Moritz von Zahn | Kevin Bauer | Cristina Mihale-Wilson | Johanna Jagow | Max Speicher | Oliver Hinz

The Smart Green Nudge: Reducing Product Returns through Enriched Digital Footprints & Causal Machine Learning

SAFE Working Paper No. 363 | October 2022

Leibniz Institute for Financial Research SAFE

Sustainable Architecture for Finance in Europe

info@safe-frankfurt.de | www.safe-frankfurt.de

The Smart Green Nudge: Reducing Product Returns through Enriched Digital Footprints & Causal Machine Learning

Moritz von Zahn[*] Kevin Bauer[†] Cristina Mihale-Wilson[*]
Johanna Jagow[‡] Max Speicher[‡] Oliver Hinz[*]

October 28, 2022

Abstract

With free delivery of products virtually being a standard in E-commerce, product returns pose a major challenge for online retailers and society. For retailers, product returns involve significant transportation, labor, disposal, and administrative costs. From a societal perspective, product returns contribute to greenhouse gas emissions and packaging disposal and are often a waste of natural resources. Therefore, reducing product returns has become a key challenge. This paper develops and validates a novel *smart green nudging* approach to tackle the problem of product returns during customers' online shopping processes. We combine a green nudge with a novel data enrichment strategy and a modern causal machine learning method. We first run a large-scale randomized field experiment in the online shop of a German fashion retailer to test the efficacy of a novel green nudge. Subsequently, we fuse the data from about 50,000 customers with publicly-available aggregate data to create what we call *enriched digital footprints* and train a causal machine learning system capable of optimizing the administration of the green nudge. We report two main findings: First, our field study shows that the large-scale deployment of a simple, low-cost green nudge can significantly reduce product returns while increasing retailer profits. Second, we show how a causal machine learning system trained on the enriched digital footprint can amplify the effectiveness of the green nudge by "smartly" administering it only to certain types of customers. Overall, this paper demonstrates how combining a low-cost marketing instrument, a privacy-preserving data enrichment strategy, and a causal machine learning method can create a win-win situation from both an environmental and economic perspective by simultaneously reducing product returns and increasing retailers' profits.

[*]Goethe University Frankfurt, Frankfurt am Main, Germany
[†]Leibniz Institute for Financial Research SAFE, and University of Mannheim, 68131 Mannheim. KB and OH gratefully acknowledge research support from the Leibniz Institute for Financial Research SAFE.
[‡]Jagow Speicher Consulting, Düsseldorf, Germany

1 Introduction

Over the past few years, we witnessed a significant and continuous surge in global E-commerce. Despite the numerous benefits of E-commerce for customers and retailers, lax, and frequently even free, product return policies render the boom of E-commerce bittersweet. The National Retail Federation (NRF) reports that in the U.S. alone, average return rates of products bought online equaled 18.1% in 2020 (National Retail Federation 2020) and 20.8% in 2021 (National Retail Federation 2022). Conditional on the product category, return rates for online purchases can be up to 50% (Ofek et al. 2011).

These high return rates are problematic for both online retailers and society at large. For retailers, product returns induce costs that curtail their overall profits. On the one hand, for retailers, product returns entail high expenses for creating and maintaining product return infrastructures (Ofek et al. 2011) including transportation costs (Zhou and Hinz 2016), and costs for salvaging returned products to make them ready for resale (Ambilkar et al. 2021). Even worse, because retailers can frequently not directly resell returned products at full price and returns management is highly cost-intensive, they often destroy products rather than resell them (Shead 2021, Ofek et al. 2011). Hence, there exists tremendous potential for online retailers to decrease product returns as a means to enhance profits.

On the other hand, for society as a whole, product returns create considerable negative environmental externalities (Calma 2019). Product returns contribute to increased pollution through their carbon dioxide footprint and additional waste (Tian and Sarkis 2022). In the U.S. alone, transportation of returns generates over 15 million metric tons of additional carbon dioxide annually, and over five billion pounds of returned products and packaging end up in landfills (Calma 2019). From a sustainability point of view, with only about 50% of the returned products being resold and the other half being destroyed (Shead 2021), product returns represent a significant waste of natural resources. Against this background, reducing product returns is a pressing goal from a societal perspective.

Despite the financial and environmental damage caused by product returns, online retailers still lack an effective tool that reliably reduces the number of product returns while being economically viable—i. e., not detrimental to business performance. The implementation of more restrictive return policies or increased return costs for customers, for instance, often lead to fewer returns at the expense of net sales and profits (Altug et al. 2021, Zhou and Hinz 2016). Similarly, financially rewarding customers not to return products (Gelbrich et al. 2017), improving product descriptions or reviews (Zhou and Hinz 2016), using web technologies to display alternative product photos (De et al. 2013), and even the provision of virtual fitting rooms (Walsh and Möhring 2017) all show limited success in cost-effectively reducing product returns (Akturk et al. 2021).

Our paper further contributes to mitigating the problem of product returns by proposing and empirically validating a novel *smart green nudging* approach. We bring together insights and methods from the fields of Psychology, Behavioral Economics, and Artificial Intelligence to reduce the environmental damage associated with product returns while increasing profits for online retailers. Specifically, we construct what we call an *enriched digital footprint* and leverage state-of-the-art Causal Machine Learning (CML) (Athey et al. 2019) to personalize the administration of a *green nudge* (Thaler and Sunstein 2009, Carlsson et al. 2019)—a minimally invasive, low-cost informational stimulus that raises customer awareness of the adverse environmental impact of returning (unnecessarily) purchased products.

We developed and validated our smart green nudging approach in a two-stage strategy. First, we designed and implemented the green nudge in a large-scale field experiment that we conducted together with a large German fashion retailer. In seven weeks, half of the visitors of the retailer's German online shop randomly encountered our green nudge. During the field experiment, we precisely tracked customers' shopping, purchasing, and return behaviors. Second, we fused the data collected from the field study with additional aggregate privacy-preserving data that is publicly-available to create customers' *enriched digital footprint* and train a CML system. The trained system addresses pivotal customer-level treatment heterogeneities (Weinmann et al. 2016) and identifies individuals that can successfully be targeted by our green nudge leading to a reduction of product returns without adversely affecting profits. We then test whether the CML system, by smartly delivering the green nudge only to certain customers, could reliably reduce returns while increasing retailer's overall profit.

Two main findings demonstrate the considerable business and societal benefits associated with our proposed smart green nudging approach. First, our large-scale field experiment shows that "naively" administering the green nudge to every customer is, on an aggregate level, successful in reducing product returns. In fact, the naive green nudge turned out to be so effective from a business perspective that our industry partner decided to implement the nudge in the online shop permanently. A more nuanced analysis on the customer level revealed that the naive green nudge led to noticeable reductions in net profits for certain customer segments and, more surprisingly, increases in product returns, i.e., the green nudge backfires for certain customers. The backfiring undermines the efficacy of the intervention from a business and societal perspective.

Second, our analyses reveal a remarkable superiority of *smart green nudging* over *naive green nudging*. Administering the green nudge according to the prediction of the CML system can not only further increase the retailer's profits. It can also lead to an additional reduction in product returns suggesting that our proposed smart green nudge creates a win-win situation for the retailer and the environment. Importantly, we further show that the smart green nudge can only tap its full potential and outperform the naive green nudge when we train the CML system on our enriched digital footprint. Using only conventional digital footprint (Berg et al. 2020) or shopping data (Xia et al. 2019) to train the CML system does not lead to substantial improvements over the naive green nudge.

The remainder of the paper unfolds as follows. Section 2 outlines the theoretical foundation and related literature of our work. Section 3 presents our empirical strategy. In section 4, we present the results of our field study and simulation. Finally, section 5 concludes with a discussion of our works' theoretical and practical contributions, limitations, and avenues for future research.

2 Theoretical foundations and related literature

Our work builds on three streams of prior research. In the following, we outline all three of them and delineate how our study complements existing work.

Product returns. The ever-growing number of product returns, especially in E-Commerce, increasingly attracts the attention of marketing researchers (Abdulla et al. 2019)—not least because it creates adverse effects for the environment (Pålsson et al. 2017). Researchers in this domain have identified different motives behind return decisions: legitimate, opportunistic, and fraudulent product returns (e.g., Pei and Paswan 2018, Harris 2008). Fraudulent returns do not comply with the legal rules of the retailer. They include returning stolen products or consumer-damaged goods (Zhang et al. 2023). Legitimate and opportunistic returns both comply with retailers' return conditions and are different from each other in one significant facet. Customers initially ordered legitimately returned products with a genuine intention to keep them. By contrast, opportunistic returns involve ordered products with a pre-made intention to return, e.g., when customers engage in bracketing and order the same item in different colors and sizes (Altug et al. 2021), even though they know that they will keep only one of them. The National Retail Federation estimated that in 2015, 4%-6% of returns were opportunistic (Altug et al. 2021). Therefore, opportunistic returns represent a non-negligible returns category that is worth avoiding.Instruments and tools that aim to reduce product returns typically target fraudulent and opportunistic returns, however, as previous work demonstrates, there exists considerable heterogeneity in their effectiveness (see, e.g., El Kihal and Shehu 2022). In this work, we present an instrument able to reduce legitimate and opportunistic product returns.

A considerable fraction of work on how to reduce product returns revolves around return policies (Ambilkar et al. 2021, Janakiraman et al. 2016). The literature distinguishes between various return policies, ranging from "no-questions-asked" full money-back guarantees to policies imposing restocking fees and nonrefundable handling charges (Su 2009). At first glance, it might seem that abandoning lenient return policies can solve the problem of increasing product returns. However, prior research presents mixed results. Several studies show that lenient return policies have considerable benefits regarding sales and profits. Moorthy and Srinivasan (1995), for instance, showed that money-back guarantees are essential to signal product quality and are thus beneficial to increasing sales. Relatedly, Davis et al. (1995) suggest that money-back guarantees can increase the sales volume and, therefore, retailers' profits. Other research efforts corroborate the importance of lenient return policies as a signaling instrument for high quality and low risk (e.g., Altug et al. 2021). By contrast, there is also some research on the downsides of generous return policies (e.g., Su 2009, Shehu et al. 2020), providing evidence that lenient return policies directly promote excessive product returns. Against the background of this mixed evidence and the considerable economic risk of imposing overly restrictive return policies, it is not surprising that return policies have not become the most important lever to reduce product returns.

Another stream of work in this literature focuses on the customer side of the product return problem, i.e., on customers' return behavior patterns. By examining the decision-making process before and after the purchase, researchers seek to understand what factors can encourage (discourage) customers to keep (return) products bought online. One can broadly distinguish between three types of returns-preventing instruments: monetary, procedural, and customer-based (Walsh and Möhring 2017). Monetary instruments provide material incentives to keep customers from returning products, .e.g, discounts on future purchases. Procedural instruments entice customers to keep a product through design choices related to the delivery or return process (Walsh and Möhring 2017). Zhou et al. (2018), for instance, show that well-designed, colorful packaging and extra gifts can induce more positive emotions in consumers and, thereby, significantly reduce consumers' return intentions and actual returns. Conversely, retailers can also make the return process less convenient by limiting the potential return channels, or "print yourself" return tickets (Walsh and Möhring 2017). This approach rests on the notion that the more cumbersome the return process, the more likely customers are to keep the product rather than go through the hassle of returning it. Similar to the return policy, procedural prevention tools are also located in the post-order phase (Walsh and Möhring 2017).

By contrast, customer-based instruments work before and during the order process (Walsh and Möhring 2017). Early on Heiman et al. (2001) argued that customers are more likely to return products when there is a high gap between customer expectations and actual product characteristics. Hence, many existing customer-based instruments work based on the rationale that additional information on the characteristics of a product helps individuals form more realistic expectations of the product and thus reduce legitimate and opportunistic returns associated with expectancy-reality dissonances (Zhou and Hinz 2016). Frequently proposed customer-based instruments entail virtual try-on sessions, alternative product photos, customer reviews of products, and online consulting assistance or demonstrations (Walsh and Möhring 2017). Notably, on an individual level, product returns depend on idiosyncratic preferences and perceptions such as risk (Petersen and Kumar 2015).

According to the specifications of prior research, our proposed approach mainly aims to reduce opportunistic product returns by informing customers about the environmental ramifications of returning ultimately unwanted products during the shopping process, i.e., is a customer-based instrument. We seek to reduce returns by enhancing individuals' awareness about the environmental problems associated with high product returns. To the best of our knowledge, we are the first to employ a large-scale minimally-invasive green nudging intervention to reduce product returns. Thereby, we expand the repertoire of customer-based interventions that help alleviate organizational and environmental problems of product returns. Additionally, by making the green nudge smart using causal machine learning, we also showcase the considerable benefits that the combination of traditional psychology-inspired marketing interventions and state-of-the-art machine learning technology can deliver. In that sense, our work constitutes a complement to the nascent stream of work developing technology-based solutions to reduce product returns (Ambilkar et al. 2021, Kamble et al. 2019, Dzyabura et al. 2019).

Nudging theory and research on green nudging. Nudging posits that non-coercive interventions can help steer individuals to voluntarily make (socially) desirable choices they would not make otherwise (Thaler and Sunstein 2009). A nudge is a design element in the choice architecture intending to steer individuals' behavior in a certain predictable direction without limiting their choice set, significantly changing material incentives, or coercing them (Mirsch et al. 2017, Thaler and Sunstein 2009). It grounds on the widely-accepted notion that human decision-making depends on unconscious biases (Mirsch et al. 2017), the circumstances of our choice (Hinz et al. 2022), and the presentation of choices (Kasperbauer 2017). Common forms of nudges include setting a default option, making social norms salient, negligibly adjusting the ease of choosing certain options, creating a psychological anchor, attracting individuals' attention to certain options, and providing individuals certain information (Hummel and Maedche 2019, Weinmann et al. 2016). Nudging is applicable in many contexts and can contribute to achieving different objectives. For instance, previous research has demonstrated the effectiveness of nudging in improving individuals' eating and health-related choices, their financial decisions, or their philanthropic and environmental commitment (e.g., Goswami and Urminsky 2016, Mirsch et al. 2017, Thaler and Sunstein 2009).

Traditionally, nudging has been applied to offline contexts. However, with the ever-increasing number of decisions that individuals have to make online, digital nudging is rapidly gaining in relevance (Mirsch et al. 2017). Digital nudging refers to *"[...] the use of user-interface design elements to guide people's behavior in digital choice environments"* (Weinmann et al. 2016, p. 433). Compared to their physical counterparts, digital nudges are much cheaper, faster, and easier to implement (Mirsch et al. 2017). Although digital nudges primarily aim to alter individuals' choices in online environments, they can also affect outcomes in the offline world. For instance, reminders and statistics on a fitness tracker can significantly increase individuals' exercise and activity levels in the real world (Weinmann et al. 2016). Relatedly, we propose using a digital nudge in the online shopping context to reduce product returns offline.

More recently, researchers started to explore nudges as a means to steer individuals towards choices that minimize the adverse effects of their lifestyles and consumption on the environment (Carlsson et al. 2019)—a form of nudging referred to as "green nudging". More broadly, green nudging is part of research on Pro-Environmental Behavior (PEB) exploring interventions that aim to reduce the adverse environmental ramifications of human behaviors (White et al. 2019).

Research on green nudging primarily investigates areas frequently considered to have the most significant room for improvement (Lehner et al. 2016). Such areas of investigation are predominantly offline contexts, including water and energy conservation (e.g., Andor et al. 2020, Carroll et al. 2014, Chabe-Ferret et al. 2019, Momsen and Stoerk 2014), the use of public transportation (e.g., Kormos et al. 2015) and the purchase of greener products (e.g., Grinstein and Riefler 2015, Loschelder et al. 2019). The employment of green nudges to encourage PEB has proven to be highly effective in many domains, especially regarding energy consumption (e.g., Allcott and Rogers 2014). Schubert (2017) broadly distinguishes between three main forms of green nudges making use of different psychological desires: (i) green nudges based on individuals' desire to maintain a positive self-image through green behavior, (ii) green nudges based on individuals' inclination to adhere to social norms and convictions, and (iii) green nudges exploiting individuals' tendency to go with default options.

While (green) nudges can powerfully impact decision-making processes (Weinmann et al. 2016) and help address environmental challenges (Carlsson et al. 2019), they are no panacea (Richter et al. 2018). The intended effects of (green) nudging do not always materialize as expected (Lehner et al. 2016) because individuals' responses are highly heterogeneous in direction and size (Hummel and Maedche 2019). The way green nudging, and PEB interventions more generally, affects individuals' behaviors, depends on a variety of factors including social influences (i.e., social norms, social identities, social desirability, self-commitment, self-efficacy perceptions), persona (i.e., personality, self-efficacy perceptions, consumption preferences, environmental concerns, demographics, loss aversion), and habits (White et al. 2019). These heterogeneities cause nudges to create systematically different outcomes conditional on how and where they are in use. For instance, while Momsen and Stoerk (2014) find no significant effect of green nudging for renewable energy adoption, Ruokamo et al. (2022) report that green nudging may foster prudent energy consumption. The effectiveness of the green nudge hinges on various aspects such as information content and households' interest in energy issues (Ruokamo et al. 2022). Settings that leverage nudging to promote the purchase of greener products present similarly mixed results. Loschelder et al. (2019) and Demarque et al. (2015), for instance, find a positive effect of green nudges in the form of descriptive norms about sustainable consumption on actual consumers' product choices. By contrast, Richter et al. (2018) and Liu et al. (2016) report that green nudging can steer behaviors in unanticipated directions and backfire. The considerably heterogeneous effects of green nudging across settings demonstrate that insights from one context are not directly transferable to other settings or broader populations than the ones they stem from (Hummel and Maedche 2019, Lehner et al. 2016).

Our work builds on prior literature on (green) nudging. We design and implement a green nudge in an online environment to reduce product returns, i. e., eventually change offline behaviors. Our green nudge aims to capitalize on customers' desire to maintain a positive self-image through green behavior by increasing the salience of the adverse environmental externalities associated with returning purchased products.We complement the literature on (green) nudging and explore the effectiveness of green nudging in an online shop as a means to increase profits by reducing product returns. Despite the high environmental damage incurred by product returns in E-Commerce, there are, to the best of our knowledge, no studies doing so. Additionally, we embrace and actively exploit heterogeneities in individuals' responses to being nudged by utilizing state-of-the-art machine learning methods to improve the effectiveness of the intervention. This way we demonstrate the complementary nature of (green) nudging in online environments and machine learning. Combining these two instruments provides researchers and practitioners with an unprecedented capability to personalize minimally invasive green nudges at scale and optimize the promotion of environmentally benign behaviors.

Personalizing interventions. Prior literature suggests that nudging can substantially backfire due to customer heterogeneities. To overcome such heterogeneities and improve the effectiveness of the (nudge) interventions, one can exploit the digital environment and personalize the administration of the nudge instead of employing a one size fits all strategy.

The increasing availability of large amounts of data has recently spawned the development of machine learning systems that predict individuals' responses to certain interventions. For example, machine learning predict individuals' price sensitivity (Arevalillo 2021), forecast consumer behavior (Hu et al. 2019, Xia et al. 2019), and personalize product recommendations (Anitha and Kalaiarasu 2021, Ramzan et al. 2019). To do so, trained machine learning models typically leverage data to produce highly accurate, individual-level predictions about a certain variable of interest. Combined with subject matter expertise, these predictions can inform intervention decisions and address pivotal behavioral heterogeneities.

However, this way of using machine learning to personalize interventions has a critical shortcoming: it does not generally provide insights into the causal effects of specific interventions because these methods rely on associative inference (Jordan and Mitchell 2015). Understanding how an intervention affects certain variables of interest, e.g., how a customer's product return behavior changes in response to a green nudge, is pivotal to assessing the effectiveness of possible interventions and determine whether a customer should be treated. A growing literature at the intersection of econometrics and machine learning aims to address this shortcoming by developing so-called "causal machine learning" methods (e.g., Wager and Athey 2018, Athey et al. 2019, Schölkopf et al. 2021, Athey and Wager 2021). Under certain assumptions about the availability and structure of data, these methods try to estimate the change in the variable of interest if an individual was to be treated by an intervention. Put differently, they aim to predict counterfactuals.

The literature on using CML methods such as causal forests (Athey et al. 2019) is in its infancy. Only a handful of studies apply these methods to estimate the efficacy of treatment interventions on an individual level and personalize them to account for heterogeneity. Cagala et al. (2021) employ CML methods for the optimal distribution of gifts among potential donors in a fundraising campaign. Zhang and Luo (2022) rely on causal forests to understand the impact of social media postings on restaurant survival. Narang et al. (2019) use causal forests to predict heterogeneities in purchase behavior changes under mobile app failures. Deryugina et al. (2019) leverage CML methods to estimate the impact of air pollution exposure on life expectancy. In the context of a field experiment on electricity consumption, Murakami et al. (2020) apply causal forests to estimate heterogeneous treatment effects at the household level. Langen and Huber (2022) show the potential merits of CML in the context of coupon promotions in business analytics and marketing research. We extend this literature by employing causal forests to personalize a green nudge intervention and optimize its overall effectiveness to reduce product returns.

Crucially, just as other machine learning applications, CML methods' success hinges upon the data available during the training process (Athey and Wager 2021). In our E-Commerce context, one straightforward data source is customers' online shopping behavior, such as their product basket. This customer-level "basket data" can encode information tremendously valuable for training high-performing CML methods. While the basket data allow insights into the number and type of products in shopping carts (e.g., organic, sustainable, sports, and female fashion items), the "digital footprint" is another valuable source of data often readily available in online contexts. A basic digital footprint typically contains technical information about customers' internet providers, used browsers, running operating systems, and IP locations. Despite its technical nature, the digital footprint can represent information that considerably improves machine learning predictions. In their seminal paper Berg et al. (2020) demonstrated that variables from a basic digital footprint can provide insights into individuals' income, character, and creditworthiness.

In our paper, we propose going beyond leveraging customers' baskets and digital footprint information to train a (causal) machine learning system. Specifically, we connect the individual-level customer data readily available to retailers (basket data and the basic digital footprint) with publicly available, aggregate data, thereby creating a more granular data set we refer to as "enriched digital footprint". The enriched digital footprint serves as the basis to train our CML system. The selection of additional data to generate an enriched digital footprint is guided by insights from the literature on factors that determine the effectiveness of green nudges (e.g., Hummel and Maedche 2019). That is, we focus on aggregate-level public data that can encode information about customers' environmental preferences and, thus, their propensity to respond to the green nudge in the intended way while preserving their privacy concerning sensitive characteristics such as their gender or age.

With this strategy we add to the nascent literature on leveraging digital footprints and aggregate data to increase the accuracy of individual-level predictions (Berg et al. 2020, De Bruyn and Otter 2022). The main difference to previous approaches is that we do not use aggregate data to directly infer customers' characteristics like gender, level of education, or age and include this information in the training process of our CML system. Instead, our approach preserves customers' data privacy, at least to a considerably larger extent. Numerous data aggregating companies—often referred to as "data brokers"—such as Acxiom or eBureau offer retailers expensive and frequently inaccurate (Neumann et al. 2019), individual-level data to profile their customers based on aggregate-level data (Federal Trade Commission 2021). The inaccuracy of this information originates from limited amounts of data and lack of insights leading to segment assignment and consumer profiling problems (De Bruyn and Otter 2022). Our research illustrates that novel CML methods such as causal forests (Athey et al. 2019) may eventually render the intermediate estimation of individual-level data from aggregate-level sources obsolete and limit the need to purchase expensive data sets from data brokers. Additionally, by not inferring personal characteristics of customers, our data enrichment strategy arguably maintains a higher degree of customers' privacy.

3 Empirical strategy

In this section we outline the structure of our strategy to develop and validate our proposed smart green nudging approach. We proceed as follows. First, we introduce a simple theoretical framework that aims to illustrate the workings of our green nudge and the potential benefits of making it "smart". Second, we outline the design of our large-scale field study. Third, we explain our data enrichment procedure, the causal machine learning method, and our simulation approach.

3.1 Theoretical framework

This subsection develops a simple theoretical framework that has two objectives. On the one hand, it aims to formally illustrate the effects of our green nudge, i. e., increasing the salience of the adverse environmental externalities related to product returns. On the other hand, the framework intends to formally explain why our proposed smart green nudge—a green nudge smartly administered through a CML system—can improve the performance of our product return intervention. Notably, our simple framework is by no means a complete theoretical model that we derive hypotheses from. Instead, we want the model to help the reader better fathom the fundamental notions underlying our proposed smart green nudging.

Let the retailer's profit Π_i for an individual customer $i \in N$ be given by

$$\Pi_i = \overline{p}_i \cdot x_i(\omega) \cdot [1 - r_i(\omega)] - c_p \cdot x_i(\omega) - c_r \cdot r_i(\omega) \cdot x_i(\omega). \tag{1}$$

$\overline{p}_i > 0$ represents the average price of the products $x_i(.)$ that customer i purchased. $r_i(.) \in [0, 1]$ describes the share of purchased products that the customer eventually returned to the retailer. $c_p > 0$ represents the average marginal costs associated with producing, offering, and delivering the products the customer purchased whereas $c_r > 0$ depicts costs of handling a product return that include any reselling value of returned products. Without loss of generality let $\omega \in \mathbb{R}_+$ denote a customer's inclination for PEB. We assume that this inclination shapes customers' demand and return behaviors. Conditional on individual i's preferences, the type of product, and other situational factors, the relationship between demand x_i (product return $r_i(.)$) and ω can be positive or negative.

Following Schneider et al. (2018) and Thaler and Sunstein (2009), we model our green nudge as a cognitive stimulus $s > 1$ that (temporarily) shifts customers' inclination ω to $\hat{\omega} = s \cdot \omega$. That is, the green nudge makes customers more aware of the environmental ramifications of their behavior. According to our theoretical framework (1), this intervention will affect profits through changing customers' demand $x_i(.)$ and the fraction of returned products $r_i(.)$. In our setting, the heterogeneity that previous literature documents regarding the success of (green) nudging originates from individual differences in the direction and relative size of demand and product return changes. Formally, it follows that the retailer's profit for customer i increases $\frac{\partial \Pi_i}{\partial \omega} \geq 0$ if and only if

$$\left| \frac{1}{x_i} \left[\frac{\overline{p}_i - c_p}{\overline{p}_i + c_r} - r_i \right] \right| \cdot \left| \frac{\partial x_i}{\partial \omega} \right| \leq \left| \frac{\partial r_i}{\partial \omega} \right|. \tag{2}$$

Following condition (2), the intervention is profitable when the (absolute) PEB elasticity in returns $\frac{\partial r_i(.)}{\partial \omega}$ is larger than the (absolute) PEB elasticity in demand $\frac{\partial x_i(.)}{\partial \omega}$ scaled by the factor $\left| \frac{1}{x_i} \left[\frac{\overline{p}_i - c_p}{\overline{p}_i + c_r} - r_i \right] \right| < 1$. According to the scaling factor, the likelihood that the intervention is profitable increases with a customer's initial demand x_i, initial return rate r_i, and the costs associated with the production and return of products c_p, c_r.

To maximize overall profits $\sum_i \Pi_i$ for a given price and cost structure, the retailer needs to account for heterogeneous treatment effects and administer the intervention only to the subset of customers $G \subset N$ (i) whose absolute PEB product return elasticity is relatively larger than their absolute PEB demand elasticity, (ii) who make a sufficiently large purchase, and (iii) who originally have a high product return rate (e.g., because they engage in bracketing). By administrating the intervention only to this type of customers, the retailer can ensure that the environmentally desirable intervention increases profits and is, therefore, viable from a business perspective.

Naturally, the question is how to identify the subset of customers G? This is where our proposed enriched digital footprint and the CML system come in. We first combine customer level with publicly-available aggregate data. Using this "enriched" data set, we then train a CML system capable of predicting whether $\frac{\partial \Pi_i}{\partial \omega} \geq 0$ holds true for individual customers allowing us to smartly administer the green nudge only to this subset of customers.

3.2 Large-scale field study

The first step in the development of our smart green nudging approach was the design and implementation of a large-scale, randomized field experiment. The purpose of the field experiment is twofold. First, we needed to test the efficacy of a green nudge to reduce product returns and collect data to train the CML system. Second, the successful training of our CML system (a causal forest (Athey et al. 2019)) effectively requires us to possess data where the treatment is randomly assigned (Wager and Athey 2018). To conduct a randomized field study, we collaborated with an industry partner—a large European fashion retailer.

Together with our industry partner, we designed a green nudge that we implemented in the German online shop by means of a randomized A/B test. To design the nudge, we closely collaborated with our partner's UX department allowing us to combine the highly complementary practical and academic expertise. During the process of developing the green nudge, we had to balance theory-based notions of how to design nudges and test isolated mechanisms with practical feasibility and business restrictions. For instance, while the researcher part of our team in a first best case would have tested a variety of different green nudges to better understand distinct mechanisms, this approach was practically infeasible. Hence, our team of researchers and practitioners eventually had to agree on one design that we expected to perform best when first put live in the online shop. While being a protracted process, the open discussions around the selection of a specific green nudge helped both the researchers and practitioners of our team to learn how to communicate with each other and carefully navigate research and business interests. Figure 1 depicts the structure of the green nudge we eventually implemented in our field experiment.

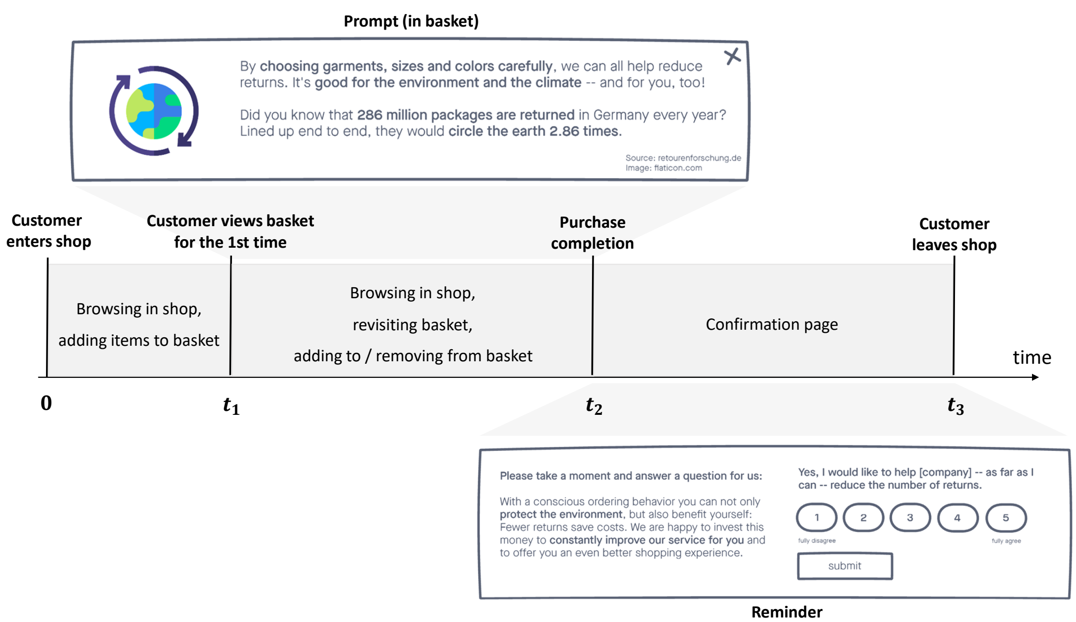

Figure 1: Two-component green nudge to reduce returns. The prompt appears when consumers check their baskets. The reminder is shown after payment.

Our implemented green nudge comprised two components that we automatically administered throughout customers' online shopping process: (i) an informational prompt that occurred whenever customers checked their shopping basket, and (ii) a reminder that popped up after customers completed their purchase and obtained the final confirmation that products are being sent. The prompt simply informed customers about the existence of the environmental issue of product returns and that considerate shopping decisions can help alleviate these negative side effects. The reminder again emphasized that product returns are a problem for the environment and asked customers to indicate their personal commitment to reduce product returns on a 5-point scale. We included the self-commitment component upon the suggestion of the practitioners in our team because they had received positive customer feedback about self-commitment instruments used during other successful A/B tests. The addition of this design component is one example of the immensely valuable insights practitioners in the team shared with researchers to adjust the green nudge to the idiosyncrasies of their customer base, emphasizing the importance of the engagement of and with practitioners in research-industry collaborations. Our team came up with a prompt and a reminder for two reasons: (i) because research has shown to be effective due to their low-cost and easy-to-implement nature (e.g., Goldstein et al. 2007, Mirsch et al. 2017, White et al. 2019), and (ii) because our industry partner had used them in a different context before, i.e., was already familiar with both instruments. Notably, we combined the prompt and the reminder to enhance the effectiveness of the green nudge and thus increase the likelihood that our industry partner would already see positive results regarding product returns during the A/B test.

We tested the designed green nudge in a randomized field experiment that ran for seven weeks (for reasons of confidentiality, we cannot give more precise information on run-time) in the German online shop of our industry partner.[1] Website visitors from stationary devices were automatically and randomly assigned to treatment (green nudge) and control groups (no nudge) when they first checked their shopping basket. In total, a seven-figure number of visitors have unknowingly participated in the field study. Several randomization checks (see Figure 7 in Appendix) show that the randomization process was successful along elicited characteristics. Across the treatment and control groups, 49,600 visitors made a purchase. The conversion rate is typical for the fashion retailer industry. When analyzing the efficacy of our green nudge to reduce product returns, we naturally focus on website visitors who made a purchase.

For each of these customers, our industry partner logged purchase information such as the

[1]Please note that we are unable to report certain details to maintain confidentiality and preserve the anonymity of our industry partner.

overall demand value, the returned products, the number of duplicate products in the basket, and product categories bought (e.g., female fashion, kids fashion, organic products). We also obtained customers' digital footprint, including their browser type, web reference, and their IP address. A considerable obstacle our team needed to overcome regarding the collected data was (i) the logging structure and (ii) the dispersion of online shopping, purchasing, and product return information throughout the organization of our industry partner. Gathering the required data and transforming it for analyses and the development of the CML system required extensive guidance and effort by our team's practitioners who had to engage with different business units in their company.

3.3 Data enrichment, causal forests, and simulation design

The enriched digital footprint. As illustrated by our theoretical framework, the smart green nudging approach seeks to administer the green nudge selectively to customers for whom it increases profits. To identify these customers, we train a CML system to predict conditional average treatment effects for each customer.

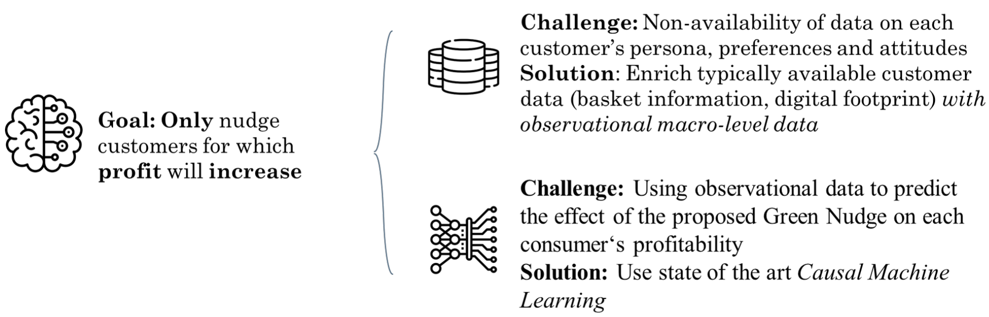

Figure 2: Challenges to the decision whom to nudge

The ability of the CML system to make accurate predictions depends critically on the richness and granularity of the data available during the training process (Guo et al. 2020). To understand what data we can leverage to train the system and eventually make predictions, we need to consider the shopping process and the point in time at which the administration of green nudging needs to be determined. As Figure 1 illustrates, the initial touch point with the green nudge is customer's first visit to her shopping basket (t_1). At this stage of the shopping process, our industry partner possesses (i) current data on customers' shopping basket composition, and (ii) customers' digital footprints. That is, without any additional efforts, a CML system can only leverage this customer-level information to predict whether it is profitable to administer the green nudge. While the digital footprint is typically equally informative across customers, the quality of the basket data depends on customers' shopping habits and can vary considerably. Some customers may like to browse and add many products to their baskets before checking it for the first time whereas others may like to add one or two items and then immediately check their shopping baskets. Therefore, the quality of the available basket data may not suffice for the CML to learn underlying behavioral regularities and eventually produce sufficiently accurate predictions.[2]

To overcome these obstacles, we enrich the customer data available when deciding whether to administer the nudge with publicly available aggregated data (see Figure 2 for an illustration of the challenges and solutions). Our industry partner can store these data in the backend of the online store and use it to make predictions as needed. We identified relevant aggregate data building on our theoretical framework and prior literature suggesting that an individual's likelihood of engaging in more PEB relates to their socio-economic environment (Green and Peloza 2014, White et al. 2019). We select publicly available data sources on socioeconomic status, environmental attitudes, and ecological behaviors at the regional, and district levels. Notably, we only use aggregate data and refrain from leveraging data that is potentially privacy-invasive, such as the home address or the browser history. Specifically, we used data from the German Socio-Economic Panel (SOEP),[3] and publicly available government-provided data on the income of private households per capita and share of settlement areas as population density indicator within a region of interest[4] to enrich the available customer data. We refer to the data resulting from fusing customer-level data (basket data and the digital footprint) with aggregate third-party data as the *enriched digital footprint*.

Table 1 provides an overview of the resulting data structure and the level of aggregation: the basket and digital footprint data from the 49,600 individual shopping sessions stem from 2,636 unique geolocations in Germany; the data from SOEP is regional (96 German regions); the other public statistics are on the district level (401 German districts), whereby districts uniquely map to regions.

To join the different levels of aggregation, we leverage the IP address included in the digital footprint. Specifically, we use the 2,636 unique geolocations of shopping sessions to compute a Voronoi diagram (cf. Burrough et al. 2015)—a common spatial resolution for

[2]In line with this argumentation, our data ablation study reported below reveals that merely relying on basket (and digital footprint) data does not suffice for the smart green nudge to outperform the naive green nudge substantially.

[3]The SOEP comprises responses from more than 30,000 representative individuals in Germany in 2019. Among many other dimensions, SOEP captures demographics, preferences, and attitudes of German citizens. While the SOEP data is generally a publicly available resource, it can only be used for scientific purposes. We refrain from using SOEP data in developing an operational system and merely leverage it as a proof of concept that we can showcase.

[4]See `https://www.statistikportal.de/de/vgrdl/ergebnisse-kreisebene/einkommen-kreise` and `https://www.landatlas.de/daten.html`

Source	Aggregation	Data	Type
Initial basket	Customer	Basket value	Numerical
		Number of products	Numerical
		Number of eco-friendly products	Numerical
		Contains duplicate products	Binary
		Category "Women"	Binary
		Category "Children"	Binary
		Category "Sports"	Binary
		Category "Hipster"	Binary
		Category "Organic"	Binary
Digital footprint	Customer	Internet provider	Categorical
		Internet browser	Categorical
		Online during working hours	Binary
Public Statistics	Aggregate (District)	Income private households	Numerical
		Share of settlement area	Numerical
SOEP	Aggregate (Region)	Worry about environment	Numerical
		Members environmental groups	Numerical

Table 1: Data for Smart Green Nudging.

connecting different granular data sets by partitioning Germany into cells (see Figure 3).[5] Creating the enriched digital footprint unlocks substantial additional business opportuni-

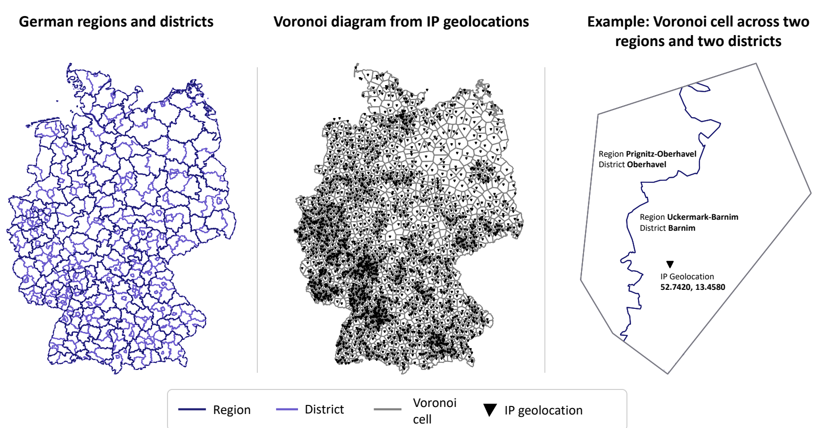

Figure 3: Visualization of different data aggregation levels.

Notes: From left to right the maps reveal (i) the partition of Germany into 96 regions and 401 districts, (ii) the 2,636 unique geolocations and their corresponding Voronoi cells, and (iii) examples of geolocations and the Voronoi cell intersecting two districts and regions. In map (iii) all shopping sessions originating from IP addresses associated with the given location will be assigned the weighted average of the third-party variables from the two districts/regions.

ties. Where individual-level data is scarce or unreliable (e.g., De Bruyn and Otter 2022), the enriched digital footprint offers additional insights that can help tailoring marketing interventions. Our approach showcases that businesses can enrich their customer data without relying on privacy-invasive data or on costly third-party data from brokers. Instead, they may draw from publicly available sources only, which makes the enriched digital footprint a

[5]Each cell uniquely maps to one of the 2,636 IP locations. A cell consists of all points of Germany closer to that IP location than to any other IP location. We assume the cell around an IP location to be an appropriate approximation of the customer's location. Before we join the aggregate data with the individual shopping sessions, we also compute the aggregate variables as weighted averages for each IP location. Specifically, each district or region overlapping with a Voronoi cell contributes to the weighted average, whereby the share of overlap in the total area of the cell constitutes the weight.

highly accessible tool for personalizing (marketing) interventions.

The causal forest. To make our green nudge smart, we apply causal forest (CF) estimation techniques to the data from our randomized field experiment that we enriched as outlined above (Athey et al. 2019, Wager and Athey 2018). The CF aims to predict the conditional average treatment effect $E(Y_{i,1} - Y_{i,0}|X = x_i)$ (CATE), i.e., the expected change in the outcome variable $Y_{i,j}$ for an individual i with certain characteristics $X = x_i$ when being exposed to some treatment intervention $j = 1$ compared to not being exposed $j = 0$. In our context, the CF estimates how the expected retailer profit obtained from customer i would change if she observes the green nudge. The problem CFs solve is that standard fit measures like mean squared error are naturally not applicable in the training process of supervised machine learning models, because it is impossible to observe $Y_{i,1}$ and $Y_{i,0}$ simultaneously for the same individual—after all, one customer can either observe the nudge or not. Athey and Imbens (2016) first showed that maximizing the variance of treatment effects across leafs in decision trees (minus some penalty term) is mathematically equivalent to minimizing the expected mean squared error of predicted treatment effects. As a result, a tree trained using the variance maximization criterion can estimate the CATE for individual i as the treatment-control difference between average outcomes $\overline{Y}_{1,L} - \overline{Y}_{0,L}$ in the tree's leaf L where individual i is contained. A CF generalizes this idea to an ensemble of multiple decision trees forming a forest. Notably, the training of CFs is typically "honest" in the sense that disjoint subsamples of the training data are used to build the CF and estimate the CATE.

In our scenario, we train a CF to estimate customers' CATE with respect to profit, and based on that, decide whether it is the profit-maximizing choice to apply green nudging to a particular customer. We wish to treat only customers whose estimated CATE is non-negative, i.e., green nudging does not backfire and for whom the previously outlined condition (2) holds. We implement the CF in Python using the popular EconML and scikit-learn libraries.[6] We follow common data science conventions and randomly split the data into different sets of training (80%) and testing (20%) partitions (Hastie et al. 2009). We apply one-hot encoding to categorical variables and standardize numerical variables based on the training set. We then use the training set to construct, tune, and estimate the CF that predicts the effect of administering the green nudge on the marginal profit of an individual shopping session. The hyperparameters for the causal forest are determined by applying a grid search in a 5-fold cross-validation on the training set based on out-of-sample R-score performance (Schuler et al. 2018). We report tuning ranges in Table 6 in the Appendix. Our implementation of the CF is combined with double machine learning-based residualization (Oprescu et al. 2019) to further boost its predictive performance.

[6]See `https://github.com/microsoft/EconML` and `https://scikit-learn.org/stable/`

Assessment design. Due to time and organizational constraints on the part of our industry partner, we needed to evaluate our smart green nudging approach using the data we already possess. In our analyses, we use the training data set to build an optimized CF. The optimized CF predicts the CATE for each observation in the test data set. We administer the green nudge to all sessions for which the predicted CATE is non-negative. As a result, for each customer, we now know (i) whether the CF would have administered the treatment (i.e., "CF, no nudge" or "CF, green nudge"), and (ii) whether this person was actually treated in the field experiment (i. e., "no nudge" or "green nudge").

To measure the effect of smart green nudging, we conservatively consider the subset of the test data for which the hypothetical CF-administered treatment and the actual treatment coincide. For example, for a given customer j, the CF may predict that the nudge increases (decreases) profits and administers the treatment "CF, green nudge" ("CF, no nudge"). If the actual treatment of customer j was "green nudge" ("no nudge"), then the CF and actual treatment coincide. We compute the average product returns, and profits across all customers for which the hypothetical CF treatment and the actual field treatment coincide. To assess the effectiveness of the smart green nudge we compare the resulting smart green nudge measures with the corresponding measures averaged across all observations in the current test set, i. e., the outcomes for the naive green nudge.

4 Results

In this section we present the impact of our green nudge intervention and smart green nudging approach. We first depict the benefits of our green nudge in the large-scale field study. Subsequently, we demonstrate the additional value that our smart green nudging approach can deliver. Note that we linearly transformed the main variables of interest (i.e., product returns and profits) due to their confidential nature.

4.1 Green nudging

Overall treatment effects. Table 2 shows summary statistics for the main variables of interest in our field experiment. We depict results for our treatment and control conditions separately. The results reveal that our green nudge had a significant impact on our industry partner's product returns and profits. A comparison between the value of returned products in our two study conditions shows that our green nudge entailed a statistically significant decrease in product returns by 3.8% (Kruskal-Wallis test: $p < 0.05$). Considering the green nudge's impact on average profits, we find a statistically and economically significant increase by 8.7% (Kruskal-Wallis test: $p < 0.05$). In relation to our theoretical framework, these results depict that $\sum_i \Pi_i(\hat{\omega}) > \sum_i \Pi_i(\omega)$ with, most importantly from an environmental perspective, $\sum_i r_i(\hat{\omega}) < \sum_i r_i(\omega)$.

Group	Obs.	**Returns** Mean (Std. Dev)	**Profits** Mean (Std. Dev)
Control	24,319	23.59 (48.86)	8.15 (32.84)
Treatment	25,281	22.69 (48.24)	8.86 (33.59)
Difference in %		-3.8**	+8.7**

*Kruskal-Wallis test significance levels: * p <0.10, ** p <0.05, *** p <0.01*

Table 2: Overall effect of green nudging

These first results demonstrate the remarkable impact that our designed green nudge had. Importantly, the low-cost, minimally invasive intervention comprising two simple pieces of information simultaneously reduces product returns and increases profits. Therefore, it is beneficial for both the society at large (i. e., by reducing environmental externalities associated with product returns) and the retailer (i. e., by increasing net profits). Considering the example of annual carbon dioxide emissions attributable to the logistics of returns in the US alone (e. g., Calma 2019), the intervention could reduce the carbon footprint by 570,000 metric tons of carbon dioxide emissions annually. This amount is roughly equivalent to the annual emissions from the electricity consumption of 110,000 homes in the US[7].

From a business perspective, according to internal figures of our industry partner, the estimated annual profit increase quickly amortizes costs associated with the development, implementation, and maintenance of the green nudge. In fact, the developed green nudge turned out to be so successful that our industry partner opted to keep the green nudge running in the German online shop. Apart from the immediate business benefits in terms of decreased costs and increased profits, our industry partner received positive customer feedback in their regular surveys. Additionally, the industry partner observed their net promoter score increase in the immediate aftermath of the green nudge intervention. We provide customer quotes and further details in the Appendix (see Table 7). While these insights merely represent anecdotal evidence, these results may constitute an indication that the introduction of the green nudge has additional positive effects on company image and long-term growth. One may, for instance, speculate that our simple intervention leads to pronounced customer loyalty or even a change in the customer base due to self-selection.

Treatment heterogeneities. Despite the green nudge's overall effectiveness in simultaneously reducing product returns and increasing profits (entailing positive average treatment

[7]https://www.epa.gov/energy/greenhouse-gas-equivalencies-calculator

effects) more nuanced analyses on treatment heterogeneities reveal that the nudge appears to backfire for certain customer segments (i. e., it sometimes creates negative conditional average treatment effects), i.e., in line with our theoretical framework there exist a subset G with $\Pi_j(\hat{\omega}) < \Pi_j(\omega)$ with $r_j(\hat{\omega}) > r_j(\omega)$ for $j \in G \subset N$. In this subsection, we will depict the existence of considerable treatment heterogeneities by means of an example. Notably, while we only provide one example for brevity and readability, our analyses revealed similar results for different customer segmentations, too (see Table 5 in Appendix).[8]

	Obs	**Returns** Mean (Std. Dev)	Effect	**Profits** Mean (Std. Dev)	Effect
Dortmund					
Control	382	25.50 (53.61)	⇓ **	5.65 (34.04)	⇑ **
Treatment	422	15.26 (31.83)		12.17 (30.71)	
Hamburg					
Control	10.72	22.80 (46.41)	⇓ **	7.70 (30.29)	⇑
Treatment	1,092	17.66 (40.51)		10.16 (27.16)	
Middle Franconia					
Control	475	20.74 (42.54)	⇑ *	9.82 (33.59)	⇓
Treatment	521	30.06 (66.61)		5.38 (38.53)	

*Kruskal-Wallis test significance levels: * p <0.10, ** p <0.05, and *** p <0.01*

Table 3: Effect of green nudging in various locations

Table 3 depicts treatment heterogeneities for three different customer segments based on the geolocation derived from their IP address. These results prove that the green nudge can have a considerably heterogeneous impact for customers from different regions. While the green nudge leads to a significant decrease in product returns in Dortmund (Kruskal-Wallis test: $p < 0.05$) and Hamburg (Kruskal-Wallis test: $p < 0.05$), we observe that it significantly increased the product returns in Middle Franconia (Kruskal-Wallis test: $p < 0.1$), i. e., the intervention evoked opposite responses for customers and backfired regarding product returns. This statistically and economically (+45.4%) significant impact on product returns by itself constitutes a non-negligible negative impact from the societal perspective. Notably, even though the impact on profits for the Middle Franconia segment is statistically insignificant ($p = 0.18$), the economic significance is remarkable (-44.9%).

These findings are in line with prior literature on potential dangers of using (green) nudges as an intervention (e.g., Liu et al. 2016, Weinmann et al. 2016). Previous research suggests that these backfire effects may represent a reactance effect (Reich and Robertson 1979) creating negative emotional appraisals (Zemack-Rugar et al. 2017, Gonzalez-Arcos et al.

[8]Note: given the variability in latent customer attributes, our analyses arguably only scratched the tip of the figurative treatment heterogeneity iceberg.

2021). The observed backfire effects for certain customer segments render the green nudge (if naively administered to all customers) as a helpful but not optimal solution from both the societal and business perspective. Inspired by our theoretical framework and previous literature, we next show how to address this heterogeneity by utilizing a causal forest (Athey et al. 2019) that we trained on customers' enriched digital footprint to identify the segment of customers G for which $\Pi_{j\in G}(\hat{\omega}) < 0$.

4.2 Smart green nudging

Impact of smart green nudging. In this subsection, we outline the additional impact of making the green nudge "smart" by means of a representative train-test split. All reported results are robust to variations in this split. For simplicity, throughout our analyses, we refer to administering the green nudge to all customers, i.e., not accounting for outlined treatment heterogeneities, as "naive green nudging". Figure 4 contrasts the overall impact of naive and smart green nudging relative to the control condition in our field experiment. We depict results for the overall number of returned products and profits. We report associated summary statistics in Table 4.

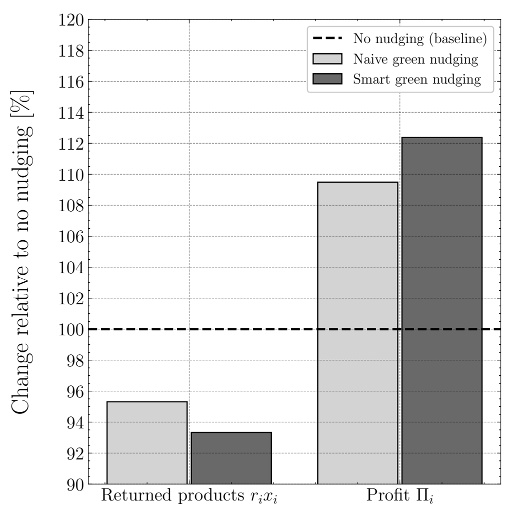

Figure 4: Impact of smart green nudging.

Notes: We depict the overall changes in returned products (left) and profit (right) of naive and smart green nudging relative to no nudging (control).

Our results paint a clear picture: making the green nudge smart using causal machine learning methods has a considerable impact on both product returns and profits. More specifically, our smart green nudging approach compared to the naive green nudging leads to an additional decrease in product returns by 2.1%. Compared to not nudging customers at all, the smart approach entails an overall decrease of product returns by 6.7%. Examining the impact of smart green nudging on our industry partner's profits we find that making the intervention smart additionally increases profits by 2.6%. Compared to having no nudge in place, the profit increase associated with the smart green nudge equals 12.5%. Notably, the optimized causal forest administers the green nudge to approximately 95% of the customers. The decrease in product returns and increase in profits associated with treating fewer customers provide strong evidence that our optimized causal forest potently addresses treatment heterogeneities at the customer level that mitigate the naive green nudging overall effectiveness. Following the example of annual carbon dioxide emissions attributable to the logistics of returns in the US alone (e. g., Calma 2019), smart green nudging could further reduce the carbon footprint by an additional 315,000 metric tons of carbon dioxide annually.

Group	Obs	**Returns** Mean (Std. Dev)	**Profit** Mean (Std. Dev)
(i) No nudging (baseline)	4,921	23.33 (48.76)	8.65 (33.36)
(ii) Naive green nudging	4,999	22.23 (46.24)	9.48 (36.43)
(iii) Smart green nudging	4,956	21.77 (45.57)	9.73 (36.32)
Difference (ii) - (i) in %		−4.71	+9.60
Difference (iii) - (i) in %		−6.69	+12.49
Difference (iii) - (ii) in %		−2.06	+2.64

Table 4: Summary statistics from the simulation of smart green nudging.

In sum, our analyses depict that there exists considerable potential to switch from naively administering the green nudge to all customers to smartly nudging only those that our causal forest identifies as "worthy" to treat. In terms of our theoretical framework, our smart green nudging approach leads to $\sum_{i \in N \setminus G} \Pi_i(\hat{\omega}) + \sum_{j \in G} \Pi_j(\omega) > \sum_{i \in N} \Pi_i(\hat{\omega}) > \sum_{i \in N} \Pi_i(\omega)$ with $\sum_{i \in N \setminus G} r_i(\hat{\omega}) + \sum_{j \in G} r_j(\omega) < \sum_{i \in N} r_i(\hat{\omega}) < \sum_{i \in N} r_i(\omega)$.

Inside the black box. At this point, the reader may naturally wonder about two closely connected things: (i) what is the value of using the enriched digital footprint, and (ii) why does the smart green nudge work? To answer these questions, in the final step of our analyses, we provide insights into the inner workings of the causal machine learning system we developed using methods from the field of eXplainable Artificial Intelligence (e.g., Bauer et al. 2022).

In response to the first question, Figure 5 presents results of a data ablation exercise where we repeated our analyses under the restrictions that the causal forest only learned based on customer-level data (i. e., basket data alone, and basket data together with the basic digital footprint). We show the results for product returns and profits relative to the naive green nudging.

The data ablation study reveals the considerable business and societal value associated

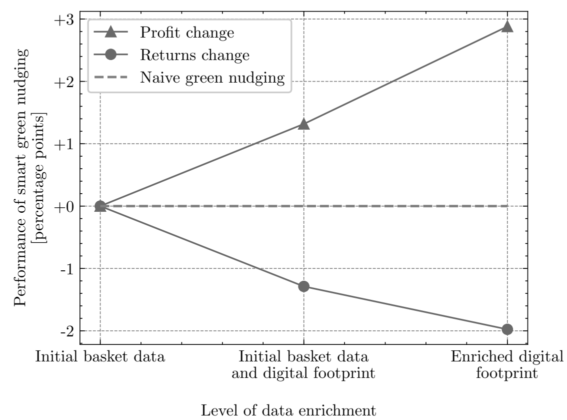

Figure 5: The value of the enriched digital footprint.

Notes: We depict the impact on the change in returns (dots) and change in profit (triangles) when moving from naive to smart green nudging for different datasets.

with the construction of our enriched digital footprint. We find that training the causal forest on basket data alone does not lead to an improvement of the smart green nudge over the naive one. In line with prior literature on the potential value of digital footprints (e.g., Berg et al. 2020), we find that the combination of basket and digital footprint data achieves a better result than the basket data alone. Using this combined data to train the causal forest causes our smart green nudging approach to outperform the naive green nudging (product returns: −1.29%; profits: +1.32%). However, fusing aggregate data to the customer-level data leads to an even better outcome (product returns: −1.98%; profits: +2.88%), illustrating the remarkable additional value that the construction of an enriched digital footprint can have.

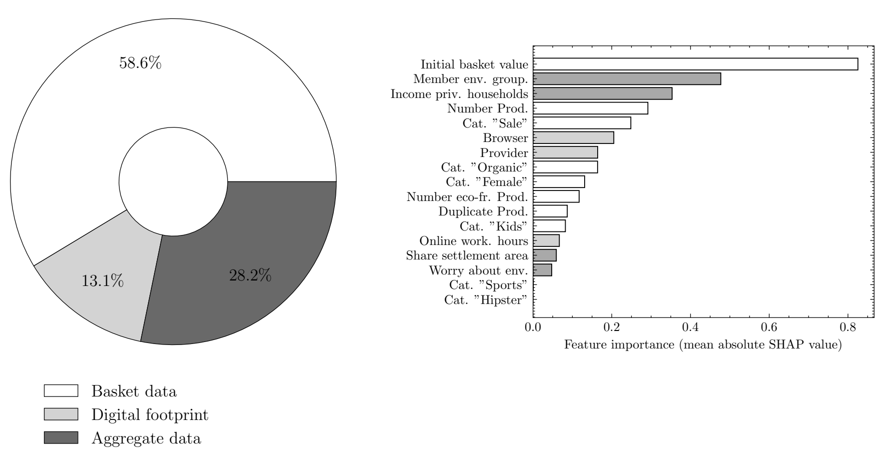

Figure 6: Feature importance aggregated by the type of data (on the left) and on an individual level (on the right). Feature importance is computed based on mean absolute SHAP values (Lundberg and Lee 2017).

To answer the second question, Figure 6 depicts the aggregated feature importance (left) during the prediction process of the optimized causal forest, i.e., how much weight the features of different types of data have for a typical prediction (basket data in white, digital footprint in light gray, aggregate data in dark gray). On a more granular level (right), we show the importances of individual features. The Figure provides two main insights. First, we gain a better understanding of the predictors of the green nudge's success in increasing profits (and reducing returns). Somewhat intuitively, and in line with our theoretical framework under (2), we find that the size of a customer's basket when first checking it (i. e., initial basket value) is the strongest single predictor of the nudge's efficacy. The feature importance plot also reveals that the shopping basket composition at this point in time is relevant, e.g., how many duplicate products, or green products it contains. We discussed these insights with experts from our industry partner to ensure the validity of the causal forest's internal working and alleviate concerns that it ultimately picked up non-generalizable feature-label relations in the data. The feature importances are largely in line with the consulted experts' expectations. Interestingly, some of the results, while ultimately plausible, came as a surprise to experts and inspired them to reconsider some of their persona definitions (e.g., the feature "Online during work hours"). This anecdotal evidence illustrates that the employment of causal machine learning methods to address treatment heterogeneities (in A/B tests) may even constitute a source of inspiration to understand customer segments better (or define new ones) and improve the design of future interventions. Second, further corroborating the value of adding aggregate data and creating the enriched digital footprint, we observe that the aggregate data features have relatively high importance in the prediction process. In particular, the share of environmental activists in a region and the average regional household income are respectively the second and third most important feature to predict the change in the retailer's profit from a specific customer.

5 Discussion and conclusion

In this paper, we propose and validate a novel interdisciplinary *smart green nudging* approach for effectively reducing product returns in E-commerce. Building on prior literature from Psychology and Behavioral Economics we first design a minimally invasive, low-cost green nudge that we test in a seven-week large-scale randomized field experiment in the online shop of a leading German fashion retailer. Overall, administering the green nudge to all customers can significantly reduce product returns and increase retailer profits. By doing so, our simple green nudge has a positive environmental and economic impact. Due to customer heterogeneity, however, the green nudging can backfire and lead to non-negligible increases in product returns and decreases in profits for certain customer segments that adversely affect the intervention's impact. To account for customers' heterogeneous responses, we construct what we call an *enriched digital footprint* and utilize it to train a CML system that "smartly" determines the administration of the green nudge on the individual level to optimize its efficacy. Further analyses show the superiority of this smart green nudging approach over naive nudging of all customers in terms of both reducing product returns and increasing profits. Hence, the construction of the enriched digital footprint and employment of state-of-the-art machine learning methods simultaneously benefits society at large and the retailer, i.e., contributing to the marketing for a better world agenda (e.g., Luo and Bhattacharya 2009, 2006).

Our work has several theoretical and practical implications. First, from a marketing research perspective, our research emphasizes the importance of addressing customers' heterogeneous responses to marketing interventions even if they are overall successful. Despite the already positive impact of our green nudge in the field study, thorough analyses revealed the existence of backfire effects for specific customer segments. By addressing these heterogeneities and making the nudge smart, we are able to reinforce the intervention's impact on product returns and retailer profits. Despite the knowledge that the effectiveness of online marketing interventions may significantly improve when addressing customer heterogeneities (Mills 2022, Peer et al. 2020, Sunstein 2013), it is typically common practice to rely on a "one-size-fits-all strategy". That is because personalizing online interventions, especially in real-time, requires granular data that is often unavailable (Mills 2022). As a result, the development and employment of algorithms to personalize interventions is frequently too costly, impracticable, or assumed to have a negligible additional impact (Peer et al. 2020). Our work not only showcases the possibly substantial benefits of addressing heterogeneities and, therefore, warns marketing researchers to neglect individual-level differences. With our data enrichment and causal machine learning strategy, we additionally present a scalable, data-driven approach to optimize the effectiveness of marketing interventions and even learn about previously unknown factors shaping customer behaviors. In that sense, our work also indicates how novel machine learning technologies may render the development and implementation of marketing interventions more beneficial as they allow for an automatic optimization of these measures on an individual level.

Second, our work adds to the discussion on granularity, and amount of data needed to personalize marketing intervention effectively (e. g., Mills 2022). Understanding customer segments is pivotal to optimizing marketing interventions. To characterize customer segments researchers and practitioners often revert to obtaining customers' personal characteristics, e.g., using statistical models on aggregate data (De Bruyn and Otter 2022) or buying them from data brokers at high costs (Federal Trade Commission 2021). Our work demonstrates the value of merging customer-level with publicly-available, aggregate data—creating an enriched digital footprint—to personalize interventions without relying on expensive third-party services. Importantly, by training a machine learning model on the enriched digital footprint to predict a business figure of interest, our way of enriching the data preserves customers' privacy in regard to sensitive characteristics such as age, gender, or ethical background. Considering increasingly tight regulations on data-privacy (e.g., EU 2016, 2021) our strategy appears to provide a sensible middle-ground to personalize marketing interventions, especially when they contribute to the notion of a better marketing for a better world (Luo and Bhattacharya 2009). Following our insights, the construction and maintenance of aggregate-level databases from publicly available sources constitutes an important complement to the employment of state-of-the-art machine learning methods to personalize marketing interventions.

Third, for practitioners, our data enrichment and analytical methodology to test the effectiveness of making the green nudge smart provide a blueprint to re-examine the effectiveness of previously discarded marketing interventions. In practice, conducting an A/B test and finding an intervention to be overall ineffective at improving a specific variable of interest frequently leads to its abandonment—this is also true for our industry partner. With our approach of enriching customer data, developing a causal machine learning system, and running simulations, practitioners have a tool at their disposal to test whether discarded marketing interventions positively affect specific customer segments that they previously overlooked. To do so, they need the data of a given intervention's A/B test and enrich it utilizing publicly-available aggregate data sets. Novel insights into the existence of pivotal heterogeneities may help practitioners improve the effectiveness of (marketing) interventions and improve the business figure of interest in a more nuanced way.

Fourth, from a practitioner's point of view, our work also illustrates how causal machine learning methods can serve as a novel source for customer insights and inspiration. Specifically, using explainable artificial intelligence methods to render the internal logic of causal machine learning methods interpretable (e.g., Bauer et al. 2022) provides the opportunity to explore and understand how specific customer characteristics contribute to the estimated treatment effect, i.e., understand the implicitly learned customer segmentation. In our case, for instance, feature importances indicate that the employed Causal Forests deems the composition of the average income of private households on the district level as a pivotal information to estimate treatment effects. This observation implies that our green nudge's success somehow depends on customers' wealth. Similar to our industry partner, based on such insights, practitioners may conduct new customer research, re-examine current assumptions and segment of their customers, refine future treatment interventions, and also justify or explain interventions' effectiveness to managers in their organization.

Finally, our work emphasizes the importance of marketing instruments—even low-cost, minimally invasive ones like nudges—to address societal problems that originate from people's consumption behaviors. In our study, even the simple green nudge naively administered to all customers helps to address the environmentally intricate issue of growing product returns in E-Commerce. The use of a data enrichment strategy and causal machine learning methods even enable us to enhance the positive environmental impact of the green nudge. Against this background, our work indicates that modern digital technology may facilitate the adherence to the principles of "Better Marketing for a Better World" (e.g., Luo and Bhattacharya 2009, 2006) by enabling researchers and practitioners to better target small interventions that have a positive impact for society.

A concluding remark is worth making. In our work, we leverage a data enrichment strategy and state-of-the-art machine learning (ML) to optimize an ethically and socially desirable outcome, showcasing how novel digital technologies and Big Data can contribute to solving contemporary problems through personalizing treatment interventions. However, it is vital to keep in mind that the high availability of (customer) data (e.g., customer footprint, aggregated data on regional preferences and behaviors) combined with powerful machine learning technologies can also be misused for undesirable purposes (e.g., targeting vulnerable groups or discriminatory practices). Against this background, we want to emphasize that our proposed approach is not meant to be an argument, much less a plea, for exploiting novel data sources and ML technologies to personalize all kinds of interventions that maximize profits at the expense of ethically, socially, or environmentally objectives. Instead, we comprehend our approach as an example of how technologies enable us to align business and societal goals. We hope that our work inspires researchers and practitioners to leverage available data and increasingly powerful machine learning methods as a means to improve their sustainability measures and make them more viable from a business perspective.

References

Abdulla H, Ketzenberg M, Abbey JD (2019) Taking stock of consumer returns: A review and classification of the literature. *Journal of Operations Management* 65(6):560–605.

Akturk MS, Ketzenberg M, Yıldız B (2021) Managing consumer returns with technology-enabled countermeasures. *Omega* 102:102337.

Allcott H, Rogers T (2014) The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review* 104(10):3003–37.

Altug MS, Aydinliyim T, Jain A (2021) Managing opportunistic consumer returns in retail operations. *Management Science* 67(9):5660–5678.

Ambilkar P, Dohale V, Gunasekaran A, Bilolikar V (2021) Product returns management: A comprehensive review and future research agenda. *International Journal of Production Research* 60(12):3920–3944.

Andor MA, Gerster A, Peters J, Schmidt CM (2020) Social norms and energy conservation beyond the US. *Journal of Environmental Economics and Management* 103:102351.

Anitha J, Kalaiarasu M (2021) Optimized machine learning based collaborative filtering (omlcf) recommendation system in e-commerce. *Journal of Ambient Intelligence and Humanized Computing* 12(6):6387–6398.

Arevalillo JM (2021) Ensemble learning from model based trees with application to differential price sensitivity assessment. *Information Sciences* 557:16–33.

Athey S, Imbens G (2016) Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences* 113(27):7353–7360.

Athey S, Tibshirani J, Wager S (2019) Generalized random forests. *The Annals of Statistics* 47(2):1148–1178.

Athey S, Wager S (2021) Policy learning with observational data. *Econometrica* 89(1):133–161.

Bauer K, von Zahn M, Hinz O (2022) Expl(ai)ned: The impact of explainable artificial intelligence on users' information processing URL `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3872711`.

Berg T, Burg V, Gombović A, Puri M (2020) On the rise of fintechs: Credit scoring using digital footprints. *The Review of Financial Studies* 33(7):2845–2897.

Burrough PA, McDonnell RA, Lloyd CD (2015) *Principles of geographical information systems* (Oxford university press).

Cagala T, Glogowsky U, Rincke J, Strittmatter A (2021) Optimal targeting in fundraising: A causal machine-learning approach. *arXiv preprint arXiv:2103.10251* .

Calma J (2019) Free returns come with an environmental cost. URL `https://www.theverge.com/2019/12/26/21031855/`

free-returns-environmental-cost-holiday-online-shopping-amazon, [Online; accessed 2022-04-17].

Carlsson F, Gravert CA, Kurz V, Johansson-Stenman O (2019) Nudging as an environmental policy instrument. Working Paper.

Carroll J, Lyons S, Denny E (2014) Reducing household electricity demand through smart metering: The role of improved information about energy saving. *Energy Economics* 45:234–243.

Chabe-Ferret S, Le Coent P, Reynaud A, Subervie J, Lepercq D (2019) Can we nudge farmers into saving water? Evidence from a randomised experiment. *European Review of Agricultural Economics* 46(3):393–416.

Davis S, Gerstner E, Hagerty M (1995) Money back guarantees in retailing: Matching products to consumer tastes. *Journal of Retailing* 71(1):7–22.

De P, Hu Y, Rahman MS (2013) Product-oriented web technologies and product returns: An exploratory study. *Information Systems Research* 24(4):998–1010.

De Bruyn A, Otter T (2022) Bayesian consumer profiling: How to estimate consumer characteristics from aggregate data. *Journal of Marketing Research* 59(4):755–774.

Demarque C, Charalambides L, Hilton DJ, Waroquier L (2015) Nudging sustainable consumption: The use of descriptive norms to promote a minority behavior in a realistic online shopping environment. *Journal of Environmental Psychology* 43:166–174.

Deryugina T, Heutel G, Miller NH, Molitor D, Reif J (2019) The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review* 109(12):4178–4219.

Dzyabura D, El Kihal S, Hauser JR, Ibragimov M (2019) Leveraging the power of images in managing product return rates. *Available at SSRN 3209307* .

El Kihal S, Shehu E (2022) It's not only what they buy, it's also what they keep: Linking marketing instruments to product returns. *Journal of Retailing* 98(3):558–571.

EU (2016) Regulation EU 2016/679 of the european parliament and of the council of 27 april 2016, article 22. *Official Journal of the European Union L 119* 59.

EU (2021) Proposal for a regulation EU of the european parliament and of the council of 21 April 2021, laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts. *Official Journal of the European Union* L 119.

Federal Trade Commission (2021) Data brokers: A call for transparency and accountability. URL https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf, [Online; accessed 2022-20-10].

Gelbrich K, Gäthke J, Hübner A (2017) Rewarding customers who keep a product: How reinforcement affects customers' product return decision in online retailing. *Psychology & Marketing* 34(9):853–867.

Goldstein NJ, Griskevicius V, Cialdini RB (2007) Invoking social norms: A social psychology perspective on improving hotels' linen-reuse programs. *Cornell Hotel and Restaurant Administration Quarterly* 48(2):145–150.

Gonzalez-Arcos C, Joubert AM, Scaraboto D, Guesalaga R, Sandberg J (2021) "how do i carry all this now?" Understanding consumer resistance to sustainability interventions. *Journal of Marketing* 85(3):44–61.

Goswami I, Urminsky O (2016) When should the ask be a nudge? the effect of default amounts on charitable donations. *Journal of Marketing Research* 53(5):829–846.

Green T, Peloza J (2014) Finding the right shade of green: The effect of advertising appeal type on environmentally friendly consumption. *Journal of Advertising* 43(2):128–141.

Grinstein A, Riefler P (2015) Citizens of the (green) world? cosmopolitan orientation and sustainability. *Journal of International Business Studies* 46(6):694–714.

Guo R, Cheng L, Li J, Hahn PR, Liu H (2020) A survey of learning causality with data: Problems and methods. *ACM Computing Surveys (CSUR)* 53(4):1–37.

Harris LC (2008) Fraudulent return proclivity: an empirical analysis. *Journal of Retailing* 84(4):461–476.

Hastie T, Tibshirani R, Friedman JH (2009) *The elements of statistical learning: data mining, inference, and prediction*, volume 2 (Springer).

Heiman A, McWilliams B, Zilberman D (2001) Demonstrations and money-back guarantees: market mechanisms to reduce uncertainty. *Journal of Business Research* 54(1):71–84.

Hinz O, Hill S, Sharma A (2022) Second screening—the influence of concurrent tv consumption on online shopping behavior. *Information Systems Research* 33(3):809–823.

Hu M, Dang C, Chintagunta PK (2019) Search and learning at a daily deals website. *Marketing Science* 38(4):609–642.

Hummel D, Maedche A (2019) How effective is nudging? a quantitative review on the effect sizes and limits of empirical nudging studies. *Journal of Behavioral and Experimental Economics* 80:47–58.

Janakiraman N, Syrdal HA, Freling R (2016) The effect of return policy leniency on consumer purchase and return decisions: A meta-analytic review. *Journal of retailing* 92(2):226–235.

Jordan MI, Mitchell TM (2015) Machine learning: Trends, perspectives, and prospects. *Science* 349(6245):255–260.

Kamble S, Gunasekaran A, Arha H (2019) Understanding the blockchain technology adoption in supply chains-indian context. *International Journal of Production Research* 57(7):2009–2033.

Kasperbauer TJ (2017) The permissibility of nudging for sustainable energy consumption. *Energy Policy* 111:52–57.

Kormos C, Gifford R, Brown E (2015) The influence of descriptive social norm information on sustainable transportation behavior: A field experiment. *Environment and Behavior* 47(5):479–501.

Langen H, Huber M (2022) How causal machine learning can leverage marketing strategies: Assessing and improving the performance of a coupon campaign. *arXiv preprint arXiv:2204.10820* .

Lehner M, Mont O, Heiskanen E (2016) Nudging–a promising tool for sustainable consumption behaviour? *Journal of Cleaner Production* 134:166–177.

Liu CW, Agarwal R, Gao G (2016) The dark side of positive social influence. *International Conference on Information Systems (ICIS).*

Loschelder DD, Siepelmeyer H, Fischer D, Rubel JA (2019) Dynamic norms drive sustainable consumption: Norm-based nudging helps café customers to avoid disposable to-go-cups. *Journal of Economic Psychology* 75:102146.

Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. *Neural Information Processing Systems (NeurIPS).*

Luo X, Bhattacharya CB (2006) Corporate social responsibility, customer satisfaction, and market value. *Journal of Marketing* 70(4):1–18.

Luo X, Bhattacharya CB (2009) The debate over doing good: Corporate social performance, strategic marketing levers, and firm-idiosyncratic risk. *Journal of Marketing* 73(6):198–213.

Mills S (2022) Personalized nudging. *Behavioural Public Policy* 6(1):150–159.

Mirsch T, Lehrer C, Jung R (2017) Digital nudging: Altering user behavior in digital environments. *Internationale Tagung Wirtschaftsinformatik (WI).*

Momsen K, Stoerk T (2014) From intention to action: Can nudges help consumers to choose renewable energy? *Energy Policy* 74:376–382.

Moorthy S, Srinivasan K (1995) Signaling quality with a money-back guarantee: The role of transaction costs. *Marketing Science* 14(4):442–466.

Murakami K, Shimada H, Ushifusa Y, Ida T (2020) Heterogeneous treatment effects of nudge and rebate: Causal machine learning in a field experiment on electricity conservation. *Graduate School of Economics Discussion Paper Series* Discussion Paper No. E–20–003.

Narang U, Shankar V, Narayanan S (2019) The impact of mobile app failures on purchases in online and offline channels. *Working Paper.*

National Retail Federation N (2020) Customer returns in the retail industry 2020. Technical report, [Online; accessed 2022-04-10].

National Retail Federation N (2022) Customer returns in the retail industry 2021. Technical report, [Online; accessed 2022-04-10].

Neumann N, Tucker CE, Whitfield T (2019) Frontiers: How effective is third-party consumer profiling? evidence from field studies. *Marketing Science* 38(6):918–926.

Ofek E, Katona Z, Sarvary M (2011) "bricks and clicks": The impact of product returns on the strategies of multichannel retailers. *Marketing Science* 30(1):42–60.

Oprescu M, Syrgkanis V, Wu ZS (2019) Orthogonal random forest for causal inference. *International Conference on Machine Learning (ICML)*.

Pålsson H, Pettersson F, Hiselius LW (2017) Energy consumption in e-commerce versus conventional trade channels-insights into packaging, the last mile, unsold products and product returns. *Journal of Cleaner Production* 164:765–778.

Peer E, Egelman S, Harbach M, Malkin N, Mathur A, Frik A (2020) Nudge me right: Personalizing online security nudges to people's decision-making styles. *Computers in Human Behavior* 109:106347.

Pei Z, Paswan A (2018) Consumers'legitimate and opportunistic product return behaviors in online shopping. *Journal of Electronic Commerce Research* 19(4):301–319.

Petersen JA, Kumar V (2015) Perceived risk, product returns, and optimal resource allocation: Evidence from a field experiment. *Journal of Marketing Research* 52(2):268–285.

Ramzan B, Bajwa IS, Jamil N, Amin RU, Ramzan S, Mirza F, Sarwar N (2019) An intelligent data analysis for recommendation systems using machine learning. *Scientific Programming* 2019.

Reich JW, Robertson JL (1979) Reactance and norm appeal in anti-littering messages. *Journal of Applied Social Psychology* 9(1):91–101.

Richter I, Thøgersen J, Klöckner CA (2018) A social norms intervention going wrong: Boomerang effects from descriptive norms information. *Sustainability* 10(8):2848.

Ruokamo E, Meriläinen T, Karhinen S, Räihä J, Suur-Uski P, Timonen L, Svento R (2022) The effect of information nudges on energy saving: Observations from a randomized field experiment in finland. *Energy Policy* 161:112731.

Schneider C, Weinmann M, Vom Brocke J (2018) Digital nudging: Guiding online user choices through interface design. *Communications of the ACM* 61(7):67–73.

Schölkopf B, Locatello F, Bauer S, Ke NR, Kalchbrenner N, Goyal A, Bengio Y (2021) Toward causal representation learning. *Proceedings of the IEEE* 109(5):612–634.

Schubert C (2017) Green nudges: Do they work? are they ethical? *Ecological economics* 132:329–342.

Schuler A, Baiocchi M, Tibshirani R, Shah N (2018) A comparison of methods for model selection when estimating individual treatment effects. *arXiv preprint arXiv:1804.05146* .

Shead S (2021) Amazon plans to cut waste following backlash over the destruction of unused products. URL `https://www.cnbc.com/2021/08/04/amazon-plans-to-cut-waste-following-backlash.html`.

Shehu E, Papies D, Neslin SA (2020) Free shipping promotions and product returns. *Journal of Marketing Research* 57(4):640–658.

Su X (2009) Consumer returns policies and supply chain performance. *Manufacturing & Service Operations Management* 11(4):595–612.

Sunstein CR (2013) Impersonal default rules vs. active choices vs. personalized default rules: A triptych. URL `https://ssrn.com/abstract=2171343`.

Thaler RH, Sunstein CR (2009) *Nudge. Improving decisions about health, wealth and happiness* (London: Penguin Books).

Tian X, Sarkis J (2022) Emission burden concerns for online shopping returns. *Nature Climate Change* 12(1):2–3.

Wager S, Athey S (2018) Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association* 113(523):1228–1242.

Walsh G, Möhring M (2017) Effectiveness of product return-prevention instruments: Empirical evidence. *Electronic Markets* 27(4):341–350.

Weinmann M, Schneider C, Brocke vJ (2016) Digital nudging. *Business & Information Systems Engineering* 58(6):433–436.

White K, Habib R, Hardisty DJ (2019) How to shift consumer behaviors to be more sustainable: A literature review and guiding framework. *Journal of Marketing* 83(3):22–49.

Xia F, Chatterjee R, May JH (2019) Using conditional restricted boltzmann machines to model complex consumer shopping patterns. *Marketing Science* 38(4):711–727.

Zemack-Rugar Y, Moore SG, Fitzsimons GJ (2017) Just do it! Why committed consumers react negatively to assertive ads. *Journal of Consumer Psychology* 27(3):287–301.

Zhang D, Frei R, Senyo P, Bayer S, Gerding E, Wills G, Beck A (2023) Understanding fraudulent returns and mitigation strategies in multichannel retailing. *Journal of Retailing and Consumer Services* 70:103145.

Zhang M, Luo L (2022) Can consumer-posted photos serve as a leading indicator of restaurant survival? Evidence from yelp. *Management Science* forthcoming.

Zhou W, Hinz O (2016) Determining profit-optimizing return policies–a two-step approach on data from taobao. com. *Electronic markets* 26(2):103–114.

Zhou W, Hinz O, Benlian A (2018) The impact of the package opening process on product returns. *Business Research* 11(2):279–308.

Online Appendix

Randomization checks

In our field experiment, visitors were randomly assigned to treatment (green nudging) and control groups (no nudging) when they first checked their shopping basket. To test whether our randomization process was successful, we performed various randomization checks. Figure 7 presents the randomization checks for exemplary shopping session characteristics.

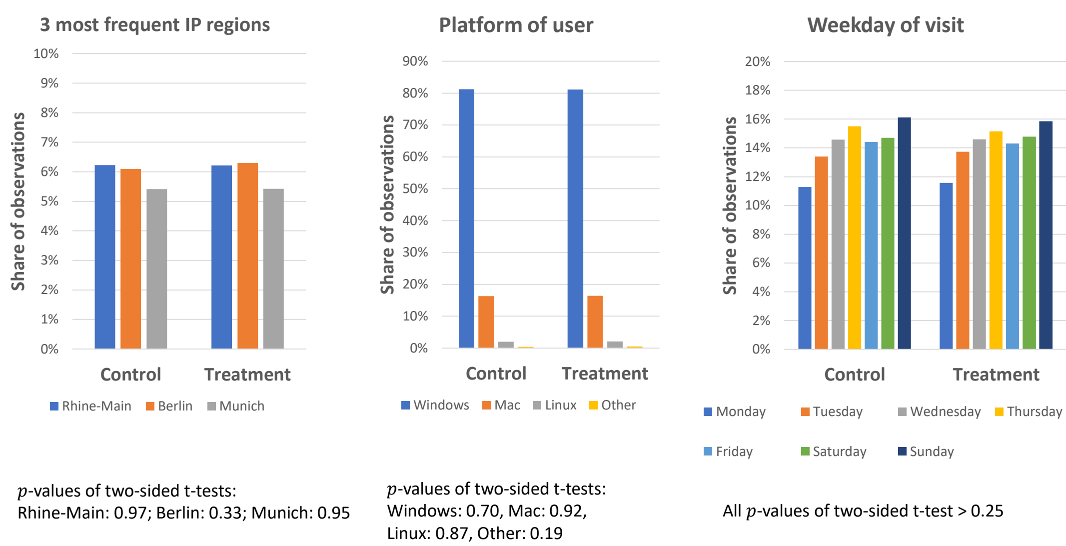

Figure 7: Examples of randomization checks.

Heterogeneous effects of green nudging in different customer subgroups

Table 5 illustrates other exemplary customer segments where the green nudge backfires. Evidently, the green nudge works particularly well for customers who already have adopted a certain level of PEB and like to purchase sustainable products. In this group, the green nudge significantly decreases returns by 4.9% (Kruskal-Wallis test: $\chi^2 = 9.246$, $p < 0.01$) and increases profits by 10.9% (Kruskal-Wallis test: $\chi^2 = 9.7$, $p < 0.01$). Similarly, the green nudge is highly effective in reducing opportunistic product returns by addressing bracketing behavior. In our field experiment, the green nudge reduces product returns for customers who engage in bracketing by i. e., by 2.8% (Kruskal-Wallis test: $\chi^2 = 3.846$, $p < 0.05$) and even manages to turn such previously not profitable customers into lucrative ones (Kruskal-Wallis test: $\chi^2 = 6.513$, $p < 0.05$). By contrast, the effect of the green nudge is limited —statistically and economically—for customers (i) who did not shop sustainable products (ii) or did not order duplicate products (i. e., who did not engage in bracketing). Overall, the illustrated effects further corroborate the existence of considerable treatment heterogeneities for different customer segments.

		Returns		**Profits**	
	Obs.	Mean (Std. Dev.)	Effect	Mean (Std. Dev.)	Effect
Sustainable products in Basket: None					
Control	6,224	1.646 (3.689)	⇑	0.460 (2.685)	⇓
Treatment	6,546	1.657 (3.784)		0.444 (2.573)	
Sustainable products in Basket: One or more					
Control	18,095	2.605 (5.213)	⇓ ***	0.937 (3.458)	⇑ ***
Treatment	18,735	2.483 (5.121)		1.040 (3.581)	
Number of duplicate products in Basket: None					
Control	16,286	1.135 (2.600)	⇓	1.225 (2.265)	⇑
Treatment	17,135	1.109 (2.528)		1.237 (2.244)	
Number of duplicate products in Basket: One or more					
Control	8,033	4.841 (7.028)	⇓ *	-0.018 (4.606)	⇑ **
Treatment	8,146	4.709 (7.071)		0.145 (4.860)	

*Kruskal-Wallis test significance levels: * p <0.10, ** p <0.05, *** p <0.01*

Table 5: Effect of green nudging in various locations

Hyperparameter tuning

We determine the hyperparameters of the causal forest by applying a grid search in a 5-fold cross-validation on the training set based on out-of-sample R-score performance (Schuler et al. 2018). We report the tuning ranges of the grid search in Table 6.

Hyperparameter	Values	Meaning
max_samples	0.3, 0.45	number of samples in training set for each tree
min_balancedness_tol	0.3, 0.5	imbalance tolerated between child nodes
min_var_fraction_leaf	None, 0.01	minimum variation of treatment vector in leaf
min_samples_leaf	5, 50, 100	minimum number of samples in leaf
max_depth	5, 7, None	maximum depth of trees
ntree	500	number of tree estimators in forest

Table 6: Hyperparameter grid for training of causal forest.

Excerpts from customer feedback

In the weeks following the field experiment, our industry partner received feedback indicating that customers value its efforts for greater sustainability. Based on more than 16,000 net promoter scores and search poll feedbacks in the weeks after the experiment, customers exhibited a slightly higher overall satisfaction and explicitly praised the retailer's growing focus on sustainability. We report selected quotes of customer feedback in Table 7. While these insights merely represent anecdotal evidence, they nonetheless indicate that the introduction of green nudging has additional positive effects on company image and possibly long-term growth.

	Customer quotes
What customers praised	"The new look and the focus on sustainability! Great job!"
	"Variety in clothing, recycling and sustainability offerings"
	"Buying sustainable products"
	"The eco-friendliness. I would wish that every person pays more attention to this. [...]"
	"The products [associated with] sustainability"
What customers suggested	"Even more sustainability"
	"Become even more environmentally friendly in every way :)"
	"[The retailer] should produce more fairly and sustainably."
	"Even more organic/sustainable (it's already great, but there's always more)"
	"More sustainable products."

Table 7: Exemplifying quotes from customer feedback collected at the online retailer in the weeks following the field experiment.

Recent Issues

No. 362	Tabea Bucher-Koenen, Andreas Hackethal, Johannes Kasinger, Christine Laudenbach	Disparities in Financial Literacy, Pension Planning, and Saving Behavior
No. 361	Ata Can Bertay, José Gabo Carreño Bustos, Harry Huizinga, Burak Uras, Nathanael Vellekoop	Technological Change and the Finance Wage Premium
No. 360	Alfons J. Weichenrieder	A Note on the Role of Monetary Policy When Natural Gas Supply Is Inelastic
No. 359	Spencer Yongwook Kwon, Yueran Ma, Niklas Kaspar Zimmermann	100 Years of Rising Corporate Concentration
No. 358	Matteo Bagnara, Ruggero Jappelli	Liquidity Derivatives
No. 357	Huynh Sang Truong, Uwe Walz	Spillovers of PE Investments
No. 356	Markus Eyting	Why do we Discriminate? The Role of Motivated Reasoning
No. 355	Stephan Jank, Emanuel Moench, Michael Schneider	Safe Asset Shortage and Collateral Reuse
No. 354	Sebastian Steuer	Common Ownership and the (Non-)Transparency of Institutional Shareholdings: An EU-US Comparison
No. 353	Olga Balakina, Claes Bäckman, Andreas Hackethal, Tobin Hanspal Dominique M. Lammer	Good Peers, Good Apples? Peer Effects in Portfolio Quality
No. 352	Monica Billio, Michele Costola, Loriana Pelizzon, Max Riedel	Creditworthiness and buildings' energy efficiency in the Italian mortgage market
No. 351	Markus Dertwinkel-Kalt, Johannes Kasinger, Dmitrij Schneider	Skewness Preferences: Evidence from Online Poker
No. 350	Ruggero Jappelli, Konrad Lucke, Loriana Pelizzon	Price and Liquidity Discovery in European Sovereign Bonds and Futures
No. 349	Monica Billio, Michele Costola, Iva Hristova, Carmelo Latino, Loriana Pelizzon	Sustainable Finance: A journey toward ESG and climate risk