

# And then what happened? A longitudinal study of the long-term effects of adding an online channel to an offline hypermarket

**Emelie Fröberg**

Stockholm School of Economics

**Sara Rosengren**

Stockholm School of Economics

## Acknowledgements:

This work was supported by Hakon Swenson Stiftelsen.

## Cite as:

Fröberg Emelie, Rosengren Sara (2020), And then what happened? A longitudinal study of the long-term effects of adding an online channel to an offline hypermarket. *Proceedings of the European Marketing Academy*, 49th, (63881)

Paper from the 49th Annual EMAC Conference, Budapest, May 26-29, 2020.



## **And then what happened? A longitudinal study of the long-term effects of adding an online channel to an offline hypermarket**

### **Abstract**

This study investigates the long-term effects of adding an online sales channel to an existing offline grocery retail offer. Based on previous literature, it is hypothesized that the addition of an online channel could have long-term positive effects thanks to tilted shopping behaviors, negative effects due to more planned shopping, or no effect because of cannibalization. Findings show that the offline channel still attracts the largest customer base, but the online channel is growing—to a large extent at the expense of existing offline customers. Using propensity score matching, we also test short- and long-term effects of becoming a multichannel household. Our results show that multichannel customers spend more (per visit and in weekly average), purchase larger volumes, and create higher margins both in the short- and long-term.

**Keywords:** *Multichannel, retailing, propensity score matching*

**Track:** Retailing & Omni-Channel Management

## **1 Introduction**

This study aims to explore the long-term effects of adding an online channel to an existing offline grocery retail offer. More specifically, we complement previous research showing how grocery shopping behaviors change when an online offer is introduced (e.g., Melis et al., 2016), by exploring what happens to shopping patterns over time.

Several studies have found that multichannel customers are more loyal and more profitable (e.g., Kumar and Venkatesan, 2005; Wallace et al., 2004). We ask whether such effects primarily occur at the time of introduction of the new channel or if they are sustained over time. We explore this question using five years of point of sales data from customers of one of the largest grocery retailers in the Nordics, zooming in on one hypermarket in a mid-sized city. Our study contributes to previous studies that have evaluated short- and mid-term effects of the online channel based on aggregated data (Cao and Li, 2015; Herhausen et al., 2015; Pauwels and Neslin, 2015) or panel data (Melis et al., 2016), in that we can follow individual households' actual shopping behaviors over a time.

## **2 Previous literature**

Adding an additional online channel to an offline grocery retail offer can be hypothesized to have positive, negative or no effect on household long-term shopping behaviors. An online channel offers convenience to customers (e.g., Bhatnagar and Ratchford, 2004; Verhoef et al., 2007, 2015). Multiple channels also extend the distribution and can thus be argued to increase sales simply because there are more distribution channels available (Neslin et al., 2006). Several studies have found that multichannel customers are more loyal and more profitable (e.g., Kumar and Venkatesan, 2005; Wallace et al., 2004). Product category characteristics have also been found to moderate this positive effect in that hedonic products evoke more impulsive and variety-seeking behaviors with multiple channels available (Kushwaha and Shankar, 2013). As grocery retailers offer both utilitarian and hedonic products these types of behaviors could be expected to occur also in this context. Overall, this would suggest a long-term positive effect of an online channel on grocery shopping behaviors.

However, many retailers have been hesitant to add an online channel to an existing offline offer as they worry about cannibalization (e.g., Hernant and Rosengren, 2017); that

sales would only shift from one channel to another (Alba et al., 1997) and that price competition online would force prices down (Brynjolfsson and Smith, 2000). Overall this would suggest a long-term negative effect of an online channel on grocery shopping behaviors.

There is, however, mixed empirical evidence for how these effects play out over time in a grocery retail setting and whether they cancel each other out. Using two years of panel data, Melis et al. (2015) show that although the initial effect is positive, after online experience increased customers tend to choose only the most efficient channel and, in turns, reduce cross-channel purchasing over time. What is more, the empirical evidence in terms of cannibalization is inconclusive (cf. e.g., Biyalogorsky and Naik, 2003; Hernant and Rosengren, 2017; Lee and Grewal, 2004). To further complicate things, many of these studies tend to compare multi-channel customers with pure offline customers. This is problematic as the two channels tend to offer different value (for example, the online channel offer search convenience, while the offline channel offer better service, assortment, after-sales support, and lower risks, cf. Verhoef et al., 2007) and we would expect different customer groups to be appealed by the different channels (e.g., Grewal et al., 2004). This self-selection challenges research on long-term effects of adding an online channel, because certain customers could self-sort to use multiple channels making it hard to discern whether it is the multichannel offering that actually drive effects such as higher customer profitability (Neslin et al., 2006). To tackle similar problems a matching procedure has been found useful (e.g., Rishika et al., 2013; Wang et al., 2015) and thus we will use this procedure to assess the long-term effects in this study.

### **3 Method**

We use point of sales data based on loyalty cards for a Nordic grocery retailer. Data is collected from households at one specific hypermarket in a mid-sized city. We also include any purchase a household has made in any of the other stores that is run by the same grocery chain. We use a period of five years in total, the first date is March 16, 2014 and the last date is March 16, 2019, where the online channel was introduced in March 16, 2015. Due to data unavailability there is no data available for December 1, 2016 to July 31, 2017. In total, we have access to data from 11,094 unique households and their 2,475,298 unique receipts. The yearly number of unique customers range between a minimum of 4,837 households, in 2014,

and a maximum of 7,145 households, in 2019. In this study, we use the terms customers and households interchangeably.

To deal with self-selection to the online channel we use a matching procedure, meaning that each multi-channel household (the treated group) is matched to a household that only shops offline (the untreated household) using propensity scores estimated by logistic regression in the time period before the event took place (i.e. online channel opened, see e.g., Ho et al., 2007). We use Recency, Frequency and Monetary values (cf. McCarty and Hastak, 2007) as well as Volume (cf. Hernant and Rosengren, 2017) as input to estimate propensity scores. Statistics were done using R 3.6.1 (R Core Team, 2018) and we used the *MatchIt* (Ho et al., 2011) package for the propensity score matching.

When we run the matching procedure, we include the 1,729 households that made at least one purchase during four of the five years for which we have access to data. In the dataset 1,536 households made at least one purchase before the online channel was introduced. Of these, 60 percent later made at least one purchase online. We take a random sample of 300 households that only made purchases offline (about 50 percent) to ensure near matching and then make individual matches (based on Recency, Frequency, Monetary and Volume) before the introduction of the online-channel from a pool of 912 households that purchased online later.

In Table 1 we briefly describe our main variables of interest. We run matched sample t-tests on all of these variables to verify that our matching was successful. There are no statistically significant differences between the treated group (households purchasing online) and the untreated group (households only purchasing in-store) before the online channel opened (the lowest p-value is 0.40, except for volume where the average difference is 0.43 products,  $t(299) = 2.49$ ,  $p = 0.013$ ).

---

Table 1

*Description of variables*

---

<i>Variable</i>	<i>Description</i>
Recency	Weeks since last purchase
Frequency	Total number of visits divided by total weeks since first purchase
Monetary	Total spend per visit (in SEK)
Volume per visit	Total number of distinct products per visit
Sales	Total spend (in SEK)
Margin	Total spend plus compensation less discounts, VAT and purchase value (in SEK)

---

## 4 Results

In Table 2 we see that the majority of households (58 percent) prefer to shop groceries offline. But the customer base offline is unchanged, whereas the online- and multichannel customer base is growing. The great majority of online households have first shopped from the retail offline offer. Among the 3,841 households that ever made an online purchase, we estimate that 2,239 households have changed habits (872 households shop online once and 418 households shop online twice; after a third time we expect a new pattern).

Table 2. <i>Three-way cross-tabulation frequencies of multichannel vs. single channel households, channel of first purchase (online vs. offline) and time window of evaluation (three year vs. five year)</i>			
Multichannel user?	First purchase	Time window	
		Three year	Five year
Yes, household has shopped in both channels	Online	39	208
	Offline	1160	2755
No, household has shopped in only one channel	Online	48	139
	Offline	4296	4203

*Note.* The cell values show the number of households. Three year time window includes one year before and first two years after the introduction of an online channel. Five year time window includes one year before and four years after the introduction of an online channel; here we require at least one purchase in the last two years (i.e. after 2017-03). A period of eight months is missing. The online channel was introduced in March 2015, we lack data for Dec-2016 to July-2017. Our empirical results are based on membership data only. All online customers are required to hold a membership, whereas this is optional in the offline retail offer. A purchase before the period of interest is not taken into account (i.e. there may exist a first purchase offline that took place over a year before the introduction of the online channel, but this is not taken into account here).

Moving forward, we focus on those households that made their first purchase offline and use a propensity score matching procedure to explore the short- and long-term effects of adding an online channel on their shopping behaviors. In Figure 1, we show comparisons of all variables of interest between offline households and multichannel households, including a period from the first online purchase of the multichannel household until the last day of the dataset. Note that the first online date for each multichannel household can differ. We use the

first individual date for every matched pair of households. We also make these comparisons (available in Figure 1) by running matched sample t-tests. The following variables are statistically significantly different between the treated and untreated households: Frequency,  $t(298) = 2.73, p = 0.007$ ; Monetary,  $t(298) = 4.40, p < 0.0001$ ; Volume,  $t(298) = 3.85, p = 0.0001$ ; Sales,  $t(298) = 6.19, p < 0.0001$ ; and Margin,  $t(298) = 6.22, p < 0.0001$ . There is no significant difference in Recency,  $t(298) = 1.43, p = 0.15$ .

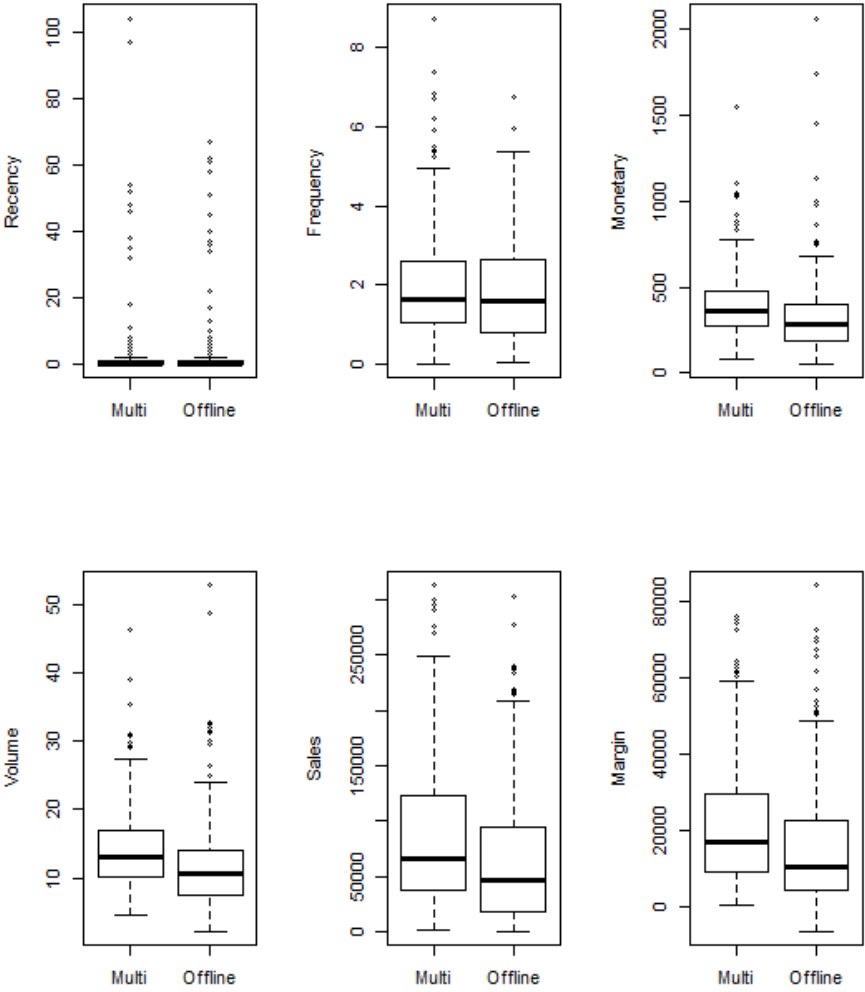


Figure 1. *Boxplot matrix comparisons of offline vs. multichannel households*

We also split our data into a short period (the immediate two years from making the first purchase), and a long period (all purchases after the immediate two years). Note again that the first online purchase will differ between households, but for each matched pair we use one cutoff. We exclude matched cases ( $n = 100$ ) where first online purchase is within two years before the final date. We evaluate average weekly sales and margins and run a mixed-design ANOVA. In Figure 2, we show that multichannel households have higher volumes and spend per visit and that margins and sales are higher than for offline households. There is no difference between short- and long-term effects. Mixed-design ANOVAs (with channel as

between factor, period as within factor, using household fixed effects, and channel and period as an interaction factor) show that channel is the only statistically significant factor in explaining differences in volume, monetary spend, sales, and margin. This suggests that shopping habits are tilted for households who use the online channel, and that this effect holds both in short- and long-term.

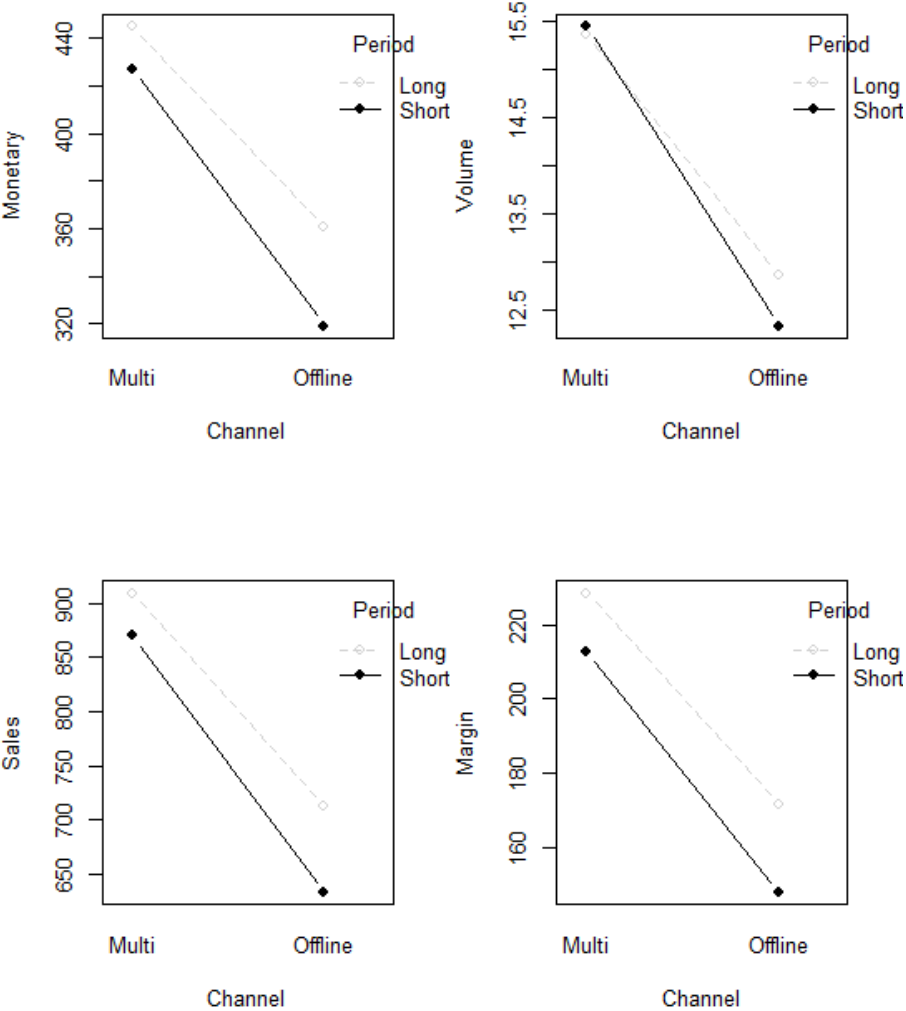


Figure 2. Two-way interaction plots with channel and period factors

Finally, in Figure 3 we show the effects of converting households to multichannel households on margins over time. Note that the red line is growing over time because the first purchase online takes place in different times depending on household. Overall margins are growing as more and more households becomes multichannel users—more than the consumer price index in the same period. Also note here that the matching methodology enables us to overcome self-selection bias, in that each pair had similar shopping patterns before the introduction of the online channel. Still, we might have an issue with survivorship bias, because we demand at least one purchase per four years in a period of five years and we rely on loyalty membership data. We thus do not make any claims about the overall profitability



effects on a hypermarket, but we do see that there are only positive short- and long-term effects of converting hypermarket customers to multichannel hypermarket customers.

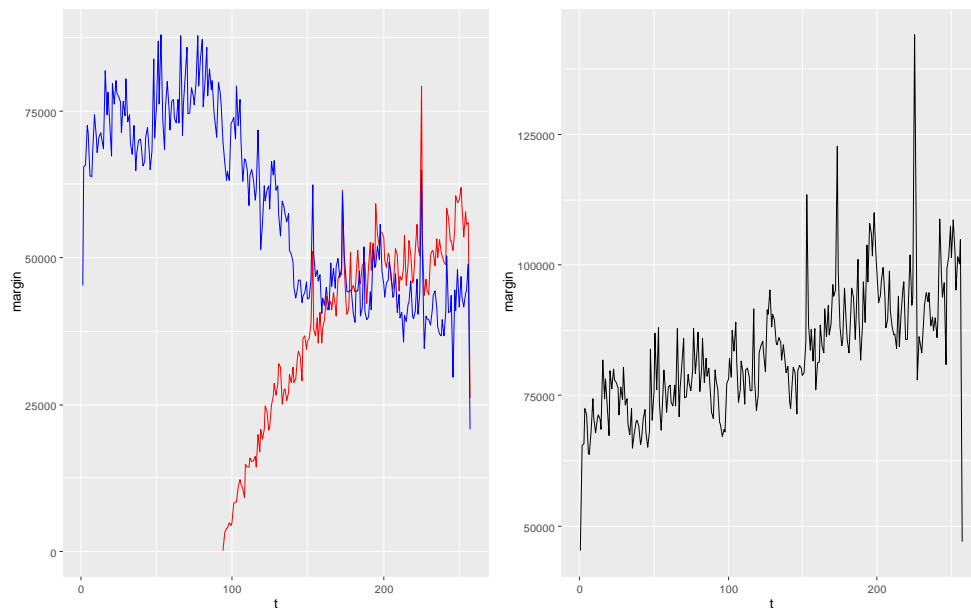


Figure 3. *Margin over time, left shows multichannel households (red) vs. offline households (blue) and right shows total effects, both based on propensity scored matched pairs*

## 5 Concluding remarks

What are long-term effects of adding an online channel to an offline hypermarket? Based on five years of actual shopping behaviors from three hundred matched pairs of households, we found that—both short- and long-term—effects are positive. After becoming a multichannel household, customers buy higher volumes, in higher frequency, and with higher monetary spend per visit. This in turn is reflected in higher sales and margins. Our results support that shopping behaviors are changed upon starting using a new shopping channel, and that consumption stays at this level (rather than diminishing after a period of learning). We also note that the great majority of households are still mainly shopping offline and multichannel households are most likely to have experience with the offline retail offer first. Managerial implications are to attract existing customers to try online shopping, to break existing habits and thereby grow in terms of sales and margins. It is our hope that we can discuss this further at the conference.

## References

Alba J, Lynch J, Weitz B, et al. (1997) Interactive Home Shopping: Consumer, Retailer, and

- Manufacturer Incentives to Participate in Electronic Marketplaces. *Journal of Marketing* 61(3): 38–53. DOI: 10.1177/002224299706100303.
- Bhatnagar A and Ratchford BT (2004) A model of retail format competition for non-durable goods. *International Journal of Research in Marketing* 21(1): 39–59. DOI: 10.1016/j.ijresmar.2003.05.002.
- Biyalogorsky E and Naik P (2003) Clicks and Mortar: The Effect of On-Line Activities on Off-Line Sales. *Marketing Letters* 14(1). Springer: 21–32. Available at: <http://www.jstor.org/stable/40216489>.
- Brynjolfsson E and Smith MD (2000) Frictionless Commerce? A Comparison of Internet and Conventional Retailers. *Management Science* 46(4): 563–585. DOI: 10.1287/mnsc.46.4.563.12061.
- Cao L and Li L (2015) The Impact of Cross-Channel Integration on Retailers' Sales Growth. *Journal of Retailing* 91(2): 198–216. DOI: 10.1016/j.jretai.2014.12.005.
- Grewal D, Levy M and Lehmann DR (2004) Retail Branding and Customer Loyalty: an overview. *Journal of Retailing* 80(4): ix–xii. DOI: 10.1016/j.jretai.2004.10.001.
- Herhausen D, Binder J, Schoegel M, et al. (2015) Integrating Bricks with Clicks: Retailer-Level and Channel-Level Outcomes of Online–Offline Channel Integration. *Journal of Retailing* 91(2): 309–325. DOI: 10.1016/j.jretai.2014.12.009.
- Hernant M and Rosengren S (2017) Now what? Evaluating the sales effects of introducing an online store. *Journal of Retailing and Consumer Services* 39: 305–313. DOI: 10.1016/j.jretconser.2017.08.010.
- Ho DE, Imai K, King G, et al. (2007) Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis* 15(3): 199–236. DOI: 10.1093/pan/mpi013.
- Ho DE, Imai K, King G, et al. (2011) MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software* 42(8): 1–28. Available at: <http://www.jstatsoft.org/v42/i08/>.
- Kumar V and Venkatesan R (2005) Who are the multichannel shoppers and how do they perform?: Correlates of multichannel shopping behavior. *Journal of Interactive Marketing* 19(2): 44–62. DOI: 10.1002/dir.20034.
- Kushwaha T and Shankar V (2013) Are Multichannel Customers Really more Valuable? The Moderating Role of Product Category Characteristics. *Journal of Marketing* 77(4): 67–85. DOI: 10.1509/jm.11.0297.
- Lee RP and Grewal R (2004) Strategic Responses to New Technologies and Their Impact on

- Firm Performance. *Journal of Marketing* 68(4): 157–171. DOI: 10.1509/jmkg.68.4.157.42730.
- McCarty JA and Hastak M (2007) Segmentation approaches in data-mining: A comparison of RFM, CHAID, and logistic regression. *Journal of Business Research* 60(6): 656–662. DOI: 10.1016/j.jbusres.2006.06.015.
- Melis K, Campo K, Breugelmans E, et al. (2015) The Impact of the Multi-channel Retail Mix on Online Store Choice: Does Online Experience Matter? *Journal of Retailing* 91(2): 272–288. DOI: 10.1016/j.jretai.2014.12.004.
- Melis K, Campo K, Lamey L, et al. (2016) A Bigger Slice of the Multichannel Grocery Pie: When Does Consumers' Online Channel Use Expand Retailers' Share of Wallet? *Journal of Retailing* 92(3): 268–286. DOI: 10.1016/j.jretai.2016.05.001.
- Neslin SA, Grewal D, Leghorn R, et al. (2006) Challenges and Opportunities in Multichannel Customer Management. *Journal of Service Research* 9(2): 95–112. DOI: 10.1177/1094670506293559.
- Pauwels K and Neslin SA (2015) Building With Bricks and Mortar: The Revenue Impact of Opening Physical Stores in a Multichannel Environment. *Journal of Retailing* 91(2): 182–197. DOI: 10.1016/j.jretai.2015.02.001.
- R Core Team (2018) R: A Language and Environment for Statistical Computing. Vienna, Austria. Available at: <https://www.r-project.org/>.
- Rishika R, Kumar A, Janakiraman R, et al. (2013) The Effect of Customers' Social Media Participation on Customer Visit Frequency and Profitability: An Empirical Investigation. *Information Systems Research* 24(1): 108–127. DOI: 10.1287/isre.1120.0460.
- Verhoef PC, Neslin SA and Vroomen B (2007) Multichannel customer management: Understanding the research-shopper phenomenon. *International Journal of Research in Marketing* 24(2): 129–148. DOI: 10.1016/j.ijresmar.2006.11.002.
- Verhoef PC, Kannan PK and Inman JJ (2015) From Multi-Channel Retailing to Omni-Channel Retailing. *Journal of Retailing* 91(2): 174–181. DOI: 10.1016/j.jretai.2015.02.005.
- Wallace DW, Giese JL and Johnson JL (2004) Customer retailer loyalty in the context of multiple channel strategies. *Journal of Retailing* 80(4): 249–263. DOI: 10.1016/j.jretai.2004.10.002.
- Wang RJ-H, Malthouse EC and Krishnamurthi L (2015) On the Go: How Mobile Shopping Affects Customer Purchase Behavior. *Journal of Retailing* 91(2): 217–234. DOI: 10.1016/j.jretai.2015.01.002.