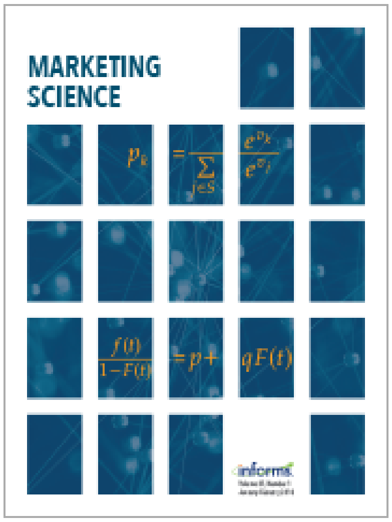


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Early Adoption of Modern Grocery Retail in an Emerging Market: Evidence from India

Vishal Narayan, Vithala R. Rao, K. Sudhir

MARKETING SCIENCE
Vol. 34, No. 6, November–December 2015, pp. 825–842
ISSN 0732-2399 (print) | ISSN 1526-548X (online)

http://dx.doi.org/10.1287/mksc.2015.0940

Early Adoption of Modern Grocery Retail in an Emerging Market: Evidence from India

Vishal Narayan
NUS Business School, National University of Singapore, Singapore 119245, bizvn@nus.edu.sg

Vithala R. Rao
Johnson Graduate School of Management, Cornell University, Ithaca, New York 14853, vrr2@cornell.edu

K. Sudhir
Yale School of Management, New Haven, Connecticut 06520, k.sudhir@yale.edu

This paper investigates early stage "modern" grocery retail adoption in an emerging market using primary household-level panel data on grocery purchases in India's largest city, Mumbai. Specifically, we seek insight on which socioeconomic class is more likely to adopt, and why. We model adoption as a two-stage process of modern retail choice followed by category expenditures within a shopping trip. We find a nonmonotonic (V-shaped) relationship between socioeconomic class and preferences for modern retail; specifically, modern retail spending and relative preference are greater among the upper and lower middle classes, relative to the middle middle class. Upper middle class preference of modern retail is driven by credit card acceptance, shorter store distance (relative to other segments), and higher vehicle ownership; whereas lower prices and low travel costs drive the preferences of the lower middle class. Modern retail is preferred more for branded and less for perishable categories. Interestingly, the lower middle class share of modern grocery retail's revenues is largest, and this share is projected to grow as prices fall and store density increases. To address concerns of endogeneity and generalizability, we replicate the key results with a "conjoint" type study with *exogenous variation* in price and distance in two cities—Mumbai and Bangalore. We discuss implications for targeting and public policy in emerging markets.

Keywords: emerging markets; retailing; segmentation; socioeconomic status; middle class
History: Received: July 5, 2012; accepted: May 20, 2015; Preyas Desai served as the editor-in-chief and Yuxin Chen served as associate editor for this article. Published online in *Articles in Advance* September 18, 2015.

> India has about 200 million households today. Our extensive study of consumption segments in India shows that only the top 14 million households have the income, attitude, and confidence to patronize modern retail.
>
> —Haden and Vittal (2008)

> Mr. Biyani [founder and chairman of Future Group, India's largest modern retailer] divides India into three types of consumers. Where he sees the greatest sales potential is among India Two—roughly 55% of Indians, with much lower incomes than India One—the top 14% of the population.
>
> —Bellman (2007)

1. Introduction

The "modern" self-service format in grocery retailing originated in the United States almost a century ago in 1916. Unlike "traditional" full-serve stores, wherein the storekeeper interacts with each customer and personally fulfills his merchandise requirements, modern stores are characterized by customers walking down aisles with carts or baskets, picking merchandise on their own and paying at checkout counters. The self-serve modern retail format typically offers wider and/or deeper assortments, because it is feasible for consumers to learn about products and make choices according to their idiosyncratic preferences, without interacting with a storekeeper.

Modern grocery retail diffused across Western Europe, Japan, and Australia in the 1950s. Since the 1980s, rapid economic growth and a growing middle class has led to substantial growth of modern retail in countries such as Korea, Taiwan, Brazil, and most recently China. However, modern retail is still at an early stage of diffusion, with less than 10% share in a large number of emerging markets (e.g., India, Indonesia, and the Philippines; see Diaz et al. 2012). Despite experience in the growth and diffusion of modern retail across many countries, there is little research on factors affecting adoption, spending, and preferences in the early stages of penetration of modern retail. Our goal in this paper is to gain empirical insight into early stage modern grocery retail adoption by the emerging middle classes as they choose between new, modern retailers and the ubiquitous traditional "mom and pop" stores.[1] To this end, we estimate a two-stage model of modern retail choice and category expenditure within a shopping trip. Our empirical setting is India, of academic interest in its own right because of its huge size and fast pace of growth.

We address three research questions about consumer adoption in the early stage of modern grocery retail diffusion. First, which segment of consumers is likely to adopt modern grocery retail in an emerging market? There is general agreement that poor consumers at the "bottom of the pyramid" do not currently have the purchasing power to support modern retail (Joseph et al. 2008). Hence, we focus on three socioeconomic segments at the top of the pyramid: the upper socioeconomic class (SEC; the very rich and the upper middle class, which, for convenience, we simply denote as the "upper" middle class), the middle middle class, and the lower middle class. But there are contrasting views, summarized in quotes at the beginning of the paper, over which of these segments drive the early growth of modern retail. Whereas the McKinsey Report (Haden and Vittal 2008) argues that modern retail will be primarily patronized by India's upper middle class, Kishore Biyani of the Future group, India's largest modern retailer, believes that the opportunity for modern retail arises from India Two, i.e., the middle and lower middle classes.

Second, we seek to understand *why* modern grocery retail shopping behavior differs across the three socioeconomic segments. We build a two-stage model of household choice of modern retail followed by category-level expenditure at modern retail for each shopping trip. We identify relevant store format attributes (e.g., prices, distances to the stores, services) and category-specific attributes (e.g., perishable/nonperishable) and study whether modern retail offers a relative advantage (e.g., lower prices). We allow for these effects to be different across distinct SECs to account for heterogeneity in attitudes and preferences toward the format and format attributes. This enables us to understand the extent to which differences in sensitivities to these attributes across segments (e.g., the upper middle class might be less sensitive to price than other segments) explain shopping decisions. So we are able to link shopping behavior to both differences in levels of attributes (e.g., prices) and differences in sensitivities to attributes across segments (e.g., lower price sensitivity). In doing this, we accommodate the fact that some households in emerging markets might be trying modern retail for the first time. We estimate the model using primary household-level and store-level panel data.

[1] In India, the number of mom and pop stores is estimated to be over 12 million, with a retail density of 10.3 stores per thousand persons. The retail density for China is 11.5 per thousand persons (Matsui et al. 2005).

Third, we apply the results to project the evolution of modern grocery retail's revenues across the three segments (upper, middle, and lower middle socioeconomic classes) as modern grocery retail expands its footprint by opening more stores (increasing access) and gains in scale and efficiency (lowering prices). These questions have implications for targeting and positioning decisions of modern retailers, the defensive strategies of traditional retailers, and more broadly in developing policy guidelines for the retail sector for policy makers.

With an estimated annual size of $450 billion, and growing at over 10% per annum (Ramola and Bhasin 2011), Indian retail is among the largest and fastest growing retail markets in the world. India is at an early stage of modern retail penetration (Sood and Jashnani 2012), making it an ideal setting to learn about early stage adoption. Furthermore, unlike developed markets, food and grocery account for as much as 70% of the Indian retail sector, making the focus on the food and grocery sector even more important in India and emerging markets.[2] Finally, the Indian economy is currently witnessing a vigorous debate on policy questions relating to the impact of regulations on the retail sector; our research is not only of managerial relevance, but can also inform the policy debate.[3]

We next elaborate on the theoretical underpinnings of our research questions. Which segments of the middle class drive early adoption of modern retail? One viewpoint is that modern retail will be initially embraced by the rich upper middle class, and the lower classes will follow.[4] There are many arguments underlying this view. First, only the rich have the money to buy the wide assortments of products offered in a modern retail store; the less affluent will spend a significant share of their incomes only on basic necessities, making the broader assortments offered by modern retail less valuable (Venkatesh 2008). Second, the rich are more likely to value the use of services such as credit card acceptance by stores, since credit card penetration is much greater among the rich. The innovation adoption and diffusion literature also suggests that the higher socioeconomic class is attitudinally more inclined and open to changing established shopping routines to

[2] See A.T. Kearney (2006). The second largest category, clothing and textiles, accounts for just 7% of retail expenditure.

[3] We note that even though our study is focused on India, by linking adoption behavior to differences in the store attributes of traditional and modern retailers, our study results can serve as plausible initial hypotheses for testing modern retail adoption behavior within the growing middle class in other emerging markets with similar institutional and socioeconomic characteristics.

[4] Karabon and Sukharevsky (2011) attribute the growth in share of modern grocery retail in Turkey from 25% to 41% in 2005–2010 to the growth in the number of high-income households, suggesting the importance of targeting high-income consumers.

shop at a new retail format (e.g., Horsky 1990, van den Bulte and Stremersch 2004).

The opposing viewpoint that the lower middle class will adopt modern retail earlier is founded on the price advantage of modern retail and the relatively low opportunity cost of travel for the lower middle classes. This view seems consistent with large-format self-serve retailing in developed markets (e.g., Walmart), which, through more efficient retail operations, can offer lower prices and is typically adopted disproportionately by the lower middle class. This view assumes that on at least some store attributes, modern retail's value proposition can be superior to that of traditional retail to address the needs of the lower socioeconomic class. Without favoring either of these views, we discuss how specific features of emerging markets impact the relative advantage of modern retail on three attributes (price, distance to store, and service) across different socioeconomic segments.

1.1. Price

In developed markets, large self-serve retailers are universally more efficient and have lower operating costs than full-serve stores, arguably because of better use of technology and management. They obtain lower wholesale prices because of greater buying power with respect to suppliers. Thus, modern retail has a relative price advantage, attracting the more price-sensitive lower middle class. Such pricing advantages are harder to realize in emerging markets because of the lack of an adequate logistics and communications infrastructure required for effective supply chain management. The regulatory environment with restrictions on transportation across states raises the cost even further. Locating in the richer neighborhoods of large urban cities also raises rental expenses (or investments in real estate) for modern retailers. Finally, traditional retailers may not pay taxes, and pay below minimum wages to employees (if any), potentially further increasing the relative cost of modern retail (Kohli and Bhagwati 2015, Sudhir and Talukdar 2015). For these reasons, modern retail might start with a higher total cost of operation than traditional retail. So whether modern retail is able to attract the lower middle class by offering lower prices in emerging markets is an empirical question we seek to address.[5]

1.2. Distance to Store

The consumer's cost of shopping at a store is not just the retail price; it includes the cost of traveling to the store. Given opportunity costs of time, consumers generally prefer stores that are closer to their homes (Reynolds 1953, Singh et al. 2006, Briesch et al. 2009). But modern retail, with its larger self-serve format, requires larger sales volumes to break even, and therefore has lower density relative to traditional retail in emerging markets. Consequently, consumers will generally have to travel farther to shop at modern retailers. Travel costs can have ambiguous effects on which socioeconomic class is more likely to adopt modern retail. The upper middle class has greater opportunity costs of time (Blattberg et al. 1978), and therefore is *less* likely to shop at modern retailers than the lower middle class. Yet, two features of emerging markets suggest otherwise. First, the upper middle class typically has household help, which makes their opportunity costs lower or irrelevant (Goldman et al. 1999). Second, the lower middle class in emerging markets do not have personal transportation (unlike in developed markets, where car penetration is close to universal); hence, their effective opportunity costs can be greater than for the upper middle class.

1.3. Service

Services offered by the store affect household choice of which format to shop. Owners of traditional stores typically have deeper relationships with their customers and might use this knowledge to offer preferential service such as store credit, more generous return policies, etc. Given their lower labor costs (typically with unemployed family members or low wage workers), traditional retailers are also able to provide home delivery.[6] In that sense, traditional retailers can provide superior customer service, relative to modern retailers. However, traditional retailers are less technologically sophisticated and less likely to accept credit cards. The need for these services could differ across socioeconomic classes. Home delivery might be more important to households without private transportation or domestic help—the lower middle class. Store credit might also be more important among poorer segments. So the importance of service-related attributes might vary across socioeconomic segments. The net effect of service on adoption by SEC is therefore an empirical question.

1.4. Category Characteristics

Consumers' shopping behavior can depend on specific characteristics of the categories of products they wish

[5] In China, where the government provides land to retailers at favorable locations at subsidized rates to encourage modern format diffusion, land costs might be favorable to modern retailers.

[6] India has an estimated one retail store for every 98 people; hence, most retailers typically know their regular customers. Moreover, these storeowners are part of a social fabric with which shoppers have social communications. Social interactions are an important motivator for some consumers to visit retailers (Evans et al. 1996), and remembering a shopper's name facilitates adherence to purchase requests made by the person who remembers (Howard et al. 1995). Greater interpersonal communication leads to greater perceived levels of relationship investments, which might affect store loyalty (De Wulf et al. 2001).

to buy (Inman et al. 2009). Certain categories are associated with specific store formats, and these associations can vary across segments (Inman et al. 2004). We focus on the effects of (i) the level of branding in a category and (ii) the perishability of the category on retail format choices and expenditures.

Modern stores can carry more brands (or a greater proportion of branded merchandise) than traditional stores, and hence may appeal more to the relatively more quality-conscious upper middle class. On the other hand, it is possible that, given the assurance of quality through brands, consumers may be indifferent between purchasing at modern or traditional retailers in branded categories.

Perishable categories (e.g., vegetables) need to be purchased more frequently than nonperishables. Because modern retail tends to be farther than traditional stores, consumers may prefer traditional stores to purchase perishable categories. On the other hand, modern retailers may be able to better store perishables and deliver, on average, better quality perishables to consumers, in which case consumers may prefer modern retail for perishables.

Given the trade-offs above, the preference for modern retail adoption in branded and perishable categories is an empirical question. Furthermore, the three segments may respond to these trade-offs differently. We therefore allow differential effects across segments for these category-format preferences.

Beyond the qualitative direction of the effects of prices, distances, services, and category characteristics across SECs on modern retail adoption, it is also important to quantitatively assess the relative importance of each store attribute across segments on trip choice and spending. For instance, what would be the revenue increase by reducing the travel distance to each segment by 20%, or by reducing prices by the same proportion? Such quantitative insights can aid managers in deciding the relative importance of improving access by opening more stores or improving supply chain management, which helps lower prices.

Given our research questions, we use socioeconomic status to segment middle class consumers into "upper," "middle," and "lower" income classes. The use of socioeconomic status for segmentation in the marketing literature is well established (Coleman 1983); it has been used in marketing (Dahl and Moreau 2007, Rucker and Galinsky 2008), psychology (Rogler 1996, Jayakody et al. 1998), and development economics (Sumarto et al. 2007). It is also widely acknowledged by practitioners in India as the most prevalent and useful segmentation approach across industries. SEC segmentation classifies households into segments depending on the occupation and education of the chief wage earner of the household. SEC A comprises the very rich and upper middle class; SEC B comprises the middle middle class; and SEC C comprises the lower middle class. Other SECs comprise the "bottom of the pyramid" and, as discussed earlier, are excluded from our analysis. SECs are preferred over just income-based segmentation, because occupation and education shape not just an individual's earning capacity but also family self-image and social status and set the tone and tenor of how they live (Bijapurkar 2008). Furthermore, SEC predicts the consumption of various goods better than income alone (Bijapurkar 2008).[7] Given the relatively low penetration of modern retail in this early stage of diffusion, we only analyze SECs at the first level of granularity (A, B, C) and abstract away from the more granular subdivisions in this paper.[8]

In studying shopping decisions across retail formats, we focus only on the modern/traditional dichotomy. More granular levels of store format classification are possible for modern retail stores. For example, retailers experiment with larger and smaller retail formats (e.g., supermarkets and hypermarkets). However, given the low levels of penetration of modern retail as a whole, there are very few store visits to specialized formats; for example, in our data, only 2.3% of store trips are to hypermarkets. We believe at the early stage of diffusion, insights at the broad level of categorization (i.e., modern and traditional) are more valuable.[9]

A major challenge in answering questions about modern retail adoption in emerging markets is the absence of readily available scanner data as in developed markets. We therefore conduct two field studies to collect primary data on grocery shopping behavior from households and marketing mix data from grocery stores in India. After describing the model in §2, we describe the first study in §3, along with the results we obtain from it. We present the second study in §4. Section 5 concludes our paper.

2. Model

We model the household's patronage of modern retail as a two-stage decision process within each shopping trip. In the first stage, the household decides whether to visit a modern retail store or not on the trip. This process is modeled as a probit model. This decision

[7] As Kamakura and Mazzon (2013) discuss, (a) stated income measures are prone to large reporting errors (Hentschel and Lanjow 2000), (b) current income might be a poor indication of consumption potential for retired individuals, and (c) consumption correlates poorly with income since people smooth their consumption over time by borrowing or drawing on savings during times of low income and investing/saving in times of high income.

[8] We provide additional details of the SEC classification system for urban India in Online Appendix 1 (available as supplemental material at http://dx.doi.org/10.1287/mksc.2015.0940).

[9] In our econometric analysis, we include fixed effects for different store formats within modern retail, but our main results are not affected by such disaggregation.

is based on category-level factors (which categories are to be purchased, their prices, level of perishability, etc.), store format-level factors (e.g., distances), and household characteristics. In the second stage, conditional on the choice of modern retail for the trip, the household decides on the level of expenditure within each category, with the possibility of not purchasing in a category. We model this second stage involving choice of expenditures across multiple categories at a modern retail store as a multivariate Tobit model. We account for correlations in preferences across the two stages and across categories at the household level.

In terms of notation, let s denote store format. It can take two values, "modern" ($s = M$) or "traditional" ($s = T$). We denote the probability that household i ($i = 1, \ldots, N$) chooses a modern store in trip t as $\text{Prob}(I_{iMt} = 1)$, where I_{iMt} is an indicator variable equal to 1 if modern store is chosen and 0 otherwise. Conditional on choosing a modern store in trip t, let E_{icMt} be the expenditure on modern retail of household i for category c ($c = 1, \ldots, C$). The household needs to overcome two thresholds to spend on each category in a trip to a modern store: a common threshold related to the process of choosing modern retail on a focal trip, and a category-specific threshold for the process governing spending in that category at a modern retailer.

First, consider trip-specific format choice. Let u_{ist} be the utility of household i from shopping from store format s in trip t. Let $U_{it} = u_{iMt} - u_{iTt}$ be the difference in utility between modern ($s = M$) and traditional ($s = T$) formats. We specify this difference as follows:

$$\begin{aligned} U_{it} = {} & \lambda_{i0} + S_{it}\beta_0 + H_i\beta_1 + H_{it}\beta_2 + \Delta D_{it}\beta_{3i} \\ & + \sum_{c=1}^{C} I_{ict}\bar{Q}_{ic}\Delta p_{ct}\beta_{4ic} + \sum_{c=1}^{C} I_{ict}\bar{Q}_{ic}a_c\beta_{5i} \\ & + \sum_{c=1}^{C} I_{ict}\bar{Q}_{ic}b_c\beta_{6i} + \varepsilon_{it}. \end{aligned} \tag{1}$$

When $U_{it} \geq 0$, a modern store is chosen, and when $U_{it} \leq 0$, a traditional store is chosen. The error term ε_{it} is distributed standard normal. The term λ_{i0} captures the effects of unobserved household format-specific characteristics (e.g., differences in exposure levels of household i to advertising across the two store formats). The term S_{it} is an indicator of trial of modern retail by household i prior to trip t. This household-specific variable can capture the effect of prior awareness and exposure to modern retail on current format choice. We operationalize S_{it} as follows: it takes the value 1 if household i has chosen modern retail at least once before trip t in our panel data or responded "yes" to an initial survey prior to our panel data collection in which they were asked if they had visited a modern retail grocery store in the previous three months.[10]

The term H_i is a vector of household characteristics such as vehicle ownership, credit card ownership, presence of domestic help, respondent gender, and apartment size; these can affect the relative value associated with modern retail formats. H_{it} is a vector of household trip-specific needs or reasons for store format choice (store ambience, product variety, product quality, home delivery, store credit, and personal relationships with storeowners). Such variables are not commonly available to researchers and help to model for intrahousehold heterogeneity in needs across shopping trips, which is typically abstracted or absorbed into the logit/probit error in choice models. We later describe how we measured them. ΔD_{it} is a vector of differences in store format characteristics that are invariant across categories: difference in distance to the two store formats from the residence of household i and a dummy variable to account for diversity in modern retail formats (1 if hypermarket, 0 if other format). Difference in distance is given by $(d_{iMt} - d_{iTt})$, where d_{iMt} is the distance traveled to modern retail by household i in trip t (if it was to a modern store) or the distance traveled during the most recent trip to a modern store (if trip t is to a traditional store). d_{iTt} is defined likewise.[11] For a household that never visited modern retail in the data window, it is defined as the mean distance traveled to modern stores by other households of the same SEC.

Next we discuss how price enters the format choice model. We denote by Δp_{ct} the difference in price indexes for category c in trip t across formats. The price index for category c in trip t for format s is the mean price per unit quantity across stockkeeping units (SKUs) in a store and across all stores in format s. The category price difference between the two formats is relevant for a household only if the household purchases the category in trip t (e.g., the price of meat will always be irrelevant for a vegetarian household). So we weigh the category price difference by the trip-specific indicator for category purchase incidence I_{ict} (1 if category c is purchased in trip t, 0 otherwise). This indicator is not format specific. Also, consider two households with 1 and 10 members, respectively, both of whom buy category c. To the extent that a larger household

[10] Only 5.2% of households in our sample began trial of modern retail during our 10-week panel data collection period. Hence, this variable remains stable at 1 or 0 for most households in our analysis.

[11] Our measure of $d_{iMt}(d_{iTt})$ varies across trips, since a household chooses different modern (traditional) stores in different trips. Our results are robust to two alternative specifications: (a) replacing the distance traveled during the most recent trip with the mean distance across all trips and (b) replacing d_{iMt} for households that never visited modern retail with the mean distance traveled to modern stores by *all* other households, not just households of the same SEC.

is likely to buy more of category c, the price of category c is likely more important for this household than for the single-member household. So we further weigh the category price difference by $\bar{Q}_{ic}$, the average quantity per trip (across all trips) of category c that the household i purchases, conditional on purchase. The terms a_c and b_c refer to measures of the level of perishability and extent of branding of category c, respectively. Similar to price, we weigh these variables by purchase incidence and average purchase quantity for that household in the particular trip.

Finally, consider the category expenditure on modern retail, conditional on choice of modern retail in trip t. Let E^*_{icMt} be the associated latent variable, interpretable as the utility from the expenditure E_{icMT}. It relates to the observed expenditure as per a multivariate Tobit specification

$$E_{icMt}=0 \quad \text{if } E^*_{icMt}\le 0, \quad E_{icMt}=E^*_{icMt} \quad \text{if } E^*_{icMt}>0, \quad (2)$$

where $E^*_{icMt}=\bar{E}_{icMt}+\varepsilon_{icMt}$, $\varepsilon_{icMt}=[\varepsilon_{i1Mt},\varepsilon_{i2Mt},\ldots,\varepsilon_{iCMt}]$ $\sim \text{MVN}(0,\Sigma_1)$, and

$$\begin{aligned}\bar{E}_{icMt} &= \lambda^*_{ic}+S_{it}\delta_0+H_i\delta_1+H_{it}\delta_2+\Delta D_{it}\delta_{3i}\\ &\quad +I_{ict}\bar{Q}_{ic}\Delta p_{ct}\delta_{4ic}+I_{ict}\bar{Q}_{ic}a_c\delta_{5i}+I_{ict}\bar{Q}_{ic}b_c\delta_{6i}. \quad (3)\end{aligned}$$

We estimate how expenditure of modern retail for each category is affected by the same set of variables as with format choice in Equation (1), with the exception that we include differences in price indexes, level of perishability, and extent of branding of category c (and not of all categories). Differences in price indexes, level of perishability, and extent of branding of category c are weighed by category purchase incidence (I_{ict}) and average category quantity purchased ($\bar{Q}_{ic}$) across all trips—not just those to modern retail. The error terms are distributed multivariate normal, allowing for all category expenditures to be correlated due to both observed covariates and unobservables. The component λ^*_{ic} serves to control for category-specific and individual-specific unobserved factors. The effects of trial of modern retail, household characteristics, trip-specific needs, and distances to stores of different formats are assumed to be identical across categories.

Next we write the joint likelihood of the data. Consider an observation where household i does not spend on category c in trip t from modern retail, i.e., $E_{icMt}=0$. There could be two mutually exclusive possibilities why this would happen, each with different likelihoods. First, household i decides to shop from traditional retail during trip t. The likelihood of this event is $1-\text{Prob}(I_{iMt}=1)$. The other possibility is that household i decides to shop from modern retail in trip t, but decides not to buy category c in this trip. The likelihood of this event is given by $\text{Prob}(I_{iMt}=1)\text{Prob}(E_{iMt}=0)$. So the likelihood of observing a zero category expenditure on modern retail in an observation is given by

$$\begin{aligned}L(E_{icMt}=0) &= 1-\text{Prob}(I_{iMt}=1)\\ &\quad +\text{Prob}(I_{iMt}=1)\text{Prob}(E_{icMt}=0). \quad (4)\end{aligned}$$

Now consider an observation where the expenditure of household i on category c in trip t from modern retail is E^*_{icMt} such that $E^*_{icMt}>0$. For this to happen, the household should have chosen modern retail in that trip. This likelihood is given by

$$L(E_{icMt}=E^*_{icMt})=\text{Prob}(I_{iMt}=1)\text{Prob}(I_{icMt}=E^*_{icMt}). \quad (5)$$

2.1. Heterogeneity, Correlation Across Stages and Across Categories

Given our interest in understanding heterogeneity across SECs, we specify the household-specific coefficients as a function of SECs, and allow them to be correlated due to unobservables, as follows:

$$\begin{aligned}\theta_{1ic} &= \left[\beta_{4ic}\ \ \delta_{4ic}\right]=\Pi_1\left[1\ \ SECA_i\ \ SECB_i\ \ a_c\ \ b_c\right]+\varepsilon_{ic};\\ \theta_{2i} &= \left[\lambda_{io}\ \ \lambda_{ic}\ \ \beta_{3i}\ \ \beta_{5i}\ \ \beta_{6i}\ \ \delta_{3i}\ \ \delta_{5i}\ \ \delta_{6i}\right]\\ &= \Pi_2\left[1\ \ SECA_i\ \ SECB_i\right]+\varepsilon_i \quad \text{where}\\ &\qquad \varepsilon_{ic}\sim\text{MVN}(0,\Sigma_2) \quad \text{and} \quad \varepsilon_i\sim\text{MVN}(0,\Sigma_3). \quad (6)\end{aligned}$$

$SECA_i$ and $SECB_i$ are dummy variables indicating membership of SEC A and SEC B respectively, and Π_1 and Π_2 are matrices of parameters. We allow for category-specific price sensitivity parameters β_{4ic} and δ_{4ic} to vary with both household characteristics and category characteristics. This allows us to estimate how price sensitivity varies across SECs and across categories with different levels of perishability and branding. The self-stated household trip-specific needs (H_{it}) capture not only heterogeneity in preferences across households but also within households for trips; hence, we do not model heterogeneity in parameters for H_{it}.[12]

2.2. Model Estimation

The model is estimated on data for all trips, both modern and traditional. The likelihood of observations of zero category expenditure (Equation (4)) allows for the possibility that the observation pertains to a trip to a traditional store, or that it pertains to a trip to a modern store but with no purchase of

[12] Estimation of a more general model that allowed the coefficients of self-stated needs to vary across SECs revealed that the effects of stated needs did not vary significantly across SECs. Conceptually, the household needs data we collected are the top three reasons (out of a list of 12 possible reasons) for store format choice. Since all three stated reasons are relatively very important for the respondent, we did not expect to find SEC-level differences in levels of their importance.

the category. All other observations have positive category expenditure at modern retailers, with their likelihood given by Equation (5). We use Bayesian estimation to estimate the parameters. This involves specifying proper but diffuse prior distributions for the model parameters, and then deriving their posterior conditional distributions. Given the set of conditional distributions and priors, we draw recursively from the posterior distribution of the model parameters. We use the Metropolis–Hastings algorithm (random walk) for drawing from the posterior distribution of parameters without conjugate priors (i.e., posterior distributions without closed-form expressions). The exact estimation procedure is described in Online Appendix 2.

3. Empirical Analysis

3.1. Data

There are several data challenges in understanding grocery shopping behavior of customers in emerging markets. Unlike developed markets, most stores (including modern retailers) do not electronically capture shopper-level data on visits and transactions. Traditional stores do not maintain *any* transaction records (electronic or otherwise). Even store-level data on product assortments and prices are not collected. Moreover, given the low penetration of modern retail and the large number of traditional retailers, an accurate picture of the shopping behavior requires recording of household purchases across large numbers of stores. Even though a small number of (traditional) stores may capture a significant share of a household's share of spending, each traditional store itself caters to only a small number of households. Hence, one needs information across a large number of retailers to accurately capture the behavior of consumers and the shopping environment (e.g., price levels) within even a relatively small urban geographical area. We partnered with Nielsen India to leverage their expertise in on-the-ground data collection procedures in India and their localized data audit and validation processes.

Given our interest in understanding differences in modern retail adoption across different SECs, we randomly sampled households until we reached a quota of 90 households for each of the three SECs. These households resided in an 80 square mile radius in suburban Mumbai, which accounts for a population of approximately 4 million (about 29% of the city population). We note that this market is substantially larger than typical cities in the United States, which have been studied using scanner panel data. The interviewer briefed an adult respondent from each household on the differences between modern and traditional retailers before administering the survey, to ensure understanding. For each household, we collected the following data for all trips to stores for grocery purchase (along with home deliveries and visits by domestic help) made during a 10-week period starting March 16, 2011. Data for each trip included (1) date of the trip; (2) chosen store format; (3) category-level purchase quantity, price, and expenditure; (4) distance from home to chosen store; and (5) reasons for choosing the store on that trip (detailed later in this section). To prevent errors associated with household self-classification of the store format as modern or traditional, we required store names to be recorded and later verified by Nielsen India to ensure proper classification. Stores classified as modern have checkout counters, provide aisle space to enable shoppers to touch all displayed merchandise, and provide computerized receipts of purchase.

We collected data for eight product categories, which Nielsen identified based on their proprietary research as prototypical categories for (a) frequently/infrequently purchased items, (b) staple items consumed every day, and (c) categories that were significant in terms of expenditures in the consumption basket. These are onions, tomatoes, unbranded rice, branded rice, biscuits, cooking oil, toilet soap, and non-alcoholic beverages. Whereas onions and tomatoes are frequently purchased food categories, rice is a staple consumption item in Mumbai. Cooking oil and toilet soap are infrequently purchased. But cooking oil and rice account for a significant proportion of food expenditure. Furthermore, both modern and traditional food and grocery retailers typically stock all of these categories, allowing us to study share of expenditures from a consumer choice perspective, without supply constraints imposed on consumers.

We used a diary-based method. Households were instructed to fill in details for a trip immediately after shopping, to minimize recall bias. To cross-check data quality, Nielsen field personnel visited each household twice after the beginning of the data collection to inspect the diaries and to verify reported data on categories purchased with purchase receipts (wherever available) and stock in household pantries (wherever permitted). Households were reminded by frequent telephone calls to complete the diaries as instructed. Owing to attrition, we finally obtained panel data from 266 households, who reported having made 7,965 trips to stores, of which 478 were to modern retailers.

Nielsen field personnel manually collected weekly data on prices and assortments for the eight categories from 20 stores (10 modern and 10 traditional) during the same 10-week period where we collected household data. This involved personnel visiting each of the 20 stores each week and listing all SKUs available for sale for each category and their pack sizes, selling prices, and promotional details, if any. To ensure our store sample reflected the stores where our households shopped, respondents were asked to indicate the names of all stores from which they had purchased food and groceries during the past month. This list was collated across respondents to obtain a census of stores in the area, and 20 stores were sampled conditional on their acceptance of Nielsen personnel collecting price data from their store. Since traditional stores in emerging markets can be quite heterogeneous in terms of store size, Nielsen ensured that different types of traditional stores were adequately represented in the sample of 10 traditional stores.

To obtain category-level price indices for each store format, we followed the procedure of Briesch et al. (2009). For each store in the sample, we adjusted the price for each SKU to account for price promotions, if any (e.g., a 10% price reduction).[13] The adjusted price for each SKU (per unit quantity) in a category was then averaged across SKUs to obtain a store-level price index for that category. Store-level price indexes across all traditional stores in the sample were averaged to obtain a category-level price index for this store format. This procedure was repeated weekly for both store formats, so that price indexes are measured at a weekly level. We define a category as perishable if its shelf life is less than a month. Given the set of categories in our data, the two vegetable categories (onions and tomatoes) are perishable ($a_c = 1$) and other categories are not ($a_c = 0$). Furthermore, onions, tomatoes, and unpackaged rice are sold as unbranded commodities ($b_c = 0$) by weight in India. All other categories are available as branded items ($b_c = 1$).

We collected data on 12 possible trip-specific reasons for choosing a retail format:[14] price, distance, product quality, store ambience, product variety, home delivery, store credit, relationship with the storekeeper, "knowledge and courtesy of the storekeeper," "knowledge and courtesy of the store employee," "return policy," and "air conditioning." We asked respondents to indicate which of these 12 were among their top three reasons for choosing a retail format for each trip. Of these 12 reasons, we did not include the following as covariates in the model because less than 1% of trips indicated these as being among the top three: knowledge and courtesy of the storekeeper, knowledge and courtesy of the store employee, return policy, and air conditioning. Of the remaining, we did not include price and distance as covariates, because we used objective measures that we collected on those. The remaining six reasons were included as covariates in the vector H_{it}. We coded a reason as 1 for trip t if it was indicated as one of the top three reasons for that trip by household i and 0 otherwise.

[13] Price promotions were rare in this market in 2011; only 1.6% of all SKUs across all eight categories were promoted in an average week.

[14] We began with a list of 22 possible reasons for store format choices, based on a survey of the store choice literature and in consultation with Nielsen. After pretesting with shoppers in Mumbai, we reduced the list to 12 reasons in the final survey.

Table 1 Distribution of Visits to Modern Retail

Proportion of store trips to modern retail (%)	Proportion of households (%)			
	All	SEC A	SEC B	SEC C
0	49.3	45.1	58.6	44.3
0.1–5	19.2	17.6	16.1	23.9
5.1–10	12.0	12.1	9.2	14.8
10.1–15	7.9	12.1	5.8	5.7
Above 15	11.7	13.2	10.3	11.3

Notes. Each cell in this table represents the proportion of households in a segment who have a certain proportion of store trips to modern retail in the first study. For example, 19.2% of all households chose modern retail stores in 0.1%–5% of all store trips.

3.2. Descriptive Analysis

We find that modern grocery retail accounted for only 6% of the 7,965 trips. Only 50.7% of households tried modern retail at least once in our data.[15] The share of trips to modern retail is highest for SEC A at 7.1%, but lowest for SEC B at 5.2%. SEC C is in the middle at 5.7%. Similarly, trial rates across SECs A, B, and C are 54.9%, 41.4%, and 55.7%, respectively. More details on the distribution of share of visits to modern retail can be found in Table 1. Overall, this suggests a V-shaped relationship between modern retail adoption and socioeconomic status.

The relationship extends to expenditure shares and absolute levels of category expenditures too. Modern retail accounted for 8.2% of total expenditure (across both formats) in the eight categories, with the breakdown across SECs A, B, and C being 9.6%, 5.9%, and 9.0%, respectively.[16] We find that for six out of eight categories, SEC B spent less on modern retail in the data period than SECs A and C (Table 2), in both absolute monetary terms and in terms of share of total retail expenditure. The ratio of the number of households in SECs A, B, and C in Mumbai is 1:1.57:1.86.[17] Based on this distribution, we estimate that 42.7% of modern retail revenues for these eight categories come from SEC C consumers, followed by 29.2% from SEC A consumers, making SEC B the least revenue generating segment of the population.

[15] To alleviate the initial conditions issue (i.e., households who never tried modern retail in 10 weeks might have tried it earlier), we conducted an exploratory cross-sectional survey prior to panel data collection. Each respondent indicated whether they had visited a modern store in the past three months.

[16] Prior to collection of the panel data, Nielsen identified 21 major categories that accounted for a large majority of the shopping baskets of urban Indian households. We measured category purchase incidence from modern retailers as a response to whether the category was "generally" purchased from a modern retailer. Category purchase incidence from modern retail was 33.2% for SEC A, followed by 23.2% for SEC C, and only 19.8% for SEC B. We infer that the V-shaped relationship is generalizable to the entire shopping basket.

[17] http://articles.economictimes.indiatimes.com/2004-01-05/news/27381554_1_nrs-national-readership-survey-council-sample-size.

Table 2 Mean Category Expenditure (Mean Share) per Household on Modern Retail (Across 10 Weeks, in Rupees)

Category	All households	SEC A	SEC B	SEC C
Onion	3.32 (7.3%)	3.49 (6.7%)	1.39 (4.3%)	5.03 (11.7%)
Tomato	3.17 (7.1%)	4.23 (6.4%)	1.21 (4.1%)	4.03 (11.5%)
Cooking oil	87.07 (10.5%)	118.51 (8.4%)	55.08 (7.3%)	86.18 (13.8%)
Beverages	45.53 (12.1%)	66.01 (11.1%)	32.75 (6.6%)	36.98 (16.0%)
Loose rice	22.36 (5.3%)	27.16 (7.4%)	11.54 (3.6%)	28.10 (7.2%)
Biscuit	18.72 (6.7%)	22.14 (10.2%)	15.49 (4.1%)	18.37 (8.2%)
Toilet soap	19.82 (4.9%)	19.86 (9.7%)	22.11 (3.9%)	17.51 (3.2%)
Packaged rice	45.77 (11.3%)	44.96 (17.7%)	47.87 (8.2%)	44.52 (8.1%)
All 8 categories	245.75 (8.2%)	306.35 (9.6%)	187.44 (5.9%)	240.73 (9.0%)

Note. Figures in parentheses represent the share of expenditure on modern retail (as a proportion of total retail expenditure).

Table 3 Mean, Minimum, and Maximum Price Indexes in Modern and Traditional Retail (Across 10 Weeks, in Rupees/Unit)

	Traditional stores			Modern stores			
Category	Mean	Min	Max	Mean	Min	Max	Difference in means (t-stat)
Onion	12.26	10.19	13.11	8.73	6.41	9.19	6.8***
Tomato	12.63	9.71	13.66	10.96	8.79	12.45	3.0***
Cooking oil	100.48	91.36	110.73	89.09	78.05	101.43	6.6***
Beverages	51.33	47.08	62.89	46.3	44.18	48.06	4.3***
Loose rice	46.49	44.65	49.62	47.09	46.04	48.15	−0.5
Biscuit	121.09	117.84	126.77	131.71	122.08	139.54	−3.1***
Toilet soap	240.62	232.89	265.06	256.57	221.02	274.70	−1.7*
Packaged rice	78.46	73.46	82.84	86.58	78.1	95.57	−3.8***

Note. For each store format, the mean price index is averaged across weeks. All price indexes are in rupees per kilogram except those for beverages and cooking oil, which are in rupees per liter. At the time of data collection, USD 1 = INR 52, approximately.
$^{*}p < 0.1$; $^{***}p < 0.01$.

Next consider the role of prices and distances (Table 3). We find that prices at modern retailers are lower in only four of the eight categories (onions, tomatoes, cooking oil, and beverages), suggesting that modern stores do not offer uniformly lower prices across all categories. For distance, as expected, the average distance traveled to a traditional retailer is only 0.62 km, and does not vary much across segments (0.65 km, 0.60 km, and 0.62 km across SECs A, B, and C, respectively). By contrast, the average distance traveled to modern retail is much higher at 2.71 km (2.11 km, 2.79 km, and 3.12 km across SECs A, B, and C, respectively).

Table 4 summarizes the stated reasons why households chose trip traditional or modern retail in each trip. Whereas price, product variety, and product quality favor modern retail, distance to the store and relationship with the storekeeper favor traditional retail. SEC C is more often likely to cite price and store credit relative to other SECs. Households differ in their stated needs across time, indicating significant intrahousehold heterogeneity; the mean share of the most cited reason was only 62%. Furthermore, only 21.1% of the variation in format choice is driven by across-household variation in choice, whereas 78.9% of the variance in the choice of modern retail comes from within-household variation. Within-household variation accounts for 50.6%–72.5% of the variation in trip expenditures.[18]

3.3. Model Comparison and Endogeneity

We assess the value of two features of our model: (1) allowing for the processes of modern retail choice and category expenditure to be correlated and (2) allowing for two different processes for choice of modern retail and category expenditure. To do so, we compare the performance of our model (Model 1) with two benchmark models. In Model 2, we do not allow for the two processes to be correlated, that is, the off-diagonal elements of Σ_1, Σ_2, and Σ_3 are assumed to be zero. The second benchmark model (Model 3) is a pure multivariate Tobit model of category expenditure on modern retail specified by Equation (2), with the heterogeneity distributions of its individual-level parameters specified by Equation (6), where we abstract away from the first process modeling choice of modern retail. Like the

[18] We regressed household trip-specific format choice (I_{iMt}) as a probit on the household-level mean (I_{iM}) of format choices; the variance accounted for by I_{iM} is the across-household variance. Similarly, we use the variance explained by household-level mean expenditure in category (E_{icM}) to explain across-household variance in trip expenditure at modern retailers in the category (E_{icMt}). More details of the regression and the variance decomposition are presented in Online Appendix 3.

Table 4 Stated Reasons for Trips

	Proportion of trips for which a need was cited (by store format) (%)		Proportion of trips for which a need was cited (by socioeconomic class) (%)		
	Traditional store trips	Modern store trips	SEC A	SEC B	SEC C
Price	7.9	58.6	5.9	10.0	16.5
Distance	30.8	15.9	30.7	37.6	22.1
Store ambience	0.2	5.2	0.7	0.7	0.1
Product variety	2.9	28.7	4.6	6.3	2.6
Product quality	6.6	35.4	7.4	7.6	9.9
Home delivery	7.9	11.3	6.2	7.7	10.2
Store credit	6.9	5.0	4.8	6.3	9.0
Relationship with storekeeper	27.2	2.3	19.0	29.5	29.3

Notes. Values are based on the top three stated reasons (out of 12 possible listed in §3.1) for choosing a store format. We do not report numbers for four of the reasons that were rarely chosen.

Table 5 Model Comparison

	Model 1 (two-stage model with correlation across stages)	Model 2 (two-stage model without correlation across stages)	Model 3 (single-stage model)
Log marginal density	−3,129.3	−3,140.4	−3,171.8
DIC	6,263.2	6,274.1	6,296.7
RMSE (holdout sample)	1.96	1.97	2.13

proposed model, this model is estimated on data for all trips, both traditional and modern. To compare model performance, we employ the log marginal density and the deviance information criterion (DIC; Spiegelhalter et al. 2002). The DIC takes model complexity into account by penalizing additional model parameters and generalizes the Akaike and Bayesian information criteria for hierarchical modeling. Models with smaller values of the DIC are preferred. We find that both measures favor the proposed model (Table 5). To compare out-of-sample predictive power, we classified 797 randomly selected trip-level observations (approximately 10% of the data) as the holdout, estimated the model on the remaining data, and predicted category expenditure at modern retailers for all observations in the holdout sample. Root mean squared error (RMSE) for the holdout sample for the proposed model is the lowest. Model performance increases more due to modeling the two-stage process than due to allowing for the processes to be correlated.

Next we discuss the issue of price endogeneity, i.e., the possibility that category prices in our model might be correlated with unobserved factors that affect store format choices (Equation (1)) and category expenditures (Equation (3)), potentially leading to biased estimates of price coefficients in these equations. For example, retail prices may reflect store unobservables such as availability of car parking, longer store hours, etc. To account for price endogeneity emanating from store-specific unobservable factors, we employ instruments in the spirit of Berry et al. (1995), Sudhir (2001), and Nevo (2001). Specifically, we replace the price index of format s and category c in week t with the mean (across stores of format s) of all store-level prices for that category-week except for the store that was visited by the focal household which we use as an instrument. This instrument is correlated with the category price level of the visited store (due to competition given substantial overlap in assortments across stores), but uncorrelated with the error term (since unobservables in the error term for a store visit would pertain to the visited store only). We find little difference in the estimates by employing these instruments. Estimates without using instruments are available from the authors.[19]

3.4. Parameter Estimates

We report the results of the proposed model in Table 6. We use the estimates to discuss why some SECs are more likely to choose and spend more on modern retail.

3.4.1. Role of Prices. The price coefficient is negative for both trip choice and category expenditure (Table 6). Hence, lower prices offered by modern retail will increase trip choice and category expenditure. The interaction effects of price with SEC A and SEC B are

[19] Another issue is whether retailers set prices for categories taking into account segment-specific preferences for different formats in categories and whether this may cause endogeneity bias on the price coefficient. Given that we include household-category intercepts for trip-category expenditures at modern retailers and household intercepts in the format choice model, segment preferences are accounted for in the demand model at the category level. We therefore do not have the potential for endogeneity bias on this count. We thank a reviewer for encouraging us to clarify this issue.

Table 6 **Effects of Price Differences Between Modern and Traditional Retail (Π_1): Posterior Means and SDs**

	Trip-specific choice of modern retail (Equation (1))	Category expenditure on modern retail (Equation (3))
Main effect of price difference	**−0.940 (0.028)**	**−2.305 (0.472)**
Interactions effects of price difference		
With SEC A	**0.386 (0.062)**	**1.073 (0.278)**
With SEC B	**0.163 (0.049)**	**0.866 (0.385)**
With perishability (a_c)	**−0.157 (0.052)**	**−0.370 (0.116)**
With branding (b_c)	0.006 (0.010)	0.037 (0.028)

Note. Parameter estimates whose posterior 95% credible interval does not include 0 are in bold.

both positive; hence, SEC C households are the most price sensitive. The higher price sensitivity of SEC C, coupled with lower prices of modern retail, explains the greater proportion of modern store trip choices and category expenditures by SEC C households.

3.4.2. Role of Distance. The distance to store affects the trip-specific format choice, but not expenditures at the modern format (Table 7). These results are intuitive: the difference in travel distance between modern and traditional stores affects whether to make a trip to a modern or traditional store; once in the store it should not affect expenditures. The interaction effect of SEC and distance shows that distance affects trip-specific choices for both SEC A and SEC B, given their higher opportunity cost of time, relative to SEC C. SEC B not only has a slightly greater sensitivity to distance than SEC A but also has to travel, on average, 32% farther than SEC A to shop from modern grocery retail. Thus, modern retail is most disadvantaged in terms of distance among SEC B. Although SEC C has to travel even longer distances to modern retail, it is less sensitive to distance.

3.4.3. Category Characteristics. We find that traditional stores are preferred for perishable categories (Table 6) over modern stores. This is consistent with the argument that given the greater frequency with which perishables have to be purchased, traditional retail will be preferred because of closer proximity to households. We also find that SEC A prefers to visit traditional stores even more for perishables relative to other SECs, further supporting the opportunity cost of shopping time argument.

By contrast, respondents choose modern stores for branded categories more often and spend more at modern stores on these categories. This is consistent with the argument that modern retail, which can usually store a greater variety of brands, is preferred in branded categories. Furthermore, SEC A has the greatest preference for expenditure on modern retail in branded categories. This further supports the variety argument, because SEC A might prefer variety more than other SECs. This result is robust to a household-level measure of the extent of branding of a category: the proportion of all SKUs purchased by household i in category c, which are branded.

3.4.4. Other Attributes. Vehicle ownership is positively associated with modern retail choice (Table 8). Because vehicle ownership is much more common in SEC A households (64.3% of all households that own cars are SEC A), vehicle ownership makes modern retail more favorable to SEC A. Similarly, credit card ownership, which is more prevalent among SEC A, has a positive effect on modern retail choice. By contrast, domestic help, more likely to be used by SEC A, suppresses trip-specific modern retail choice, reflecting the unwillingness of the lower classes to visit modern retail for each trip. But conditional on choosing modern retail for a trip, domestic help has no impact on expenditures.

Table 7 **Effects of SEC, Distances, and Category-Specific Factors (Π_2): Posterior Means and SDs**

	Trip-specific choice of modern retail (Equation (1))			Category expenditure on modern retail (Equation (3))		
	Main effect	Interaction with SEC A	Interaction with SEC B	Main effect	Interaction with SEC A	Interaction with SEC B
SEC A	**0.327** (**0.310**)	—	—	−14.415 (17.323)	—	—
SEC B	−0.156 (0.114)	—	—	**−10.492** (**3.311**)	—	—
Distance difference	**−0.751** (**0.105**)	**−0.238** (**0.083**)	**−0.257** (**0.094**)	−0.834 (0.491)	−0.940 (0.663)	−0.952 (0.719)
Hypermarket (1 if yes)	0.011 (0.010)	0.003 (0.003)	0.004 (0.003)	0.009 (0.011)	−0.001 (0.004)	−0.002 (0.003)
Perishability (a_c)	**−0.017** (**0.003**)	**−0.004** (**0.002**)	−0.003 (0.002)	**−0.083** (**0.029**)	**−0.016** (**0.005**)	0.008 (0.007)
Branding (b_c)	**0.014** (**0.003**)	0.005 (0.003)	0.004 (0.003)	**0.063** (**0.019**)	**0.017** (**0.003**)	0.019 (0.011)

Note. Parameter estimates whose posterior 95% credible interval does not include 0 are in bold.

Table 8 **Effects of Trial (S_{it}), Household Characteristics (H_i), and Trip-Specific Needs (H_{it}): Posterior Means and SDs**

	Trip-specific choice of modern retail (Equation (1))	Category expenditure on modern retail (Equation (3))
Household characteristics (β_1 and δ_1)		
Vehicle ownership	**0.204 (0.060)**	−9.978 (11.841)
Domestic help	**−1.761 (0.422)**	1.258 (2.902)
Credit card ownership	**1.894 (0.648)**	−14.436 (18.719)
Apartment size	−0.055 (0.052)	−4.781 (6.409)
Gender (1 if male)	0.263 (0.137)	−1.920 (1.563)
Trip-specific needs (β_2 and δ_2)		
Store ambience	0.020 (0.051)	6.804 (6.359)
Product variety	0.090 (0.076)	**21.734 (9.576)**
Product quality	**0.094 (0.041)**	13.570 (9.751)
Home delivery	−0.042 (0.031)	9.396 (10.918)
Store credit	−0.005 (0.016)	−1.447 (6.482)
Relationship with storekeeper	**−0.154 (0.067)**	1.956 (2.284)
Trial of modern retail (β_0 and δ_0)	**1.093 (0.068)**	**51.417 (12.403)**

Note. Parameter estimates whose posterior 95% credible interval does not include 0 are in bold.

Home delivery, store ambience, and store credit do not seem to affect trip-specific choice of modern retail. Product variety leads to greater expenditure, but does not affect trip choice. Strong relationships with traditional retailers prevent modern retailers from gaining high shares. Overall, these results provide guidance on which variables modern retailers should focus on in their competition with traditional retailers. We discuss detailed managerial implications in §5.

A key conclusion is that SEC A and SEC C households both patronize modern retail relatively more than SEC B households. SEC C households are attracted by lower prices, and do not mind traveling longer distances as much as the other SEC households. This suggests the possibility that modern retail might be differentiated such that within modern retailers, stores offering high prices attract SEC A households, and stores offering low prices attract SEC C households. We analyzed average prices per week across all SKUs stocked by all modern retailers and found that variance in prices across modern retailers is not high. This suggests that in these early stages of modern retail in India, price-based differentiation within modern retailers is not very significant.

We find that SEC C households are most price sensitive, suggesting that they might cherry-pick, i.e., buy only those categories that are cheaper from modern retailers. On studying the share of modern retail expenditure across categories and SECs, we find that SEC C consumers do not restrict their spending only to cheaper categories. Specifically, out of the total expenditure on modern retail (on the eight categories) by SEC C consumers, 45.1% is on the four categories that are priced at least as high at modern retailers (loose rice, toilet soap, biscuits, and packaged rice). This suggests that modern retail might be able to strategically employ a loss leader strategy by reducing prices of some "key value items" such as onions, tomatoes, and cooking oil. A more detailed examination of the profitability of each segment would require margin data—an issue we suggest be addressed in future work.

4. Robustness Check: Endogeneity and Generalizability

Next we assess the robustness of our results to two concerns. The first is endogeneity in retailer choices; i.e., there could be some unobserved variables that affect retailers' decisions about marketing mix variables (prices, distances, etc.) and consumer choices of store format that can potentially bias parameter estimates. Although we employ instruments for prices, instruments for distance are unavailable. The second is generalizability. Our data are from households in suburban Mumbai, India's largest city, covering an 80 square mile radius and four million people. Even though this market is much larger than traditional markets studied using U.S. scanner panel data, the generalizability of our results to other parts of India is unclear. To alleviate the endogeneity and generalizability concerns, we conducted a follow-up conjoint study in July 2013 in Mumbai and in a second city, Bangalore.

4.1. Conjoint Data

To address endogeneity concerns in our first study, we collected "conjoint" choice data on retail format choice from 741 consumers (roughly split across the three SECs) where we exogenously varied the marketing mix variables. To address the generalizability concern, the new data cover two cities: Mumbai ($N = 368$), where we did our primary study, and Bangalore ($N = 373$). Bangalore—widely known as India's Silicon Valley—has a high concentration of affluent SEC A consumers and the greatest penetration of modern grocery retailers in urban India. It is therefore interesting in its own right. Respondents for this study were screened by asking individuals if they (a) shopped for grocery requirements for their household and (b) made the most trips in their household to buy groceries. Only those individuals who responded "yes" to both questions were included in the final sample, so that they are knowledgeable and provide meaningful answers to the subsequent questions. All such respondents stated that they had shopped for groceries at least four times in the previous month.

Our goal in the conjoint task was to test how consumers choose store format when there is exogenous variation in distance to store, regular price level, and price promotion. Because the size of the shopping list can affect store format choice (e.g., modern stores preferred for infrequent stock-up, and traditional stores for more frequent fill-ins), we also varied the number of items on the shopping list.

In each conjoint choice task, respondents were asked to place themselves in a situation in which they had to shop for groceries from a shopping list for their household. In each choice task, the respondent was provided with one randomly selected short or long shopping list described below.

- Shopping list 1: 2 kg of onions and five packets of biscuits of 200 g each.
- Shopping list 2: 2 kg of onions, five packets of biscuits of 200 g each, 5 kg of rice, 500 ml of cooking oil, and three bars of toilet soap.[20]

The respondent was then presented with a description of two stores (one modern and one traditional) and asked to choose which they would prefer for their (randomly assigned) shopping list. We described each store in terms of three attributes: distance of the store from where the respondent lives, price of each category to be purchased, and promotions. Although providing information on more attributes would better mimic real-life shopping, we decided on employing fewer attributes but more choice tasks, so we could control for unobserved heterogeneity across respondents, while making robust inferences on the two key attributes of greatest interest—price and distance.

For the traditional store, the distance attribute was manipulated across two levels: 0.2 km or 0.8 km. For the modern store, these levels were set at 1.7 km or 2.7 km. This ensures that the mean distance to each store format mimics the mean travel distance reported in Study 1 (the panel data). Since the effect of absolute prices on store choice is not identified in a binary choice model, prices at the traditional store were held constant (over choice tasks), and the price difference across the two stores was manipulated. Specifically, prices at the traditional store were fixed (based on 2011 market prices) as follows: 100 rupees for 5 kg of rice, 20 rupees for 2 kg of onions, 50 rupees for five packets of biscuits, 50 rupees for 500 ml of cooking oil, and 70 rupees for three bars of toilet soap. Category prices at the modern store were fixed at a price x% lower than prices at the traditional store for all categories, where x can take one of four values: 0%, 10%, 20%, or 30%.[21]

Promotions were manipulated by offering a discount (none or 40%) on the price of rice at the modern store over the price of rice at the traditional store. For choice tasks where the modern store offered a promotion, rice was offered at a 40% discount, irrespective of the regular price difference. Promotions were made distinct from regular prices by labeling the promoted price "special promotion price." Finally, although we did not label any store as traditional or modern, we did provide the following textual descriptions:

- *Traditional store*. "In this store you would speak with a storekeeper, who would bring the groceries you need to you."
- *Modern store.* "In this store you would find and pick your own groceries and check out yourself."

We now discuss how this experimental setup alleviates the endogeneity problem. In the real world, consumers make store format choices based on several attributes, some of which are observed by researchers, whereas others are not. Correlation between the observed endogenous attribute (e.g., price) and unobserved attributes that affect consumer decisions could potentially lead to biased estimates of the effect of the endogenous attribute. In the experimental setup, we can control the attributes that subjects consider to make decisions. In each choice task of our study, we instructed respondents to assume that any information that was not provided to them about the two stores was invariant across the two stores. Assuming that subjects adhered to this instruction, there is no unobserved factor in this experimental setup that varies across store formats and influences store format choice. The absence of unobserved factors rules out the possibility that observed attributes are systematically correlated with unobserved factors, thus alleviating any bias due to endogeneity.

We constructed 10 sets of 16 choice tasks such that the three store-specific attributes (distance, prices, and promotions) are orthogonal to each other and to the number of categories being purchased. We assigned each respondent randomly to one of the 10 sets. By randomizing the presentation order of the 16 choice tasks for each respondent, we eliminate the possibility of order bias. We asked questions on demographics before the choice tasks. We assumed the attributes to be interval scaled to conserve degrees of freedom.

4.2. Results

We estimated probit models of choice of modern retail with randomly distributed individual-specific coefficients to assess whether the following results from the

[20] Nielsen research suggests that the shorter shopping list is representative of fill-in trips in that market, and the longer list is representative of a stock-up trip.

[21] Our manipulations do not allow for lower distances and greater prices at modern stores. Our conjoint design was intended to understand how variations in distance and price differentials with respect to modern retail affect format choices, conditional on keeping the relative sign of advantage/disadvantage of modern retail constant. One disadvantage of this design choice is that our conjoint results are not generalizable to situations where prices are greater at modern stores or distances to modern stores are lower.

first study are robust when faced with exogenously manipulated prices, distances, promotions, and number of purchase categories: (a) the nonlinear relationship of modern retail adoption across the three segments, (b) the greater price sensitivity of lower SECs, and (c) the greater distance sensitivity of upper SECs. The first model (Model 1) includes the following covariates: dummies for SECs A and B, difference in weighted prices between modern and traditional store, difference in distance between modern store and traditional store (in kilometers), a dummy variable for city, the number of categories purchased (two or five), and a dummy variable for promotion. Consistent with the model proposed for Study 1, the difference in weighted prices is given by $\sum_{c=1}^{C} Q_{ict} \Delta p_{ict}$, where Δp_{ict} is the difference in unit price of category c across store formats in choice task t faced by respondent i, and Q_{ict} is the corresponding quantity purchased. Although $\sum_{c=1}^{C} Q_{ict} \Delta p_{ict}$ is a function of the number of categories purchased, we include the number of categories as a separate covariate to capture any residual effect of fill-in versus stock-up trips on modern store choices. Similarly, the promotion dummy captures any transaction utility from promotion in addition to the effect of lower price. The vector of individual-specific coefficients of price, distance, the number of categories, and promotion are distributed as follows: $\beta_i \sim \text{MVN}(\bar{\beta}, \Sigma)$. Other covariates do not vary within respondents. Their coefficients are assumed invariant across respondents. Table 9 presents the parameter estimates.

The negative coefficient of SEC B again provides robust evidence of the nonlinear relationship of modern retail adoption across SECs. Respondents in Bangalore were more likely to choose a modern retail store, perhaps because of greater familiarity and exposure, and consistent with greater usage of the format in real life. The coefficient of number of categories is not significant, suggesting that number of purchase categories affect modern retail choice only to the extent that buying more categories leads to greater monetary savings if modern retail offers lower prices. Similarly, the effect of promotion is sufficiently captured by accounting for its effect on price. Although price, promotions, and number of categories purchased might be endogenous in store choice models estimated with market data, parameter estimates in this study are not prone to such biases.

To investigate how sensitivities to prices and distances vary across segments, we extended the previous model to allow for individual-specific coefficients of price and distance to vary across SECs. We label this Model 2. Consistent with Study 1, we find that SEC C respondents are most sensitive to price and least sensitive to traveling longer distances. This provides convergent evidence from a bigger sample in a larger geography and a different time period that differences in modern retail adoption across SECs are driven by differences in sensitivities to prices and distances. One somewhat surprising finding is the statistically insignificant coefficient for distance, suggesting that SEC C respondents are insensitive to differences in distances between modern and traditional stores. Although this result could be driven by limited degrees of freedom (only 16 observations per respondent) and limited variation in the manipulated distances in the conjoint study, it is consistent with the notion that SEC C consumers can be effectively targeted by modern retail by focusing on lowering prices without making substantial investments in more stores to increase access. We also estimated this model separately for each city and found consistent results. This suggests that the patterns of price and distance sensitivities across SECs hold despite significant differences in aggregate level penetration across cities. Overall, we conclude that our results from the primary study are robust.

We note that purchase data in both studies are limited to eight narrowly defined categories, which potentially cover a small share of the grocery shopping

Table 9 **Factors Affecting Modern Store Choice in Study 2: Posterior Means and Posterior SDs**

Covariate	Model 1 (Mumbai and Bangalore)	Model 2 (Mumbai and Bangalore)	Model 3 (Mumbai only)	Model 4 (Bangalore only)
SEC A	−0.043 (0.032)	0.237 (0.161)	0.284 (0.197)	0.180 (0.182)
SEC B	**−0.090 (0.032)**	0.315 (0.194)	0.328 (0.201)	0.304 (0.223)
City (1 if Mumbai, 0 if Bangalore)	**−0.597 (0.024)**	**−0.615 (0.027)**	—	—
Promotion (1 if yes)	0.173 (0.098)	0.260 (0.186)	0.413 (0.238)	0.175 (0.207)
Number of purchase categories	0.030 (0.021)	−0.058 (0.046)	−0.097 (0.057)	0.008 (0.054)
Price difference (modern − traditional)	**−0.033 (0.001)**	**−0.036 (0.001)**	**−0.033 (0.002)**	**−0.038 (0.002)**
Distance difference (modern − traditional)	**−0.069 (0.018)**	−0.031 (0.041)	−0.036 (0.079)	−0.027 (0.041)
Price × SEC A		**0.010 (0.001)**	**0.008 (0.002)**	**0.012 (0.002)**
Price × SEC B		**0.008 (0.002)**	**0.005 (0.003)**	**0.010 (0.004)**
Distance × SEC A		**−0.116 (0.051)**	**−0.124 (0.053)**	**−0.108 (0.051)**
Distance × SEC B		**−0.157 (0.063)**	**−0.129 (0.064)**	**−0.171 (0.065)**

Note. Parameter estimates whose posterior 95% credible interval does not include 0 are in bold.

basket for most households. To understand whether our results extend to the entire grocery shopping basket, we collected additional survey data using broader category definitions that cover the entire basket. The share of modern retail continues to be lower among SEC B relative to SEC A and SEC C, even when considering the entire grocery shopping basket. Details are provided in Online Appendix 4.

5. What-If Analysis and Implications

5.1. What-If Analysis

Our results show differences in modern format patronage among the different socioeconomic classes. Our key result is that SEC B has lower patronage than SEC A and SEC C. Would the differences persist or attenuate as modern retail expands store density and gain in efficiency, leading to lower prices? To this end, we conduct a what-if analysis using model estimates.

First, as modern retail increases store density, it improves accessibility for consumers relative to traditional retail by reducing the distance that consumers have to travel to the store. Over time, store expansion leads to greater economies of scale and greater bargaining power with respect to suppliers. Furthermore, improvements in back-end and logistics infrastructure over time also lead to lower costs. Overall, these lead to lower relative prices for modern retail. We predict changes at the SEC level as the market evolves.

We consider three scenarios: (1) prices alone fall, (2) travel distances alone fall, and (3) both prices and travel distances fall.[22] Specifically, we simulate revenue shares at current prices and distances, and the decline of 20% of modern retail prices and distances relative to traditional retail prices and distances. Figure 1 presents the projected weekly revenue shares for the eight categories by SEC under these scenarios. The share of revenues of a segment is simply the ratio of the expenditure at modern retailers by all households in the catchment area in this segment to the expenditure at modern retailers by all households in the catchment area.[23]

As mentioned earlier, we find that SEC C provides the largest share of modern retail revenues, at 42.7%. As prices fall, SEC C gains in share, given the greater price sensitivity of this segment. If prices were lowered by 20%, the proportion of modern retail revenue derived from SEC C would jump from 42.7% to 51.6%. This is due to greater price sensitivity of SEC C households. The effect of travel distance, however, is greater on SEC A, given its greater opportunity cost. The effect, however, is much smaller than that of price reductions. The proportion of modern retail revenue from SEC A would jump from 29.2% to 30.6% if distances to modern retail were decreased by 20%. Finally, the reality is likely to be a combination of declines in prices and travel distances. In combination, we find that a substantial increase in revenue comes from SEC C households. If prices and travel distances both decrease by 20%, the share of modern retail revenue from SEC C is estimated at 50.4%. We conclude that not only is SEC C the major contributor of current modern retail revenues, but the importance of SEC C for modern retail will continue to increase over time. Next we discuss the implications of these and other results for retailers and policy makers.

5.2. Implications for Retailing and Policy

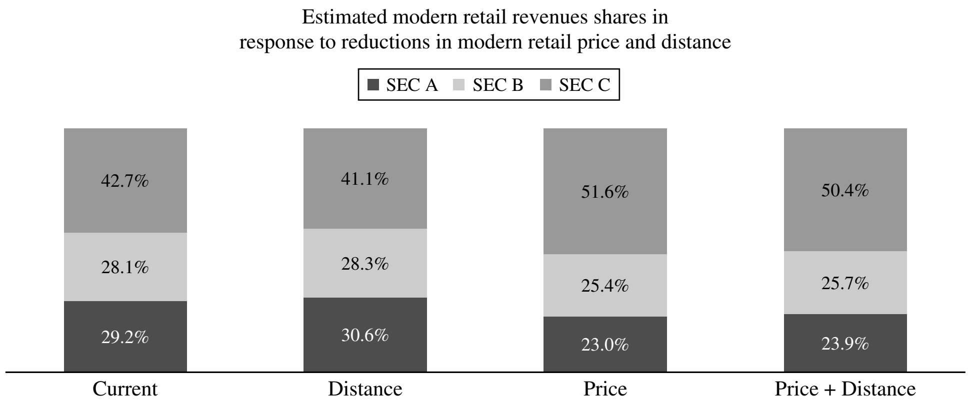
Our results have implications for modern retailers, traditional retailers, and regulators. First, from a targeting standpoint, SEC C seems more attractive than SECs A and B for modern retailers. Greater modern retail adoption by SEC C households and a larger proportion of such households in the population make them the greatest revenue generating segment for modern retail. Furthermore, the share of revenue from SEC C is projected to increase in the face of lowering prices. Overall, these results suggest that Mr. Biyani's focus on India Two (Bellman 2007), from the lower socioeconomic classes, appears to be a good bet for modern retailers.

Second, SEC C consumers are most price sensitive and least sensitive to travel. So in terms of the "four P's," price is therefore more important than place in expanding revenues for modern retail within their largest share segment. Modern retailers could attract this segment by offering lower prices, without making large investments in opening more stores to improve access. Third, SEC A consumers are attractive since they are the least price sensitive and spend substantially more on high-priced categories (Tables 1 and 2). Hence, the current modern retail strategy of opening stores closer to SEC A neighborhoods appears to be the right strategy. Given our estimates of credit card and vehicle ownership, modern retail should accept credit cards and offer convenient vehicle parking to attract SEC A households. Last, beyond targeting specific SECs, modern retailers can boost revenues from existing consumers by investing more in reducing prices and offering more variety than on improving store ambience and offering potentially expensive services such as home delivery and store credit. More generally, targeting different segments, whether based on SEC (A, B, or C) or not, requires focusing on different

[22] Although it is typical for prices (at modern retail stores) and travel distances (to modern retail stores) to fall as this format diffuses, there is limited understanding of how prices of specific categories might evolve. We assume the same price difference across all categories for this analysis.

[23] We estimate revenues per household for each SEC and multiply those with estimates of the number of households in each SEC in the catchment area. The population density of Mumbai is 53,000 per square mile (http://censusindia.gov.in/).

Figure 1 **What-If Analysis of Impact of Lowering Prices and Opening Modern Retail Stores on Revenue Share Across SECs**



Note. The revenue share of an SEC is the ratio of the expenditure at modern retailers by all households of this SEC in the catchment area in this segment to the expenditure at modern retailers by all households in the catchment area.

elements of the retailer's marketing mix. Our results provide insights into these elements.

From the point of view of traditional retailers, the middle middle class is the most desirable middle-class segment and will continue to remain so. Distance to modern stores looms larger for this segment, and the price advantage of modern retail is not attractive enough given the distance disadvantage. Also, they tend to value the relationships with the store most—an area that traditional retailers have a distinct advantage, relative to the somewhat impersonal nature of the modern retail shopping experience.

From a policy standpoint, our results show that the value of modern retail is not restricted to the rich and elite, but to lower middle classes as well—a larger constituency from a political perspective. In recent years, there has been vociferous opposition to the expansion of modern retail, and to foreign direct investment that is necessary to support the cash intensive sector. However, there has been little discussion of the benefits of modern retailers to consumers. For several categories, SEC C consumers enjoy the benefits of lower prices at modern retailers. Therefore, the lack of consumer support for modern retail is not because its benefits accrue to only the rich elite, as much conventional wisdom in this area suggests, but because of a lack of political mobilization on behalf of consumers.

6. Conclusion

Rapid globalization over the last two decades has led to increased growth and rising incomes in many major economies across Asia, Eastern Europe, Latin America, and Africa (Burgess and Steenkamp 2006, Sheth 2011). The resultant growing middle class in these countries has served as an impetus for the entry and growth of modern retail in many markets. Despite the large size of this middle class, there is little research documenting how subsegments of the middle class differ in their adoption behavior toward new retailing formats. This paper contributes to the literature on retailing/marketing in emerging markets by addressing the following questions of interest to managers, researchers, and policy makers: First, which segment of the growing middle class is more likely to adopt modern retail? Second, why does modern retail adoption differ across socioeconomic segments? Third, as modern retail expands its footprint by opening more stores (increasing access) and gains in scale and efficiency (lowering prices), what will the relative shares of the three segments be of modern retail's revenues?

A vast stream of literature in international marketing has studied market entry strategies *across* countries, yet there has been little work on market entry strategies *within* a country. When making targeting decisions, firms could adopt a hierarchical (or waterfall) strategy of new product introduction across countries, wherein firms introduce products sequentially from one country to another. The alternative is a simultaneous (or sprinkler) strategy of targeting, wherein firms introduce products in several countries at the same time. Researchers have studied market conditions favoring each model (Kalish et al. 1995, Tellis et al. 2003). Irrespective of when a firm enters a country, it needs to decide which consumer segment(s) in that country to focus its resources on. In this paper, we shed light on the issue of whether firms should enter a country by targeting one segment first followed by other segments (akin to the waterfall strategy), or whether firms should target several segments (akin to the sprinkler strategy) simultaneously. Our results favor a simultaneous targeting strategy. We find that modern retail penetration is nonlinear in socioeconomic class. The upper and lower middle classes, however, value different elements of modern retail. Our findings underscore the importance of socioeconomic status as an important segmentation variable to understand middle class consumption behavior in emerging markets.

We conclude with a discussion of certain limitations of the studies that suggest possibilities for further research. Given the extensive and costly primary panel data collection, by necessity, we restricted our data and analysis to two cities. To the extent that factors like credit card ownership, distance to modern retail, price sensitivity, etc., are similar in other markets across the upper, middle, and lower middle class segments, we believe our results are potentially generalizable to other new emerging markets and also make theoretical sense. But future research needs to assess the robustness of our findings with additional data across other cities and countries. Furthermore, food and grocery categories are characterized by low search behavior; it remains unclear whether our results are generalizable for shopping in categories involving greater search, such as jewelry and apparel. It will also be important to study the distinctions in retail adoption behavior across rural and urban markets. Urban and rural retailers differ significantly in their cost structures because of the differences in the intensity of distribution, lower store density, etc. These differences are not merely important from a consumer adoption perspective, but also from a supply side perspective. The second study can be extended to see whether greater prices at modern retail affect store format choices, and to study differences in purchase behaviors across more categories.

We study consumer choices between modern and traditional retail. As the market develops, it might be fruitful to study choices *within* various modern retail formats (e.g., large hypermarkets versus small supermarkets). Also, socioeconomic classifications are static. With longer time series of data, the effect of increasing incomes and purchasing power of households on modern store adoption could be studied. Last, online retailing is a rapidly growing channel in several emerging markets, especially in second-tier cities, where it is not cost-effective to set up a physical presence. The effect of this channel on the patronage of brick-and-mortar retail formats might be another interesting avenue of research. From a methodological standpoint, longer time series of consumer choice data might enable researchers to study dynamics in choice behavior, e.g., learning, loyalty, and habit formation. Finally, future work can relax our modeling assumption that the category purchase decision is unaffected by store format choice.

In summary, our research is a first step to begin a systematic study of modern retail adoption in emerging markets that aids with address targeting and positioning questions related to modern retail. We hope our approach of (a) filling the gap in availability of secondary data by collecting primary panel data from consumers and stores and (b) augmenting behavioral panel data with experimental field data to alleviate endogeneity concerns will make feasible new research on other substantively important questions, leading to a strong conceptual and empirical foundation for a deeper understanding of these explosive high-growth markets.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mksc.2015.0940.

Acknowledgments

The authors thank Ahmed Khwaja, Jiwoong Shin, Subrata Sen, Baohong Sun, and participants at the 2011 Marketing Science Conference, the 2012 China India Insights Conference, the 2012 Marketing Science in Emerging Markets Conference at Wharton, and marketing seminars at the City University of Hong Kong, Emory University, Hong Kong University of Science and Technology, Indiana University, Indian School of Business, Nanyang Technological University, National University of Singapore, Singapore Management University, and University of Connecticut for helpful comments on this paper. The authors are grateful to Cornell's Institute for Emerging Markets, the Yale China India Insights Program, and the Marketing Science Institute for their generous funding to conduct this research.

References

A.T. Kearney (2006) Retail in India: Getting organized to drive growth. Report, Confederation of Indian Industry, Haryana, India. https://www.atkearney.com/consumer-products-retail/ideas-insights/featured-article/-/asset_publisher/KQNW4F0xInID/content/retail-in-india-getting-organized-to-drive-growth/10192.

Bellman E (2007) In India a retail finds key to success is clutter. *Wall Street Journal* (August 8).

Berry S, Levinsohn JL, Pakes A (1995) Automobile prices in market equilibrium. *Econometrica* 63(4):841–890.

Bijapurkar R (2008) *Winning in the Indian Market: Understanding the Transformation of Consumer India* (John Wiley & Sons, Singapore).

Blattberg R, Buesing T, Peacock P, Sen S (1978) Identifying the deal prone segment. *J. Marketing Res.* 15(3):369–377.

Briesch RA, Chintagunta PK, Fox EJ (2009) How does assortment affect grocery store choice? *J. Marketing Res.* 46(2):176–189.

Burgess MJ, Steenkamp J-BEM (2006) Marketing renaissance: How research in emerging markets advances marketing science and practice. *Internat. J. Res. Marketing* 23(4):337–356.

Coleman RP (1983) The continuing significance of social class to marketing. *J. Consumer Res.* 10(3):265–280.

Dahl DW, Moreau CP (2007) Thinking inside the box: Why consumers enjoy constrained creative experiences. *J. Marketing Res.* 44(3):357–369.

De Wulf K, Odekerken-Schröder G, Iacobucci D (2001) Investments in consumer relationships: A cross-country and cross-industry exploration. *J. Marketing* 65(4):33–50.

Diaz A, Magni M, Poh F (2012) From Oxcart to Wal-Mart: Four key ways to reaching emerging market consumers. *McKinsey Quart.* (October). http://www.mckinsey.com/insights/winning_in_emerging_markets/from_oxcart_to_wal-mart_four_keys_to_reaching_emerging-market_consumers.

Evans KR, Christiansen T, Gill JD (1996) The impact of social influence and role expectations on shopping center patronage intentions. *J. Acad. Marketing Sci.* 24(3):208–218.

Goldman A, Krider R, Ramaswami S (1999) The persistent competitive advantage of traditional food retailers in Asia: Wet markets' continued dominance in Hong Kong. *J. Macromarketing* 19(2):126–139.

Haden P, Vittal I (2008) The great Indian bazaar: Organized retail comes of age in India. McKinsey & Company.

Hentschel J, Lanjow P (2000) Household welfare measurement and the pricing of basic services. *J. Internat. Development* 12(1):13–27.

Horsky D (1990) A diffusion model incorporating product benefits, price, income and information. *Marketing Sci.* 9(4):342–365.

Howard DJ, Gengler C, Jain A (1995) What's in a name? A complimentary means of persuasion. *J. Consumer Res.* 22:200–211.

Inman JJ, Shankar V, Ferraro R (2004) The roles of channel-category associations and geodemographics in channel patronage. *J. Marketing* 68(2):51–71.

Inman JJ, Winer RS, Ferraro R (2009) The interplay among category characteristics, customer characteristics, and customer activities on in-store decision making. *J. Marketing* 73:19–29.

Jayakody R, Danziger S, Kessler R (1998) Early-onset psychiatric disorders and male socioeconomic status. *Soc. Sci. Res.* 27: 371–387.

Joseph M, Soundararajan N, Gupta M, Sahu S (2008) Impact of organized retailing on the unorganized sector. Indian Council for Research on International Economic Relations, http://dipp.nic.in/english/publications/reports/icrier_report_27052008.pdf.

Kalish S, Mahajan V, Muller E (1995) Waterfall and sprinkler new-product strategies in competitive global markets. *Internat. J. Res. Marketing* 12:105–119.

Kamakura WA, Mazzon JA (2013) Socioeconomic status and consumption in an emerging economy. *Internat. J. Res. Marketing* 30(1):4–18.

Karabon K, Sukharevsky A (2011) Goodbye, mom-and-pops? The modern journey of Turkish consumers. McKinsey & Company, http://www.mckinseyonmarketingandsales.com/good-bye-mom-and-pops-the-modern-journey-of-turkish-consumers.

Kohli R, Bhagwati J (2015) Organized retailing in India: Issues and outlook. Bhagwati J, Panagariya A, eds. *Indian Economic Policies in the Twenty-First Century* (Cambridge University Press), Forthcoming.

Matsui K, Lu S, Nariu T, Yukimoto T (2005) Marketing channels and retail store density in East Asia. *Asian Econom. J.* 19(4):407–422.

Nevo A (2001) Measuring market power in the ready-to-eat cereal industry. *Econometrica* 69(2):307–342.

Ramola BG, Bhasin R (2011) Winning in India's retail sector: Factors for success. Report, PricewaterhouseCoopers.

Reynolds RB (1953) A test of the law of retail gravitation. *J. Marketing* 17(3):273–277.

Rogler LH (1996) Increasing socioeconomic inequalities and the mental health of the poor. *J. Nervous Mental Disease* 184(12): 719–722.

Rucker DD, Galinsky AD (2008) Desire to acquire: Powerlessness and compensatory consumption. *J. Consumer Res.* 35: 257–267.

Sheth JN (2011) Impact of emerging markets on marketing: Rethinking existing perspectives and practices. *J. Marketing* 75: 166–182.

Singh VP, Hansen KT, Blattberg RC (2006) Market entry and consumer behavior: An investigation of a Wal-Mart supercenter. *Marketing Sci.* 25(5):457–476.

Sood D, Jashnani P (2012) India's food retail sector growing. Report IN2059, Global Agricultural Information Network, U.S. Department of Agriculture, Washington, DC.

Spiegelhalter DJ, Best NC, Carlin BP, van der Linde A (2002) Bayesian measures of model complexity and fit. *J. Roy. Statist. Soc. Ser. B* 64:583–639.

Sudhir K (2001) Competitive pricing behavior in the US auto market: A structural analysis. *Marketing Sci.* 20(1):42–60.

Sudhir K, Talukdar D (2015) The "Peter Pan syndrome" in emerging markets: The productivity transparency tradeoff in IT adoption. *Marketing Sci.* 34(4):500–521.

Sumarto S, Suryadarma D, Suryahadi A (2007) Predicting consumption poverty using non-consumption indicators: Experiments using Indonesian data. *Soc. Indicators Res.* 81(3):543–578.

Tellis GJ, Stremersch S, Eden Y (2003) The international take-off of new products: The role of economics, culture, and country innovativeness. *Marketing Sci.* 22(2):188–208.

van den Bulte C, Stremersch S (2004) Social contagion and income heterogeneity in new product diffusion—A meta-analytic test. *Marketing Sci.* 23(4):530–544.

Venkatesh U (2008) The case for modern agri-retail—The Indian imperative. *J. Services Res.* 8(1):91–107.