Tilburg University

Purchase behavior of consumers in emerging markets

van Ewijk, Bernadette

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Purchase Behavior of Consumers in Emerging Markets

Bernadette J. van Ewijk

Purchase Behavior of Consumers in Emerging Markets

Proefschrift

ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof.dr. E.H.L. Aarts, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de Aula van de Universiteit op woensdag 28 november 2018 om 16.00 uur door

Bernadette Johanna van Ewijk

geboren op 15 mei 1985 te Naarden.

Promotores:	Prof. dr. Els Gijsbrechts
	Prof. dr. Jan-Benedict E.M. Steenkamp
Promotiecommissie:	Prof. dr. Yubo Chen
	Prof. dr. Inge Geyskens
	Dr. Katrijn Gielens
	Prof. dr. Harald J. van Heerde

Acknowledgements

Although only my name is printed on the cover, the truth is, I received a lot of help, guidance, and support while writing this dissertation. Therefore I would like to express my most sincere thanks to the following people.

First and foremost, my two supervisors: Els and Jan-Benedict. Dear Els, I cannot thank you enough for guiding me through this tough process. Your kindness, your endless patience, your commitment, your honesty, and your stringency have helped me in many different ways. You gave me the feeling I was never alone and could always fall back on you. Even when you were super busy, you somehow always managed to make time for me: thanks a lot for that. Dear Jan-Benedict, thank you so much for all of your effort and support. Having a call or physical meeting with you usually resulted (besides a long to-do list) in me seeing new possibilities and looking at things from a totally different perspective. Your energy, humor, and positive vibe always cheered me up. Els and Jan-Benedict, I feel honored that you were willing to supervise me. I could not have been luckier with you. I have learned so much from your extensive knowledge, expertise, and experience. I am indebted to the way in which both of you prioritized the finalization of my dissertation: I am aware that if you had not done that, the whole process would have taken me much longer. The support that I received from you both has been invaluable. Thanks a million for everything you have done for me. I am looking forward to our collaboration in the future.

Harald, Inge, Katrijn, and Yubo, I am truly honored that you are willing to serve on my doctoral committee. Thanks for all the comments, remarks, and suggestions you have made. I am grateful for the time and effort you have spent on doing this: your comments, remarks, and suggestions really improved the papers. Harald, thanks a lot for the detailed suggestions you shared with me, not only on my complete dissertation, but also on earlier versions of the papers during conferences and research visits. Also, thank you for your interesting PhD tutorials. Especially the one on endogeneity was very helpful for my dissertation (and most likely also for future projects). Inge, it was after having conversations with you that I dared to start the research master and, a couple of years later, the PhD program. You hired me as your teaching assistant, a time in which I started to learn how much fun it can be to do scientific research. Although the period in which you supervised me was relatively short, due to health issues, I will never forget that you planted the seed for this dissertation. Katrijn, you have helped me in so many different ways. Your critical remarks and suggestions on my dissertation helped me to see things differently and to better think things through. Besides, I learned so many things from you during my time at AiMark: how to use different kinds of estimation techniques, but also (and maybe even more importantly) how to handle large datasets, operationalize variables and construct them in an efficient way – skills that definitely sped up the data preparation time of my papers. Yubo, thank you for all the insights you shared with me about how things work in China. I learned a lot from attending your presentations at conferences and from the conversations we had during these events – how nice that one such time took place in your hometown Beijing! The elaborate answers you gave to my questions and all the background information you provided me with really helped me to better understand Chinese society.

Alfred, I learned so much from you during my time at AiMark: things I still reap the benefits of. Thank you for encouraging me to start a PhD. I am grateful for the space and support you gave me to follow this program alongside of my job. Of course, thanks a lot for the data too: they are of great value for my dissertation.

Ana, Anouk, Astrid, Constant, Esther, Georgi, Jonne, Kristopher, Max, Nick, Suzanne, Yan, and Yufeng, thank you for all of the help and support that I have received from you. Also, a big thank you to all the other colleagues at Tilburg: even though I was an external PhD candidate, you really gave me the feeling I was part of the group. Thanks for all the fun drinks, lunches, dinners, and small talks we had: I enjoyed every single one of them. Special thanks to you Anouk. Over the years you have not only been a very nice colleague, you have also become a dear friend.

Abhishek, Andrea, Frauke, Freddy, Jonne, Karin, Mark, Marlene, Roger, Ruta, Umut, and Willemijn, it is because of you that I quickly got used to my new place in Amsterdam. Starting to teach while having to finish my PhD at the same time was a tough, though interesting, experience for me. Thanks a lot for the help, advice, tips, and tricks I received from each and every one of you: although there is still a lot I need to learn, it really helped me to feel more comfortable while teaching. Jonne, I am happy you introduced me to this group of people, and that you are my office mate now. Thanks for everything you have helped me out with (which is a lot).

Papa, I know you would have been proud of me. Mama, I am grateful for everything I have learned from you. Thank you for every moment you took to inquire about the progress of my PhD. Hetty, Marjan, Gert, Han, and Saskia, thanks for your love and support. Saar, Elon, and Florijn, thank you for bringing so much joy to my life. You help me to realize how important my family is to me.

Margreet, I am so blessed you are my best friend. I am super grateful that you are a part of my life for so many years already. Thank you so much for letting me share with you all the ups and downs I experienced during the PhD program. Pien, thank you for all the interesting conversations we have had: I learned so much from you. Thanks for the love and understanding I have received from you. Florieke, Marielle, and Willeke, thank you for supporting me throughout the whole process. I am happy to be part of such a nice group of friends. To all my friends at the riding school: thank you for providing the necessary distraction, not only every Tuesday during our training sessions, but also at the lunches, dinners, and parties we have every now and then. Special thanks to Arjan, Janine, Jytte, and Marit: riding such a nice horse as Volmoed gives me a great feeling, every time.

Patrick, words cannot express how happy I am to have you by my side. Thank you for your endless understanding when I was working like crazy, especially during the last couple of months in which we also moved to our new home. Thank you for listening to me, not only when I was enthusiastically talking about my work, but also when I was facing difficulties. Thanks for being there for me and for putting a smile on my face. You are deep in my heart.

‘s-Hertogenbosch, September 2018

Table of Contents

Chapter 1 | General Introduction

Emerging markets (EMs) like Brazil, Russia, India, and China are becoming increasingly important for global economic growth. While historically, developed markets (DMs) like France, Germany, Japan, the U.K. and the U.S. had the greatest economic power, this is no longer the case. In 2014, EMs took over DMs based on gross domestic product in purchasing power parity terms. Moreover, EMs are expected to become even more powerful in the future (PwC 2015, 2017): by 2050, six of the seven largest economies in the world could be EMs. Among the EMs, China is by far the most important country: it has undergone a dramatic evolution in the last three decades and has already overtaken the U.S. to be the largest economy in the world (PwC 2017). Speed and change define China – in 1980, China's gross domestic product was $306 billion; in 2015, it exceeded $11 trillion. No country in world history has experienced such a dramatic shift in its economic fortune in such a short time span.

With DMs maturing, global consumer-packaged-goods (CPG) manufacturers know that being successful in EMs like China is crucial for their overall firm performance and growth. Brand manufacturers like The Coca-Cola Company, Procter & Gamble, Nestlé, Colgate-Palmolive and Unilever nowadays generate 40% to more than half of their revenues from EMs. However, operating in EMs comes with important challenges. As pointed out by Burgess and Steenkamp (2006), EMs differ from the Western world in several ways (e.g., socio-economically, demographically, and culturally), and these differences are likely to affect consumers' purchase behavior. As a result, despite their economic attractiveness, many CPG manufacturers experienced that performing well in these markets is far from easy.

Unable to meet expectations, global players like Revlon, L'Oréal's Garnier and Danone Nutricia's Karicare even withdrew their products from the Chinese market. Especially in more recent years, brands increasingly struggled as EMs faced a slowdown in growth (although EMs still grew at a much faster pace than DMs), and competition intensified due to a growing number of players on the CPG market. When making marketing decisions to improve a brand's EM performance, managers can hardly rely on academic research executed on EMs, as the vast majority of academic consumer studies took place in DMs. As such, though some notable exceptions exist (e.g., Batra et al. 2000; Pauwels, Erguncu, and Yildirim 2013; Zhou, Su, and Bao 2002), strikingly little rigorous empirical evidence exists in the area of what drives consumers' purchase behavior in EMs – leaving a large gap within the marketing field.

This dissertation contributes to filling this gap by studying the purchase behavior of consumers in the largest EM in the world, namely China. Throughout the chapters, we develop insights into the effectiveness of the marketing mix, across brands/categories, consumers, and time. By doing so, our goal is to guide brand managers that operate in EMs in setting up successful marketing mix strategies for their brands. In addition, for scholars, we answer the call for more research on EMs to further advance marketing as an academic discipline (Burgess and Steenkamp 2006; Narasimhan, Srinivasan, and Sudhir 2015; Sheth 2011).

To the best of our knowledge, with the three chapters of this dissertation we are the first to empirically analyze the CPG purchase behavior across a diverse set of brands and categories of Chinese consumers. As indicated by Burgess and Steenkamp (2006), compared to DMs, obtaining data from EMs is quite challenging. For the three studies of this dissertation, we have access to unique data on the Chinese market. That is, we have a dataset that tracks the purchases of a large sample (n=40,000) of Chinese urban households for hundreds of brands in a comprehensive set of categories, across multiple years (i.e., between 2011 and 2015). For these brands and categories, advertising spend data is available for the same time span as well. In addition, for a selection of brands and categories, we have access to survey data of 2,764 urban Chinese consumers that was collected in 2014. Combining these datasets allows us to study how marketing mix instruments as well as consumer perceptions influence the decisions Chinese consumers make when buying brands in CPG categories.

In Chapter 2 – “Price Elasticities for CPG Brands in China: Empirical Generalizations from a Large Scale Study” – we focus on one of the most important issues in marketing, namely pricing. Numerous studies have reported price elasticities, leading to empirical generalizations summarized in two important meta-analyses (Bijmolt, Van Heerde, and Pieters 2005; Tellis 1988). However, almost all these studies pertain to developed (Western) markets, not to EMs like China. Success in China has become crucial for Western companies, which requires knowledge of marketing mix elasticities, including first and foremost pricing: competition in China has intensified – leading to a stronger focus on pricing decisions. Yet, it is unclear whether ‘Western’ empirical generalizations apply to China: established brand- and category moderators of price elasticities in DMs may play out differently in China, and other drivers may come into play. Therefore, we conduct a comprehensive analysis of price elasticities for 376 brands in 50 CPG categories over the period 2011-2015 in China. We theorize on, and quantify the moderating effect of, eight category and brand factors, and assess the relative importance of price vs. three other key marketing instruments – advertising, distribution, and line length.

In Chapter 3 – “Consumer Learning about Quality of Global and Local Brands in the CPG Industry in China” – our interest is in better understanding the brand choices that EM consumers make over time. Brands are deemed to play a large role in EMs. However, the drivers of consumers' brand choice in these dynamic and heterogeneous markets are not yet well understood. We study the effects of brand quality and quality uncertainty on brand choice behavior, for global vs. local brands. In particular, we study whether Chinese consumers attach different quality beliefs and/or uncertainties to global vs. local brands, and we also investigate how important quality and uncertainty are in driving brand choice, compared to other marketing mix instruments such as distribution and price. In addition, we explore whether differences exist across consumers with different geographic and sociodemographic profiles with respect to both their global vs. local brand quality (uncertainty), as well as to the importance of quality (uncertainty) and other marketing mix instruments when making a brand choice. To this end, we use our scanner panel dataset of urban Chinese households over the period 2011-2014 to estimate a Bayesian learning model on five product categories.

Chapter 4 – "The Rise of Online Grocery Shopping: Which Brands Will Benefit?" – studies how the rise of e-commerce in grocery affects brand performance. With China being one of the most important countries fueling the worldwide online grocery trend, we look at how brand managers can make sure to benefit from this trend. We derive how a brand's total (online plus offline) sales change as the fraction of groceries sold online goes up, and show that it critically depends on two indices: (i) the brand's online index (BOI) and (ii) the category's online index (COI). While the former indicates how the brand's relative position within the category will evolve, the latter indicates how the category's overall CPG share will contribute to (or hamper) brand sales as the online CPG channel grows. We then identify brand and category factors that drive these indices. We estimate our model on 448 brands in 60 product categories, using 2011-2015 data – a period in which the online channel took off in China.

Chapter 5 summarizes the main research findings. In addition, it reflects on the most important implications, and formulates recommendations for brand managers operating in China. It also discusses the limitations of our analyses, and suggests potential directions for further research.

Chapter 2 | Price Elasticities for CPG Brands in China: Empirical Generalizations from a Large Scale Study

Introduction

Price is among the most important and widely studied areas of marketing scholarship (Gordon, Goldfarb, and Li 2013, p. 4). Two influential meta-analyses (Bijmolt, Van Heerde, and Pieters 2005; Tellis 1988) develop empirical generalizations on the overall level of price elasticity and its moderators. However, all pricing studies in their meta-analyses present empirical findings for Western countries. While historically, that might be understandable given the overwhelming economic preponderance of the West, this is no longer the case. Since 2000, the share of emerging markets (EMs) in global GDP has increased from less than 40% to nearly 60%. Along the way, EMs have become ever more important for the Western companies. Companies like P&G, Nestlé and Unilever derive 40% to more than half of their sales from EMs. Faced with declining sales at home, Coca Cola and Pepsi Cola are more than ever looking to EMs for growth. According to CSPI (Center for Science in the Public Interest), the soft-drink companies are “spending several billions of dollars a year in such countries as Brazil, China, India, and Mexico to build bottling plants, create distribution networks, and advertise their products” (Center for Science in the Public Interest 2016, p. VII).

However, as observed by Narasimhan, Srinivasan, and Sudhir (2015), our field knows strikingly little about the effectiveness of price in EMs, including China, which is the focus of our study. While EMs as a whole have become economically important, China looms larger than any other. Within an unprecedented short period of 35 years, China has become the world's largest economy in purchasing power parity terms (PwC 2017). It is unclear whether received 'Western' empirical generalizations on the magnitude and the moderators of price elasticity are applicable to China. Perhaps there is little difference in overall price sensitivity, which is an important finding in its own right. Alternatively, the difference may be substantial, which is also noteworthy. How do price elasticities vary in function of category and brand characteristics in China? What is the effect of 'established' moderators (i.e., documented in research conducted in developed markets)? Might there be moderators that are more or less unique to China, or EMs in general? If so, what is their effect, both in an absolute sense and relative to 'Western' moderators? Another question that emerges is: How important is price vs. other marketing mix instruments in affecting brand market share? Is price more or less influential than instruments like advertising, assortment (line length), or distribution?

These questions motivated the present study. The overriding goal of the study is to provide empirical generalizations about the magnitude and moderators of price elasticity in China. Our study covers 376 brands in 50 consumer packaged goods (CPG) categories, which should provide a promising basis to derive empirical generalizations on price elasticity in China. To compare, the Tellis (1988) (Bijmolt, Van Heerde, and Pieters 2005) meta-analysis included 367 (1,851) price elasticities of which 76% (98%) were for CPG brands, spanning the U.S., Canada, various European countries, Australia, and New Zealand. In our work, we combine scanner panel and advertising data over a five-year period, with consumer and expert survey data. In a first stage, we assess the market share-price elasticities for each brand, using a modeling approach that takes into account the inherently dynamic emerging-market setting (Burgess and Steenkamp 2006). Next, we examine how the brand price elasticities differ systematically in function of brand and category-specific characteristics.

The remainder of the paper is organized as follows. We first introduce our research framework, and briefly outline the expected effect of the moderators of price elasticity in China. Next, we describe the modeling approach and the data. Then, we present our findings. We conclude with a discussion of the results where we also compare our findings with the predicted average price elasticity (taking into account study characteristics) in the U.S. using the parameter estimates presented by Bijmolt, Van Heerde, and Pieters (2005, Table 2). We provide managerial implications, and reflect upon how the results for China can be modified to approximately gauge (in the spirit of Raju 2005, p. 18) what magnitude of price elasticity managers can broadly expect in the other three BRIC countries (Brazil, Russia, and India). We conclude with limitations and give directions for further research.

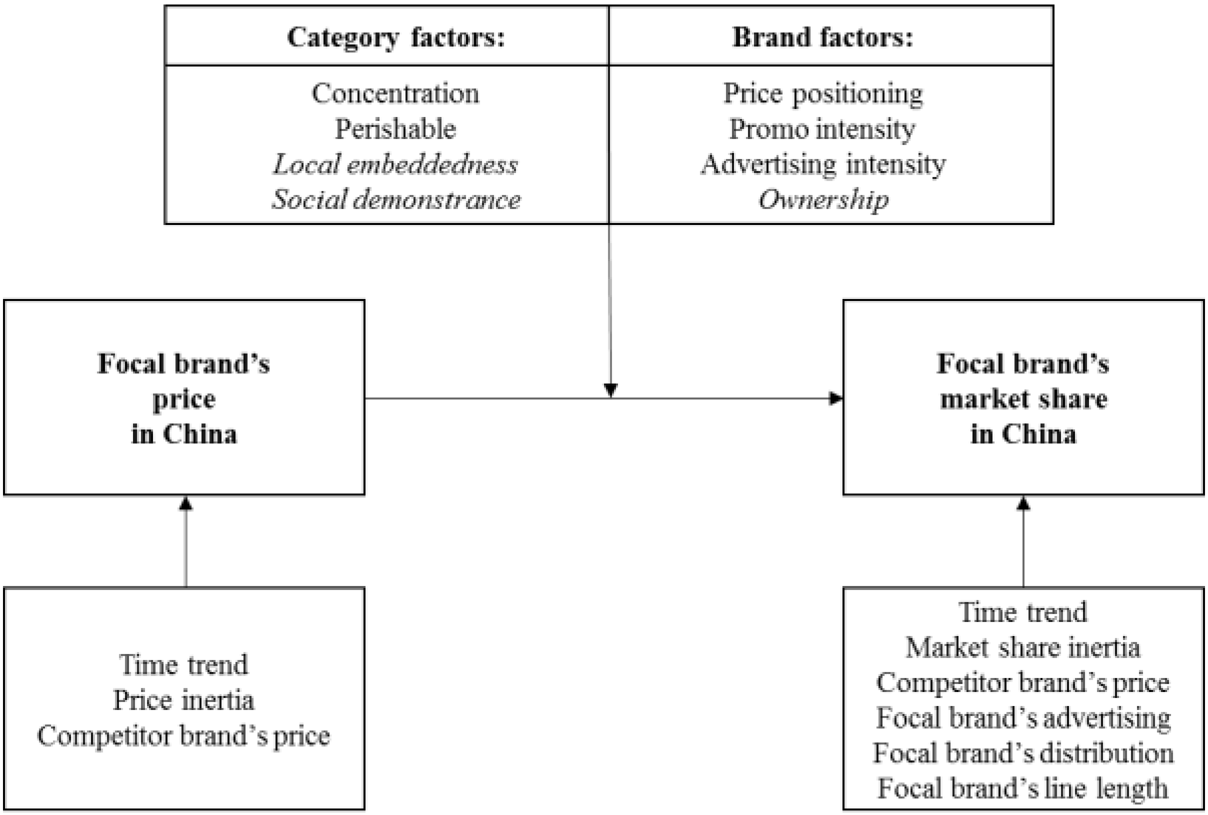
Research Framework

Figure 2.1 provides a schematic overview of the major aspects of our study. We begin by discussing price sensitivity in China. Next, we develop a rationale for the moderating effects of category and brand characteristics.

Price elasticity in China

Extant literature and industry reports provide mixed signals on the price sensitivity of Chinese consumers. On the one hand, one could expect Chinese consumers to be strongly price focused. Tighter budgetary constraints may command them to seek out low prices (Burgess and Steenkamp 2006). Moreover, because Chinese markets are often less efficient (a market being efficient if all relevant and ascertainable information is widely available to participants and all the information changes are reflected in price changes), consumers may possess weaker price-quality schemas and use price to a lesser extent to infer product quality (see Zhou, Su, and Bao 2002 for evidence on China vs. the U.S.).

Figure 2.1: Research framework[a]



[a] In italics are category and brand factors that are more or less uniquely relevant to EMs.

On the other hand, other market characteristics may induce a focus away from price. In China, brands are seen as an important sign of quality and status (Kantar Millward Brown 2010), and social signaling is often deemed important in China (Zhu 2013). The importance of brands may also follow from risk avoidance. Deceptive advertising, trademark violation and the practice of selling poor-quality products at high prices have occurred in China at a large scale and, though the situation has improved over time, consumer protection against low-quality products is still lower (Sudhir et al. 2015). In such a context, the guarantee provided by a well-known brand name is highly valued (Batra 1999). Finally, increased purchasing power has made many Chinese consumers develop a preference for premium brands. For example, BCG (2008) reports that 50% of Chinese consumers who purchase premium products state they buy a product because of its brand name (compared to 33% in the U.S. and only 20% in Western Europe).

Factors affecting the magnitude of price elasticity

Building on the framework developed by Bijmolt, Van Heerde, and Pieters (2005), we consider several category and brand characteristics as moderators of price elasticity (note that, because our large-scale study uses a unified data set and modeling approach across all brands and categories, we do not need to control for methodological differences here).

At the category level, we will study market concentration, perishability, embeddedness in local (Chinese) culture, and social demonstrance. The first two moderators have been examined before in Western studies. However, the last two have not been considered, perhaps because they may not be seen as particularly relevant in mature markets. At the brand level, we will consider three key marketing mix instruments – price positioning, promotional intensity, and advertising intensity. We add to this set brand ownership (foreign vs. domestic), which might be especially pertinent in EMs (Batra et al. 2000). In the discussion below, we will provide a more elaborate rationale for the three moderators that have not been considered in detail in Western studies on price elasticities – local embeddedness, social demonstrance, and brand ownership – while only briefly discussing the more established moderators.

Category moderators

Market concentration. Under the assumption that low-concentration markets consist of homogenous goods and that consumers are fully informed about prices, an economist would argue that in low-concentration markets, demand will be more price elastic (as full competition will increase the importance of price) than in high-concentration markets (where suppliers can set the price). However, in our current study we look at CPG markets in which, even in case of low concentration, manufacturers still strive (and manage) to differentiate their products from competitors, and consumers' price knowledge is less than perfect. In such a setting, several studies have shown that highly concentrated categories experience stronger price effects than less concentrated categories (e.g., Narasimhan, Neslin, and Sen 1996; Nijs et al. 2001) because processing price information in such categories is easier, or because high concentration is actually indicative of more homogenous product (taste)s. There is no obvious reason why this would be different in the CPG industry in China. Hence, we expect price elasticities in China to be larger in magnitude in categories where market concentration is higher.[1]

Perishability. Consumers generally respond more weakly to price changes of perishable (compared to non-perishable) products, because these cannot be stockpiled (Narasimhan, Neslin, and Sen 1996). Therefore, we expect price elasticities of brands in perishable categories to be smaller in magnitude.

Local embeddedness. Local embeddedness of the category is the extent to which consumers perceive the category to be typically Chinese and originating from China. For example, tea and baijiu (distilled alcoholic beverage – Moutai being the most famous brand) have been around for ages and are more deeply embedded in Chinese society than coffee or wine. Serge Dumont, vice chairman at the advertising company Omnicom Group, described China in 1985: “People in those days didn't eat chocolate, they didn't know what a contact lens was. So it was not just trying to convince them to buy this brand versus another, you had to educate about what the product was” (Doland 2015).

While CPG categories ranging from laundry detergents and shampoo to coffee and chocolate have been part of the Western marketing scene for many decades, these anecdotal examples illustrate that in China, this is often not the case. Indeed, consumers have only recently begun to adopt some CPG categories, such that the distinction between vested and newer-to-the-country categories is potentially important. For one, consumers will be more

[1] In this paper, we use terminology based on the (absolute) magnitude of price elasticities. For example, we label a change in price elasticity from -.5 to -1 as in increase in magnitude.

familiar with categories that are deeply locally embedded and have been part of the Chinese consumptionscape for many decades, if not centuries. Familiarity with a product category is associated with lower risk (Song and Schwarz 2009), and consumers are more price-conscious (and thus more price sensitive) in categories they are more knowledgeable about (Bronnenberg et al. 2015). Moreover, categories with deeper embeddedness typically show more intense competition – players having been around for a long time, and category expansion often being lower – which may increase the focus on price in the firms' marketing mix, and heighten the price responsiveness of consumers. As such, we propose that the magnitude of the price elasticity is larger in categories with a stronger local embeddedness.

Social demonstrance. Social demonstrance refers to the use of brands as a symbolic device to project and communicate one's self-concept (Fischer, Völckner, and Sattler 2010). Fischer and colleagues document that higher levels of social demonstrance of a category render brands in that category more relevant to consumers, and increase their willingness to buy the preferred brand at a higher price. Social demonstrance has not played a major role in Western research on price elasticities. Most research on price elasticities involves CPG (Bijmolt, Van Heerde, and Pieters 2005) and Western consumers see little social demonstrance value in CPG brands (Fischer, Völckner, and Sattler 2010, Table 5; Kumar and Steenkamp 2007).

The situation in China is different. Anecdotal evidence suggests that CPG can have significant social demonstrance value in EMs (Dawar and Chattopadhyay 2002). For example, in an award-winning case study, Guimaraes and Chandon (2007) describe how detergent brands play an important social signaling role for many Brazilian consumers. Later in the paper, we will present evidence that CPG brands also have a high social demonstrance value in China. Moreover, China can be characterized as a collectivistic culture (Hofstede, Hofstede, and Minkov 2010), in which purchase decisions are heavily influenced by opinions of friends and family members, and 'face' and social status are crucial (Zhu 2013). De Jong, Steenkamp, and Fox (2007) reported that out of 11 countries, China rates highest on average on susceptibility to normative influences – the need to enhance one's image in the opinion of significant others through the acquisition and use of products and brands (Bearden, Netemeyer, and Teel 1989). Brands' 'signaling utility' may be an important consumption driver (Sudhir et al. 2015). This may dissuade consumers from purchasing cheap brands. Building on these considerations, we expect a magnitude-decreasing effect of social demonstrance on price elasticity.

Brand moderators

Brand price positioning. The brand's price positioning (i.e., its price level relative to the average price of other brands) distinguishes cheaper from more expensive brands in a category. To consumers, a price decrease (increase) of a more expensive brand might have a greater effect, because it may bring the brand within (out of) economic reach. Indeed, studies in developed markets report stronger effects of price changes for brands with a high price level (e.g., Fok et al. 2006). We therefore expect price elasticities of more expensive brands to be more negative.

Brand promotion intensity. High promotion activity makes consumers more price sensitive (Van Heerde et al. 2013; Mela, Gupta, and Lehmann 1997). One reason is that the intensive use of promotions decreases consumers' reference prices: consumers expect to obtain the brand for a reduced price and are willing to pay less (Mazumdar, Raj, and Sinha 2005). Also, promotions might affect the salience of the brand's price and thus the price sensitivity (Boulding, Lee, and Staelin 1994). There is no compelling a priori reason why these mechanisms would not apply to China as well. Therefore, we anticipate that promotion intensity has a magnitude-increasing effect on brand-price elasticity.

Brand advertising intensity. Brand advertising generally leads to lower price sensitivity (Mela, Gupta, and Lehmann 1997). Advertising could work as a shield against price competition: through advertising, a brand can differentiate itself from its competitors by emphasizing its unique benefits (Boulding, Lee, and Staelin 1994). Following this line of research, we expect that brand advertising has a magnitude-decreasing effect on brand-price elasticity.

Brand ownership. Foreign brands are brands owned by a manufacturer that originates from outside the country, whereas domestic brands are owned by a domestic manufacturer. Especially in EMs like China, consumers might respond differently to price changes of foreign vs. domestic brands, though it is not clear a priori whether the response will be weaker or stronger. On the one hand, to the extent that foreign brands are generally stronger brands (Steenkamp 2014) that enjoy a ‘status preference’ (Batra et al. 2000), Chinese consumers may be willing to pay a premium for these brands (Bain & Company and Kantar Worldpanel 2012). Moreover, many Chinese consumers are first-time buyers in a product category. These consumers often gravitate towards big brand names and demonstrate lower price elasticity (Heilman, Bowman, and Wright 2000).

On the other hand, because consumers generally have a home-country bias (i.e., a more positive attitude to their own country; Shimp and Sharma 1987), one could expect them to be less price sensitive for brands perceived to originate from their own country. Also, local brands might attract consumers with a ‘local identity’ (i.e., interested in local culture and identifying with people in their local community), who are shown to have lower price sensitivity (Gao, Zhang, and Mittal 2017). Finally, Chinese consumers may be less familiar with foreign brands, which may decrease their willingness to pay for these brands (Erdem, Swait, and Louviere 2002). Because it is not clear upfront which of these forces dominates, we formulate no expectations on the direction of the effect.[2]

Table 2.1 summarizes our expectations.

Table 2.1: Expected moderating effects on the brand price-market share relationship

DRIVER	EXPECTED SIGN[a]
CATEGORY	
Concentration	-
Perishable	+
Local embeddedness	-
Social demonstrance	+
BRAND	
Price positioning (High end)	-
Promotion intensity	-
Advertising intensity	+
Ownership (Foreign vs. domestic)	+/-

[a] A positive sign means we expect the price elasticity to become less strong, i.e., less negative.

Methodology

Developing an approach to answer our research questions comes with several challenges. First, even for CPG products, EMs are still evolving, and this inherently dynamic market setting calls for a methodology that accommodates possible non-stationarity of our focal time series. Second, especially in these markets, sellers may still be experimenting with price/adjust prices in response to a change in performance, and accommodating these changes is necessary to obtain unbiased estimates of the price effects. Third, given these dynamics, it is important to account for longer-term effects (i.e., delayed reactions and inertia) in the model specifications. Fourth, as brands are still fighting to establish their position, our focus is on market share, and the (possibly complex) interplay between brands should be accounted

[2] Note that local embeddedness and brand ownership are not necessarily intertwined. Bain & Company and Kantar Worldpanel (2012) report that foreign manufacturers are doing well in traditional Chinese categories like candy and biscuits, while a lot of domestic manufacturers are entering or even starting up less locally embedded categories like ice cream and liquid detergents. For example, while laundry detergent typically used to be in powder or bar form in China, the domestic brand Bluemoon started the liquid detergent category quite some time before foreign brands like Omo and Tide entered this category. Empirical evidence that these two factors do not strongly overlap in our data can be found in the 'Results' section.

for. Fifth, although we seek to measure the impact of price changes, other factors might change over time as well, and must be controlled for. Finally, our aim is to develop empirical generalizations, so our approach should be able to handle a large number of categories and brands.

To address these challenges, our methodology consists of two stages. In the first stage, we obtain the brand-price elasticities by estimating a system of equations for each brand in each category, using weekly observations. Our dependent variables are the brand's market share within the category (measured in volume units), and its price (per volume unit). By estimating the two equations as a 'structured' system of equations (see also below), we control for possible price endogeneity and are sure to separate the price effect from common unobserved (price and market-share) drivers (see Van Heerde, Gijsbrechts, and Pauwels 2015 for a similar approach). In the second stage, we explore the link between these price elasticities and several category and brand characteristics. Below, we discuss these stages in turn.

First-stage Analysis

Unit roots. Before setting up the system of market share-price equations, we test for the presence of unit roots in each brand's performance and marketing mix variables, using the Enders procedure (Enders 2004). For variables with a unit root, we use first differences, other variables are expressed in levels. Thus, if some variables in an equation have a unit root and others do not, the equation is a mixture of levels and differences.

Market share equation. To ensure logical consistency (market shares ranging from 0 to 1, and summing to 1 across brands), we use an attraction specification. This model expresses a brand's market share as the ratio of its attraction divided by the attraction of all brands in the category (i.e., our selected brands, plus the 'outside option' comprising all other brands grouped in a 'rest' brand; Besanko, Gupta, and Jain 1998). We include a trend and two lagged dependent variables to control for deterministic long-term changes and inertia. To flexibly capture the price effects, we use a 'fully extended' specification in which a brand's attraction depends not only on its own price, but also on that of competitors – thereby allowing for differential cross-effects between brand pairs (Cooper and Nakanishi 1988). Both own- and cross-prices have an immediate and a lagged impact (i.e., we allow consumers to have a delayed response to price changes). To obtain valid estimates for the price effects, and to assess its impact relative to other important marketing mix instruments, we also control for advertising, distribution, and line length. While endogeneity in price (which is our focal variable, and one that can be easily adjusted) is accommodated through the system of market-share and price equations, we deal with possible endogeneity in the other marketing mix variables through the Gaussian copula method (Park and Gupta 2012).[3]

We linearize the model using the ratio method, with the market share of the 'outside option' or 'rest brand' (market $share_{ot}$) as the reference (see the 'Data' section for a description of which brands are selected). If none of the brand's (market share and marketing mix) variables has a unit root, this leads to the following expression (Equation 2.1):

$$\text{(2.1)} \quad \log m_{jt} - \log m_{0t} = \beta_{m0j} + \beta_{m1j} \log trend_t + \beta_{m2j} \log m_{jt-1} + \beta_{m3j} \log m_{jt-2} + \beta_{m4j} \log p_{jt} + \beta_{m5j} \log p_{jt-1} + \sum_{i,i\neq j} \beta_{m4ji} p_{it} + \sum_{i,i\neq j} \beta_{m5ji} p_{it-1} + \psi_{m1j} a_{jt} + \psi_{m2j} \log d_{jt} + \psi_{m3j} \log l_{jt} + \sum_k \delta_{kj}\, copula_{kjt} + \varepsilon_{jt}$$

where i and j are brand indicators, and

m_{jt} = volume market share brand j in week t;

m_{0t} = volume market share outside option in week t;

β_{m0j} = brand-specific intercept for brand j;

[3] The Gaussian copula for marketing mix variable K_{jt} of brand j in week t, is defined as $copula_{kjt} = \Phi^{-1}(H(K_{jt}))$, where Φ^{-1} is the inverse distribution function of the standard normal, and $H(\cdot)$ is the empirical cumulative distribution function of K_j. The Gaussian copula method requires that the endogenous regressors are not normally distributed. Shapiro-Wilk tests at $p < .10$ formally confirm this for 93% of the cases.

p_{jt} = price of brand j in week t;

a_{jt} = advertising (measured as Adstock) of brand j in week t;

d_{jt} = distribution of brand j in week t;

l_{jt} = line length of brand j in week t;

$copula_{kjt}$ = Gaussian copula for marketing mix variable k of brand j in week t;

ε_{jt} = normally distributed error term for brand j in week t.

Price equation. Our interest is in the market share equation, but we need the price equation to control for endogeneity. Bijmolt, Van Heerde, and Pieters (2005) documented that failure to control for price endogeneity can lead to serious underestimation of the magnitude of the price elasticity. For price, we use a log-log specification, in which own- and cross-prices as well as market share have two lags (i.e., we allow sellers to change prices in response to a change in own price, competitor's price, or market share, one or two weeks ago). In case of stationary variables, the price equation then looks as follows:

(2.2) $\log p_{jt} = \beta_{p0j} + \beta_{p1j} \log trend_t + \beta_{p2j} \log m_{jt-1} + \beta_{p3j} \log m_{jt-2} + \beta_{p4j} \log p_{jt-1} + \beta_{p5j} \log p_{jt-2} + \sum_{i,i\neq j} \beta_{p4ji} p_{it-1} + \sum_{i,i\neq j} \beta_{p5ji} p_{it-2} + \xi_{jt}$

If some variables pertaining to a brand have a unit root, expressions (2.1) and (2.2) are maintained, but after replacing those variables by their 'differenced' counterpart (that is: same-week minus last-week level). Appendix 2.A provides the exact expressions for different combinations of (stationary and non-stationary) price and market share settings. Details on the operationalization of the variables will be given in the 'Data' section.

Estimation approach. We estimate the equations using a Seemingly Unrelated Regressions (SUR) approach, i.e., allowing ε_{jt} and ξ_{jt} to be correlated. To account for possible autocorrelation (within each brand over time), we use Feasible Generalized Least Squares (FGLS). To avoid overparameterization, we use Carpenter et al. (1988)'s three-step procedure to identify which cross-brand price effects to include in the final model: we (i) first estimate the model as shown in equations (2.1) and (2.2) but excluding cross-price effects; then (ii) regress the residuals of that model on all possible competitor prices, and determine which cross-effects reach significance, and then (iii) re-estimate the model after retaining only the significant cross-price effects.

Second-stage Analysis

Having estimated the market share and price models, we examine the pattern of price effects across categories and brands. For each brand, the price elasticity is calculated as follows (Cooper and Nakanishi 1988):

(2.3) $\eta_{m_j p_j} = \beta_{m4j}(1 - \overline{m_j}) - \sum_{i,i \neq j} \beta_{m4ij} \overline{m_i}$

where $\overline{m_j}$ ($\overline{m_i}$) is the average market share of brand j (competitor i) across the data period.

Next, we 'stack' these elasticities, across all brands, and use them as dependent variable in a second-stage regression to link them to brand and category characteristics. More specifically, we estimate the following regression:

(2.4) $\eta_{m_j p_j} = \gamma_o + \gamma_1 co_{c(j)} + \gamma_2 pe_{c(j)} + \gamma_3 le_{c(j)} + \gamma_4 sd_{c(j)} + \gamma_5 pp_j + \gamma_6 pi_j + \gamma_7 ai_j + \gamma_8 fo_j + e_j$

where

$co_{c(j)}$ = concentration of category c to which brand j belongs;

$pe_{c(j)}$ = whether category c to which brand j belongs is perishable (1) vs. non-perishable (0);

$le_{c(j)}$ = local embeddedness of category c to which brand j belongs;

$sd_{c(j)}$ = social demonstrance of category c to which brand j belongs;

pp_j = price positioning brand j;

pi_j = promo intensity brand j;

ai_j = advertising intensity brand j;

fo_j = whether owner of brand j is foreign (1) vs. domestic (0);

e_j = random component.

Because the brand's price elasticity is an estimated quantity, the random component e_j comprises two parts: (i) the measurement (sampling) error r_j – the variance of which ω_j^2 is brand-specific and can be calculated based on the variance-covariance matrix of the brand's parameter estimates in the first stage – and (ii) the part of the elasticity not explained by the brand- and category-drivers v_j – with unknown variance σ^2. Or: $e_j = r_j + v_j$. To account for this error structure, we use the FGLS estimation approach proposed by Lewis and Linzer (2005), which is efficient and produces consistent standard errors, irrespective of the size of σ^2 and the ω_j^2's.[4]

Data

Sources

We obtained our data through Kantar Worldpanel, Kantar Media, and GfK. The purchase data come from a Chinese urban household panel (n=40,000) that tracks the panelists' purchases in CPG categories between 2011 and 2015. In addition, for a selection of 62 categories, we obtained monthly advertising spending data on the top (15) brands. From these data, we retain brands based on the following criteria: (i) the brand has to be sold nationwide, (ii) it has to be present in the market across the entire data period[5], (iii) the brand has to be sold in at least 90% of the weeks, and (iv) the category has to have a minimum of three brands. This leaves us with 377 brands in 50 categories for which the market share-price elasticities will be estimated. For an overview of the selected categories and number of selected brands per category, see Appendix 2.B. In addition, 46 categories were part of a

[4] This approach is a refinement of the commonly used WLS procedure with observation weights $\frac{1}{\omega_j}$. We used this WLS procedure as a robustness check, and found the pattern of results to be similar.

[5] In total, 19 brands were not present in the market across the entire data period: 18 brands in 17 categories entered and 1 brand left the market.

consumer survey administered by GfK in 2014 to 2,764 urban Chinese consumers. The four social demonstrance items were part of the survey. On average, 92 respondents rated each category on social demonstrance. In addition, we surveyed experts about category characteristics. We use these survey measures averaged across respondents/experts to quantify some of the category- and brand-drivers of price elasticity.

Measurement

The operationalization of the variables is described in Table 2.2. In the first stage, market share is calculated based on volume sales (e.g., milliliters, grams). For the price variable, we use price per volume unit (converted into real prices using China's category-specific consumer price index). The advertising variable measures share of voice, that is: the % Adstock that a brand captures relative to the category's Adstock in a certain week, where Adstock is a weighted average of previous Adstock and current Ad spending, with weights equal to λ and $(1 - \lambda)$, respectively (where ad spending is converted into real prices using China's consumer price index). The smoothing constant λ is obtained via a grid search in the first model estimation step, as the one that provides the highest R^2. Distribution is calculated as the percentage of offline retailers that carry the brand, weighted by the retailers' market share. Line length measures the percentage of the number of stock keeping units (SKUs) in the category that belong to the brand.

In the second stage, local embeddedness in China is coded by 5 (native Chinese) judges (Cronbach's alpha .94); social demonstrance was part of the consumer survey and is available for 46 out of the 50 categories under study (Cronbach's alpha .88). Whether the brand's owner is Chinese (domestic) or not (foreign) is coded by consulting the brands' websites. Category concentration is calculated as the sum of the market shares of the top 3 brands in the category across 2011-2015; perishable vs. non-perishable is coded by 7 (Dutch) judges.

The brand's price positioning is obtained as the average of a brand's weekly price index across 2011-2015; we use an index to allow for meaningful comparison of brand prices between categories with different volume units (e.g., milliliters, grams). The index is calculated by dividing the weekly brand prices by the average category price in a base week (the first observation week). A price index above (below) unity thus indicates that the brand is more (less) expensive than the category average in the base week.

Advertising intensity equals the average weekly spending (in ¥) on all media across 2011-2015. Because this variable is highly skewed, we use its log-transform in the second-stage analysis (after adding a small number to accommodate cases with zero advertising). Finally, promo intensity is quantified as the average % (across retailers and weeks) of a brand's SKUs on promotion at a top 3 retailer in a given week, with retailer weights equal to their market share.

Results

Descriptives

Table 2.3, Panel A displays summary statistics across brands, for the outcome variable (market share) and our focal marketing-mix instrument (price), as well as the other marketing mix instruments (advertising, distribution, and line length). As the table shows, our data cover a wide variety of brands, both in terms of market position (with a market share average of 8.62%, and standard deviation of 10.57%, across brands) and price level relative to other brands in the category (the price index for included brands is 1.04 on average, with a standard deviation of .45). Also, within each brand, market share and price vary over time (as indicated by the coefficient of variation, which amounts to .32 for market share, and to .11 for price) – corroborating the dynamic nature of the market. Table 2.3, Panel B, displays summary statistics and a correlation table for the drivers of price elasticity. Again, these measures show quite some variation across our categories and brands, and relatively little overlap – making them suitable for our second-stage analysis.

Unit root tests

Appendix 2.C provides a summary of the unit-root test outcomes for the different variables, across the studied brands. Zooming in on our focal constructs (price and market share), we find both variables to be stationary in only 39.0% of the cases, while 25.5% have a unit root for price but not market share, 17.8% have a unit root for market share but not price, and 17.8% of the brands have a unit root for both variables. As indicated in Appendix 2.A, this results in four specifications of the market share-price equations.

Marketing mix elasticities in the Chinese market

Table 2.4 provides an overview of the elasticities based on the estimation results for the brand-specific market share-price systems of equations.[6] Our interest is in the market share equation. The price equation was included to control for endogeneity. We find that the average price elasticity in China is -.51, indicating that a 1% increase in brand price entails a .5% decrease in the brand's category share within the same week[7]. Yet, there is large heterogeneity in price elasticities, as shown in Figure 2.2. For 18% of the brands, demand is elastic. This heterogeneity suggests the presence of moderators, to which we will turn in the next subsection.

[6] One brand was removed from the analysis because its dynamic effects lacked face validity. The reason for this outlier might be a combination of 1) the brand having a very dominant position in the category (i.e., market share of about 70%) and 2) having very low variation in price over time (coefficient of variation equals .04).
[7] The lagged dependent variables in the market share equation allow us to calculate the price elasticity in the longer term. The average price elasticity in the medium term (12 weeks, i.e., 1 quarter) equals -.62. The average long term price elasticity can only be calculated for a subset of brands (i.e., for brands that have a unit root in both log market share and log price or that have no unit root in log market share nor log price) and equals -.59.

Table 2.2: Operationalization market share and marketing mix variables

VARIABLE	SOURCE	OPERATIONALIZATION	REFERENCE
FIRST STAGE			
Market share (m_{jt})	Kantar Worldpanel	Total volume sales (e.g., milliliters) of brand j in week t relative to category total volume sales in week t.	
Price (p_{jt})	Kantar Worldpanel	Absolute price, calculated as price (in ¥) per volume unit (e.g., per milliliter), of brand j in week t (converted into real prices using China's category-specific consumer price index, source: National Bureau of Statistics China).	
Advertising (a_{jt})	Kantar Media	Share of Voice, calculated as Adstock of brand j in week t ($Adstock_{jt}$) relative to the Adstock of the category to which brand j belongs in week t ($Adstock_{c(j)t}$), where: $Adstock_{jt} = (1-\lambda)*Advertising_{jt} + \lambda*Adstock_{jt-1}$; and $Adstock_{c(j)t} = (1-\lambda)*Advertising_{c(j)t}+ \lambda*Adstock_{c(j)t}$ (where advertising spend by the brand ($Advertising_{jt}$) or category ($Advertising_{c(j)t}$) is converted into real prices using China's consumer price index, source: National Bureau of Statistics China). The optimal λ is found in the first step of the estimation approach via grid search (on the interval [0, .9] in increments of .1).	Datta, Ailawadi, and van Heerde (2017)
Distribution (d_{jt})	Kantar Worldpanel	Weighted average of indicators of availability (0 vs. 1) for brand j in the four-weekly period to which week t belongs across all offline retailers, weighted by the retailers' market shares in the four-weekly period to which week t belongs.	Sotgiu and Gielens (2015)
Line length (l_{jt})	Kantar Worldpanel	Total number of unique SKUs that brand j offers in the four-weekly period to which week t belongs, relative to category total number of unique SKUs in the four-weekly period to which week t belongs.	Srinivasan et al. (2004)

SECOND STAGE			
Concentration ($co_{c(j)}$)	Kantar Worldpanel	Sum of market shares of top 3 brands in category c across 2011-2015.	Steenkamp and Geyskens (2014)
Perishable ($pe_{c(j)}$)	Expert survey	Dummy variable equal to 1 if majority of judges coded category c as perishable, 0 otherwise.	
Local embeddedness($le_{c(j)}$)	Expert survey	Average of 3 items that were rated from 1=very strongly disagree to 7=very strongly agree: - This category does not originate from China (reversed before calculation) - This category is typically Chinese - This category has been around in China for a long time (Cronbach's alpha: .94).	
Social demonstrance ($sd_{c(j)}$)	GfK consumer survey (subset of 46 categories only)	Average of 4 items that were rated from 1=very strongly disagree to 7=very strongly agree): When I make a purchase in category c… - the brand is important because I believe other people judge me on the basis of it - I purchase particular brands because I know that other people notice them - I purchase particular brands because I have much in common with other buyers of that brand - I pay attention to the brand because its buyers are just like me (Cronbach's alpha: .88).	Fischer, Völckner, and Sattler (2010)
Price positioning (pp_j)	Kantar Worldpanel	Average price index of brand j across 2011-2015, where the index is calculated as the price per volume unit of brand j in week t, relative to the average price per volume unit of the category to which brand j belongs in the base week (i.e., week 1 of 2011).	Van Heerde, Gijsbrechts, and Pauwels (2008)
Promo intensity (pi_j) [a]	Kantar Worldpanel	Average % (across retailers and weeks) of brand j's SKUs on promotion at a top 3 retailer in a given week, with retailer weights equal to their market share.	Srinivasan et al. (2004)

Advertising intensity (ai_j)	Kantar Media	Brand j's average weekly advertising spending (in ¥) across 2011-2015 (converted into real spending using China's consumer price index, source: National Bureau of Statistics China).	
Brand ownership (fo_j)	Brand's websites	Coded as 1=foreign (i.e., brand owner is not Chinese), 0=domestic (i.e., brand owner is Chinese).	

[a] In our purchase data, no promotional information is present, therefore we work with a proxy measure.

Table 2.3: Data descriptives

PANEL A: SUMMARY OF MARKET SHARE AND MARKETING MIX ACROSS BRANDS (N=377)					
VARIABLE	STATISTIC[a]	MEAN	SD	LOWER QUARTILE	UPPER QUARTILE
Market share	Average	8.62%	10.57%	2.26%	11.08%
	Coefficient of variation	.32	.21	.18	.41
Price index[b]	Average	1.04	.45	.78	1.18
	Coefficient of variation	.11	.09	.05	.13
Advertising (Share of voice of Adstock)	Average	.02%	.08%	.08%	.01%
	Coefficient of variation[c]	1.89	1.52	.85	2.50
Distribution	Average	.72	.18	.65	.86
	Coefficient of variation	.09	.11	.03	.10
Line length	Average	4.69%	6.08%	1.68%	5.52%
	Coefficient of variation	.18	.10	.11	.23

PANEL B: CORRELATION TABLE AND SUMMARY STATISTICS DRIVERS IN SECOND-STAGE ANALYSIS									
VARIABLE	MEAN (STANDARD DEVIATION)	CORRELATIONS (NUMBER OF OBSERVATIONS)							
		$co_{c(j)}$	$pe_{c(j)}$	$le_{c(j)}$	$sd_{c(j)}$	pp_j	pi_j	ai_j	fo_j
Concentration ($co_{c(j)}$)	48.56% (17.47%)	1.00 (377)							
Perishable ($pe_{c(j)}$)	15 perishable categories/95 brands	.09 (377)	1.00 (377)						
Local embeddedness ($le_{c(j)}$)	3.95 (1.29)	-.29 (377)	.18 (377)	1.00 (377)					
Social demonstrance $sd_{c(j)}$)	4.92 (.28)	.03 (319)	.10 (319)	-.41 (319)	1.00 (319)				
Price positioning (pp_j)	1.04 (.45)	.01 (377)	.01 (377)	.07 (377)	-.02 (319)	1.00 (377)			
Promo intensity (pi_j)	7.09% (2.11%)	-.07 (377)	.19 (377)	.10 (377)	-.09 (319)	.13 (377)	1.00 (377)		
Log advertising intensity (ai_j)[d]	.87 (18.37)	-.11 (377)	.12 (377)	.06 (377)	.06 (319)	.13 (377)	.33 (377)	1.00 (377)	
Brand ownership (fo_j)	143 foreign brands	.04 (377)	-.04 (377)	-.26[e] (377)	.17 (319)	.22 (377)	.18 (377)	.20 (377)	1.00 (377)

[a] For market share, ‘average’ is the average, across all brands, of the brand’s mean market share over time (i.e., across 251 weeks); ‘coefficient of variation’ is the average, across all brands, of the brand’s [market share standard deviation over time] divided by its [mean market share over time]. The statistics for price, advertising, distribution and line length are defined in a similar way.

[b] Because prices are expressed per volume unit, and volume units differ across categories (e.g., milliliters for shampoo, grams for potato crisps), we display summary statistics of the price index to ensure comparability across brands in different categories (see also Van Heerde, Gijsbrechts, and Pauwels 2008). To obtain the price index, we divide weekly brand prices by the average category price in a base week (i.e., week 1 in 2011). A price index above (below) one thus indicates that the brand is more (less) expensive than the category average in the base week.

[c] The coefficient of variation is only calculated for the 255 out of 377 brands in our sample that advertised across 2011-2015.

[d] Log advertising intensity represents the log-transform of average weekly spending (in ¥) on all media across 2011-2015 (the log of 1.00E-11 is taken for the 122 out of 377 brands in our sample with zero ad spending across 2011-2015. Average weekly ad spending of the 255 out of 377 brands that did advertise across 2011-2015 is ¥5,666,511.59 with a standard deviation of ¥14,124,482.58).

[e] Note that local embeddedness and brand ownership do not strongly overlap in our data.

While our focus is on price, our extensive dataset and marketing mix coverage allows for additional findings that are of managerial importance in their own right. Distribution emerges as the most important marketing instrument (in terms of elasticity), with an average elasticity of .84. Line length matters too, the elasticity being .49. Advertising on the other hand has a negligible impact on market share, echoing results for Western CPG markets (Van Heerde et al. 2013). Increasing relative ad spending enhances market share only for 11% of the brands.

We observe inertia in brand shares and prices – indicating that consumers' brand preferences tend to be persistent, and that pricing history is an important driver of current brand prices. Finally, we find that many brands exhibit a significant (deterministic) trend in market share (39% of the brands) and price (36%), underscoring the dynamics in the market. The trend averages across brands are very small (while their standard deviations are not), indicating that some of the brands exhibit market share (price) increases, while others show decreases.

Moderators of price elasticity

As noted above, there is large heterogeneity in price elasticities. We now turn to examining the effect of the moderators. Table 2.5, Panel A shows the results of our second-stage analysis including all moderators. Market concentration is a major moderator, like in Western markets. More concentrated markets (two standard deviations (SDs) above the mean) are considerably more price elastic than fragmented markets (two SDs below the mean), the difference being .56.[8] Compared to brands in non-perishable categories, perishable brands are marginally less price elastic (Δ = .09).

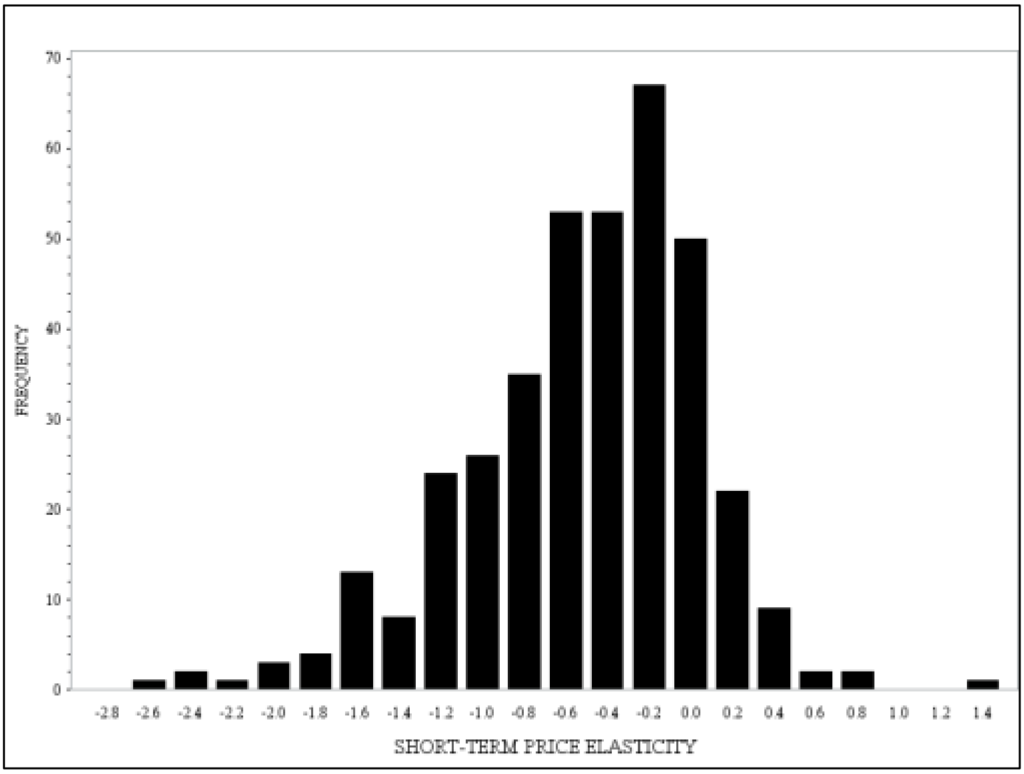
[8] Unless noted otherwise, high versus low refers to two standard deviations above versus below the mean.

Table 2.4: Summary of estimation results (elasticities)[a]

VARIABLE	MEAN ELASTICITY[b]	SD[a]	NUMBER OF BRANDS WITH *P* < .05
MARKET SHARE EQUATION			
Price	-.51***	.55	53.72%
Advertising	-.002	.05	11.37%
Distribution	.84***	4.77	11.44%
Line length	.49***	1.99	9.57%
Market share inertia	.24***	.22	60.37%
Trend	-.01**	.13	38.56%[c]
PRICE EQUATION			
Market share	.003	.05	6.91%
Price inertia	.33***	.19	81.65%
Trend	-.0006***	.02	36.44%[c]

[a] To ensure comparability across brands and specifications (variables in levels or differences), the table reports the elasticities instead of the 'raw' coefficients. The marketing mix elasticities (i.e., for advertising, distribution, line length and price) in the market share equation are the % change in market share from a 1% change in the marketing mix instrument in the same week, calculated at the average level of the brand's market share in the observation period. The trend elasticity represents the % change in market share and price from moving up one week in time. The market share (price) inertia elasticity is the % change in current market share (price) from a one percent increase in market share (price) one week ago. The market share elasticity in the price equation is the % change in current price from a 1% change in market share one week ago.
[b] Means and standard deviations across 376 brands in 50 categories (255 brands in 43 categories for advertising). Significance of the mean based on a meta-analysis of the (one-sided) *p*-values of the individual brand elasticities, using the method of adding Zs (Rosenthal 1991) across brands. ****p*<.01, ***p*<.05, **p*<.1.
[c] Two-sided p-value.

Figure 2.2: Histogram of price elasticities



Turning to the two category characteristics that have not been studied extensively in Western research, we find that the higher the social demonstrance of a product category, the less consumers respond to the price weapon. A product category with high social demonstrance has a predicted price elasticity that is .46 smaller in magnitude than a product category low on social demonstrance. In high social-demonstrance categories, the predicted price elasticity is a modest -.29, versus -.75 in low social-demonstrance categories. This suggests that for Chinese consumers, when the social aspect comes into play, the loss-of-face from consuming cheap brands partly overshadows the financial consequences. Brands in categories that are deeply embedded in Chinese society have on average a predicted price elasticity of -.63, versus -.41 in 'new' categories.

Turning to the brand factors, highly promoted brands have a price sensitivity that is .23 larger in magnitude than brands that are hardly promoted. Advertising has a dampening effect on price elasticity: highly advertised brands have a price elasticity that is .18 smaller in magnitude than brands that receive no advertising support. We find no evidence for the moderating role of price positioning. Finally, brand ownership matters. Foreign brands are more price elastic than domestic brands: -.65 versus -.44.

Table 2.5: Results of moderator analysis

VARIABLE		ESTIMATE	P-VALUE (ONE-SIDED)	EXPECTATION CONFIRMED?
PANEL A: MAIN ANALYSIS (N=318; R^2=.19)				
Intercept		-1.76	.002	
Category	Concentration	-.80	<.0001	Yes
	Perishable (1=Perishable)	.09	.10	Marginally
	Local embeddedness	-.04	.04	Yes
	Social demonstrance	.41	<.0001	Yes
Brand	Price positioning	.05	.80	No
	Promo intensity	-2.87	.02	Yes
	Log Advertising intensity	.005	.002	Yes
	Ownership (1=Foreign)	-.21	.0004[a]	n.a.[b]

PANEL B: ADDITIONAL ANALYSIS, EXCLUDING SOCIAL DEMONSTRANCE (N=376; R^2=.13)				
Intercept		.43	.003	
Category	Concentration	-.75	<.0001	Yes
	Perishable (1=Perishable)	.18	.001	Yes
	Local embeddedness	-.10	<.0001	Yes
Brand	Price positioning	.07	.91	No
	Promo intensity	-3.28	.005	Yes
	Log Advertising intensity	.004	.007	Yes
	Ownership (1=Foreign)	-.19	.0006[a]	n.a.[b]

[a] Two-sided *p*-value.
[b] n.a. = not applicable (no prior expectation formulated).

To check the stability of these findings, and because social demonstrance is measured for only 46 (out of the 50) categories, we re-run the second stage analysis on the full set of categories and brands (i.e., 376 instead of 318 brands), after dropping social demonstrance[9]. As Table 2.5, Panel B shows, the results are replicated in direction. However, the magnitude of the effect of local embeddedness (high vs. low) on price elasticity increases substantially, from -.11 to -.25. This is because of the negative correlation between social demonstrance and local embeddedness of -.41 (Table 2.3B). Categories that are newer to China tend to have higher social demonstrance. By eliminating social demonstrance from the model, this aspect of a category is picked up by local embeddedness.

Discussion

EMs, and China in particular, constitute an ever more important source of business for many companies. With a slowdown in growth, and the number of players increasing, competition in China has intensified – leading to a stronger focus on pricing decisions. Yet, empirical generalizations on price elasticity and its moderators are based on developed markets, leaving it unclear whether these Western findings apply to China too. Perhaps they do – which is important to know. Perhaps there are differences, which is also important to

[9] We also re-ran the second stage analysis with the medium term price elasticity as dependent variable, and found the pattern of results to be similar.

know. The goal of this study is to provide an initial set of empirical generalizations on brand price elasticities for China, the world's largest EM by far. To allow for more precise results, we use a unified data, modeling, and estimation framework. Below, we discuss our findings around three themes that guided our research: average price elasticity, brand- and category moderators of price elasticities, and the importance of price relative to other marking mix instruments.

Average brand price elasticity in China

On average, the price elasticity in China is -.51, implying that a 1% price increase leads to a drop in market share of half a percent. Thus, CPG markets in China are generally price inelastic. How does this finding compare to Western markets? Is China more or less elastic than the U.S.? For this, we turn to the meta-analysis of Bijmolt et al. (2005). These authors provide a detailed overview of the estimates of the effects of market and methodology characteristics on price elasticity (Table 2 of their paper). We use their results to arrive at an average predicted U.S. price elasticity for a modeling context that resembles our context as closely as possible. This yields a price elasticity of -.90.[10] So, after controlling for study characteristics, we find no evidence that that price sensitivity in China is higher than in the U.S. However, economic theory suggests that lower income is associated with higher price elasticity. Clearly, that is not the case here. There appears to be a countervailing force operating. We propose that countervailing force is in differences between the U.S. and China

[10] More specifically, China's CPG markets are generally in the introduction or growth stage (effect = 0 in Bijmolt et al., Table 2), we use household panel data (+.22), temporal aggregation is weekly (+.51), estimates are at the brand level (+.47), our criterion variable is market share (0), we use an attraction model (-.21), duration of the effect is short term (0), we use the actual price (0), we account for price endogeneity (-1.27), we include distribution (+.68) and advertising (+.84), and we use SUR as estimation method (+.26). Finally, if we take perishability as proxy for stockpilability, 25% of our brands are in the category 'grocery, low stockpiling' (0) and 75% in 'groceries, high stockpiling' (+1.39), leading to an aggregate effect of +1.04 for CPG. The intercept is -3.79. Adding all effects yields a predicted elasticity in North America of -1.25. Moreover, Bijmolt et al. (2005) found a slight effect for mean-centered time trend (+.01 per year). The mean year in their series was 1978 (Harald van Heerde, personal communication), which means that the price elasticity in 2013 (mid-point in our time series) is .35 smaller in magnitude (i.e., +.35). We thus arrive at a final estimate of the average U.S. price elasticity of about -.90.

on susceptibility to normative influences and social demonstrance. Recall that susceptibility to normative influences refers to the need to enhance one's image in the opinion of significant others through the acquisition and use of products and brands (Bearden, Netemeyer, and Teel 1989). This need can best be fulfilled with brands in categories that are high in social demonstrance (Fischer, Völckner, and Sattler 2010). After all, if the brand I buy in category X says something about the kind of person I am, and if others judge me on the basis of the brand I buy, and I have a need to enhance my image in the eyes of others, this should reduce price sensitivity and foster a brand focus. Now if, on average, 1) Chinese consumers are much more susceptible to normative influences than Americans and 2) CPG are much higher on social demonstrance in China than the U.S., this could provide an explanation for our finding that the average price sensitivity in CPG is not more negative in China than in the U.S. despite significant differences in disposable income.

De Jong et al. (2007) report country averages for susceptibility to normative influences for multiple countries, including the U.S. and China. Further, as part of a global study, Kantar Worldpanel and GfK administered the four social demonstrance items as part of a larger survey in the U.S., China, and the other three BRIC nations – Brazil, Russia, and India. Sample size was around N=1,600 in each country (except for China, where the sample size was larger, as mentioned earlier). Table 2.6 reports country means, based on the partial scalar invariance model (Steenkamp and Baumgartner 1998). The mean for the reference country (U.S.) is fixed to zero. We see that, indeed, China is much higher than the U.S. on susceptibility to normative influences, and attributes a much higher social signaling function to CPG.

Table 2.6: Country comparisons: U.S. and BRIC countries

COUNTRY	SNI[a]	SOCIAL DEMONSTRANCE	MONTHLY DISPOSABLE INCOME
U.S.	0	0 (reference)	$3259
China	2.130	1.040	$731
India	n.a.[b]	.874	$452
Brazil	.798	.459	$757
Russia	1.204	.090	$686

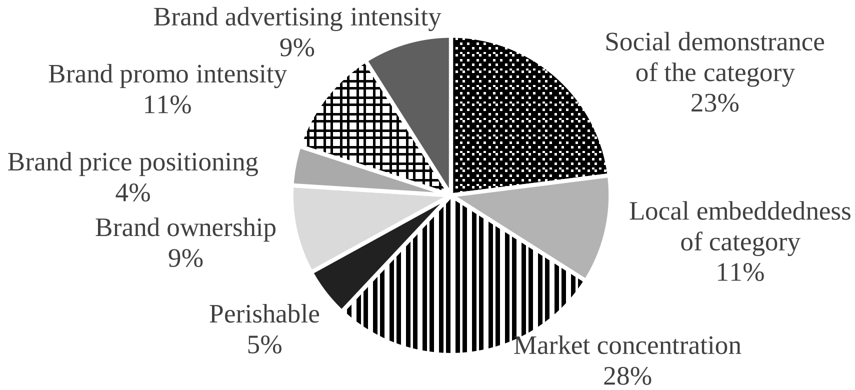
[a] Note: Susceptibility to normative influence (SNI) taken from De Jong et al. (2007), social demonstrance calculated by authors (N=10,289), and monthly disposable income (2014) taken from www.nationmaster.com/country-info/stats/Cost-of-living/Average-monthly-disposable-salary/After-tax.
[b] n.a. = not available.

Comparing the elasticities of our control variables to previous large scale studies based on Western CPG data (that used similar variable operationalizations as we did), reveals that the low advertising elasticity is in line with the results of Ataman, Mela, and Heerde (2008), Ataman, Van Heerde, and Mela (2010), and Van Heerde et al. (2013) – where the latter study provides the most fair comparison as that study is also based on aggregated household panel data, whereas the former two are based on aggregated store panel data. While the distribution and line length elasticities obtained by Ataman and colleagues are much smaller in magnitude (i.e., .76 and .15 in Ataman, Mela, and Heerde (2008) for new brands and .13 and .08 in Ataman, Van Heerde, and Mela (2010) for existing brands), the order of importance is the same as what we find: distribution ranks highest, followed by line length and advertising.

Moderators of brand price elasticity in China

Beneath this overall picture, however, we uncover important differences in price elasticities across product categories and brands. Figure 2.3 presents a pie chart of the relative effect of the moderators, where the effect is defined as the difference in price elasticity between ± 2 standard deviations on the moderator (except for brand ownership and perishable vs. nonperishable, which are dummies, and advertising where no advertising is the low option).

Figure 2.3: Relative effects of brand and category moderators of brand price elasticities in China

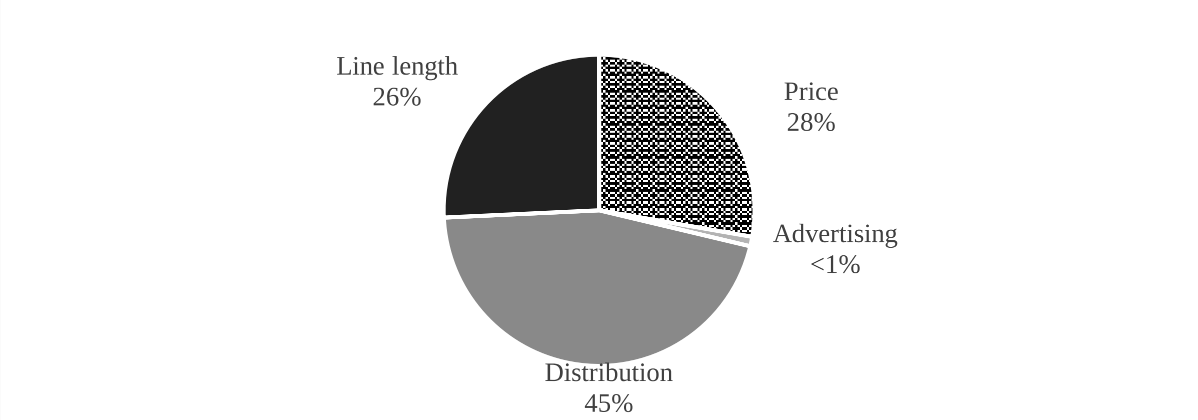


We find that category factors account for two-thirds of the total effect of the moderators. Clearly the category in which a brand competes has an important effect on its price elasticity – in fact considerably more than brand-specific actions. Figure 2.3 further reveals that the three moderators that have not been studied much in previous price elasticity research – possibly because they are not deemed relevant, at least in DMs – have a strong combined relative effect on price elasticity of 43%. The predicted price elasticity of a foreign brand in a category that has been around in China for a long time and is of low social demonstrance, is .90 larger in magnitude than the predicted price elasticity of a domestic brand in a ‘new’ category of high social demonstrance.

Relative elasticities across the marking mix

What about the elasticities of other instruments in China? Distribution matters the most. An increase in brick-and-mortar distribution of 1% increases brand share by .84%. Expanding the brand assortment with SKUs is another powerful instrument, the elasticity being .49. On the other hand, advertising’s effect is on average non-significant as well as negligible. However, as we have seen, advertising has an appreciable moderating effect on brand price elasticities. Figure 2.4 presents a pie chart of the relative marketing mix elasticities. As we can see, though price has an important part (28%), it is not the dominant instrument.

Figure 2.4: Relative elasticities across the mix



Managerial implications

Our findings offer insights for CPG managers operating in China. Some may believe that success in China is first and foremost about price – not an unreasonable assumption given that Chinese average monthly disposable income per capita in 2014 was only $731 vs. $3,258 for the U.S. While price obviously matters, assortment decisions are about equally important and distribution matters substantially more. The strong effect of 'old-fashioned' brick-and-mortar distribution is noteworthy as nowadays, much of managerial attention is directed to the potential of the online channel. Indeed, the penetration of the online channel for CPG is higher than anywhere in the world and online now accounts for 7% of all CPG sales in China (Bain & Company and Kantar Worldpanel 2017). While we do not argue to ignore online, we caution not to neglect investing in the offline channel, which remains very important for brand building in China.

Not surprisingly, we find large heterogeneity in price elasticities in China. Perhaps more surprising is the relative weight of social demonstrance, local embeddedness, and foreign vs. domestic brand ownership – three factors that CPG managers in the U.S. or

Europe might not readily consider as particularly relevant. The most important of these factors is social demonstrance. Price elasticity is considerably lower in categories that have a high social demonstrance function. Market leaders in categories that are low on social demonstrance could attempt to increase the symbolic value of the category. This reduces (category) price sensitivity, which is attractive for brands in a leading position. Luxurious packaging and advertising that emphasizes social consumption or sharing might be ways to change consumer perceptions. For example, toothpaste rates low on social demonstrance in our survey. Market leaders like Crest, Colgate, and Darlie could communicate that their buyers are just like the target audience and develop and advertise unique flavors that signal the use of a prestigious brand. The adverse social effects of not using the ‘right’ brand are easy to convey in advertising and, as our results show, move consumers’ focus away from price.

The honeymoon for foreign brands in China appears to be over. In general, one would expect that strong brands have a smaller absolute price elasticity than weaker brands. Historically, strong brands in China used to be foreign brands (Steenkamp 2014). But nowadays, in China, foreign brands are more price elastic. This could be due to home-country bias of Chinese consumers (Shimp and Sharma 1987), the appeal of local identity (Gao, Zhang, and Mittal 2017), and/or lower familiarity with foreign brands (Erdem, Swait, and Louviere 2002). Industry evidence has documented that foreign brands are struggling. According to Bain & Company, local companies grew by 8.4%, while foreign brands grew only by 1.5%. Bain offered several reasons including local players’ entrepreneurial governance, their knowledge of local taste, and their ability to make quick decisions and just as quickly execute those decisions – including those that help them innovate or embrace digital opportunities (Bain & Company and Kantar Worldpanel 2017). For foreign brands to get back in the game, it appears crucial to push decision-making authority from corporate headquarters to local managers in China.

We document how managers can further reduce Chinese consumers' price focus for their brands, using principles established in developed markets. While we obtain no direct positive effect of advertising on market share, our findings show that also in EMs, ad spending, and the familiarity that comes with it, reduces consumers' brand-price sensitivity. Similarly, like in Western markets, brand managers should be wary of over-using promotions – higher deal-intensity increasing brand price sensitivity.

Finally, what can we say about price elasticity in other BRIC nations? Is its magnitude likely to be higher or lower than in China? We cannot give a precise answer, but in the spirit of (Raju 2005, p.18) who emphasized the value of approximate answers to important issues, we can provide an approximate direction by taking into account the country's susceptibility to normative influences, social demonstrance of CPG, and disposable income per capita. The disposable income per capita of Indian consumers is significantly below China's while social demonstrance is also slightly lower (Table 2.6). This suggests, as a benchmark for managers, that the price elasticity is substantially larger in magnitude in the world's second largest EM. The average monthly disposable income in Russia ($686) and Brazil ($757) are not very different from China's ($731) but both countries score much lower on susceptibility to normative influences, and the social demonstrance function of CPGs is also low. Thus, we tentatively predict that CPG markets in both countries are more price elastic than China, with the difference being especially large for Russia.

Limitations and future research

Our study is not without limitations, which offer opportunities for future research. First, our empirical analysis only pertains to China. Any generalization to other emerging markets is subject to further research. Though we used Chinese findings to suggest approximate answers for the other three BRIC nations, these findings should be verified with primary research in these countries. Second, our empirical context is the CPG industry. We are not unique in this respect (cf. Bijmolt, Van Heerde, and Pieters 2005). Available evidence reported by Bijmolt and colleagues suggests that demand for durables may be more price elastic. It remains to be tested whether this is also the case for China, or other EMs. Given the importance of social demonstrance, at least in EMs, it would be beneficial to include this measure in this research. Third, though we are among the first to analyze actual purchase behavior of Chinese consumers, using household panel data has limitations too. Estimating a multiplicative model to aggregated household panel data may lead to aggregation bias and possibly a lower price elasticity (see Christen et al. 1997). In addition, the panel covers urban households (which represent the bulk of the Chinese market potential: only 25% of China's GDP comes from rural households, China Daily 2015). One could argue that these households will be more similar to Western consumers, which could partly explain why the average price elasticity we obtain is very comparable to what one would expect when using similar data and methodology for a Western country (Bijmolt, Van Heerde, and Pieters 2005). Future research based on store panel data should verify the robustness of our results with other data types.

Fourth, we focused on market share as the key performance metric. Future studies could consider brand-sales and category-expansion effects of price changes. Fifth, though we documented the effects of several category- and brand-drivers of price elasticity in CPG that are more or less idiosyncratic to China – and perhaps to other EMs – other factors remain to be explored. For example, while private labels currently hardly play a role in EMs and it may take years for these products to get a foothold in these markets, the extent to/speed with which private labels will develop may be related to differences in price elasticities. Finally, like previous large-scale studies, we documented market-level price response, which made it feasible to cover a large number of brands and categories. Given that EMs are often heterogeneous, it may be useful to study price reactions at the household level – something we leave for future analysis.

Chapter 3 | Consumer Learning About Quality of Global and Local Brands in the CPG Industry in China

Introduction

It is generally accepted that consumers are imperfectly informed and thus uncertain about brand quality. Research has shown that this uncertainty persists even after consumption of the brand, because use experience may provide only noisy information (Erdem, Zhao, and Valenzuela 2004). Even in mature categories like consumer packaged goods (CPG), consumers are not perfectly knowledgeable about the quality of brands because the quality effects take time or multiple consumption experiences to materialize (Erdem 1998) and it may be difficult to isolate the quality of the brand from other confounding factors (Hoch and Deighton 1989). Further, consumers generally buy and consume products sequentially rather than simultaneously, which hampers effective brand comparisons and learning (Warlop, Ratneshwar, and van Osselaer 2005). Finally, consumers need to update their knowledge as their consumption patterns evolve (Du and Kamakura 2006).

Brand-quality learning in CPG has been studied extensively in Western markets and is by now quite well understood (Ching, Erdem, and Keane 2013; Erdem and Keane 1996; Szymanowski and Gijsbrechts 2013). This is not surprising as brands have been widely available in developed markets (DMs) for as long as one can remember – for example, Coca Cola, Kellogg, Gillette, Hershey, Colgate, Wrigley, and Campbell were already the leading brand in their category in 1925 (Aaker 1991). The market situation is very different in emerging markets (EMs). In these countries, many consumers encountered their first brand of breakfast cereal or toothpaste after they had already come of age. For example, such global stalwart brands as Lay's, Crest, and Pampers were introduced in China only in the late 1990s. Consequently, as a field, we know relatively little about brands, the role of brands, and brand learning processes in EMs.

It has been suggested that today, brands still play a larger role in consumer behavior in EMs (Dawar and Chattopadhyay 2002). While in most CPG categories the majority of brands is present for multiple years in EMs, we know from research on Western markets that even for existing brands consumers may be uncertain (and learn) about the quality of brands (Erdem and Keane 1996; Erdem and Sun 2002). In EMs, people may rely more on brands to reduce the risk of making the wrong choice because the institutional infrastructure provides less protection and fewer opportunities to get legal redress (Burgess and Steenkamp 2006; Narasimhan, Srinivasan, and Sudhir 2015). EM consumers might also be more prone to use CPG brands as symbolic devices to project their self-image (Guimaraes and Chandon 2007).

As motivation for our study, we tested whether brands are indeed (even) more relevant in CPG industry in EMs than in the West. In collaboration with the global market research organizations Europanel and GfK, we administered Fischer, Völckner, and Sattler's (2010) 12-item brand relevance in category (BRiC) scale to approximately 8,000 consumers in Brazil, India, China, and the U.S. The BRiC instrument consists of three 4-item subscales measuring the brand functions of risk reduction and social demonstrance, and overall brand relevance in category. Each respondent scored the 12 BRiC items for one CPG category using a seven-point Likert scale. In total 30-50 categories were included. We analyzed the data with multigroup factor analysis using Mplus. Table 3.1 provides the latent means for the partial scalar invariance model (Steenkamp and Baumgartner 1998). The findings support the idea that CPG brands are more relevant, and perform a more pertinent role in risk reduction and in projecting one's self-image in EMs than in DMs like the U.S.

Table 3.1: Latent construct means of BRiC and brand functions in the U.S., Brazil, India, and China[a]

COUNTRY	BRiC	RISK REDUCTION	SOCIAL DEMONSTRANCE	SAMPLE SIZE
U.S.	4.25	4.62	3.40	1,557
Brazil	4.85*	5.41*	4.20*	1,695
China	5.36*	5.30*	5.14*	2,994
India	5.34*	5.32*	4.88*	1,503

[a] Model fit of the partial scalar invariance model: χ^2 (253) = 1452.15, RMSEA = .05, CFI = .97, TLI = .97, SRMR = .05.
* = significantly different from the U.S.

This suggests that investigating and quantifying brand learning in EMs is managerially very relevant. With Western markets being largely saturated and with stagnant populations, CPG multinationals from Colgate-Palmolive and Procter & Gamble to Unilever and Nestlé are ever more dependent on success in emerging markets to grow. EMs also provide an exciting opportunity from an academic point of view. As stated by Narasimhan, Srinivasan, and Sudhir (2015, p.473), "…research on emerging markets can contribute to richer theoretical and substantive understanding of markets and marketing." The heterogeneity in market environments allows academics to test not only the effects of socio-demographics, but also of regions and economic subunits such as city tiers (BCG 2008), which previously have not received much academic attention because they (are perceived to) matter less in developed markets. The dynamic nature of EMs presents another academic opportunity. In the early period, global brands were at an advantage because they were generally of higher quality, promoted with more sophisticated marketing, and – in those cases where the consumer knew the brand was global – benefited from the prestige and esteem associated with global brands (Batra et al. 2000). More recently though, local brands appear to have made a comeback. They have improved product performance and are sometimes seen as being more aligned with the needs and aspirations of local consumers (Steenkamp and de Jong 2010). Whether, and if so how, Chinese consumers nowadays learn differently from consuming global vs. local brands remains unclear.

These observations provided the impetus for the present study. Its purpose is to empirically study brand-quality perceptions, and the impact of quality uncertainty and learning, in the largest emerging market, China – a country that has undergone a dramatic evolution in the last three decades. Speed and change define China – in 1980, China's GDP was $306 billion; in 2015, it exceeded $11 trillion. No country in world history has experienced such a dramatic shift in its economic fortune in such a short time span. Serge Dumont, vice chairman at the advertising company Omnicom Group, described China in 1985: "There were no billboards. For foreign brands, the country was a blank slate. People in those days didn't eat chocolate, they didn't know what a contact lens was. So it was not just trying to convince them to buy this brand versus another, you had to educate about what the product was." (Doland 2015). The situation in 2015 is very different: Chinese consumers no longer want global brands just because they are foreign. For example, Revlon withdrew because it failed to connect with local consumers (Chan 2014). The competition between global and local brands is tough because China's consumers have gotten savvier (Doland 2015). Therefore, we will give special attention in this paper to global vs. local brands.

In this paper, we study the effects of brand quality and consumer uncertainty about brand quality, for global vs. local brands, on their brand choice behavior. We assess the importance of these factors relative to traditional marketing mix instruments – product line length, price, advertising, promotion, and distribution – for different geographic and socio-demographic consumer profiles. We estimate a brand-choice model with Bayesian learning, on purchases obtained in a scanner panel of 40,000 Chinese urban households over a four-year period: 2011-2014, operated by Kantar Worldpanel. We test our model on five CPG categories that cover foods (breakfast cereals), snacks (potato chips), hair care (shampoo), skin care (body creams & skin care), and fabric care (laundry detergent).

Specifically, our research addresses the following research questions. First, how do consumers in China perceive and learn about the quality of brands over time? How does this differ between global and local brands? Second, how do the effects of quality uncertainty on brand choice compare to those of (other) marketing mix instruments? Which are the key drivers of brand success in China – in particular: what is the role of consumers' attitude toward risk? Third, is there systematic heterogeneity in quality beliefs about global vs. local brands, and sensitivity towards quality uncertainty and (other) marketing mix instruments, in function of the geographic and sociodemographic makeup of the consumer (city tier, region, income, and age)? Can we begin to derive some generalizable insights on the makeup of different target groups that marketing managers and academics need to take into account when studying EMs?

Methodology

Our methodology is rooted in the literature that treats consumers as Bayesian learners (e.g., Erdem and Keane 1996; Erdem, Swait, and Valenzuela 2006; Mehta, Rajiv, and Srinivasan 2003; Narayanan and Manchanda 2009; Shin, Misra, and Horsky 2012; Szymanowski and Gijsbrechts 2012). Consumers have imperfect knowledge about the quality of (local and global) brands in a given category and, hence, make choices based on their quality beliefs. These beliefs are not static but dynamic: consumers update their prior beliefs to posterior beliefs via signals they obtain through consumption. Before we move to the formal model, we will briefly clarify how we define quality (uncertainty) and global vs. local brands.

First, following extant studies, we conceptualize quality as "a summary statistic that captures any intangible and tangible attributes of a product that may be imperfectly observable by consumers" (Erdem, Zhao, and Valenzuela 2004, p.87). The better the brand is able to match the needs of a consumer, the higher its quality for that consumer (Zeithaml

1988). Brand quality is thus subjective and consumer-specific. Consumers may differ in their brand perceptions and relative importance of objectively measurable attributes (e.g., cool water cleanability vs. dust sebum for laundry detergents) and in their brand perceptions and relative importance of intangible attributes (e.g., preference for prestigious brands). For intangible attributes objective levels may not even exist. It follows that to the researcher, who does not readily observe the consumer's preferences, quality is a latent construct. Moreover, because of its experiential nature, consumers have only imperfect information about the quality of a brand. When making a choice, they will rely on brand-quality beliefs, which they gradually update and become less uncertain about through actual consumption. These consumer-specific brand qualities and uncertain quality beliefs will be central to our analysis of brand learning and choice in an EM.

Second, the quality that consumers attach to a brand and the way they update their quality beliefs, may differ between global and local brands. We define brand type (global vs. local) based on actual availability. Global brands are sold in multiple regions of the world while local brands are generally available in only one country (Steenkamp 2014). Global brands are, by definition, produced by global manufacturers, but local brands can be produced by local or global firms. Procter & Gamble, for example, sells in China the global shampoo brands Pantene, Head & Shoulders, and Vidal Sassoon as well as the local brand Rejoice. In the same category, Unilever offers the global brands Clear, Lux, and Dove as well as the local brand Hazeline. Consumer evaluation and uncertainty may differ between global and local brands, without consumers necessarily being aware of the brand's global or local status (Steenkamp 2014). Global brands will often – albeit not always – be higher in objective quality on key attributes because R&D can be leveraged globally with the best minds being put to work to develop a product of superior functional performance (Yip and Hult 2012). Global brands can also benefit from leveraging the best marketing ideas from around the world (Kotabe and Helsen 2010). Some consumers may even be aware that the brand is global, which in and of itself is associated with higher perceived quality (Holt, Quelch, and Taylor 2004; Steenkamp, Batra, and Alden 2003). At the same time, in their efforts to create a broadly consistent image around the world, some global brands may not be as closely aligned to the specific bundle of tangible and intangible attributes sought by consumers in different locales. Local brands are typically better able to satisfy unique local needs, both in terms of attributes offered and the way these are communicated (Kotabe and Helsen 2010). Thus, for some consumers, a global brand's attributes (especially superior performance on key tangible attributes) may better fit their preferences, while for other consumers, local brands are a better match to their particular needs. Our model below will accommodate this.

Model Specification

Consider a market with consumers i = 1, 2, …, I, who have the option to choose from a set of brands j = 1, 2, …, J across a number of subsequent acquisition occasions in a product category. As further explained in the 'Data' section, consumers in our setting can either purchase the brand or receive it as gifts. Let t = 1, 2, …, T_i denote subsequent acquisition (i.e., purchase or gift) occasions for consumer i in the considered category. Brands are split into two types: local brands (which form the subset J^{lb}) vs. global brands (subset J^{gb}).

Let q_{ijc} be the quality of brand j in category c for consumer i. Because quality may not only depend on the brand's global or local nature but also on other aspects idiosyncratic to the brand, we let it be brand (and not just brand-type) specific. As indicated above, given that a brand's characteristics may better match the needs of some consumers than others, we further allow this brand quality to differ between consumers. Moreover, in line with extant learning models, we assume that consumers' knowledge is imperfect, and that they are uncertain about the quality of both local and global brands. Consumers' brand choices will thus depend on their quality beliefs at the time of purchase (Q_{icjt}), which we assume to be normally distributed with mean μ_{icjt} and variance σ^2_{icjt}.

Because their beliefs are uncertain, consumers choose the brand that maximizes the following utility expression:

$$(3.1) \quad U_{ijct} = \mu_{ijct} + r_{ic} * \left(\sigma^2_{ijct}\right) + \beta_{ic}P_{jct} + \gamma_{ic}D_{jct} + \delta_{ic}PM_{jct} + \zeta_{ic}PLL_{jct} + \theta_{ic}AD_{jct} + \varepsilon_{ijct},$$

$$= V_{ijct} + \varepsilon_{ijct},$$

where

μ_{ijct} is consumer i's mean belief about the quality of brand j in category c on purchase occasion t;

σ^2_{ijct} is consumer i's quality belief variance of brand j in category c on purchase occasion t;

P_{jct} is the relative price of brand j in category c on purchase occasion t;

D_{jct} is the distribution of brand j in category c on purchase occasion t;

PM_{jct} is the promotion of brand j in category c on purchase occasion t;

PLL_{jct} is the product line length of brand j in category c on purchase occasion t;

AD_{jct} is the relative advertising stock of brand j in category c on purchase occasion t;

r_{ic}, β_{ic}, γ_{ic}, δ_{ic}, ζ_{ic}, θ_{ic} are the risk aversion, price, distribution, promotion, product line length, and advertising parameters, respectively, for consumer i in category c; and

ε_{ijct} are random, i.i.d. (gumbel-distributed) utility components unobserved to the researcher, but observed by the consumer.

As expression (3.1) shows, we use a mean-variance framework to incorporate consumers' quality belief uncertainty in their utility expression. This framework allows to freely estimate consumers' attitude towards risk, while still being consistent with random utility maximization[11].

[11] Unlike the CARA (constant absolute risk aversion) specification, which is less flexible in that it excludes risk seeking behavior. As indicated by Meyer (1987), with normally-distributed quality beliefs, the mean-variance framework is consistent with RUM. Moreover, being separable in the mean and variance of the beliefs, it allows to directly assess the impact of bringing in uncertainty, over and above a model like that used by Shin, Misra, and Horsky (2012).

Consumer i's initial belief (i.e., at time 0) about the quality of brand j is given by:

(3.2) $Q_{ijc0} = N(\mu_{jc0}, \sigma^2_{jc0})$,

where μ_{jc0} is the initial mean quality belief for brand j in category c, and σ^2_{jc0} is the initial belief variance for brand j in category c (i.e., the initial uncertainty). Because estimating brand-specific initial variances is empirically challenging and may lead to unstable parameter estimates, we follow Erdem, Zhao, and Valenzuela (2004) and specify a separate initial variance for each brand type (local vs. global) and category, but restrict it to be the same for brands within a brand type and category: $\sigma^2_{jc0} = \sigma^2_{c0}|gb$ if j is a global brand, and $\sigma^2_{jc0} = \sigma^2_{c0}|lb$ if j is a local brand.

Every acquisition occasion (purchase or gift) provides the consumer with experiences (consumption signals) that allow them to learn about the quality of the brand. Like previous authors, we assume that on each acquisition occasion t, the consumer acquires only one brand. Let y_{ijct} (g_{ijct}) be an indicator variable that equals 1 if consumer i acquired (bought or received) brand j in category c on occasion t, and 0 otherwise.

Brands can be acquired in different package sizes, and large packs entail more consumption experiences than small packs. Following Szymanowski and Gijsbrechts (2013), we accommodate this by defining a 'consumption unit' in each category, (i.e., the most popular volume size in that category), and determining the number of consumption units M_{ict} corresponding to each acquisition by the consumer in that category. Each unit m then provides a (new) quality experience, which we refer to as a consumption signal s_{ijtm}. We assume that the consumption signals are i.i.d. normally distributed with mean equal to the brand quality q_{ijc} and variance $\sigma^2_{v_{jc}}$. A series of signals s_{ijtm} for the M_{ict} consumption units acquired at time t can be summarized as:

(3.3) $S_{ijct} = \frac{\sum_{m=1}^{M_{ict}} s_{ijtm}}{M_{ict}} \sim N(q_{ijc}, \frac{\sigma^2_{v_{jc}}}{M_{ict}})$.

Based on these consumption signals, consumers update prior beliefs at time t – 1 to posterior beliefs at time t, in a Bayesian fashion. Using the standard Bayesian updating equations (DeGroot 1970), the mean quality belief μ_{ijct} and the perceived quality uncertainty (belief variance) σ^2_{ijct} become:

$$(3.4) \quad \mu_{ijct} = \left(\frac{\mu_{ijct-1}}{\sigma^2_{ijct-1}} + \frac{(y_{ijct}+g_{ijct})s_{ijct}}{\frac{\sigma^2_{v_{jc}}}{M_{ict}}} \right) * \left(\frac{1}{\sigma^2_{ijct-1}} + \frac{(y_{ijct}+g_{ijct})}{\frac{\sigma^2_{v_{jc}}}{M_{ict}}} \right)^{-1}$$

and

$$(3.5) \quad \sigma^2_{ijct} = \left(\frac{1}{\sigma^2_{ijct-1}} + \frac{(y_{ijct}+g_{ijct})}{\frac{\sigma^2_{v_{jc}}}{M_{ict}}} \right)^{-1}$$

which feed into the expected-utility expression (3.2).

Assuming that the error terms ε in the utilities follow i.i.d. extreme value distributions, the probability of consumer i choosing brand j in category c at time t takes the form of a multinomial logit choice probability (McFadden 1974):

$$(3.6) \quad \Pr_{ijct} = \frac{e^{(V_{ijct})}}{\sum_{l=1}^{J} e^{(V_{ijct})}}.$$

Identification and Estimation

We estimate the model given by (3.1)-(3.6) separately for each product category. To ensure model identification, we set the quality for one brand (the market leader in the first year of the data) equal to zero. We assume that each consumer has 'rational expectations,' i.e., that his/her initial quality beliefs equal the mean quality of the brand across consumers in the same category $\bar{q}_{jc}$ (Crawford and Shum 2005; Narayanan and Manchanda 2009). In each product category, we fix (the log of) the initial uncertainty for local brands to one (similar to Szymanowski and Gijsbrechts 2012), and estimate the initial uncertainty of global brands as a separate parameter. To allow for differences in the brand's ability to satisfy the needs of individual consumers, we let brand quality be normally distributed across consumers, with mean $\bar{q}_{jc}$ and variance σ^2_{qjc}. To accommodate consumer heterogeneity in risk aversion and in the effect of marketing mix instruments on their choice behavior, we use a random-effects model with normal mixing distributions: $r_{ic} \sim N(\bar{r}_c, \sigma^2_{r_c})$; $\beta_{ic} \sim N(\bar{\beta}_c, \sigma^2_{\beta_c})$; $\gamma_{ic} \sim N(\bar{\gamma}_c, \sigma^2_{\gamma_c})$; $\delta_{ic} \sim N(\bar{\delta}_c, \sigma^2_{\delta_c})$; $\zeta_{ic} \sim N(\bar{\zeta}_c, \sigma^2_{\zeta_c})$; and $\theta_{ic} \sim N(\bar{\theta}_c, \sigma^2_{\theta_c})$.

To summarize, for each product category, we estimate the following parameters: (1) mean and variance of the quality distribution across consumers for each brand except the reference brand ($\bar{q}_{jc}$ and σ^2_{qjc} for j≠reference brand); (2) (log of the) consumption signal variance ($\log(\sigma^2_{v_{jc}})$); (3) (log of the) initial variance of global brands ($\log(\sigma^2_{c0|gb})$); and (4) means and variances of the mixing distribution for the risk and marketing mix parameters: $\bar{r}_c, \sigma^2_{r_c}$, $\bar{\beta}_c, \sigma^2_{\beta_c}$, $\bar{\gamma}_c, \sigma^2_{\gamma_c}$, $\bar{\delta}_c, \sigma^2_{\delta_c}$, $\bar{\zeta}_c, \sigma^2_{\zeta_c}$, and $\bar{\theta}_c, \sigma^2_{\theta_c}$. We estimate the model with simulated maximum likelihood, using 100 shuffled Halton draws from the household mixing distributions, combined with 100 consumption-signal draws on each category purchase or gift. In each product category, the simulated log-likelihood for a given category c thus becomes:

$$(3.7) \quad LL_c = \sum_{i=1}^{I} \ln\left[\sum_{f=1}^{100}\left(\prod_{t=T_{init,i}+1}^{T_i}\prod_{j=1}^{J}\left(Pr_{icjt}|\Omega^f_{ic}, E^f_{ict}\right)^{y_{icjt}}\right)/100\right],$$

where Ω^f_{ic}, E^f_{ict} represent arrays of random draws for the heterogeneous household parameters and consumption signals, respectively. In equation (3.7), two points are worth noting. First, though quality belief updating in equations (3.4) and (3.5) starts at the first observed consumption occasion for each household (t=0), in the calculation of the log-likelihood we only include trips after a household's initialization period ($t>T_{init,i}$) to deal with left truncation (i.e., the fact that we do not observe households' purchase histories from their first purchase in the category onward; Szymanowski and Gijsbrechts 2012). Second, whereas belief updating in (3.4) and (3.5) naturally includes all acquisition occasions (purchases *and* gifts), the probabilities in (3.7) pertain to purchase occasions only. As such, the fact that our data comprise purchases as well as gifts helps to identify the model, and separate the parameters that govern quality learning from those that drive choice.

Data

To estimate our model we have access to a unique dataset, provided by the global market research agency Kantar Worldpanel. The data pertain to a Chinese household panel (n=40,000) that tracks the panelists' purchases and gifts in CPG categories between 2011 and 2014. The panel covers urban (i.e., not rural) households: a very important part of the population of China, as urban households represent about 75% of China's GDP (China Daily 2015). For every shopping trip panelists undertake, we know which brands they bought, how many, and at what price and retailer. As mentioned earlier, an interesting phenomenon in this market is that of CPG 'gift giving:' consumers often receiving (small packages of) products for free (from their employer or friends/family) rather than buying them at the store. This is something that is uncommon in developed countries like the U.S. Across all occasions on which a panelist acquires a product in the category, 87.4% actually refer to a purchase by the household, and 12.6% are gifts (10.4% from family/friends, 2.2% from employer). We also know panelists' age (in years), household income (32 classes, where 1= ¥400 or less; 2=¥401–¥600; …; 31=¥50,001–¥60,000; 32=¥60,001 or more), the region where they live (i.e., East, South, West, or North), and their city tier. Kantar Worldpanel distinguishes between the 'high' city tier (covering 26 cities: municipality cities, provincial capital cities, and Shenzhen, Dalian, and Qingdao) and the 'low' city tier (comprising 228 prefecture-level cities, 322 county-level cities, and 1300 counties).[12]

[12] Fifty panelists (24 of which living in high-tier cities, 26 in low-tier cities) moved from one city tier to another during the data period. To ensure a clean effect of city tier when exploring consumer heterogeneity, they are removed from the estimation sample.

We conduct our analyses in five categories: breakfast cereals, potato chips, shampoo, body creams & skin care, and laundry detergent (washing powder). All five categories have both local and global brands, and they represent a diverse set in terms of use (i.e., food, personal care and household care), purchase frequency, category penetration, type of buyers, and market concentration. See Table 3.2 for an overview of the characteristics of the categories. We selected all brands in the category with >2% market share.

The household panel data provide us with information to construct measures on brand level: prices, distribution intensity, promotion activity, and product line length[13]. In addition, we bought advertising data from Kantar Media, comprising monthly gross total ad spending (per brand and per category) across all media types. Table 3.3 provides an overview of the operationalization of the marketing mix measures.

Descriptives

Before turning to the estimation results, we provide some model-free insights. To save space, we discuss one category (breakfast cereals) that is more or less representative for our set and indicate deviations in other categories only where needed. Descriptives for the four remaining categories are given in Appendices 3.A, 3.B, and 3.C.

Brand Share Evolution

Table 3.4 shows the performance of the top seven brands in the breakfast cereals category, for each year in our observation period. Considering the combined shares of local vs. global brands, we see that both brand types account for a substantial portion of category sales.

[13] In emerging markets like China, price could be an indicator of quality. We aim to obtain clean estimates of our quality (uncertainty) parameters by separately controlling for price as well as promotion.

Table 3.2: Characteristics analyzed product categories

	BREAKFAST CEREALS	POTATO CHIPS	SHAMPOO	BODY CREAMS & SKIN CARE	LAUNDRY DETERGENT
Number of brands with 2011-2014 market share > 2% (global vs. local)	7 (2 vs. 5)	8 (3 vs. 5)	10 (6 vs. 4)	8 (4 vs. 4)	8 (3 vs. 5)
Combined market share (in volume, over period 2011–2014)	57%	83%	68%	34%	89%
Number of brand manufacturers (global vs. local)	7 (2 vs. 5)	8 (3 vs. 5)	5 (4 vs. 1)	7 (4 vs. 3)	6 (2 vs. 4)
Average purchase frequency (per year per panelist)	3.1	5.7	3.8	5.3	3.5
Average volume received as gift as % of volume consumed	12%	8%	11%	12%	8%
Average category penetration (per year)	7 (2 vs. 5)	8 (3 vs. 5)	10 (6 vs. 4)	8 (4 vs. 4)	8 (3 vs. 5)

Table 3.3: Operationalization marketing mix variables

VARIABLE	OPERATIONALIZATION	REFERENCE
Price	Relative price, calculated as price (in ¥) per volume (e.g., per milliliter), of brand j in category c in week w, relative to the average price per volume in category c in week w.	Cleeren, van Heerde, and Dekimpe (2013)
Distribution	Weighted average of indicators of availability for brand j in category c in week w across all retailers, weighted by the retailers' market shares in week w-1.	Sotgiu and Gielens (2015)
Promotion[a]	Percentage of stock keeping units of brand j in category c that are in price promotion in week w.	Srinivasan et al. (2004)
Line length	Total number of unique stock keeping units that brand j offers in category c in four-weekly period f.	Ataman, Van Heerde, and Mela (2010); Geyskens, Gielens, and Gijsbrechts (2010)
Advertising	Share of Voice, calculated as Adstock of brand j in category c in week w relative to the Adstock of category c in week w, where: - Adstock of brand j in category c in week w = (1-λ)*Advertising + λ*Adstock of brand j in category c in week w-1; and - Adstock of category c in week w = (1-λ)*Advertising + λ*Adstock of category c in week w-1. Following George, Mercer, and Wilson (1996), λ equals .8.	Gijsenberg et al. (2011); Luan and Sudhir (2010)

[a] In our purchase data, no promotional information is present, therefore we work with a proxy measure.

For instance, at the start of our observation period (2011), 36.8% of breakfast cereal sales go to global brands and 63.2 % to local brands. Moreover, over time, global brands are growing at the expense of local brands (e.g., for breakfast cereals their share increases by almost 20%, from 36.8% to 44.1%). These patterns are found in all categories except body creams & skin care, where local brands are not only bigger but also growing more strongly. At the same time, in all five categories, we see differences between brands within a given brand type. For example, while GB2 in breakfast cereals is gaining share, the market share of GB1 is getting smaller every year.[14]

[14] Brand numbers are randomly assigned.

In Western markets, such large temporal variations in market shares are uncommon in the mature CPG industry. However, in EMs, brands are a more recent phenomenon and market shares are subject to greater fluctuation, which leaves less room for complacency by brand managers. Hence, it is not surprising that EMs are increasingly seen as contexts in which managers of Western corporations have to prove their mettle before they are promoted to senior positions.

Table 3.4: Market shares top brands in breakfast cereals over period 2011-2014

BRAND NUMBER (GB=Global Brand; LB=Local Brand)	MARKET SHARE (VOLUME)[a]			
	2011	2012	2013	2014
GB1	5.80%	4.82%	4.40%	3.49%
GB2	31.04%	37.12%	39.56%	40.66%
LB1	11.05%	9.72%	8.54%	7.51%
LB2	5.11%	3.59%	3.04%	3.07%
LB3	5.68%	6.32%	7.31%	7.65%
LB4	36.13%	32.93%	31.70%	31.80%
LB5	5.18%	5.51%	5.44%	5.84%
Sum GBs	36.84%	41.93%	43.96%	44.15%
Sum LBs	63.16%	58.07%	56.04%	55.85%

[a] Rescaled to sum to 100%.

Brand (Type) Switching

The brand-share evolutions may be due to consumers building up experience with (different) brands over time, and shifting their preferences as a result. Table 3.5, Panel A indicates how many unique top brands Chinese people buy or receive as gift in the breakfast cereals category (see Appendix 3.B for similar tables of the other four categories). The average consumer buys about two different cereal brands and 30.8% stick to one and the same brand. Moreover, even single-brand buyers may enjoy multiple-brand experiences because of gift giving. As the table shows, a non-negligible fraction (i.e., 15.8%) of consumers receives gifts of two or more different cereal brands and, for those who get gifts of only one brand (28.3%), quite often this is a brand outside their purchase set (9.3%). Taken together the data thus reveal that, indeed, most consumers consume more than one cereal brand. As shown in Appendix 3.B, this pattern is even stronger in the other categories, where 4.6% to 9.7% of the panelists are single-brand buyers.

Table 3.5, Panel B sheds more light on consumers' underlying purchase switches. For breakfast cereals, in more than one out of four shopping trips, consumers buy a brand different from that bought on the previous occasion (26.5%). Many of the switches involve brands of different types, where the fraction moving from a local to a global cereal brand (7.9%), and is somewhat higher than the reverse (i.e., only 7.2% switching from global to local). In the other categories, the overall degree of switching is higher, but the pattern is by and large the same. In sum, consumers build up experience with multiple brands, and with global as well as local brands. This enables them to learn about the quality of these brands which, in turn, may influence their subsequent choices and fuel some of the observed dynamics[15].

As EM conditions tend to be more fluid than those in DMs (Sudhir et al. 2015), such high degree of brand switching is not unusual. Thus, the Chinese market offers a rich and variegated empirical setting for studying brand learning.

Brand Marketing Mix

The brand dynamics in the marketplace may be affected by the brands' marketing mix. Table 3.6 displays the price, distribution, promotion, line length, and advertising levels for breakfast cereals across time, and brands/brand types. The marketing mix pressure generally differs between global and local brands. Global brands typically have larger line lengths, and are more expensive, more widely distributed, and more heavily advertised and promoted than local brands.

[15] One may argue that people may switch brands not with the intention to learn about quality but just because they like variety. The model-free evidence on brand switching is provided to show that households in our data have the opportunity to learn about the quality of different brands through consumption. Even when households are switching because of variety seeking, they still will learn about the brands they switch between such that over time their quality beliefs about these different brands will come closer to their true qualities.

The level of marketing mix instruments changes over time, and differently so across brand types. In breakfast cereals, while global brands on average increased their prices, line length, and advertising, this pattern does not hold for local brands, which rather stepped up their promotion activity while lowering their advertising expenses.

Important differences can be observed between brands of a given type. In breakfast cereals, GB1 is higher-priced and less widely available than GB2, and the increase in line length is much stronger for GB2 than for GB1. LB1 on the other hand is more expensive than the category average, whereas all other local brands are cheaper. Similar observations hold in the other categories (see Appendix 3.C): the marketing mix pressure differs between but also within brand types, and brands change the level of their instruments over time. These differences in the marketing mix between brands, brand types, and over time may already explain some of the changes in brand (type) share.

To conclude this section: the data illustrate the inherent dynamism and heterogeneity in EMs noted by previous authors (Burgess and Steenkamp 2006; Sudhir et al. 2015) even in CPG categories. We see important changes over time in brands' share and marketing mix, and substantial switching within individual consumers' purchase histories. Global brands are gaining market share in general, and with quite some heterogeneity between brands within a brand type. The question remains what drives this (heterogeneous) evolution. Our model will shed light on this.

Table 3.5: Brand (Type) Switching in breakfast cereals over period 2011-2014[a]

PANEL A: CONSUMPTION VARIETY										
Number of top brands		0	1	2	3	4	5	6	7	Weighted average[b]
% of households buying/receiving top 7 brands	Purchase	n.a.[c]	30.8%	38.6%	21.7%	7.1%	1.6%	.2%	.0%	2.1
	Gift	55.9%	28.3%	11.5%	3.5%	.7%	.1%	.0%	.0%	.7

PANEL B: RELATIVE FREQUENCY OF % OF SHOPPING TRIPS IN WHICH A HOUSEHOLD SWITCHED BRAND TYPES					
	Global Brand to Global Brand	Local Brand to Global Brand	Global Brand to Local Brand	Local Brand to Local Brand	Total
% of shopping trips where one switched brands	2.4%	7.9%	7.2%	9.1%	26.5%

[a] Note: only panelists that belong to the samples on which our models are estimated, are included.
[b] The 'weighted average' number of purchased brands is obtained as: 1*.308+2*.386+3*.217+4*.071+5*.016+6*.0002=2.1.
[c] n.a. = not applicable.

Table 3.6: Level marketing mix instruments top brands in breakfast cereals over period 2012-2014[a]

BRAND NUMBER (GB=Global Brand; LB=Local Brand)	PRICE			DISTRIBUTION			PROMOTION			LINE LENGTH[b]			ADVERTISING		
	2012	2013	2014	2012	2013	2014	2012	2013	2014	2012	2013	2014	2012	2013	2014
GB1	1.44	1.40	1.50	.72	.74	.74	.11	.09	.09	.13	.14	.16	.0007	.0006	.00
GB2	.67	.73	.74	.86	.87	.88	.09	.08	.08	.34	.46	.60	.14	.36	.52
LB1	1.19	1.18	1.20	.80	.80	.79	.06	.05	.07	.14	.14	.14	.39	.16	.18
LB2	.79	.65	.63	.50	.47	.45	.02	.03	.02	.23	.19	.17	.05	.003	.0001
LB3	.99	.86	.83	.71	.73	.75	.06	.08	.15	.16	.18	.17	.00	.00	.00
LB4	.81	.77	.76	.84	.85	.85	.10	.10	.09	.33	.36	.41	.15	.07	.07
LB5	.93	.86	.83	.67	.69	.71	.05	.07	.07	.14	.14	.14	.00	.00	.00
Average GBs	1.05	1.07	1.12	.79	.80	.81	.10	.09	.08	.23	.30	.38	.07	.18	.26
Average LBs	.94	.86	.85	.70	.71	.71	.06	.07	.08	.20	.20	.21	.12	.05	.05

[a] 2011 is the initialization year, so is not shown here.
[b] In 100s.

Model-Based Results

Model Selection

In each product category, we estimate a sequence of three models. We start with a multinomial logit model (M0), with random effects for the brand constants, marketing mix instruments, last brand acquired, and the household's brand acquisition share in the initialization period. Next, we take the learning aspect into account: we estimate a Bayesian learning model (M1), in which the initial variance is equal across brands. Then, we introduce the brand-type factor (M2) by allowing the initial variance to differ between local and global brands.

Households have an initialization period of one year, starting from their first observed purchase in the category after January 1st 2011; all periods after the initialization year belong to the estimation dataset. Households eligible for estimation are those that spent 70% or more of their category budget on the selected top brands (brands with a market share exceeding 2%); bought or received a top brand in their initialization year at least twice; and bought a top brand in their estimation period at least twice. For reasons of tractability (Rossi, McCulloch, and Allenby 1996), we estimate the models on a randomly selected subset of these eligible households.[16] Table 3.7 reports the model fit. Our learning models clearly win in 4 out of 5 categories: going from M0 to M1 or M2 improves fit in all categories, except one (i.e., laundry detergent). Because model M0 is a very 'stringent' benchmark (incorporating both households' initial purchase shares and previous-trip purchases), this is quite a strong result that underscores the importance of accommodating learning. Compared to M1, model M2 (which accommodates a different initial-uncertainty parameter for global than local brands)

[16] Depending on the number of eligible households in the category, we estimate the model on a randomly chosen subsample of 25% (potato chips, shampoo, laundry detergent) or 75% (breakfast cereals, body creams & skin care) of all eligible households. This keeps estimation manageable while ensuring an estimation sample of sufficient size in each category and city tier.

leads to a lower AIC and BIC in two out of five categories (potato chips and body creams & skin care). For consistency, we discuss the results of M2 for all categories (which, for categories where it does not offer a fit improvement, are very similar to those of M1 anyway).[17]

Quality of Global vs. Local Brands

Table 3.8, Panel A displays the parameter estimates for M2 per category (the results for M1 can be found in Appendix 3.D: the table shows that the estimates are very similar). Many of the quality estimates are significant relative to the reference brand (chosen to be the brand with the largest market share in the category), with positive as well as negative values – pointing to (mean) quality differences between brands in a given category. Figure 3.1 shows the individual brand qualities, mean-centered by category. It can be seen that the quality of global brands is generally higher than the category average, whereas local brands' quality is often lower. Brand type does not tell the whole story however. For instance in potato chips, the quality of global brand GB3 is below average, whereas that of local brands LB1, LB3, and LB4 exceeds the category average. At the same time, the standard deviations of quality mixing distributions often are significant, indicating that consumers differ in their quality assessment for a given brand.

To examine how Chinese consumers perceive the quality of global vs. local brands, Table 3.8 (Panel B) provides the 'Global-to-Local Brand Quality Ratio' for each category. In the spirit of Erdem, Zhao, and Valenzuela (2004), we define this ratio as the average quality of global brands minus that of the worst-performing brand in the category, divided by the average quality of local brands minus that of the same worst-performing category brand.

[17] Note that learning models are not all about fit improvement, but more about understanding the underlying processes (Chintagunta 2018).

Table 3.7: Model fit

	M0: Multinomial logit model	M1: Bayesian learning model	M2: M1 + separate initial variance for local vs. global brands
Breakfast Cereals			
Log-likelihood	-16,167.0	-15,568.7	-15,567.8
AIC[a]	32,386.1	**31,187.3**	31,187.5
BIC[b]	32,446.8	**31,252.7**	31,255.6
Potato Chips			
Log-likelihood	-59,527.4	-58,777.8	-58,759.1
AIC[a]	119,110.7	117,609.5	**117,574.2**
BIC[b]	119,186.8	117,689.7	**117,657.4**
Shampoo			
Log-likelihood	-51,254.5	-50,588.7	-50,588.7
AIC[a]	102,573.0	**101,239.5**	101,241.3
BIC[b]	102,653.3	**101,325.7**	101,330.3
Body Creams & Skin Care			
Log-likelihood	-27,731.3	-27,056.0	-27,035.9
AIC[a]	55,518.5	54,166.0	**54,127.9**
BIC[b]	55,584.2	54,236.1	**54,200.6**
Laundry Detergent			
Log-likelihood	-55,832.1	-55,979.2	-55,978.1
AIC[a]	**111,720.1**	112,012.4	112,012.2
BIC[b]	**111,793.7**	112,090.5	112,093.2

[a] AIC = Akaike information criterion, bold numbers indicate lowest value across models.
[b] BIC = Bayesian information criterion, bold numbers indicate lowest value across models.

In all cases, this ratio significantly exceeds unity, meaning that global brands on average are seen as being of higher quality than local brands. The average quality ratio across product categories is 1.84, which means that on average, global brands enjoy a substantial quality advantage over local brands in China.

Quality Uncertainty of Global vs. Local Brands

While the mean quality of global brands is generally higher than the mean quality of local brands, to what extent are consumers uncertain about their quality? If higher quality is combined with lower uncertainty, that would put global brands in a very strong position in China, given that the risk parameter in Table 3.8 (Panel A) points to significant risk aversion on average. As the estimates of the initial variances can only be interpreted relative to one another, Table 3.8 (Panel B) provides the ‘Global-to-Local Initial-Uncertainty Ratio’: the ratio of the initial variance in quality beliefs of global brands relative to local brands. Values greater than one indicate that consumers are more uncertain initially about the quality of global brands than about that of local brands; values lower than one point to the opposite.

In four out of five categories, consumers are more uncertain about the quality of global brands than local brands (potato chips is the only exception, with a ratio of .89). However, in only two categories the quality uncertainty of global brands differs significantly from the quality uncertainty of local brands (as was already expected based on the fit comparison between M1 and M2). The mean ratio is 1.02, showing that on average, global brands do not particularly suffer or benefit lower or higher quality uncertainty than local brands. Thus, on the whole, consumers experience global brands to be of higher quality, but are about equally certain about the quality of these brands. It might however be that differences exist in quality (uncertainty) beliefs across different consumer groups, something we will look at in the last part of this section.

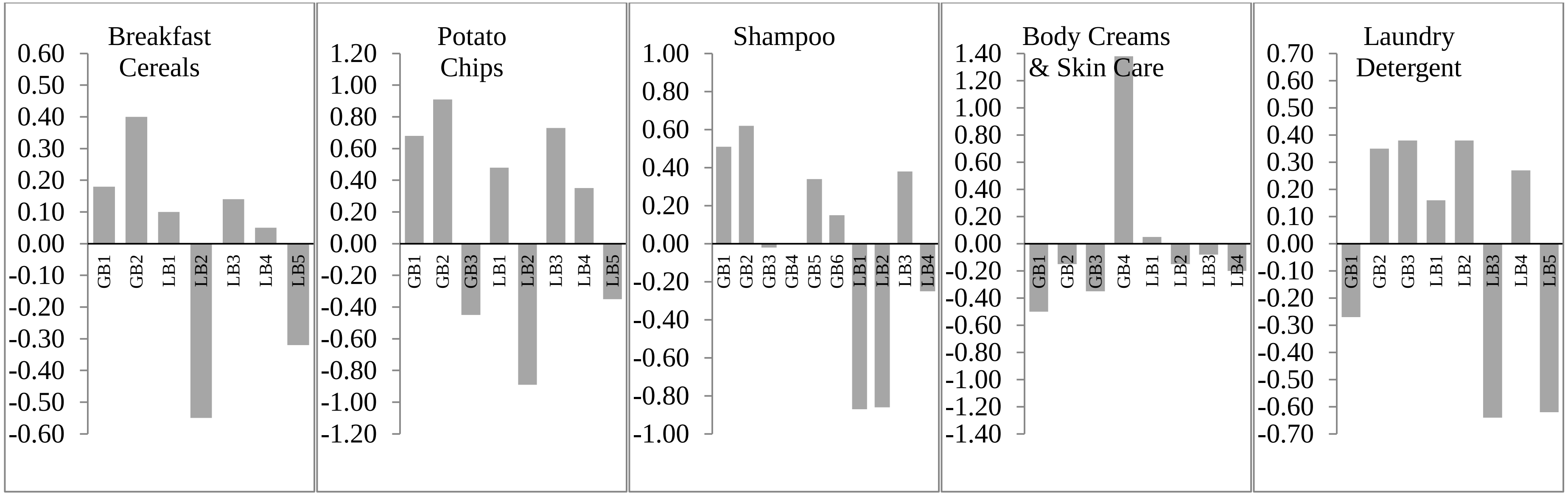
Table 3.8: Result learning model M2

PANEL A: PARAMETER ESTIMATES					
BRAND NUMBER (GB=Global Brand; LB=Local Brand)	BREAKFAST CEREALS	POTATO CHIPS	SHAMPOO	BODY CREAMS & SKIN CARE	LAUNDRY DETERGENT
Brand Quality					
GB1[a]	.13 (.24)	0[b]	.14 (.98*)	-.54* (1.64*)	-.64* (1.94*)
GB2[a]	.35* (.51*)	.23 (.70*)	.24* (.84*)	-.20* (1.04*)	-.03 (1.42*)
GB3[a]	.	-1.13* (1.80*)	-.39* (1.76*)	-.40 (1.15*)	0[b]
GB4[a]	.	.	-.38* (1.00*)	1.34* (.77*)	.
GB5[a]	.	.	-.03 (.74*)	.	.
GB6[a]	.	.	-.23 (1.46*)	.	.
LB1[a]	.06 (1.10*)	-.19 (.84*)	-1.24* (1.87*)	0[b]	-.22 (1.21*)
LB2[a]	-.59* (1.19*)	-1.57* (1.36*)	-1.23* (1.44*)	-.19* (1.00*)	.01 (1.09*)
LB3[a]	.10 (1.59*)	-1.40* (1.29*)	0[b]	-.13 (.81*)	-1.02* (2.09*)
LB4[a]	0[b]	-.33 (.90*)	-.63* (1.25*)	-.25* (1.37*)	-.11 (1.09*)
LB5[a]	-.37* (1.47*)	-1.02* (.98*)	.	.	-1.00* (1.76*)
Brand quality uncertainty					
Log Initial variance GBs	1.03*	.88*	1.02*	1.14*	1.03*
Log Initial variance LBs	1[b]	1[b]	1[b]	1[b]	1[b]
Log Signal variance	1.51*	1.26*	.66*	.61*	-.18*
Risk[a]	-1.71* (.34*)	-1.13* (.26*)	-.78* (.34*)	-1.11* (.40*)	-1.76* (.70*)
Marketing mix					
Price[a]	-.56* (1.16*)	-1.48* (1.22*)	-.64* (1.11*)	-.60* (.06)	-.37 (1.53*)
Distribution[a]	1.61* (.98*)	1.49* (1.36*)	1.13* (.24)	2.30* (.96)	.36 (.47)
Promotion[a]	.33 (.68)	.08 (.28)	.62 (.14)	.62* (1.71*)	1.14* (1.60*)
Line length[a]	.20 (2.49*)	.49* (.11*)	.27* (.31*)	.20 (.23*)	.38* (.57*)
Advertising[a]	-.14 (1.10*)	.29* (.61*)	-.31* (1.11*)	-1.97* (.44)	-.04 (.90*)
Number of households	2,911	4,587	4,427	2,772	5,423
Number of observations	289,933	747,496	602,240	315,808	621,928

PANEL B: RATIOS DERIVED FROM PARAMETER ESTIMATES					
	BREAKFAST CEREALS	POTATO CHIPS	SHAMPOO	BODY CREAMS & SKIN CARE	LAUNDRY DETERGENT
Global-to-Local Brand Quality Ratio: (mean Quality GBs – minimum Quality) to (mean Quality LBs – minimum Quality)	1.94[†]	1.91[†]	2.43[†]	1.48[†]	1.44[†]
Global-to-Local Initial-Uncertainty Ratio: Initial Variance GBs to LBs	1.03	.89[†]	1.02	1.15[†]	1.03

[a] Mean across households; SD across households in parentheses.
[b] Parameter fixed.
* Significant at $p < .05$.
[†] Global significantly different from local at $p < .05$.

Figure 3.1: Quality of global and local brands[a]



[a] Quality scores are mean-centered within categories.

Table 3.9: Brand choice elasticities for quality uncertainty, price, distribution, promotion, line length, and advertising (averaged across brands in category) by product category

	BREAKFAST CEREALS	POTATO CHIPS	SHAMPOO	BODY CREAMS & SKIN CARE	LAUNDRY DETERGENT	MEAN ACROSS CATEGORIES
Quality uncertainty[a]	-4.01	-2.60	-0.85	-2.85	-4.23	-2.91
Price[b]	-0.46	-1.22	-0.55	-0.24	-0.33	-0.56
Distribution[b]	1.00	1.04	0.88	1.81	0.27	1.00
Promotion[b]	0.02	0.01	0.02	0.04	0.04	0.03
Line length[b]	0.04	0.19	0.27	0.14	0.27	0.18
Advertising[b]	-0.01	0.02	-0.02	-0.05	-0.003	-0.01

[a] Quality uncertainty elasticity is calculated as: $\eta(\text{quality})_{jc} = \hat{r_c}(1 - \overline{\text{Pr}}_{cj})\sigma_{jc0}$, where $\hat{r_c}$ is the mean estimate of the risk parameter, $\overline{\text{Pr}}_{cj}$ is the brand's average choice share in the category, and σ_{jc0} is the initial uncertainty of brand j's brand type.

[b] The marketing mix (i.e., price, distribution, promotion, line length, and advertising) elasticities are calculated as: $\eta(k)_{jc} = \widehat{k_c}(1 - \overline{\text{Pr}}_{cj})\overline{k_{cj}}$ where $\widehat{k_c}$ is the mean estimate of marketing mix parameter k, $\overline{\text{Pr}}_{cj}$ is the brand's average choice share in the category, and $\overline{k_{cj}}$ is the average level of brand j's marketing mix variable k in category c.

Importance of Brand-Quality Uncertainty and Other Marketing Mix Activity for Brand Success

Having quantified EM consumers' brand-quality (learning) processes and the effects of the other marketing mix instruments, the question becomes: How important are these for brand success? Which ones are very important, which ones less important? Since the different coefficients are scale-dependent, we calculate scale-free elasticities. Table 3.9 reports the percentage change in the brand's choice probability due to a one percent increase in the brand quality uncertainty[18], price, distribution, promotion, line length, and advertising.

Our results show that quality uncertainty is a key driver of brand choice, in each category. That is, quality uncertainty elasticity always has the highest elasticity (except for shampoo, where the quality uncertainty elasticity is slightly lower than the distribution elasticity). On average, a 1% decrease in perceived quality uncertainty increases the likelihood of brand purchase by 2.91%. When comparing absolute effect sizes, the second most important marketing mix instrument is distribution, with an average elasticity of 1.00. Distribution always has the second highest impact on brand choice, except for body creams & skin care, where line length and price have about the same effect sizes as distribution. Across categories, with average elasticities of -.56 and .18 respectively, price and line length only have a moderate effect on choice. The effects of promotion and advertising are generally much smaller and often do not reach significance.

Consumer heterogeneity: The Role of Geographics and Sociodemographics

The significant standard deviations of the parameter mixing distributions (see Table 3.8, Panel A) point to household differences. To explore the sources of this heterogeneity, we run auxiliary regressions on pooled data across households and categories, with household

[18] See the legend of Table 3.9 for details on the calculations.

geographics and sociodemographics (region, city tier, income, and age) and category dummies as explanatory variables. As dependent variables, we use the households' posterior estimates for two ratios that reflect the difference in quality, and in quality uncertainty, between global and local brands; as well as the households' posterior estimates (expressed as elasticities to ensure comparability across categories) for risk attitude and (other) marketing mix response (see Table 3.10, Panel A for the operationalization of the dependent variables). More specifically, for each dependent variable $\eta_{k_{ic}}$, we estimate the following regression:

(3.8) $$\eta_{k_{ic}} = \varphi_{k_0} + \varphi_{k_1} ct_i + \varphi_{k_2} rg_i + \varphi_{k_3} ag_i + \varphi_{k_4} ic_i + \sum_{c=2}^{5} \vartheta_{k_c} cat_c + e_{k_{ic}}$$

where

ct_i = city tier in which consumer i lives (1 if high tier; -1 if low tier);

rg_i = region in which consumer i lives (1 if East; -1 if South, West, or North);

ag_i = age of consumer i (in 10 years);

ic_i = income of consumer i (in 1000 ¥);

cat_c = category dummy (1 if category c; -1 otherwise);

e_{ic} = random component.

Because the dependent variables are estimated quantities, the random component $e_{k_{ic}}$ comprises two parts: (i) the measurement (sampling) error $r_{k_{ic}}$ – the variance of which $\omega_{k_{ic}}^2$ is category-specific and can be calculated based on the variance-covariance matrix of the category's parameter estimates in the first stage[19] – and (ii) the part of the elasticity not explained by the drivers $v_{k_{ic}}$ – with unknown variance σ_k^2. Or: $e_{k_{ic}} = r_{k_{ic}} + v_{k_{ic}}$. To account for this error structure, we use the FGLS estimation approach proposed by Lewis and Linzer (2005), which is efficient and produces consistent standard errors, irrespective of the size of

[19] The variance of the measurement error $\omega_{k_{ic}}^2$ is calculated by performing Monte Carlo simulations on 100 draws.

σ_k^2 and the $\omega_{k_{ic}}^2$'s[20]. In addition, we use clustered-error regression to account for household replications across categories. Table 3.10, Panel C, summarizes the results.

Heterogeneity in Quality Perceptions of Global vs. Local Brands. The table shows that households' mean quality beliefs and belief uncertainty for global vs. local brands, significantly depends on where they live and their sociodemographics. Compared to high-income inhabitants of high-tier cities in the East of China, low-income inhabitants of low-tier cities elsewhere attach a lower quality premium to global brands, and are more uncertain about these brands. Although older consumers do not differ from younger consumers with respect to their mean quality perceptions about global vs. local brands, they feel more uncertain about the quality of global brands.

Our findings can help managers understand which consumers value their brands more highly, and/or which consumers to target with sampling promotion efforts because they are uncertain. Consider a manager who wants to know which consumers more strongly favor global over local brands. Using the regression estimates reported in Table 3.10, we find that for people with higher incomes who live in high-tier cities in the East, the Global-to-Local Quality Ratio equals 2.04 on average, compared to 1.96 for residents from low-tier cities with lower incomes who live elsewhere in China – a difference of about 4.5%. A similar type of analysis reveals that differences in uncertainty are larger: younger people with higher incomes from the high tier in Eastern China score more than 16% lower on initial uncertainty about global brands relative to local brands than older people with lower incomes from the low tier living elsewhere (1.33 vs. 1.15).

[20] This approach is a refinement of the commonly used WLS procedure with observation weights $\frac{1}{\omega_{k_{ic}}}$. We used this WLS procedure as a robustness check, and found the pattern of results to be similar.

Table 3.10: Results regression analyses

PANEL A: OPERATIONALIZATION DEPENDENT VARIABLES	
VARIABLE (ABBREVATION)	OPERATIONALIZATION
Quality Ratio (Qual Ratio)[a]	Posterior Global-to-Local Quality Ratio for household h in product category c, defined as: (average posterior quality of global brands for household h in category c – minimum posterior quality of all brands for household h in category c); divided by (average posterior quality of local brands for household h in category c – minimum posterior quality of all brands for household h in category c).
Uncertainty Ratio (Unc Ratio)	Posterior Global-to-Local Initial Variance Ratio for household h in category c, defined as: average initial variance of global brands at end of initialization period for household h in category c, divided by average initial variance of local brands at end of initialization period for household h in category c.
Uncertainty (Unc)	Posterior uncertainty elasticity for household h in category c, i.e., the average of [the product of a) the posterior risk estimate for household h in category c, b) the initial uncertainty of brand j in category c, and c) one minus the average market share of brand j across weeks in category c].
Price (Price)	Posterior price elasticity for household h in category c, i.e., the average of [the product of a) the posterior price estimate for household h in category c, b) the average level of price of brand j across weeks in category c, and c) one minus the average market share of brand j across weeks in category c].
Distribution (Distr)	Posterior distribution elasticity for household h in category c, i.e., the average of [the product of a) the posterior distribution estimate for household h in category c, b) the average level of distribution of brand j across weeks in category c, and c) one minus the average market share of brand j across weeks in category c].
Promotion (Promo)	Posterior promotion elasticity for household h in category c, i.e., the average of [the product of a) the posterior promotion estimate for household h in category c, b) the average level of promotion of brand j across weeks in category c, and c) one minus the average market share of brand j across weeks in category c].
Line length (LLength)	Posterior line length elasticity for household h in category c, i.e., the average of [the product of a) the posterior line length estimate for household h in category c, b) the average level of line length of brand j across weeks in category c, and c) one minus the average market share of brand j across weeks in category c].
Advertising (Adv)	Posterior advertising elasticity for household h in category c, i.e., the average of [the product of a) the posterior advertising estimate for household h in category c, b) the average level of advertising of brand j across weeks in category c, and c) the average market share of brand j across weeks in category c].

<table>
<tr><th colspan="7">PANEL B: DESCRIPTIVES GEOGRAPHIC AND SOCIODEMOGRAPHIC VARIABLES</th></tr>
<tr><th colspan="4">B1: FREQUENCIES REGION*CITY TIER</th><th colspan="3" rowspan="2">B2: MEANS AND STANDARD DEVIATIONS AGE AND INCOME</th></tr>
<tr><th></th><th>Low city tier</th><th>High city tier</th><th>Total</th></tr>
<tr><td>East</td><td>5,262
(32.02%)</td><td>6,525
(39.71%)</td><td>11,787
(71.74%)</td><td></td><td>MEAN</td><td>SD</td></tr>
<tr><td>South, West, North</td><td>2,259
(13.75%)</td><td>2,385
(14.52%)</td><td>4,644
(28.26%)</td><td>Age (in 10 years)</td><td>4.56</td><td>1.14</td></tr>
<tr><td>Total</td><td>7,521
(45.77%)</td><td>8,910
(54.23%)</td><td>16,431
(100%)</td><td>Income (in 1000 ¥)</td><td>6.87</td><td>3.17</td></tr>
</table>

PANEL C: PARAMETER ESTIMATES REGRESSION ANALYSES[b]

	Qual Ratio	Unc Ratio	Risk	Price	Distr	Promo	LLength	Adv
Intercept	2.00**	1.24**	-1.39**	-.91*	2.32**	.008**	.05*	-.09**
City tier (ct_i)	.01**	-.04**	.02**	.0007	-.002**	.0003**	-.0004	.0002*
Region (rg_i)	.01**	-.01**	-.002	-.003	-.0001	-.0001	.0007	.0002*
Age (ag_i)	.0001	.008**	-.02**	-.003	.00005	-.00001	-.001**	.00003
Income (ic_i)	.003**	-.004**	.003**	.002**	.00005	-.00002	.0001	.0001**
$Cat_{breakfast\ cereals}$	.04**	.15**	.36**	-.07**	.58**	-.002**	-.11**	-.07*
$Cat_{potato\ chips}$	.23**	-.10**	.60**	-.43**	.39**	-.02**	-.04**	.01*
$Cat_{shampoo}$	.34**	.09**	.69**	-.10**	.31**	-.01**	.002	-.009*
$Cat_{body\ creams\ \&\ skin\ care}$	.04**	.15**	.33**	.05**	.77**	-.003**	-.07**	-.02*
Number of observations	20,120	20,120	20,120	20,120	20,120	20,120	20,120	20,120
Number of clusters	16,431	16,431	16,431	16,431	16,431	16,431	16,431	16,431
R^2	.25	.33	.59	.52	.92	.50	.48	.91

[a] See Erdem, Zhao, and Valenzuela (2004) for a similar measure in the comparison of national brand and private label qualities.
[b] The independent variables age and income are mean-centered.
*significant at $p < .10$; **significant at $p < .05$.

Thus, global brands generally enjoy a quality premium over local brands, with relatively small differences across consumers with different sociodemographic profiles. In contrast, while on average consumers feel almost equally (un)certain about the quality of global vs. local brands, some consumer groups may feel much more uncertain about the quality of global vs. local brands than others.

Heterogeneity in Risk Aversion and Marketing Mix Response. Do Chinese consumers differ in their sensitivity towards brand quality uncertainty in general? Table 3.10, Panel C, shows that older consumers with lower incomes living in low-tier cities are significantly more risk-averse than younger consumers with higher incomes from high-tier cities (predicted elasticities equal -1.47 vs. -1.32, a difference of 10%).

Brand quality (uncertainty) is of course an integral element of the marketing mix the firm employs to make its offering attractive to the different consumers in the market. But what can managers expect of the effectiveness of other marketing mix instruments across consumers groups? It can be seen that high income households are less price sensitive, but somewhat more sensitive to advertising. The former is consistent with previous research that showed that price sensitivity decreases with income (Erdem and Sun 2002; Gao, Zhang, and Mittal 2017), while the latter is consistent with the economist's prediction that low-income consumers make more rational vs. emotional choices (and thus will be less influenced by advertising). In addition, people living in high-tier cities are less sensitive to distribution, but more sensitive to promotional activities. This makes sense as people from smaller (i.e., low-tier) cities (still) have less access to products and brands than their counterparts from larger cities, making them more sensitive to changes in distribution (McKinsey 2013a). Moreover, these people are believed to have "a stronger appetite to spend" – resulting in a lower promotion sensitivity (Morgan Stanley 2018). Consistent with the finding that older people like less choice (Rozin et al. 2006), we find that younger consumers are more sensitive for line length changes. Finally, consumers from high tier cities in the East are more sensitive to advertising, in line with the idea that people from the coastal areas are more likely to be influenced by emotional factors (McKinsey 2012).

Conclusion

To the best of our knowledge, we are the first to empirically document how EM consumers assess the quality of global and local brands, and how this affects their actual purchases. We quantify these processes, using Chinese scanner-panel data on five CPG categories that cover foods (breakfast cereals), snacks (potato chips), hair care (shampoo), skin care (body creams & skin care), and fabric care (laundry detergent) over four years in different city tiers. We structure our discussion around the three contributions set out for this study.

First, how do consumers in China perceive (and learn about) the quality of brands over time? How does this differ between global and local brands? We find that global brands are favored on quality. Despite earlier contentions and anecdotal evidence (BCG 2008), our results reveal that Chinese consumers generally attach considerably higher quality to global brands than to local brands. At the same time, Chinese consumers are not necessarily more certain about the quality of global brands. This suggests that on average, global brands have a decided quality advantage but no uncertainty advantage over local brands.

Second, what are the key drivers of brand success in China? We find that uncertainty about quality is a key factor in brand choice behavior. Quality uncertainty is a significant drag on brand success, given the strong risk aversion among Chinese consumers. Our findings indicate that it is the most important factor in brand choice behavior in China, followed by distribution. The role of price promotion and advertising is minor, while price and line length play generally moderate roles. These results underscore that while brand managers should continue to invest in securing physical shelf space, another (and even more critical) point is to reduce consumers' uncertainty about the quality of their brand. One way to achieve this, and to stimulate usage/trial, is to provide free samples to current non-users. This could be done via the 'worn path' of in-store sampling as well as generous refund policies if the brand does not live up to its expectations. Facilitating gift giving is another option in China. Chinese employers often reward their employees by paying them in kind. Providing the brand to employers at a reduced rate so that they distribute it as a reward to their employees could reduce uncertainty about the quality of your brand substantially.

Third, is there systematic heterogeneity in quality beliefs and uncertainty of global vs. local brands? Although generally, Chinese consumers assign a higher quality premium to global than to local brands, while being equally uncertain about the quality of these brands, this is less so for consumers with lower incomes who live in low-tier cities in the West, North or South of China. That is, these consumers have lower quality beliefs about global vs. local brands, and feel more uncertain about the quality of these brands than their counterparts in high-tier cities in the East. Older consumers, too, feel more uncertain about the quality of global vs. local brands. This corroborates the notion that younger, richer people from the coastal areas of China (i.e., the East) have longer experience in consumer markets (McKinsey 2012) and, therefore, may be less uncertain about the quality of the relatively new global brands. Furthermore, consumers' sociodemographic profiles also affect their sensitivity to quality uncertainty; especially older consumers with lower incomes living in low-tier cities being more risk averse. Hence, brand managers who seek to enhance the perceived quality level and reduce the quality uncertainty for their brands, should particularly focus on older, less affluent consumers from low-tier cities in the non-Eastern parts of China.

Limitations and Directions for Further Research

Our study has limitations that open up several new research questions. First, we only analyzed one EM. Though China is the biggest emerging-market economy, and shares features with other EMs, it also has distinct characteristics. Future studies should verify to what extent our results hold in other EMs.

Second, like extant literature on consumer learning processes, we used data on the household rather than the individual level. As indicated by Bruno, Cebollada, and Chintagunta (2018) individuals within a household may largely differ in their purchase behavior. It is questionable whether this issue plays a large role in our case. Group harmony is very important in a collectivistic country like China, increasing the likelihood that these people make brand choices that are in line with what their fellow household members would choose. Moreover, to the extent that differences in purchase behavior decrease with household size, they will play less of a role in the current study because, due to China's one-child policy, Chinese households typically consist of few members.[21] Nevertheless, future research should verify if and to what extent the issue exists.

Third, our analysis of the sources of consumer heterogeneity in the quality and marketing parameters was limited to the sociodemographic variables collected by Kantar Worldpanel. In all, the impact of these sociodemographics on consumers' marketing mix responsiveness appears rather small, which may be linked to the fact that we consider an urban panel, and/or suggest that other consumer characteristics may play a role. Future research should consider a richer set of variables, including consumer traits.

Fourth, though we found clear patterns across the categories studied, there were also differences. These could be due to several factors, such as the newness (or 'foreignness') of the category to Chinese consumers, its expensiveness, or whether it is publicly or privately consumed (BCG 2008). Given the rather small number of categories studied, we could only speculate on the underlying reasons. Future studies could consider a broader range of CPG

[21] China replaced the one-child policy into a two-child policy in 2016: our data ranges to 2014 and 80% of our households have a maximum of 3 members.

categories.

Fifth, as our main interest was in how EM consumers learn differently from consuming global vs. local brands, and how this uncertainty influences brand choice, we did not take into account whether consumers were new to a category or already had quite some experience with it. Heilman, Bowman, and Wright (2000) study how brand preferences and responses to marketing activities evolve for consumers that start consuming a new category: first-time parents who learn from using diapers and baby towels. Future research could look at what drives EM consumers to start consuming a new (possibly less locally-embedded) category, and how category learning in such cases takes place.

Finally, while our model incorporates that consumers' brand-quality perceptions change over time, like extant brand-choice models with Bayesian learning, it assumes that brand quality itself does not change. In reality, new SKU introductions may improve the brands' true quality, which, in turn, may affect consumers' quality beliefs. Incorporating these effects represents a modeling challenge that we leave for future study.

Much remains to be studied before we can offer definitive guidelines and empirical generalizations regarding brand learning and marketing mix effectiveness in emerging markets. We hope that this study will spark additional research on brand learning processes in such markets.

Chapter 4 | The Rise of Online Grocery Shopping: Which Brands Will Benefit?

Introduction

Online grocery shopping is on the rise. According to Nielsen and FMI (2018), by 2022 consumers worldwide will be spending 100 billion dollars a year on online grocery. A report of Kantar Worldpanel (2017) indicates that in 2015, only 4.6% of total groceries were sold online, but that this share is expected to reach 10% by 2025. While the trend in online grocery is visible in almost every country of the world, Asia is leading the way. Half of the top six countries in online grocery share growth are coming from this continent: 2016 shares increased to 19.7% in South Korea, 7.5% in the UK and Japan, 6.2% in China, 5.6% in France and 1.5% in the US (Kantar Worldpanel 2017). With more than 23 billion dollars spent on online CPG categories between September 2015 and August 2016, China has the biggest online grocery market in the world (Nielsen 2017).

For CPG brands, it is unclear how this increase in online grocery share will affect their total sales. While industry sources point to growth opportunities ("Online shoppers spend more," Kantar Worldpanel 2015, p.11), which brands will benefit from these opportunities (or, in contrast, suffer from the rise of online CPG), and why, has not been pinned down yet. In addition, stark differences exist in online vs. offline performance between brands. While some brands appear to hold similar market shares online and offline, others enjoy a dominant positon in their category in the offline channel, but do not seem able to capture a large portion of online category sales, or vice versa (see for example Kantar Worldpanel 2015, table on p.4). Moreover, category sales themselves may evolve differently as online gains way. What drives these differences? And how can brand managers make sure to be on the winning end?

What factors will influence brands' online relative to offline sales performance is not clear upfront. Academic studies to date have investigated brand success in both channels, but were interested in a non-monetary metric like loyalty (Danaher, Wilson, and Davis 2003) or considered online and offline choice shares for only a small set of categories and brands (Chu, Chintagunta, and Cebollada 2008; Degeratu, Rangaswamy, and Wu 2000). Moreover, most of these studies only focus on a small set of drivers, such as price (Chu, Chintagunta, and Cebollada 2008) or pack size (Chu et al. 2010). While Campo and Breugelmans (2015) looked at a large set of marketing mix instruments and intrinsic market characteristics, they focused on the online vs. offline performance of categories, not brands.

Industry reports hint at a large set of factors that may influence the success of brands and categories in the increasingly digital world. For example, Kantar Worldpanel (2015) and McKinsey (2013a) investigated consumers' motivations for purchasing online across multiple countries. Although these surveys provide interesting insights, they do not tell us anything about actual purchase behavior. More importantly, while these reports explore consumers' willingness to purchase in the online channel, they have little to say about online relative to offline performance. Some of the motivations mentioned were price, availability, assortment size, and heaviness: factors that are not only likely to play a role when purchasing online, but also in physical stores. Thus, how these factors increase or decrease a brand's or category's relative sales in the online vs. offline channel remains unclear.

The current study aims to fill this gap by answering two research questions. First: How do total brand sales change as the overall share of CPG sold online goes up? Using a decomposition approach, we show that this change critically depends on two easy-to-calculate indices, namely: the BOI ('brand online index', which is the brand's online category share relative to its offline category share) and the COI (‘category online index’: the category’s grocery share online relative to its grocery share offline). While the BOI metric captures to what extent the brand’s relative position in the category will improve or deteriorate with the rise of online; the COI metric indicates whether it will benefit (suffer) from the category’s (lack of) propensity to ‘bloom’ online. Together, our BOI and COI metrics can be used by managers as additional indicators (i.e., next to existing metrics like overall share and sales) of their ‘value at risk’ in a world where the online channel becomes more important. Our second research question is: What are the drivers of brands’ overall performance in such an increasingly digital market? To address this question, we consider a comprehensive set of brand and category characteristics that may affect the brand’s BOI and COI, respectively. While the brand factors are the direct result of managerial decisions, the category factors generally cannot be manipulated directly but have to be reckoned with, and can help managers to properly allocate their resources and anticipate future sales levels. We empirically test the impact of these drivers using a unique dataset that tracks the purchases of Chinese panelists for over 440 CPG brands in 60 categories, between 2011 and 2015 – a period in which the online CPG market started to take off in China.

The results of our study are relevant for academics and practitioners alike. We deepen academic knowledge on how CPG brand sales change as a result of the increasing popularity of the online channel, by identifying key underlying metrics, and studying a comprehensive set of brand and category factors driving these metrics. Moreover, we provide empirical evidence on their impact through a large-scale study covering a broad set of CPG categories and brands. From a practitioner perspective, our BOI and COI metrics can be easily added to existing dashboards to help brand managers gauge how the rise of online will threaten or further their relative position in the category, and to what extent they will benefit (suffer) from ‘riding the category waves’. We also provide actionable insights, by revealing which buttons managers can press to get the most out of the online grocery trend.

The remainder of the paper is structured as follows. First, we formally derive the link between brands’ sales change as the overall share of CPG sold online goes up, and their BOI and COI metrics. We then conceptualize what brand and category factors influence these metrics. Next, we discuss the methodology to estimate these effects, followed by a description of our empirical setting and data. Having presented the estimation results, we discuss implications and directions for future research.

Impact of Online Growth on Brand Sales

In this section, we start by uncovering the key metrics that underlie brands’ sales evolution as CPG online becomes more important. Next, we describe the possible drivers of overall brand performance, and formulate expectations on the direction of their effects.

Change in brand sales as overall share of CPG sold online goes up

Let b be a brand indicator, c a category indicator and t a period (e.g., year). We start by considering the baseline sales of a brand in a ‘pure offline’ world (denoted by superscript ‘0’), in which there are no online grocery stores and thus no online purchases yet. These baseline brand sales $S^0_{c,b,t}$ can be written as the product of the brand’s category share, the category’s grocery share, and the total grocery sales:

$$(4.1) \quad S^0_{c,b,t} = \left[\frac{S^0_{c,b,t}}{S^0_{c,.,t}}\right] * \left[\frac{S^0_{c,.,t}}{S^0_{.,.,t}}\right] * S^0_{.,.,t}$$

where $S^0_{c,.,t}$ denotes total sales of the category to which the brand belongs in a given period in a pure offline setting, and $S^0_{.,.,t}$ total grocery sales in such setting.

The moment that next to the offline channel, the online channel arises, the setting becomes a mixed-channel world, in which total brand sales $S^T_{c,b,t}$ consist of offline brand sales plus online brand sales, which can be written as follows:

(4.2) $$S^{T}_{c,b,t} = \left[\frac{S^{Off}_{c,b,t}}{S^{Off}_{c,.,t}}\right] * \left[\frac{S^{Off}_{c,.,t}}{S^{Off}_{.,.,t}}\right] * \left[\frac{S^{Off}_{.,.,t}}{S^{T}_{.,.,t}}\right] * S^{T}_{.,.,t} + \left[\frac{S^{On}_{c,b,t}}{S^{On}_{c,.,t}}\right] * \left[\frac{S^{On}_{c,.,t}}{S^{On}_{.,.,t}}\right] * \left[\frac{S^{On}_{.,.,t}}{S^{T}_{.,.,t}}\right] * S^{T}_{.,.,t}$$

where $S^{T}_{c,b,t}$ denotes the total (i.e., offline and online) sales of category's c brand b in time period t and, similar to before, offline (online) brand sales are the product of brand's offline (online) category share, the category's offline (online) grocery share, the share of total grocery sold offline (online), and the total grocery sales.

We are interested in how total brand sales evolve with the advent of the online channel. Therefore, we consider the ratio of brand sales in the mixed-channel world to the baseline:

(4.3) $$\frac{S^{T}_{c,b,t}}{S^{0}_{c,b,t}} = \left[\frac{\frac{S^{Off}_{c,b,t}}{S^{Off}_{c,.,t}}}{\frac{S^{0}_{c,b,t}}{S^{0}_{c,.,t}}}\right] * \left[\frac{\frac{S^{Off}_{c,.,t}}{S^{Off}_{.,.,t}}}{\frac{S^{0}_{c,.,t}}{S^{0}_{.,.,t}}}\right] * \left[\frac{S^{Off}_{.,.,t}}{S^{T}_{.,.,t}}\right] * \left[\frac{S^{T}_{.,.,t}}{S^{0}_{.,.,t}}\right] + \left[\frac{\frac{S^{On}_{c,b,t}}{S^{On}_{c,.,t}}}{\frac{S^{0}_{c,b,t}}{S^{0}_{c,.,t}}}\right] * \left[\frac{\frac{S^{On}_{c,.,t}}{S^{On}_{.,.,t}}}{\frac{S^{0}_{c,.,t}}{S^{0}_{.,.,t}}}\right] * \left[\frac{S^{On}_{.,.,t}}{S^{T}_{.,.,t}}\right] * \left[\frac{S^{T}_{.,.,t}}{S^{0}_{.,.,t}}\right]$$

When quantifying the effect of the growth in online grocery on total brand sales, it is important to take into account that the trend towards the online channel might lead to a change in the overall grocery business. That is, the online channel could have an effect on total CPG sales: because of the new channel, people may buy more (expansion), or buy less (reduction). Expansion could occur for example because people have better or easier access to products via the online channel, whereas reduction could occur for example because with regular home deliveries time-constrained people can manage their inventories better, leading to less waste. If we let every dollar sold offline lead to an equivalent of g dollars sold online – where values of g smaller (larger) than 1 indicate reduction (expansion) – then:

(4.4) $$\frac{S^{T}_{.,.,t}}{S^{0}_{.,.,t}} = \left(1 - \left[\frac{S^{On}_{.,.,t}}{S^{T}_{.,.,t}}\right] + g * \left[\frac{S^{On}_{.,.,t}}{S^{T}_{.,.,t}}\right]\right) = \left(1 + (g-1) * \left[\frac{S^{On}_{.,.,t}}{S^{T}_{.,.,t}}\right]\right)$$

Our key question is: How does a change in the fraction of groceries sold online influence brand sales? In Appendix 4.A, we show the calculations and prove that if online does not lead to overall grocery expansion or contraction (g=1), the total sales of a brand will increase as a result of the online grocery share going up if

(4.5) $\left[\mathrm{BOI}_{c,b,t}\right] * \left[\mathrm{COI}_{c,.,t}\right] > 1$

where

$\mathrm{BOI}_{c,b,t} = \left[\frac{\left[\frac{s^{On}_{c,b,t}}{s^{On}_{c,.,t}}\right]}{\left[\frac{s^{Off}_{c,b,t}}{s^{Off}_{c,.,t}}\right]}\right]$ = brand’s online index, the brand’s category sales share online vs. offline,

and

$\mathrm{COI}_{c,.,t} = \left[\frac{\left[\frac{s^{On}_{c,.,t}}{s^{On}_{.,.,t}}\right]}{\left[\frac{s^{Off}_{c,.,t}}{s^{Off}_{.,.,t}}\right]}\right]$ = category’s online index, the category’s sales share online vs. offline.

If g>1 (expansion), even brands for which $\left(\left[\mathrm{BOI}_{c,b,t}\right] * \left[\mathrm{COI}_{c,.,t}\right] - 1\right) < 0$ can still gain from online growth. If g<1 (contraction), the condition becomes more stringent (i.e., even brands for which $\left(\left[\mathrm{BOI}_{c,b,t}\right] * \left[\mathrm{COI}_{c,.,t}\right] - 1\right) > 0$ can still lose sales).

Our decomposition shows that an increase in the CPG share online will affect total brand sales through two key metrics: BOI, the brand’s category share online relative to the brand’s category share offline, and COI, the category’s share in total grocery online relative to offline. A BOI higher (lower) than 1 indicates that the brand’s relative position in the category will improve (deteriorate) with the rise of online; a COI higher (lower) than 1 indicates that the category will achieve a larger (smaller) portion of consumers’ grocery wallet as the online channel grows.

Apart from the mathematical logic, looking at metrics like BOI and COI also makes intuitive sense when one is interested in how a brand’s overall performance will improve in a world where the online channel is growing. That is, selling through a channel that is growing is important when a brand’s aim is to (at least) maintain sales. First, a brand that operates in a category that sells relatively more through the offline than the online channel (i.e., COI is smaller than 1, for example because people perceive a lack of control when buying the category online), may be at risk because the total pie shrinks. For example, instead of buying fresh milk, consumers may choose long-life milk in case that is more safely bought via the online channel, reducing the potential for fresh-milk brands. Second, for brands that sell relatively more through the offline than the online channel (i.e., BOI is smaller than 1, for example because the brand offers only few SKUs online), growth of the online channel represents a threat (e.g., because consumers buy a product of a competitor within the category that does offer lots of SKUs online). In both cases, brands lose sales to competitors as consumers gravitate to the online channel – unless they take appropriate action.

Keeping track of the product of BOI and COI next to monitoring these indices separately, make sense intuitively too. That is, a low BOI will not lead a manager to believe that the online channel is necessarily a bad development: a high COI may compensate the low BOI such that the brand may still benefit from the online trend. On the other hand, a large BOI will also not lead a manager to draw the short-sighted conclusion that the online channel automatically brings prosperity: a low COI may actually result in the brand losing sales as online grows.

Drivers of overall brand performance

The next question then becomes: what drives these metrics? Building on extant literature in online and offline channels, we propose a set of relevant brand and category factors.

Drivers of BOI.

We discuss a comprehensive set of drivers that are the outcome of managerial decisions, namely variables related to the brand’s pack sizes, distribution (online and offline availability), price (online/offline price ratio and price position within the category), and communication (advertising spending), as well as brand ownership (foreign vs. domestic), brand trust, and to what extent the brand is a ‘fun’ brand.

Pack size. Brands that offer larger package sizes than usually sold in the category, are expected to have an advantage in the online channel (and a higher BOI). Ordering online (and having the products delivered at home, or into the car trunk at a pickup point) may avoid the physical burden of handling large package sizes – having to place them in a shopping cart and take them home. Hence, we expect large packs to be relatively more appealing in an online setting (Campo and Breugelmans 2015; Chu, Chintagunta, and Cebollada 2008). Moreover, people may find it hard to estimate the real size of a pack from a screen which, according to Burke et al. (1992), results in larger sizes being purchased more frequently online.

Online and offline availability. As for availability, we distinguish between availability in the online and offline channel. For most CPG brands, being available in a large number of online stores or marketplaces will, most likely, strongly drive online sales but not enhance offline performance. Hence, increasing online availability will be an important driver of BOI. The effect of offline availability on BOI is less clear upfront. On the one hand, being highly visible in offline stores might aid performance in the online channel (the so-called billboard effect, Avery et al. 2012). Indeed, research has shown that brands with a strong offline presence do better in the online environment (Danaher, Wilson, and Davis 2003). On the other hand, substitution effects may occur in that offline availability might reduce the consumer's propensity to buy the brand via the online channel. Which of these forces prevails is not clear upfront, so we leave the impact of offline availability on BOI as an empirical question.

Price position. Brands that are among the more expensive brands in the category are expected to do relatively better in the online compared to the offline channel. Several studies have shown that online, consumers are less price sensitive (e.g., Chu, Chintagunta, and Cebollada 2008; Degeratu, Rangaswamy, and Wu 2000; Lynch and Ariely 2000). This may be because they are more convenience than price oriented, or use price as a quality signal to make up for not being able to physically inspect the product. Moreover, expensive brands may more easily justify the payment of a delivery fee. Therefore, we anticipate that more expensive brands will have higher BOIs.

Online to offline price ratio. Common knowledge dictates that for a given brand, charging higher prices online than offline is expected to result in lower performance online relative to offline. Even if price sensitivity is lower in an online than in an offline setting (Chu, Chintagunta, and Cebollada 2008; Degeratu, Rangaswamy, and Wu 2000; Lynch and Ariely 2000), this does not mean that online shoppers do not pay attention to price at all. Given that the online channel facilitates price search (Häubl and Trifts 2000), consumers may notice online-offline price differences for a given brand and act upon them. Industry reports highlight that, indeed, finding lower prices online than in store is one of the motivations for Chinese consumers to shop online (China Internet Watch 2015; Kantar Worldpanel 2015). Thus, we expect the online to offline price ratio of a brand to have a negative effect on BOI.

Advertising spending. Though the online environment generally provides consumers with lots of easily accessible information (Shankar, Smith, and Rangaswamy 2003), especially information on sensory attributes (like freshness) or more abstract attributes (like quality) may be less available, which may increase perceived risk (Danaher, Wilson, and Davis 2003; Degeratu, Rangaswamy, and Wu 2000). One way to reduce this kind of risk is to signal quality via advertising (Erdem, Keane, and Sun 2008). We therefore expect more intense advertising by brands to enhance BOI.

Brand trust. Hernandez (2002) argues that brand trust plays a particularly important role in the decision process of online consumers. Consistent with this, Danaher, Wilson, and Davis (2003) show that high market-share brands (which are typically more familiar, and trusted; Chaudhuri and Holbrook 2001), enjoy a loyalty advantage online. Thus, we expect more trusted brands to have higher BOIs.

Fun brand. This construct measures the hedonic aspect of brands (Voss, Spangenberg, and Grohmann 2003)[22]. The more fun a brand is perceived to be, the higher (lower) the emotional (functional) benefits a brand has to offer (Steenkamp 2014). As the more ‘functional’ shopping environment will make people buy less on impulse online (Campo and Breugelmans 2015), we expect fun brands to have a disadvantage online compared to offline. Therefore, we postulate that brands perceived as being ‘fun’ will have lower BOIs.

Brand ownership (foreign vs. domestic). Foreign brands are brands owned by a manufacturer that originates from outside the country (in our case China), whereas domestic brands are owned by a domestic (i.e., Chinese) manufacturer. Especially in a country like China, performance of foreign vs. domestic brands may be different online vs. offline. First, online buyers often gravitate towards big brand names (Heilman, Bowman, and Wright 2000); brands that are typically owned by non-Chinese manufacturers. Also, the online environment assists brands to guard against counterfeiting, something especially foreign brands suffer from in China. For example, JD.com (one of the biggest B2C websites in China) has a zero tolerance policy with regard to counterfeit products, and Alibaba (owner of C2C website Taobao and B2C website Tmall), lets brands pay for prominence, using banners to assure consumers of the brand’s authenticity (Kantar Worldpanel 2015). Second, online marketplaces (such as Tmall Global and JD Worldwide) primarily invite foreign brands to sell through the venue, which provides them with an online-channel advantage over domestic competitors (China Briefing 2015). Hence, we expect that compared to domestic brands, foreign brands will perform relatively better online than offline (i.e., have higher BOIs).

[22] We look at ‘fun’ as a brand factor rather than as a category factor (see Steenkamp 2014 for a similar view): within a category, some brands may focus more on hedonic aspects while others may focus more on utilitarian aspects. For example, in shampoo, some brands may use fun scents like “apple pie” or “coconut island” whereas others may focus more on sensitive scalp or anti-dandruff.

Drivers of COI.

Expensiveness. Category expensiveness, i.e., the average amount paid on a typical category purchase (see e.g., Lourenço, Gijsbrechts, and Paap 2015), can have a dual effect on online relative to offline performance. On the one hand, buying expensive categories online (where the possibilities for physical inspection are limited) may be more risky. On the other hand, expensive categories may more easily justify the payment of a fee associated with online ordering and home delivery. So, though we expect category expensiveness to affect COI, the direction of the effect is not clear upfront.

Risk reduction function of brands. When buying from a category, consumers may choose well-known, trusted brands to reduce the risk of making the wrong purchase. The 'risk reduction function of brands' measures the extent to which a category's brands reduce the consumer's (perceived) risk of making a purchase mistake (Fischer, Völckner, and Sattler 2010). Categories that score highly on this construct (i.e., in which brands strongly act as 'risk-reducers') should have a benefit in the digital channel because many consumers shop online to find high-quality, branded products (Bain & Company and Kantar Worldpanel 2015), and, with less information on other attributes available (Degeratu, Rangaswamy, and Wu 2000), rely more heavily on brand names. Hence, we expect these categories to have a higher COI.

Advertising. For reasons similar to those of brand advertising, heavily advertised categories are expected to do relatively better online. Advertising messages may reduce perceived category risk and make consumers rely more strongly on brand cues available in an online setting. So, we expect higher COI for categories with high advertising spending.

Assortment size. We expect categories with large assortments to do especially well in the online channel. Kantar Worldpanel (2015) reported a 'wider range' as one of the most important motivations for Chinese shoppers to make online purchases. While consumers can enjoy the benefit of choice variety, they are less likely to experience choice overload. Search costs for products and product-related information are much lower online (Lynch and Ariely 2000), and consumers have several tools at their disposal (e.g., search bars and filters) to reduce their consideration set and identify the product that best satisfies their needs with relatively small effort (Häubl and Trifts 2000). Therefore, we expect that the larger a category's assortment size, the higher its COI.

Bulkiness/Heaviness. Categories that consist of bulky products (e.g., kitchen paper) and/or heavy items (e.g., cooking oil) are generally found to have an advantage in the online channel (Chu, Chintagunta, and Cebollada 2008). These categories are more likely to be purchased online for convenience reasons: consumers can, for a large part, outsource the handling and transportation of the products from the store to their homes (Campo and Breugelmans 2015). As a result, we expect more bulky and heavy categories to have larger COIs.

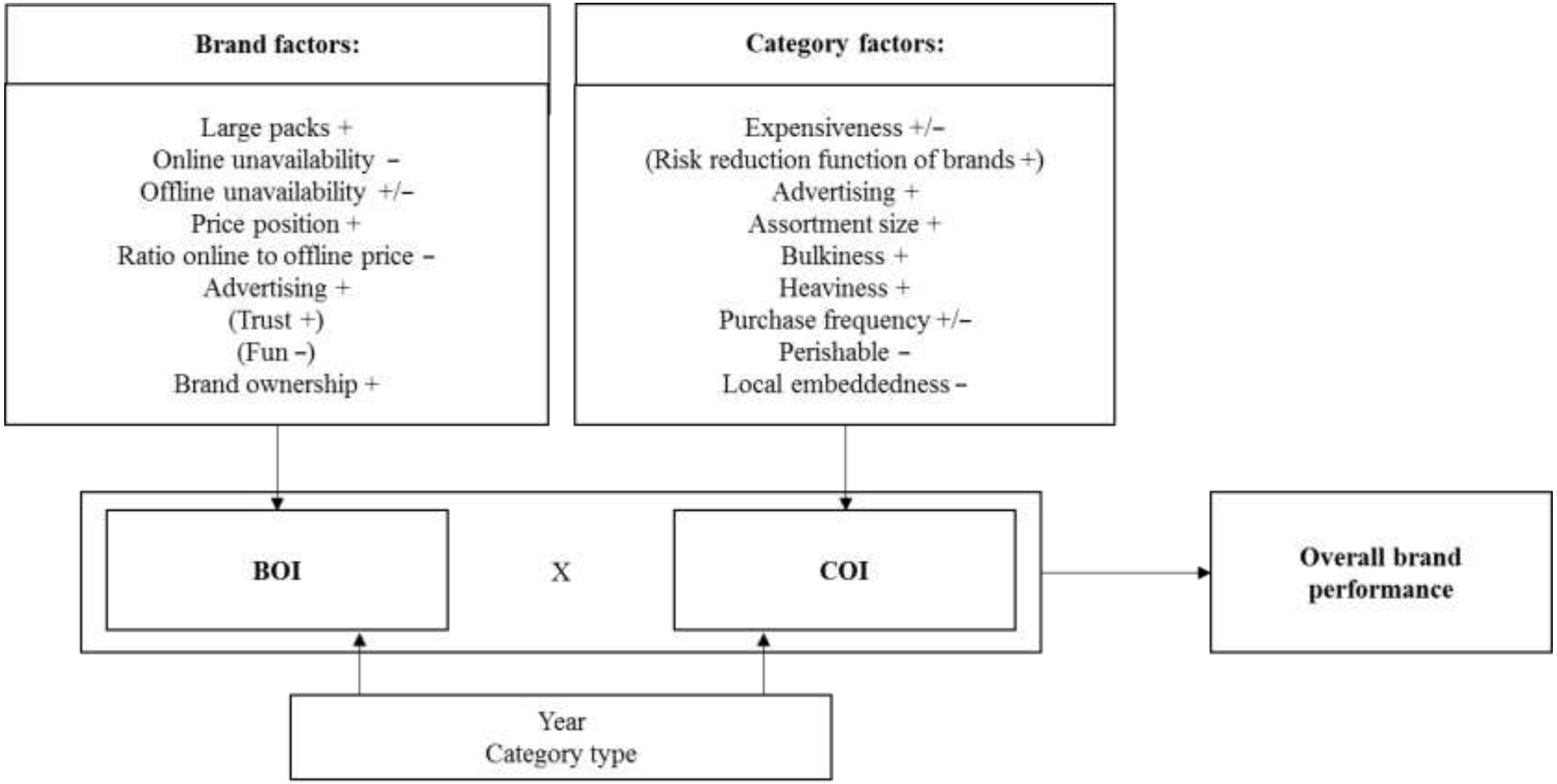
Purchase frequency. Category purchase frequency may have a dual impact on COI. On the one hand, grocery websites often create custom-made shopping lists for consumers based on previously-bought items. According to Kantar Worldpanel (2015), these online shopping lists are quite popular: more than half of online shoppers use them. Frequently bought items may be more likely to show up on the online shopping list and, therefore, to be bought online. On the other hand, to avoid delivery fees, consumers may predominantly shop online for large-basket, stock-up trips, including a larger proportion of less frequently needed products. Moreover, online shopping may reduce the purchase of unplanned items (Babin and Darden 1995), which typically belong to categories with low inter-purchase times (Inman, Winer, and Ferraro 2009). Because of these countervailing forces, we have no a priori expectation on the effect of purchase frequency on COI, but leave it as an empirical question.

Perishability. Consumers might feel a stronger need to physically inspect products from perishable than non-perishable categories prior to purchase. For example, shoppers may want to choose cheese that looks ‘fresh’ or buy milk with an expiry date that is still remote. Because the online setting offers no opportunity for physical inspection, perishable products are more likely to be bought offline (Chu, Chintagunta, and Cebollada 2008). As a result, we expect perishable categories to have lower COIs.

Local embeddedness. Local embeddedness reflects the extent to which consumers perceive the category as typical for, or originating from, the home country or region. For China, examples include tea and baijiu (distilled alcoholic beverage – Moutai being the most famous brand): these categories have been around for ages and are more deeply embedded in Chinese society than for instance coffee or wine. We expect less locally-embedded categories to better fit with the online channel than more locally-embedded categories. For one, the online channel is still relatively new in China (World Economic Forum 2016). People that score high on innovativeness are more likely to both adopt new channels (Arts, Frambach, and Bijmolt 2011) and be more open to try products from categories that are less ingrained in their culture. In addition, Chinese people use the online channel to explore and discover new products (BCG 2017), something less locally-embedded categories may benefit from. In sum, we expect categories with high local embeddedness to have a lower COI.

Figure 4.1 depicts our research framework and summarizes the expected effects. In the next section, we discuss how we will empirically test these effects.

Figure 4.1: Research framework[a]



[a] Factors between brackets are only available for a subset of brands and categories.

Methodology

Model Setup

To test our hypotheses, we run regressions with (the logarithm of) BOI and COI as the dependent variables[23], and the brand- and category drivers from Figure 1 as explanatory variables. Specifically, for BOI we estimate the following model:

(4.6) $\ln(\mathrm{BOI}_{c,b,t}) = \beta_0 + \beta_1 \mathrm{la}_{bt} + \beta_2 \mathrm{av}_{bt}^{on} + \beta_3 \mathrm{av}_{bt}^{off} + \beta_4 \mathrm{pp}_{bt} + \beta_5 \mathrm{rp}_{bt} + \beta_6 \mathrm{ad}_{bt} + \beta_7 \mathrm{tr}_b + \beta_8 \mathrm{fu}_b + \beta_9 \mathrm{fb}_b + \sum_k \gamma_k \,\mathrm{copula}_{kt} + \sum_{y=2}^{4} \delta_t \mathrm{year}_t + \sum_{p=2}^{6} \theta_p \mathrm{cattype}_{cb} + \varepsilon_{bt}$

where

la_{bt}	=	Brand b's % large packs in year t;
av_{bt}^{on}	=	Online availability of brand b in year t;
av_{bt}^{off}	=	Offline availability of brand b in year t;
pp_{bt}	=	Price position brand b in year t;
rp_{bt}	=	Ratio online to offline price of brand b in year t;
ad_{bt}	=	Adstock brand b in year t;
tr_b	=	Trust brand b (survey measure; only available for a subset of brands);
fu_b	=	Fun brand b (survey measure; only available for a subset of brands);
fb_b	=	Ownership brand b (foreign vs. domestic);
copula_{kt}	=	Gaussian copula for driver k in year t;
year_t	=	Year dummy (equal to 1 for year t, and -1 otherwise);
$\mathrm{cattype}_{cb}$	=	Category type dummy (equal to 1 if brand b's category c is in category type p, and -1 otherwise);
ε_{bt}	=	normally distributed error term for brand b in year t.

For COI, the following equation is estimated:

[23] Because both BOI and COI have skewed distributions (see the 'Data' section), we use a log transform.

(4.7) $\ln(\mathrm{COI}_{c,.,t}) = \alpha_0 + \alpha_1 \mathrm{cx}_{ct} + \alpha_2 \mathrm{rr}_c + \alpha_3 \mathrm{ad}_{ct} + \alpha_4 \mathrm{as}_{ct} + \alpha_5 \mathrm{bu}_c + \alpha_6 \mathrm{he}_c + \alpha_7 \mathrm{fr}_c + \alpha_8 \mathrm{pe}_c + \alpha_9 \mathrm{le}_c + \sum_k \gamma_k \,\mathrm{copula}_{kt} + \sum_{y=2}^{4} \delta_t \mathrm{year}_t + \sum_{p=2}^{6} \theta_p \mathrm{cattype}_c + \varepsilon_{ct}$

where

cx_{ct} = Expensiveness category c in year t;

rr_c = Risk reduction function of brands in category c (survey measure; only available for a subset of categories);

ad_{ct} = Adstock category c in year t;

as_{ct} = Category c's assortment size in year t;

bu_c = Bulkiness category c;

he_c = Heaviness category c;

fr_c = Average yearly purchase frequency category c;

pe_c = Perishability category c (perishable vs. non-perishable);

le_c = Local embeddedness category c;

$copula_{kt}$ = Gaussian copula for driver k in year t;

$year_t$ = Year dummy (equal to 1 for year t, and -1 otherwise);

$cattype_c$ = Category type dummy (equal to 1 if category c is in category type p, and -1 otherwise);

ε_{ct} = normally distributed error term for category c in year t.

Endogeneity concerns

When estimating equations (4.6) and (4.7), we face several endogeneity concerns.

First, because we estimate the models on a subset of brands (i.e., those already available online, see the 'Data' section), we face a selection problem. We resolve this through the 'control function' approach proposed by Dubin and McFadden (1984). In a first step, we estimate a binary logistic 'selection model' that explains whether and when brands are offered online, using observations on all brands available in the data set. Based on the estimates of this model we calculate correction factors, which then enter the main model to control for unobservables associated with both BOI (or COI) and online presence. As instruments in the selection model, we use (i) a variable that proxies for the (shipping) costs of home-delivery, namely whether the brand is sold nationally in China or only regionally[24], and (ii) two variables that reflect the costs/difficulties of offline channel presence, namely category rotation and manufacturer power (see Table 4.1 for more information on the operationalization of these variables: category rotation measures the rate at which the packs of the SKUs offered in a category are renewed, while manufacturer power measures in how many categories the manufacturer is present). More details on the setup of the selection model are given in Appendix 4.B.

Second, both for COI and BOI, the marketing drivers may be endogenous. This may be due to reversed causality within brands and categories over time (e.g., brands might set their marketing mix instruments depending on how well they performed online vs. offline in the same period). Moreover, there may be 'cross-sectional endogeneity': unobserved brand or category characteristics driving both their marketing mix and BOI/COI. Finally, unobserved temporal factors may influence both the marketing mix drivers and online vs. offline outcome metrics. When unaccounted for, these phenomena may bias the estimates in our BOI and COI models. For lack of good instruments[25], we accommodate this potential endogeneity by using time fixed effects, and adding Gaussian copula-based control variables for the marketing mix drivers in equations (4.6) and (4.7) (see, e.g., Datta, Ailawadi, and van Heerde 2017)[26][27].

---

[24] Another variable that could account for (shipping) costs of setting up home-delivery is 'need for refrigeration'. However, only 5 out of 62 categories would need to be transported refrigerated or frozen (for an overview of categories: see Appendix 4.D), so we decided not to add this variable in the selection model.
[25] Because of autocorrelation, lagged values of these variables do not qualify as instruments. Using the brands' marketing mix levels in other countries is also a problem because not all brands are available there (and data for comparable countries are not available for some variables (like advertising)).
[26] Gaussian copulas only partially safeguard against cross-sectional endogeneity. The best way to control for that would be with category and brand fixed effects but that that would not be a good option as then all degrees of freedom would be absorbed.
[27] The Gaussian copula for marketing mix variable K_{jt} (K_{ct}) of brand j in year t (category c in year t), is defined as $copula_{kjt} = \Phi^{-1}(H(K_{jt}))$ ($copula_{kct} = \Phi^{-1}(H(K_{ct}))$), where Φ^{-1} is the inverse distribution function of the standard

Data

Data sources and sample selection

We obtained our data through Kantar Worldpanel, Kantar Media, and GfK. The purchase data come from a Chinese urban household panel (n=40,000) that tracked the panelists' purchases made through the online and offline channel in 62 CPG categories between 2011 and 2015 (all categories were sold online in these years). For every category, we select brands that belong to the top 10 in at least one of the five years, dropping 13 brands with 'holes' in their time series (e.g., for which we observe sales in 2011-2012 and 2014-2015, but not in 2013). This leaves us with 617 brands in 62 categories[28]. Across the years, 32 brands in our set entered their category, while 13 left. We use the first year of a brand's data as initialization period, the remaining years belong to the estimation sample. For each brand and category, we obtained monthly advertising spending data at the brand level as well as the total category level.

To estimate our BOI and COI models, we retain brands that meet two criteria. First, to avoid problems due to data sparseness, we select brands with an overall (i.e., offline and online combined) volume share within the category of at least 1% in the estimation sample. Second, the brand needs to have both offline and online presence in the estimation sample: we retain brands sold via both channels for at least two consecutive years.[29] This leaves us with 448 brands in 60 categories. The majority of brands is present for all four years of the estimation sample (for only 35 brands, we have less than four years of data)[30].

---

normal, and H(·) is the empirical cumulative distribution function of K_j (K_c). The Gaussian copula method requires that the endogenous regressors are not normally distributed. Shapiro-Wilk tests at $p < .10$ formally confirm this for all cases.

[28] Among these brands, no private labels were present.

[29] In addition, one brand's BOI in 2012 is unusually high while in 2015 its offline sales are reported to be zero, so we decided to drop this brand from the analysis.

[30] For 6 small brands (present in 6 categories, with an average overall market share of 2%), we observe 'missings' in the online sales. For example, we do not observe sales via the online channel in 2012, though we do observe online sales in 2011, and from 2013 onwards. We assume that the brand was still offered online in that year, but did not sell anything online. Therefore, BOI is set to zero in these cases, while drivers related to

Next to the purchase and marketing mix data, 45 categories and 154 brands in 43 categories were part of a consumer survey administered by GfK in 2014 to 2,764 urban Chinese consumers. Four risk reduction items, as well as the trust and fun constructs were part of the survey (and are available for those brands and categories only). On average, 92 respondents rated each category and brand. For an overview of the (survey) categories and number of selected (survey) brands per category, see Appendix 4.D. Finally, we surveyed experts about characteristics of all 62 categories, namely local embeddedness and perishability. We use these consumer and expert survey measures, averaged across respondents/experts, to quantify the corresponding drivers of COI and BOI.

Measurement

Table 4.1 provides details on the variable operationalizations. Panels A and B describe the variables of our main (BOI and COI) models, while Panel C describes the variables used in the selection model.

To calculate *BOI*, we use online and offline market share based on volume sales (e.g., milliliters, grams). Because some brands having zero online sales in few years (and thus BOI equal to zero), we add the value one and multiply with 100 before log-transforming it. For the *online to offline price ratio*, we use the brand's average channel prices per volume unit. *Price position* measures the price per volume unit of a typical pack of the brand relative to the price per volume unit of a typical pack of the category. The *advertising* variable is operationalized as Adstock (log-transformed due to skewness, after adding a small number to accommodate cases with zero advertising). Specifically, it is a weighted average of previous Adstock and current Ad spending (on all media), with weights equal to λ and $(1 - \lambda)$, respectively, where ad spending is converted into real prices using China's consumer price index. *Online (offline)*

the online channel (i.e., online unavailability and ratio online/offline price) are imputed with the average value of the previous and next year.

availability is calculated as the percentage of websites (offline retailers) that carry the brand, weighted by the website's (retailer's) market share. In addition, offline availability is corrected for brand's regional character (i.e., calculated in regions where the brand is physically marketed). *Large packs* measures the percentage of the number of stock keeping units of the brand that have a pack size larger than the category average. Whether the brand's owner is Chinese (domestic) or not (foreign) is coded by consulting the brands' websites. *Brand trust* and *Fun brand* were part of the consumer survey. *Brand trust* is the average of two items ('Brand b is a brand I trust', and 'Brand b delivers what it promises'), while *Fun brand* is measured with one item ('Brand b is a fun brand').

To calculate COI, online and offline market share are based on volume sales expressed in 'equivalent monetary value' (cfr. Ma et al. 2011) to ensure comparability across categories. Specifically, we multiply the volume sales (e.g., milliliters for shampoo, grams for potato crisps) with the average price per volume unit in the category across 2011-2015. For the *online to offline price ratio*, we use the category's average channel prices per volume unit. *Category expensiveness* is the average amount paid for the typical quantity selected when the category is bought (log-transformed due to skewness); the weight of that typical quantity measures *heaviness*; and the volume measures *bulkiness*[31]. Similar to brand advertising, *category advertising* is operationalized as Adstock (log-transformed due to skewness, after adding a small number to accommodate cases with zero advertising): a weighted average of previous Adstock and current Ad spending across all brands in the category on all media, converted to real prices. *Assortment size* measures the unique number of stock keeping units offered in the category, and *purchase frequency* is calculated as the average number of purchase events by households who bought the category. *Perishable* vs.

[31] Note that heaviness and bulkiness are two distinct measures. For example, milk might be quite heavy though not bulky. Toilet tissue on the other hand generally does not weigh much but can be pretty bulky.

non-perishable is coded by 7 (Dutch) judges; *local embeddedness* in China is coded by 5 (native Chinese) judges (Cronbach's alpha .94); *risk reduction* (available for 45 out of the 60 categories in the survey) is the average across the four risk items in the consumer survey (Cronbach's alpha .87).

For the selection model, the dependent variable equals one (zero) for the years in which the brand was (not or not yet) present online. As discussed above, as drivers of online presence, we include variables used in the main models (lagged one period, and excluding the survey constructs that are available for only a subset of brands and categories)[32], next to a set of instruments. Table 4.1, Panel C, explains the operationalization of these instruments. Whether the brand is *regional* (sold in one region) or not (sold nationwide) is coded based on the regional shares of the brand's volume sales. *Category rotation* is calculated by dividing the number of units sold, by the category's unique number of stock keeping units (lagged one year), while *manufacturer power* measures the number of categories in which the brand's manufacturer is active. Due to skewness, both category rotation as well as manufacturer power are log-transformed.

Results

Model free evidence

Figure 4.2, Panel A displays the histogram of the brands' BOI (averaged across years), while Table 4.2, Panel A reports summary statistics across brands, for BOI as well as its drivers. As Figure 4.2 shows, the BOI distribution is highly skewed, with a mean equal to 1.30, and a median of .65. At the same time, the figure shows large variation in BOI across brands, as is also reflected in the standard deviation (SD: 2.13).

[32] In the selection model, advertising is only taken into account at the brand level, not at the category level.

Table 4.1: Operationalization brand and category variables

VARIABLE	SOURCE	OPERATIONALIZATION
PANEL A: (DRIVERS) BOI		
BOI (BOI_{bt})	Kantar Worldpanel	[Total volume sales of brand b in year t in online channel relative to category total volume sales in year t in online channel] divided by [Total volume sales of brand b in year t in offline channel relative to category total volume sales in year t in offline channel].
Large packs (la_{bt})	Kantar Worldpanel	Total number of 'large' stock keeping units that brand b sold in year t relative to total number of stock keeping units that brand b sold in year t, where 'large' means the pack size in volume (e.g., milliliters) of brand b's stock keeping unit sold in year t is larger than the average volume per pack sold in brand b's category in year t.
Online availability (av_{bt}^{on})	Kantar Worldpanel	Weighted average of indicators of availability (0 vs. 1) for brand b in year t across all websites, weighted by the websites' market shares in year t (Sotgiu and Gielens 2015).
Offline availability (av_{bt}^{off})[b]	Kantar Worldpanel	Weighted average of indicators of availability (0 vs. 1) for brand b in year t across all offline retailers, weighted by the retailers' market shares in year t and brand b's regional presence, where weight regional presence equals .125 (.25; 1) if brand is sold in 1 (2; all 8) region(s) of China (Sotgiu and Gielens 2015).
Price position (pp_{bt})[a]	Kantar Worldpanel	[Average price 'typical' pack (in ¥) per volume unit of brand b in year t] divided by [Average price 'typical' pack (in ¥) per volume unit in brand b's category in year t].
Ratio online price to offline price (rp_{bt})	Kantar Worldpanel	[Average price (in ¥) per volume unit, of brand b in year t in online channel] divided by [Average price (in ¥) per volume unit, of brand b in year t in offline channel].
Advertising (ad_{bt})	Kantar Media	Adstock of brand b in year t, calculated as $Adstock_{jt} = (1-\lambda)*Advertising_{jt} + \lambda*Adstock_{jt-1}$ (Datta, Ailawadi, and van Heerde 2017), where advertising spend on all media by the brand ($Advertising_{jt}$) in ¥ is converted into real prices using China's consumer price index (source: National Bureau of Statistics China). Following George, Mercer, and Wilson (1996), λ is set to .8.
Brand trust (tr_b)	GfK consumer survey	Average of 2 items that were rated from 1=very strongly disagree to 7=very strongly agree (Chaudhuri and Holbrook 2001; Erdem and Swait 2004):

	(subset of 154 brands in 43 categories only)	Brand b… - Is a brand I trust - Delivers what it promises (Cronbach's alpha: .81).
Fun brand (fu_b)	GfK consumer survey (subset of 154 brands in 43 categories only)	Item that was rated from 1=very strongly disagree to 7=very strongly agree (Voss, Spangenberg, and Grohmann 2003): Brand b is a fund brand.
Brand ownership (fb_b)	Brand's websites	Coded as 1=foreign (i.e., brand owner is not Chinese), -1=domestic (i.e., brand owner is Chinese).
PANEL B: (DRIVERS) COI		
COI (COI_{ct})[c]	Kantar Worldpanel	[Total volume sales (expressed in monetary values) of category c in year t in online channel relative to grocery total volume sales (expressed in monetary values) in year t in online channel] divided by [Total volume sales (expressed in monetary values) of category c in year t in offline channel relative to grocery total volume sales (expressed in monetary values) in year t in offline channel].
Category expensiveness (cx_c)[d]	Kantar Worldpanel	Average price paid for a 'typical' quantity (in ¥) in category c (Lourenço, Gijsbrechts, and Paap 2015), where prices are converted into real spending using China's category-specific consumer price index (source: National Bureau of Statistics China).
Risk reduction function of brands (rr_c)	GfK consumer survey (subset of 45 categories only)	Average of 4 items that were rated from 1=very strongly disagree to 7=very strongly agree (Fischer, Völckner, and Sattler 2010): When I make a purchase in category c… - I purchase mainly brand name products because that reduces the risk of aggravation later - I purchase brand name products because I know that I get good quality - I choose brand name products to avoid disappointment - I purchase brand name products because I know that the performance promised is worth its money (Cronbach's alpha: .87).
Adstock (ad_{ct})	Kantar Media	Adstock of category c in year t, calculated as $Adstock_{ct} = (1-\lambda)*Advertising_{ct} + \lambda*Adstock_{ct-1}$ (Datta, Ailawadi, and van Heerde 2017), where advertising spend

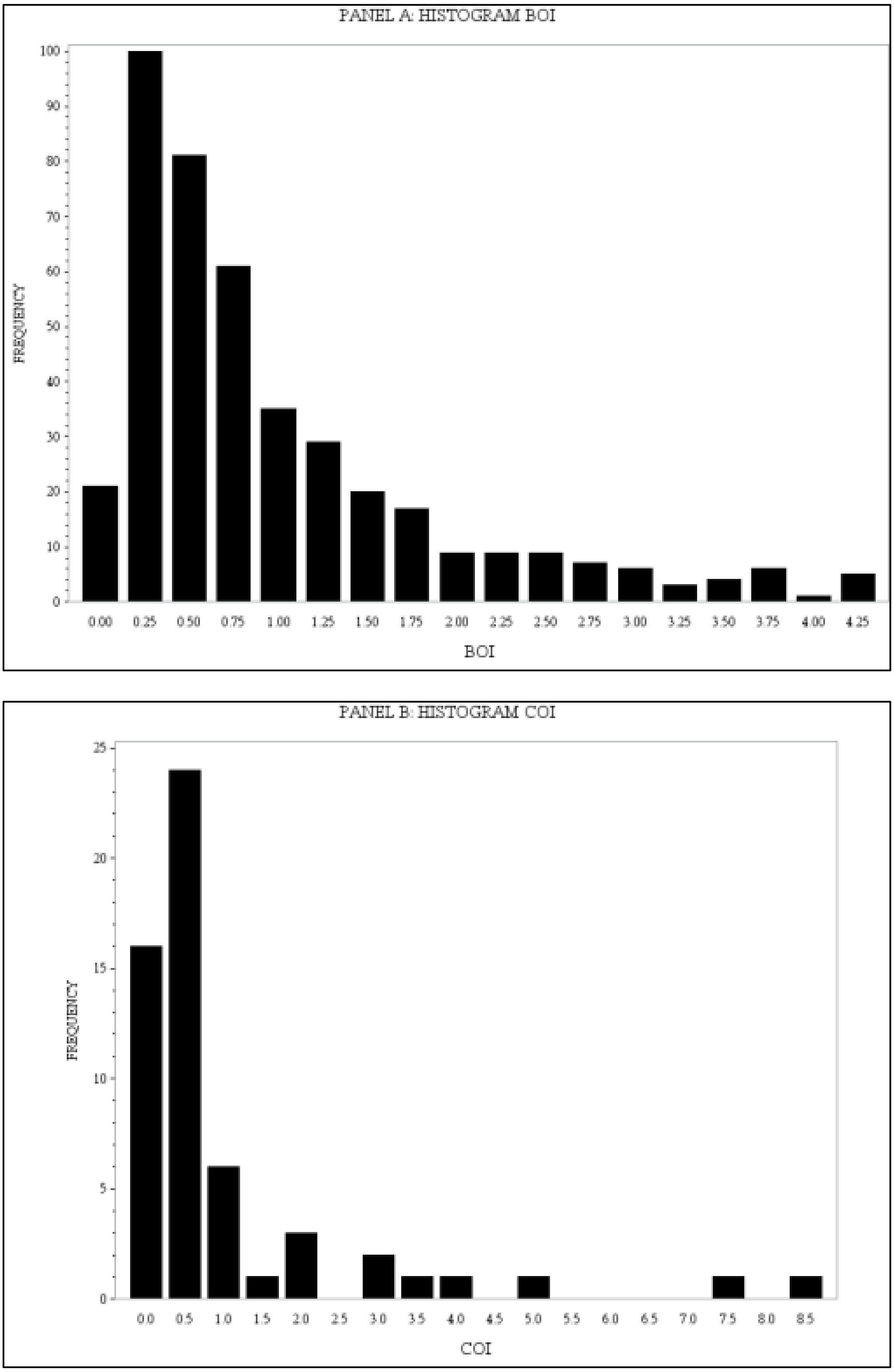
		on all media by the category ($Advertising_{ct}$) in ¥ is converted into real prices using China's consumer price index (source: National Bureau of Statistics China). Following George, Mercer, and Wilson (1996), λ is set to .8.
Assortment size (as_{ct})	Kantar Worldpanel	Unique number of stock keeping units offered in category c in year t.
Category bulkiness (bu_c)[d]	Kantar Worldpanel	Bulkiness of a 'typical' quantity (in cubic inches) in category c.
Category heaviness (he_c)[d]	Kantar Worldpanel	Weight of a 'typical' quantity (in pounds) in category c.
Purchase frequency (fr_c)	Kantar Worldpanel	Average yearly number of purchase events made by households who purchased in category c.
Perishable (pe_c)	Expert survey	Dummy variable equal to 1 if majority of judges coded category c as perishable, -1 otherwise.
Local embeddedness (le_c)	Expert survey	Average of 3 items that were rated from 1=very strongly disagree to 7=very strongly agree: - This category does not originate from China (reversed before calculation) - This category is typically Chinese - This category has been around in China for a long time (Cronbach's alpha: .94).
PANEL C: (INSTRUMENTS) ONLINE PRESENCE		
Online presence (pr^{on}_{bt})	Kantar Worldpanel	Coded as 1=brand b was present online in year t, 0=brand b was not (yet) present online in year t.
Regional brand (rb_b)[b]	Kantar Worldpanel	Coded as 1=regional (i.e., regional share of total brand b's volume sales < 2% in 6 or 7 regions), -1=not regional (i.e., regional share of total brand b's volume sales ≥ 2% in 6 or 7 regions).
Category rotation (ro_{ct})	Kantar Worldpanel	Number of units sold in category c in year t relative to unique number of stock keeping units offered in category c in year t.
Manufacturer power (po_b)	Kantar Worldpanel	Number of categories in which brand b's manufacturer is active.

[a] A 'typical' pack in the category means the total volume (e.g., milliliters) bought divided by total units bought in the category during a shopping trip, averaged across all observed shopping trips.
[b] In total, Kantar Worldpanel distinguishes 8 regions in China.
[c] Because sales are expressed per volume unit, and volume units differ across categories (e.g., milliliters for shampoo, grams for potato crisps), we express category volume sales in monetary values to ensure comparability across different categories (see also Ma et al. 2011). To obtain volume sales in monetary value, we multiply the volume sales with the average price per volume of category c across 2011-2015 (note that for brands, we do not express volume sales in equivalent monetary values as it would not change our BOI measure).
[d] A 'typical' quantity in the category means the total volume (e.g., milliliters) bought in the category during a shopping trip, averaged across all observed shopping trips.

The possible drivers of this variation also differ widely across brands. For example, we observe large variation in brands' availability offline and online (SDs of .28 and .22). Our selection of brands is generally positioned in the medium to somewhat higher price tiers (average price position is 1.23), though the first and upper quartile show that both brands from the lower price tier as well as more premium brands are represented in our data. In addition, on average, brands are slightly more expensive online than offline.

The histogram of the categories' COI (averaged across years) is depicted in Figure 4.2, Panel B; while descriptives of COI and its drivers can be found in Panel B of Table 4.2. Again, while the average COI (1.51) exceeds unity (meaning that on average, our categories perform somewhat better online vs. offline), the COI distribution is highly skewed, with a median value that is much lower (.43). Like for BOI, the COI standard deviation (2.66) and the lower and upper quartiles (.22 vs. 1.06) reveal that categories largely differ in relative online to offline performance. The category characteristics show quite some variation as well. Moreover, as can be seen from the correlation tables (Table 4.2, Panels C and D), there is relatively little overlap among the (brand- and category-) drivers – making them suitable for our regression analyses. As for online presence (the dependent variable in the selection model), the majority of brands were present online in 2012 already, or entered in one of the three years after. Still, 36 brands in 16 categories were not present online by 2015. Descriptives and correlations of online presence and its drivers can be found in Appendix 4.C.

Figure 4.2: BOI and COI histograms[a]



[a] The histograms are based on the brands' average BOI and the categories' average COI, with averages calculated across the full data period. For presentation purposes, we drop the top 5% observations for these plots.

Table 4.2: Data descriptives

PANEL A: DESCRIPTIVE STATISTICS (DRIVERS) BOI					
VARIABLE	NUMBER OF OBSERVATIONS[a]	MEAN	SD	LOWER QUARTILE	UPPER QUARTILE
BOI (BOI_{bt})	1,746 (35% of brand-year combinations has BOI>1)	1.30	2.13	.30	1.39
Large packs (la_{bt})	1,746	0.44	.29	.20	.63
Online availability (av_{bt}^{on})	1,746	.82	.22	.77	.97
Offline availability (av_{bt}^{off})	1,746	.78	.28	.75	.97
Price position (pp_{bt})	1,746	1.22	.70	.80	1.48
Ratio online to offline price (rp_{bt})	1,746	1.03	.35	.89	1.10
Adstock (ad_{bt})[b]	1,746	195.85	531.76	0	107.55
Brand trust (tr_b)	611	5.37	.25	5.20	5.56
Fun brand (fu_b)	611	5.25	.27	5.04	5.45
Brand ownership (fb_b)	1,746 (39.96% of 448 brands are foreign)	n.a.[c]	n.a.[c]	n.a.[c]	n.a.[c]
PANEL B: DESCRIPTIVE STATISTICS (DRIVERS) COI					
COI (COI_{ct})	240 (26% of category-year combinations has COI>1)	1.51	2.66	.22	1.06
Category expensiveness (cx_c)	240	24.57	33.26	10.30	22.94
Risk reduction function in category (rr_c)	180	5.53	.21	5.41	5.65
Adstock (ad_{ct})[b]	240	35232.31	91563.06	179.94	25683.10
Assortment size (as_{ct})[d]	240	31.95	37.31	3.03	47.27
Bulkiness (bu_c)	240	233.56	345.97	26.30	297.43
Heaviness (he_c)	240	2.70	3.88	.60	3.55
Purchase frequency (fr_c)	240	4.02	2.32	2.14	5.33
Perishable (pe_c)	240 (33% of 62 categories are perishable)	n.a.[c]	n.a.[c]	n.a.[c]	n.a.[c]
Local embeddedness (le_c)	240	3.77	1.34	2.60	4.83

PANEL C: CORRELATIONS BOI AND ITS DRIVERS (NUMBER OF OBSERVATIONS[a])										
	$\ln(BOI_{bt})$	la_{bt}	av^{on}_{bt}	av^{off}_{bt}	pp_{bt}	$\ln(rp_{bt})$	$\ln(ad_{bt})$	tr_b	fu_b	fb_b
$\ln(BOI_{bt})$[e]	1.00 (1,746)									
la_{bt}	.15 (1,746)	1.00 (1,746)								
av^{on}_{bt}	-.32 (1,746)	-.01 (1,746)	1.00 (1,746)							
av^{off}_{bt}	-.0003 (1,746)	.12 (1,746)	.50 (1,746)	1.00 (1,746)						
pp_{bt}	.27 (1,746)	.07 (1,746)	-.12 (1,746)	-.06 (1,746)	1.00 (1,746)					
$\ln(rp_{bt})$	-.16 (1,746)	.06 (1,746)	-.05 (1,746)	-.07 (1,746)	-.06 (1,746)	1.00 (1,746)				
$\ln(ad_{bt})$[f]	-.06 (1,746)	-.07 (1,746)	-.30 (1,746)	-.34 (1,746)	.06 (1,746)	.12 (1,746)	1.00 (1,746)			
tr_b	.14 (611)	.04 (611)	.04 (611)	.09 (611)	.03 (611)	-.05 (611)	-.10 (611)	1.00 (611)		
fu_b	.08 (611)	.01 (611)	.08 (611)	.05 (611)	.04 (611)	-.09 (611)	-.08 (611)	.85 (611)	1.00 (611)	
fb_b	.25 (1,746)	.01 (1,746)	-.22 (1,746)	-.16 (1,746)	.13 (1,746)	-.07 (1,746)	.15 (1,746)	.20 (611)	.23 (611)	1.00 (1,746)
PANEL D: CORRELATIONS COI AND ITS DRIVERS (NUMBER OF OBSERVATIONS[a])										
	$\ln(COI_{ct})$[e]	$\ln(cx_c)$	rr_c	$\ln(ad_{ct})$	as_{ct}	bu_c	he_c	fr_c	pe_c	le_c
$\ln(COI_{ct})$[e]	1.00 (240)									
$\ln(cx_c)$	0.59 (240)	1.00 (240)								

rr_c	0.41 (180)	0.28 (180)	1.00 (180)							
ln (ad_{ct})[f]	-0.19 (240)	0.02 (240)	0.01 (180)	1.00 (240)						
as_{ct}	-0.10 (240)	0.25 (240)	-0.16 (180)	0.35 (240)	1.00 (240)					
bu_c	0.02 (240)	0.26 (240)	-0.05 (180)	0.13 (240)	0.23 (240)	1.00 (240)				
he_c	-0.13 (240)	0.28 (240)	-0.15 (180)	0.12 (240)	0.21 (240)	0.60 (240)	1.00 (240)			
fr_c	-0.33 (240)	0.00 (240)	-0.14 (180)	0.35 (240)	0.38 (240)	0.06 (240)	0.18 (240)	1.00 (240)		
pe_c	-0.12 (240)	-0.09 (240)	0.07 (180)	-0.04 (240)	-0.19 (240)	-0.18 (240)	-0.07 (240)	0.26 (240)	1.00 (240)	
le_c	-0.49 (240)	-0.03 (240)	-0.38 (180)	0.23 (240)	0.47 (240)	0.08 (240)	0.35 (240)	0.43 (240)	0.09 (240)	1.00 (240)

[a] For brand-year (category-year) combinations, the number of observations equal to 1,746 (240) represents all 448 brands (60 categories) used in our main analyses; the number of observations equal to 611 (180) represents the 154 (45) survey brands (categories).
[b] In 100,000s.
[c] n.a. = not applicable.
[d] In 100s.
[e] Some brands were not sold online in certain years, so yearly BOI equals zero in these cases. Therefore, ln(BOI) represents the natural logarithm of [BOI*100+1]. For reasons of consistency, the same is done with yearly COI.
[f] Log Adstock represents the log-transform of Adstock (the log of 1.00E-11 is taken for the 668 out of 1,746 brand-year combinations and the 14 out of 240 category-year combinations in our sample with Adstock equal to zero).

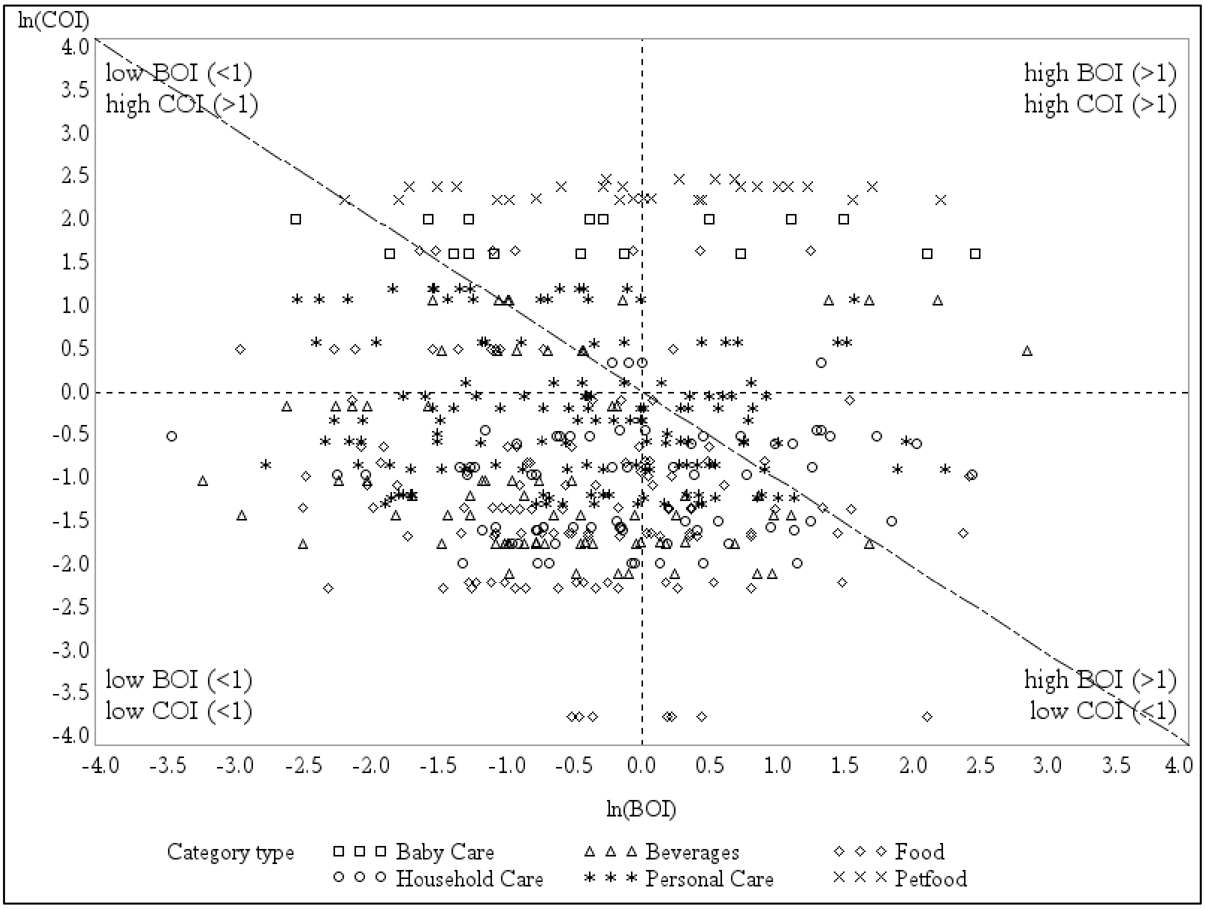
How does brand performance change as the overall share of CPG sold online goes up?

Figure 4.3 plots the logarithm of BOI against the logarithm of COI for all brands (calculated across all observation years). As the graph shows, brands are quite spread in the [ln(BOI)*ln(COI)] space. The diagonal line indicates where the product of BOI and COI equals unity: if the rise of online would not affect total grocery outlay (value expansion nor contraction, g=1), brands positioned above the line see their overall sales increase as the share of CPG sold online goes up, whereas brands below the line face a decline in total sales.

Comparing the average drivers for brands below vs. above the line by means of t-tests provides some first model-free insights. Compared to brands with sales reduction (i.e., that are located below the diagonal line), brands with sales expansion (i.e., that are located above the diagonal line) generally sell larger packs ($M_{expansion}$=.49 vs. $M_{reduction}$=.42, p=.03), do not have significantly larger online availability ($M_{expansion}$=.83 vs. $M_{reduction}$=.81, p=.27), yet have lower offline availability ($M_{expansion}$=.69 vs. $M_{reduction}$=.81, p<.0001), are positioned in the higher price tier ($M_{expansion}$=1.39 vs. $M_{reduction}$=1.16, p=.0002), have somewhat lower online to offline price ratios ($M_{expansion}$=1.00 vs. $M_{reduction}$=1.04, p=.09), but do not significantly advertise more ($M_{expansion}$=1,928 vs. $M_{reduction}$=1,906, p=.97). For the subset of survey brands, the t-tests reveal that compared to brands with sales reduction, brands with sales expansion are generally highly trusted brands ($M_{expansion}$=5.48 vs. $M_{reduction}$=5.30, p < .0001), or fun brands ($M_{expansion}$=5.35 vs. $M_{reduction}$=5.18, p<.0001)[33]. Although these t-tests provide some first insights, they do not tell us anything about the combined influence of the drivers. Moreover, as shown in our decomposition, the extent to which sales expansion will occur is due to two separate factors, with different implications, viz. BOI and COI.

[33] For the non-survey measures, the means are calculated across brands' averages of the full data period (expressed in 100,000s for advertising).

Figure 4.3: Brand distribution in BOI-COI space (n=448 brands)[a]



[a] Note: BOI and COI are calculated across the full data period.

As such, it is instructive to consider the brands' position in Figure 4.3, along the two separate dimensions. Brands in the upper-right quadrant are expected to flourish as online gains way: not only does the category in which they operate fare particularly well online (COI >1), their relative position within the category is expected to improve as well (BOI >1). Examples are Friso in infant milk powder, Ferrero Rocher in chocolates and Bawang in hair coloring products. The opposite holds for brands in the lower-left quadrant, which are expected to suffer from double jeopardy: they operate in a category that agrees less with the online channel (COI <1) and, within that category, do worse online than competing brands (BOI <1). This would, for instance, hold for Lux in toilet soap, Capico in potato chips, and Yili in yoghurt. Brands in the upper-left quadrant, though they do worse online than their immediate competitors (BOI <1), may still maintain or increase sales by riding on the waves of the online success of the category (COI >1). Exemplars would be Friskies in dry cat food, Nivea in facial cleaning products, and Pampers in Nappies and Diapers. Conversely, while brands in the lower-right quadrant are bound to improve their market share (BOI >1), their sales may be dampened because they operate in a category that loses share online. This holds, for example, for Arawana in rice, Tide in laundry detergent, and Heinz in ketchups. The key question, then, is what characterizes categories that do better (or worse) online than offline and, more importantly, which actions can help brand managers secure their relative position as the online channel gains importance. Our regression results shed light on this.

What drives overall brand performance?

Before zooming in on the results of our BOI and COI models, we briefly discuss the estimates of our selection model, which, although not the focus of this paper, are interesting in their own right (See Panel C of the table in Appendix 4.C for the full set of results). The probability to be present online is higher for more expensive, regionally sold brands that also have a stronger offline presence, advertise more, and/or are produced by a foreign manufacturer that operates in multiple categories. In addition, brands in more expensive, more frequently purchased, more perishable and less locally-embedded categories with large assortments are more likely to enter the online channel. Controlling for these selection effects will result in cleaner estimates of our BOI and COI models, which are discussed next.

BOI model. The estimation results of our BOI models can be found in Table 4.3, Panel A. As expected, foreign brands perform relatively better online vs. offline than domestic brands (β_9=.22, p<.01). In addition, brands that are more widely available online and charge relatively lower prices in that channel, will improve their position within the category as the online channel grows (β_2=1.39, p<.01; β_5=-.61, p<.05). Interestingly, we find no significant impact of offline availability on BOI (β_3=.03, p>.10), possibly because the positive billboard effect and the negative substitution effect cancel out. Brands that sell relatively more large packs generally do not enjoy higher market shares online than offline (β_1=-.46, *p*<.05). This might be caused by consumers' attempts to minimize shipping costs, something we will turn to when discussing the results of the COI model. The estimate of (high) price positioning does not reach significance (β_4=-.06, *p*>.10), refuting the premise that expensive brands fare relatively better in the digital channel. Finally, although the selection model revealed that heavily advertised brands are more likely to be present online, they do not seem to enjoy higher BOIs (β_6=.000003, *p*>.10).

Moving to the subset of brands for which survey data are available (typically the somewhat larger brands), we see that the previous pattern of effects remains largely similar (except for the online to offline price ratio and large packs, the coefficients of which are no longer significant). In addition, we find significant effects for the two survey measures: while more trusted brands generally have higher online vs. offline market shares (β_7=-.69, *p*<.01), the opposite holds for fun brands (β_8=-.61, *p*<.05).

COI model. Panel B of Table 4.3 displays the estimates of the COI models (for the full set and the subset of survey categories). As expected, more expensive, less perishable and less frequently purchased categories capture a larger CPG share online than offline (α_1=.69, *p*<.01; α_8=-.22, *p*<.01; α_7=-.11, *p*<.01). The same holds for categories with large assortments (that can be easily searched-through online; α_4=.0002, *p*<.01), that are less locally embedded (α_9=-.17, *p*<.01). Opposite to our expectations, categories that consist of less heavy products perform relatively better online than offline (α_6=-.07, *p*<.01). This may have to do with the structure of shipping costs, which decrease in the total amount spent and increase with the weight of the shopping basket. Advertising nor bulkiness of a category has a significant influence (α_3=.005, *p*>.10; α_5=-.00007, *p*>.10).

Looking at the regression outcomes for the subset, where risk reduction is added as a driver, we find the effects to be very robust – the significance, sign and magnitude remaining the same. Furthermore, categories in which brand names serve as important risk-reduction cues, take a larger share of groceries sold online vs. offline (α_2=.59, p<.05); in line with our proposition.

Discussion

Online grocery shopping is ready for takeoff, and this trend will unavoidably affect total brand performance. In this paper, we consider CPG brands' change in total (online plus offline) sales as a function of the fraction of groceries sold online. We show that, apart from total CPG expansion or contraction effects, this sales change critically depends on two indices: (i) the brand's online index (BOI), which reflects the brand's relative market position within the category, in the online vs. the offline channel and (ii) the category's online index (COI), which captures the category's share of total grocery sales online vs. offline. Brands with high BOI and COI will experience a double whammy as the online channel grows, being situated in a category that lends itself well to online buying, and doing better than their immediate competitors within that category. For brands in high COI categories yet with low BOI, our decomposition acts as a warning signal: though these brands appear to maintain high sales levels as the online channel becomes more popular, they simply ride on the category waves, yet lose position relative to other players in the category. Conversely, for brands in low COI categories, the sales erosion is likely attributable to factors outside of the brand managers' control. Importantly, our decomposition model, which allows researchers to disentangle the effects of online on sales, is applicable in any market, not only China.

Table 4.3: Estimation results BOI and COI models

PANEL A: ESTIMATION RESULTS BOI ($\ln(BOI_{bt})$)[a]				
DRIVERS	FULL SAMPLE		SURVEY SAMPLE	
	ESTIMATE	P-VALUE[b]	ESTIMATE	P-VALUE[b]
Intercept	4.12	<.0001	3.63	<.0001
Brand factors				
Large packs (la_{bt})	-.46	.0201	.69	.2274
Online availability (av_{bt}^{on})	1.39	<.0001	2.67	<.0001
Offline availability (av_{bt}^{off})	.03	.8910	-.30	.2533
Price position (pp_{bt})	-.06	.4622	-.39	.0053
Ratio online to offline price $\ln(rp_{bt})$	-.61	.0147	-.40	.2780
Adstock $\ln(ad_{bt})$	.000003	.9987	-.0005	.8243
Brand trust (tr_b)			.69	.0095
Fun brand (fu_b)			-.61	.0196
Brand ownership (fb_b)	.22	<.0001	.19	<.0001
Instruments				
Correction factor	-.56	<.0001	-1.14	.0039
Copula Large packs	.22	<.0001	.02	.8949
Copula Online availability	.38	<.0001	.27	<.0001
Copula Offline availability	-.21	.0002	-.21	.0005
Copula Price position	.31	<.0001	.34	.0003
Copula Ratio online to offline price	.008	.9013	.0006	.9946
Copula Adstock	-.26	<.0001	-.09	.1931
Covariates[c]				
Category type: Baby care	-.15	.0224	-.23	.0029
Category type: Beverages	-.17	<.0001	-.06	.2555
Category type: Household care	.06	.1229	-.12	.0669
Category type: Personal care	-.13	<.0001	-.28	<.0001
Category type: Pet food	-.02	.6964	-.33	.0097
Year: 2013	.15	<.0001	.13	.0067
Year: 2014	.17	<.0001	.17	.0007

Year: 2015	.17	<.0001	.20	.0002
Number of observations				
Total	1,746		611	
Number of brands	448		154	
Number of categories	60		43	
Model fit				
R^2	.35		.38	
Adjusted R^2	.34		.36	
PANEL B: ESTIMATION RESULTS COI ($\ln(COI_{ct})$)[a]				
DRIVERS	FULL SAMPLE		SURVEY SAMPLE	
	ESTIMATE	P-VALUE[b]	ESTIMATE	P-VALUE[b]
Intercept	6.23	<.0001	6.25	<.0001
Category factors				
Category expensiveness $\ln(cx_c)$	.69	<.0001	.60	<.0001
Risk reduction function of brands (rr_c)			.59	.0139
Adstock $\ln(ad_{ct})$	.005	.3449	.009	.2505
Assortment size (as_{ct})	.0002	<.0001	.0002	<.0001
Bulkiness (bu_c)	-.00007	.7014	-.00005	.7991
Heaviness (he_c)	-.07	<.0001	-.06	.0004
Purchase frequency (fr_c)[e]	-.11	<.0001	-.08	.0013
Perishable (pe_c)	-.22	<.0001	-.333	<.0001
Local embeddedness (le_c)	-.17	<.0001	-.14	.0038
Instruments				
Copula Adstock	-.20	.0055	-.15	.0576
Copula Assortment size	-.53	<.0001	-.51	<.0001
Covariates[c]				
Category type: Baby care	.91	<.0001	.94	<.0001
Category type: Beverages	.28	<.0001	.28	<.0001
Category type: Household care	-.19	.0043	-.12	.1321
Category type: Personal care	.002	.9810	.03	.6532
Category type: Pet food	1.14	<.0001	1.31	<.0001
Year: 2013	.15	.0036	.14	.0188

Year: 2014	.21	<.0001	.19	.0012
Year: 2015	.29	<.0001	.30	<.0001
Number of observations				
Total	240		180	
Number of categories	60		45	
Model fit				
R^2	.84		.84	
Adjusted R^2	.83		.83	

[a] Mean-centered estimates are reported, dependent variables equal the logarithm of [BOI*100+1] and the logarithm of [COI*100+1].
[b] Two-sided p-value are reported.
[c] Covariates are effect-coded (e.g., ‘Category type: Baby care’ equals 1 if baby care, and -1 otherwise); base categories are category type=food and year=2012.

Calculating the BOI and COI for a set of 448 brands present in 60 categories in the Chinese CPG market, we observe large variation in these metrics across categories and brands. In a next step, we identify brand- and category drivers that may underlie this variation. We then empirically assess the impact of these drivers, after controlling for possible selection bias in our brand (and category) set. We find that, overall, these drivers explain a large portion of the variability in our key metrics: up to almost 40% of the variation in BOI, and over 85% of the variation in COI[34].

How many brands and which brands benefit?

Across the 448 brands under study, the product of BOI and COI exceeds unity for 119 brands (i.e., 27%), whereas [BOI*COI] drops below unity for 329 brands (i.e., 73%). Two points are worth noting, though. First, much of the effect is due to the dominance (large COI) of specific category types: especially brands in baby care and pet food seem to benefit from the trend toward online grocery, while for the majority in food and beverages it is much harder to reap the benefits of this trend.[35] Second, industry reports generally agree that as a result of the growth in online, total consumption in China goes up (i.e., g in equation (4.4) is believed to be larger than one). In their report on the CPG industry, Bain & Company and Kantar Worldpanel (2015)[36] indicate that "new demand made up 60% of the value growth in e-commerce, the other 40% came from substitution of purchases shoppers previously made offline" (p.13). This would imply a value for g equal to 1.6. According to this same report, online share of CPG would amount to 3.3% in 2014 (which is halfway our estimation period).

[34] Variance explained of models with only year and category type fixed effects equal 1.9% and 56.5% respectively, while variance explained of models with solely year (category type) fixed effects equal .07% (1.8%) and 2.3% (54.2%).
[35] We also note that our total CPG comprises the 62 categories delivered by the data provider, which may not cover all CPG purchases.
[36] McKinsey (2013) examined 2011 data from 266 Chinese cities and reported that "a dollar of online consumption replaces roughly 60 cents of sales in offline stores and generates around 40 cents of incremental consumption." China Internet Watch (2015) reports a more conservative number, namely that "78% online shopping consumption in China are alternatives to the traditional consumption, and 22% are new demands stimulated by online shopping market in 2014". These reports however took into account all products and services, i.e., groceries, but also for example apparel, furniture, health care products, and mobile phones.

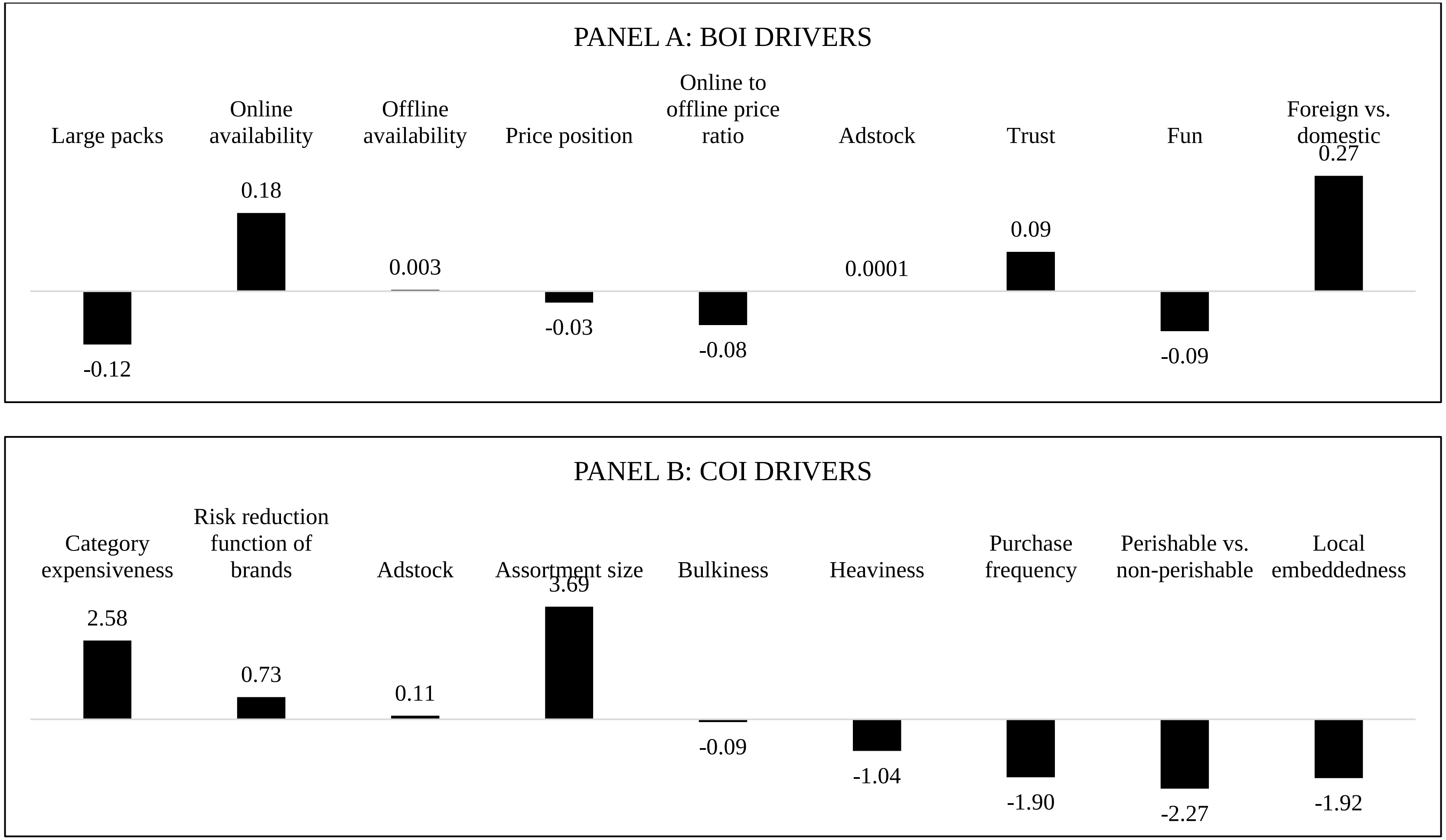
Plugging in these values in expression (4.4) (g=1.6 and $\left[\frac{S^{On}_{.,.,t}}{S^{T}_{.,.,t}}\right]$ = .033) reveals that [BOI*COI] would need to be larger than .42 for sales expansion to occur, which would hold true for about half (i.e., 46%) of the brands in our study.

Having documented the levels of BOI and COI, we also provide insights into what drives these indices (and to what extent). Building on the model estimates (the magnitude of which is not directly interpretable or comparable across drivers), Figure 4.4 presents the importance of the drivers of BOI and COI. We calculated these effect sizes as the difference between the predicted value of the index if the driver equals (i) its upper-quartile level, vs. (ii) its lower-quartile level (except for foreign vs. domestic and perishable vs. non-perishable, which are dummies and for which we consider values of one vs. zero).

Zooming in on Figure 4.4, Panel A, we find that brand ownership and online availability have the largest effects on BOI: being owned by a foreign manufacturer, and being widely available on the web, substantially enhances online vs. offline performance. Specifically, a foreign brand's BOI is generally .27 higher than that of a domestic brand, and moving from low to high online availability (i.e., from .77, lower quartile, to .97, upper quartile), increases BOI by .18. Offering small packs also helps to benefit from the online trend: brands that sell a percentage of large packs equal to the upper quartile, on average have BOIs that are .12 lower than brands at the lower quartile.

Though the online to offline price ratio matters, it is not a dominant factor: changing the brand's price from being 10% more expensive online vs. offline to being 11% cheaper increases BOI by only .08. Similar effect sizes are obtained for 'fun' and 'trusted' brands: moving from the lower to upper quartiles of these variables leads to an average difference in BOI of -.09 and .09 respectively.

Figure 4.4: Relative importance of drivers of BOI and COI[a]



[a] Numbers in Panel A (Panel B) show the difference between the predicted average values of BOI (COI) if the driver equals the upper quartile vs. the lower quartile (see last two columns of Panels A and B of Table 4.2 for the ranges). The effects of Trust, Fun, and Risk reduction are calculated based on the estimates for the reduced data set (including only the survey brands and categories).

As can be seen from Figure 4.4, Panel A, the remaining drivers are much less influential. A change in the brand's price position in the category from the lower quartile (.80) to the upper quartile (1.48) has an effect of only -.03 on a brand's online index. Similarly, the figure shows that changes in a brand's offline availability and advertising have only negligible effects on BOI.

As for COI, Figure 4.4, Panel B shows that category assortment size and expensiveness are the most important drivers of a category's online to offline performance. Categories with large assortment sizes (i.e., equal to the upper quartile) have COIs 3.69 above those of small-sized (lower quartile) categories. Moreover, categories where a typical quantity costs ¥10.30 (lower quartile) have COIs 2.58 below categories where a typical quantity costs ¥22.94 (upper quartile).

Whether the category is perishable or not, how locally embedded the category is, and how often the category is purchased, have similar-sized effects on COI: moving from the lower to the upper quartiles of local embeddedness or purchase frequency, or from non-perishable to perishable, leads to COIs that are about 2.00 lower on average.

The impact of heaviness and the risk reduction function of brands is somewhat smaller. The COI difference between light categories (about to -.60 pounds for a typical quantity) and heavy categories (3.55 pounds) amounts to -1.04, and that for categories where brands have a strong vs. weak risk-reduction function to .73. Finally, as Figure 4.4, Panel B shows, the effect sizes for advertising spending and bulkiness are the lowest (.11 and -.09, respectively), corroborating their insignificant effects on COI.

Managerial implications

How does brands' relative position in the category change as online grows? As we show, brand managers can use two simple (and easy-to-calculate) metrics, COI and BOI, to get a first indication of how their relative market position and that of their category is likely to evolve, and how this will affect their brand sales. Together with metrics already present on a brand managers' dashboards (e.g., overall share and sales), these new metrics can be added such that a more complete overview arises for brands operating in increasingly digital markets.

Moreover, we also indicate how this ties in with the characteristics of the category and the brand. While domestic players are generally believed to be on the winning hand in China (Bain & Company and Kantar Worldpanel 2017), our results suggest that this trend may be curbed by the growth of the online channel, where foreign brands appear to outperform domestic brands. As expected, managers that aim to benefit from the online grocery trend should focus on becoming or remaining present on a large number of websites, while focusing not so much on reducing offline unavailability. Moreover, though it is not a dominant factor, brand's online to offline price ratio does influence BOI. Thus, despite claims that consumers' online price sensitivity is lower (Chu, Chintagunta, and Cebollada 2008; Degeratu, Rangaswamy, and Wu 2000; Lynch and Ariely 2000), brands should avoid online prices to become higher than offline prices.

Against expectations, we find a negative effect of pack size. That is, brands that sell larger pack sizes than what is typically sold in the category, generally have lower BOIs, most likely because consumers may see their shipping fee increase as such a large (i.e., heavy) pack is added to the shopping basket. Contrary to common wisdom, brands in the higher price tiers should not expect to particularly benefit in the online vs. offline channel, while highly trusted brands do benefit, and fun brands suffer from the online grocery trend. Thus, in order to increase BOI, brand managers should invest in marketing programs that instill trust. In contrast, managers of fun brands should not put too much effort in the online channel. Rather, they would better invest their resources in the offline channel, where chances of being bought on impulse are much higher. Finally, advertising does not significantly influence BOI.

This may be the result of the fact that advertising plays more important roles in influencing long term, rather than short term, success (Ataman, Van Heerde, and Mela 2010).

How about category characteristics? It appears that categories with large assortment sizes flourish online. Having access to a wide range of products is generally believed to play an important role in the online channel, in which consumers can benefit from the varied offer without incurring the search costs or suffering the overload they would experience offline. This aspect might become even more important as online shoppers become more experienced (Melis et al. 2015). While expensive categories (for which each purchase represents a high monetary amount) are relatively more often bought online, the opposite holds for heavy categories. In line with the effect of large packs on BOI, these effects might be caused by the way delivery fees are set: shipping costs charged to consumers are a function of the order value (amount spent) and weight.

Another important category characteristic is perishability, which, as postulated, negatively affects COI. Though not directly under control of specific brand managers, online retailers may put effort in downsizing this problem. For example, Wal-Mart invested heavily in logistic systems in China, which enables the U.S. giant to do home deliveries of fresh products within an hour. Furthermore, locally embedded categories with high purchase frequencies are likely to struggle with the online channel, whereas for categories in which brand cues play an important risk-reduction role, the online channel provides excellent opportunities.

Limitations and future research

Our study opens up important avenues for future research. First, our empirical analysis pertained to only one market, China. Though China is very important in terms of total CPG, and leads the way when it comes to online share, some of the effects may be idiosyncratic to the country. Future studies should verify generalizability to other markets.

Second, our measures of BOI and COI represent a snapshot, and even their drivers may change over time, as consumers become more accustomed to buying online, and the modalities of buying in the digital channel evolve. Though we expect this to be a gradual process, it implies that the impact of drivers needs to be revisited as time evolves.

Third, the BOI and COI metrics, and the underlying brand- and category determinants, may depend on the setting (e.g., urban vs. rural local market) and online format. For instance, placing the online order through different devices (e.g., desktop vs. mobile phone), or using different types of online order fulfillment (e.g., home delivery vs. click and collect) may trigger different marketing mix responses, and favor some categories more than others. As these different formats become more important, a separate study of their impact on brands' relative market position and categories' sales shares is warranted.

Chapter 5 | Conclusion

With the growing global economic power of emerging markets (EMs), and China in particular, performing well in these markets has become ever more important for CPG manufacturers. However, becoming successful requires knowledge of marketing mix effectiveness in EMs, which may well differ from that in developed markets (DMs), and is not well documented in academic literature to date. Though obtaining data from EMs may be quite challenging, we have access to a unique and rich dataset that allows us to study the use of marketing mix instruments and brand learning processes in the (online) Chinese CPG industry (by taking into account differences across brands/categories, consumers, and time). With this dissertation, we aim to provide both academics and practitioners with a better understanding of the purchase behavior and marketing-mix responsiveness of consumers living in EMs. The next section provides a summary of each essay, followed by a discussion on the implications of these results, and recommendations for brand managers that operate in EMs. The last section discusses the limitations of the essays and directions for future research.

Summary of Findings

This dissertation consists of three research-based chapters, the findings of which are summarized below.

Chapter 2 | Price Elasticities for CPG Brands in China: Empirical Generalizations from a Large Scale Study

In Chapter 2, we studied the relationship between price and market share (while controlling for other marketing mix instruments), and conceptualized which category and brand factors moderate this relationship in EMs. To test our propositions, we conducted a comprehensive analysis of price elasticities for 376 brands in 50 CPG categories over the period 2011-2015 in China. We assessed the moderating effect of eight category and brand factors, and the relative importance of price vs. three other key marketing instruments – advertising, distribution, and line length. We find that CPG markets in China are generally price inelastic (and for less than one-fifth of the brands, demand is elastic). Yet, across the brands and categories under study, there is large heterogeneity in price elasticities, which suggests the presence of moderators. Some of these moderating effects mimic established findings in DMs. Like in Western markets, price sensitivity is higher in more concentrated categories and in less-perishable categories. Turning to the brand factors, highly promoted brands have larger price sensitivities than brands that are hardly promoted, while advertising has a dampening effect on price elasticity. Interestingly, the three moderators that have hardly received attention in previous price elasticity research and are deemed unimportant in DMs, have a strong combined effect in EMs. The predicted price elasticity of a foreign brand in a category that has been around in China for a long time and is of low social demonstrance, is .90 larger in magnitude than the predicted price elasticity of a domestic brand in a 'new' category of high social demonstrance. Finally, we find that, across the marketing mix instruments, price is important, but it is not the dominant instrument. In fact, we find that distribution matters the most. Expanding the brand assortment with stock keeping units (SKUs) is another powerful instrument, while advertising's effect is on average non-significant as well as negligible.

Chapter 3 | Consumer Learning about Quality of Global and Local Brands in the CPG Industry in China

In chapter 3, we studied the effects of brand quality and quality uncertainty on brand choice behavior, for global vs. local brands. We used a unique scanner panel dataset of urban

Chinese households over the period 2011-2014 to estimate a Bayesian learning model on five product categories. Our study reveals that Chinese consumers, in general, attach higher quality to global than to local brands, and are not necessarily more uncertain about the quality of global brands. Yet, this overall pattern conceals geographic and sociodemographic differences. For instance, in contrast to consumers with higher incomes that live in high-tier cities in the East of China, less affluent people from low-tier cities elsewhere attach lower quality premiums to global brands, and are more uncertain about these brands. Furthermore, the results show that, next to distribution, quality uncertainty is a key driver of brand success in China, especially when targeting older, less affluent consumers from low-tier cities.

Chapter 4 | The Rise of Online Grocery Shopping: Which Brands Will Benefit?

In chapter 4, we derived how a brand's total (online plus offline) sales change as the fraction of groceries sold online goes up, and showed that it critically depends on two simple (and easy-to-calculate) metrics: (i) the brand's online index (BOI) and (ii) the category's online index (COI). While the former indicates how the brand's relative position within the category will evolve, the latter indicates how the category's overall CPG share will contribute to (or hamper) brand sales as the online CPG channel grows. Combining COI and BOI thus provides managers a first indication of how their relative market position and that of their category is likely to evolve, and how this will affect their brand sales. We then identify brand and category factors that drive the two indices. Our analyses show that BOI not only increases with higher levels of online availability and lower online to offline price ratios, it is also higher for foreign and 'trusted' brands, yet lower for 'fun' brands. As for COI, less-frequently bought, expensive, and locally-embedded categories with large assortment sizes benefit from the shift towards the online channel, whereas the opposite holds for perishable categories.

Implications and Recommendations

The results of the three essays that make up this dissertation have implications for brand managers operating in the Chinese CPG industry and offer important guidelines. Some may believe that success in China is first and foremost about price – not an unreasonable assumption given that Chinese average monthly disposable income per capita in 2014 was only $731 vs. $3,258 for the U.S. As shown across the three chapters, price obviously plays a role, but assortment decisions are about equally important while distribution and especially quality uncertainty matter substantially more. Moreover, to be successful in China, brand managers have to reckon with factors that are less important in DMs. For example, the extent to which the category is deeply embedded in Chinese society and has a social demonstrance function will influence the effectiveness of the price instrument, as well as the brand's online vs. offline success. Likewise, whether the brand is owned by a foreign or Chinese manufacturer, or sold globally vs. only in China, makes a difference. At the same time, brand managers have to be aware that they cannot always rely on factors that work well in DMs. For example, while premium brands in DMs have higher price elasticities and perform better online than offline, positioning the brand towards the higher end of the market hardly has an influence on these metrics in EMs like China. Finally, some factors play similar roles in DMs and EMs. Examples include how concentrated a category is, and whether it contains perishable vs. non-perishable products. Below, we outline what effects a manager may expect when pressing different buttons in China.

Chapter 3 has shown that reducing consumers' uncertainty about the quality of their brand should be one of the main focus points of managers operating in China: stimulating usage/trial via sampling, refund policies or by facilitating gift giving, could reduce uncertainty about the brand's quality substantially. A side effect from investing in such trust-enhancing marketing programs will be that the brand will also benefit more from the online grocery trend in China (as shown in chapter 4). In line with this, for brand managers that operate in categories in which brand cues play an important risk-reduction role, the online channel provides excellent opportunities. For managers of global brands, we find that these brands are favored on quality, but also that Chinese consumers are not necessarily more certain about the quality. Therefore, if global brands could combine higher quality with lower uncertainty, that would put them in a very strong position in China. For global brands, reducing uncertainty via stimulating usage/trial would be especially meaningful for older consumers with lower incomes who live in low-tier cities in the West, North or South of China as these consumers feel more uncertain about the quality of global vs. local brands than their counterparts in high-tier cities in the East.

Investing in distribution, both offline and online, is crucial too for healthy brand performance in China. The results of chapters 2 and 3 show that brands with higher distribution levels perform considerably better compared to brands that are less widely distributed, especially in low-tier cities. Chapter 4 further reveals that for managers that particularly aim to benefit from the online grocery trend, it would be wise to put most effort in becoming or remaining present on a large number of websites, while maintaining offline unavailability.

Across the marketing mix, we find the price instrument to be generally of moderate importance at best. Still, chapter 3 showed that less affluent consumers are generally somewhat more price elastic, and chapter 2 has shown that important differences exist across brands and categories. Of special interest are the roles that social demonstrance, local embeddedness, and foreign vs. domestic brand ownership play in influencing price sensitivity – three factors that CPG managers in the U.S. or Europe might not readily consider as particularly relevant, but that play important roles in EMs like China. The most important factor is social demonstrance: price elasticity is considerably lower in categories that have a high social demonstrance function. Market leaders in categories that are low on social demonstrance could thus attempt to increase the symbolic value of the category. This reduces (category) price sensitivity, which is attractive for brands in a leading position. Furthermore, compared to buying 'new' categories, people are generally more price sensitive when buying brands in categories that are deeply embedded in Chinese society. Also, although foreign brands outperform domestic brands in the online channel; they have larger absolute price elasticities than domestic brands. One way to reduce price sensitivity for brands in locally embedded categories and/or foreign brands, is to use principles established in developed markets that will move consumers' focus away from price: increase advertising spending and focus less on price promotions – factors that hardly influence the brand's online and/or offline performance in China. Finally, brands in more expensive categories (i.e., for which each purchase represents a high monetary amount) will help online shoppers to reduce shipping costs, and may therefore expect to benefit from the online grocery trend in China. However, whether a manager chooses to position its brand towards the higher-priced end of the category seems to play a minor role: it generally does not affect the brand's price elasticity (chapter 2), nor its online to offline performance (chapter 4).

Though chapters 2 and 3 showed that a brand's line length generally only plays a moderate role, chapter 4 reveals that a category's line length is the most important factor influencing COI (which, together with BOI, determines to what extent brand sales will increase as a result of the online grocery trend). Thus, brands should not aim to decrease the number of SKU offerings in an attempt to prevent choice overload in the category: with the growing importance of the online channel, and the tools that this channel offers to consumers to select the products they prefer, it seems that the more choice they have, the better. Brands should however be cautious not to offer many 'large' SKUs (i.e., that are larger than what is typically sold in the category), especially when present in a category that already consists of heavier products (like beer, olive oil, or laundry detergent): weight increases shipping costs when shopping online, which may lead consumers to add smaller (i.e., less heavy) packs to their shopping baskets.

Limitations and Suggestions for Future Research

The essays of this dissertation have several limitations that provide interesting avenues for future research. First, our empirical analyses only pertain to China. Though China is the biggest EM economy (in terms of GDP, as well as in (online) CPG), and shares features with other EMs, it also has distinct characteristics. Any generalization to other EMs is subject to further research that should verify to what extent our results hold in other EMs. Though, especially in chapter 2, we used Chinese findings to suggest approximate answers for the other three BRIC nations (i.e., Brazil, Russia, and India), these findings should be verified with primary research in these countries.

Second, our empirical context is the CPG industry. It remains to be tested whether, and if yes, to what extent, our findings on the price instrument, brand learning processes, and online shopping, also hold for other industries like services and durables. The relative importance of ‘core’ marketing mix instruments like quality (uncertainty), distribution, price, line length, and advertising may well differ for these products. Moreover, given the importance in EMs of social demonstrance, local embeddedness, and brand ownership, it would be beneficial to include these measures in this research as well.

Third, we mostly focused on (online and/or offline) choice or market share as the key performance metrics. Future studies could consider brand-sales and category-expansion effects of marketing efforts in EMs. Moreover, we mainly focused on short term effects. Given the highly dynamic nature of EMs, future studies could explore how market shares (or sales) can be influenced in the longer term.

Fourth, only one essay of this dissertation was using a model at the household level, the other two essays were based on large-scale analyses at the market level. Though this enabled us to provide insights on a diverse set of brands and categories, given the large heterogeneity of EMs, it may be useful to further study marketing responses at the household level, for a larger set of categories than the five product categories we looked at. When doing so, special attention should be given to the heterogeneity across consumers in urban vs. rural areas (as our study was limited to the urban parts of China), and to a richer set of variables, including consumer traits (as we only focused on limited geographic and sociodemographic information).

Fifth, though we documented the effects of multiple category- and brand-drivers on price elasticity and online vs. offline brand performance, other factors remain to be explored. For example, we have looked at the effects of line length in general, but future research could study how innovation programs can be effectively managed in EMs. Also, the effectiveness of the drivers we looked at may well change over time. For instance, as consumers become more accustomed to buying online, especially factors that hamper online performance (like perishability) may become less important. In addition, with the speed and change that define China, factors that do not play a role now, may become very important in the foreseeable future: for example, while private labels currently hardly play a role in EMs, they may start to grow over time, in a way that may differ from how they developed in DMs.

Much remains to be studied before we can offer definitive guidelines and empirical generalizations regarding brand learning and marketing mix effectiveness of brands operating in the CPG industries of EMs. We hope that this dissertation will spark additional research on how to set and execute marketing strategies in such markets.

References

Aaker, David A. (1991), *Managing Brand Equity*, New York: Free Press.

Arts, Joep W.C., Ruud T. Frambach, and Tammo H.A. Bijmolt (2011), “Generalizations on consumer innovation adoption: A meta-analysis on drivers of intention and behavior,” *International Journal of Research in Marketing*, 28 (2), 134–44.

Ataman, M. Berk, Harald J. Van Heerde, and Carl F. Mela (2010), “The Long-Term Effect of Marketing Strategy on Brand Sales,” *Journal of Marketing Research*, 47 (5), 866–82.

———, Carl F. Mela, and Harald Van Heerde (2008), “Building Brands,” *Marketing Issues in Transitional Economies*, 27 (6), 1036–54.

Avery, Jill, Thomas J. Steenburgh, John Deighton, and Mary Caravella (2012), “Adding Bricks to Clicks: Predicting the Patterns of Cross-Channel Elasticities Over Time,” *Journal of Marketing*, 76 (3), 96–111.

Babin, Barry J. and William R. Darden (1995), “Consumer Self-Regulation in a Retail Environment,” *Journal of Retailing*, 71 (1), 47–70.

Bain & Company and Kantar Worldpanel (2012), “What Chinese Shoppers Really Do But Will Never Tell You.”

——— (2015), “Winning Over Shoppers in China’s ‘New Normal.’”

——— (2017), “China’s two-speed growth: in and out of the home.”

Batra, Rajeev (1999), “Marketing Issues and Challenges in Transitional Economies,” in *Marketing Issues in Transitional Economies*, R. Batra, ed., Boston, MA: Springer, 3–35.

———, Venkatram Ramaswamy, Dana L. Alden, Jan-Benedict E.M. Steenkamp, and S. Ramachander (2000), “Effects of Brand Local and Nonlocal Origin on Consumer Attitudes in Developing Countries,” *Journal of Consumer Psychology*, 9 (2), 83–95.

BCG (2008), “Foreign or Local Brands in China? Rationalism Trumps Nationalism.”

——— (2017), “The Chinese Consumer’s Online Journey From Discovery To Purchase.”

Bearden, William O., Richard G. Netemeyer, and Jesse E. Teel (1989), “Measurement of Consumer Susceptibility to Interpersonal Influence,” *Journal of Consumer Research*, 15 (4), 473–81.

Besanko, David, Sachin Gupta, and Dipak Jain (1998), “Logit Demand Estimation Under Competitive Pricing Behavior: An Equilibrium Framework,” *Management Science*, 44 (11–1), 1533–47.

Bijmolt, Tammo H.A., Harald J. Van Heerde, and Rik G.M. Pieters (2005), “New Empirical Generalizations on the Determinants of Price Elasticity,” *Journal of Marketing Research*, 42 (2), 141–56.

Boulding, William, Eunkyu Lee, and Richard Staelin (1994), “Mastering the Mix: Do Advertising, Promotion, and Sales Force Activities Lead to Differentiation?,” *Journal of Marketing Research*, 31 (2), 159–72.

Bronnenberg, Bart J., Jean-Pierre H. Dubé, Matthew Gentzkow, and Jesse M. Shapiro (2015), “Do Pharmacists Buy Bayer? Informed Shoppers And The Brand Premium,” *The Quarterly Journal of Economics*, 130 (4), 1669–1726.

Bruno, Hernán A., Javier Cebollada, and Pradeep K. Chintagunta (2018), “Targeting Mr. or Mrs. Smith: Modeling and Leveraging Intrahousehold Heterogeneity in Brand Choice Behavior,” *Marketing Science*.

Burgess, Steven Michael and Jan-Benedict E.M. Steenkamp (2006), “Marketing Renaissance:

How Research in Emerging Markets Advances Marketing Science and Practice," *International Journal of Research in Marketing*, 23 (4), 337–56.

Burke, Raymond R., Bari A. Harlam, Barbara E. Kahn, and Leonard M. Lodish (1992), "Comparing Dynamic Consumer Choice in Real and Computer-simulated Environments," *Journal of Consumer Research*, 19 (1), 71–82.

Campo, Katia and Els Breugelmans (2015), "Buying Groceries in Brick and Click Stores: Category Allocation Decisions and the Moderating Effect of Online Buying Experience," *Journal of Interactive Marketing*, 31, 63–78.

Carpenter, Gregory S., Lee G. Cooper, Dominique M. Hanssens, and David F. Midgley (1988), "Modeling Asymmetric Competition," *Marketing Science*, 7 (4), 393–412.

Center for Science in the Public Interest (2016), "Carbonating the World: The Marketing and Health Impact of Sugar drinks in Low-and middle-income Countries."

Chan, Jenny (2014), "Chinese consumers no longer dazzled," *Campaign Asia-Pacific*, March, 6.

Chaudhuri, Arjun and Morris B. Holbrook (2001), "The Chain of effects from brand trust and brand affect to Brand Performance: The Role of Brand Loyalty," *Journal of Marketing*, 65 (2), 81–93.

China Briefing (2015), "Tmall, Yihaodian and JD: A Comparison of China's Top E-Commerce Platforms for Foreign Enterprises."

China Daily (2015), "Rural-Urban Income Gap Narrows."

China Internet Watch (2015), "22% Online Shopping Spend Are Newly Created Demand."

Ching, Andrew T., Tülin Erdem, and Michael P. Keane (2013), "Learning Models: An Assessment of Progress, Challenges and New Developments," *Marketing Science*, 32 (6), 913–38.

Chintagunta, Pradeep K. (2018), "Structural Models in Marketing," in *Handbook of Marketing Analytics*, N. Mizik and D. M. Hanssens, eds., Northampton, M.A.: Edward Elgar, 200–223.

Christen, Markus, Sachin Gupta, John C. Porter, Richard Staelin, and Dick R. Wittink (1997), "Using Market-Level Data to Understand Promotion Effects in a Nonlinear Model," *Journal of Marketing Research*, 34 (3), 322–34.

Chu, Junhong, Marta Arce-Urriza, José-Javier Cebollada-Calvo, and Pradeep K. Chintagunta (2010), "An Empirical Analysis of Shopping Behavior Across Online and Offline Channels for Grocery Products: The Moderating Effects of Household and Product Characteristics," *Journal of Interactive Marketing*, 24 (4), 251–68.

———, Pradeep K. Chintagunta, and Javier Cebollada (2008), "Research Note—A Comparison of Within-Household Price Sensitivity Across Online and Offline Channels," *Marketing Science*, 27 (2), 283–99.

Cleeren, Kathleen, Harald J. van Heerde, and Marnik G Dekimpe (2013), "Rising from the Ashes: How Brands and Categories Can Overcome Product-Harm Crises," *Journal of Marketing*, 77 (2), 58–77.

Cooper, Lee G. and Masao Nakanishi (1988), *Market-Share Analysis*, Boston Dordrecht London: Kluwer Academic Publishers.

Crawford, Gregory S. and Matthew Shum (2005), "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 73 (4), 1137–73.

Danaher, Peter J., Isaac W. Wilson, and Robert A. Davis (2003), "A Comparison of Online and Offline Consumer Brand Loyalty," *Marketing Science*, 22 (4), 461–76.

Datta, Hannes, Kusum L. Ailawadi, and Harald J. van Heerde (2017), "How Well Does Consumer-Based Brand Equity Align with Sales-Based Brand Equity and Marketing-Mix Response?," *Journal of Marketing*.

Dawar, Niraj and Amitava Chattopadhyay (2002), "Rethinking Marketing Programs for

Emerging Markets," *Long Range Planning*, 35 (5), 457–74.
Degeratu, Alexandru M., Arvind Rangaswamy, and Jianan Wu (2000), "Consumer choice behavior in online and traditional supermarkets: The effects of brand name, price, and other search attributes," *International Journal of Research in Marketing*, 17 (1), 55–78.
DeGroot, Morris H. (1970), *Optimal Statistical Decisions*, New York: McGraw-Hill.
Doland, Angela (2015), "Why Foreign Brands Have to Figure Out China All Over Again," *Advertising Age*, (December 7), 1–11.
Du, Rex Y. and Wagner A. Kamakura (2006), "Household Life Cycles and Lifestyles in the United States," *Journal of Marketing Research*, 43 (1), 121–32.
Dubin, Jeffrey A. and Daniel L. McFadden (1984), "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption," *Econometrica*, 52 (2), 345–62.
Enders, Walter (2004), *Applied Econometric Time Series*, New York, NY: J. Wiley, 3rd ed.
Erdem, Tülin (1998), "An Empirical Analysis of Umbrella Branding," *Journal of Marketing Research*, 35 (3), 339–51.
——— and Michael P. Keane (1996), "Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science*, 15 (1), 1–20.
———, ———, and Baohong Sun (2008), "A Dynamic Model of Brand Choice When Price and Advertising Signal Product Quality," *Marketing Science*, 27 (6), 1111–25.
——— and Baohong Sun (2002), "An Empirical Investigation of the Spillover Effects of Advertising and Sales Promotions in Umbrella Branding," *Journal of Marketing Research*, 39 (4), 408–20.
——— and Joffre Swait (2004), "Brand Credibility, Brand Consideration, and Choice," *Journal of Consumer Research*, 31 (1), 191–98.
———, ———, and Jordan Louviere (2002), "The Impact of Brand Credibility on Consumer Price Sensitivity," *International Journal of Research in Marketing*, 19 (1), 1–19.
———, ———, and Ana Valenzuela (2006), "Brands as Signals: A Cross-Country Validation Study," *Journal of Marketing*, 70 (1), 34–49.
———, Ying Zhao, and Ana Valenzuela (2004), "Performance of Store Brands: A Cross-Country Analysis of Consumer Store-Brand Preferences, Perceptions, and Risk," *Journal of Marketing Research*, 41 (1), 86–100.
Fischer, Marc, Franziska Völckner, and Henrik Sattler (2010), "How Important Are Brands? A Cross-Category, Cross-Country Study," *Journal of Marketing Research*, 47 (October), 823–39.
FMI and Nielsen (2018), "Digitally Engaged Food Shopper Study," in *FMI Midwinter Conference*, Miami, FL.
Fok, Dennis, Csilla Horváth, Richard Paap, and Philip Hans Franses (2006), "A Hierarchical Bayes Error Correction Model to Explain Dynamic Effects of Price Changes," *Journal of Marketing Research*, 43 (3), 443–61.
Gao, Huachao, Yinlong Zhang, and Vikas Mittal (2017), "How Does Local–Global Identity Affect Price Sensitivity?," *Journal of Marketing*, 81 (3), 62–79.
George, Jennifer, Alan Mercer, and Helen Wilson (1996), "Variations in price elasticities," *European Journal of Operational Research*, 88 (1), 13–22.
Geyskens, Inge, Katrijn Gielens, and Els Gijsbrechts (2010), "Proliferating Private-Label Portfolios: How Introducing Economy and Premium Private Labels Influences Brand Choice," *Journal of Marketing Research*, 47 (5), 791–807.
Gijsenberg, Maarten J., Harald J. Van Heerde, Marnik G. Dekimpe, and Vincent R. Nijs (2011), "Understanding the Role of Adstock in Advertising Decisions."
Gordon, Brett R, Avi Goldfarb, and Yang Li (2013), "Does Price Elasticity Vary with

Economic Growth? A Cross-Category Analysis," *Journal of Marketing Research*, 50 (1), 4–23.
Guimaraes, Pedro Pacheco and Pierre Chandon (2007), *Unilever in Brazil (1997-2007): Marketing Strategies for Low-Income Consumers, INSEAD*, 1–22.
Häubl, Gerald and Valerie Trifts (2000), "Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Decision Aids," *Marketing Science,* 19 (1), 4–21.
Van Heerde, Harald J., Els Gijsbrechts, and Koen Pauwels (2008), "Winners and Losers in a Major Price War," *Journal of Marketing Research*, 45 (5), 499–518.
———, ———, and ——— (2015), "Fanning the Flames? How Media Coverage of a Price War Affects Retailers, Consumers, and Investors," *Journal of Marketing Research*, 52 (5), 674–93.
———, Maarten J. Gijsenberg, Marnik G. Dekimpe, and Jan-Benedict E.M. Steenkamp (2013), "Price and Advertising Effectiveness over the Business Cycle," *Journal of Marketing Research,* 50 (2), 177–93.
Heilman, Carrie M., Douglas Bowman, and Gordon P. Wright (2000), "The Evolution of Brand Preferences and Choice Behaviors of Consumers New to a Market," *Journal of Marketing Research*, 37 (2), 139–55.
Hernandez, José Mauro C. (2002), "Brand Trust and Online Consumer Behavior," in *Advances in Consumer Research,* S. M. Broniarczyk and K. Nakamoto, eds., Valdosta, GA, 255–56.
Hoch, Stephen J. and John Deighton (1989), "Managing What Consumers Learn from Experience," *Journal of Marketing,* 53 (2), 1–20.
Hofstede, Geert, Gert-Jan Hofstede, and Michael Minkov (2010), *Cultures and Organizations-Software of the Mind*, New York, NY: McGraw Hill Companies.
Holt, Douglas B., John A. Quelch, and Earl L. Taylor (2004), "How Global Brands Compete," *Harvard Business Review*, 82 (9), 68–75.
Inman, J .Jeffrey, Russel S. Winer, and Rosellina Ferraro (2009), "Characteristics , and Customer Activities on In-Store Decision," *Journal of Marketing*, 73 (5), 19–29.
De Jong, Martijn G., Jan-Benedict E. M. Steenkamp, and Jean-Paul Fox (2007), "Relaxing Measurement Invariance in Cross-National Consumer Research Using a Hierarchical IRT Model," *Journal of Consumer Research*, 34 (2), 260–78.
Kantar Millward Brown (2010), "The Differences Between Mature and Emerging Markets."
Kantar Worldpanel (2015), "Accelerating the E-Commerce in FMCG."
——— (2017), "The Future of E-Commerce in FMCG."
Kotabe, Masaaki and Kristiaan Helsen (2010), *Global Marketing Management*, Hoboken, NJ: Wiley.
Kumar, Nirmalaya and Jan-Benedict E.M. Steenkamp (2007), *Private Label Strategy,* Cambridge, MA: Harvard Business Press.
Lewis, Jeffrey B. and Drew A. Linzer (2005), "Estimating Regression Models in Which the Dependent Variable is Based on Estimates," *Political Analysis*, 13 (4), 345–64.
Lourenço, Carlos J.S., Els Gijsbrechts, and Richard Paap (2015), "The Impact of Category Prices on Store Price Image Formation: An Empirical Analysis," *Journal of Marketing Research,* 52 (2), 200–216.
Luan, Y. Jackie and K. Sudhir (2010), "Forecasting Marketing-Mix Responsiveness for New Products," *Journal of Marketing Research,* 47 (3), 444–57.
Lynch, John G. and Dan Ariely (2000), "Wine Online: Search Costs Affect Competition on Price, Quality, and Distribution," *Marketing Science*, 19 (1), 83–103.
Ma, Yu, Kusum L. Ailawadi, Dinseh K. Gauri, and Dhruv Grewal (2011), "An Empirical Investigation of the Impact of Gasoline Prices on Grocery Shopping Behavior," *Journal*

of Marketing, 75 (2), 18–35.
Mazumdar, Tridib, S.P. Raj, and Indrajit Sinha (2005), "Reference Price Research: Review and Propositions," *Journal of Marketing*, 69 (4), 84–102.
McKinsey (2012), "From Mass to Mainstream: Keeping Pace With China's Rapidly Changing Consumers."
——— (2013a), "China's e-tail revolution."
——— (2013b), "The Future of Online Grocery in Europe."
Mehta, Nitin, Surendra Rajiv, and Kannan Srinivasan (2003), "Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation," *Marketing Science*, 22 (1), 58–84.
Mela, Carl F., Sunil Gupta, and Donald R. Lehmann (1997), "The Long-Term Impact of Promotion and Advertising on Consumer Brand Choice," *Journal Of Marketing Research*, 34 (2), 248–61.
Melis, Kristina, Katia Campo, Els Breugelmans, and Lien Lamey (2015), "The Impact of the Multi-channel Retail Mix on Online Store Choice: Does Online Experience Matter?," *Journal of Retailing*, 91 (2), 272–88.
Meyer, Jack (1987), "Two-Moment Decision Models and Expected Utility Maximization," *The American Economic Review*, 77 (3), 421–30.
Morgan Stanley (2018), "Bullish on China's Lower-Tier Cities."
Narasimhan, Chakravarthi, Scott A. Neslin, and Subrata K. Sen (1996), "Promotional Elasticities and Category Characteristics," *Journal of Marketing*, 60 (April), 17–30.
Narasimhan, Laxman, Kannan Srinivasan, and K. Sudhir (2015), "Editorial — Marketing Science in Emerging Markets," *Marketing Science*, 34 (4), 473–79.
Narayanan, Sridhar and Puneet Manchanda (2009), "Heterogeneous Learning and the Targeting of Marketing Communication for New Products," *Marketing Science*, 28 (3), 424–41.
Nielsen (2017), "What's In-Store for Online Grocery Shopping."
Nijs, Vincent R., Marnik G. Dekimpe, Jan-Benedict E.M. Steenkamp, and Dominique M. Hanssens (2001), "The Category-Demand Effects of Price Promotions," *Marketing Science*, 20 (1), 1–22.
Park, Sungho and Sachin Gupta (2012), "Handling Endogenous Regressors by Joint Estimation Using Copulas," *Marketing Science*, 31 (4), 567–86.
Pauwels, Koen, Selin Erguncu, and Gokhan Yildirim (2013), "Winning Hearts, Minds and Sales: How Marketing Communication Enters the Purchase Process in Emerging and Mature Markets," *International Journal of Research in Marketing*, 30 (1), 57–68.
PwC (2017), "The Long View - How Will the Global Economic Order Change by 2050?"
Raju, Jagmohan S. (2005), "Revitalizing the Role of Marketing in Business Organizations: What Can Poor Academics Do to Help?," *Journal of Marketing*, 69 (4), 17.
Rosenthal, Robert (1991), *Meta-Analytic Procedures for Social Research*, Beverly Hills, CA: Sage, 2nd ed.
Rossi, Peter E., Robert E. McCulloch, and Greg M. Allenby (1996), "The Value of Purchase History Data in Target Marketing," *Marketing Science*, 15 (4), 321–40.
Rozin, Paul, Claude Fischler, Christy Shields, and Estelle Masson (2006), "Attitudes Towards Large Numbers of Choices in the Food Domain: A Cross-Cultural Study of Five Countries in Europe and the USA," *Appetite*, 46 (3), 304–8.
Shankar, Venkatesh, Amy K. Smith, and Arvind Rangaswamy (2003), "Customer satisfaction and loyalty in online and offline environments," *International Journal of Research in Marketing*, 20 (2), 153–75.
Shimp, Terence A. and Subhash Sharma (1987), "Consumer Ethnocentrism: Construction and Validation of the CETSCALE," *Journal of Marketing Research*, 24 (3), 280.

Shin, Sangwoo, Sanjog Misra, and Dan Horsky (2012), "Disentangling Preferences and Learning in Brand Choice Models," *Marketing Science*, 31 (1), 115–37.

Song, Hyunjin and Norbert Schwarz (2009), "If It's Difficult to Pronounce, It Must Be Risky," *Psychological Science,* 20 (2), 135–38.

Sotgiu, Francesca and Katrijn Gielens (2015), "Suppliers Caught in Supermarket Price Wars: Victims or Victors? Insights from a Dutch Price War," *Journal of Marketing Research,* 52 (6), 784–800.

Srinivasan, Shuba, Koen Pauwels, Dominique M. Hanssens, and Marnik G. Dekimpe (2004), "Do Promotions Benefit Manufacturers, Retailers, or Both?," *Management Science,* 50 (5), 617–29.

Steenkamp, Jan-Benedict E.M. (2014), "How Global Brands Create Firm Value: The 4V Model," *International Marketing Review*, 31 (1), 5–29.

———, Rajeev Batra, and Dana L. Alden (2003), "How Perceived Brand Globalness Creates Brand Value," *Journal of International Business Studies,* 34 (1), 53–65.

——— and Hans Baumgartner (1998), "Assessing Measurement Invariance in Cross-National Consumer Research," *Journal of Consumer Reserach*, 25 (1), 78–107.

——— and Inge Geyskens (2014), "Manufacturer and Retailer Strategies to Impact Store Brand Share: Global Integration, Local Adaptation, and Worldwide Learning," *Marketing Science*, 33 (1), 6–26.

——— and Martijn G. de Jong (2010), "A Global Investigation into the Constellation of Consumer Attitudes Toward Global and Local Products," *Journal of Marketing,* 74 (6), 18–40.

Sudhir, K., Joe Priester, Matt Shum, David Atkin, Andrew Foster, Ganesh Iyer, Ginger Jin, Daniel Keniston, Shinobu Kitayama, Mushfiq Mobarak, Yi Qian, Ishani Tewari, and Wendy Wood (2015), "Research Opportunities in Emerging Markets: An Inter-disciplinary Perspective from Marketing, Economics, and Psychology," *Customer Needs and Solutions*, 2 (4), 264–76.

Szymanowski, Maciej and Els Gijsbrechts (2012), "Consumption-Based Cross-Brand Learning: Are Private Labels Really Private?," *Journal of Marketing Research*, 49 (2), 231–46.

——— and ——— (2013), "Patterns in consumption-based learning about brand quality for consumer packaged goods," *International Journal of Research in Marketing,* 30 (3), 219–35.

Tellis, Gerard J. (1988), "The Price Elasticity of Selective Demand: A Meta-Analysis of Econometric Models of Sales," *Journal of Marketing Research*, 25 (4), 331–41.

Voss, Kevin E., Eric R. Spangenberg, and Bianca Grohmann (2003), "Measuring the Hedonic and Utilitarian Dimensions of Consumer Attitude," *Journal of Marketing Research,* 40 (3), 310–20.

Warlop, Luk, S. Ratneshwar, and Stijn M.J. van Osselaer (2005), "Distinctive brand cues and memory for product consumption experiences," *International Journal of Research in Marketing*, 22 (1), 27–44.

World Economic Forum (2016), "3 Great Forces Changing China's Consumer Market."

Yip, George S. and Tomas M. Hult (2012), *Total Global Strategy,* Pearson.

Zeithaml, Valarie A. (1988), "Consumer Perceptions of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence," *Journal of Marketing*, 52 (3), 2–22.

Zhou, Kevin Zheng, Chenting Su, and Yeqing Bao (2002), "A Paradox of Price-Quality and Market Efficiency: A Comparative Study of The US and China Markets," *International Journal of Research in Marketing*, 19 (4), 349–65.

Zhu, Juliet (2013), "Understanding Chinese Consumers," (accessed April 1, 2018), [available at https://hbr.org/2013/11/understanding-chinese-consumers].

Appendices

Appendix Chapter 2

Appendix 2.A: Model Specifications

As indicated in the main text, the unit root tests on market share and price can give rise to four combinations, resulting in the following systems of equations (for ease of exposition, we group the advertising, distribution and line length controls here. Note that, depending on the outcomes of their unit root tests, those variables, as well as the cross-price effects, can enter the equations below in levels or differences):

Equation 2.1 and 2.2 if log market share has no unit root and log price has no unit root.

(2.1) $\log \mathrm{m_{jt}} - \log \mathrm{m_{0t}} = \beta_{\mathrm{m0j}} + \beta_{\mathrm{m1j}} \log \mathrm{trend_t} + \beta_{\mathrm{m2j}} \log \mathrm{m_{jt-1}} + \beta_{\mathrm{m3j}} \log \mathrm{m_{jt-2}} + \beta_{\mathrm{m4j}} \log \mathrm{p_{jt}} + \beta_{\mathrm{m5j}} \log \mathrm{p_{jt-1}} + \sum_{i,i\neq j} \beta_{\mathrm{m4ji}} \mathrm{p_{it}} + \sum_{i,i\neq j} \beta_{\mathrm{m5ji}} \mathrm{p_{it-1}} + \sum_{\mathrm{v}} \psi_{\mathrm{vj}}\, \mathrm{control_{vjt}} + \sum_{\mathrm{v}} \delta_{\mathrm{vj}}\, \mathrm{copula_{vjt}} + \varepsilon_{\mathrm{jt}}$

(2.2) $\log \mathrm{p_{jt}} = \beta_{\mathrm{p0j}} + \beta_{\mathrm{p1j}} \log \mathrm{trend_t} + \beta_{\mathrm{p2j}} \log \mathrm{m_{jt-1}} + \beta_{\mathrm{p3j}} \log \mathrm{m_{jt-2}} + \beta_{\mathrm{p4j}} \log \mathrm{p_{jt-1}} + \beta_{\mathrm{p5j}} \log \mathrm{p_{jt-2}} + \sum_{i,i\neq j} \beta_{\mathrm{p4ji}} \mathrm{p_{it-1}} + \sum_{i,i\neq j} \beta_{\mathrm{p5ji}} \mathrm{p_{it-2}} + \xi_{\mathrm{jt}}$

Equation 2.1 and 2.2 if log market share has a unit root and log price has a unit root.

(2.1a) $(\log \mathrm{m_{jt}} - \log \mathrm{m_{jt-1}}) - (\log \mathrm{m_{0t}} - \log \mathrm{m_{0t-1}}) = \beta_{\mathrm{m0j}} + \beta_{\mathrm{m1j}} \log \mathrm{trend_t} + \beta_{\mathrm{m2j}}(\log \mathrm{m_{jt-1}} - \log \mathrm{m_{jt-2}}) + \beta_{\mathrm{m3j}}(\log \mathrm{m_{jt-2}} - \log \mathrm{m_{jt-3}}) + \beta_{\mathrm{m4j}}(\log \mathrm{p_{jt}} - \log \mathrm{p_{jt-1}}) + \beta_{\mathrm{m5j}}(\log \mathrm{p_{jt-1}} - \log \mathrm{p_{jt-2}}) + \sum_{i,i\neq j} \beta_{\mathrm{m4ji}}(\mathrm{p_{it}} - \mathrm{p_{it-1}}) + \sum_{i,i\neq j} \beta_{\mathrm{m5ji}}(\mathrm{p_{it-1}} - \mathrm{p_{it-2}}) + \sum_{\mathrm{v}} \psi_{\mathrm{vj}}\, \mathrm{control_{vjt}} + \sum_{\mathrm{v}} \delta_{\mathrm{vj}}\, \mathrm{copula_{vjt}} + \varepsilon_{\mathrm{jt}}$

(2.2a) $(\log \mathrm{p_{jt}} - \log \mathrm{p_{jt-1}}) = \beta_{\mathrm{p0j}} + \beta_{\mathrm{p1j}} \log \mathrm{trend_t} + \beta_{\mathrm{p2j}}(\log \mathrm{m_{jt-1}} - \log \mathrm{m_{jt-2}}) + \beta_{\mathrm{p3j}}(\log \mathrm{m_{jt-2}} - \log \mathrm{m_{jt-3}}) + \beta_{\mathrm{p4j}}(\log \mathrm{p_{jt-1}} - \log \mathrm{p_{jt-2}}) + \beta_{\mathrm{p5j}}(\log \mathrm{p_{jt-2}} - \log \mathrm{p_{jt-3}}) + \sum_{i,i\neq j} \beta_{\mathrm{p4ji}}(\mathrm{p_{it-1}} - \mathrm{p_{it-2}}) + \sum_{i,i\neq j} \beta_{\mathrm{p5ji}}(\mathrm{p_{it-2}} - \mathrm{p_{it-3}}) + \xi_{\mathrm{jt}}$

Equation 2.1 and 2.2 if log market share has no unit root and log price has unit root.

(2.1b) $\log \mathrm{m_{jt}} - \log \mathrm{m_{0t}} = \beta_{\mathrm{m0j}} + \beta_{\mathrm{m1j}} \log \mathrm{trend_t} + \beta_{\mathrm{m2j}} \log \mathrm{m_{jt-1}} + \beta_{\mathrm{m3j}} \log \mathrm{m_{jt-2}} + \beta_{\mathrm{m4j}}(\log \mathrm{p_{jt}} - \log \mathrm{p_{jt-1}}) + \beta_{\mathrm{m5j}} (\log \mathrm{p_{jt-1}} - \log \mathrm{p_{jt-2}}) + \sum_{i,i\neq j} \beta_{\mathrm{m4ji}}(\mathrm{p_{it}} - \mathrm{p_{it-1}}) + \sum_{i,i\neq j} \beta_{\mathrm{m5ji}}(\mathrm{p_{it-1}} - \mathrm{p_{it-2}}) + \sum_{\mathrm{v}} \psi_{\mathrm{vj}} \mathrm{control_{vjt}} + \sum_{\mathrm{v}} \delta_{\mathrm{vj}} \mathrm{copula_{vjt}} + \varepsilon_{\mathrm{jt}}$

(2.2b) $(\log \mathrm{p_{jt}} - \log \mathrm{p_{jt-1}}) = \beta_{\mathrm{p0j}} + \beta_{\mathrm{p1j}} \log \mathrm{trend_t} + \beta_{\mathrm{p2j}} \log \mathrm{m_{jt-1}} + \beta_{\mathrm{p3j}} \log \mathrm{m_{jt-2}} + \beta_{\mathrm{p4j}} (\log \mathrm{p_{jt-1}} - \log \mathrm{p_{jt-2}}) + \beta_{\mathrm{p5j}}(\log \mathrm{p_{jt-2}} - \log \mathrm{p_{jt-3}}) + \sum_{i,i\neq j} \beta_{\mathrm{p4ji}}(\mathrm{p_{it-1}} - \mathrm{p_{it-2}}) + \sum_{i,i\neq j} \beta_{\mathrm{p5ji}}(\mathrm{p_{it-2}} - \mathrm{p_{it-3}}) + \xi_{\mathrm{jt}}$

Equation 2.1 and 2.2 if log market share has a unit root and log price has no unit root.

(2.1c) $(\log \mathrm{m_{jt}} - \log \mathrm{m_{jt-1}}) - (\log \mathrm{m_{0t}} - \log \mathrm{m_{0t-1}}) = \beta_{\mathrm{m0j}} + \beta_{\mathrm{m1j}} \log \mathrm{trend_t} + \beta_{\mathrm{m2j}}(\log \mathrm{m_{jt-1}} - \log \mathrm{m_{jt-2}}) + \beta_{\mathrm{m3j}}(\log \mathrm{m_{jt-2}} - \log \mathrm{m_{jt-3}}) + \beta_{\mathrm{m4j}} \log \mathrm{p_{jt}} + \beta_{\mathrm{m5j}} \log \mathrm{p_{jt-1}} + \sum_{i,i\neq j} \beta_{\mathrm{m4ji}}\mathrm{p_{it}} + \sum_{i,i\neq j} \beta_{\mathrm{m5ji}}\mathrm{p_{it-1}} + \sum_{\mathrm{v}} \psi_{\mathrm{vj}} \mathrm{control_{vjt}} + \sum_{\mathrm{v}} \delta_{\mathrm{vj}} \mathrm{copula_{vjt}} + \varepsilon_{\mathrm{jt}}$

(2.2c) $\log \mathrm{p_{jt}} = \beta_{\mathrm{p0j}} + \beta_{\mathrm{p1j}} \log \mathrm{trend_t} + \beta_{\mathrm{p2j}}(\log \mathrm{m_{jt-1}} - \log \mathrm{m_{jt-2}}) + \beta_{\mathrm{p3j}} (\log \mathrm{m_{jt-2}} - \log \mathrm{m_{jt-3}}) + \beta_{\mathrm{p4j}} \log \mathrm{p_{jt-1}} + \beta_{\mathrm{p5j}} \log \mathrm{p_{jt-2}} + \sum_{i,i\neq j} \beta_{\mathrm{p4ji}}\mathrm{p_{it-1}} + \sum_{i,i\neq j} \beta_{\mathrm{p5ji}}\mathrm{p_{it-2}} + \xi_{\mathrm{jt}}$

Re-write equations in levels.

For each of these four cases, we can re-write the estimated market share and price equations in levels:

(2.1d) $\log \mathrm{m_{jt}} - \log \mathrm{m_{0t}} = \beta_{\mathrm{m0j}} + \beta_{\mathrm{m1j}} \log \mathrm{trend_t} + \zeta_{\mathrm{m1j}} \log \mathrm{m_{jt-1}} + \zeta_{\mathrm{m2j}} \log \mathrm{m_{jt-2}} + \zeta_{\mathrm{m3j}} \log \mathrm{m_{jt-3}} + \zeta_{\mathrm{m4j}} \log \mathrm{m_{0t-1}} + \zeta_{\mathrm{m5j}} \log \mathrm{p_{jt}} + \zeta_{\mathrm{m6j}} \log \mathrm{p_{jt-1}} + \zeta_{\mathrm{m7j}} \log \mathrm{p_{jt-2}} + \sum_{\mathrm{I,i\neq j}} \beta_{\mathrm{m4ji}}\mathrm{p_{it}} + \sum_{\mathrm{I,i\neq j}} \beta_{\mathrm{m5ji}}\mathrm{p_{it-1}} + \sum_{\mathrm{v}} \psi_{\mathrm{vj}} \mathrm{control_{vjt}} + \sum_{\mathrm{k}} \delta_{\mathrm{kj}} \mathrm{copula_{kjt}} + \varepsilon_{\mathrm{jt}}$

(2.2d) $\log p_{jt} = \beta_{p0j} + \beta_{p1j} \log trend_t + \theta_{p1j} \log m_{jt-1} + \theta_{p2j} \log m_{jt-2} + \theta_{p3j} \log m_{jt-3} +$

$\theta_{p4j} \log p_{jt-1} + \theta_{p5j} \log p_{jt-2} + \theta_{p6j} \log p_{jt-3} + \sum_{i,i\neq j} \beta_{p4ji} p_{it-1} + \sum_{i,i\neq j} \beta_{p5ji} p_{it-2} + \xi_{jt}$

For example, if log market share has no unit root and log price has no unit root, in Equation 2.1d $\zeta_{m1j} = \beta_{m2j}$, $\zeta_{m2j} = \beta_{m3j}$, $\zeta_{m3j} = 0$, $\zeta_{m4j} = 0$, $\zeta_{m5j} = \beta_{m4j}$, $\zeta_{m6j} = \beta_{m5j}$, and $\zeta_{m7j} = 0$. However, if log market share has unit root and log price has unit root, in Equation 2.1d $\zeta_{m1j} = \beta_{m2j}+1$, $\zeta_{m2j} = -\beta_{m2j}+\beta_{m3j}$, $\zeta_{m3j} = -\beta_{m3j}$, $\zeta_{m4j} = -1$, $\zeta_{m5j} = \beta_{m4j}$, $\zeta_{m6j} = -\beta_{m4j} + \beta_{m5j}$, and $\zeta_{m7j} = -\beta_{m4j} + \beta_{m5j}$.

Appendix 2.B: Overview Selected Categories: Number of Selected Brands per Category (Percentage of Category Sales that These Brands Cover)

Beer (pilsner, lager): 6 (58.92%)
Bleach: 7 (74.60%)
Body Creams and Skin Care (body milk, lotion, and oil): 9 (36.91%)
Breakfast Cereals (oatmeal): 9 (62.11%)
Butter: 3 (50.10%)
Candy Bars (chocolate candy bars): 6 (92.62%)
Chewing Gum/Bubble Gum/Throat Drops: 11 (83.52%)
Chocolates: 11 (75.40%)
Concentrated Fruit Squash (concentrated fruit juices): 9 (75.78%)
Cooking Fats and Oils – Liquid: 8 (64.92%)
Dentifrice and Toothpaste: 11 (80.88%)
Dry Cat Food: 3 (49.39%)
Dry Dog Food: 4 (63.30%)
Fabric Conditioners – Liquid: 4 (85.95%)
Facial Cleaning Products: 9 (38.68%)
Facial Tissues: 11 (63.61%)
Flavored Carbonates (CSD’s): 8 (88.02%)
Hair Coloring Products (hair dye, color rinse): 8 (53.73%)
Hair Conditioning Products: 12 (66.91%)
Household Cleaners (liquids to clean the house): 7 (51.40%)
Household Cleaning (utensils to clean the house): 7 (39.33%)
Ice Cream: 7 (56.23%)
Infant Milk Powder: 11 (70.94%)
Instant Coffee: 9 (86.44%)
Instant Noodles: 10 (91.84%)
Laundry Soap (bars to clean clothes): 7 (79.25%)
Lavatory Cleaners (liquid to clean the toilet): 7 (71.41%)
Lemonades (non-carbonated soft drinks): 4 (24.34%)
Liquid Soap: 8 (70.18%)
Milk: 5 (61.60%)
Nappies and Diapers: 9 (77.66%)
Paper Towels: 3 (24.60%)

Potato Crisps: 8 (86.56%)
Processed Cheese (cream cheese): 5 (74.13%)
Rice: 7 (17.65%)
Salad Dressings: 3 (90.16%)
Sanitary Protection – Pads: 10 (58.15%)
Shampoo: 12 (73.24%)
Shower and Bath Additives (shower gel, bath foam): 7 (43.60%)
Soup and Bouillons – Wet (wet hot pot): 7 (42.94%)
Soy Sauces: 8 (67.35%)
Still Mineral Water: 9 (76.60%)
Sweet Biscuits (cookies): 4 (72.16%)
Tea (dry tea): 9 (58.28%)
Toilet Soap (soap bars): 9 (85.17%)
Toilet Tissues: 6 (42.25%)
Toothbrushes: 12 (64.06%)
Laundry detergent (powder detergent): 9 (92.39%)
Washing Up Liquids (liquid hand dishwashing detergent): 4 (70.08%)
Yoghurt: 5 (62.65%)

Appendix 2.C: Outcome Unit Root Tests

PANEL A: RESULTS ENDERS PROCEDURE[a]				
VARIABLE	NO UNIT ROOT		UNIT ROOT	
	Number of brands (%)	Average number of lags (SD)	Number of brands (%)	Average number of lags (SD)
log market share	243 (64.46%)	6.41 (3.76)	134 (35.54%)	10.21 (2.55)
log price	214 (56.76%)	6.35 (3.75)	163 (43.24%)	9.43 (2.70)
advertising	188 (73.73%)	8.02 (3.94)	67 (26.27%)	8.45 (4.05)
log distribution	189 (50.13%)	9.17 (3.21)	188 (49.87%)	10.75 (2.19)
log line length	162 (42.97%)	9.25 (3.17)	215 (57.03%)	10.17 (2.51)

PANEL B: CROSS TABLE UNIT ROOT LOG MARKET SHARE AND LOG PRICE				
		UNIT ROOT LOG PRICE		
		no	yes	total
UNIT ROOT LOG MARKET SHARE	no	147 (38.99%)	96 (25.46%)	243 (64.46%)
	yes	67 (17.77%)	67 (17.77%)	134 (35.54%)
	total	214 (56.76%)	163 (43.24%)	377 (100%)

[a] The appropriate lag length was selected by starting with a maximum of 13 lags and paring down the model (i.e., dropping lags) until the highest order lag in the model is significantly different from zero.

Appendix Chapter 3

Appendix 3.A: Market Shares Top Brands over Period 2011-2014

Table 3.A1: Market shares top brands in potato chips over 2011-2014

BRAND NUMBER (GB=Global Brand; LB=Local Brand)	MARKET SHARE (VOLUME)[a]			
	2011	2012	2013	2014
GB1	41.72%	45.48%	49.20%	50.78%
GB2	18.98%	22.35%	23.17%	22.16%
GB3	2.95%	2.16%	2.17%	2.59%
GB4	6.72%	5.42%	3.86%	3.26%
LB1	11.13%	11.31%	10.34%	11.02%
LB2	4.10%	2.53%	1.70%	1.48%
LB3	8.64%	5.24%	5.00%	3.98%
LB4	5.77%	5.51%	4.56%	4.73%
Sum GBs	70.37%	75.41%	78.40%	78.78%
Sum LBs	29.63%	24.59%	21.60%	21.22%

[a] Rescaled to sum to 100%.

Table 3.A2: Market shares top brands in shampoo over 2011-2014

BRAND NUMBER (GB=Global Brand; LB=Local Brand)	MARKET SHARE (VOLUME)[a]			
	2011	2012	2013	2014
GB1	7.24%	9.69%	10.63%	12.23%
GB2	20.58%	18.43%	18.86%	20.04%
GB3	2.31%	3.12%	4.27%	5.16%
GB4	7.47%	7.45%	6.26%	5.86%
GB5	15.33%	14.02%	14.70%	14.18%
GB6	3.21%	3.79%	4.45%	5.44%
LB1	5.54%	4.63%	3.55%	3.02%
LB2	4.62%	4.26%	3.14%	2.44%
LB3	27.48%	8.05%	27.51%	24.99%
LB4	6.22%	6.54%	6.64%	6.66%
Sum GBs	56.14%	56.51%	59.16%	62.90%
Sum LBs	43.86%	43.49%	40.84%	37.10%

[a] Rescaled to sum to 100%.

Table 3.A3: Market shares top brands in body creams & skin care over 2011-2014

BRAND NUMBER (GB=Global Brand; LB=Local Brand)	MARKET SHARE (VOLUME)[a]			
	2011	2012	2013	2014
GB1	13.93%	11.79%	11.11%	11.03%
GB2	10.85%	9.77%	8.61%	8.11%
GB3	9.38%	8.29%	7.04%	6.82%
GB4	7.83%	8.17%	7.89%	7.73%
LB1	27.56%	25.53%	24.82%	23.48%
LB2	7.86%	9.60%	10.20%	10.31%
LB3	10.20%	12.43%	16.24%	17.84%
LB4	12.39%	14.41%	14.09%	14.68%
Sum GBs	42.00%	38.03%	3465%	33.69%
Sum LBs	58.00%	61.97%	65.35%	66.31%

[a] Rescaled to sum to 100%.

Table 3.A4: Market shares top brands in laundry detergent over 2011-2014

BRAND NUMBER (GB=Global Brand; LB=Local Brand)	MARKET SHARE (VOLUME)[a]			
	2011	2012	2013	2014
GB1	6.31%	6.15%	6.31%	6.06%
GB2	17.33%	20.59%	22.28%	21.86%
GB3	25.89%	26.08%	25.61%	27.17%
LB1	4.85%	5.55%	5.97%	6.41%
LB2	17.22%	16.24%	15.22%	14.30%
LB3	4.43%	3.70%	3.48%	3.51%
LB4	18.15%	17.02%	17.39%	17.05%
LB5	5.82%	4.67%	3.74%	3.64%
Sum GBs	49.52%	52.82%	54.21%	55.09%
Sum LBs	50.48%	47.82%	45.79%	44.91%

[a] Rescaled to sum to 100%.

Appendix 3.B: Brand (Type) Switching Over Period 2011-2014

Table 3.B1: Brand (type) switching in potato chips over period 2011-2014[a]

PANEL A: CONSUMPTION VARIETY											
Number of top brands		0	1	2	3	4	5	6	7	8	Weighted average
% of households buying/receiving top 8 brands	Purchase	n.a.[b]	6.4%	20.6%	28.4%	25.6%	13.4%	4.5%	1.1%	.02%	3.4
	Gift	62.5%	18.2%	9.6%	5.8%	2.7%	1.1%	.2%	.0%	.0%	.7

PANEL B: RELATIVE FREQUENCY OF % OF SHOPPING TRIPS IN WHICH A HOUSEHOLD SWITCHED BRAND TYPES					
	Global Brand to Global Brand	Local Brand to Global Brand	Global Brand to Local Brand	Local Brand to Local Brand	Total
% of shopping trips where one switched brands	13.2%	14.44%	12.9%	7.8%	48.3%

[a] Note: only panelists that belong to the samples on which our models are estimated, are included.
[b] n.a. = not applicable.

Table 3.B2: Brand (type) switching in shampoo over period 2011-2014[a]

PANEL A: CONSUMPTION VARIETY													
Number of top brands		0	1	2	3	4	5	6	7	8	9	10	Weighted average
% of households buying/ receiving top 10 brands	Purchase	n.a.[b]	4.6%	16.2%	27.2%	24.8%	15.8%	7.8%	2.4%	.9%	.2%	.02%	3.7
	Gift	58.8%	19.3%	10.7%	6.2%	3.0%	1.4%	.4%	.2%	.1%	.0%	.0%	.8

PANEL B: RELATIVE FREQUENCY OF % OF SHOPPING TRIPS IN WHICH A HOUSEHOLD SWITCHED BRAND TYPES					
	Global Brand to Global Brand	Local Brand to Global Brand	Global Brand to Local Brand	Local Brand to Local Brand	Total
% of shopping trips where one switched brands	21.3%	15.5%	14.3%	5.1%	56.2%

[a] Note: only panelists that belong to the samples on which our models are estimated, are included.
[b] n.a. = not applicable.

Table 3.B3: Brand (type) switching in body creams & skin care over period 2011-2014[a]

<table>
<tr><td colspan="12">PANEL A: CONSUMPTION VARIETY</td></tr>
<tr><td colspan="2">Number of top brands</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>Weighted average</td></tr>
<tr><td rowspan="2">% of households buying/receiving top 8 brands</td><td>Purchase</td><td>n.a.[b]</td><td>9.7%</td><td>26.3%</td><td>29.6%</td><td>20.6%</td><td>9.2%</td><td>3.3%</td><td>1.2%</td><td>.1%</td><td>3.1</td></tr>
<tr><td>Gift</td><td>71.3%</td><td>18.0%</td><td>6.8%</td><td>2.4%</td><td>1.0%</td><td>.3%</td><td>.1%</td><td>.0%</td><td>.0%</td><td>.5</td></tr>
<tr><td colspan="12">PANEL B: RELATIVE FREQUENCY OF % OF SHOPPING TRIPS IN WHICH A HOUSEHOLD SWITCHED BRAND TYPES</td></tr>
<tr><td></td><td colspan="2">Global Brand to Global Brand</td><td colspan="2">Local Brand to Global Brand</td><td colspan="2">Global Brand to Local Brand</td><td colspan="2">Local Brand to Local Brand</td><td colspan="3">Total</td></tr>
<tr><td>% of shopping trips where one switched brands</td><td colspan="2">7.4%</td><td colspan="2">13.2%</td><td colspan="2">13.6%</td><td colspan="2">14.6%</td><td colspan="3">48.85%</td></tr>
</table>

[a] Note: only panelists that belong to the samples on which our models are estimated, are included.
[b] n.a. = not applicable.

Table 3.B4: Brand (type) switching in laundry detergent over period 2011-2014[a]

<table>
<tr><td colspan="12">PANEL A: CONSUMPTION VARIETY</td></tr>
<tr><td colspan="2">Number of top brands</td><td>0</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>Weighted average</td></tr>
<tr><td rowspan="2">% of households buying/receiving top 8 brands</td><td>Purchase</td><td>n.a.[b]</td><td>5.9%</td><td>21.5%</td><td>30.3%</td><td>24.8%</td><td>12.2%</td><td>4.3%</td><td>.9%</td><td>.07%</td><td>3.3</td></tr>
<tr><td>Gift</td><td>67.2%</td><td>18.8%</td><td>8.0%</td><td>3.8%</td><td>1.5%</td><td>.5%</td><td>.1%</td><td>.0%</td><td>.0%</td><td>.6</td></tr>
<tr><td colspan="12">PANEL B: RELATIVE FREQUENCY OF % OF SHOPPING TRIPS IN WHICH A HOUSEHOLD SWITCHED BRAND TYPES</td></tr>
<tr><td></td><td colspan="2">Global Brand to Global Brand</td><td colspan="2">Local Brand to Global Brand</td><td colspan="2">Global Brand to Local Brand</td><td colspan="2">Local Brand to Local Brand</td><td colspan="3">Total</td></tr>
<tr><td>% of shopping trips where one switched brands</td><td colspan="2">8.2%</td><td colspan="2">13.9%</td><td colspan="2">13.2%</td><td colspan="2">13.4%</td><td colspan="3">48.8%</td></tr>
</table>

[a] Note: only panelists that belong to the samples on which our models are estimated, are included.
[b] n.a. = not applicable.

Appendix 3.C: Level Marketing Mix Instruments Top Brands over Period 2012-2014

Table 3.C1: Level marketing mix instruments top brands in potato chips over period 2012-2014[a]

BRAND NUMBER (GB=Global Brand; LB=Local Brand)	PRICE			DISTRIBUTION			PROMOTION			LINE LENGTH[b]			ADVERTISING		
	2012	2013	2014	2012	2013	2014	2012	2013	2014	2012	2013	2014	2012	2013	2014
GB1	1.14	1.12	1.10	.91	.92	.91	.14	.12	.12	1.56	1.60	1.58	.37	.53	.84
GB2	1.11	1.07	1.07	.87	.86	.87	.10	.16	.12	.26	.28	.31	.08	.03	.03
GB3	1.35	1.30	1.29	.73	.71	.74	.03	.03	.03	.39	.40	.39	.00	.00	.00
GB4	.93	.92	.92	.86	.86	.85	.06	.10	.07	.48	.41	.48	.19	.08	.01
LB1	.82	.79	.75	.74	.68	.64	.10	.13	.09	.20	.17	.17	.0001	.00	.00
LB2	.64	.65	.66	.81	.78	.76	.04	.03	.03	.36	.44	.40	.19	.17	.04
LB3	1.00	1.00	.98	.85	.84	.82	.05	.06	.06	.52	.43	.39	.001	.02	.05
LB4	.76	.80	.79	.83	.81	.79	.08	.11	.07	.30	.27	.25	.05	.001	.00
Average GBs	1.20	1.16	1.15	.84	.83	.84	.09	.10	.09	.74	.76	.76	.15	.19	.29
Average LBs	.83	.83	.82	.82	.79	.77	.07	.08	.06	.37	.35	.33	.09	.05	.02

[a] 2011 is the initialization year, so is not shown here.
[b] In 100s.

Table 3.C2: Level marketing mix instruments top brands in shampoo over period 2012-2014[a]

BRAND NUMBER (GB=Global Brand; LB=Local Brand)	PRICE			DISTRIBUTION			PROMOTION			LINE LENGTH[b]			ADVERTISING		
	2012	2013	2014	2012	2013	2014	2012	2013	2014	2012	2013	2014	2012	2013	2014
GB1	1.13	1.14	1.09	.91	.91	.90	.04	.03	.03	1.38	1.49	1.42	.21	.29	.24
GB2	1.22	1.18	1.14	.95	.96	.95	.03	.03	.05	2.03	2.29	1.98	.20	.35	.28
GB3	1.11	1.08	1.02	.82	.84	.85	.08	.06	.04	.47	.65	.81	.07	.002	.05
GB4	.88	.87	.84	.88	.86	.86	.05	.03	.02	.70	.67	.85	.08	.06	.09
GB5	1.00	.94	.92	.94	.93	.93	.04	.04	.04	1.59	1.58	1.45	.05	.01	.0001
GB6	1.44	1.40	1.36	.84	.87	.88	.04	.03	.03	.59	.66	.68	.0002	.0001	.00003
LB1	.82	.78	.76	.81	.78	.74	.02	.02	.04	.67	.58	.48	.05	.01	.001
LB2	.71	.69	.69	.81	.79	.75	.03	.02	.02	.78	.64	.47	.00002	.00	.00
LB3	.69	.67	.64	.96	.95	.96	.05	.05	.05	1.96	1.99	1.88	.14	.13	.16
LB4	.89	.86	.81	.86	.85	.83	.03	.02	.02	1.26	1.16	1.11	.001	.0001	.01
Average GBs	1.13	1.10	1.06	.89	.89	.89	.05	.04	.04	1.12	1.22	1.20	.10	.12	.11
Average LBs	.78	.75	.72	.86	.84	.82	.03	.03	.03	1.17	1.09	.98	.05	.03	.04

[a] 2011 is the initialization year, so is not shown here.
[b] In 100s.

Table 3.C3: Level marketing mix instruments top brands in body creams & skin care over period 2012-2014[a]

BRAND NUMBER (GB=Global Brand; LB=Local Brand)	PRICE			DISTRIBUTION			PROMOTION			LINE LENGTH[b]			ADVERTISING		
	2012	2013	2014	2012	2013	2014	2012	2013	2014	2012	2013	2014	2012	2013	2014
GB1	.21	.21	.20	.90	.88	.88	.06	.06	.05	.77	.83	.82	.01	.001	.001
GB2	.07	.07	.07	.88	.86	.84	.03	.03	.04	.59	.55	.50	.002	.002	.002
GB3	.46	.48	.47	.87	.84	.83	.03	.03	.02	.93	.93	.94	.03	.03	.04
GB4	2.09	1.91	1.79	.93	.92	.91	.03	.02	.02	1.76	1.90	1.76	.16	.17	.19
LB1	.14	.15	.16	.92	.91	.90	.11	.13	.13	.30	.28	.30	.003	.01	.01
LB2	.12	.12	.13	.88	.86	.86	.07	.06	.08	.57	.56	.57	.01	.003	.003
LB3	.29	.39	.54	.88	.90	.91	.08	.06	.05	.60	.78	.83	.005	.003	.003
LB4	.09	.09	.09	.89	.88	.88	.12	.11	.12	.25	.23	.26	.0005	.001	.0001
Average GBs	.71	.67	.63	.90	.87	.86	.04	.03	.03	1.01	1.05	1.01	.05	.05	.06
Average LBs	.16	.19	.23	.89	.89	.89	.10	.09	.09	.43	.46	.49	.003	.004	.004

[a] 2011 is the initialization year, so is not shown here.
[b] In 100s.

Table 3.C4: Level marketing mix instruments top brands in laundry detergent over period 2012-2014[a]

BRAND NUMBER (GB=Global Brand; LB=Local Brand)	PRICE			DISTRIBUTION			PROMOTION			LINE LENGTH[b]			ADVERTISING		
	2012	2013	2014	2012	2013	2014	2012	2013	2014	2012	2013	2014	2012	2013	2014
GB1	1.15	1.11	1.09	.85	.83	.81	.03	.03	.03	.55	.61	.53	.22	.39	.014
GB2	.96	.93	.92	.90	.90	.88	.07	.06	.04	.87	.96	.91	.07	.09	.01
GB3	.85	.86	.85	.93	.93	.91	.05	.05	.05	1.57	1.50	1.40	.44	.33	.40
LB1	1.46	1.44	1.42	.86	.85	.85	.03	.05	.06	.36	.36	.35	.00	.00	.00
LB2	.93	.92	.91	.91	.89	.88	.05	.06	.06	1.11	.93	.78	.05	.002	.00007
LB3	.88	.87	.87	.78	.77	.73	.02	.02	.02	.69	.60	.63	.00	.003	.009
LB4	1.06	1.07	1.07	.91	.89	.87	.04	.03	.04	1.45	1.35	1.23	.02	.0002	.00004
LB5	.98	1.02	1.03	.82	.80	.77	.06	.05	.03	.32	.30	.29	.00	.00	.00
Average GBs	.99	.97	.95	.89	.88	.86	.05	.05	.04	1.00	1.02	.95	.24	.27	.14
Average LBs	1.06	1.06	1.06	.86	.84	.82	.04	.04	.04	.79	.71	.66	.02	.001	.002

[a] 2011 is the initialization year, so is not shown here.
[b] In 100s.

Appendix 3.D: Results Learning Model M1

PANEL A: PARAMETER ESTIMATES					
BRAND NUMBER (GB=Global Brand; LB=Local Brand)	BREAKFAST CEREALS	POTATO CHIPS	SHAMPOO	BODY CREAMS & SKIN CARE	LAUNDRY DETERGENT
Brand Quality					
GB1[a]	-..04 (1.11*)	0[b]	.09 (.99*)	-.71* (1.71*)	-.66* (1.97*)
GB2[a]	.31* (.48*)	.21 (.66*)	.18 (.87*)	-.35* (1.08*)	-.04 (1.44*)
GB3[a]	.	-1.09* (1.69*)	-.42* (1.76*)	-.57* (1.23*)	0[b]
GB4[a]	.	.	-.39* (1.02*)	1.21* (.91*)	.
GB5[a]	.	.	-.06 (.75*)	.	.
GB6[a]	.	.	-.29 (1.50*)	.	.
LB1[a]	.05 (1.09*)	-.35* (.92*)	-1.20* (1.86*)	0[b]	-.18 (1.18*)
LB2[a]	-.58* (1.18*)	-1.77* (1.52*)	-1.19* (1.42*)	-.14* (1.02*)	.03 (1.08*)
LB3[a]	.11 (1.58*)	-1.58* (1.42*)	0[b]	-.07 (.76*)	-.98* (2.07*)
LB4[a]	0[b]	-.52* (1.02*)	-.62* (1.24*)	-.18* (1.31*)	-.08 (1.08*)
LB5[a]	-.36* (1.49*)	-1.21* (1.05*)	.	.	-.96* (1.74*)
Brand quality uncertainty					
Log Initial variance	1[b]	1[b]	1[b]	1[b]	1[b]
Log Signal variance	1.50*	1.23*	.64*	.56*	-.18*
Risk[a]	-1.73* (.35*)	-1.05* (.24*)	-.79* (.35*)	-1.18* (.42*)	-1.79* (.71*)
Marketing mix					
Price[a]	-.53* (1.17*)	-1.44* (1.15*)	-.55* (1.11*)	-60* (.06)	-.37 (1.53*)
Distribution[a]	1.62* (.99*)	1.48* (1.35*)	1.20* (.005)	2.25* (.89)	.36 (.54)
Promotion[a]	.34 (.53)	.08 (.24)	.58 (.14)	.61* (1.81*)	1.14* (1.60*)
Line length[a]	.22 (2.53*)	.47* (.11*)	.28* (.32*)	.18 (.20*)	.39* (.57*)
Advertising[a]	-.14 (1.07*)	.28* (.63*)	-.31* (1.08*)	-1.89* (.02)	-.04 (.90*)
Number of households	2,911	4,587	4,427	2,772	5,423
Number of observations	289,933	747,496	602,240	315,808	621,928

PANEL B: RATIOS DERIVED FROM PARAMETER ESTIMATES					
	BREAKFAST CEREALS	POTATO CHIPS	SHAMPOO	BODY CREAMS & SKIN CARE	LAUNDRY DETERGENT
Global-to-Local Brand Quality Ratio: (mean Quality GBs – minimum Quality) to (mean Quality LBs – minimum Quality)	1.78†	2.16†	2.34†	.99	1.37†

[a] Mean across households; SD across households in parentheses.
[b] Parameter fixed.
* Significant at $p < .05$.
† Global significantly different from local at $p < .05$.

Appendix Chapter 4

Appendix 4.A: Calculations and Proofs How Total Brand Sales Will Change as the Overall Share of CPG Sold Online Goes up

Our key question is: How does a change in the fraction of groceries sold online influence brand sales? To see this, we rewrite equation (4.3) as follows:

$$(4.A1)\quad \frac{S^{T}_{c,b,t}}{S^{0}_{c,b,t}} = \left[\frac{S^{T}_{.,.,t}}{S^{0}_{.,.,t}}\right] * \left(\left[\frac{\frac{S^{Off}_{c,b,t}}{S^{Off}_{c,.,t}}}{\frac{S^{0}_{c,b,t}}{S^{0}_{c,.,t}}}\right] * \left[\frac{\frac{S^{Off}_{c,.,t}}{S^{Off}_{.,.,t}}}{\frac{S^{0}_{c,.,t}}{S^{0}_{.,.,t}}}\right] * \left(1 - \left[\frac{S^{On}_{.,.,t}}{S^{T}_{.,.,t}}\right]\right) + \left[\frac{\frac{S^{On}_{c,b,t}}{S^{On}_{c,.,t}}}{\frac{S^{0}_{c,b,t}}{S^{0}_{c,.,t}}}\right] * \left[\frac{\frac{S^{On}_{c,.,t}}{S^{On}_{.,.,t}}}{\frac{S^{0}_{c,.,t}}{S^{0}_{.,.,t}}}\right] * \left[\frac{S^{On}_{.,.,t}}{S^{T}_{.,.,t}}\right]\right)$$

and calculate the following derivative:

(4.A2)

$$\frac{\partial\left(\frac{S^{T}_{c,b,t}}{S^{0}_{c,b,t}}\right)}{\partial\left[\frac{S^{On}_{.,.,t}}{S^{T}_{.,.,t}}\right]} = \left(1 + (g-1) * \left[\frac{S^{On}_{.,.,t}}{S^{T}_{.,.,t}}\right]\right)\left(\left[\frac{\frac{S^{On}_{c,b,t}}{S^{On}_{c,.,t}}}{\frac{S^{0}_{c,b,t}}{S^{0}_{c,.,t}}}\right] * \left[\frac{\frac{S^{On}_{c,.,t}}{S^{On}_{.,.,t}}}{\frac{S^{0}_{c,.,t}}{S^{0}_{.,.,t}}}\right] - \left[\frac{\frac{S^{Off}_{c,b,t}}{S^{Off}_{c,.,t}}}{\frac{S^{0}_{c,b,t}}{S^{0}_{c,.,t}}}\right] * \left[\frac{\frac{S^{Off}_{c,.,t}}{S^{Off}_{.,.,t}}}{\frac{S^{0}_{c,.,t}}{S^{0}_{.,.,t}}}\right]\right) + (g-1) * (\left[\frac{\frac{S^{Off}_{c,b,t}}{S^{Off}_{c,.,t}}}{\frac{S^{0}_{c,b,t}}{S^{0}_{c,.,t}}}\right] * \left[\frac{\frac{S^{Off}_{c,.,t}}{S^{Off}_{.,.,t}}}{\frac{S^{0}_{c,.,t}}{S^{0}_{.,.,t}}}\right] + (\left[\frac{\frac{S^{On}_{c,b,t}}{S^{On}_{c,.,t}}}{\frac{S^{0}_{c,b,t}}{S^{0}_{c,.,t}}}\right] * \left[\frac{\frac{S^{On}_{c,.,t}}{S^{On}_{.,.,t}}}{\frac{S^{0}_{c,.,t}}{S^{0}_{.,.,t}}}\right] - \left[\frac{\frac{S^{Off}_{c,b,t}}{S^{Off}_{c,.,t}}}{\frac{S^{0}_{c,b,t}}{S^{0}_{c,.,t}}}\right] * \left[\frac{\frac{S^{Off}_{c,.,t}}{S^{Off}_{.,.,t}}}{\frac{S^{0}_{c,.,t}}{S^{0}_{.,.,t}}}\right]) * \left[\frac{S^{On}_{.,.,t}}{S^{T}_{.,.,t}}\right])$$

For this expression to be positive, we need that:

(4.A3) $\left(1 + (g-1) * \left[\frac{S^{On}_{.,.,t}}{S^{T}_{.,.,t}}\right]\right)\left(\left[\frac{S^{On}_{c,b,t}}{S^{On}_{c,.,t}}\right] * \left[\frac{S^{On}_{c,.,t}}{S^{On}_{.,.,t}}\right] - \left[\frac{S^{Off}_{c,b,t}}{S^{Off}_{c,.,t}}\right] * \left[\frac{S^{Off}_{c,.,t}}{S^{Off}_{.,.,t}}\right]\right) + (g-1) * \left(\left[\frac{S^{Off}_{c,b,t}}{S^{Off}_{c,.,t}}\right] * \left[\frac{S^{Off}_{c,.,t}}{S^{Off}_{.,.,t}}\right] + \left(\left[\frac{S^{On}_{c,b,t}}{S^{On}_{c,.,t}}\right] * \left[\frac{S^{On}_{c,.,t}}{S^{On}_{.,.,t}}\right] - \left[\frac{S^{Off}_{c,b,t}}{S^{Off}_{c,.,t}}\right] * \left[\frac{S^{Off}_{c,.,t}}{S^{Off}_{.,.,t}}\right]\right) * \left[\frac{S^{On}_{.,.,t}}{S^{T}_{.,.,t}}\right]\right) > 0$

Or:

(4.A4) $\left(1 + (g-1) * \left[\frac{S^{On}_{.,.,t}}{S^{T}_{.,.,t}}\right]\right)\left([BOI_{c,b,t}] * [COI_{c,.,t}] - 1\right) + (g-1) * \left(1 + \left([BOI_{c,b,t}] * [COI_{c,.,t}] - 1\right) * \left[\frac{S^{On}_{.,.,t}}{S^{T}_{.,.,t}}\right]\right) > 0$

Or:

(4.A5) $\left([BOI_{c,b,t}] * [COI_{c,.,t}] - 1\right) * \left(1 + 2(g-1) * \left[\frac{S^{On}_{.,.,t}}{S^{T}_{.,.,t}}\right]\right) > (1-g)$

where

$BOI_{c,b,t} = \left[\frac{\left[\frac{S^{On}_{c,b,t}}{S^{On}_{c,.,t}}\right]}{\left[\frac{S^{Off}_{c,b,t}}{S^{Off}_{c,.,t}}\right]}\right]$ = brand’s online index, the brand’s category sales share online vs. offline, and

$\mathrm{COI}_{c,.,t} = \left[\frac{\left[\frac{s^{On}_{c,.,t}}{s^{On}_{.,.,t}}\right]}{\left[\frac{s^{Off}_{c,.,t}}{s^{Off}_{.,.,t}}\right]}\right]$ = category's online index, the category's sales share online vs. offline.

It is easy to see that if online does not lead to overall grocery expansion or contraction (g=1), the derivative in expression (4.A4) reduces to

(4.A6) $[\mathrm{BOI}_{c,b,t}] * [\mathrm{COI}_{c,.,t}] > 1$

If g>1 (expansion), even brands for which $([\mathrm{BOI}_{c,b,t}] * [\mathrm{COI}_{c,.,t}] - 1) < 0$ can still gain from online growth. If g<1 (contraction), the condition becomes more stringent (i.e., even brands for which $([\mathrm{BOI}_{c,b,t}] * [\mathrm{COI}_{c,.,t}] - 1) > 0$ can still lose sales).

Appendix 4.B: Setup Selection Model

The selection model we use to control for endogeneity (i.e., selection bias) is estimated with a logistic regression. The probability that brand b in category c is offered on or enters the online channel in year t is given by:

(4.B1) $\mathrm{Pr}^{on}_{b,c,t} = \frac{e^{\alpha+\varphi_1 la_{bt-1}+\varphi_2 av^{off}_{bt-1}+\varphi_3 pp_{bt-1}+\varphi_4 ad_{bt-1}+\varphi_5 fb_b+\varphi_6 cx_c+\varphi_7 as_{ct-1}+\varphi_8 bu_c+\varphi_9 he_c+\varphi_{10} fr_c+\varphi_{11} pe_c+\varphi_{12} le_c+\varphi_{13} rb_b+\varphi_{14} ro_{ct-1}+\varphi_{15} po_b}}{1+e^{\alpha+\varphi_1 la_{bt-1}+\varphi_2 av^{off}_{bt-1}+\varphi_3 pp_{bt-1}+\varphi_4 ad_{bt-1}+\varphi_5 fb_b+\varphi_6 cx_c+\varphi_7 as_{ct-1}+\varphi_8 bu_c+\varphi_9 he_c+\varphi_{10} fr_c+\varphi_{11} pe_c+\varphi_{12} le_c+\varphi_{13} rb_b+\varphi_{14} ro_{ct-1}+\varphi_{15} po_b}}$

where

la_{bt-1} = Brand b's % large packs in year t-1;

av^{off}_{bt-1} = Offline availability of brand b in year t-1;

pp_{bt-1} = Price position brand b in year t-1;

ad_{bt-1} = Adstock brand b in year t-1;

fb_b = Ownership brand b (foreign vs. domestic);

cx_c = Expensiveness category c;

as_{ct-1} = Category c's assortment size in year t-1;

bu_c = Bulkiness category c;

he_c = Heaviness category c;

fr_c = Average yearly purchase frequency category c;

pe_c = Perishability category c (perishable vs. non-perishable);

le_c = Local embeddedness category c;

rb_b = Brand b is regional (yes vs. no);

ro_{ct-1} = Category rotation of brand b's category c in year t-1;

po_b = Power of brand b's manufacturer.

Based on these estimates, correction factors are calculated. If brand b in category c is available online in year t:

(4.B2) $$\mathrm{CF}_{b,c,t} = \widehat{\mathrm{pr}^{\mathrm{off}}_{b,c,t}} \ln\left(\widehat{\mathrm{pr}^{\mathrm{off}}_{b,c,t}}\right) \Big/ 1 - \widehat{\mathrm{pr}^{\mathrm{off}}_{b,c,t}} + \ln\left(\widehat{\mathrm{pr}^{\mathrm{on}}_{b,c,t}}\right)$$

and

$$\text{(4.B3)} \quad \mathrm{CF_{b,c,t}} = \left. \widehat{\mathrm{pr}^{\mathrm{on}}_{\mathrm{b,c,t}}} \ln\left(\widehat{\mathrm{pr}^{\mathrm{on}}_{\mathrm{b,c,t}}}\right) \middle/ 1 - \widehat{\mathrm{pr}^{\mathrm{on}}_{\mathrm{b,c,t}}} \right. + \ln\left(\widehat{\mathrm{pr}^{\mathrm{off}}_{\mathrm{b,c,t}}}\right)$$

if brand b in category c is not available online in year t.

Appendix 4.C: Data Descriptives and Estimation Results Selection Model

PANEL A: DESCRIPTIVE STATISTICS (INSTRUMENTS) ONLINE PRESENCE[a]					
VARIABLE	NUMBER OF OBSERVATIONS	MEAN	SD	LOWER QUARTILE	UPPER QUARTILE
Online presence ($\mathrm{pr}^{\mathrm{on}}_{\mathrm{bt}}$)	2,395 (89.85% of brand-year combinations present online)	n.a.[b]	n.a.[b]	n.a.[b]	n.a.[b]
Regional brand ($\mathrm{rb_b}$)	2,395 (18.96% of 617 brands are regional)	n.a.[b]	n.a.[b]	n.a.[b]	n.a.[b]
Category rotation ($\mathrm{ro_{ct-1}}$)	2,395	74.29	86.90	22.52	93.22
Manufacturer power ($\mathrm{po_b}$)	2,395	3.24	3.65	1.00	4.00

PANEL B: CORRELATIONS ONLINE PRESENCE AND ITS DRIVERS (NUMBER OF OBSERVATIONS: 2,395[c])																
	$\mathrm{pr}^{\mathrm{on}}_{\mathrm{bt}}$	$\mathrm{la_{bt-1}}$	$\mathrm{av}^{\mathrm{off}}_{\mathrm{t-1}}$	$\mathrm{pp_{bt-1}}$	ln ($\mathrm{ad_{bt-1}}$)	$\mathrm{fb_b}$	ln ($\mathrm{cx_c}$)	$\mathrm{as_{ct-1}}$	$\mathrm{bu_c}$	$\mathrm{he_c}$	$\mathrm{fr_c}$	$\mathrm{pe_c}$	$\mathrm{le_c}$	$\mathrm{rb_b}$	ln ($\mathrm{ro_{ct-1}}$)	ln ($\mathrm{po_b}$)
$\mathrm{pr}^{\mathrm{on}}_{\mathrm{bt}}$	1.00															
$\mathrm{la_{bt-1}}$	-0.11	1.00														
$\mathrm{av}^{\mathrm{off}}_{\mathrm{t-1}}$	0.40	-0.19	1.00													
$\mathrm{pp_{bt-1}}$	0.12	-0.01	0.10	1.00												

ln (ad_{bt-1})[d]	0.23	-0.13	0.44	0.07	1.00											
fb_b	0.16	-0.004	0.21	0.15	0.16	1.00										
ln (cx_c)	0.17	0.17	0.05	-0.001	0.13	0.17	1.00									
as_{ct-1}	0.22	0.04	0.37	0.01	0.38	-0.03	0.23	1.00								
bu_c	0.09	0.09	0.08	-0.004	0.09	-0.10	0.29	0.20	1.00							
he_c	0.09	0.04	-0.01	-0.05	0.03	-0.14	0.29	0.18	0.60	1.00						
fr_c	0.24	-0.11	0.23	0.03	0.30	-0.12	0.0005	0.40	0.08	0.18	1.00					
pe_c	-0.03	-0.07	-0.21	-0.04	-0.02	-0.01	-0.07	-0.20	-0.17	-0.06	0.24	1.00				
le_c	0.03	-0.01	0.13	0.01	0.22	-0.24	-0.05	0.47	0.10	0.34	0.41	0.05	1.00			
rb_b[e]	-0.32	0.11	-0.76	-0.14	-0.19	-0.23	-0.19	-0.15	-0.05	0.03	-0.03	0.17	0.12	1.00		
ln (ro_{ct-1})	0.04	-0.19	0.05	0.02	0.05	-0.13	-0.52	-0.16	-0.19	-0.05	0.58	0.33	0.27	0.16	1.00	
ln (po_b)	0.18	-0.11	0.34	0.12	0.14	0.33	0.07	0.07	-0.05	-0.07	-0.02	-0.20	-0.08	-0.25	-0.08	1.00

PANEL C: ESTIMATION RESULTS ONLINE PRESENCE (pr_{bt}^{on})

DRIVERS	ESTIMATE	P-VALUE[f]
Intercept	-4.08	<.0001
Brand factors		
Large packs (la_{bt-1})	.04	.8548
Offline availability (av_{bt-1}^{off})	4.02	<.0001
Price position (pp_{bt-1})	.62	.0003
	.02	.0086

Adstock $\ln(ad_{bt-1})$	.66	.0074
Brand ownership (fb_b)		
Category factors		
Category expensiveness $\ln(cx_c)$	1.00	<.0001
Assortment size (as_{ct-1})	.0002	.0219
Bulkiness (bu_c)	-.00004	.9504
Heaviness (he_c)	.04	.3863
Purchase frequency (fr_c)	.40	.0001
Perishable (pe_c)	.43	.0322
Local embeddedness (le_c)	-.42	<.0001
Instruments		
Regional brand (rb_b)	.48	.0009
Category rotation $\ln(ro_{ct-1})$	.24	.1054
Manufacturer power $\ln(po_b)$	.35	.0156
Number of observations		
Total	2,395	
Number of brands	617	
Number of categories	62	
Model fit		
Nagelkerke R^2	.48	

[a] For operationalization of the variables, see Table 4.1.
[b] n.a. = not applicable.
[c] The number of observations equal to 2,395 represents brand-year combinations of 617 brands in 62 categories used in our selection model.
[d] Log Adstock represents the log-transform of Adstock (the log of 1.00E-11 is taken for the 1,160 out of 2,395 brand-year combinations in our sample with Adstock equal to zero).
[e] Note that regional brands not necessarily need to be produced by a domestic manufacturer. Sedrin for example, is a beer brand that is mainly sold in the south of China, and is owned by AB InBev (that acquired the brand of Fujian Sedrin Brewery in 2006).
[f] Two-sided p-value are reported.

Appendix 4.D: Overview Selected Categories and Number of Selected Brands per Category

Beer (pilsner, lager)[b,g]: 7 (6)
Bleach[d]: 6
Body Creams and Skin Care (body milk, lotion, oil)[e,g]: 10 (9)
Breakfast Cereals (oatmeal)[c,g]: 8 (3)
Butter[c]: 7
Candy Bars (chocolate candy bars)[c,g]: 4 (2)
Chewing Gum/Bubble Gum/Throat Drops[c,g]: 8 (2)
Chocolates[c,g]: 11 (3)
Concentrated Fruit Squash (concentrated fruit juices)[b,g]: 9 (1)
Cooking Fats and Oils – Liquid[c,g]: 8 (3)
Dentifrice and Toothpaste[e,g]: 11 (3)
Dry Cat Food[f]: 9
Dry Dog Food[f,g]: 12 (2)
Fabric Conditioners[d,g]: 5 (2)
Facial Cleaning Products[e,g]: 9 (3)
Facial Tissues[e]: 10
Flavored Carbonates (CSD's)[b,g]: 7 (7)
Hair Coloring Products (hair dye, color rinse)[e]: 11
Hair Conditioning Products[e,g]: 12 (3)
Hairsprays[e,g]: 6 (2)
Household Cleaners (liquids to clean the house)[d,g]: 9
Household Cleaning (utensils to clean the house)[d,g]: 7 (3)
Ice Cream[c,g]: 7 (3)
Infant Milk Powder[a,g]: 11 (6)
Instant Coffee[b,g]: 8 (3)
Instant Drinking Chocolate[b]: 7
Instant Noodles[c,g]: 7 (3)
Ketchups[c]: 4
Kitchen Papers[d,g]: 4 (2)
Laundry Soap (bars to clean clothes)[d,g]: 8 (3)
Lavatory Cleaners (liquid to clean the toilet)[d,g]: 8 (3)
Lemonades (non-carbonated soft drinks)[b,g]: 5 (3)
Liquid Soap[e]: 5
Milk[b,g]: 8 (5)
Nappies and Diapers[a,g]: 9 (8)
Paper Towels[d]: 5
Potato Chips[c,g]: 9 (9)
Processed Cheese (cream cheese)[c]: 5
Razor Blades[e,g]: 2 (2)
Rice[c,g]: 8 (3)
Salad Dressings[f]: 2
Sanitary Protection – Pads[e,g]: 11 (3)
Sanitary Protection – Tampons[e]: 1
Shampoo[e,g]: 12 (12)

Shower and Bath Additives (shower gel, bath foam)[e,g]: 10 (3)
Soup and Bouillons – Wet (wet hot pot)[c]: 7
Soy Milk[b,g]: 5 (1)
Soy Sauces[c,g]: 10 (3)
Still Mineral Water[b,g]: 10 (3)
Sweet Biscuits (cookies)[c,g]: 4
Tea (dry tea)[b,g]: 7 (3)
Toilet Soap (soap bars)[e,g]: 9 (2)
Toilet Tissues[e,g]: 7 (1)
Toothbrushes[e,g]: 11 (3)
Laundry Detergent (powder detergent)[d,g]: 8 (3)
Washing Up Liquids (hand dishwashing detergent)[d,g]: 5 (1)
Wet Cat Food[f,g]: 4 (1)
Wet Dog Food[f]: 4
Window Cleaners[d]: 5
Yoghurt[c,g]: 10 (8)

[a-f] Category types: [a] baby care; [b] beverages; [c] food; [d] household care; [e] personal care; [f] pet food.
[g] Category was part of the consumer survey (number of survey brands between brackets).