

Going green: the effect of green labels on delivery time slot choices

Niels Agatz, Yingjie Fan*, and Daan Stam

*Rotterdam School of Management, Erasmus University Rotterdam, Burgemeester
Oudlaan 50, 3062 PA Rotterdam, The Netherlands*

Abstract

In this paper, we study the effectiveness of incentives on delivery service time slot choices. In particular, we focus on the use of green labels that specify time slot as environmentally friendly and that *intrinsically* motivate customers to choose a specific delivery time slot in lieu of price incentives based on *extrinsic* motivation. We argue this is important since green labels' intrinsic nature affects customer choice in fundamentally different ways than price incentives. We conduct two experiments and two simulation studies to study effects of using green labels. Our experimental findings suggest that: (1) green labels are an effective tool to steer shoppers toward a certain delivery option, (2) green labels are more effective for people who are more eco-conscious, (3) green labels remain effective in the presence of price incentives, while price incentives offer little added value beyond that of just green labels, and (4) the effectiveness of green labels versus price discounts remains high when time slots are less appealing (longer). Our simulation findings suggest that green slots, compared to price incentives or no incentives, offer providers a way to effectively steer consumer time slot choices to yield shorter routes, fewer delivery vehicles used, and more per-customer revenue. We thus conclude that steering individuals to select delivery time slots through intrinsic motivation via green labels may be a promising, no-cost direction for (online) retailers and an important topic for research.

Keywords: customer behavior, green label, intrinsic incentive, attended home delivery

*corresponding author

1 Introduction

In many home-service settings, the customer must be present to accept the provided services. Examples include the delivery of groceries, furniture and white goods, as well as home repair and installation services. To coordinate such businesses, it is common for service providers to offer customers a choice of narrow delivery time slots via an online booking system.

In this context, customer time-slot choices directly impact the operational efficiencies of service providers (Boyer et al. 2009, Agatz et al. 2008). For example, a delivery truck may have to visit the same street multiple times daily to serve customers scheduling different delivery times. Thus, attended service may generate twice the vehicle miles as unattended home delivery systems (Punakivi and Tanskanen 2002, Lin and Mahmassani 2002, Agatz et al. 2011).

Inefficiencies in the last-mile not only yield higher costs for the service provider, but also impose substantial societal costs such as carbon emissions, air pollution and traffic congestion. These are important issues since air pollution is associated with millions of premature deaths across the globe (WHO 2016), and traffic congestion costs hundreds of billions of dollars in lost productivity and fuel waste (INRIX 2018).

Companies have typically focused on reducing inefficiencies through *supply-side* process improvements related to network design (de Vries et al. 2020), inventory deployment, and route-dispatch decisions (Dayarian and Savelsbergh 2020, Emadikhiav et al. 2020). Another approach, at least as powerful, focuses on influencing *demand* as a way to enhance system performance. Regarding attended home delivery, such demand-side management can help steer customers toward service options expected to cut system-wide vehicle miles.

The prevailing demand-management paradigm uses price to steer behavior (Amornpetchkul et al. 2018, Ceryan et al. 2018, Shen and Yu 2019). For instance, service providers often enlist price incentives to encourage customers to select delivery time slots that facilitate efficient route plans. Such incentives are typically presented as discounts on the delivery fee. The question of how to dynamically offer such price incentives to maximize profits has been much studied in recent years (Campbell and Savelsbergh 2006, Yang et al. 2014, Yang and Strauss 2017). However, price incentives pose several major drawbacks. First, price incentives kill margins. In low-margin business models, such as delivered groceries, small discounts can greatly harm profitability (Srinivasan et al. 2002). Second, customers may come to expect lower prices (DelVecchio et al. 2007) and delay purchases hoping for more discounts (Zhang et al. 2019). This type of strategic behavior is a known bane to effectiveness in dynamic pricing policies (Liu

and Zhang 2013, Su 2007). Third, dynamic pricing can be perceived as unfair by the customer, undermining trust in the online provider (Weisstein et al. 2013, Xia et al. 2004).


Alternatively, one can manage demand based on differentiating the product or service offering (Agatz et al. 2013). For example, service providers can limit the number of time slot options in a given geographical area to cluster customer orders and reduce the distance traveled per order. Several papers have focused on developing prescriptive methods to decide which slots to open and in which areas, both static, prior to the order in-take (Agatz et al. 2011) and dynamic, based on already accepted orders (Campbell and Savelsbergh 2005, Ehmke and Campbell 2014, Köhler et al. 2019). The disadvantage of slot-based approaches is that they limit the number of time slots choices for the customer which may make the service proposition less attractive.

Fortunately, there are other ways to steer demand that are not based on pricing and slotting. Specifically, we argue that companies can also encourage customers to choose delivery times that reduce routing distance *by featuring environmental benefits*. Several online grocery retailers, including Ocado and Peapod, are appealing to the customer environmental concerns by indicating which delivery times are associated with lower fuel consumption and emissions. We refer to such non-price incentives as *green labels* where time slots are green and indicate their associated environmental savings (see Table 1).

Green labels fundamentally depart from price incentives in how they steer customer behavior. While price incentives are extrinsic motivators offering financial reward, green labels are *intrinsic* motivators featuring an alternative that feels “right” attuned to customers’ values. Green labels thus steer customers into specific slots free of price incentive problems. This makes green label usage an appealing strategy for service providers. Beyond being an alternative to price discounts, green labels also exert very different influences on customer behavior due to its intrinsic nature. As such, well-known banes linked to price incentives not apply to green labels. Although online retailers in practice are starting to use green labels as part of their service offerings, this topic remains largely unexamined in the literature. This paper aims to fill this void. As far as we know, this is the first study to scrutinize the effectiveness of green labels in the context of last-mile delivery.

In this paper, we focus on several key questions related to the use of green labels in this context. In experimental studies, we investigate some of these questions: First and foremost, we examine the effectiveness of green labels in steering customers toward more sustainable service options compared to using price incentive versus no discount. Second, to demonstrate that green

Table 1: The use of green labels at UK-based e-grocer Ocado

 These little vans show our greener delivery slots.

MORNING ^a	Friday 3rd May	Saturday 4th May	Sunday 5th May
5:30 - 6:30am	£0.00	£0.00	£0.00
6:00 - 7:00am	£0.00	£0.00	£0.00
⋮	⋮	⋮	⋮
12:30 - 1:30pm	£0.00 	£0.00	£0.00
1:00 - 2:00pm	£0.00	£0.00	£0.00
2:00 - 3:00pm	£0.00 	£0.00	£0.00
⋮	⋮	⋮	⋮
7:00 - 8:00pm	£0.00 	£0.00 	£0.00
⋮	⋮	⋮	⋮
10:00 - 11:00pm	£0.00	£0.00	£0.00
10:30 - 11:30pm	£0.00	£0.00 	£0.00

^aThis table shows a part of a delivery slot menu of Ocado in 2019.

labels are intrinsic motivators, we study whether green labels especially appeal to customers that value the natural environment. Third, we test the impact of combining green labels with price incentives to argue, based on the intrinsic nature of green labels, that price incentives offer no extra value. Fourth, we investigate whether the effects of green slots persist when targeted options become less convenient (i.e. longer windows versus other available slots) to argue, again based on the intrinsic nature of green labels, that green label outperform price incentives.

Finally, we simulate the impact of green-label steering of customer behavior on the efficiency of the last-mile delivery system: are green labels truly pro-environment, and how do they affect delivery as a whole? In particular, these simulations consider two different demand-management settings where green labels can be used in last-mile delivery. In the *dynamic* setting, customers choose from a set of time slots equal in length, e.g., one-hour time slots. The service provider dynamically determines, as each customer engages, delivery time slots likely associated with the least routing distance adding that customer to the given schedule. Here, we investigate the effects of green labels versus price incentives. In the *static* setting, the customer can choose among time slots of different lengths, e.g., one- or five-hour slots. Longer time slots offer more

routing flexibility and less travel (Köhler et al. 2019, Strauss et al. 2020). Here, longer slots are incentivized for all customers *before* routing. In this setting, too, we investigate the effects of green labels versus price incentives.

Our main contributions to the Operations Management (OM) literature can be stated as follows. First, we debut empirically testing of green-label effectiveness in steering service-choice behavior. We provide evidence that green labels intrinsically motivate to work in fundamentally different ways commonly studied in extrinsic motivators such as price promotions. Moreover, we show that green labels and price promotions/price incentives do not work well together. While prior OM literature (Chen and Chen 2015) has primarily focused on price-based approaches, we show that green labels can be equally or more effective (albeit different) in steering behavior. We hope to spark future work in this area. Second, we pair a simulation model based on state-of-the-art time slotting and routing heuristics with our empirically validated customer-choice model to quantify the potential impact of steering demand. This means we complement extant studies that explore the impact of demand management approaches in the last-mile based on artificial and stylized customer demand data (Campbell and Savelsbergh 2006, Ehmke and Campbell 2014) or those chiefly based on price incentives (Yang et al. 2014).

Finally, our study merges the advantages of experimental study and simulation to overcome some of the limitations these methods have in isolation to complement existing frameworks that combine surveys or case studies with simulation. Specifically, Chandrasekaran et al. (2018) overview the latest possible uses for the combining empirical and simulation methods. Absent from their model is our use of experiments offering causal evidence of effects joined with simulation that quantify these effects on delivery operations systems (Chandrasekaran et al. 2018). We believe this mode of merging experiment with simulation is widely applicable and useful in operations management.

The remainder of this paper is organized as follows: Section 2 provides our theoretical foundation and the hypotheses; Section 3 and 4 feature results from empirical plus simulation study of the *dynamic* setting for equal length slots; Sections 5 and 6 discuss results from empirical plus simulation study of the *static* setting for slots of different lengths; and Section 7 concludes.

2 Theory and hypotheses

The effect of green labels

The core premise of this research is that the use of green labels in last-mile delivery is an effective tool to steer demand. The assumption behind this premise is that people prefer “green” services. There are several streams of literature that testify to this idea.

First, we present evidence concerning “green” consumption. Over recent decades, companies have increasingly developed and produced eco-sustainable products and services to reduce their environmental impact while appealing to the burgeoning number of eco-conscious consumers. Organizations and individuals concerned about the environment are willing to take steps toward eco-protection (Peattie and Ratnayaka 1992). Research suggests that consumers generally prefer “green” products (Salzman 1991, Ottman 1993, Rana and Paul 2017). In a recent European study, three-quarters of participants claimed *willingness to pay up* for eco-friendly products. In a similar vein, surveys in the U.S., Japan, and India show that consumers are willing to pay more for “green” electricity (Roe et al. 2001, Nomura and Akai 2004, Yadav and Pathak 2017, Rana and Paul 2017).

Second, apart from “green” consumption, there is also a large body of empirical research on stimulating people to display pro-environmental behavior, i.e., behavior that incurs little harm to the environment, or even benefits it (Steg and Vlek 2009). More than three decades of research suggest that (near) real-time feedback can change household-level behavior in terms of energy and water consumption. However, recent meta-studies show that the results and effect sizes differ widely across and even within studies (Delmas et al. 2013, Vine et al. 2013). Moreover, much debate remains about the long-term effectiveness of feedback interventions (Lynham et al. 2016, Buchanan et al. 2015). Interestingly, several studies based on field experiments show that listing environmental information can more effectively elicit conservation behavior than offering mere financial information. Jain et al. (2013), for example, saw tallying environmental externality units (e.g., the number of “trees needed to offset emissions” of customer consumed energy) to be more effective in cutting energy use than other information strategies, both short- and long-term.

To summarize, providing green information works. It may, in some circumstances, even be more effective than the alternatives since green information may *intrinsically* motivate people to buy a product, to display a behavior, or in our study, to choose a certain delivery service option.

Ryan and Deci (2000) defined intrinsic motivation as the motivation to conduct an activity simply for the enjoyment, pleasure, or satisfaction of the activity itself. Extrinsic motivation refers to acts seeking some separable outcomes, e.g., perceived utility, getting rewards, and avoiding penalties. Providing green information about certain activities may help customers to see the value of these and, to the extent that sustainability fits their value systems, intrinsically motivate them to pursue these actions. As a consequence, the consumer may thus perceive the buying of green products to be more satisfying or the using of less energy to be more important.

We argue that the same holds for green delivery services that provide attended home delivery. Using green labels to highlight the sustainable delivery times can allow customers to opt for slots that better align with their values. Consequently, we argue that the use of green labels are effective in steering customers towards these specific slots. This leads to the following hypothesis:

Hypothesis 1 *Delivery time slots featuring a green label are more likely to be selected than those without a green label.*

One important premise underlying green labels is that they work through intrinsic motivation. This, we propose, has several important consequences for the effectiveness and use of green labels. Specifically, extending prior research on intrinsic motivation, we propose that green labels, due to their intrinsic nature, will *i)* be more effective for eco-conscious customers, *ii)* not synergize with price incentives and *iii)* be less sensitive to the degree of inconvenience (length) in slot options. The following sections explain these ideas.

Green labels and eco-consciousness

The notion that green labels work through intrinsic motivation also implies varying effectiveness across different customer segments. The idea behind the intrinsic value of green labels is that they match the value system of customers to create intrinsic motivation. Indeed, values and norms are widely known as important drivers of behavior. For instance, the theory of planned behavior posits that individual behavior is influenced by attitudes toward it, by subjective norms, and by perceived control (Ajzen 1991). Wilson and Dowlatabadi (2007) have argued that behavior is determined by psychological factors such as values, attitudes, and personal norms, as well as by contextual factors such as available choices, economic incentives and social norms.

Buying green products or displaying green behavior is altruism influenced by personal norms (Schwartz 1977). More specifically, one important factor in green behavior is eco-consciousness (Laroche et al. 2001, Gleim et al. 2013, Rana and Paul 2017). Eco-consciousness is an individual’s stable inclination to favor sustainability and value the natural environment. Research suggests that eco-conscious customers tend to buy green products (Kim and Chung 2011). For instance, a survey on consumers in Iran has shown a positive correlation between consumer eco-concern and green purchase intention (Vazifehdoust et al. 2013). An empirical study by Suki (2016) showed similar results. Green labels thus stimulate eco-conscious customers to act with environmental responsibility (Menon and Menon 1997, D’Souza et al. 2007, Yazdanifard and Mercy 2011, Leonidou and Leonidou 2011).

We argue that green labels work best when they align with customer values and concerns that are pro-environment. Indeed, according to Steg et al. (2016), obligation-based intrinsic motivation underlies pro-environment behavior. This implies that green labels may heighten intrinsic motivation in customers who care more about the environment. This leads to the next hypothesis:

Hypothesis 2 *Customers who are more eco-conscious are more likely to choose a delivery slot featuring a green label.*

Green labels and price incentives

We earlier suggested that the use of green labels benefit companies. Given the predominance of price incentives, this begs the question of how green labels and price incentives may interact.

Specifically, price incentives are known to be an effective way to increase customer purchase intention and choices (Chen et al. 1998, Duvvuri et al. 2007). Price incentives can be effective in shaping customer time slot choices as well (Yang et al. 2014). Some companies, e.g., Albert Heijn Online and Peapod, now merge price incentives with green labels. Yet, it is unclear how a combination of price incentives with green labels affects slot choices: Is the combination of these parts equal to, bigger or smaller than its summation? Research on extrinsic and intrinsic motivation speaks to this issue. Prior research has predominantly implied that external rewards may actually suppress intrinsic motivation (Deci 1971, Lepper et al. 1973). When extrinsically motivated tasks are linked with external rewards, touting external rewards may curb people’s intrinsic motivation for the task (Schwartz et al. 2015). This is often referred to as the “crowding-out” effect (Lepper et al. 1973).

The idea behind the crowding-out effect is that feelings of control and autonomy induce more intrinsic motivations (Ryan and Deci 2000). For instance, autonomy-supportive teachers yield more intrinsic-motivated students (Ryan and Grolnick 1986). External rewards, though, are often perceived as autonomy-diminishing or arm-twisting, thus desensitizing any intrinsic motivation (Bowles 2008). Meta-analyses show rewards attenuating intrinsic motivations (Deci et al. 1999).

When a price incentive is offered for a behavior, it implies that conduct can be bought and sold, a commodity view rarely associated with social norms. In this way, price incentives reduce the intrinsic motivation for a specific behavior (Gneezy et al. 2011). For instance, a penalty on parents who pick up children late at the day-care center can actually increase the number of late pickups (Gneezy and Rustichini 2000). A price incentive can be perceived as distrust, even disrespect. That is, an individual’s intrinsic motivation is deemed too weak to undertake the task without a price incentive. Since people exert more pro-social efforts when they feel trusted and respected (Ellingsen and Johannesson 2007), the price incentive crowds out intrinsic motivation.

This suggests that the impact of simultaneously pairing a price incentive with a green label may underperform their separate effects. The price incentives (as a form of external rewards) may undermine (crowd out) the intrinsic motivation triggered by the green labels. Customers may feel that choosing a delivery option that has both a price incentive and a green label may not constitute a behavior in tune with their values toward sustainability. However, this occurs only when price incentives are seen as “controlling” by customers, i.e., when shoppers feel that companies are steering them toward certain slots via discounts. Assuming that choosy customers understand the value of demand management for companies, we believe the crowding-out effect really occurs. This analysis leads to our third hypothesis:

Hypothesis 3 *The effect of the green label is less strong if the same slots are also associated with a price incentive.*

Green labels and longer time slots

Another consequence of the intrinsic motivation from green labels is that they may still work even when the green service is not appealing. This is important in the area of last-mile delivery since slot length is often crucial to both the costs and effectiveness of demand management systems (Campbell and Savelsbergh 2005, Solomon 1987). Longer time windows, e.g., five

hours versus one hour, are much less appealing for customers waiting at home for an uncertain delivery time. Encouraging customers to choose longer delivery slots via price incentives may prove difficult and costly. Indeed, when price incentives are used to prompt certain behavior, an individual's effort positively correlates to compensation amount (Heyman and Ariely 2004). In other words, price incentives trigger an economic transaction rationale where cost and benefit are weighted. In order to steer customers to longer time slots, companies must offer *greater* price incentives or, put differently, the same price incentive will be less effective for longer time slots.

This may be less so for green labels and their intrinsic motivating function. There is evidence that intrinsic motivation may be more resilient than extrinsic motivation in the face of obstacles or difficulties in the transaction. For instance, Amabile and Pratt (2016) have noted that intrinsic motivation is a stronger predictor of creative behavior than extrinsic motivation, eliciting greater persistence in the face of obstacles in creative tasks. Persistence correlates positively to creativity (Nijstad et al. 2010). Moreover, when customers are intrinsically motivated to pro-environmental action, they willingly exert more effort than the extrinsically motivated ones (Heyman and Ariely 2004). In other words, intrinsic motivation may not trigger an economic rationale, and it remains more motivating even for less appealing time slots. Therefore, we propose the following hypothesis:

Hypothesis 4 *The longer the time slot, the stronger the effect of green labels as compared to the effects of financial incentives.*

Overview studies

We conducted two experimental studies and two simulation studies to test our hypotheses and examine their impacts on the operational system in last-mile delivery. The first empirical experiment tests the first three hypotheses in a setting where participants select from time slots of equal duration, some of which are incentivized. We used the effects found in this study in a simulation experiment enlisting a dynamic approach to incentivize time slots and assess impacts on the operating system. A second empirical experiment aimed to replicate the effects found in the first experiment and also test Hypothesis 4 in a setting where participants opt between short and longer, incentivized time slots. A second simulation translates these findings into effects on the operating system.

3 Empirical study 1 – incentivizing identical time slots

Experimental design

In this study, we tested our first three hypotheses: whether adding green labels makes time slots more likely to be selected (Hypothesis 1); whether participants who are more eco-conscious are more likely to select a green-label time slot (Hypothesis 2); and whether the effect of the green label fades when merged with a price incentive (Hypothesis 3). We operationalized the “green label” by using a green-colored label accompanied by the following explanation: *“In the green time slots a delivery van will already be in your area. So, booking it will help to save fuel and reduce emissions.”*

We designed the study for high external validity and encouraged participants to reveal their actual views. We staged our study as a validation study of a new survey (for the eco-consciousness survey, see below). Before showing participants the survey, we told them that by participating in the study, they also enter a lottery where they may win an actual home delivery in the time slot of their choice. Time-slot selection embodied the actual experiment. We emphasized that they must be at home during the entire time slot to receive the goods. To make the incentives as realistic as possible, participants could win \$100 of groceries. When selecting a discounted slot, the discount would be added to the shopping voucher. This ensured that the study concerned choosing the delivery time for an actual (probabilistic) delivery.

In particular, we used the following text: *“Thank you for filling out our survey. As a token of our appreciation, you will have the chance to win a \$100,- voucher to shop for groceries at a major online grocery retailer in your area. To schedule the delivery of your groceries in case you win, we now ask you to select a day (from Monday to Friday) and a one-hour time window from a menu (from 10am to 8pm). Note that if you win, the groceries of your choice will be delivered in this time window so make sure that you are home then. A standard delivery fee (\$8) will be paid from your voucher. ”*

Participants had to select one time slot from a set of ten non-overlapping one-hour time slots for delivery from 10am to 8pm. This stylized, discrete-choice experiment reflects the main service platform of an online grocery retailer. We designed different experimental conditions where we randomly assign labels or price incentives for two of the ten time slots. We refer to these manipulated slots as our *focal* slots. The other time slots displayed no labels or price incentives. We then recorded the number of participants who selected one of the focal slots

in each condition. Figure 1 shows an example of the time-slot menu where the two focal slots feature green labels.



















Figure 1: Example for condition “Green, Green”

In this study, we enlisted nine different experimental conditions. To avoid carryover effects, we used a between-subjects design with each subject randomly assigned to one of the conditions. In the “No Label” control condition, no slots displayed a label or a price incentive. Both focal slots featured a green label in the “Green, Green” condition. We used three discount conditions offering a price incentive on the deliver fee of \$2, \$5, or \$8. The \$2 (25%) discount is common in practice. Since the default delivery fee in practice is about \$8, we added the “-\$8, -\$8” condition to represent a setting with free delivery on the focal slots (at a very high discount). We also add a discount condition between them. We further included a condition where both focal slots combine the green label *and* a price incentive of \$2 (Green-\$2) to investigate any synergy between these types of incentives. Note that in these six conditions, all focal slots have the same label or price incentive. However, given that actual consumers often have a choice among various incentives, we added three more conditions for two focal slots featuring different mixes of green labels and \$2 price incentives (since \$2 price incentives are most common, and further discounts would have created a large, unwieldy study). Table 2 presents an overview of the conditions array.

To avoid sampling bias in setting the focal slots, we randomly chose two out of ten slots as focal slots for each participant. For each condition, we have either 45 different focal slot settings when participants were offered with focal slots of the same label or 90 different focal slot settings when participants saw focal slots with two different labels.

At the end of the survey, participants were asked to fill out a questionnaire to assess their eco-consciousness. The questionnaire comprised nine questions based on related measurement methods in Rivera-Becerra and Lin (1999) and Alsmadi (2007). Participants were asked to rate the extent of their care for the environment on a five-point scale, i.e., “strongly agree” (5) to

Table 2: Overview of conditions for study 1

Condition	Focal 1		Focal 2	
Green, Green	Green		Green	
-\$2, -\$2	-\$2		-\$2	
-\$5, -\$5	-\$5		-\$5	
-\$8, -\$8	-\$8		-\$8	
Green-\$2, Green-\$2	Green and -\$2		Green and -\$2	
Green, -\$2	Green		-\$2	
Green-\$2, Green	Green and -\$2		Green	
Green-\$2, -\$2	Green and -\$2		-\$2	

“strongly disagree” (1).

Participants

We operated our experiment in Qualtrics¹ and distributed it via Mechanical Turk in January 2020. Each participant was randomly assigned to one of the conditions. A total of 1032 subjects participated in this study. All of the participants were based in the United States with an average age of 37 and 54% being female. Table 3 lists the characteristics of the participants in each condition.

Results

To test our hypotheses, we conducted two types of tests. First, we directly compared the percentages of participants choosing focal slots among the various conditions. The number of participants opting for one of the focal slots in the different conditions is shown in Figure 2.

Second, to study the choice behavior in more detail, we ran a multinomial (MNL) regression analysis to estimate the utilities for the different slots featured with variety of labels and price incentives. The nominal dependent variable corresponds to ten possible time slot choices. We enlisted binary variables to model the different labels and/or conditions.

Table 4 presents results for three distinct MNL model specifications. Model 1 used a binary variable for the focal slot manipulations per condition, e.g., one variable for “Green, Green” and

¹<https://www.qualtrics.com/>

Table 3: Study 1 - Sample composition per condition

Conditions	Sample	Female	Avg. Age
Green, Green	118	58%	38
-\$2, -\$2	117	55%	39
-\$5, -\$5	106	51%	36
-\$8, -\$8	105	51%	37
Green-\$2, Green-\$2	118	48%	36
Green, -\$2	118	64%	38
Green-\$2, Green	118	55%	39
Green-\$2, -\$2	117	52%	37
No Label, No Label	115	54%	35

one for “-\$2, -\$2”. Model 2 used a binary variable per label type, e.g., one for “Green” and one for “-\$2”. Note that we use one variable for the three discount labels in Model 2. We controlled for varied time-slot popularity in Model 1 and 2. Intercepts can be found in the appendix. We fit Model 3 on the data of condition “Green, Green” and “-\$2, -\$2” without intercepts per slot choice. This simplified the model by reducing the number of variables. We then used the output of this model as input for the simulation study in Section 4.

For ease of exposition, we later discuss the results for the different conditions associated with the different hypotheses in separate subsections.

The effect of the green label

Figure 2 shows that substantially more participants selected a focal slot with a green label (81.4%) versus no label (22.6%). This result is statistically significant ($\chi^2; p < 0.0001$). We also observe that the green-labeled slots were selected more often than those with price incentives of any amount ($\chi^2; p < 0.001$). Moreover, in a setting where participants can choose between a slot with a green label and a price incentive (the “Green, -\$2” condition), we see that more chose the slot featuring the green label (49.2%) versus the \$2 discount (22.9%). This difference is statistically significant ($\chi^2; p < 0.0001$).

The results from the MNL model are similar. Even while explicitly controlling for the popularity differences among slots (Model 1 and 2), we see that the green labels yielded a significant uplift in the utility of a time slot. This again supports the notion that green labels

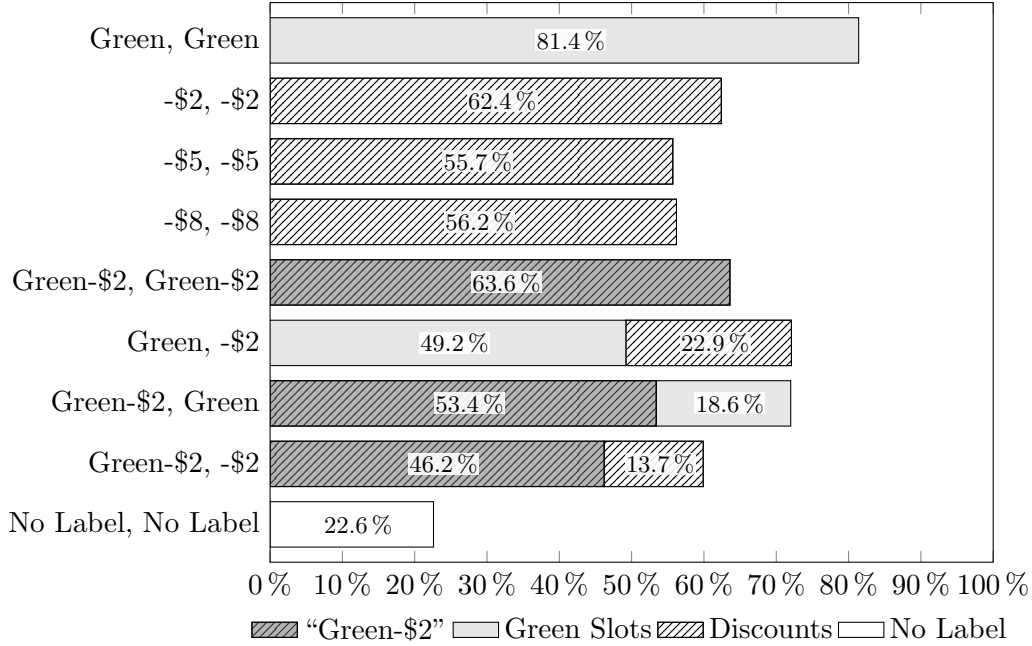


Figure 2: Number of participants that selected a focal slot

raise the likelihood of slot selection. Moreover, we observe that the utility of a green label outperforms that of a price incentive. These results strongly support Hypotheses 1.

The effect of the combined green label and price incentive

Figure 2 shows that fewer participants selected a focal slot in the “Green-\$2, Green-\$2” condition (63.6%) than in the “Green, Green” condition (81.4%) under statistical significance ($\chi^2; p < 0.01$). This validates results from MNL Model 1 (Table 4) showing the utility of the “Green-\$2” label (1.94) underperforming that of the green-only label (2.88). Furthermore, results show that the number of participants choosing a focal slot was roughly the same for the mirrored ‘Green-\$2, Green-\$2’ and ‘-\$2, -\$2’ conditions where the difference is not statistically significant ($\chi^2; p > 0.1$). These results support Hypothesis 3.

Notably, when participants can choose between a combined slot and a slot with a green label or price incentive (Green-\$2, Green vs. Green-\$2, -\$2), we observe more participants choosing the combined slot ($\chi^2; p < 0.0001$). This is consistent with the results of MNL Model 2 in Table 4 showing lower utility for the green label (2.30) versus the ‘Green-\$2’ label (2.51).

Eco-Consciousness

We next investigated the impact of eco-consciousness on participant time slot choice. We grouped the participants into a high eco-conscious group (comprising those outscoring the me-

Table 4: MNL models for time slot choices

Conditions	model 1	model 2	model 3 ^a
Green, Green	2.88***	-	2.86***
-\$2, -\$2	1.91***	-	1.89***
-\$5, -\$5	1.62***	-	-
-\$8, -\$8	1.65***	-	-
Green-\$2, Green-\$2	1.94***	-	-
Green, -\$2	2.32***	-	-
Green-\$2, Green	2.32***	-	-
Green-\$2, -\$2	1.78***	-	-
No Label, No Label	0.06	-	-
Labels			
Green	-	2.30***	-
Green-\$2	-	2.51***	-
Discount ^b	-	1.64***	-
R^2	0.20	0.20	0.29

* p-value < 0.05; ** p-value < 0.01; *** p-value < 0.001

^aThis model is based on the data of the condition green and the condition -\$2.

^bIt includes -\$2, -\$5, and -\$8 discounted labels.

dian eco-scores) and a low eco-conscious group (comprising those scoring lower than the median in eco-scores). We present participant choices for the green label, the \$2 price incentive, and the “Green-\$2” label under the six conditions in Figure 3.

Figure 3 depicts \$2 price incentives as chosen significantly more often by low eco-conscious participants (71.7%) than by highly eco-conscious (54.7%) ($\chi^2; p < 0.05$) - even though the highly eco-conscious chose green labels just a bit more often than low eco-consciousness scorers. Compared to participant choices in the price-only “-\$2, -\$2” condition, adding a green label (Green-\$2) significantly prompted more high eco-conscious participants to choose focal slots (71.2%, up from 54.7%) ($\chi^2; p < 0.05$). The green-label addition also significantly cut the likelihood of low eco-conscious participants choosing focal slots (dropping from 71.7% to 55.9%) ($\chi^2; p < 0.05$). The “Green, -\$2” condition makes clear that highly eco-conscious participants

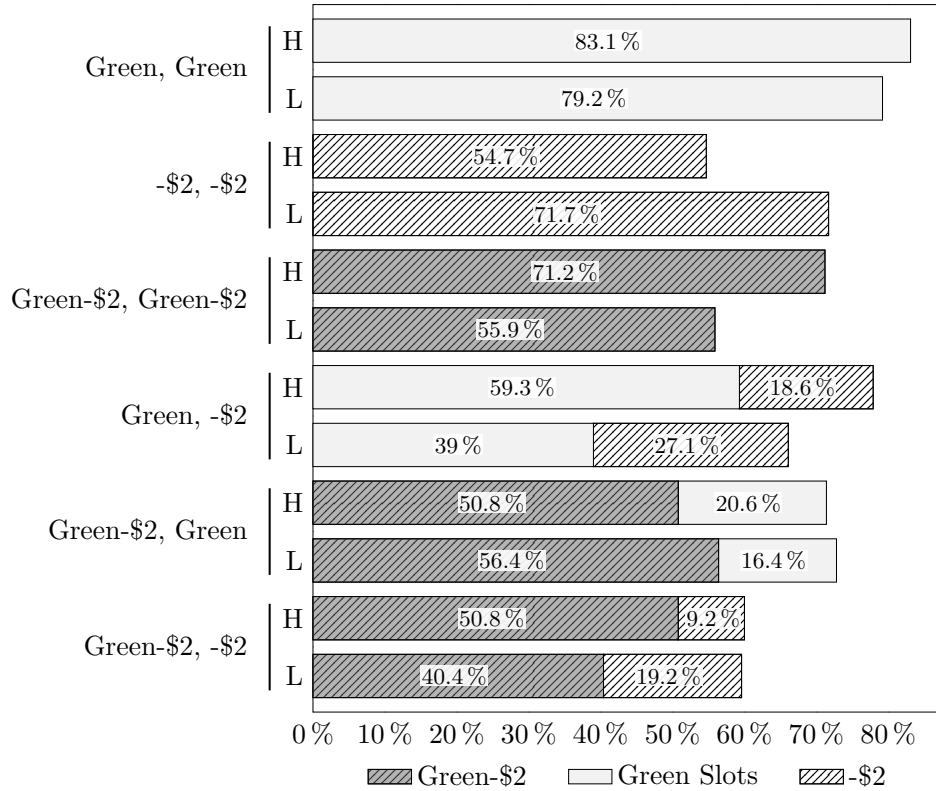


Figure 3: Eco-consciousness^a and participant choice

^aIn this figure, we use “H” and “L” to indicate results from high eco-conscious participants and low eco-conscious participants, respectively.

(59.3%) are more likely to choose a green slot versus low eco-conscious counterparts (39%) ($\chi^2; p < 0.05$). The “Green, -\$2” condition also shows that highly eco-conscious participants (59.3%) choose green slots significantly more often than \$2 price incentives (18.6%) ($\chi^2; p < 0.001$). Yet this is not the case for low eco-conscious participants (39% versus 27.1%) ($\chi^2; p > 0.05$). These results strongly support Hypothesis 2.

Discussion

The findings of Study 1 highlight the potential influence of green labels and attest their relation to intrinsic motivation. First, the results show green labels exerting significant impact on slot choice. Adding green labels not only greatly boosts slot appeal over those without incentive; green labels also outperform very substantial price incentives! These results demonstrate Hypothesis 1. Second, analysis of the effects of eco-consciousness identify highly eco-conscious participants as more prone to select green-labelled slots. These results support Hypothesis 2, further suggesting, as outlined in the theory section, that the effects of green labels may work via

intrinsic motivation for people having strong intrinsic motivation to improve the environment.

Third, we find there is little to no added value in merging green labels with price incentives. Adding price incentives to one or both of the green labelled slots does not encourage more participants to select a focal slot, thus confirming Hypothesis 3. Interestingly, even though the condition with one green label and one green label with price incentive lead to less people choosing focal slots than the condition with two green labels, in the prior condition a majority of participants choose the green label with price incentives over the green label. This suggests that although adding price incentives to green labels will not persuade more people to choose a focal slot, people already preferring a focal slot will be more likely to go for the combined green label - price incentivized slot.

Two things are noteworthy. First, the advantage of green labels over price incentives is that they are cost-free to the provider, while a two- to eight-euro price incentive evaporates small profit margins. This advises that, given their influence on outcomes, it may be better for companies to “go green.” Before making definite conclusions, though, we need more comprehensive insight into the effects of green labels on the delivery system. The next section thus details a simulation to input findings from this experiment that mimics a delivery system to test the potential effects of using green labels. Second, if intrinsic motivation is the driving force behind green labels, we may expect their effects (versus extrinsic motivation via price incentives) to hold even when slots become less attractive (see Hypothesis 4). We test this idea in a second experiment.

4 Simulation study 1 - dynamic time slots of equal length

Here, we use simulation experiment to study the impact of the green slots and price incentives that dynamically steer customers to reduce system-wide vehicle usage for last-mile delivery operations. We draw a customer time slot choice model based on Model 3 (see Table 4) to study the effect of using the green label or the \$2 price incentive versus no incentives in reducing distances of the last-mile delivery. We choose these two incentive types since they act as the single best incentives (pairing them seems clearly less beneficial than using green labels alone).

4.1 Dynamically incentivizing time slots

In this section, we describe how to dynamically determine which slots to incentivize based on already accepted orders. The key idea is to steer customers toward the (focal) delivery time

slots to help minimize the system-wide vehicle kilometers that trigger proportional variable costs (e.g., labor, maintenance, and fuel) and environmental impact (e.g., emissions). Based on the tentative route plans for the already accepted orders, we feature for each new customer time slots associated with the smallest detour. This means that the set of focal slots differs per customer and over time.

Customers continuously book their delivery time slots. During a booking period, we maintain a route plan with vehicles serving prior-scheduled orders within their selected time windows. When a new customer selects a time slot, we know the delivery location and shopping basket size. To quickly assess the feasibility and costs of accepting a customer in each of the different time window, we use cheapest insertion (Campbell and Savelsbergh 2004). Since customers do not like waiting and this approach is very fast, it is often used to dynamically solve the underlying routing problems in real-time (Campbell and Savelsbergh 2005, 2006, Ehmke and Campbell 2014).

We determined the focal slots by using the insertion approach as follows. Assign $c_{i,j}$ as the distance from i to j . The additional routing distance of inserting customer u between customers i and j is given by $c_{i,u} + c_{u,j} + c_{i,j}$. We found the cheapest insertion by trying out all positions in each route and taking the feasible position with the shortest distance. We featured two slots with the lowest insertion values as the focal slots, unless differences among insertion distances for other slots were less than α . In our experiments, $\alpha = 3$ kilometers reflected cost savings of approximately \$2 equaling the loss revenue of the \$2 price incentive.

The route plan is optimized each iteration by re-planning all routes using greedy cheapest insertion after each customer has selected a certain time slot. When it is impossible to find a new feasible schedule in this way, we inserted the customer into the existing route plan.

4.2 Simulation setup

Inspired by real-world e-grocery order data, we randomly generated 100 base instances, each having 400 customers uniformly distributed over an area of 30 km \times 30 km (900 sq.km). Customers arise singly in a predetermined queue to choose a time slot that maximized shopper utility, scored for each customer referencing base utilities for each time slot estimated in Section 3, *plus* a random error term per customer slot. We drew these random error terms from the standard extreme value Gumbel distribution (Ben-Akiva and Bierlaire 1999).

We varied instance densities (low, medium, high) by selecting different subsets of customers

from the base instances. The highest density used all customers in the base instances. The medium density used the first 200 customers, and the low case used the first 100 customers. Here, we reduced variance among the different experiments since the set of customers (and associated preferences) of the lower density instances was simply a subset of the higher density instances. This effectively meant that the first 100 customers were the same in each density scenario.

For each instance, we ran three computational experiments using green labels, \$2 price incentives, or no labels on two of ten one-hour time slots between 10am and 8pm. Customer selected one time slot from the ten options, consistent with the setting used in our empirical studies.

Next, we ran computational experiments on labels for different effects in driving customers toward focal slots. We tallied a ratio of customers selecting labeled slots to indicate the effect level. Effect ratios for green labels, \$2 price incentives, and no labels are 81%, 62%, and 20%, respectively. In order to gain insights into additional levels of incentive effectiveness, we also ran experiments featuring extremely high effect level (100%) and low effect level (40%). We enlisted the 100 medium customer density setting for these experiments.

Each vehicle's delivery capacity was 20 customers traveling at constant speed of 30 km/hour. Vehicles were deployed in two shifts: an early shift (from 9am to 4pm) and a late shift (2pm - 9pm). The service time for each customer was 10 minutes.

4.3 Results

We rated the final route plans using the distance per customer and the number of vehicles used. We compared the percentage savings to the reference no label setting. Table 5 shows the average results for the different customer densities and labels over 100 random instances.

Results clearly show that steering customers helped to reduce both travel distances and number of vehicles used. As more customers chose focal slots, more distance savings accrued. We also note that green labels outperformed price incentives - the added influence on customer behavior from green labels pays off! Distance savings likely translates to costs savings in terms of fuel and labor. The green label thus yields greater cost savings without any loss in revenue using price incentives. Offering \$2 price incentives cuts revenues by \$1.12-\$1.15 per customer on average.

Results suggest that dynamically steered time-slot choices work better when densities are

low. For green labels, savings are higher (12.6%) for low density than for high-density instances (8.9%). A similar pattern is observed for price incentives. One explanation for this is that the average intra-customer distances (and savings) fall with density. The marginal gain abates when increasing the spatial and temporal clustering of customer where distances between them are already low.

While Table 5 shows the average savings, Figure 4 specifies per-instance savings for the different runs. At one end of the spectrum, savings reach 400 km (30%) for low densities. On the other end, some cases of customer steering can actually *raise* the total system-wide travel distances, especially when label effect levels weaken. When realizing that slotting decisions were made for each customer without knowing future demand, we expect uncertain impacts in decisions. We still observe, given green label's higher effectiveness, the negative impacts were less frequent (and even non-existent in high density conditions) with green labels than with price incentives. Overall, the simulation provides strong support for the positive effect of green labels on the operational performance.

To test how sensitive results are to the incentive efficacy, we ran another simulation study for medium-density instances. Figure 5 shows results from four different effect-sizes: 62% corresponding to the effectiveness of the price incentive, 81% to green-label effectiveness, 100% level for all customers selecting the incentivized focal slot, and the less-strong 40% effect. As expected, we see that savings increase with incentive effective size. However, results further suggest that savings can still be achieved even when the effect is modest (i.e., 40%).

Table 5: Simulation results for different densities and incentives

Instance	Dens ^a	KM ^b	Δ KM ^c	$\Delta\%$ KM ^d	Veh ^e	Δ Veh ^f	$\Delta\%$ Veh ^g
Green, Green	100	10.43	1.50	12.6%	9.77	1.00	9.3%
Green, Green	200	8.44	0.88	9.4%	15.89	1.06	6.3%
Green, Green	400	6.67	0.65	8.9%	26.33	1.46	5.3%
-\$2, -\$2	100	10.99	0.95	7.9%	10.10	0.67	6.2%
-\$2, -\$2	200	8.57	0.75	8.0%	15.88	1.07	6.3%
-\$2, -\$2	400	6.82	0.51	6.9%	26.69	1.10	4.0%

^athe number of customers in each instance

^bthe travel distance per customer in average

^cthe average travel distance savings per customer (taking the results of “No Label” as the benchmark)

^dthe percentage of travel distance saving (taking the results of “No Label” instances as the benchmark)

^ethe number of vehicles used in instances in average

^fvehicle savings of instances in average

^gthe percentage of vehicle savings in average

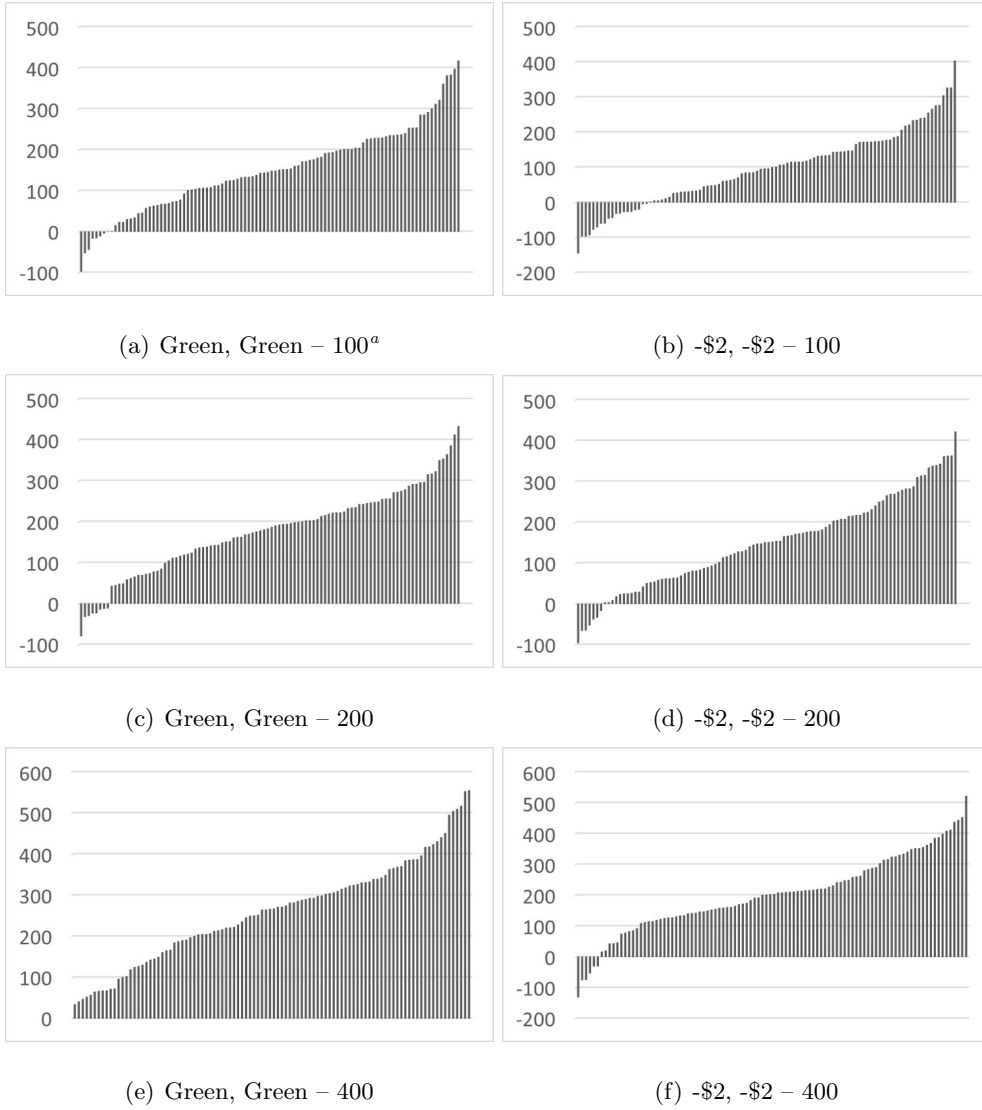


Figure 4: System-wide savings in vehicle distance (km) per instance

^a“Green, Green - 100” indicates vehicle distances of the “Green, Green” instances with 100 customers which corresponds to the first row in Table 5.

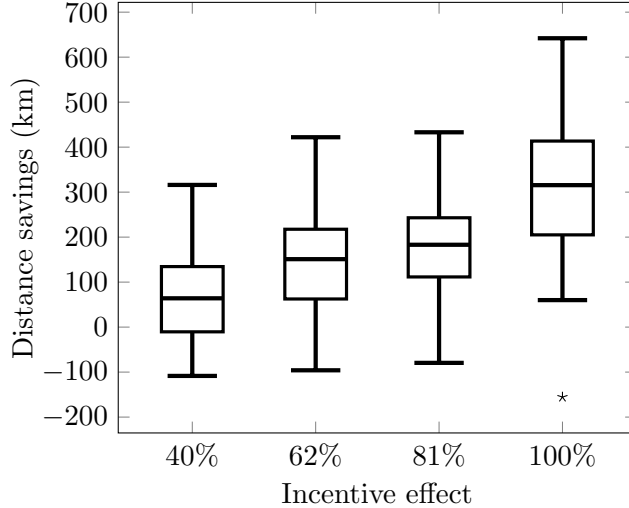


Figure 5: System-wide distance savings for different incentive effects

5 Empirical study 2: incentivizing time slot with different lengths

Experiment design

Study 1 tested the effect of green labeled delivery time slots of equal length. Study 2 investigates the effect of using green labels to steer people to longer, less convenient time slots. More specifically, we study whether green labels would be more effective than price incentives in drawing participants to longer time slots. We use the same design as in Study 1 except for the time slot manipulation. In this study, we offer participants the choice of unincentivized slots of one hour versus incentivized slots of varying lengths (two-, three-, or five-hour options). We predict that the longer the incentivized time slot offered, the less strong the effectiveness of price incentives versus green labels (Hypothesis 4). In other words, “green labels” will outperform discounts for longer time slots.

As with Study 1, we enlisted a between-subjects design to avoid carryover effects in this study. We asked each participant to choose one from a menu of slots for receiving groceries. The menu featured ten, one-hour slots (as in Study 1) and a number of overlapping longer slots (two-, three-, or five-hour; see Table 6). We next manipulated the slot length of incentivized time slots for each participant. Referencing a standard delivery fee of €8, we offered long slots using either a green label or a €2 price incentive. Overall, six different conditions were featured. Figure 6 shows examples of time slot menus used in this study.

As in Study 1, we presented a green-colored label explained this study: “*By choosing a longer*

Table 6: Study 2 – Overview of conditions

SlotID	2-Hour	3-Hour	5-Hour
11	10:00 - 12:00	10:00 - 13:00	10:00 - 15:00
12	12:00 - 14:00	13:00 - 16:00	15:00 - 20:00
13	14:00 - 16:00	16:00 - 19:00	13:00 - 18:00
14	16:00 - 18:00	11:00 - 14:00	-
15	18:00 - 20:00	14:00 - 17:00	-
16	-	17:00 - 20:00	-

time window, you provide the retailer with more routing flexibility which allows them to save fuel and reduce emissions.” At the end of the survey, we asked participants to fill out the same questionnaire as in Study 1 to assess levels of eco-consciousness.

Participants

We implemented our experiments in Qualtrics.¹ A total of $n = 442$ subjects participated in this study (see Table 7). The average age was 20 with 57% participants female. Participants in this study were students of a large University in the Netherlands.

Table 7: Study 2 - Sample composition per instance

Conditions	Sample	Female	Avg. Age
Green, 2-Hour	76	45%	22
-€2, 2-Hour	75	65%	22
Green, 3-Hour	69	54%	19
-€2, 3-Hour	74	55%	20
Green, 5-Hour	73	60%	19
-€2, 5-Hour	75	61%	19

¹<https://www.qualtrics.com/>

10:00 – 11:00 €8	11:00 – 12:00 €8	12:00 – 13:00 €8	13:00 – 14:00 €8	14:00 – 15:00 €8	15:00 – 16:00 €8	16:00 – 17:00 €8	17:00 – 18:00 €8	18:00 – 19:00 €8	19:00 – 20:00 €8
10:00 – 12:00 €8		12:00 – 14:00 €8		14:00 – 16:00 €8		16:00 – 18:00 €8		18:00 – 20:00 €8	

(a) One and 2-hour time slots

10:00 – 11:00 €8	11:00 – 12:00 €8	12:00 – 13:00 €8	13:00 – 14:00 €8	14:00 – 15:00 €8	15:00 – 16:00 €8	16:00 – 17:00 €8	17:00 – 18:00 €8	18:00 – 19:00 €8	19:00 – 20:00 €8
10:00 – 13:00 €8			13:00 – 16:00 €8			16:00 – 19:00 €8			
	11:00 – 14:00 €8			14:00 – 17:00 €8			17:00 – 20:00 €8		

(b) One and 3-hour time slots

10:00 – 11:00 €8	11:00 – 12:00 €8	12:00 – 13:00 €8	13:00 – 14:00 €8	14:00 – 15:00 €8	15:00 – 16:00 €8	16:00 – 17:00 €8	17:00 – 18:00 €8	18:00 – 19:00 €8	19:00 – 20:00 €8
10:00 – 15:00 €6					15:00 – 20:00 €6				
	13:00 – 18:00 €6								

(c) One and 5-hour time slots

Figure 6: Example of conditions with longer slots

Results

Figure 7 shows the booked time slot choices. We see that both green labels and price incentives helped to steer participants toward the longer slots. The effect decreased when slots became longer and less convenient.

Importantly, we see varied effectiveness of the incentive options over different slot widths. While green labels and price incentives worked equally well in steering participants toward two-hour slots (price incentives did even slightly better), we see that green labels outperform price incentives in steering participants to three-hour and five-hour slots. In general, the pattern suggests that the effectiveness of green labels versus price incentives is less influenced by slot length, in line with hypothesis 4. This effect is most manifest when comparing differences in the effectiveness for green labels versus price incentives for the 2-hour and 3-hour conditions. Green labels were as effective in promoting 3-hour slots (81.2%) as 2-hour slots (85.5%; $\chi^2; p > 0.05$). Price incentives, though, seem obviously less effective in steering into 3-hour slots (71.6%) versus 2-hour slots (86.7%; $\chi^2; p < 0.05$). This effect faded statistically in the 5-hour slot conditions where significantly fewer participants chose 5-hour slots over the 2-hour in both the discount condition (61.3% and 86.7% respectively; $\chi^2; p < 0.001$) and the green labels condition (65.8% and 85.5% respectively; $\chi^2; p < 0.001$), with slightly greater drop in the discount incentive

efficacy for longer time slots.

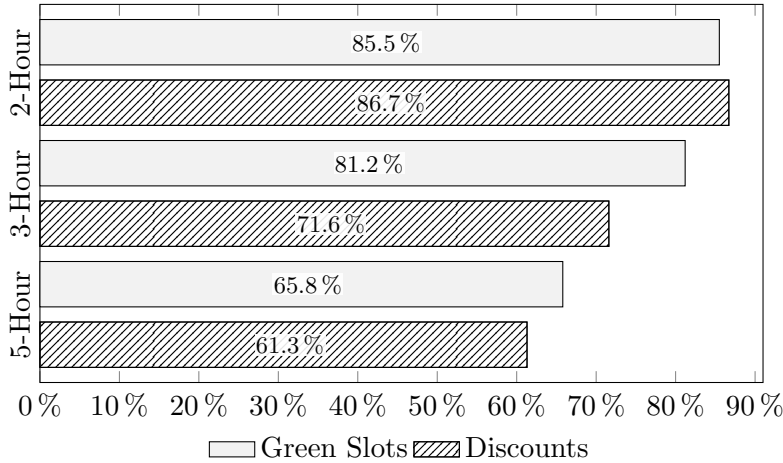


Figure 7: Number of participants that selected a longer slot

Eco-consciousness

To show the impact of eco-consciousness on participant choice, we grouped participants into high eco-conscious (ranking above the median) and low eco-conscious groups (ranking below median) according to their eco-scores. We present slot choices of the two groups in Figure 8.

The results for *two-hour slots* show *no significant difference* in choices of high eco-conscious versus low eco-conscious participants, either for the green label or discount incentive.

Again, however, this pattern changed for the three-hour slot. Here, green labels significantly steered high eco-conscious (93.9%) over low eco-conscious participants (69.4%) to choose the three-hour slots ($\chi^2; p < 0.01$). No such effect arose for the discount incentive. Results also show that high eco-conscious participants significantly chose longer slots more often in the green label (93.9%) versus discount condition (74.3%) ($\chi^2; p < 0.05$). No such effect emerged for low eco-conscious participants.

We did not detect significant differences in longer slot (five-hour) choices between high and low eco-conscious participants, either for green labels or price discounts.

In sum, two-hour slots were indifferent to incentive types. The three-hour slots clearly favored green labels, while price steered notably fewer low eco-conscious shoppers to tolerate the five-hour slot. Together, these results provide some support for Hypothesis 4.

Next, we present MNL models of participant choice for different labels and varied slot lengths in Table 8. For two-, three-, and five-hour slots, different MNL models were enlisted for

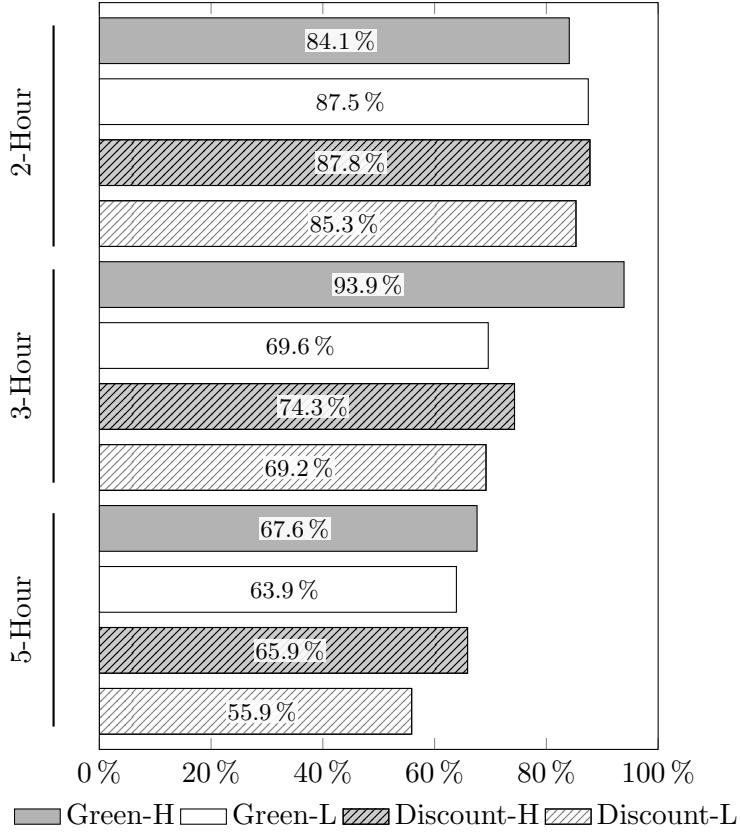


Figure 8: Eco-consciousness^a and participant choice

^aIn this figure, we use “H” and “L” to indicate high and low eco-conscious participants, respectively, e.g., “Green-H” indicates higher eco-conscious participant choice on the green labels; “Discount-L” indicates low eco-conscious participant choice on the price incentives.

estimating different label utilities. The MNL models for two-hour slots are quite similar, and we forged a joint MNL model for green labels versus price incentives in Appendix 2.

We note that participants disclosed different preferences for specific long slots. The latest slots in the afternoon especially (including slot 15 among 2-hour options, slot 16 among 3-hour options, and slot 12 among the 5-hour choices) proved most popular. Accordingly, we favored estimating separate label utilities for each of the longer slots.

From the MNL models in Table 8, we see that the two-hour utilities *as a whole* generally exceeded those of the three- and five-hour slots, as expected. Green label utility mostly prevailed over price-incentive utility for three- and five-hour slots. These results confirm those in Figure 7.

Table 8: MNL models for time slot choice of different lengths

SlotID ^a	2-Hour		3-Hour		5-Hour	
	Green	-€2	Green	-€2	Green	-€2
11	2.95***	3.00***	1.82***	1.34**	1.97***	1.57***
12	2.21***	2.40***	1.68***	0.36	2.08***	1.77***
13	2.39***	1.61**	1.53**	1.46***	1.39***	1.64***
14	1.51**	2.40***	1.53**	1.46***	-	-
15	2.74***	2.89***	1.68***	0.64	-	-
16	-	-	2.83***	2.25***	-	-

^aSee the indexes of time slots in Table 6

Discussion

In Study 2, we investigated how effective green labels and price incentives were in steering people to pick two-, three- or five-hour slots over available one-hour slots. Results show notable insights. First, we find green labels and price incentives are both effective in driving participants to longer time slots, but that incentives abate for longer slots. Second, we find that this pattern of lower effectiveness for longer slots to diverge for green labels versus price incentives. Specifically, green labels remained relatively more effective than discounts when incentivized time slots lengthened. However, this green label advantage vanished in the longest (5-hour) time slots. This incentive pattern supports hypothesis 4, albeit with the caveat that the effect disappears with very long time slots. A reason is that waiting at home for five hours is too inconvenient that many people cannot make it.

Our analysis of eco-consciousness reveals that participants high in eco-consciousness respond better to green labels, but only in the three-hour time slot condition. This offers some, but not full support for Hypothesis 2. Our study reveals green labels to be useful, in some cases more so than using price incentives, in steering people toward longer time slots - an effect highly relevant in practice. In order to assess the practical impact of these effects in a delivery system, we finally conducted a second simulation experiment described in Section 4.

6 Simulation study 2: static time slot with different lengths

In this section, we simulate customer time slot choices of *different* widths based on MNL models in Table 8 to study the impact of green labels and price incentives. The goal remains to minimize system-wide vehicle kilometers for last-mile delivery.

Longer time slots offer more planning flexibility and thus outperform the tighter time slots. Contrasting the dynamic labeling policy in Section 4, we used static labeling in this experiment. Only *after all customers have selected* a time slot do we finalize a route plan attaining the cheapest insertion heuristic. As with the prior study, each participant again received a time slot menu with one-hour slots amid longer slots incentivized for all customers.

6.1 Simulation setup

We ran 100 instances with medium customer density (200 customers in a 30 km \times 30 km area) from Section 4 using the slot settings of Study 2 (including one-, two-, three-, and five-hour slots). Customer slot choice models were based on the MNL models in Table 8. As with Section 4, we used Gumbel random variables to simulate the random utilities of slots to customers.

We ran six computational conditions: 1) a joint experiment for two-hour slots to simulate both green labels and price incentives (due to similar effects), 2) using only green labels for three-hour slots, 3) using only price incentives for three-hour slots, 4) using only green labels for five-hour slots, 5) using only price incentives for five-hour slots, with all five of these against 6) unlabeled one-hour slots serving as a control reference for comparison.

Parameters were the same per Section 4: every vehicle (either in the first shift from 9am - 4pm, or in the late shift from 2pm - 9pm) served at most 20 customers with 10 minutes as a fixed service time for each customer, traveling at a constant speed of 30 km/hour.

6.2 Results

We use the same metrics per Section 4 to assess the performance of different labels on long slots (versus the controlled only one-hour slot benchmark). Table 9 lists the average results from different labels for long slots of different widths in over 100 random instances.

As expected, when more participants selected a longer slot, this improved the performance of the system in terms of the total travel distances. However, slightly more vehicles were needed for the two-hour case versus the control group, since offering two-hour slots clustered a large

portion of customers into a few time slots, e.g., 32% of the customers booked time slots between 10:00 and 12:00 (including Slot 1, Slot 2, and Slot 11). Here, more vehicles were needed than where time slots might have been more equally selected by customers.

Results show that featuring the green label on three- and five-hour slots enabled shorter distances than using the price incentive. We find that using the green label reduced travel distance for three-hour slots 7.1% more (20.4% savings) than the price incentive (13.3%). Using the green label for five-hour slots cut travel distance by 3.3% more (21.8%) than the price incentive (18.5%). In contrast, the €2 price incentives to customers choosing longer slots *lost revenues* of €1.72 for the two-hour option, €1.43 for the three-hour case, and €1.22 for the 5-hour instance, on average.

Table 9: Simulation results for different time slot lengths and incentives

Instance	KM	Δ KM	$\Delta\%$ KM	Veh	Δ Veh	$\Delta\%$ Veh
Green, 2-Hour	8.83	0.83	8.9%	17.06	-0.04	-0.2%
-€2, 2-Hour ^a	8.83	0.83	8.9%	17.06	-0.04	-0.2%
Green, 3-Hour	7.56	1.91	20.4%	14.16	2.86	16.8%
-€2, 3-Hour	8.13	1.24	13.3%	15.24	1.78	10.5%
Green, 5-Hour	7.33	2.04	21.8%	14.30	2.72	16.0%
-€2, 5-Hour	7.64	1.73	18.5%	14.65	2.37	13.9%

^aThe same customer choice model based on the MNL model in Appendix 2 is used for the instances in “Green, 2-Hour” and “-€2, 2-Hour” due to similar effects between the two.

7 Discussion and conclusions

In this paper, we describe two experimental studies and two simulation studies investigating the role of green labels in steering demand in last-mile delivery. Our studies provide five main findings. First, green labels are an effective tool to manage demand and work as well as commonly used price incentives. Second, eco-conscious customers are more responsive to green labels. Third, combining green labels and price incentives yields little added effect over green labels alone. Fourth, green labels work better than price incentives in steering people toward longer time slots. Fifth, the varied effects of green labels exert quite strong impact on downstream operational system performance. These results contribute in several important

ways.

First, the research offers a theoretical contribution toward the field of demand management in operations management and beyond. This is, to our knowledge, the first study to demonstrate that green labels are effective in steering demand in last-mile delivery and strongly impacts the operating system of home delivery. This is notable amid all the research on the customer side of demand management currently focused on price incentives. Our studies demonstrate that there are alternatives to price incentives that are just as effective. Equally important, however, are the findings that pertain to the underlying processes at work as green labels steer customers toward specific time slots. Green labels strongly engage more eco-conscious customers, demonstrating a value-alignment process that intrinsically motivates. This has several key implications, prompting us to suggest, for instance, that price incentives and green labels would not synergize, and that green labels would be especially effective in steering towards “inconvenient” (longer) time slots. In this sense, our studies more broadly contribute to general literature on incentives that steer customer behavior.

In pondering and testing the nature of green labels, our studies more broadly contribute to general literature on incentives that steer customer behavior. Although extant literature casts green incentives as more embedded, our categorization of incentives as intrinsic versus extrinsic motivators helps us to better understand different incentives. In particular, our research findings provide a strong contribution to this broad field by suggesting (with supporting evidence) that *i*) incentives that steer customer behavior can be categorized as extrinsic motivators versus intrinsic motivators, *ii*) intrinsic and extrinsic motivators do not mix, and *iii*) intrinsic motivators outperform extrinsic motivators in steering customer behavior especially as customer business costs rise.

Our second contribution is more methodological in terms of modeling and research strategy. In terms of modeling, we designed a simulation model based on state-of-the-art time slotting and routing heuristics for an empirically validated customer choice model that quantified potential impacts in steering demand on the operating system. This means that we complemented extant studies exploring the impact of demand management approaches in the last-mile based on artificial and stylized customer demand data (Campbell and Savelsbergh 2006, Ehmke and Campbell 2014).

In terms of research strategy, our set of studies also contributes. Indeed, there has been recent emphasis in operations management on merging empirical studies with simulation in

research. Specifically, Chandrasekaran et al. (2018) have developed an overview of how simulations and empirical studies can synergize. In our type of research, theory testing research, Chandrasekaran et al. (2018) discuss how simulation can tackle issues of endogeneity in empirical (survey) studies. Our study combines the advantages of experimental study and simulation to thus overcome limitations of these methods in isolation and complement existing frameworks that blend surveys or case studies with simulation studies. While experiments do not suffer endogeneity issues, they can incur validity problems in large, complex systems. This is precisely where simulation can benefit. We thus recommend the complementary use of experimental and simulation studies in future operations management papers.

A third contribution is practical in nature. Operations management practitioners can well exploit of our findings. Indeed, our research strongly advises that using green slots either dynamically or by incentivizing longer slots yields strong positive effects on last-mile delivery operations in terms of costs and environmental criteria. Our simulation findings further caution that there may be instances where green labeling actually incurs more costs and emissions, but these cases are rare. Overall, using green slots is beneficial. Our empirical results suggest that green slots especially work well with more eco-consciousness customers. This insight, together with insights from the research on the relationship between eco-consciousness and demographic and cultural characteristics (Golob and Kronegger 2019, Gray et al. 2019), may help practitioners to predict the effect of green labels for specific customer populations. The way companies implement green slots, beyond dynamic or to incentivized longer slots, may also be important. When going “green” to steer demand to improve operations, we show that it is not beneficial to combine green slots and price incentives for the same slot. However, this does not imply that we advise against using pricing altogether. From a marketing perspective, pricing is still a powerful tool to exploit differences in willingness between different market segments increasing revenues. The design of the pricing and green incentives for last-mile delivery should take these different considerations into account.

In sum, the current research investigated the effects of green labels on customer slot choice for last-mile delivery. We find strong effects from green slots, providing a theoretical model and research results that suggest interesting nuances as to these effects: they work better for eco-friendly customer, they do not synergize with discounts, and they work especially well when the costs of slot choice rise to real impact on operating systems in last-mile delivery. We hope that this theory and these results not only provide important insights in the field of demand

management in last-mile delivery, and the wider field of green incentives, but will also stimulate other scholars to continue this line of research to unleash the full potential of green labels or incentives.

As this is the first paper that studies the use of green labels in last-mile delivery operations, we see several directions for future research at the interface between marketing and operations management. First, we would welcome any field research that would replicate our findings. Although the strength of experimental designs such as ours is that they yield causal evidence of effects, the downside may invoke validity issues. Using a field experiment to study the impact of green labels provides various methodological challenges given the many confounding factors that are associated with a dynamic and fast growing business such as online grocery. Second, the advent of omnichannel retailing provides relevant new applications for using green labels to steer customers between sales channels. This gives rise to new research questions on how to design and dynamically execute these incentives across different channels while taking into account both marketing and operational considerations. Third, we believe that the general idea of our paper - steering customers by providing more insight into the environmental consequences of their choices - may provide fertile ground for future research in many Operations Management applications.

Acknowledgement

We would like to thank the anonymous reviewers for their constructive feedback and efforts towards improving our paper. The co-author *** was funded by NWO as part of the Complexity in Transport and Logistics Programme.

Appendix 1. Intercepts of MNL models in Table 4

SlotID	model 1	model 2
1	0.24	0.23
2	0.30*	0.31*
3	0.16	0.18
4	0.09	0.09
5	0.00	0.00
6	-0.07	-0.05
7	0.22	0.24
8	0.26	0.26
9	0.13	0.13
10	-0.36*	-0.34*

* p-value < 0.05

Appendix 2. MNL model for “Green, 2-Hour” and “-€2, 2-Hour” instances of Table 9

SlotID	2-Hour
11	2.97***
12	2.30***
13	2.09***
14	2.03***
15	2.81***

*** p-value < 0.001

References

- Agatz, N., A. Campbell, M. Fleischmann, and M. Savelsbergh. 2011. Time slot management in attended home delivery. *Transportation Science*, 45(3), 435–449.
- Agatz, N., A. M. Campbell, M. Fleischmann, J. Van Nunen, and M. Savelsbergh. 2013. Revenue management opportunities for internet retailers. *Journal of Revenue and Pricing Management*, 12(2), 128–138.
- Agatz, N., A. M. Campbell, M. Fleischmann, and M. Savelsbergh. 2008. Challenges and opportunities in attended home delivery. In *The Vehicle Routing Problem: Latest Advances and New Challenges*, Springer, pages 379–396.
- Ajzen, I. 1991. The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211.
- Alsmadi, S. 2007. Green marketing and the concern over the environment: Measuring environmental consciousness of jordanian consumers. *Journal of Promotion Management*, 13(3-4), 339–361.
- Amabile, T. M. and M. G. Pratt. 2016. The dynamic componential model of creativity and innovation in organizations: Making progress, making meaning. *Research in Organizational Behavior*, 36, 157–183.
- Amornpetchkul, T., H.-S. Ahn, and Ö. Şahin. 2018. Conditional promotions and consumer overspending. *Production and Operations Management*, 27(8), 1455–1475.
- Ben-Akiva, M. and M. Bierlaire. 1999. Discrete choice methods and their applications to short term travel decisions. In *Handbook of Transportation Science*, Springer, pages 5–33.
- Bowles, S. 2008. Policies designed for self-interested citizens may undermine “the moral sentiments”: Evidence from economic experiments. *Science*, 320(5883), 1605–1609.
- Boyer, K. K., A. M. Prud’homme, and W. Chung. 2009. The last mile challenge: evaluating the effects of customer density and delivery window patterns. *Journal of Business Logistics*, 30(1), 185–201.
- Buchanan, K., R. Russo, and B. Anderson. 2015. The question of energy reduction: The problem (s) with feedback. *Energy Policy*, 77, 89–96.
- Campbell, A. M. and M. Savelsbergh. 2004. Efficient insertion heuristics for vehicle routing and scheduling problems. *Transportation science*, 38(3), 369–378.
- Campbell, A. M. and M. Savelsbergh. 2006. Incentive schemes for attended home delivery services. *Transportation Science*, 40(3), 327–341.
- Campbell, A. M. and M. W. Savelsbergh. 2005. Decision support for consumer direct grocery initiatives. *Transportation Science*, 39(3), 313–327.
- Ceryan, O., I. Duenyas, and O. Sahin. 2018. Dynamic pricing and replenishment with customer upgrades. *Production and Operations Management*, 27(4), 663–679.

- Chandrasekaran, A., K. Linderman, and F. J. Sting. 2018. Avoiding epistemological silos and empirical elephants in om: How to combine empirical and simulation methods? *Journal of Operations Management*, 63, 1–5.
- Chen, M. and Z.-L. Chen. 2015. Recent developments in dynamic pricing research: multiple products, competition, and limited demand information. *Production and Operations Management*, 24(5), 704–731.
- Chen, S.-F. S., K. B. Monroe, and Y.-C. Lou. 1998. The effects of framing price promotion messages on consumers' perceptions and purchase intentions. *Journal of Retailing*, 74(3), 353–372.
- Dayarian, I. and M. Savelsbergh. 2020. Crowdshipping and same-day delivery: Employing in-store customers to deliver online orders. *Production and Operations Management*, DOI: <https://doi.org.eur.idm.oclc.org/10.1111/poms.13219>.
- de Vries, H., J. van de Klundert, and A. P. Wagelmans. 2020. The roadside healthcare facility location problem a managerial network design challenge. *Production and Operations Management*, 29(5), 1165–1187.
- Deci, E. L. 1971. Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, 18(1), 105–115.
- Deci, E. L., R. Koestner, and R. M. Ryan. 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6), 627–668.
- Delmas, M. A., M. Fischlein, and O. I. Asensio. 2013. Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012. *Energy Policy*, 61, 729–739.
- DelVecchio, D., H. S. Krishnan, and D. C. Smith. 2007. Cents or percent? the effects of promotion framing on price expectations and choice. *Journal of Marketing*, 71(3), 158–170.
- D'Souza, C., M. Taghian, P. Lamb, and R. Peretiatko. 2007. Green decisions: Demographics and consumer understanding of environmental labels. *International Journal of Consumer Studies*, 31(4), 371–376.
- Duvvuri, S. D., A. Ansari, and S. Gupta. 2007. Consumers' price sensitivities across complementary categories. *Management Science*, 53(12), 1933–1945.
- Ehmke, J. F. and A. M. Campbell. 2014. Customer acceptance mechanisms for home deliveries in metropolitan areas. *European Journal of Operational Research*, 233(1), 193–207.
- Ellingsen, T. and M. Johannesson. 2007. Paying respect. *Journal of Economic Perspectives*, 21(4), 135–150.
- Emadikhiav, M., D. Bergman, and R. Day. 2020. Consistent routing and scheduling with simultaneous pickups and deliveries. *Production and Operations Management*, DOI: <https://doi.org.eur.idm.oclc.org/10.1111/poms.13200>.

- Gleim, M. R., J. S. Smith, D. Andrews, and J. J. Joseph Cronin. 2013. Against the green: A multi-method examination of the barriers to green consumption. *Journal of Retailing*, 89(1), 44–61.
- Gneezy, U., S. Meier, and P. Rey-Biel. 2011. When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives*, 25(4), 191–210.
- Gneezy, U. and A. Rustichini. 2000. A fine is a price. *The Journal of Legal Studies*, 29(1), 1–17.
- Golob, U. and L. Kronegger. 2019. Environmental consciousness of european consumers: A segmentation-based study. *Journal of Cleaner Production*, 221, 1–9.
- Gray, S. G., K. T. Raimi, R. Wilson, and J. Árvai. 2019. Will millennials save the world? the effect of age and generational differences on environmental concern. *Journal of environmental management*, 242, 394–402.
- Heyman, J. and D. Ariely. 2004. Effort for payment: A tale of two markets. *Psychological Science*, 15(11), 787–793.
- INRIX. 2018. 2018 global traffic scorecard. Technical report, INRIX, <http://inrix.com/scorecard/>.
- Jain, R. K., J. E. Taylor, and P. J. Culligan. 2013. Investigating the impact eco-feedback information representation has on building occupant energy consumption behavior and savings. *Energy and Buildings*, 64, 408–414.
- Kim, Y. H. and J.-E. Chung. 2011. Consumer purchase intention for organic personal care products. *Journal of Consumer Marketing*, 28(1), 40–47.
- Köhler, C., J. F. Ehmke, and A. M. Campbell. 2019. Flexible time window management for attended home deliveries. *Omega*, DOI: <https://doi.org/10.1016/j.omega.2019.01.001>.
- Laroche, M., J. Bergeron, and G. Barbaro-Forleo. 2001. Targeting consumers who are willing to pay more for environmentally friendly products. *Journal of Consumer Marketing*, 18(6), 503–520.
- Leonidou, C. N. and L. C. Leonidou. 2011. Research into environmental marketing/management: A bibliographic analysis. *European Journal of Marketing*, 45(1/2), 68–103.
- Lepper, M. R., D. Greene, and R. E. Nisbett. 1973. Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality and social Psychology*, 28(1), 129–137.
- Lin, I. and H. Mahmassani. 2002. Can online grocers deliver? Some logistics considerations. *Transportation Research Record: Journal of the Transportation Research Board*, 1817(1), 17–24.
- Liu, Q. and D. Zhang. 2013. Dynamic pricing competition with strategic customers under vertical product differentiation. *Management Science*, 59(1), 84–101.
- Lynham, J., K. Nitta, T. Saijo, and N. Tarui. 2016. Why does real-time information reduce energy consumption? *Energy Economics*, 54, 173–181.
- Menon, A. and A. Menon. 1997. Enviropreneurial marketing strategy: The emergence of corporate environmentalism as market strategy. *The Journal of Marketing*, 61(1), 51–67.

- Nijstad, B. A., C. K. De Dreu, E. F. Rietzschel, and M. Baas. 2010. The dual pathway to creativity model: Creative ideation as a function of flexibility and persistence. *European Review of Social Psychology*, 21(1), 34–77.
- Nomura, N. and M. Akai. 2004. Willingness to pay for green electricity in Japan as estimated through contingent valuation method. *Applied Energy*, 78(4), 453–463.
- Ottman, J. A. 1993. *Green Marketing: Challenges & Opportunities for the New Marketing Age*. NTC Publishing Group.
- Peattie, K. and M. Ratnayaka. 1992. Responding to the green movement. *Industrial Marketing Management*, 21(2), 103–110.
- Punakivi, M. and K. Tanskanen. 2002. Increasing the cost efficiency of e-fulfilment using shared reception boxes. *International Journal of Retail & Distribution Management*, 30(10), 498–507.
- Rana, J. and J. Paul. 2017. Consumer behavior and purchase intention for organic food: A review and research agenda. *Journal of Retailing and Consumer Services*, 38, 157–165.
- Rivera-Becerra, A. and L. Lin. 1999. Measuring environmental consciousness in product design and manufacturing. *Concurrent Engineering: Research and Applications*, 7(2), 123–138.
- Roe, B., M. F. Teisl, A. Levy, and M. Russell. 2001. US consumers' willingness to pay for green electricity. *Energy Policy*, 29(11), 917–925.
- Ryan, R. M. and E. L. Deci. 2000. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54–67.
- Ryan, R. M. and W. S. Grolnick. 1986. Origins and pawns in the classroom: Self-report and projective assessments of individual differences in children's perceptions. *Journal of Personality and Social Psychology*, 50(3), 550–558.
- Salzman, J. 1991. Green labels for consumers. *The OECD Observer*, 169, 28–30.
- Schwartz, D., W. Bruine de Bruin, B. Fischhoff, and L. Lave. 2015. Advertising energy saving programs: The potential environmental cost of emphasizing monetary savings. *Journal of Experimental Psychology: Applied*, 21(2), 158–166.
- Schwartz, S. H. 1977. Normative influences on altruism. *Advances in Experimental Social Psychology*, 10(1), 221–279.
- Shen, X. and Y. Yu. 2019. Capacity allocation with multiple suppliers and multiple demand classes. *Production and Operations Management*, 28(11), 2792–2807.
- Solomon, M. M. 1987. Algorithms for the vehicle routing and scheduling problems with time window constraints. *Operations Research*, 35(2), 254–265.
- Srinivasan, S., K. Pauwels, D. Hanssens, and M. Dekimpe. 2002. Who benefits from price promotions? *Harvard Business Review*, 80(9), 22.

- Steg, L., S. Lindenberg, K. Keizer, et al. 2016. Intrinsic motivation, norms and environmental behaviour: The dynamics of overarching goals. *International Review of Environmental and Resource Economics*, 9(1–2), 179–207.
- Steg, L. and C. Vlek. 2009. Encouraging pro-environmental behaviour: An integrative review and research agenda. *Journal of Environmental Psychology*, 29(3), 309–317.
- Strauss, A., N. Gülpınar, and Y. Zheng. 2020. Dynamic pricing of flexible time slots for attended home delivery. *European Journal of Operational Research*, DOI: <https://doi.org/10.1016/j.ejor.2020.03.007>.
- Su, X. 2007. Intertemporal pricing with strategic customer behavior. *Management Science*, 53(5), 726–741.
- Suki, N. M. 2016. Green product purchase intention: Impact of green brands, attitude, and knowledge. *British Food Journal*, 118(12), 2893–2910.
- Vazifehdoust, H., M. Taleghani, F. Esmaeilpour, and K. Nazari. 2013. Purchasing green to become greener: Factors influence consumers’ green purchasing behavior. *Management Science Letters*, 3(9), 2489–2500.
- Vine, D., L. Buys, and P. Morris. 2013. The effectiveness of energy feedback for conservation and peak demand: a literature review. *Open Journal of Energy Efficiency*, 2(1), 7–15.
- Weisstein, F. L., K. B. Monroe, and M. Kukar-Kinney. 2013. Effects of price framing on consumers’ perceptions of online dynamic pricing practices. *Journal of the Academy of Marketing Science*, 41(5), 501–514.
- WHO. 2016. Ambient air pollution: A global assessment of exposure and burden of disease. Technical report, World Health Organization, <https://www.who.int/phe/publications/air-pollution-global-assessment/en/>.
- Wilson, C. and H. Dowlatabadi. 2007. Models of decision making and residential energy use. *Annual Review of Environment and Resources*, 32, 169–203.
- Xia, L., K. B. Monroe, and J. L. Cox. 2004. The price is unfair! a conceptual framework of price fairness perceptions. *Journal of Marketing*, 68(4), 1–15.
- Yadav, R. and G. S. Pathak. 2017. Determinants of consumers’ green purchase behavior in a developing nation: Applying and extending the theory of planned behavior. *Ecological Economics*, 134, 114–122.
- Yang, X. and A. K. Strauss. 2017. An approximate dynamic programming approach to attended home delivery management. *European Journal of Operational Research*, 263(3), 935–945.
- Yang, X., A. K. Strauss, C. S. Currie, and R. Eglese. 2014. Choice-based demand management and vehicle routing in e-fulfillment. *Transportation Science*, 50(2), 473–488.
- Yazdanifard, R. and I. E. Mercy. 2011. The impact of green marketing on customer satisfaction and environmental safety. In *2011 International Conference on Computer Communication and Management*. volume 5, pages 637–641.

Zhang, D. J., H. Dai, L. Dong, F. Qi, N. Zhang, X. Liu, Z. Liu, and J. Yang. 2019. The long-term and spillover effects of price promotions on retailing platforms: Evidence from a large randomized experiment on alibaba. *Management Science*, 66(6), 2291–2799.