victor.
ACKNOWLEDGEMENT

During the four years I spent in Wharton’s Ph.D. program, I met some very talented people that I wish to acknowledge for their help and support. I could have not completed the PhD program without them, and it is not without nostalgia that I am preparing to leave Wharton to start a new adventure.

First and foremost, I would like to thank my co-advisors Raghu Iyengar and Eric Bradlow for their support, endless energy and availability. In our weekly meetings, we have had insightful discussions and I have learned a lot about research.

I also want to extend warm thanks to my other committee members David Bell and Jean-François Houde. David gave me thorough advice on my dissertation idea early in the program, and the directions he suggested have guided me in the development of my dissertation. Jean-François gave me crucial advice on methodological issues and his expertise in industrial organization has been invaluable.

Other faculty members gave me insightful feedback at various stages of my dissertation, whose names deserve to be acknowledged: Josh Eliashberg, Dave Reibstein, Pete Fader and Bob Meyer. More generally, the current dissertation benefited from comments that were brought up by the faculty and students during my presentations in Wharton’s Marketing Department and I am very grateful for these.

This work would have not been possible without the funding I received from the Jay H. Baker Retailing Center, for which I am very grateful. I am also indebted to the James M. Kilts Center for Marketing at Chicago Booth for giving me access to the data I used in my dissertation. I also want to thank my former boss Steve Cohen and the team at In4mation Insights, who first introduced me to marketing research and supported my decision to pursue an academic career.

On a personal note, I want to extend special thanks to all the PhD students (current and
past) that I have met during my time at Wharton, who have contributed to make each day more enjoyable in the office.

Last but not least, I want to thank my wife Valeria for her support. I could have not completed this work without her encouragement.
ABSTRACT

ESSAYS ON THE ESTIMATION OF DEMAND FOR COMPLEMENTARY GOODS

Ludovic Stourm

Raghuram Iyengar

Eric T. Bradlow

This dissertation investigates the problem of demand estimation across complementary goods at the level of individual purchase decisions made by consumers. At this micro level, package sizes may restrict the amount of goods that can be purchased, and temporary price discounts may induce consumers to stockpile goods in anticipation of their future consumption needs, leading to dynamics in purchasing behavior over time. We address these issues in two essays by proposing new structural models of demand. In the first essay, we develop a new approach to modeling consumer preferences for complements which is based on household production theory, and we study the importance of accounting for package size restrictions in modeling cross-category price effects. In the second essay, we embed this approach in a model of consumer forward-looking behavior to study the consequences of stockpiling behavior on the spillover effect of prices across complementary products that are storable.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACKNOWLEDGEMENT</strong></td>
<td>iii</td>
</tr>
<tr>
<td><strong>ABSTRACT</strong></td>
<td>v</td>
</tr>
<tr>
<td><strong>LIST OF TABLES</strong></td>
<td>viii</td>
</tr>
<tr>
<td><strong>LIST OF ILLUSTRATIONS</strong></td>
<td>ix</td>
</tr>
<tr>
<td><strong>CHAPTER 1</strong>: The importance and challenges of modeling demand for complements</td>
<td>1</td>
</tr>
<tr>
<td><strong>CHAPTER 2</strong>: A Flexible Demand Model for Complements Using Household Production Theory</td>
<td>4</td>
</tr>
<tr>
<td>2.1 Motivation</td>
<td>4</td>
</tr>
<tr>
<td>2.2 Formalization and literature</td>
<td>7</td>
</tr>
<tr>
<td>2.3 Model</td>
<td>10</td>
</tr>
<tr>
<td>2.4 Econometric model</td>
<td>19</td>
</tr>
<tr>
<td>2.5 Simulation study</td>
<td>23</td>
</tr>
<tr>
<td>2.6 Empirical application</td>
<td>27</td>
</tr>
<tr>
<td>2.7 Conclusion</td>
<td>34</td>
</tr>
<tr>
<td><strong>CHAPTER 3</strong>: A Dynamic Model of Purchase and Consumption Across Complementary Categories</td>
<td>36</td>
</tr>
<tr>
<td>3.1 Motivation</td>
<td>36</td>
</tr>
<tr>
<td>3.2 Literature</td>
<td>38</td>
</tr>
<tr>
<td>3.3 Model</td>
<td>42</td>
</tr>
<tr>
<td>3.4 Simulation study</td>
<td>52</td>
</tr>
<tr>
<td>3.5 Empirical application</td>
<td>57</td>
</tr>
<tr>
<td>3.6 Conclusion</td>
<td>64</td>
</tr>
</tbody>
</table>
LIST OF TABLES

TABLE 1: Results of a simulation study under the household-production model of complementarity ........................................ 25

TABLE 2: Description of purchase data in the tortilla chips and Mexican salsa categories .................................................. 29

TABLE 3: Reduced-form evidence of complementarity between tortilla chips and Mexican salsa ........................................... 29

TABLE 4: Estimation results on the tortilla chips and Mexican salsa purchase data .................................................................. 30

TABLE 5: Results of counterfactual analyses under alternative couponing strategies ............................................................ 33

TABLE 6: Integrating the literature on cross-category demand estimation and the literature on stockpiling behavior .............. 39

TABLE 7: Simulation results under different versions of the dynamic cross-category model of demand ................................. 54

TABLE 8: Description of purchase data in the spaghetti and tomato sauce categories ............................................................ 59

TABLE 9: Reduced-form analysis of stockpiling behavior ........................................................................................................ 59

TABLE 10: Reduced-form analysis of complementarity between spaghetti and tomato sauce .................................................. 61

TABLE 11: Estimates of the price process parameters ........................................................................................................... 62

TABLE 12: Estimation results on the spaghetti and tomato sauce purchase data ................................................................. 63
LIST OF ILLUSTRATIONS

FIGURE 1: Relationship between $\tilde{U}$, $U$ and $V$ in the household-production model of complementarity ........................................ 12

FIGURE 2: Indifference curves under various parameter values .................. 16

FIGURE 3: Consumption policy under various parameter values ............... 49

FIGURE 4: Effect of an exogenous promotion on sales across categories in simulated data ................................................................. 57

FIGURE 5: Effect of exogenous promotions on sales of spaghetti and tomato sauce over time ................................................................. 65
CHAPTER 1 : The importance and challenges of modeling demand for complements

This chapter motivates the importance of modeling demand for complementary goods in marketing, reviews the related literature, introduces some of the challenges in doing so, and gives an overview of how we tackle them in the next two chapters.

Understanding how demand responds to marketing activity is crucial for any successful marketing action. In today’s world, companies record data about individual purchases made by their customers, and can leverage this historical data to get insight about their preferences and purchasing behavior. Companies can then use this knowledge to make better decisions about their marketing mix, for example by optimizing their prices or their promotion schedules to achieve a higher profit. For retailers and manufacturers who produce multiple goods, it is important to understand how their marketing mix affects demand across related categories, since their total profit depend on their joint sales. Therefore, they can better coordinate their marketing strategy by analyzing purchase data across categories. The first part in implementing such a data-driven strategy consists in estimating an empirical model of demand on the data. It is particularly important to have a good model of demand because it is used as input for the evaluation of any alternative marketing plan: as such, the optimal marketing action derives directly from it.

Micro-level models of demand, where each individual consumer purchase decision is modeled, present at least two important advantages over macro-level models of demand, where demand is aggregated over consumers and/or time periods in some pre-defined markets. First, it is important for marketers to understand consumer heterogeneity, so that the marketing mix can be tailored to the specific needs of individual consumers. For example, companies target their customers when sending promotional coupons, or propose different versions of their products to capture different segments of consumers. Macro-level models of demand cannot give insights about such consumer heterogeneity, since they only represent consumers through their aggregate behavior. The second advantage is methodological
and relates to the estimation of the model: as explained by Lucas (1976), aggregate-level empirical models yield parameter estimates that are not invariant to policy changes. In a marketing context, this means that the parameters of the model depend on the current marketing mix: if some aspects of the marketing mix are changed, then the parameter estimates are no longer accurate. This can be problematic since the purpose of the analysis is usually to evaluate alternative marketing strategies. On the other hand, it is possible to specify a micro-level model of consumer behavior based on economic theory such that its parameters are invariant to policy changes.

In the marketing literature, the first demand model applied on micro-level scanner data was the logit model by Guadagni and Little (1983) to investigate consumer choices of brands within a product category. Many models have been developed since then to capture consumer heterogeneity (Kamakura and Russell, 1989; Rossi et al., 1996), to model consumers’ purchase incidence and quantity decisions (Chintagunta, 1993), and to capture dynamic behavior over time such as learning (Erdem and Keane, 1996) or stockpiling (Erdem et al., 2003; Sun, 2005). A stream of research has also emerged to model cross-category demand, with the goal of measuring the spillover effect of marketing activity and prices across related product categories (Mehta, 2007; Song and Chintagunta, 2007). Unlike brands within a product category, which are typically substitutes, different categories may be substitutable or complementary: under the usual definition from economics, they may exhibit positive or negative cross price effects on demand.

As noted by Berry et al. (2014), many empirical tools to study demand for substitutable goods have been developed, but the estimation of demand for complements still presents some methodological challenges. While the starting point of a micro-level demand system derived from first economic principles is that the consumer maximizes her utility or stream of utilities defined on goods purchased, the problem of specifying a parametric form for the utility function such that it can capture complementary effects, and yet is amenable for empirical research, has not been fully resolved in previous research. Instead, previous
models have relied on additional assumptions to derive a demand system, such as continuous quantities or sequential decision processes (Mehta, 2007; Song and Chintagunta, 2007; Mehta and Ma, 2012; Lee and Allenby, 2014).

This dissertation proposes a new approach to modeling demand for complements by invoking household production theory. Under that view, consumers purchase goods to use them as inputs to produce final goods, and they derive utility by consuming these final goods. In Chapter 2, we show that representing consumer preferences for latent final goods consumed instead of the goods purchased can yield a flexible utility function that can capture complementarity, using very simple parametric forms. We then study the properties of the resulting demand system in a static model with independent purchase occasions, and estimate the model on data to show how it can be used for coordinated pricing and packaging decisions across complementary categories. In Chapter 3, we embed our utility specification in a dynamic model of cross-category purchase and consumption where consumers are forward-looking and can store goods between time periods. Through simulation and in a real data application, we investigate the importance of accounting for consumer stockpiling behavior in estimating the spillover effects of prices across complementary categories.
CHAPTER 2 : A Flexible Demand Model for Complements Using Household Production Theory

2.1. Motivation

Consumers commonly purchase multiple goods to consume them jointly. For example, they buy burgers and buns to make sandwiches, they combine detergent and softener for their laundry, or eat crackers with dips. From a managerial point of view, cross-category consumption provides an opportunity for coordinated promotion and pricing across the goods. A rigorous analysis of optimal marketing strategy first requires a model of demand in a way that takes into account the volumes purchased by consumers across categories. Not surprisingly, modeling demand across categories has therefore been an important area of research in the marketing literature.

By leveraging household-level data of purchases made across categories and over time, marketing researchers have built models to estimate demand at the micro level (Manchanda et al., 1999; Song and Chintagunta, 2007; Mehta, 2007; Niraj et al., 2008; Mehta and Ma, 2012; Lee et al., 2013). Many of these econometric models take a utility-maximization paradigm to obtain parameters that are invariant to policy changes. In this paradigm, the starting point is that consumers choose to buy the quantities of goods that maximize their utility under a budget constraint. The goal of estimation is therefore to find the parameters of the model that best explain the purchases observed, under the assumption that these purchases are the result of an optimization problem. From this common setup, two approaches have been developed, which differ in what economic object is parameterized. Both approaches have advantages and limitations.

In the first approach, the researcher specifies a functional form for the indirect utility and applies Roy’s identity (Roy, 1947) to derive the demand function, which relates the parameters of the indirect utility function to the consumer’s quantity decisions as a function of prices. While previous literature has shown that this approach can yield flexible patterns of
complementarity and/or substitution across goods, Roy’s identity only applies if any continuous quantity can be demanded, and if prices are linear in quantity. These assumptions may lead to biases when demand is indivisible due to package size constraints (Lee and Allenby, 2014).

In the second approach, the researcher parameterizes the consumer’s direct utility function and solves the first-order conditions to obtain the demand function. The model can then be easily extended to include richer aspects of the world, such as discrete package sizes. However, it is difficult to find a functional form for the direct utility that is tractable (in the sense that the first-order conditions can be solved easily), regular (in the sense that it yields well-behaved preferences on the entire domain of preference parameters), and flexible (in the sense that it can accommodate different patterns of consumer preferences and demand functions). This is problematic: a non-flexible functional form may lead to a poor model of demand if it is unable to capture a pattern that exists in the data. Additively-separable specifications, which are the most tractable, cannot lead to complementarity between goods, as defined by negative cross price effects (Chintagunta and Nair, 2011). For that reason, cross-category demand models often take the indirect utility approach.

In this chapter, we propose a new model of demand for complementary goods in a way that yields flexible patterns, yet can be used in contexts where Roy’s identity does not apply. Our approach relies on household production theory (Muth, 1966): the original consumer problem is altered to distinguish between inputs that can be purchased and latent final goods that can be consumed. We represent the joint consumption of multiple inputs by a separate final good: for example, buns and burgers can be purchased (and consumed separately), but sandwiches need to be produced from buns and burgers. Formally, the consumer has a utility function defined on goods consumed (e.g. sandwiches, burgers and buns), but can only buy inputs (e.g. burgers and buns). She decides not only what inputs to purchase but also what final goods to produce for her own consumption. This conceptual model is equivalent to the original one if we assume that the final goods are identical to the
goods purchased, but it is more general if we assume that a broader set of final goods can be produced from the goods purchased.

Even though consumption may not be observable, we find that it is useful to lay out the consumer problem in this way to obtain a parametric representation of consumer preferences for observables that yields a flexible demand system. We show that a direct utility defined on inputs can always be derived from our model, although its functional form is not straightforward. In contrast, even a separately-additive functional form for the utility on final goods can capture very different patterns of preferences, including perfect complementarity, asymmetric complementarity or no complementarity. Thus, representing consumer preferences for latent final goods instead of inputs greatly facilitates the problem of finding a flexible, tractable and regular functional form.

We consider three different versions of our model: one version where the consumer can purchase any continuous quantities of goods, one version where demand is restricted to lie on a discrete grid of points due to package size constraints, and one version where demand is discrete but we shut off complementarity by removing any joint consumption as a final good. We derive some insights about the properties of our demand system by studying the continuous model since the corresponding consumer problem can be solved analytically. In a simulation study, we investigate the consequences of ignoring (or allowing for) discreteness and complementarity between goods in the estimation of own- and cross-category price effects, and we show that our model recognizes the absence of complementarity when demand is independent across goods. In an empirical application, we estimate our models on purchase data from a panel of consumers over time, focusing on the tortilla chips and Mexican salsa categories. We find that ignoring the discreteness of demand leads to very different inferences about cross-price elasticities, which is of particularly interesting since most previous multicategory models of demand take the indirect-utility approach and ignore such discreteness. Then, we run counterfactual analyses to show how our model can be used to make improved managerial decisions regarding the distribution of coupons to
consumers.

The rest of the chapter is organized as follows. In Section 2.2, we formalize the economic problem of a consumer making quantity choices across related goods, and review the econometric models that have been derived from it in the marketing literature. In Section 2.3, we lay out a more general model of consumer behavior based on household production theory, and show its usefulness for empirical work by studying its properties. In Section 2.4, we introduce stochasticity in the model, and discuss its identification and estimation. Section 2.5 describes the results of a simulation study that compares our model against extant research. In Section 2.6, we describe an application of our model on household-level purchase data, and show how it can inform managerial decisions. Section 2.7 concludes with a discussion of the main contributions and limitations of our model, and ideas for future research.

2.2. Formalization and literature

Suppose we want to model the purchase decisions of a consumer across \( J \) goods of interest. In the simplest microeconomic model, the consumer has a utility function \( U \) defined on the quantities of these goods \( x = (x_1, ..., x_J) \) and the quantity \( z \) of an outside good. After observing the prices \( p \), she chooses to buy the quantities that maximize her utility within her budget \( M \). Mathematically, the consumer problem is represented as:

\[
V(p, M) = \max_{x \in \mathcal{X}, z \geq 0} \quad U(x, z)
\]

subject to: \( p(x) + z \leq M \) \hspace{1cm} (2.1)

where \( x \) belongs to some set \( \mathcal{X} \), \( p(x) \) is the dollar amount to be paid to buy the quantities \( x \) of inside goods, and where the price of the outside good is normalized to one. In this maximization problem, \( U \) is the direct utility, \( V \) is the indirect utility, and the demand function \( x^*(p, M) \) gives the optimal quantities. From the observation of prices and the quantities purchased by consumers, the goal of demand estimation consists in characterizing the demand function so that one can measure the effect of prices on demand, and/or predict

7
demand under alternative scenarios with different prices \( p \), budget \( M \), or constraints on the set \( X \).

Empirically estimating a demand model requires one to make parametric assumptions which should lead to a tractable, flexible and regular demand system. Finding such a functional form turns out to be challenging in the case of complementary goods. Previous marketing literature has addressed it in two different ways, by either specifying a functional form for the indirect utility \( V \) or for the direct utility \( U \). The article by Chintagunta and Nair (2011) provides an excellent review of that literature.

In the indirect-utility approach, the researcher specifies a functional form for \( V \), typically using the translog function which can approximate any twice continuously differentiable function \( V \) at a second degree (Song and Chintagunta, 2007; Mehta, 2007; Mehta and Ma, 2012). Under the assumptions that prices are linear in quantity \( (p(x) = \sum_j x_j p_j) \), and that any non-negative quantities can be demanded \( (X = \mathbb{R}^J_+) \), the researcher then applies Roy’s identity to obtain the demand function: \( x^*_j = -\frac{\partial V}{\partial p_j} \frac{\partial V}{\partial M} \). The estimation problem consists in finding those parameters of \( V \) that best fit the data. This approach has been used to model demand across substitutable and complementary categories, providing an integrated framework for purchase incidence, volume and brand decisions (Song and Chintagunta, 2007; Mehta, 2007; Mehta and Ma, 2012). Parameterizing the indirect utility \( V \) bypasses the problem of specifying a functional form for the direct utility \( U \) and is thought to be a more flexible approach since the indirect utility can always be derived from a direct utility specification but the reverse is not true. However, the underlying properties of consumer preferences are not clear and may lead to regularity problems: the demand function may not always exhibit some desirable properties, such as homogeneity of degree zero. In a recent article, Mehta (2015) alleviates some of these concerns by studying the conditions on the specification that are necessary and sufficient for regularity. Nevertheless, this approach still has limitations: the assumptions of infinitely divisible demand make it impossible to study the effects of discrete package sizes.
In a second approach pioneered by Kim et al. (2002), the researcher specifies a functional form for the direct utility $U$ and derives the demand function by solving the first-order conditions of the consumer problem in Equation 2.1. This approach is appealing because it allows one to accommodate different pricing schemes and/or constraints present in the real world by changing the feasible set $\mathcal{X}$ and the price function $p$. For example, Satomura et al. (2011) allow for multiple constraints (such as a volume constraint) instead of a single budget constraint; Lee et al. (2014) capture the indivisibility of demand by including integer constraints on the feasible set. Through these extensions, researchers have gained a richer understanding of the impact of prices and packaging on demand. However, one major hurdle consists in finding a parameterization of $U$ that yields flexible patterns of demand. Chintagunta and Nair (2011) show that, under linear prices, additively separable utility functions of the form $U_x(x_1, \ldots, x_J) = \sum_{j=1}^{J} \phi_j(x_j)$ can only yield nonnegative cross effects of prices on demand, thus ruling out complementarity between the goods. On the other hand, non-additively separable utility functions may be intractable or may not exhibit desired properties of regularity for all parameter values, such as monotonicity and concavity. Thus finding a parametric form of $U$ to capture complementarity is difficult. Lee et al. (2013) circumvent this problem by assuming a sequential decision process whereby consumers make purchase decisions one category at a time: at each stage, the consumer maximizes a category-specific utility function that may be affected by purchase decisions in previous categories through an effect of inventories. Such sequential decisions are characteristic of lexicographic preferences, which cannot be represented by a joint utility function defined on the entire set of goods.

We develop a third approach, which relies on the household production theory introduced by Muth (1966). In this view, households buy inputs to produce final goods that they consume: as such, they are both producers (with a production function that transforms inputs into final goods) and consumers (with a utility function defined on final goods). In our context, we argue that complementary goods can be consumed either separately, or jointly in some combination. We represent the $J$ goods of interest as inputs that can be
purchased, and each consumption use (separate or joint) as a different final good. Thus two goods can be complementary if the consumer uses them jointly to construct a final good from which she enjoys utility.

Even though consumer preferences are only revealed through purchase behavior in the absence of consumption data, we show that it is useful to represent consumer preferences over consumption bundles to obtain a flexible system of demand for goods purchased, which is amenable for estimation in the presence of indivisibility constraints and corner solutions. Our work is closely related to the latent separability concept defined by Blundell and Robin (2000), which is a property of direct utility functions. Like the authors, we recognize that parts of the same input can be allocated to different final goods. For example, consumers may eat some burgers by themselves, and combine some with buns to make sandwiches. Unlike Blundell and Robin (2000) however, we do not require fewer final goods than inputs: in fact, we include more final goods than inputs. This is because they seek to reduce the dimensionality of the demand system by grouping goods into latent groups while we aim at flexibly estimating the demand across a narrower set of goods. Section 2.3 describes our model in more depth.

2.3. Model

We start by laying out the general model of household production and explain how it relates to the direct-utility approach and the indirect utility approach. Then we describe our parametric assumptions, study the implications of the resulting model for consumer preferences over goods purchased, and its implications for the demand function under the assumptions of Roy’s identity.

2.3.1. Consumer problem with household production

By considering household production, we make a conceptual distinction between the inputs that the consumer can purchase on the market, and the final goods that she can consume after producing them from the inputs. For example, a consumer may be able to buy burgers
and buns which are sold separately on the market, but in order to eat a sandwich, she first needs to produce it by combining a bun and a burger. In this example, burgers and buns are inputs, and sandwiches are final goods. We denote by $J$ the number of distinct inputs, by $K$ the number of distinct final goods that can be produced from the $J$ inputs, by $x = (x_1, ..., x_J)$ the volumes of inputs purchased and by $c = (c_1, ..., c_K)$ the quantities of final goods consumed. Importantly, the consumer first needs to buy enough inputs to be able to produce the final goods. To do so, she has a budget $M$ that she can spend across the inputs and to buy some quantity $z$ of an outside good whose price is normalized to one.

The consumer enjoys utility from consuming the final goods and the outside good, which is represented by a utility function $\tilde{U}(c, z)$. The consumer problem consists in buying the optimal volumes of inputs $x$ and outside good $z$, and use the inputs to produce the optimal quantities of final goods $c$ in a way that maximizes her utility$^1$:

$$V(p, M) = \max_{c, x, z} \tilde{U}(c, z)$$
subject to: \[ p(x) + z \leq M \]
\[ x \in X \]
\[ c \in C(x) \]

(2.2)

where $p(x)$ is the dollar amount to be paid by the consumer if she buys the volumes of inputs $x$, and $C(x)$ is the set of final goods that can possibly be produced from them. Clearly, the consumer makes the best use of her inputs by allocating them into the production of the optimal quantities of final goods. Thus, the value that the consumer can derive from input quantities $x$ and outside good quantity $z$ can be obtained as:

$$U(x, z) = \max_{c} \tilde{U}(c, z)$$
subject to: \[ c \in C(x) \]

(2.3)

$^1$We assume that it is not costly to produce final goods from inputs.
Under this definition of $U(x, z)$, the consumer problem in Equation 2.2 can be rewritten under the form of Equation 2.1. From the point of view of the consumer’s purchase behavior, our model is thus equivalent to the original problem considered in Equation 2.1. However, adding a production-consumption step allows us to parameterize $\tilde{U}(c, z)$ and $C(x)$ instead of $U(x, z)$ in the direct-utility approach, or $V(p, M)$ in the indirect-utility approach. The relationship between $\tilde{U}(c, z), U(x, z)$ and $V(p, M)$ is explained in Figure 1. In our approach, we specify a functional form on $\tilde{U}(c, z)$ and solve the problem in Equation 2.2 to derive the demand function $x^*(p, M)$. It should be noted that, if we assume $C(x)$ to be the identity correspondence ($C(x) = \{x\}$ $\forall x$, so that there is no difference between goods purchased and goods consumed), then Equation 2.3 trivially implies $U(x, z) = \tilde{U}(c, z)$. It is thus obvious that any parameterization of $U(x, z)$ in a direct-utility model can easily be accommodated under our approach by setting $C(x)$ to be the identity correspondence. In contrast, Equation 2.3 implies that a utility function over goods purchased $U(x, z)$ can always be inferred from a utility function over goods consumed $\tilde{U}(c, z)$; however, the derived expression of $U$ may not be straightforward. Since preferences are only revealed through the purchase of inputs in the absence of consumption data, only properties of $x^*(p, M)$ can be identified from the data. Yet, parameterizing $\tilde{U}(c, z)$ and $C(x)$ instead of $U(x, z)$ can be much more convenient as we show in the next sections, because it allows us to derive a flexible and regular demand function using very simple parametric forms.

![Figure 1: Relationship between $\tilde{U}$, $U$ and $V$ in the household-production model of complementarity](image-url)
2.3.2. Parametric assumptions

We represent the joint consumption of multiple goods as a separate final good, which is obtained by combining inputs in fixed proportions. We also allow each good to be consumed separately. Thus, the same input can be used to produce different final goods, and some final goods may require multiple inputs. We represent this by a Leontief production function that can be described by a $J \times K$ input-output table, denoted by $A$, such that $a_{jk}$ represents the volume of input $j$ that is required to make one unit of final good $k$. For example, the input-output table may look like the following in the case of burgers and buns:

$$A = \begin{pmatrix}
\text{sandwich} & \text{burg. only} & \text{bun only} \\
\text{burger} & 1 & 1 & 0 \\
\text{bun} & 1 & 0 & 1
\end{pmatrix}$$

The set $C(x)$ of final good quantities $c$ that can be produced from input volumes $x$ is the set of vectors $c$ with nonnegative entries and such that $\sum_k a_{jk}c_k \leq x_j$ for all $j$, or in matrix form: $Ac \leq x$.

Next, we parameterize the utility function for goods consumed as follows:

$$\tilde{U}(c, z) = \tilde{U}_c(c) + U_z(z)$$

with:

$$\tilde{U}_c(c) = \sum_{k=1}^K \alpha_k \log(c_k + 1)$$

$$U_z(z) = \alpha_0 z$$

where $\alpha_1, ..., \alpha_K$ and $\alpha_0$ are parameters whose values are nonnegative. This form is commonly used to parameterize the utility for goods purchased $U(x, z)$ in direct-utility models of demand (Satomura et al., 2011; Lee et al., 2013; Lee and Allenby, 2014). It has desirable properties: the log function captures monotonicity and concavity in its argument, and the intercept added inside the log allows for corner solutions. It should be noted that
this functional form is additively separable, which makes it easy to solve the consumer problem but implies that a model where this form is used to parameterize $U(x, z)$ cannot capture complementarity between the $J$ goods of interest. Since we use it to parameterize $\tilde{U}(c, z)$ instead, the resulting model can actually capture very flexible models of demand for complements, as we show in the next two sections.\(^2\)

2.3.3. Implications for consumer preferences over goods purchased

In this section, we show how our model of preferences over final goods translates into preferences over goods purchased. Using Equation 2.3 and Equations 2.5a-2.5c, our parameterization of $\tilde{U}(c, z)$, implies the following utility for purchased goods $U(x, z)$:

$$U(x, z) = U_x(x) + U_z(z)$$  \hspace{1cm} (2.6a)

where

$$U_x(x) = \max_c \sum_{k=1}^{K} \alpha_k \log(c_k + 1)$$  \hspace{1cm} (2.6b)

subject to:

$$\sum_{k} a_{jk} c_k \leq x_j \forall j$$  \hspace{1cm} (2.6c)

$$c_k \geq 0 \forall k$$  \hspace{1cm} (2.6d)

Equations 2.6b-2.6d define the consumer’s problem of optimally allocating inputs to construct final goods. This problem is mathematically very similar to the multiple-constraint model by Satomura et al. (2012), except that there is no outside good added for each linear constraint in Equation 2.6c. The first-order conditions of the optimal allocation problem

\(^2\)In previous literature, the utility for the outside good is sometimes specified as $U_z(z) = \alpha_0 \log(z)$, to allow for substitution between the goods purchased. While we could allow for substitution in a similar way, we assume a linear specification since our contribution is in the way we model complementarity between the goods purchased.
are as follows:

\[
\frac{\alpha_k}{c_k + 1} + \lambda_k - \sum_j \mu_j a_{jk} = 0 \quad \forall k
\]

\[
\lambda_k c_k = 0 \quad \forall k
\]

\[
\mu_j \left( x_j - \sum_k a_{jk} c_k \right) = 0 \quad \forall j
\]

\[
c_k, \mu_k, \lambda_j \geq 0
\]

where the coefficients \( \lambda_k \) and \( \mu_j \) are Lagrange multipliers corresponding to the non-negativity constraints on the quantities of final goods \( c_k \) and the constraints on the amount of inputs available, respectively. In a case with \( J = 2 \) inputs and \( K = 3 \) final goods such that

\[ A = \begin{pmatrix}
1 & 0 & 1 \\
0 & 1 & a_{23}
\end{pmatrix} \]

we can easily obtain the expression of the optimal quantities of final goods consumed as a function of the quantities of goods purchased:

\[
c_3^*(x) = \min \left( \max \left( 0, c_{3}^{(int)}(x) \right), x_1, \frac{x_2}{a_{23}} \right)
\]

\[
c_1^*(x) = x_1 - c_3^*
\]

\[
c_2^*(x) = x_2 - a_{23} c_3^*
\]

where \( c_{3}^{(int)} \) is the solution to a second-degree equation, as shown in Appendix A.1. The resulting utility over inputs purchased is then \( U_x(x) = U_c(c^*(x)). \)

Using the expression of \( U_x \), we illustrate in Figure 2 several patterns of indifference curves obtained from our model with different parameter values, holding the outside good constant and focusing on \( U_x \). As we can see, our model is very flexible: it can accommodate very different patterns of preferences from perfect complementarity to no complementarity, with symmetry or asymmetry.

---

3One may argue that we could have directly parameterized the utility for inputs purchased \( U_x(x) \) using that expression, without laying out the household production problem. While it is true that the resulting demand system would be the same, we emphasize that the derived expression of \( U_x(x) \) is not straightforward and cannot be simplified, while our parameterization of \( U_c(c) \) has a closed-form expression. Thus setting up the household production problem greatly facilitates the problem of specifying a convenient functional form.
Figure 2: Indifference curves under various parameter values

The above plots represent indifference curves over two goods purchased, holding the quantity of outside good constant. Different values of the parameters give very different patterns, from no complementarity (subplot a) to perfect complementarity (subplot d), with symmetric complementarity (subplot b) or asymmetric complementarity; (subplot c)).
2.3.4. Implications for demand under continuous quantities

In this section, we study the properties of the demand function \( x^*(p, M) \) under the assumption of linear prices (such that \( p(x) = \sum_j p_j x_j \)), and assuming that the \( J \) inputs can be purchased in any nonnegative, continuous quantities (such that \( \mathcal{X} = \mathbb{R}_+^J \)). In this case, the demand function can be solved easily and we can obtain intuitive insights into its properties. Since any continuous quantities of inputs can be purchased, the consumer will buy the quantities that are exactly necessary to make the optimal quantities \( c^*_k \) of final goods: buying more inputs would only come at a monetary cost. Thus the demand for inputs must be such that:

\[
    x^*_j = \sum_{k=1}^{K} a_{kj} c^*_k \quad \forall j
\]  

Furthermore, we can define the full price \( f_k \) of each final good \( k \) as the dollar amount that the consumer needs to pay to produce one unit of that final good, by buying all the necessary inputs:

\[
    f_k = \sum_{j=1}^{J} a_{kj} p_j
\]

By using this definition, the consumer problem in Equation 2.2 can be simplified and rewritten in terms of the final goods only:

\[
    \max_{c,z} \quad \tilde{U}(c, z) \\
    \text{s.t.} \quad \sum_{k=1}^{K} f_k c_k + z \leq M
\]

The first-order conditions can then be solved to obtain the optimal quantities of final goods under our parameterization of \( \tilde{U}(c, z) \):

\[
    c^*_k = \max \left( 0, \frac{\alpha_k}{f_k} - 1 \right)
\]
The demand function is obtained by combining Equations 2.9 and 2.3.4. We can then use these results to study the effect of input prices on demand for the $J$ inputs:

$$\frac{\partial x^*_j}{\partial p_{j'}} = \frac{\partial}{\partial p_{j'}} \left( \sum_{k=1}^{K} a_{kj} c^*_k \right)$$

$$= \sum_k a_{kj} \frac{\partial c^*_k}{\partial p_{j'}}$$

$$= \sum_k a_{kj} \frac{\partial c^*_k}{\partial f_k} \frac{\partial f_k}{\partial p_{j'}}$$

$$= \sum_k a_{kj} a_{kj'} \frac{\partial c^*_k}{\partial f_k}$$

$$= -\sum_k a_{kj} a_{kj'} \alpha_k \frac{\partial c^*_k}{\partial f_k} I(p_k < \alpha_k)$$

(2.13)

where $I(.)$ is the indicator function. If $j = j'$, then clearly the derivative is negative, which implies that own price effects are negative. If $A$ is the identity matrix (which means that $C(x)$ is the identity correspondence, and there is no final good representing a joint consumption), then the cross price effects are zero, since $a_{jj'} = 0$ if $j \neq j'$. However, if it is not the case, then we should expect negative cross price effects when two goods $j$ and $j'$ are jointly used as inputs to produce at least one final good $k$, such that $a_{kj} > 0$ and $a_{kj'} > 0$. Thus, by specifying a consumer problem in the space of final goods and then aggregating demand across these final goods to obtain the demand for inputs, we capture complementarity between the goods purchased.

Let us consider again the case with $J = 2$ inside goods and $K = 3$ final goods, where the first two final goods correspond to the separate consumption of each inside good and the third one is a composite final good that corresponds to their joint consumption. Let us assume that the composite final good is obtained by combining the inputs in the same proportions, such that $a_{23} = 1$. When the utility parameter for the composite final good is positive and the utility parameters for the separate consumptions are low ($\alpha_3 > 0; \alpha_1 < p_1$; $\alpha_2 < p_2$, as in Figure 2.d), Equation implies that no quantity of the inputs is purchased to be consumed separately: $c^*_1 = c^*_2 = 0$. Thus the inputs are only purchased to produce the composite final
good, and the consumer simply considers the “full price” of the composite final good and
decides what quantity $c_3$ to produce and consume of it. In that case, a one-dollar increase
in the price of one input leads to a one-dollar increase in the full price of the composite final
good, which may lead the consumer to reduce the quantity $c_3^*$ of composite final good and
therefore the quantities of both inputs purchased by the same amount. Mathematically,
Equation 2.13 implies that $\frac{\partial x_1^*}{\partial p_1} = \frac{\partial x_2^*}{\partial p_2} = \frac{\partial x_3^*}{\partial p_1} = \frac{\partial x_3^*}{\partial p_2} = -\frac{\alpha_3}{f_3}$ in that case. Thus a one-dollar
price change of each input should have the same effect on demand across all inputs: the
own- and cross- price effects should be the same and should all be equal across the $J$ goods.

Alternatively, if the consumer has positive utilities for some of the inputs consumed sep-
arrately ($\alpha_1 > p_1$ and/or $\alpha_2 > p_2$), then the own- price effects and the cross- price effects
may be different because some of the inputs may be consumed by themselves. In that case,
the price of good 1 impacts the demand for good 2 only to the extent that it impacts their
joint consumption (whose utility is captured by $\alpha_3$). On the other hand, the price of good 2
impacts the demand for good 2 both because it impacts the consumption of good 2 by itself
(whose utility is captured by $\alpha_2$) and the joint consumption with good 1. If the separate
consumption of the inputs give different utilities (such that $\alpha_1 \neq \alpha_2$), then the cross- price
effects may be asymmetric. Thus, our model of household production provides an intuitive
explanation for the existence of asymmetric cross- price effects between goods.

2.4. Econometric model

In this section, we describe how we apply our model to data. We first introduce stochastic
elements to allow for consumer heterogeneity and variation of preferences over time. Then
we focus on the estimation of the model of demand under the assumption of continuous
demand. Next, we relax the assumption of infinitely divisible demand. Finally, we discuss
the identification or our model.
2.4.1. Distributional assumptions

We now introduce the subscript $i$ to refer to a consumer, and the subscript $t$ to indicate a specific purchase occasion. Suppose that we follow $N$ consumers, each over $T_i$ purchase occasions: we observe the prices $p_{it}$ that they face, and the volumes of goods $x_{ijt}$ that they purchase. We take a random-utility approach to rationalize these purchase decisions under the consumer problem in Equation 2.2, by including random shocks in the consumer’s utility function. Specifically, we assume that the $\alpha$ parameters of consumer $i$ at time $t$ are such that:

$$\alpha_{ikt} = \alpha_{ik} e^{\epsilon_{ikt}}$$

where $\epsilon_{ikt} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$ (2.14)

The values $\alpha_{ik}$ represents the stable part of the consumer’s preferences and the shocks $\epsilon_{ikt}$ capture variation over time. The model only makes sense if the preference parameters $\alpha_{ik}$ are non-negative: thus we consider their log. We collect all individual-level preference parameters into a vector $\omega^i$, and model unobserved consumer heterogeneity through random coefficients:

$$\omega^i \overset{iid}{\sim} \text{MVN}(\bar{\omega}, V)$$

where $\omega^i = [\log(\alpha_{i1}), ..., \log(\alpha_{ik})]$ (2.15)

The resulting model is a hierarchical Bayesian model where consumer $i$’s contribution to the likelihood can be written as follows:

$$L_i(\omega^i, \sigma^2) = \prod_{t=1}^{T_i} \int \phi \left( \frac{1}{\sigma} \epsilon_{ikt} \right) \left[ \prod_k \phi \left( \frac{1}{\sigma} \epsilon_{ikt} \right) \right] I \{ x^*(\omega^i, \epsilon_{it}|p_{it}) = x_{it} \} d\epsilon_{it}$$

where $I(.)$ is the indicator function, $\epsilon_{it}$ is a collection of all the shocks $\epsilon_{ikt}$, $x^*(\omega^i, \epsilon_{it}|p_{it})$ is the optimal demand under price function $p_{it}$ and shocks $\epsilon_{it}$, and $\phi$ is the pdf of the standard Normal distribution. The difficult part in evaluating the likelihood consists in computing the integral, which does not have a closed-form expression. Instead, we rely on data augmentation to estimate the model (Tanner and Wong, 1987), as we discuss in the next two sections.
2.4.2. Model with continuous quantities

In this section, we discuss the estimation of the model considered in Section 2.3.4, where each good \( j \) has a unit price equal to \( p_j \) and can be purchased in any nonnegative, continuous quantity. As explained earlier, we can define a full price \( f_k \) for one unit of each final good \( k \) using Equation 2.11. Let us suppose for a moment that we can observe the quantities of final goods \( c_{ikt} \) produced by consumer \( i \) at time \( t \). Then, from Equations 2.3.4 and 2.14, we can derive either an upper bound or an exact value for the random utility shocks:

\[
\epsilon_{ikt} \begin{cases} 
\leq \log \left( \frac{f_{ikt}}{\alpha_{ik}} \right) & \text{if } c_{ikt} = 0 \\
= \log \left( \frac{f_{ikt}}{\alpha_{ik}} \times (c_{ikt} + 1) \right) & \text{if } c_{ikt} > 0
\end{cases}
\] (2.17)

We can use this result to compute the probability that the quantities of final goods \( c_{ikt} \) are optimal under preference parameters \( \omega^i \) and prices \( p_{it} \):

\[
Pr(c) = \left[ \prod_{k \text{ s.t. } c_k = 0} \Phi \left( \frac{1}{\sigma} \log \left( \frac{f_{ikt}}{\alpha_{ik}} \right) \right) \right] \left[ \prod_{k \text{ s.t. } c_k > 0} \frac{1}{\sigma} \phi \left( \frac{1}{\sigma} \log \left( \frac{f_{ikt}}{\alpha_{ik}} \times (c_k + 1) \right) \right) \right] \frac{1}{c_k + 1}
\] (2.18)

where \( \Phi \) is the cdf of the standard Normal distribution, and the term \( \frac{1}{c_k + 1} \) on the right side is a Jacobian due to the transformation from the random shocks \( \epsilon \) to the quantities of final goods \( c \). Since consumption is unobserved, we need to integrate this probability over the possible set of final good quantities, conditional on the observed quantities purchased \( x_j \). From Equation 2.9, we must have \( x_j = \sum_k a_{jk} c_k \), or in matrix notation: \( x = Ac \) where \( c_k \geq 0 \) for all \( k \). Thus, we have an aggregate data problem: we have a complete likelihood defined over the consumption of final goods \( c_k \) (Equation 2.18), but we only observe the volumes of inputs purchased \( x_j \), which are an aggregation of those final goods. We can rewrite consumer \( i \)'s contribution to the likelihood as:

\[
L_i(\omega^i, \sigma^2) = \prod_{t=1}^{T_i} \int_{c \geq 0 \text{ s.t. } Ac = x_{it}} Pr\left( c^* (\omega^i, p_{it}) = c \right) dc_1...dc_K
\] (2.19)
In the case of two inputs and three final goods, we only need to integrate out $c_3$ since $c_1$ and $c_2$ can be obtained directly from $c_3$ and $x$. Still, this integration is not straightforward. To estimate the model, we apply data augmentation (Tanner and Wong, 1987), treating $c_3$ as the augmented data.

2.4.3. Model with indivisible demand

We now relax the assumptions of continuous purchases, and consider a situation in which each good is sold as a package containing a volume $s_j$, with a unit price equal to $p_j$. The consumer can then choose to buy any integer number $q_j$ of packs for good $j$. Under these assumptions, we have:

$$q_j \in \{0, 1, 2, \ldots\} \forall j \quad (2.20a)$$

$$x_j = q_j s_j \quad (2.20b)$$

$$p(x) = \sum_j q_j p_j \quad (2.20c)$$

This new setup allows us to accommodate the indivisibility of demand, as the possible quantities purchased by the consumer lie on a grid of discrete points defined by the integer constraints on the quantities of packs $q_j$. The consumer problem 2.2 can no longer be simplified to obtain a closed-form expression for the demand $x^*$, like it was the case under continuous quantities. We must then evaluate the consumer’s objective function on all such points (assuming a reasonable upper bound on the number of packs $q_j$), and find that point which yields the highest value to obtain the demand $x^*$. Furthermore, the integral in Equation 2.16 cannot be simplified, thus we apply data augmentation to estimate the model, by treating $z_{ikt} = \log(\alpha_{ik}) + \epsilon_{ikt}$ as the augmented data. Our approach is similar to that used by Lee and Allenby (2014); the estimation algorithm is laid out in Appendix A.2.
2.4.4. Identification

We discuss how the parameters of our model can be identified from patterns of purchases as a function of prices, in the absence of consumption data. We assume that the consumer’s budget $M$ is large enough so that it does not restrict the amount of inputs purchased but only the amount of outside good; consequently, it drops out of the consumer problem.\footnote{This is an assumption commonly made in the literature (see Lee et al., 2013, for example).}

The $\alpha_k$ parameters capture the consumer’s utility for final goods: as such, they are defined relative to the marginal utility for the outside good $\alpha_0$, and we fix their scale by setting $\alpha_0 = 1$. Different values of $\alpha_k$ yield different own- and cross-price effects, as discussed in Section 2.3.4. Thus they are identified by the own- and cross-effects of prices on purchases; the magnitude of the price effects are smaller when the $\alpha_k$ parameters are larger. Separate from cross-price effects, correlation in purchases is captured through the random shocks $\epsilon$.

In the analyses that follow, we consider a case with $J = 2$ inputs and $K = 3$ final goods, with the assumption that $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$.\footnote{The proportions used in combining both inputs for a joint consumption, captured by $a_{23}$, can be treated as a parameter to estimate. In theory, it can be identified through the proportions in which goods are purchased but in practice, it may be difficult to identify it when there is not much variation in quantities purchased, which is the case in our empirical application. Future research should consider the consequences of estimating $a_{23}$.}

The variation in purchase behavior at the individual level can be disentangled from consumer heterogeneity since we have panel data with a time series of observations for each consumer.

2.5. Simulation study

We ran a simulation study to investigate how the assumption of continuous demand may affect the parameter estimates of our model and the inferences drawn regarding the effect of prices on demand. We also sought to investigate the possible consequences of ignoring (or allowing for) complementarity between goods when estimating demand. For this analysis, we considered three different versions of our proposed model. Model 1 is a restricted version of our discrete model where we shut off any complementarity between goods by removing the possibility of joint consumption: in this model, the final goods are equivalent to the
inputs purchased, thus \( K = J \) and \( A \) is the identity matrix. Model 2 is our discrete model that allows for complementarity, as laid out in Section 2.4.3. Model 3 is the continuous version of our model, as laid out in Section 2.4.2. For each model, we used the hierarchical Bayesian paradigm to model consumer heterogeneity.

We generated a dataset for each of the three models, using that model as the true data generating process. We defined a set of \( J = 2 \) goods purchased, each good being sold in a pack of size one under Model 1 and Model 2, or in any continuous quantities under Model 3. We generated the preference parameters of 200 consumers from a multivariate normal distribution with mean \( \bar{\omega} \) and covariance matrix \( V \), where \( \bar{\omega} \) and \( V \) were drawn in a way that yields a reasonable amount of variation in purchase behavior.\(^6\) Unit prices were drawn from a uniform distribution in interval \([0.2, 1.2] \). For each consumer, we generated 100 purchase occasions, and solved for the optimal purchase quantities using either a grid search (for Models 1 and 2) or the first-order conditions (for Model 3). Each resulting dataset contained the prices observed by consumers and the quantities purchased of each good. For each consumer, we used the first 100 purchase occasions for model estimation, and the other 100 purchase occasions for out-of-sample prediction.

\[^6\text{For completeness, } \bar{\omega}_k \text{ was drawn as } \log(b_k u_k), \text{ where } u_k \text{ was a standard uniform variable between 0 and 1, while } b_1, b_2, b_3 \text{ were set to 2, 2 and 3 respectively. } V \text{ was drawn as a diagonal matrix whose diagonal elements were drawn from a standard uniform distribution.}\]
<table>
<thead>
<tr>
<th>Model 1 (no complementarity)</th>
<th>Model 2 (discrete)</th>
<th>Model 3 (continuous)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truth</td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>$\bar{\omega}_1$ (good1)</td>
<td>0.97 (0.02)</td>
<td>0.95 (0.02)</td>
</tr>
<tr>
<td>$\bar{\omega}_2$ (good2)</td>
<td>0.72 (0.05)</td>
<td>0.68 (0.05)</td>
</tr>
<tr>
<td>$\bar{\omega}_3$ (joint)</td>
<td>0.05 (0.01)</td>
<td>0.08 (0.01)</td>
</tr>
<tr>
<td>$V_{11}$</td>
<td>0.00 (0.01)</td>
<td>0.01 (0.01)</td>
</tr>
<tr>
<td>$V_{21}$</td>
<td>0.51 (0.04)</td>
<td>0.44 (0.04)</td>
</tr>
<tr>
<td>$V_{31}$</td>
<td>0.00 (0.01)</td>
<td>0.01 (0.01)</td>
</tr>
<tr>
<td>$V_{32}$</td>
<td>0.01 (0.10)</td>
<td>0.22 (0.14)</td>
</tr>
<tr>
<td>$V_{33}$</td>
<td>0.91 (0.16)</td>
<td>5.06 (1.57)</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>0.19 (0.01)</td>
<td>0.19 (0.01)</td>
</tr>
</tbody>
</table>

Table 1: Results of a simulation study under the household-production model of complementarity
We estimated the three models on the datasets generated under Model 1 and Model 2. While demand is discrete in these datasets, Model 3 disregards package size constraints and assumes that the quantities purchased are exactly the continuous quantities desired by consumers. Reversely, Models 1 and 2 cannot accommodate the continuous quantities present in the dataset generated under Model 3, therefore we only applied Model 3 on that dataset, with the objective of verifying that we can recover the true parameters. The estimation results are compiled in Table 1. The first set of rows indicates our estimates of the average preference parameters $\bar{\omega}$ along with their posterior standard error in parentheses. The second set of rows gives in-sample and out-of-sample model fit criteria: for each good, the hit-rate of predicted purchase incidence and mean squared error of volume purchased averaged across all purchase occasions and consumers. The last set of rows gives an estimate of own- and cross-price elasticities, obtained by simulating the aggregate demand under a 10% price increase for each good.

We can make several observations from these results. First, we recovered the true parameters within their 95% posterior intervals for each dataset when estimating the “true” model, and the “true” model made predictions that were better or at least as good as the other models, both in-sample and out-of-sample. While expected, these results confirm that the parameters of our model can be empirically identified and that our estimation procedure can recover them.

The estimation results also provide insights about the effect of capturing or ignoring complementarity between the goods. The first dataset was generated according to Model 1, which does not have any complementarity between goods purchased. However, it should be noted that Model 3, which allows for such complementarity, accurately estimates the true preference parameters for separate consumptions $\bar{\omega}_1$ and $\bar{\omega}_2$ and yields a very low estimate for the preference parameter of the joint consumption $\bar{\omega}_3$. With such a low preference for joint consumptions, the consumer prefers to consume the goods separately rather than jointly, and thus the demand system can be reduced to the demand system of Model 1, where the
joint consumption does not exist. Thus the model yields accurate own-price elasticities, and cross-price elasticities equal to zero. Reversely, the second dataset was generated under the assumption of a joint consumption having a relatively high utility compared to the separate consumption of the goods, as indicated by the high value of $\bar{\omega}_3$. This leads to the existence of negative cross-price effects, as shown by the sign of the true cross-price elasticities. On that dataset, the model that does not allow for complementarity (Model 1) leads to a bias in the estimation of own-price elasticities. Taken together, these results suggest that if there is no complementarity, the model that allows for complementarity yields correct estimates of own- and cross-price effects, but if there exist complementarity, then a model that ignores it not only yields cross-price elasticities equal to zero, but also a biased estimate of own-price elasticities.

We can also get some insights about the importance of accounting for the indivisibility of demand by looking at the estimation results of the continuous model (Model 3) on the second dataset, which is generated according to the discrete model (Model 2). The only difference between the two models is that Model 3 does not take into account the discreteness implied by package size constraints. As we can see, ignoring indivisibility leads to an underestimation of all preference parameters. This result is consistent with the article by Lee and Allenby (2014): in a model of continuous demand, consumers should buy a positive quantity of each good (even very small) unless the good gives a smaller marginal utility than her marginal utility of money: consequently, such a model rationalizes a low purchase incidence rate by a very small marginal utility for the good when in fact, the low purchase incidence rate may due to large packages. In our simulation, the own- and cross-price elasticities we found were still similar under the continuous and the discrete model, which may be explained by the relatively small package size we assumed.

2.6. Empirical application

This section discusses an application of our model on a real dataset from a panel of consumers reporting their purchases of tortilla chips and Mexican salsa over time. Through
this analysis, we sought to evaluate the fit of our model against a set of competing models, and to show how the results can be used to help make better managerial decisions. We first describe the data, then we present our estimation results, and finally we illustrate the usefulness of our model through counterfactual analyses studying the effect of different types of coupons.

2.6.1. Data

The data was collected by AC Nielsen and is made of two parts. The first part contains data about a set of households who report their purchases over time using a scanner device at home: it provides the information about all shopping trips made by the households, the number of units purchased and the price paid for each item purchased. Prices are only observable when a purchase is made, therefore we combine the data about households with the second part of the data, which gives us store-level prices for each item each week.

We chose the tortilla chips and Mexican salsa categories for our empirical application, and focused on a two-year time window from 2010 to 2011. We focused respectively on the 10-ounce and 16-ounce package sizes, which are the predominant formats in these categories. We aggregated prices at the category level by taking an average of brand-specific prices weighed by market share. We restricted the data to households making at least 8 purchases in each of the two categories over the two-year period.

The resulting dataset contained 251 consumers, each consumer making 151 shopping trips on average. Table 2 gives some summary statistics of the data. Consumers bought multiple units of chips or salsa in 30% of their purchases, which indicates the need for a model of quantity. Consumers often bought both goods together, as indicated by the high percentage of purchase co-incidence relative to the marginal percentages of incidence. We ran logit models of incidence for each category to investigate the effect of prices on demand, with and without household-specific fixed effects. The results are displayed in Table 2 and suggest that prices have a negative effect on demand within and across categories. They also suggest the existence of heterogeneity across households.
(a) Purchase frequency

<table>
<thead>
<tr>
<th></th>
<th>Salsa and chips</th>
<th>Salsa only</th>
<th>Chips only</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of observations (%)</td>
<td>1187 (3.1%)</td>
<td>2218 (5.9%)</td>
<td>2878 (7.6%)</td>
<td>31625 (83.4%)</td>
</tr>
</tbody>
</table>

(b) Descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>Salsa</th>
<th>Chips</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price ($)</td>
<td>2.70</td>
<td>2.60</td>
</tr>
<tr>
<td>Purchase incidence (%)</td>
<td>8.98</td>
<td>10.72</td>
</tr>
<tr>
<td>Mean purchase quantity</td>
<td>1.34</td>
<td>1.32</td>
</tr>
</tbody>
</table>

Table 2: Description of purchase data in the tortilla chips and Mexican salsa categories

<table>
<thead>
<tr>
<th></th>
<th>Incidence of salsa</th>
<th>Incidence of salsa (with household FE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salsa price</td>
<td>-0.4381 (0.0905)</td>
<td>-0.6294 (0.1206)</td>
</tr>
<tr>
<td>Chips price</td>
<td>-0.2934 (0.0552)</td>
<td>-0.2797 (0.0782)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Incidence of chips</th>
<th>Incidence of chips (with household FE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salsa price</td>
<td>-0.2572 (0.0847)</td>
<td>0.0192 (0.1171)</td>
</tr>
<tr>
<td>Chips price</td>
<td>-0.7411 (0.0524)</td>
<td>-0.9065 (0.0749)</td>
</tr>
</tbody>
</table>

Table 3: Reduced-form evidence of complementarity between tortilla chips and Mexican salsa

2.6.2. Estimation results

We estimated several models on the data: the discrete model that rules out complementarity (Model 1), our discrete model of complementary demand based on household production theory (Model 2) and the continuous version of that model (Model 3). For each model, we used a Bayesian estimation procedure to generate 500,000 draws of the parameters from their posterior distribution. We discarded the first 100,000 draws and kept one draw every 100 draws thereafter to reduce autocorrelation. We ran three separate chains to assess convergence.

Comparing model fits is difficult because the Deviance Information Criterion (DIC), which is usually used in model selection between hierarchical Bayesian models, requires one to evaluate the likelihood at the average posterior draw of all parameters. Since we also draw the random shocks as augmented data, we would need to evaluate the likelihood at the average of the random shocks, individual- and population-level preference parameters across posterior draws to compute the DIC. The posterior random shocks are such that
<table>
<thead>
<tr>
<th></th>
<th>Model 1 (no complementarity)</th>
<th>Model 2 (discrete)</th>
<th>Model 3 (continuous)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega_1$ (salsa)</td>
<td>-2.15 (0.05)</td>
<td>-2.45 (0.05)</td>
<td>-7.82 (0.13)</td>
</tr>
<tr>
<td>$\omega_2$ (chips)</td>
<td>-1.92 (0.05)</td>
<td>-2.23 (0.06)</td>
<td>-7.17 (0.15)</td>
</tr>
<tr>
<td>$\omega_3$ (joint)</td>
<td>-</td>
<td>-2.95 (0.08)</td>
<td>-17.76 (1.65)</td>
</tr>
<tr>
<td>$V_{11}$</td>
<td>0.32 (0.03)</td>
<td>0.35 (0.04)</td>
<td>2.28 (0.25)</td>
</tr>
<tr>
<td>$V_{21}$</td>
<td>0.26 (0.03)</td>
<td>0.25 (0.04)</td>
<td>1.98 (0.25)</td>
</tr>
<tr>
<td>$V_{22}$</td>
<td>0.47 (0.05)</td>
<td>0.53 (0.06)</td>
<td>3.39 (0.36)</td>
</tr>
<tr>
<td>$V_{31}$</td>
<td>-</td>
<td>0.10 (0.04)</td>
<td>5.05 (1.38)</td>
</tr>
<tr>
<td>$V_{32}$</td>
<td>-</td>
<td>0.15 (0.05)</td>
<td>5.11 (1.98)</td>
</tr>
<tr>
<td>$V_{33}$</td>
<td>-</td>
<td>0.66 (0.10)</td>
<td>26.78 (10.07)</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>2.66 (0.06)</td>
<td>2.87 (0.07)</td>
<td>22.81 (0.50)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Model 1 (no complementarity)</th>
<th>Model 2 (discrete)</th>
<th>Model 3 (continuous)</th>
</tr>
</thead>
<tbody>
<tr>
<td>In-sample hit-rate for salsa</td>
<td>0.91</td>
<td>0.91</td>
<td>0.91</td>
</tr>
<tr>
<td>In-sample hit-rate for chips</td>
<td>0.90</td>
<td>0.90</td>
<td>0.90</td>
</tr>
<tr>
<td>In-sample volume MSE for salsa</td>
<td>44.16</td>
<td>44.11</td>
<td>5.5e+08</td>
</tr>
<tr>
<td>In-sample volume MSE for chips</td>
<td>18.52</td>
<td>18.51</td>
<td>1.0e+09</td>
</tr>
<tr>
<td>Out-of-sample hit-rate for salsa</td>
<td>0.90</td>
<td>0.90</td>
<td>0.90</td>
</tr>
<tr>
<td>Out-of-sample hit-rate for chips</td>
<td>0.87</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>Out-of-sample volume MSE for salsa</td>
<td>49.22</td>
<td>49.16</td>
<td>4.9e+07</td>
</tr>
<tr>
<td>Out-of-sample volume MSE for chips</td>
<td>24.21</td>
<td>24.28</td>
<td>8.3e+07</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Model 1 (no complementarity)</th>
<th>Model 2 (discrete)</th>
<th>Model 3 (continuous)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasticity salsa→salsa</td>
<td>-1.11</td>
<td>-1.03</td>
<td>-0.93</td>
</tr>
<tr>
<td>Elasticity salsa→chips</td>
<td>0.00</td>
<td>-0.11</td>
<td>-0.02</td>
</tr>
<tr>
<td>Elasticity chips→salsa</td>
<td>0.00</td>
<td>-0.12</td>
<td>-0.02</td>
</tr>
<tr>
<td>Elasticity chips→chips</td>
<td>-1.11</td>
<td>-1.07</td>
<td>-0.96</td>
</tr>
</tbody>
</table>

Table 4: Estimation results on the tortilla chips and Mexican salsa purchase data
the observed actions are optimal, which implies that the likelihood of an observed action is equal to one conditional on posterior random shocks and posterior preference parameters. However, this property may no longer be true when averaging across draws, leading to a likelihood equal to 0 at the “average draw” and thus a DIC equal to minus infinity.

Instead of likelihood-based criteria, we relied on in-sample and out-of-sample predictions to compare model fits. Specifically, we predicted purchase incidence and volumes purchased (in ounces) for each consumer trip, for each of the two goods. To make predictions, we drew random shocks from their prior distribution and determined the optimal purchase action conditional on our posterior draws of the consumer’s preference parameters and the prior random shocks drawn. We predicted a purchase incidence when the probability of a purchase was at least 50%. The hit-rate was obtained as the percentage of correct incidence predictions. It should be noted that the actual incidence rates are quite low in the data, and therefore it is easy to obtain a high hit-rate by simply predicting no incidence at every shopping trip. For that reason, we also report the mean squared error (MSE) of our volume predictions.

In Table 4, we have compiled for each model, our estimates of the population-level mean parameters along with the corresponding posterior standard deviation in parentheses. We have also reported the results of our predictions, as well as own- and cross-price elasticities. Price elasticities were computed by simulating a 10% increase in the price of each good and determining the aggregate demand under each alternative scenario, based on the posterior draws of the random shocks and of the individual-level preference parameters.

The continuous model’s predictions are clearly worse than those obtained from the other two models, as the high estimated variance for the random shocks leads to very large predictions. Compared to the discrete model, the average preference parameters \( \bar{\omega}_k \) are again under-estimated in the continuous model. The continuous model also yields much smaller cross-price elasticities, which is particularly interesting since most multicategory models in previous research disregard the discreteness of demand. We emphasize that this downward
bias may lead to suboptimal marketing decisions from retailers and manufacturers.

Our model and the no-complementarity model yield similar fits in terms of predictions both in-sample and out-of-sample. However, the no-complementarity model cannot explain the reduced-form evidence of cross-price effects shown in Section 2.6.1. Furthermore, the estimates of preference parameters $\bar{\omega}_1$ and $\bar{\omega}_2$ are significantly different across the two models, unlike what we obtained in the simulation study described in Section 2.5, where our discrete model recovered the true values of $\bar{\omega}_1$ and $\bar{\omega}_2$ and yielded a very low value for $\bar{\omega}_3$ when data was generated according to the no-complementarity model. This suggests that our discrete model captures complementarity that is present in the data. The own-price elasticities obtained are somewhat similar across the two models but our discrete model yields non-negligible cross-price elasticities, which could lead to different price and promotion decisions across the two categories.

2.6.3. Counterfactuals

We now illustrate how our model can be used to inform managerial decisions in the context of category-level coupons. Retailers commonly use these coupons by which they offer consumers a discount if they buy an item within a product category or a set of product categories. When a coupon can be used in multiple categories, the dollar amount paid by the consumer for her purchases in one category may not be independent of her purchases in another category, because the coupon can only be used once. For example, suppose that a consumer has a one-dollar coupon that is valid for either chips or salsa. If she only buys a $3-jar of salsa, she has to pay $2. If she only buys a $4-bag of chips, she has to pay $3. But if she buys both, then she pays $6: the price paid is not additive across categories, and we must then consider the joint purchase decisions made across categories. Our model is well-suited to simulate demand under this type of scenarios because we can use the parameters of the consumers’ utility function to solve the optimal purchase problem with non-additive prices.
We ran a counterfactual analysis to study the effect of different coupon schemes on retailer revenue. We considered three coupon types: a single-category coupon valid for chips only, a single-category valid for salsa only, and a multi-category coupon valid for either chips or salsa. For each coupon type, we simulated an alternative scenario in which a one-dollar coupon was available to consumers at each of their shopping trips. Based on the estimates we obtained for our discrete model (Model 2), we determined the optimal purchase decision made by consumers based on each posterior draw of their preference parameters and utility shocks, assuming that they would redeem their coupon if it is applicable. Then, we averaged demand and dollar expenses across draws and summed it across consumers. In Table 5, we have compiled the revenue generated by the retailer, the number of packs sold and the number of purchase incidences for each category, under each scenario.

We can make several observations from these results. First, offering single-category coupons increases demand within each category through a higher number of purchase incidences; to a lesser extent, it also increases demand for the complementary category due to complementarity between them. The increase in demand generated by the salsa coupon results in an increase in revenue, which is not obvious since a one-dollar rebate is given away for each purchase incidence of salsa that would be made even when no coupon were offered. On the other hand, the increase in demand generated by the chips coupon does not compensate the loss in revenue from consumers who would buy chips without having a coupon. Second, multicategory coupons considerably increase demand for both categories compared to the baseline scenario, reaching a level that is slightly lower than in each single-category coupon scenario. However, multicategory coupons yield a lower revenue because of the revenue lost
on already existing purchases.

Which couponing strategies is the most profitable to retailers depends on the wholesale prices paid by retailers, since an increase in demand imply an increase in procurement costs. Managers can use our model to simulate demand, revenue and costs under alternative couponing scenarios and select the most profitable strategy. Multicategory coupons, if targeted, may be an attractive strategy for retailers since they can increase demand across categories. Future research should study the importance of targeting consumers when offering coupons in complementary categories.

2.7. Conclusion

In this chapter, we have proposed a general approach to modeling micro-level demand for complementary goods by invoking household production theory. Under our approach, two goods are complementary if the consumer enjoys utility from their joint consumption, represented by a final good that she can produce from them. By assumption, she decides not only what goods to purchase, but also how to allocate them into final goods that represent consumption uses, which may be joint or separate. Therefore, the consumption utility is defined over a set of latent final goods which is larger than the set of goods purchased. The resulting demand system exhibits negative cross price effects under a linear budget constraint. It also provides an intuitive explanation for the existence of asymmetric cross price effects because of differences in utility for separate consumptions. From the point of view of purchase behavior, the new consumer problem can be reduced to the usual problem without household production, by redefining the consumer’s utility over goods purchased after recognizing that the allocation of goods purchased into consumption uses is optimal. However, laying out the consumer problem with a household production step is useful for empirical work even if consumption is unobserved, because it opens up the possibility of parameterizing the consumption utility over the set of final goods which is larger than the set of goods purchased. We can then use an additively separable specification and still accommodate very different patterns of preferences, from no complementarity to perfect
complementarity, with or without asymmetry. Thus, our approach resolves the problem of specifying a flexible functional form to capture complementarity in a direct-utility model of demand. We allow for randomness in consumer behavior through a random-coefficient approach, similar to previous direct-utility models.

Estimating the model in the absence of consumption data requires one to solve an aggregate data problem, where the observed quantities purchased are an aggregation of unobserved quantities consumed. Overall, the model has desirable properties, yields flexible patterns with very simple functional forms, and can be easily estimated on purchase data using Bayesian methods. Through a simulation study and an empirical application on purchase data from a set of consumers, we have demonstrated the usefulness of the approach in modeling demand for complements, especially in contexts where demand is indivisible because of package size constraints, and where the prices paid are not additive across goods.

The approach we propose opens new ideas for future research. First, future research should examine the scalability of our approach in a context with more than two product categories. Second, while we have focused on complementarity in this chapter, we believe that the approach can be extended to obtain a flexible demand system for substitutes. Doing so would require a change in the specification of the consumer’s production function, such that two goods are substitutable if one or the other can be used to fulfill a consumption need, represented by a final good. Third, the approach can also be used in a model of forward-looking behavior, since our specification of the consumption utility can be used as the reward function of a dynamic programming problem. In the next chapter, we develop such a model to estimate demand for storable complements.
3.1. Motivation

Many pairs of goods are typically consumed together, such as crackers and dips, laundry detergent and fabric softener, or spaghetti and tomato sauce. For manufacturers of such goods (such as Tostitos, Proctor & Gamble, and Barilla), and retailers who sell the goods to end-consumers, it is crucial to understand how prices and promotions impact the purchases made by consumers across categories, since their overall profit depends on sales across them. It is therefore not surprising that cross-category demand estimation has been an active area of research in marketing over the last two decades (Manchanda et al., 1999; Song and Chintagunta, 2007; Mehta, 2007; Mehta and Ma, 2012; Lee et al., 2013). By leveraging variation in prices at the micro level, researchers have built structural models to measure the effect of prices on demand across related categories. One premise of these models is that consumers buy the bundle of goods that maximize their immediate utility under a budget constraint; consequently, they focus exclusively on the effect of prices on purchases in the same time period.

However, we know from another stream of research that temporary price discounts may be followed by post-promotion dips in sales once prices go back to normal (Neslin and Schneider Stone, 1996). This behavior can be explained by the forward-looking behavior of consumers who take advantage of periods with low prices to stockpile quantities of goods in anticipation of their future consumption needs. By increasing their inventory at home when prices are low, they avoid buying goods when prices are high. Such stockpiling behavior has been evidenced on a wide range of product categories, both at the aggregate level and at the micro level. Moreover, it has been shown that ignoring stockpiling can lead to substantial biases in demand estimation, by dramatically overstating the effect of temporary price discounts on demand within a category (Sun et al., 2003; Hendel and Nevo, 2006a).
It is therefore important to distinguish between how goods are purchased and how they may be consumed over time: static models based on immediate utility maximization may lead to suboptimal pricing and promotion policies. While the implications of stockpiling behavior on demand within one category are well understood, its consequences on demand across related categories remain largely unknown at this point.

In this chapter, we focus specifically on complementary categories and investigate how stockpiling behavior alters the effect of prices across such categories. Unlike the usual definition of complementarity, which is based on the immediate effects of prices on purchases, we construe two goods as complementary if they provide a higher utility when consumed jointly (Topkis, 1998). Under that definition, the consumptions of two complements are likely to be interrelated: for example, a consumer may use more tomato sauce when she eats more spaghetti because the joint consumption gives her more utility. Despite this, the consumer may purchase the two complements at different times if she can store them for a while prior to consumption. In fact, purchasing complements at different times may allow the consumer to save money if price discounts happen at different times. To study the implications of asynchronous cross-category purchases, we develop a dynamic model of a forward-looking consumer making periodic purchase decisions across a set of complementary goods which can be stored from one period to the next. We endogenize the quantities of goods consumed as decision variables, so that they may be inter-related across categories. Compared to a single-category model, one major challenge consists in specifying a functional form for the consumption utility in a way that allows for flexible patterns of preferences. To overcome this challenge, we use our approach based on household production discussed in Chapter 2. Namely, we represent consumer preferences over quantities of latent final goods that the consumer can produce by combining quantities of the purchased goods in some proportions. This allows us to derive a flexible utility function over goods purchased that we can take to the data. To the best of our knowledge, our model is the first dynamic structural model of demand that allows for stockpiling of multiple goods and complementarity between them, and more generally the first one that accounts for storability across multiple categories.
Through simulation, we show that ignoring stockpiling behavior may lead to downward biases in the estimation of cross-category price effects because one ignores any lagged effect of past prices on future demand through the effect of inventories stored. We then apply our dynamic structural model on purchase data from a panel of consumers reporting their purchases across two prototypical complementary categories: spaghetti and tomato sauce. We show that our model fits the data better than restricted versions of the models that ignore complementarity or stockpiling behavior. We also compare the inferences made against inferences that one would make under an alternative static model that does not separate consumption from purchase, and under a dynamic model that treats each category independently.

The rest of this chapter is organized as follows. In Section 3.2, we review the literature on demand complementarity, consumption and stockpiling. Section 3.3 describes our model, its properties, and issues of estimation and identification. In Section 3.4, we run a simulation study to highlight the importance of accounting for stockpiling behavior in demand estimation across complementary goods. In Section 3.5, we describe an application of our model on data from a panel of consumers to measure the effect of prices and promotions on demand across categories, and discuss important conclusions from a managerial point of view. Section 3.6 concludes with a summary of the main contributions of our work, its limitations and ideas for future research.

3.2. Literature

This chapter relates to two streams of research that have separately studied demand for complementary goods on the one hand, and consumption and stockpiling behavior on the other hand.

3.2.1. Demand complementarity
<table>
<thead>
<tr>
<th>Study</th>
<th>Incidence</th>
<th>Brand choice</th>
<th>Quantity</th>
<th>Complementarity</th>
<th>Inventories</th>
<th>Forward-looking behavior</th>
<th>Endogenous consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chintagunta &amp; Haldar (1998)</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manchanda et al. (1999)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Mehta (2007)</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Niraj et al. (2008)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Song &amp; Chintagunta (2007)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Mehta &amp; Ma (2012)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Lee et al. (2013)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Neslin et al. (1985)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gupta (1988)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ailawadi &amp; Neslin (1998)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Bell et al. (1999)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Ailawadi et al. (2007)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Erdem et al. (2003)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Sun et al. (2003)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Sun (2005)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Hendel &amp; Nove (2006b)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Chan et al. (2008)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Hartmann &amp; Nair (2010)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Current study</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 6: Integrating the literature on cross-category demand estimation and the literature on stockpiling behavior
The literature has studied consumer-level demand for complements from different angles, as shown in the upper part of Table 6. Chintagunta and Haldar (1998) propose a timing model for purchases made across categories and find evidence of a positive duration dependence between spaghetti and tomato sauce. In a seminal paper, Manchanda et al. (1999) develop an incidence model to measure the cross-category effects of prices on demand. Researchers have then extended this approach to capture multiple aspects of the consumer’s decisions across categories, including brand choice and quantity. Many of these models do not take into account the inventories held by consumers, and thus rule out any dependence between time periods (Mehta, 2007; Song and Chintagunta, 2007). Other models recognize that consumers may keep inventories at home and that these inventories may influence their purchases, but assume an exogenous rate of consumption independently for each good and do not capture the stockpiling incentive through forward-looking behavior (Niraj et al., 2008; Lee et al., 2013). Overall these models have been successful at estimating demand across categories, but have not addressed the potential consequences of stockpiling behavior.

Complementarity can also be viewed as a feature of consumer preferences: two goods are complementary if the consumers gets more utility by consuming them jointly instead of separately. Formally, this can be captured by a supermodular utility function, which means that the marginal utility of one good increases with the quantity consumed of the other good. This approach has been taken by Gentzkow (2007) to show that online newspapers and print newspapers are complementary, and by Liu et al. (2010) to show complementarities between the consumption of broadband Internet services, cable TV and local phone. In these two articles, the consumer’s utility function is defined over a set of binary variables (e.g. reading the print newspaper vs not), and no inventory can be stored between time periods.

We contribute to this literature in several ways. First, we allow the inventories of the complementary goods to be interrelated, as they may be consumed jointly or separately. Second, we endogenize consumption and specify a supermodular utility on the quantities of goods consumed, which allows us to capture complementarity in the consumers’ preferences.
As a result, we formally model how cross price effects arise from consumption preferences, and our approach extends the model of Gentzkow (2007) by accommodating quantities and inventories. Third, we capture stockpiling behavior by recognizing consumers as forward-looking agents.

3.2.2. Consumption and stockpiling

Researchers have recognized for a long time that price promotions may increase sales because they may induce consumers to switch brands, to consume more, or to buy goods early in anticipation of a future consumption. Thus an increase in sales at the category level may in part be due to a consumption effect and to a stockpiling effect. An important body of literature, starting with Gupta (1988), has focused on measuring the importance of the consumption effect and of the stockpiling effect within a category. Bell et al. (1999) run a meta-analysis of purchases made by a panel of consumers across many product categories, and find that the stockpiling effect is substantial. Researchers have then built dynamic structural models to understand the implications of stockpiling behavior on demand over time (Erdem et al., 2003; Sun et al., 2003). These models are based on two main premises: the consumer can store goods from one period to the next, and she is forward-looking in the sense that she maximizes her flow of utilities over some time horizon. Thus, she may buy and store a good to reduce her future cost if she anticipates a price increase.

Building on evidence that people tend to consume at a higher rate when they have a higher inventory available (Ailawadi and Neslin, 1998) and that such behavior is rational (Assunção and Meyer, 1993), researchers have allowed the consumption rate to be flexible by endogenizing it as a decision variable (Sun, 2005; Hendel and Nevo, 2006b; Chan et al., 2008). In these models, the consumer not only decides what quantity to purchase at every time period, but also what quantity to consume. Under this approach, consumption may respond to marketing activity and to inventory levels, and the consumption effect can thus be disentangled from the stockpiling effect. While stockpiling behavior remains a very active area of research (Gordon and Sun, 2015; Haviv, 2015; Ching and Osborne, 2015), the
literature has mostly been restricted to a single category. One exception is the article by Hartmann and Nair (2010), who develop a demand model for tied goods where one good is storable but the other good is durable (such as blades and razors). In contrast, we build a demand model for two storable goods that are consumed endogenously.

3.3. Model

This section describes our dynamic model in detail. We first lay out the model of behavior at the individual level and discuss some of its properties. Then we introduce consumer heterogeneity and discuss its estimation as well as its identification.

3.3.1. General setup

We consider the behavior of a consumer who periodically goes on a shopping trip, and we specifically focus on \( J \) categories that she may be interested in buying. We model the incidence and quantity decisions at the category level and leave out any brand choice. At the beginning of a period \( t \), the consumer has some volumes \( i_t = (i_{1t}, \ldots, i_{Jt}) \) of the various goods at home, which are measured in the appropriate units (e.g. ounces). At the store, the consumer decides whether to purchase each of the goods, and in what quantities. Multiple package sizes may be available for purchase within each category. For example, spaghetti may be available in 16-ounce packs and 24-ounce packs. We denote by \( R \) the total number of all packages across the \( J \) categories, and by \( z_{r,j} \) the volume of good \( j \) that is contained in pack \( r \) (that amount is equal to zero if pack \( r \) belongs to another category than category \( j \)). We collect these package volumes in a matrix \( Z \). We further denote by \( p_t = (p_{1t}, \ldots, p_{Rt}) \) the prices of the packs at period \( t \), and by \( q_t = (q_{1t}, \ldots, q_{Rt}) \) the corresponding numbers of packs purchased by the consumer.

Buying goods is costly, as the consumer needs to pay a monetary price as well as a time cost for going to the aisle and picking the good from the shelf. We denote this acquisition cost as \( C_a(p, q) \), which is a function of the prices and the quantities purchased. After her shopping trip, the consumer goes home and stacks her new supplies onto her existing inventory: her
new inventories become equal to $i_{jt}^P = (i_{1jt}^P, ..., i_{Jjt}^P)$ such that $i_{jt}^P = i_{jt} + \sum_r z_{rjg_{rt}}$ for every good $j$. During the period, she consumes some volumes $y_t = (y_{1t}, ..., y_{Jt})$ of the various goods from which she derives a consumption utility $U_y(y_t)$. We endogenize the quantities consumed as decision variables to allow for flexible consumption rates (Sun, 2005). Finally, any remaining volumes are stored such that $i_{j,t+1} = i_{jt}^P - y_{jt}$ for each good $j$, and the next period starts.

To capture stockpiling behavior, we assume that the consumer is forward-looking with rational expectations about future prices. Therefore, she has an incentive to stockpile goods during periods with low prices: by storing quantities of goods when prices are low, she may be able to reduce her future cost of purchases when prices will be higher. However, this incentive may be diminished if her available shelf space is limited, and consequently we include a cost of storage $C_s(i_{t+1})$ in our model, which is a function of the volumes stored until the next period. Further, the consumer may not go on a shopping trip at every period. If she is uncertain about the timing of her next trip, she may have an incentive to buy and store quantities of goods even when prices are high. We take this into account by specifying a random process for store visits, and assuming that consumers have rational expectations about them. Specifically, we assume that shopping trips are exogenous and follow a Bernoulli process, such that the consumer has a constant probability $\rho$ to go on a shopping trip in every time period\(^7\). Given this setup, the consumer solves the following

\(^7\)The shopping trip decision is unlikely to be endogenous since we consider a small set of items from the larger shopping basket. The assumption of exogenous trips is common in the literature (Sun, 2005; Gordon and Sun, 2015).
dynamic optimization problem at period $t$:

$$V(i_t, p_t, s_t, \epsilon_t) = \max_{q_{j\tau}, y_{j\tau}} \mathbb{E} \left\{ \sum_{\tau = t}^{\infty} \delta^{\tau-t} \left[ -C_a(q_{j\tau}) + U_y(y_{j\tau}) - C_s(i_{\tau+1}) + \epsilon_{\tau}(q_{j\tau}) \right] \mid i_t, p_t, s_t \right\}$$

(3.1a)

subject to:

$$i_{j,\tau+1} = i_{j\tau} + \sum_r z_{rj} q_{r\tau} - y_{j\tau}$$

(3.1b)

$$q_{r\tau} \in \{0, 1, ..., \bar{q}_r\} \text{ if } s_{\tau} = 1 ; q_{r\tau} = 0 \text{ otherwise}$$

(3.1c)

$$y_{j\tau}, i_{j\tau} \geq 0$$

(3.1d)

where $s_t$ is an indicator variable equal to one if the consumer goes on a shopping trip at time $t$ and zero otherwise, $\epsilon_t(q_t)$ is a random shock that is specific to each purchase decision $q_t$, and $\delta$ is the consumer’s discount factor. To make the estimation of the model feasible, we assume an upper bound $\bar{q}_r$ on the number of units that can be purchased for each pack $r$ in any time period, as indicated by Constraint 3.1c. The second part of Constraint 3.1c simply indicates that the consumer cannot make a purchase of any positive quantity if she does not go on a shopping trip. Assuming that prices follow an exogenous first-order Markov process, we can rewrite the consumer problem under the form of a Bellman equation:

$$\bar{V}(i, p, s) = \mathbb{E}_{\epsilon} \left[ \max_{q \in \mathcal{Q}(s)} W(i + Zq, p) - C_a(q) + \epsilon(q) \right]$$

(3.2a)

where $W(i^p, p) = \max_y U_y(y) - C_s(i^p - y) + \delta \sum_{s' \in \{0,1\}} Pr(s') \int_{p'} Pr(p'|p) \bar{V}(i^p - y, p', s') dp'$

(3.2b)

where $\mathcal{Q}(s)$ is the set of possible purchase decisions under Constraint 3.1c, $p'$ and $s'$ denote respectively the prices and store visit indicator in the next period, and $Pr(p'|p)$ is the probability distribution of next period’s prices $p'$ conditional on current prices $p$. Equation 3.2a captures the tradeoff in the purchase decision: buying some packages is costly but allows the consumer to increase her inventory $i^p$ and therefore her value $W(i^p, p)$ at the following consumption stage. Equation 3.2b captures the tradeoff in the consumption decision: by
consuming some goods, she derives consumption utility and reduces her storage costs, but also reduces her inventory available and therefore her value $\bar{V}(i^p - y, p', s')$ at the beginning of the subsequent purchase stage. Following Rust (1987), we specify an extreme value distribution of type II on the random shocks $\epsilon(q)$. We can then obtain the conditional purchase probabilities as:

$$Pr(q|i, p, s) = \frac{\exp [W(i + Zq, p) + U_q(q)]}{\sum_{q' \in Q(s)} \exp [W(i + Zq', p) + U_q(q')]}$$

(3.3)

3.3.2. Model of price expectations

In every shopping trip, the consumer observes the prices of $R$ packages across all $J$ categories. We follow Erdem et al. (2003) and model the evolution of prices over time with a VAR(1) model:

$$\log (p_{j, t+1}) = \lambda_{j0} + \sum_{j'} \lambda_{jj'} \log (p_{j', t}) + \eta_{jt}$$

(3.4)

where $\eta_t \sim MVN(0, \Sigma_{\eta})$. This specification allows us to capture the dynamics of prices in a reduced-form way while accounting for correlation between prices across categories.

3.3.3. Functional forms

We now turn to the parametric specifications of the elements of the reward function, which are the cost of acquisition, the consumption utility and the storage cost.

Cost of acquisition

We assume that the consumer has a constant marginal utility for money denoted by $\gamma$. When making purchases, the consumer needs to give away money and therefore pays an opportunity cost equal to $\gamma$ times the dollar amount paid. Additionally, the consumer may incur a time cost of going to the aisle and picking the product from the shelf. This cost may be partly shared across multiple categories if they are located closeby in the store. To
capture this aspect, we include a cost that depends on purchase incidences across the $J$ categories. When prices are equal to $p$, the total cost of acquiring the quantities of packs $q$ is specified as follows:

$$C_a(q) = \gamma \sum_{j=1}^{J} p_j q_j + \sum_{j} \psi_{0j} I_j(q) + \sum_{j<j'} \psi_{jj'} I_j(q) I_{j'}(q)$$  \hspace{1cm} (3.5)$$

where $I_j(q)$ is equal to one if at least one pack is purchased in category $j$ and zero otherwise. The interaction term allows us to capture any reduction in time cost due to proximity of the categories inside the store.

**Consumption utility**

By consuming some quantities $y = (y_1, ..., y_J)$ of the goods, the consumer gets a reward which is captured by the consumption utility $U_y(y)$. We focus on complementary goods that may give a higher utility when consumed together: instead of the usual definition of complementarity based on cross price effects on demand, we construe complementarity as supermodularity of the consumption utility $U_y$ (Topkis, 1998). Our approach is similar to that taken by Gentzkow (2007), although we model utility for quantities consumed instead of binary variables for incidence. When dealing with quantities, one common approach in static models consists in parametrizing the indirect utility function and relying on Roy’s identity to derive the demand function. This approach cannot be invoked in a model of forward-looking behavior, as Roy’s identity cannot be applied in that context. Instead, we make use of the approach developed in Chapter 2, based on household production.

We assume that $K$ final goods can be obtained by combining the $J$ goods in some proportions, through a Leontief technology summarized by an input-table $A$ such that $a_{jk}$ gives the volume of good $j$ that is necessary to make one unit of final good $k$. For example, one

---

In our dynamic model, the equivalent of the indirect utility function is the value function $V(i,p,s,\epsilon)$ described in Equations 3.1a-3.1d, which takes the unobserved inventories as arguments. By parametrizing the value function instead of the reward function, it is not clear how one could capture the trade-off between the short-term utility derived from consumption and the long-term cost of inventory depletion.
burger and one bun may be combined together to obtain a sandwich, but the burger and
the bun may also be consumed separately or for some other use. Thus, we assume that the
$J$ goods are themselves final goods that can be directly consumed. We can then write $A$ as:

$$A = (I, A_2) \quad (3.6)$$

where $I$ is the identity matrix of dimension $J \times J$ and $A_2$ is a $J \times (K - J)$ matrix giving
the proportions used to produce latent composite goods. The consumer derives utility $U_c(c)$
from the quantities $c = (c_1, ..., c_K)$ of final goods, which we parametrize as follows:

$$U_c(c) = \sum_k \left[ \psi_k c_k + \psi_{kk} c_k^2 \right] \quad (3.7)$$

where $\psi_k$ and $\psi_{kk}$ are parameters. Unlike in Chapter 2 where we chose a functional form
based on logarithms, here we choose this quadratic form because it simplifies the computa-
tions to estimate of the dynamic structural model, as discussed in Appendix A.4. $^9$ There are
multiple ways a consumer may consume the ingredient quantities $y$, by using them jointly
or separately. Naturally, if she were to consume the ingredient quantities $y$, she would
combine them into the final good quantities $c$ that gives her the highest utility $U_c$. Thus
the utility of a set of ingredients $y$ is the highest utility that can be derived by combining
them into some final good quantities $c$:

$$U_y(y) = \max_{c \text{ s.t. } Ac = y} U_c(c) \quad (3.8)$$

$^9$Quadratic consumption utility functions have often been used in single-category models of endogenous
consumption (Sun, 2005; Gordon and Sun, 2015).
Storage costs

We assume linear costs of storage across categories:

\[ C_s(i) = \sum_{j=1}^{J} \kappa_j i_j \] (3.9)

where \( \kappa_j \) are parameters.

3.3.4. Consumption policy

Our model endogenizes consumption as a decision variable. As a result, inventories are depleted according to a consumption policy \( y^*(i^p, p) \), which is obtained by solving the Bellman equation laid out in Equations 3.2a-3.2b. Thus, different values of the parameters may yield different consumption patterns. The optimal consumption \( y^* \) clearly depends on the post-purchase inventories \( i^p \) available, since the consumer cannot consume more than her inventory. Current prices may also impact her consumption because they may affect her expectation of future prices: if she expects a price increase in the next period, she may prefer to consume less and store more inventory to avoid paying a high price in the future.

In Figure 3, we have illustrated some of the consumption patterns allowed by our model under different parameter values. The horizontal axes represent the volumes of both goods that are available to the consumer after the purchase stage, and the vertical axis represents the volumes consumed.

Figure 3.a) corresponds to a case where there is no consumption complementarity between the two goods: the consumer gets no utility from a joint consumption since \( \psi_1 = \psi_{11} = 0 \). Consequently, she consumes both goods independently: the volume of good 1 she consumes does not depend on her inventory of good 2, and vice versa. She consumes her entire inventories when they are low: consumption follows a forty-five degree line as a function of inventory. When she has higher inventory levels, she does not consume them entirely and prefers to keep some volumes for the next period; the amount she consumes increases
(a) No consumption complementarity
\[ \psi_1 = 0, \psi_2 = 2, \psi_3 = 2, \psi_{11} = 0, \psi_{22} = -0.5, \psi_{33} = -0.5, \alpha_{21} = 1, \gamma = 1.5, \kappa_1 = 0.1, \kappa_2 = 0.1, \rho = 1. \]

(b) Perfect consumption complementarity
\[ \psi_1 = 3, \psi_2 = 0, \psi_3 = 0, \psi_{11} = -0.1, \psi_{22} = 0, \psi_{33} = 0, \alpha_{21} = 1, \gamma = 1.5, \kappa_1 = 0.1, \kappa_2 = 0.1, \rho = 1. \]

(c) Asymmetric consumption complementarity
\[ \psi_1 = 2, \psi_2 = 2, \psi_3 = 1, \psi_{11} = -1, \psi_{22} = -1, \psi_{33} = -1, \alpha_{21} = 1, \gamma = 1.5, \kappa_1 = 0.1, \kappa_2 = 0.1, \rho = 1. \]

Figure 3: Consumption policy under various parameter values
slightly as a function of inventory available, consistent with Assunção and Meyer (1993).

Figure 3.b) corresponds to a case where there is perfect consumption complementarity between the two goods: the consumer derives utility from the joint consumption of the two goods but does not get any utility from consuming them separately since $\psi_2 = \psi_{22} = \psi_3 = \psi_{33} = 0$. In this case, the two goods are always consumed jointly in fixed proportions. Thus, a higher inventory available of good 1 does not necessarily imply a higher consumption of good 1 if there is no more inventory of good 2: in that case, the consumer prefers to keep some good 1 in her inventory and “wait” until she buys good 2 to consume them jointly in the future. This behavior gives rise to a pyramid-like shape for the consumption policy where the volumes consumed are equal to the minimum of the two inventory levels available. At high levels of inventory, the marginal utility for joint consumption decreases and the consumer keeps inventories of both goods for future consumption.

Figure 3.c) corresponds to an intermediate case in which there is some consumption complementarity between the two goods: the consumer derives utility from the joint consumption of the two goods and also from their separate consumption. In this case, she consumes some amount of each good even when she does not have any inventory available for the other good, and the volumes she consumes increase as a function of both inventories.

Clearly, our model can accommodate very different patterns where the amounts consumed of the two goods are interrelated. These patterns emerge endogenously from the model by making consumption a decision variable.

3.3.5. Heterogeneity and estimation

We assume that the parameters of the price process ($\lambda_{j0}$, $\lambda_{jj'}$, $\Sigma_\eta$) are homogenous in the population of consumers, and we estimate them in a first stage. In a second stage, we estimate the parameters of the dynamic structural model, which we allow to be heterogenous. We introduce subscript $h$ to denote a consumer, and we collect in a vector $\omega_h = (\psi_h, \gamma_h, \kappa_h, \rho_h)$ all the consumer-level parameters. We specify a discrete distribution
on these consumer-level parameters $\omega_h$, which results in a model with $S$ latent classes, each defined by a vector $\omega_s$ and a size $\pi_s$ such that $\sum_{s=1}^{S} \pi_s = 1$. The likelihood can then be written as:

$$L = \prod_{h=1}^{H} \sum_{s=1}^{S} \pi_s L_h(\omega_s)$$

where consumer $h$'s contribution to the likelihood conditional on belonging to class $s$ is:

$$L_h(\omega_s) = \prod_{t=1}^{T_h} \rho_s^{s_{ht}} (1 - \rho_s)^{1-s_{ht}} \times Pr(q_{ht}|i_{ht}, p_{ht}, s_{ht}; \omega_s)$$

In this equation, $T_h$ is the number of observable periods for consumer $h$ and the choice probabilities $Pr(q_{ht}|i_{ht}, p_{ht}, s_{ht}; \omega_s)$ are given in Equation 3.3. We estimate the dynamic structural model by maximum likelihood through a nested fixed-point algorithm (Rust, 1987). In the outer loop, we use the Nelder-Mead routine to search for the parameters $\{\pi_s, \omega_s\}_{s=1}^{S}$ that maximize the likelihood. For each value of $\omega_s$ drawn, we first solve the Bellman equation by following the approach developed by Keane and Wolpin (1994), and then we evaluate each consumer’s contribution to the likelihood. Appendix A.3 describes in depth how we solve the dynamic programming problem. In the evaluation of the likelihood, we use the law of motion in Equation 3.1b to update the inventories forward from one period to the next: this can be done easily since the quantities of goods purchased are observed and the optimal quantities consumed arise endogenously from the model, as discussed in Section 3.3.4. Like all dynamic models with unobservable stock variables, we face an initial conditions problem since the initial inventories are unobserved. To avoid any bias in estimation due to that problem, we use a warm-up window during which we update the inventories but do not evaluate the likelihood; we start evaluating the likelihood after that warm-up time window.\(^{10}\)

\(^{10}\)This approach has been used by Hendel and Nevo (2006b). In simulation, we found that a time window of four periods was sufficient to accurately initialize the inventories.
3.3.6. Identification

The identification of our model follows the same arguments as single-category inventory models found in previous literature (see Hendel and Nevo, 2006b, for a good discussion). The price parameter $\gamma$ is identified by variation in prices. The parameters of the time cost of purchases are identified through the (joint and marginal) frequencies of incidence across the goods, while the parameters of the consumption utility are identified through the distribution of quantities purchased. Some elements in the input-output matrix $A$ need to be normalized: in our simulation and empirical application, we set

$$A = \begin{pmatrix} 1 & 1 & 0 \\ a_{21} & 0 & 1 \end{pmatrix},$$

and estimate $a_{21}$ as a parameter, which is the amount of good 2 that is combined with one unit of good 2 to make one unit of the final good representing their joint consumption. The storage cost parameters are identified by state dependence through the lagged effects of past prices and past purchases on current purchases: intuitively, higher storage costs imply less stockpiling and therefore less dependence between time periods. Finally, the discount factor $\delta$ is not identified separately from the risk aversion parameters (Rust, 1987); we follow Erdem and Keane (1996) and set it equal to 0.995.

3.4. Simulation study

We ran a simulation study to investigate the biases that one may obtain when ignoring stockpiling behavior or cross-category complementarity in demand estimation. For that purpose, we considered three competing models.

Model 1 is a static version of our model in which consumers cannot store goods between periods. In that context, we cannot disentangle purchase and consumption since the goods purchased are necessarily consumed in the same period and it is therefore difficult to tease apart consumption complementarity (represented by the parameters of $U_y$) and purchase complementarity (captured by a sub-additive cost of acquisition through $\psi_{00}$) since both elements imply negative effects of prices on demand in the same time period. Therefore we shut off consumption complementarity in Model 1 by removing the possibility of joint
consumption, and captured complementarity only through parameter $\psi_{00}$.$^{11}$

Model 2 is a dynamic version of the model in which we shut off any complementarity between the goods, both through consumption (by removing the possibility of joint consumption) and through purchase (by setting $\psi_{00} = 0$). The model allows for stockpiling behavior and endogenous consumption within each category, but disregards any dependence between categories. Estimating that model is similar to estimating, for each category separately, a single-category model with stockpiling and endogenous consumption like Sun (2005).

Model 3 is the full dynamic model described in Section 3.3. It captures consumption complementarity through joint consumption utility parameters ($\psi_1, \psi_{11}, a_{21}$) and allows for sub-additive or super-additive time costs across categories through parameter $\psi_{00}$.

For each of the three models, we generated a synthetic dataset (Dataset 1, 2 and 3, respectively), assuming a simple setup so that the insights can be easily interpreted. Specifically, we assumed a homogenous population of 200 consumers followed over 100 weeks, who make purchase decisions in two product categories. The two goods are sold in package sizes equal to 16 and 26 ounces$^{12}$, respectively, and prices are independent across the two goods as well as over time. The values of price parameters and preference parameters used in data generation were chosen to yield a reasonable amount of variation in purchase behavior over time. After generating random draws of prices and random utility shocks, and solving the Bellman equation, we sequentially solved each consumer problem one period at a time, updating the inventory between periods. The resulting datasets contained, for each consumer and each period, the prices observed by the consumer, her purchase decision, her starting inventories and the quantities of both goods that she consumed.

$^{11}$We also estimated the alternative static model where we shut off purchase complementarity by setting $\psi_{00} = 0$ and kept the utility parameters for joint consumption. We obtained very similar model fits and elasticities between the two models.

$^{12}$In the simulation study as well as in the empirical application described in Section 3.5, we have converted the volumes into tens of ounces instead of ounces, so that the estimates of the consumption utility and storage cost parameters can be read more easily.
<table>
<thead>
<tr>
<th>Dataset 1</th>
<th>Dataset 2</th>
<th>Dataset 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>generated according to Model 1</td>
<td>generated according to Model 2</td>
<td>generated according to Model 3</td>
</tr>
<tr>
<td>(static)</td>
<td>(no complementarity)</td>
<td>(dynamic with complementarity)</td>
</tr>
<tr>
<td>Truth</td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Good acq (ψ₁₁)</td>
<td>1.80</td>
<td>1.78 (0.08)</td>
</tr>
<tr>
<td>Good acq (ψ₁₂)</td>
<td>1.40</td>
<td>1.51 (0.08)</td>
</tr>
<tr>
<td>Joint acq (ψ₀₀)</td>
<td>-1.50</td>
<td>-1.52 (0.05)</td>
</tr>
<tr>
<td>Price (γ)</td>
<td>1.80</td>
<td>1.87 (0.08)</td>
</tr>
<tr>
<td>Joint cons. (ψ₁)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Joint cons.² (ψ₁₁)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Good cons. (ψ₂₁)</td>
<td>2.20</td>
<td>2.22 (0.08)</td>
</tr>
<tr>
<td>Good cons.² (ψ₂₂)</td>
<td>-0.30</td>
<td>-0.30 (0.01)</td>
</tr>
<tr>
<td>Good cons.³ (ψ₃₁)</td>
<td>1.80</td>
<td>1.87 (0.06)</td>
</tr>
<tr>
<td>Good cons.² (ψ₂₂)</td>
<td>-0.20</td>
<td>-0.21 (0.01)</td>
</tr>
<tr>
<td>Proportion (a₂₁)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Good storage (κ₁₁)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Good storage² (κ₂₁)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Trip frequency (ρ)</td>
<td>0.70</td>
<td>0.71 (0.01)</td>
</tr>
</tbody>
</table>

| AIC | 81,351 | 82,542 | 81,354 | 69,617 | 68,951 | 68,954 | 84,337 | 81,987 | 81,087 |
| BIC | 81,422 | 82,621 | 81,464 | 69,688 | 69,030 | 69,065 | 84,408 | 82,066 | 81,198 |

Elast. 1 → 1 | -1.13 | -1.17 | -1.17 | -1.17 | -0.81 | -1.03 | -0.78 | -0.77 | -0.89 | -1.17 | -0.96 | -0.80 |
Elast. 1 → 2 | -0.18 | -0.19 | 0.00 | -0.19 | 0.00 | 0.00 | 0.00 | 0.00 | -0.22 | -0.06 | 0.00 | -0.22 |
Elast. 2 → 1 | -0.18 | -0.19 | 0.00 | -0.19 | 0.00 | 0.00 | 0.00 | 0.00 | -0.18 | -0.07 | 0.00 | -0.17 |
Elast. 2 → 2 | -0.73 | -0.75 | -0.75 | -0.75 | -0.73 | -0.78 | -0.69 | -0.69 | -0.27 | -1.03 | -0.32 | -0.24 |

Table 7: Simulation results under different versions of the dynamic cross-category model of demand
Next, we estimated the three competing models on each synthetic dataset. The estimation results are displayed in Table 7. For each model, we have reported the estimates of the model parameters along with their standard error in parentheses. To compare the fits of the models, we have also reported the log-likelihood of each model along with the Akaike information criterion (AIC) and the Bayesian information criterion (BIC), which penalize models with a higher number of parameters. To investigate the inferences made on the effects of prices in a way that is comparable across models, we have computed price elasticities by simulating aggregate demand across consumers and periods in counterfactual scenarios where the price of each good is increased by 10%. The lower part of Table 7 contains our estimates of own- and cross-price elasticities.

Across the three datasets, the true model fits the data better than its competing models based on the AIC and the BIC, and we recovered its true parameter values within their 95% confidence intervals. These results give support to the empirical identification of our model. We now consider the results of each dataset in turn.

In Dataset 1, demand is static since goods cannot be stored. The results of Model 3, which allows for dynamics, indicate that the parameters that are common with Model 1 (the true model) are correctly estimated in the sense that the true value is within the 95% confidence interval. For good 1, the absence of dynamics is captured by a large storage cost parameter $\kappa_1$, which prevents the consumer from storing volumes. The storage cost parameter for good 2 ($\kappa_2$) is negative, but the large standard deviation suggests that its value does not dramatically change the results. Model 3 yields the same price elasticities as the true model (Model 1) and fits the data almost as well. Thus, our dynamic model can fit the data well even in the absence of dynamics.

In Dataset 2, demand is independent across the two goods. Again, the parameters of Model 3 that are common with Model 2 (the true model) are correctly estimated. Moreover, the consumption policy $y^* (i^p, p)$ that we obtained under Model 3 imply that no inventory is carried over between periods. This is because the quadratic term $\psi_{33}$ of the consumption utility is small, leading to little satiation and therefore no reason to prefer storage for future consumption over immediate consumption.

\footnote{The consumption policy $y^* (i^p, p)$ that we obtained under Model 3 imply that no inventory is carried over between periods. This is because the quadratic term $\psi_{33}$ of the consumption utility is small, leading to little satiation and therefore no reason to prefer storage for future consumption over immediate consumption.}
Model 3 fits the data almost as well as Model 2 and yields very similar price elasticities. The absence of complementarity between the goods is captured by a non-significant value for $\psi_{00}$ suggesting that acquisition costs are additive across goods, and by very small values of $\psi_1$ and $\psi_{11}$ suggesting that consumers derive almost no utility from joint consumption of the goods. Thus, our dynamic model can give accurate results even in the absence of complementarity between goods.

In Dataset 3, the results of Model 1 (the static model) provide insight about the consequences of ignoring stockpiling behavior in cross-category demand estimation. The static model yields over-estimated own-price elasticities, which is consistent with prior literature: while it recognizes the immediate increase in demand during a temporary price discount, it fails to recognize the decrease in demand in future periods, which results in an over-estimation of long-run price effects (Hendel and Nevo, 2006b). More interestingly, the static model also largely under-estimates the cross-price elasticities, which has not been studied in previous research. We explain this bias by the following argument: a price discount for one good may lead a consumer to buy it to consume it jointly with a complement that she has in her inventory, in which case she will likely repurchase that complement earlier to replenish her inventory. Furthermore, a consumer who stockpiles a good during a promotion may be more likely to repurchase the complement in future periods to consume both goods jointly. For these reasons, a price discount for one good may increase demand for a complement not only in the same period, but also in future periods. A static model disregards any delayed effect of price discounts across categories, and only captures the immediate cross-category price effect: as such, it may only capture a fraction of the total increase in demand over time.

To illustrate our explanation, we made a small change in Dataset 3, by simulating an exogenous temporary price discount on good 1 in a given time period. On Figure 4, we plot the time series of aggregate sales of both goods, ten periods prior to the price discount to ten periods after. As we can see on the right plot, the promotion on good 1 leads to an
Figure 4: Effect of an exogenous promotion on sales across categories in simulated data

In this simulation study, we have shown that our model can yield accurate estimates of price elasticities even in the absence of stockpiling behavior or complementarity. We have also shown that static models of cross-category demand may yield biased estimates of cross-price elasticities and have proposed an explanation for that bias. In the next section, we present an application of our model on real data.

3.5. Empirical application

This section describes an application of our model on data from a panel of consumers. We describe the data, we discuss our estimation results and the results of counterfactual analyses.
3.5.1. Data

We obtained data from AC Nielsen through the Kilts Center for Marketing. The data encompasses a very large set of product categories and is made of two parts: the Consumer Panel data and the Retail Scanner data. The Consumer Panel data contains data about a panel of households who report their purchases over time: after each shopping trip, they scan each item purchased and enter the corresponding number of units they purchased as well as the price they paid. In this data, every product is identified by a universal product code (upc), and information is available about its category, brand and package size. We combine the Consumer Panel data with the Retail Scanner data which gives, for each participating store and for each product sold, the number of units sold and the dollar sales on a weekly basis. Merging both data sources allows us to obtain the prices of the products that are not purchased by a household on a given trip.

In our analysis, we focused on purchases of spaghetti and tomato sauce categories in years 2011 and 2012. Spaghetti and tomato sauce can be stored for several weeks and are a prototypical pair of complementary categories that has been analyzed in previous literature (Chintagunta and Haldar, 1998; Mehta and Ma, 2012). We focused on the 16-ounce and 25-ounce packages of spaghetti and sauce, respectively, which are the predominant formats, and we restricted our analysis to households who were part of the panel for both years and made at least ten purchases in each of the two categories of interest over the entire time window.\footnote{Given the dynamic nature of our model, we need enough observations with purchases to correctly estimate the parameters of the model.} We constructed category-level prices for each consumer-week combination by taking an average of brand-specific prices weighted by market share.

The resulting dataset contained 453 households followed over 105 weeks. For each household-week pair, the dataset indicates whether the household has gone on a shopping trip, and gives the unit price and the number of units purchased of each good by the household. In Table 8, we display some summary statistics of the data. The two goods were often pur-
(a) Purchase incidences

<table>
<thead>
<tr>
<th></th>
<th>Both</th>
<th>Spaghetti only</th>
<th>Sauce only</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of observations (%)</td>
<td>3,434 (10.5%)</td>
<td>3,132 (9.6%)</td>
<td>4,693 (14.4%)</td>
<td>21,302 (65.4%)</td>
</tr>
</tbody>
</table>

(b) Descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>Spaghetti</th>
<th>Sauce</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price ($)</td>
<td>1.35</td>
<td>2.07</td>
</tr>
<tr>
<td>Purchase incidence (%)</td>
<td>20.17</td>
<td>24.96</td>
</tr>
<tr>
<td>Mean purchase quantity</td>
<td>1.53</td>
<td>1.74</td>
</tr>
</tbody>
</table>

Table 8: Description of purchase data in the spaghetti and tomato sauce categories

<table>
<thead>
<tr>
<th></th>
<th>Spaghetti</th>
<th>Sauce</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-sale</td>
<td>24.16</td>
<td>42.59</td>
</tr>
<tr>
<td>Total</td>
<td>0.83</td>
<td>3.50</td>
</tr>
<tr>
<td>Within</td>
<td>0.78</td>
<td>1.35</td>
</tr>
<tr>
<td>Mean purchase quantity (oz)</td>
<td>(0.2007)</td>
<td>(0.30)</td>
</tr>
<tr>
<td>Weeks from previous purchase</td>
<td>6.31</td>
<td>5.03</td>
</tr>
<tr>
<td>(0.09)</td>
<td>0.05</td>
<td>0.22</td>
</tr>
<tr>
<td>(0.15)</td>
<td>-0.13</td>
<td>0.02</td>
</tr>
<tr>
<td>Weeks until next purchase</td>
<td>6.21</td>
<td>5.07</td>
</tr>
<tr>
<td>(0.09)</td>
<td>0.23</td>
<td>0.20</td>
</tr>
<tr>
<td>(0.14)</td>
<td>0.22</td>
<td>0.21</td>
</tr>
<tr>
<td>Inventory at beginning</td>
<td>-6.93</td>
<td>-10.02</td>
</tr>
<tr>
<td>(0.82)</td>
<td>-3.21</td>
<td>-7.02</td>
</tr>
<tr>
<td>(1.21)</td>
<td>0.95</td>
<td>0.56</td>
</tr>
<tr>
<td>(1.07)</td>
<td>(1.20)</td>
<td>(2.33)</td>
</tr>
</tbody>
</table>

Table 9: Reduced-form analysis of stockpiling behavior

Purchased together: when a consumer purchased spaghetti, she also purchased tomato sauce about half of the time.

3.5.2. Reduced-form evidence

In this section, we investigate the existence of stockpiling behavior within each category and the existence of complementarity between them.
Stockpiling behavior

Following Hendel and Nevo (2006a), we look for evidence of stockpiling behavior within each category by comparing the purchases made when goods are on sale (sale purchases) and purchases made when they are not on sale (non-sale purchases). Under forward-looking behavior, consumers take advantage of sales to accumulate inventory for their future consumption, and buy at the regular price only for their immediate consumption when they have little inventory left. Consequently, sale purchases tend to happen at higher level of inventories than non-sale purchases, and lead to higher levels of inventory after a consumption period. While consumer inventories are not directly observable, Hendel and Nevo (2006a) suggest using inter-purchase times as proxies to test two hypotheses. First, sale purchases should happen after a shorter inter-purchase time than non-sale purchases because consumers stockpile goods even when they still have some inventory: thus price discounts should lead to purchase acceleration. Second, sale purchases should be followed by a longer duration until the next purchase than non-sale purchases. These two hypotheses are easily testable given data on prices and purchases, and consumer heterogeneity can be easily accommodated by taking advantage of the cross-sectional aspect of the data.

We reproduced the same empirical analysis, where we define a purchase as a sale purchase if the price paid was at least 5% lower than the last five prices observed by the same household in previous shopping trips. We also constructed an inventory variable for each category, assuming a linear rate of consumption similar to Neslin et al. (1985), to test for differences in inventory levels between sale-purchases and non-sale purchases. We compile the results of our analysis in Table 9. The Non-Sale column gives the average volumes and inter-purchase times for non-sale purchases, the “Between” column gives the average difference for sale purchases compared to non-sale purchases when pooling all observations together, and the “Within” column reports the average difference across households, after first computing the difference between sale purchases and non-sale purchases separately for each household. The “Within” column allows us to test for increased purchased quantities,
Table 10: Reduced-form analysis of complementarity between spaghetti and tomato sauce

<table>
<thead>
<tr>
<th>Incidence of spaghetti</th>
<th>Incidence of spaghetti</th>
</tr>
</thead>
<tbody>
<tr>
<td>(no HH FE)</td>
<td>(with HH FE)</td>
</tr>
<tr>
<td>Price of spaghetti</td>
<td>-0.4465 (0.0857)</td>
</tr>
<tr>
<td>Price of sauce</td>
<td>-0.0365 (0.0661)</td>
</tr>
<tr>
<td></td>
<td>-0.5120 (0.1075)</td>
</tr>
<tr>
<td></td>
<td>-0.0736 (0.0993)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Incidence of sauce</th>
<th>Incidence of sauce</th>
</tr>
</thead>
<tbody>
<tr>
<td>(no HH FE)</td>
<td>(with HH FE)</td>
</tr>
<tr>
<td>Price of spaghetti</td>
<td>0.0803 (0.0779)</td>
</tr>
<tr>
<td>Price of sauce</td>
<td>-0.3190 (0.0616)</td>
</tr>
<tr>
<td></td>
<td>-0.0800 (0.0995)</td>
</tr>
<tr>
<td></td>
<td>-0.4630 (0.0932)</td>
</tr>
</tbody>
</table>

The results indicate that consumers buy larger quantities when goods are on sale versus when they are not, which is expected even in the absence of stockpiling behavior. Promotions seem to lead to some purchase acceleration in the spaghetti category (although the effect is not statistically significant) but there does not seem to be any purchase acceleration in the sauce category. Across the two categories, sales purchases tend to be followed by longer inter-purchase times and tend to happen at higher levels of inventory, although not at a statistically significant level. Together, the results give some support to the existence of stockpiling behavior, especially in the spaghetti category.

**Complementarity**

We applied logit models of purchase incidence to investigate the effect of prices on purchase and thus look for evidence of complementarity between the two categories. For each category, we ran two logit models with prices as explanatory variables: one with the same intercept for all households, and one with household-specific fixed effects to control for heterogeneity. The results are compiled in Table 10. After controlling for household heterogeneity, the price coefficients are all negative. Thus a lower price leads to a higher
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept $\lambda_{j0}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spaghetti</td>
<td>0.0452</td>
<td>(0.0030)</td>
</tr>
<tr>
<td>Sauce</td>
<td>0.2070</td>
<td>(0.0024)</td>
</tr>
<tr>
<td>Lagged prices $\lambda_{jj'}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spaghetti $\rightarrow$ Spaghetti</td>
<td>0.6357</td>
<td>(0.0038)</td>
</tr>
<tr>
<td>Sauce $\rightarrow$ Spaghetti</td>
<td>0.0869</td>
<td>(0.0045)</td>
</tr>
<tr>
<td>Spaghetti $\rightarrow$ Sauce</td>
<td>0.0532</td>
<td>(0.0030)</td>
</tr>
<tr>
<td>Sauce $\rightarrow$ Sauce</td>
<td>0.6907</td>
<td>(0.0035)</td>
</tr>
<tr>
<td>Variance-covariance matrix $\Sigma_\eta$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance Spaghetti</td>
<td>0.0090</td>
<td>(0.0001)</td>
</tr>
<tr>
<td>Variance Sauce</td>
<td>0.0057</td>
<td>(0.0001)</td>
</tr>
<tr>
<td>Cov. (Spaghetti, Sauce)</td>
<td>0.0015</td>
<td>(0.0001)</td>
</tr>
</tbody>
</table>

Table 11: Estimates of the price process parameters

probability of purchase incidence in the same category and in the other category, although the cross-category effects are not statistically significant. Together, these results provide some support to the existence of complementarity between the two categories under the usual definition based on negative cross-price effects.

3.5.3. Estimation results

We first estimated the parameters of the price process, which are later used to obtain the consumers’ expectations of future prices in our dynamic structural model. The results of this first-stage estimation are displayed in Table 11. We observe that the lagged prices have a positive effect on the next prices within each category, and small positive effects across categories.

In the second stage, we estimated the three competing models already discussed in Section 3.4: Model 1 is a static model, Model 2 is a model that disregards complementarity across categories, and Model 3 is our proposed model that captures forward-looking behavior as well as complementarity across categories. Our original results are compiled in the upper
### Unconstrained models

<table>
<thead>
<tr>
<th></th>
<th>Model 1 (static)</th>
<th>Model 2 (no complementarity)</th>
<th>Model 3 (full model)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of buying sauce ($\psi_{02}$)</td>
<td>1.45 (0.02)</td>
<td>1.04 (0.01)</td>
<td>1.42 (0.01)</td>
</tr>
<tr>
<td>Cost of buying spaghetti ($\psi_{01}$)</td>
<td>1.26 (0.04)</td>
<td>0.87 (0.01)</td>
<td>0.73 (0.01)</td>
</tr>
<tr>
<td>Extra cost of buying both ($\psi_{00}$)</td>
<td>-1.60 (0.02)</td>
<td>-</td>
<td>-1.49 (0.02)</td>
</tr>
<tr>
<td>Price ($\gamma$)</td>
<td>0.16 (0.01)</td>
<td>0.40 (0.01)</td>
<td>0.16 (0.01)</td>
</tr>
<tr>
<td>Consumption of joint ($\psi_1$)</td>
<td>-</td>
<td>-</td>
<td>-1.13 (0.01)</td>
</tr>
<tr>
<td>Consumption of joint$^2$ ($\psi_{11}$)</td>
<td>-</td>
<td>-</td>
<td>0.07 (0.01)</td>
</tr>
<tr>
<td>Consumption of sauce ($\psi_3$)</td>
<td>-0.14 (0.01)</td>
<td>0.00 (0.01)</td>
<td>-0.12 (0.01)</td>
</tr>
<tr>
<td>Consumption of sauce$^2$ ($\psi_{33}$)</td>
<td>-0.00 (0.01)</td>
<td>0.00 (0.01)</td>
<td>-0.04 (0.01)</td>
</tr>
<tr>
<td>Consumption of spaghetti ($\psi_2$)</td>
<td>-0.50 (0.03)</td>
<td>-3.72 (0.04)</td>
<td>-1.02 (0.01)</td>
</tr>
<tr>
<td>Consumption of spaghetti$^2$ ($\psi_{22}$)</td>
<td>0.02 (0.01)</td>
<td>0.03 (0.01)</td>
<td>0.07 (0.01)</td>
</tr>
<tr>
<td>Consumption proportion ($a_{21}$)</td>
<td>-</td>
<td>-</td>
<td>1.45 (0.04)</td>
</tr>
<tr>
<td>Inv. cost sauce ($\kappa_2$)</td>
<td>-</td>
<td>0.00 (0.01)</td>
<td>0.03 (0.01)</td>
</tr>
<tr>
<td>Inv. cost spaghetti ($\kappa_1$)</td>
<td>-</td>
<td>0.04 (0.01)</td>
<td>1.08 (Inf)</td>
</tr>
<tr>
<td>Trip frequency ($\rho$)</td>
<td>0.72 (0.01)</td>
<td>0.72 (0.01)</td>
<td>0.72 (0.01)</td>
</tr>
<tr>
<td>AIC</td>
<td>146,786</td>
<td>149,379</td>
<td>146,367</td>
</tr>
<tr>
<td>BIC</td>
<td>146,864</td>
<td>149,466</td>
<td>146,489</td>
</tr>
<tr>
<td>LL</td>
<td>-73,384</td>
<td>-74,680</td>
<td>-73,170</td>
</tr>
</tbody>
</table>

#### Elasticity

<table>
<thead>
<tr>
<th></th>
<th>Model 1 (static)</th>
<th>Model 2 (no complementarity)</th>
<th>Model 3 (full model)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasticity sauce → sauce</td>
<td>-0.57</td>
<td>-1.67</td>
<td>-0.65</td>
</tr>
<tr>
<td>Elasticity sauce → spaghetti</td>
<td>-0.15</td>
<td>0.00</td>
<td>-0.21</td>
</tr>
<tr>
<td>Elasticity spaghetti → sauce</td>
<td>-0.07</td>
<td>0.00</td>
<td>-0.05</td>
</tr>
<tr>
<td>Elasticity spaghetti → spaghetti</td>
<td>-0.35</td>
<td>-0.84</td>
<td>-0.34</td>
</tr>
</tbody>
</table>

### Constrained models

<table>
<thead>
<tr>
<th></th>
<th>Model 1 (static)</th>
<th>Model 2 (no complementarity)</th>
<th>Model 3 (full model)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of buying sauce ($\psi_{02}$)</td>
<td>1.54 (0.02)</td>
<td>1.27 (0.02)</td>
<td>1.87 (0.04)</td>
</tr>
<tr>
<td>Cost of buying spaghetti ($\psi_{01}$)</td>
<td>1.91 (0.02)</td>
<td>0.88 (0.01)</td>
<td>1.85 (0.04)</td>
</tr>
<tr>
<td>Extra cost of buying both ($\psi_{00}$)</td>
<td>-1.59 (0.02)</td>
<td>-</td>
<td>-1.45 (0.02)</td>
</tr>
<tr>
<td>Price ($\gamma$)</td>
<td>0.29 (0.04)</td>
<td>0.20 (0.06)</td>
<td>0.36 (0.03)</td>
</tr>
<tr>
<td>Consumption of joint ($\psi_1$)</td>
<td>-</td>
<td>-</td>
<td>0.00 (-)</td>
</tr>
<tr>
<td>Consumption of joint$^2$ ($\psi_{11}$)</td>
<td>-</td>
<td>-</td>
<td>0.00 (-)</td>
</tr>
<tr>
<td>Consumption of sauce ($\psi_3$)</td>
<td>-</td>
<td>0.00 (-)</td>
<td>0.27 (0.12)</td>
</tr>
<tr>
<td>Consumption of sauce$^2$ ($\psi_{33}$)</td>
<td>-0.00 (0.19)</td>
<td>-0.01 (0.08)</td>
<td>-0.03 (0.10)</td>
</tr>
<tr>
<td>Consumption of spaghetti ($\psi_2$)</td>
<td>-</td>
<td>0.00 (-)</td>
<td>0.11 (0.35)</td>
</tr>
<tr>
<td>Consumption of spaghetti$^2$ ($\psi_{22}$)</td>
<td>-0.05 (0.02)</td>
<td>0.00 (-)</td>
<td>-0.08 (0.08)</td>
</tr>
<tr>
<td>Consumption proportion ($a_{21}$)</td>
<td>-</td>
<td>-</td>
<td>2.17 (0.71)</td>
</tr>
<tr>
<td>Inv. cost sauce ($\kappa_2$)</td>
<td>-</td>
<td>0.08 (0.18)</td>
<td>0.11 (0.10)</td>
</tr>
<tr>
<td>Inv. cost spaghetti ($\kappa_1$)</td>
<td>-</td>
<td>0.07 (0.02)</td>
<td>0.25 (0.12)</td>
</tr>
<tr>
<td>Trip frequency ($\rho$)</td>
<td>0.72 (0.01)</td>
<td>0.72 (0.01)</td>
<td>0.72 (0.01)</td>
</tr>
<tr>
<td>AIC</td>
<td>146,920</td>
<td>149,624</td>
<td>146,727</td>
</tr>
<tr>
<td>BIC</td>
<td>146,998</td>
<td>149,711</td>
<td>146,849</td>
</tr>
<tr>
<td>LL</td>
<td>-73,451</td>
<td>-74,802</td>
<td>-73,349</td>
</tr>
</tbody>
</table>

#### Elasticity

<table>
<thead>
<tr>
<th></th>
<th>Model 1 (static)</th>
<th>Model 2 (no complementarity)</th>
<th>Model 3 (full model)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasticity sauce → sauce</td>
<td>-1.00</td>
<td>-0.70</td>
<td>-1.23</td>
</tr>
<tr>
<td>Elasticity sauce → spaghetti</td>
<td>-0.32</td>
<td>0.00</td>
<td>-0.43</td>
</tr>
<tr>
<td>Elasticity spaghetti → sauce</td>
<td>-0.16</td>
<td>0.00</td>
<td>-0.20</td>
</tr>
<tr>
<td>Elasticity spaghetti → spaghetti</td>
<td>-0.62</td>
<td>-0.43</td>
<td>-0.75</td>
</tr>
</tbody>
</table>

Table 12: Estimation results on the spaghetti and tomato sauce purchase data
While our proposed model (Model 3) fits the data better than the competing models based on the AIC and BIC, some parameters do not have the expected sign: we expect $\psi_1$, $\psi_2$ and $\psi_3$ to be positive and $\psi_{11}$, $\psi_{22}$ and $\psi_{33}$ to be negative, since we expect the consumption utility to be increasing and concave. The signs we obtained are probably due to a lack of variation in quantities purchased. We have re-run the models by constraining those parameters to have the expected signs. The results are compiled in the lower part of Table 12. Model 3 still fits the data better than the competing models based on both the AIC and the BIC. The parameter estimates suggest that complementarity arises mostly through purchase: the parameters of the joint consumption utility $\psi_1$ and $\psi_{11}$ were estimated to be zero, while our estimate of $\psi_{00}$ is negative, suggesting sub-additive costs of acquisition across categories. This may be explained by the likely proximity of the corresponding shelves in the store. According to these results, a price discount in one category may lead the consumer to buy an item in the other category but not necessarily to consume it faster.

Using the constrained parameter values of Model 3, we ran a counterfactual analysis in which we simulated aggregate demand over time under a 20% price discount occurring at a given week, for each of the two categories. We plot the time series of the aggregate volumes purchased on Figure 5. The results suggest that price promotions have an important impact on sales across categories during the promotion, but a negligible impact on sales in following periods. Future research should study the impact of price promotions on cross-category demand in other pairs of complementary categories.

3.6. Conclusion

In this chapter, we have proposed a micro-level model of stockpiling behavior and endogenous consumption across complementary goods. Our model considers the purchase and consumption of goods that are storable and that provide a higher utility when consumed jointly. In this context, a consumer may purchase two complementary goods at different
Figure 5: Effect of exogenous promotions on sales of spaghetti and tomato sauce over time times before consuming them jointly.

Our model integrates two streams of models that have been designed separately to investigate cross-category demand on the one hand, and on the other hand the consequences of stockpiling behavior for the effect of prices and promotions within a single category. Thus, it extends the existing literature on stockpiling behavior by investigating the effect of such behavior across complementary categories, and it extends the literature on cross-category demand by investigating the consequences of storability and forward-looking behavior.

Through simulation, we have shown that ignoring stockpiling behavior may lead one to
under-estimate the cross-category effect of price. To explain this, we have suggested that price promotions may have a positive lagged effect on sales of complementary goods, and that static models only partly captures the cross-category effects of prices because they ignore any lagged effect. We have validated this explanation through simulation, using our model as a normative model of behavior.

In an application of our model on purchase data from a set of consumers, we have found that our dynamic model fits the data better than a static model that disregard stockpiling behavior, or a dynamic model that disregard complementarity between goods. We have however found that the lagged effect of promotions was negligible. Further research is needed to understand the effect of price promotions across multiple pairs of complementary categories.
CHAPTER 4 : General Conclusion

This dissertation has made important progress on the problem of demand estimation across goods that are complementary. These goods, often consumed together, are characterized by the existence of negative cross-price effects on demand: as the price of one good increases, demand for its complement decreases. Knowing how the price of one good affects demand for its complement is crucial for both retailers and manufacturers that produce them, as their overall profit depend on sales across these goods.

Yet, previous methods of cross-category demand estimation are limited. On the one hand, macro-level models are not robust to policy changes, which reduces their applicability when the goal of the analysis is precisely to find the optimal pricing policy. On the other hand, the micro-level approach based on economic theory leaves the researcher with the hurdle of finding a functional form for the consumer’s utility that is tractable and leads to a demand system that is both flexible and regular.

We have proposed a new approach to overcome the challenges in estimating micro-level demand for complements. Our approach is based on the household production theory: according to that theory, consumers purchase goods and use them as inputs to produce final goods from which they enjoy utility. In Chapter 2, we have shown that laying out the consumer problem this way allows the researcher to obtain very flexible representations of consumer preferences for goods purchased using simple functional forms. We have also shown that the new consumer problem can be reduced to the original one, which means that no further assumption about the consumer’s decision process is necessary, unlike the previous approaches mentioned earlier. The household production theory is particularly well-suited to model preferences for complements, as it links the existence of negative cross-price effects between two goods to the consumer’s preferences for their joint consumption.

In Chapters 2 and 3, we have taken the household-production approach to study the implications of two micro-level aspects of demand: discreteness and stockpiling behavior. The
importance of recognizing both aspects when estimating demand within a single category has been emphasized in previous research, but their consequences on cross-category estimation were largely unknown. Our new approach to modeling complementarity has allowed us to address these questions.

In Chapter 2, we have focused on the implications of demand discreteness for the estimation of demand across complementary categories. Micro-level demand for packaged goods necessarily lies on grid of discrete points because the consumer can only buy an integer number of packs. We have found that ignoring such discreteness may lead to important biases in the estimation of own and cross-price effects on demand. This is of particular interest since previous demand systems relying on Roy’s identity disregard the discreteness of demand.

In Chapter 3, we have studied the implications of stockpiling behavior on the estimation of cross-category price effects. Stockpiling behavior is the process by which consumers take advantage of temporary price discounts to store quantities of goods in anticipation of their future consumption needs. In the case of complementary goods, it implies that two goods consumed together may be purchased at different times. Consequently, the effect of prices on demand across categories may be in part delayed: a price discount for one good in a given week may induce a consumer to purchase its complement in the following weeks. By embedding our utility specification in a dynamic structural model with forward-looking consumers, we were able to measure these price effects. We found that ignoring stockpiling behavior may lead to biases in the estimation of cross-category price effects.

We now conclude this dissertation by opening new ideas for future research.

First, while we have focused on the demand side in this dissertation, it would be interesting to study further the implications of discreteness and stockpiling behavior for the optimal policies from manufacturers and retailers. For example, how should manufacturers decide on package sizes when they produce complementary goods that are consumed together? Should they schedule their promotions simultaneously or at different times? While
Sinitsyn (2012) advances a brand loyalty argument to suggest that manufacturers should promote their complementary goods simultaneously, sequential promotions may allow them to achieve a better inter-temporal price discrimination than simultaneous promotions. Intuitively, price-sensitive consumers may be able to stockpile each good separately when it is on sale, while price-insensitive consumers may purchase complementary goods simultaneously. Consequently, price-sensitive consumers may effectively pay a lower price than price-insensitive consumers under negatively correlated prices across the complementary goods.

Second, it would be interesting to study real consumption behavior, as consumption data becomes available. In the absence of such data, our two models have been estimated on purchase data even though we consider consumption in our model. Therefore, we have made the assumption that consumers optimally allocate the goods they purchase into consumption bundles. By observing consumption data, one could then investigate how consumers really allocate the goods they purchase, how they consume them over time, and whether they anticipate correctly their consumption needs when making purchase decisions.

Finally, while we have focused on complementary goods, it may also be possible to capture flexible patterns of substitution by invoking household production theory. Previous direct-utility models of demand have used an additively separable parameterization for the utility of the focal goods and have accommodated substitution between through a decreasing marginal utility for the outside good (Satomura et al., 2011; Lee and Allenby, 2014): as the consumer spends more money on one focal good, the marginal utility of the outside good increases, inducing her to spend less on the other focal goods. While this approach allows the researcher to capture positive cross-price effects, the resulting demand system is such that the cross-price effects are determined by how much is spent on each focal good. This property may restrict the substitution patterns that one may be able to capture, similar to the way the Independence of Irrelevant Alternatives property restricts price elasticities in a logit model. Another solution to capture substitution would be to capture it directly in
the utility function for focal goods, such that the marginal utility of a focal good decreases if the amount consumed of a substitute is increased. Such an approach may be lead to a more flexible demand system. An approach based on household production may be useful to overcome the challenge of finding a tractable functional form to yield such a demand system.
A.1. Solution to the consumption problem in Chapter 2

In this section, we derive the solution to the optimal consumption problem given in Equations 2.6b-2.6d:

\[ U_x(x) = \max_c \sum_k \alpha_k \log(c_k + 1) \]

s.t. \[ c_k \geq 0 \quad \forall k \]

\[ \sum_k a_{jk} c_k \leq x_j \quad \forall j \] \hspace{1cm} (A.1)

In the case with two inputs and one composite such that \( A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & a_{23} \end{pmatrix} \), the first-order conditions given in Equation 2.7 become:

\[ \frac{\alpha_1}{c_1 + 1} + \lambda_1 - \mu_1 = 0 \] \hspace{1cm} (A.2a)

\[ \frac{\alpha_2}{c_2 + 1} + \lambda_2 - \mu_2 = 0 \] \hspace{1cm} (A.2b)

\[ \frac{\alpha_3}{c_3 + 1} + \lambda_3 - \mu_1 - \mu_2 a_{23} = 0 \] \hspace{1cm} (A.2c)

\[ \lambda_1 c_1 = 0 \] \hspace{1cm} (A.2d)

\[ \lambda_2 c_2 = 0 \] \hspace{1cm} (A.2e)

\[ \lambda_3 c_3 = 0 \] \hspace{1cm} (A.2f)

\[ \mu_1 (x_1 - c_1 - c_3) = 0 \] \hspace{1cm} (A.2g)

\[ \mu_2 (x_2 - c_2 - c_3 a_{23}) = 0 \] \hspace{1cm} (A.2h)

\[ c_k, \mu_k, \lambda_j \geq 0 \] \hspace{1cm} (A.2i)
Because $U_c$ is strictly increasing, constraints (*g) and (*h) must be binding:

\[
\begin{align*}
    c_1 &= x_1 - c_3 \\ 
    c_2 &= x_2 - a_{23}c_3
\end{align*}
\] (A.3)

We can simplify the FOC as:

\[
\begin{align*}
    \frac{\alpha_1}{x_1 - c_3 + 1} + \lambda_1 - \mu_1 &= 0 \quad (A.4a) \\
    \frac{\alpha_2}{x_2 - a_{23}c_3 + 1} + \lambda_2 - \mu_2 &= 0 \quad (A.4b) \\
    \frac{\alpha_3}{c_3 + 1} + \lambda_3 - \mu_1 - \mu_2 a_{23} &= 0 \quad (A.4c) \\
    \lambda_1(x_1 - c_3) &= 0 \quad (A.4d) \\
    \lambda_2(x_2 - a_{23}c_3) &= 0 \quad (A.4e) \\
    \lambda_3 c_3 &= 0 \quad (A.4f)
\end{align*}
\]

From equations (*d), (*e) and (*f), there can be three possible cases: $c_3 = 0$, $c_3 = U$ or $0 < c_3 < U$, where $U = \min\{x_1, \frac{x_2}{a_{23}}\}$.

**Case 1:** $c_3 = 0$ iff $\alpha_3 \leq \frac{\alpha_1}{x_1 + 1} + a_{23} \frac{\alpha_2}{x_2 + 1}$.

**Case 2a:** $c_3 = x_1 < \frac{x_2}{a_{23}}$ iff $x_1 < \frac{x_2}{a_{23}}$ and $\alpha_1 + \frac{\alpha_2 a_{23}}{x_2 - a_{23}x_1 + 1} \leq \frac{\alpha_3}{x_1 + 1}$.

**Case 2b:** $c_3 = \frac{x_2}{a_{23}} < x_1$ iff $\frac{x_2}{a_{23}} < x_1$ and $\frac{\alpha_1}{x_1 a_{23} - x_2 + a_{23}} + \alpha_2 \leq \frac{\alpha_3}{x_2 + a_{23}}$.

**Case 3 (interior solution):** $c_3$ solves: $\frac{\alpha_3}{c_3 + 1} = \frac{\alpha_1}{x_1 - c_3 + 1} + \frac{\alpha_2 a_{23}}{x_2 - a_{23}c_3 + 1}$

This leads to a second-degree equation: $Ac_3^2 + Bc_3 + C = 0$ where:

\[
\begin{align*}
    A &= a_{23}(\alpha_1 + \alpha_2 + \alpha_3) \\
    B &= (\alpha_1 + \alpha_2)a_{23} - (\alpha_1 + \alpha_3)(x_2 + 1) - (\alpha_2 + \alpha_3) a_{23}(x_1 + 1) \quad (A.5) \\
    C &= \alpha_3(x_1 + 1)(x_2 + 1) - \alpha_1(x_2 + 1) - \alpha_2 a_{23}(x_1 + 1)
\end{align*}
\]
Clearly, $A > 0$. In addition, since $\alpha_3 > \frac{a_1}{x_1+1} + a_{23} \frac{\alpha_2}{x_2+1}$, then $B < 0$ and $C > 0$. By unicity of the optimization solution, there can be at most one root between 0 and $U$: therefore it must be the smallest one since it is positive:

$$c_3^{(int)} = \frac{-B - \sqrt{B^2 - 4AC}}{2A} \tag{A.6}$$

Note that in case 1, $C < 0$ and therefore $c_3^{(int)} < 0$. In cases 2a and 2b, one can show that $c_3^{(int)} > U$. We can simplify the solution as:

$$c^*_3 = \min \left( \max \left( 0, c_3^{(int)} \right), x_1, \frac{x_2}{a_{23}} \right) \tag{A.7}$$

A.2. Estimation by Markov Chain Monte Carlo

This section describes our Bayesian estimation algorithm to estimate the discrete model laid out in Chapter 2. We use data augmentation (Tanner and Wong 1987), treating $z_{it}$ as missing data, where $z_{it} = \log(\alpha_{it}) = \omega^i + \epsilon_{it}$.

We specify the following prior:

$$\bar{\omega} \sim MVN(\omega_0, \Sigma_0)$$

$$V \sim IW(\nu, \Delta) \tag{A.8}$$

$$\sigma^2 \sim InvGamma(\alpha, \beta)$$

The algorithm of our Gibbs sampler is provided below.

1. Initialize $\{z_{it}\}$ by solving the FOC of the continuous consumer problem (i.e. without integer constraints).

2. For $i = 1$ to $N$

   (a) Draw $\omega^i|\{z_{it}\}_{t=1}^{T_i}, \bar{\omega}, V$:
\( \omega^i \mid \{ z_{it} \}_{t=1}^T, \bar{\omega}, V \sim MVN \left( S_1 \left[ \frac{1}{\sigma^2} \sum_{t=1}^T z_{it} + V^{-1} \bar{\omega} \right], S_1 \right) \)

where \( S_1 = \left[ \frac{T_i}{\sigma^2} I + V^{-1} \right]^{-1} \)

(b) Draw \( \{ z_{it} \} \mid \omega^i \) by Metropolis-Hastings step using the following likelihood and prior:

\[
Pr(x_{it} \mid \omega^i, p_{it}, z_{it}) = \begin{cases} 
1 & \text{if } x_{it} = \arg\max_x U_x(Sx | \alpha = \exp(z_{it})) - p_{it} \cdot x \\
0 & \text{otherwise}
\end{cases} \quad (A.9)
\]

\( z_{it} \mid \omega^i \sim MVN (\omega^i, \sigma^2 I) \)

3. Draw \( \bar{\omega} \mid \{ \omega^j \}, V \):

\( \bar{\omega} \mid \{ \omega^j \}_{i=1}^N, V \sim MVN \left( S_2 \left[ V^{-1} \sum_{i=1}^N \omega^i + \Sigma_0^{-1} \omega_0 \right], S_2 \right) \) where \( S_2 = (NV^{-1} + \Sigma_0)^{-1} \)

4. Draw \( V \mid \{ \omega^j \}, \bar{\omega} \):

\( V \mid \{ \omega^i \}, \bar{\lambda} \sim IW \left( N + \nu, \Delta + \sum_{i=1}^N (\omega^i - \bar{\omega})(\omega^i - \bar{\omega})' \right) \)

5. Draw \( \sigma^2 \mid \{ z_{it} \}, \{ \omega^i \} : \)

\( \sigma^2 \mid \{ z_{it} \}, \{ \omega^i \} \sim InvGamma \left( \frac{K \sum_{i=1}^N T_i}{2} + \alpha, \frac{\sum_{i=1}^N \sum_{t=1}^{T_i} \epsilon_{it}^2}{2} + \beta \right) \) where \( \epsilon_{it} = z_{it} - \omega^i \)

6. Go to 2.

A.3. Dynamic programming: general approach

In this section, we describe how we solve the Bellman equation in Chapter 3, a necessary step in evaluating the likelihood and predicting demand. The Bellman equation is as follows:

\[
\tilde{V}(i, p, s) = \mathbb{E}_x \left[ \max_{q \in Q(s)} \left( W(i + Zq, p) - C_q(q) + \epsilon(q) \right) \right] \quad (A.10a)
\]

where \( W(i^p, p) = \max_{0\leq y \leq i^p} U_y(y) - C_s(i^p - y) + \delta \mathbb{E}_{q'.s'} \left[ \tilde{V}(i^p - y, p', s') \mid p \right] \) \( (A.10b) \)

The variables of the state space are the prices \( p \), the beginning-of-period inventories \( i \) and...
the shopping trip indicator $s$. Given the continuous nature of prices and inventories, we follow Keane and Wolpin (1994) and construct a grid of points $(i, p, s)$ at which we evaluate the value function $\bar{V}$, and use a regression equation to interpolate the value function between the grid points. Specifically, we approximate $\bar{V}(i, p, s)$ by a parametric form $\hat{V}(i, p, s; \theta)$ where $\theta$ is a parameter, and use an iterative procedure to find the value of $\theta$ such that the Bellman equation is best approximated. More details about the parametric approximation are given in Section A.4.

The procedure is as follows: at each iteration, we apply the Bellman operator to evaluate $\bar{V}$ at all the grid points, using the current value of $\theta$ to interpolate $\bar{V}$ whenever needed; then we update $\theta$ by regressing the latest values of $\bar{V}$ on a set of predictors evaluated at the grid points. We repeat this process a number of times. The output of this procedure is a vector $\theta$ that allows us to approximate the value function $\bar{V}(i, p, s)$ for any state $(i, p, s)$.

Our Bellman operator has two parts, since the consumer makes a purchase decision and then a consumption decision. The consumer’s state at the consumption stage $(i^p, p)$ depends on her initial state at the purchase stage $(i, p, s)$ and her purchase decision ($q$). Since the number of possible purchase decisions is finite and the number of grid points is finite, there are a finite number of states $(i^p, p)$ that can be reached at the consumption stage. Thus we solve all consumption problems to obtain the consumption-stage values $W(i^p, p)$ and then we solve all the purchase problems, making use of the consumption value $W(i^p, p)$ to obtain $\bar{V}(i, p, s)$. Section A.4 focuses specifically on the problem of evaluating the value function $W$ at the consumption stage. The overall algorithm is laid out in Algorithm 1:
1. Initializations
1.1. Construct a grid of \( M \) grid points defined by prices \( p_m \), beginning-of-period inventories \( i_m \), and shopping trip indicator \( s_m \).
1.2. Compute the basis \( B \) which gives, for each grid point, a set of regressors for \( \bar{V} \).
1.3. Initialize \( \theta \leftarrow 0 \)
2. Run \( H \) iterations:
   \textbf{for } \( t = H \) to 1 \textbf{do}
   \hspace{1em} 2.1. Update \( \bar{V} \) at each grid point:
   \hspace{2em} \textbf{for } \( m = 1 \) to \( M \) \textbf{do}
   \hspace{3em} 2.1.1. Solve the consumption problems for all purchase actions \( q \) starting from the \( m \)th initial state:
   \hspace{4em} \textbf{for } \( q = 1 \) to \( Q \) \textbf{do}
   \hspace{5em} \( W_{m,q} \leftarrow \max_{0 \leq y \leq i^p_m} U_y(y) - C_s(i_m + q - y) + \delta \mathbb{E}_{y',s'} \left[ \bar{V}(i^p - y, p', s'; \theta)|p_m \right] \)
   \hspace{4em} \textbf{end}
   \hspace{3em} 2.1.2. Take the Emax:
   \hspace{4em} \( \bar{V}_{m} \leftarrow \text{Emax}_{q \in Q(s_m)} [W_{m,q} - C_q(q, p_m)] \)
   \hspace{2em} \textbf{end}
   \hspace{1em} 2.2. Use regression equation to update \( \theta \) by OLS:
   \hspace{2em} \( \theta \leftarrow (B'B)^{-1} B'\bar{V} \)
\textbf{end}

\textbf{Algorithm 1:} Solving the Bellman equation

A.4. Solution to the consumption problem

One important challenge in the dynamic programming algorithm consists in solving the consumption problem in step 2.1.1 of Algorithm 1. This step is the bottleneck of our algorithm as it needs to be performed \( H \times M \times Q \) times to solve a single dynamic programming problem, where \( H \) is the horizon, \( M \) is the number of grid points, and \( Q \) is the number of possible purchase decisions. With a time horizon \( H \) equal to 100, \( M = 10,000 \) grid points and \( Q = 100 \) possible actions, solving one dynamic programming problem requires one to solve 100 million consumption problem. Therefore this step needs to be performed very efficiently. Unlike in previous models of endogenous consumption where consumption is univariate (Sun, 2005; Hendel and Nevo, 2006b; Gordon and Sun, 2015), the quantities consumed \( y \) are multivariate in our model since we consider more than one category (\( J > 1 \)). Therefore we cannot apply a dichotomous search to find the optimal consumption \( y \). In addition, a grid search in a multivariate space would be too costly computationally.
Instead, we choose a quadratic approximation to the value function $\tilde{V}$, which makes the consumption problem tractable:

$$\tilde{V}(i, p, s; \theta) = \theta_1 + \sum_{r=1}^{R} \theta_{pr} p_r + \sum_{r=1}^{R} \theta_{pr^2} p_r^2 + \sum_{r<r'} \theta_{pr,p_r} p_r p_{r'}$$  \hspace{1cm} (A.11a)

$$+ \sum_{j=1}^{J} \theta_{ij} i_j + \sum_{j=1}^{J} \theta_{i_i} i_j^2 + \sum_{j<j'} \theta_{ij,i_j,i_j'}$$  \hspace{1cm} (A.11b)

$$+ \sum_{r=1}^{R} \sum_{j=1}^{J} \theta_{p_r,i_j} p_r i_j$$  \hspace{1cm} (A.11c)

$$+ s \left[ \theta_s + \sum_{r=1}^{R} \theta_{sp_r} p_r + \sum_{r=1}^{R} \theta_{sp_r^2} p_r^2 + \sum_{r<r'} \theta_{sp_r,p_r} p_r p_{r'} \right]$$  \hspace{1cm} (A.11d)

$$+ \sum_{j=1}^{J} \theta_{s+s_j} i_j + \sum_{j=1}^{J} \theta_{s+s_j^2} i_j^2 + \sum_{j<j'} \theta_{s+s_j,i_j,i_j'}$$  \hspace{1cm} (A.11e)

$$+ \sum_{r=1}^{R} \sum_{j=1}^{J} \theta_{sp_r,i_j} p_r i_j$$  \hspace{1cm} (A.11f)

Let us define $\tilde{V}_2(i^+|p; \theta) = \mathbb{E}_{p^+,s^+} \left[ \tilde{V}(i^+, p^+, s^+; \theta) \right]$, which is a quantity of interest since it appears in the expression of $W$. After taking an integral of $\tilde{V}$ over the distribution of future shopping trips $s^+$ and prices $p^+$ conditional on observing current prices $p$, we can write:

$$\tilde{V}_2(i^+|p; \theta) = \tilde{\theta}_1 + \sum_{j=1}^{J} \tilde{\theta}_{ij} i_j^+ + \sum_{j=1}^{J} \tilde{\theta}_{i_j}^+ i_j^+ + \sum_{j<j'} \tilde{\theta}_{ij,i_j,i_j'} i_j^+ i_j'^+$$  \hspace{1cm} (A.12)

where:

$$\tilde{\theta}_1 = (\theta_1 + \rho \theta_s) + \sum_r (\theta_{pr} + \rho \theta_{sp_r}) E[p_r|p]$$

$$+ \sum_r (\theta_{pr^2} + \rho \theta_{sp_r^2}) E[p_r^2|p] + \sum_{r<r'} (\theta_{pr,p_r} + \rho \theta_{sp_r,p_r'}) E[p_r p_{r'}|p]$$

$$\tilde{\theta}_{ij} = (\theta_{ij} + \rho \theta_{s+i_j}) + \sum_r (\theta_{p_r,i_j} + \rho \theta_{sp_r,i_j}) E[p_r^+|p]$$  \hspace{1cm} (A.13)

$$\tilde{\theta}_{i_j}^+ = (\theta_{i_j} + \rho \theta_{s+i_j^+})$$

$$\tilde{\theta}_{ij,i_j'} = (\theta_{ij,i_j'} + \rho \theta_{s+i_j,i_j'})$$
We can easily evaluate the moments $E[p_t|p], \ E[p_t^2|p]$ and $E[p_t p'_t|p]$ given our VAR(1) model for the price process. Using this definition and our parametric form for $U_y$, the consumption problem can be simplified as:

$$W(i^p, p; \theta) = \max_{c.s.t. Ac \leq i^p} U_c(c) - C_s(i^p - Ac) + \delta \hat{V}_2(i^p - Ac|p; \theta) \quad (A.14)$$

After some algebra, we can rewrite it under the form of a quadratic programming problem:

$$\begin{align*}
\max_{\{c_k\}} & \quad \beta_0 + \sum_k \beta_k c_k + \sum_k \beta_{kk} c_k^2 + \sum_{k<k'} \beta_{kk'} c_k c_{k'} \\
\text{subject to:} & \quad c_k \geq 0 \\
& \quad \sum_{k=1}^K a_{jk} c_k \leq i^p_j
\end{align*} \quad (A.15)$$

where:

$$\begin{align*}
\beta_0 &= \delta \left[ \tilde{\theta}_1 + \sum_j \tilde{\theta}_{ij} i^p_j + \sum_j \tilde{\theta}_{ij}^2 (i^p_j)^2 + \sum_{j<j'} \tilde{\theta}_{ij} i^p_j i^p_{j'} - \sum_j \kappa_j i^p_j - \sum_j \kappa_{jj} (i^p_j)^2 \right] - \delta \sum_j a_{jk} \kappa_j i^p_j + \delta \sum_{j<j'} \kappa_{jj} (i^p_j)^2 \\
\beta_k &= \psi_k + \sum_j a_{jk} \kappa_j + 2 \sum_j a_{jk} \kappa_{jj} i^p_j \\
& \quad - \delta \left[ \sum_j a_{jk} \tilde{\theta}_{ij} + \sum_j a_{jk} \tilde{\theta}_{ij}^2 + \sum_{j<j'} \tilde{\theta}_{ij} i^p_{j'} + \sum_{j<j'} \tilde{\theta}_{ij} i^p_j (a_{j'k} i^p_j + a_{jk} i^p_{j'}) \right] \\
\beta_{kk} &= \psi_{kk} - \sum_j a_{jk}^2 \kappa_{jj} + \delta \left[ \sum_j \tilde{\theta}_{ij}^2 (a_{jk})^2 + \sum_{j<j'} \tilde{\theta}_{ij} a_{jk} a_{j'k} \right] \\
\beta_{kk'} &= -2 \sum_j \kappa_{jj} a_{jk} a_{j'k} + \delta \left[ 2 \sum_j \tilde{\theta}_{ij} a_{jk} + \delta \sum_{j<j'} \tilde{\theta}_{ij} (a_{jk} a_{j'k'} + a_{j'k} a_{jk'}) \right]
\end{align*} \quad (A.16)
We solve the quadratic programming problem by using the first-order conditions:

\[ \beta_k + 2 \beta_{kk} c_k + \sum_{k' < k} \beta_{kk'} c'_k + \lambda_k - \sum_{j=1}^{J} \mu_j a_{jk} = 0 \forall k \]  
\[ \lambda_k c_k = 0 \ \forall k \]  
\[ \mu_j \left( i_j - \sum_{k=1}^{K} a_{jk} c_k \right) = 0 \ \forall j \]

We now develop several special cases considered in our analysis.

**Case 1:** \( J = 2 \) ingredients, \( K = 3 \), 

\[ A = \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & 0 & a_{23} \end{pmatrix} \]

The first-order conditions are the following:

\[ \beta_1 + 2 \beta_{11} c_1 + \beta_{12} c_2 + \beta_{13} c_3 + \lambda_1 - a_{11} \mu_1 - a_{21} \mu_2 = 0 \]  
\[ \beta_2 + 2 \beta_{22} c_2 + \beta_{12} c_1 + \beta_{23} c_3 + \lambda_2 - a_{12} \mu_1 = 0 \]  
\[ \beta_3 + 2 \beta_{33} c_3 + \beta_{13} c_1 + \beta_{23} c_2 + \lambda_3 - a_{23} \mu_2 = 0 \]  
\[ \lambda_1 c_1 = 0 \]  
\[ \lambda_2 c_2 = 0 \]  
\[ \lambda_3 c_3 = 0 \]  
\[ \mu_1 \left( i_1 - a_{11} c_1 - a_{12} c_2 \right) = 0 \]  
\[ \mu_2 \left( i_2 - a_{21} c_1 - a_{23} c_3 \right) = 0 \]

There are five Lagrange multipliers, which can be either zero or positive. Thus there are \( 2^5 = 32 \) cases to consider. Under each case, the optimal value of \( (c_1, c_2, c_3) \) can be obtained analytically. To solve the consumption problem, we enumerate these 32 possibilities, we remove those where the constraints are not all respected, and we keep the argmax.
BIBLIOGRAPHY


P. Rossi, R. McCulloch, and G. Allenby. The value of purchase history data in target


